2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application

2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application | 978-1-6654-9457-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/HPCC-DSS-SmartCity-DependSys53884.2021.00159

A Video Stitching System of Underwater Image

Zhenyu Wang[1,2], Haoming Wu[1,2], Tianming Yu[1,2], Mengqin Fu[1,2], Bernie Liu[2,*], Jieren Cheng[3,4,*]
Quanfeng Ye[1,2], Ni Zhang[1,2], Shuo Huang[1,2], Xingyue Lu[1,2] Xiulai Li[1,5] Kai Lu[6] and Gengquan Xie[1,7]
[1] Hainan University, Haikou, 570228, China
[2] RobAI-Lab, Hainan University
[3] School of Computer Science and Technology, Hainan University, Haikou, China
[4] Hainan Blockchain Technology Engineering Research Center, Haikou, China
[5] Hainan Hairui Zhong Chuang Technol Co Ltd, Haikou, China
[6] Hainan Vocational College of Political Science and Law, Haikou, China
[7] College of Foreign Language, Hainan University, Haikou, China
{ZhenyuWang860571, cn_whm, Tianming, by.liu}@ieee.org, {cjr22}@163.com

***Abstract*—Stitching underwater videos captured by a hand-held camera can essentially enhance the entertainment experience. However, underwater video is difficult to stitch due to its shortcomings in definition, visibility, shakiness and parallax. In this paper, we propose a video stitching system for underwater videos captured by mobile cameras. Our method consists of three components, underwater image enhancement, unified video stitching and 3D projection. The underwater image enhancement is a component based on Relative Global Histogram Stretching (RGHS). After the processing of this component, videos with higher definition and better visibility can be obtained, which will benefit the subsequent video stitching. The unified video stitching is a component based on background identification. Through background identification, we divide video stitching into two steps: background video stitching and foreground video stitching. This makes the stitched panoramic video more stable and the picture more natural, which is the key to heavy shakiness and large parallax problem. In addition, we have also studied the playback of the panoramic video to make it more immersive. We test the proposed system on videos that are captured by hand-held cameras when diving, and use SSIM, UCIQE, UIQM, and stitching score to quantitatively evaluate the generated panoramic video. The experimental results prove the effectiveness and robustness of our method in various cases, and the better video stitching results than existing methods.**

***Index Terms*—Video stitching, image enhancement, background identification, 3D projection, panoramic video**

I. Introduction

Showing clear, true and comprehensive videos of underwater image is of great significance to marine ecology, marine science, underwater biological recognition and underwater robot vision.

However, limited by the natural environment, the video of underwater image obtained by direct shooting has the disadvantages of poor visibility, low definition, and small field of view.

In order to solve the first two shortcomings, many excellent underwater image restoration methods [1-7] have been produced in the past decades, which has made the definition and visibility of underwater images qualitatively improved.

Regarding the shortcomings of videos of underwater image with small field of view, although there are many excellent video stitching methods [8-14], most of the methods default all the videos used for the experiment are shot on land. These methods are not robust enough towards videos of underwater image captured underwater.

In this paper, we try to solve the problem of stitching underwater videos. For example, a typical scenario we envision is: A diver wears a camera holder to fix a camera on his forehead, while holding a second camera on the diver's hand. And the diver takes videos while diving. After the diver has landed and exported the video, the two videos can be stitched together to get a more entertaining video with a large field of view. Stitching these two videos is very challenging for two reasons: (1) The videos captured underwater have poor visibility and low definition. (2) Because the foreground and background of the videos captured by different cameras have significant depth changes, there may be a large parallax between the videos.

Facing such two challenges, we propose a robust video stitching algorithm for the videos captured underwater. Our algorithm can make the underwater video have better visibility and higher definition, and produce a very coherent panoramic video.

As shown in Fig. 1, different from the general image stitching scheme [8-14], our method consists of two components, underwater image enhancement and unified video stitching. In order to obtain a coherent underwater panoramic video, it is necessary to perform image enhancement on all frames of the video, and then stitch corresponding frames between different videos

Based on the inherent simplicity and the physical model of underwater images, we adopted a method of relative global histogram stretching [15] for underwater image enhancement. [15] has relatively ideal results on the three indicators of SSIM [16], UCIQE [17], [18] and UIQM [19], which will benefit the subsequent video stitching work.

The processed video is stitched in the stitching system

978-1-6654-9457-1/21/$31.00 ©2021 IEEE
DOI 10.1109/HPCC-DSS-SmartCity-DependSys53884.2021.00159

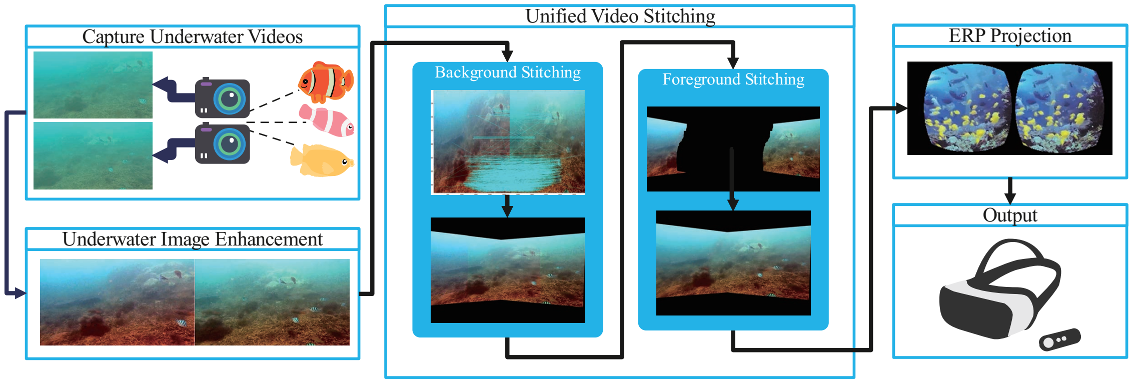

Fig. 1. Our method consists of three components. The underwater image enhancement [15] is a pre-processing stage, which increases the clarity of the video. The second stage, unified video stitching, consists of background stitching [21],[22] and foreground stitching [20], completes the stitching of underwater videos. The third stage is ERP projection and the video can be played in a way that suits the human eye after this stage.

designed by us. The system consists of two parts: background stitching and foreground stitching. When shooting a video in a complex underwater environment, the parallax between foreground objects and background objects is very large, so it is impossible to stitch the two at the same time. Therefore, we separate the foreground and background stitching. First, we identify common backgrounds for stitching between background areas to avoid the negative impact of foreground feature matching on stitching. Second, we perform seam cutting and covering strategies in the stitched common area [12], [20-22] to synthesize the transformed video to obtain panoramic video. We add these two methods to the loop of the video stitching system.

Finally, we use ERP projection to play the panoramic video in a way that suits the human eye as the output of the system.

The main contribution of this paper is that we have established a unified underwater video stitching system that facilitates image enhancement and video stitching. As far as we know, this is the first system that can optimize the processing and stitching of underwater videos. It can stitch underwater videos captured by multiple cameras. Experiments show that our new video stitching method is superior to other competitive video stitching methods (see Section IV-D).

II. Related Work

Our work is related to underwater image enhancement, image video stitching, and 360° video projection. We also reviewed the background identification works related to video stitching.

A. Underwater Image Enhancement

In previous work, a large number of methods have been proposed for image processing. He's Dark Channel Prior (DCP) dehazing method [3] is widely used in underwater image enhancement [24, 25]. Based on the enhancement algorithm of histogram equalization, Sun et al. [26] set adaptive upper and lower limits for histogram equalization. The upper limit is used to remove image noise, and the lower limit is used to protect image details. Considering the effectiveness of the physical model used in underwater images and the inherent properties of underwater images, we use relative global histogram stretching (RGHS) [15] as the method of underwater image enhancement. [15] mainly Based on the G-B channel equalization and histogram stretching in the RGB color model, the stretching range parameters are determined according to the distribution characteristics of the original image and the light absorption of different wavelengths underwater.

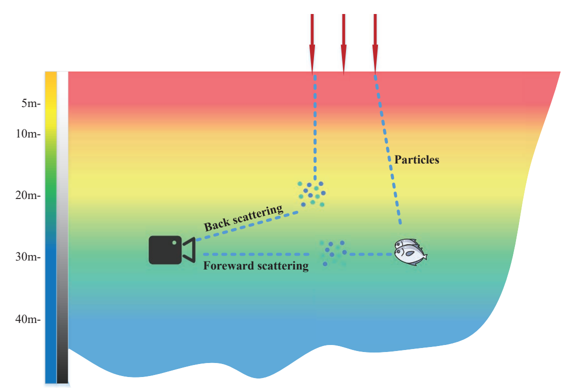

Fig. 2. The basic principles of underwater optical imaging.

B. Image Stitching

Nowadays, image stitching is a practical technology widely used in image processing. Early methods [27], [28] used a single homographies to align two images. However, these methods require the captured scene to be flat, otherwise it will fail due to obvious parallax. According to the way of handling parallax, recent image stitching work can be roughly divided into two categories. Distortion-based methods [29-32] use multiple models to represent the relationship between images. For example, Gao et al. [29] calculated two isomorphic images to stitch images with only two principal planes. Zaragoza et al. [31] divided the image into grids and calculated the spatial variation of homographies in these grids. These methods can handle small parallax, but if the parallax is very large, the misalignment will still be remained. In order to deal with more serious parallax, the seam-based method [33-35] finds the seam with the least energy in the overlapping area of the two images (considered as the stitching boundary). Large parallax is the key problem when stitching underwater videos. To solve this problem, we clearly divide objects of different depths into foreground and background objects. Then use the warping-based method to stitch the background area, and use the seam-based method to obtain the foreground object.

C. *Video Stitching*

Compared with image stitching, video stitching has received much less attention. Different methods are proposed because of the different camera settings. For example, some works [8], [9] are based on static cameras, while some [10-14] on moves the camera but the relative positions between the cameras are still fixed. In these two types of settings, the videos are stitched globally according to the determined position between the cameras, and then locally adjusted. Under the premise of the determined positions of multiple cameras, the processing method of this type of method is relatively simple and efficient. However, the works of [20-22], [36-38] take multiple cameras with changing relative positions as the research object in video stitching. The difference is that the works of [38], [39] did not take the stability of stitching into consideration. The works of [20-22], [36] considered stability for better stitching. Because the video obtained from underwater shooting is very unstable, we also consider the stability during stitching, and take the temporal feature matching between the frames of the input video as one of the cores of video stitching.

D. $360°$ *Video Projection*

Currently, the most widely used projection method for 360 video is ERP projection. Most of the shooting sequences are stored in ERP format. There are several works on ERP and CMP coding. For example, a co-domain-based face filling mechanism [41] is proposed to improve the motion compensation of CMP. Likewise, Li et al. [42] derive the pixels outside the face boundary by using a 3D fill based on the coprojection plane. To achieve the interactive goal of our system, we are inspired by the method of [42]. We project 2D video on a 3D stereoscopic surface according to the characteristics of the video format in order to accomplish interactive video browsing on mobile devices.

III. Video Stitching System of Underwater Image

Fig. 1 shows an overview of our system. The core of the system is video stitching. Without any other components, the component itself can generate panoramic video. But for better stitching effect and final display results, we proposed the second and third components to make the video stitching system more robust. Among them, the second component is to perform image enhancement on each frame of the input video. In this way, we can improve the clarity and visibility of the picture, which is more conducive to the subsequent image stitching. And the third component brings users a better sense of immersion through a special projection. In the following, we first introduce our main components, and then introduce the second and third components.

A. *Unified Video Stitching*

Video Stitching relies on two kinds of feature matching: (1) the temporal feature matching between consecutive frames of the input video; (2) the spatial feature matching between the corresponding frames of the two input videos. The former matching is used for video stabilization, while the latter is used to stitch two input videos together. In order to make our main components more robust, we divide the stitching work into two parts: background stitching and foreground stitching.

a) Background Video Stitching: We look back the methods of [21], [22]. We first use the KLT tracker to calculate the characteristic trajectory of the input video. Then divide the input video into a series of overlapping temporal windows, each window has 40 frames, and two adjacent windows have 20 overlapping frames. In the second row, each time window is clustered into several regions according to the characteristic trajectory in the window by the motion segmentation method [40]. Finally, a grid graph is constructed: each column in the graph corresponds to each time window; one column contains multiple nodes, which correspond to the segmented area of the corresponding time window; when there is at least one characteristic trajectory passing through the corresponding area, Add edges between nodes in adjacent columns. The weight of the edge is defined as:

$$\omega^{i,j}_{k,k+1} = exp(\alpha S^{i,j}_{k,k+1}) \times r^{i,j}_{k,k+1} \tag{1}$$

In this formula, k and k+1 indicate trellis columns, i and j are indices of nodes in the columns, is the number of common feature trajectories passing both the regions of nodes n^i_k and n^j_{k+1}, and $r^{i,j}_{k,k+1}$ is the rank of the matrix formed by the set of common trajectories (lower rank means the two nodes are more likely being background [40], [43]). The parameter α is used to adjust the influence of $S^{i,j}_{k,k+1}$ which is set as 0.01. Given a clearly defined trellis graph, dynamic programming can be used to efficiently calculate the energy minimum path of the trellis graph. The video area corresponding to the node of the minimum energy path is regarded as the background area. Using a similar idea, after adding SIFT [43] matching, the common background area is obtained by traversing the node combination between different videos and calculating the corresponding path. After dividing the image into a background area and a foreground area, in order to perform global video stitching and stabilization for the background area as the object, we adopted the bundled camera path stabilization method of [20], [21] to divide the input video into small grids It also stabilizes the shakiness path of each small grid, and calculates the global mosaic homographies to stitch the distorted grid for matching the temporal characteristics better.

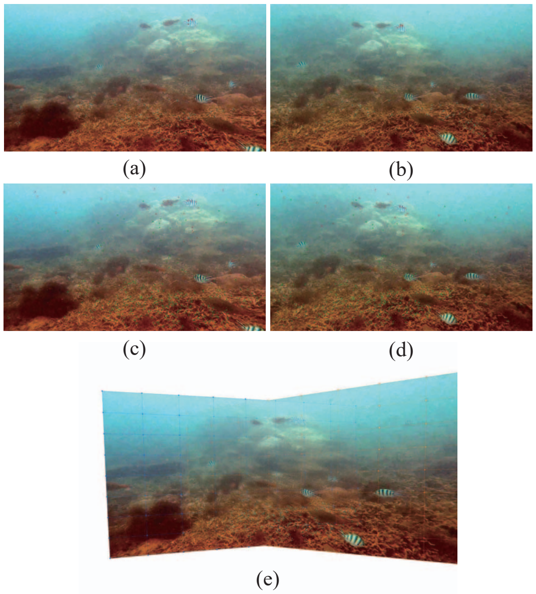

Fig. 3. (a)(b) Background identified in underwater image. (c)(d) Common Background identified in underwater image. (e) Video stitching based only on the feature points of the background.

b) Foreground Video Stitching: After video stitching with the background area as the object, we found that the background stitching in the non-overlapping area is satisfactory, but the foreground area has ghost images. In order to seamlessly stitch the two videos, we adopted the schemes in [12] and [20]. First, we find a seam in the overlapping area. Second, we keep the part of the left video on the left side of the seam, and keep the part of the right video on the right side of the seam, and stitch the reserved parts together to get the final panoramic video. The seams of adjacent frames should maintain time consistency.

B. Underwater Image Enhancement

The steps of background identification and stabilization in video stitching will be affected by the image quality. According to the comparison of multiple schemes [1-7], we used Relative Global Histogram Stretching (RGHS), which has the best performance in the three evaluation indicators of SSIM [16], UCIQE [17], [18], UIQM [19]. This is an underwater image enhancement method based on Adaptive-stretching in CIE-Lab Color Model, which greatly improves the final video stitching effect.

C. ERP Projection

In the previous sections, we have completed the generation of the underwater panoramic video. The generated panoramic video can be displayed on a flat screen in 2D form. In order to achieve the interactive goal of our system, we project the 2D video on the 3D stereo surface according to the characteristics of the video format, in order to complete the interactive browsing of the video on the mobile device. The projection method we use is ERP projection. In ERP projection, the image is projected on a two-dimensional plane in a cylindrical projection. ERP is a simple projection method that maps the meridians to vertical lines with constant spacing and the latitude lines to horizontal lines with constant spacing. At present, the ERP projection is most widely used in panoramic video, and most of the shooting sequences are stored in ERP format.

IV. Experiments

A. Evaluation Metrics

a) Image Enhancement: We first introduce three metrics used to evaluate the results of underwater image enhancement:SSIM compares local patterns of pixel intensities that have been normalized for luminance and contrast [16]. UCIQE is the latest non-reference model for underwater color image quality evaluation with a comprehensive indicator of chroma, saturation and contrast [17], [18]. UIQM can effectively evaluate the quality of underwater images in accordance with human perception [19].

b) Video Stitching: Stitching score evaluates the quality of stitching. We use a similar metric as Guo et al. [35], which is widely used in the field of video stitching. There is a stitching boundary (i.e., seam) in each frame of a panoramic video. For each frame, we find pairs of matched features within a small region that is near the boundary, and compute the distances between the feature matches. The average distance of all the matches is viewed as the stitching score of the frame. The average score of all the frames is viewed as the stitching score of the panoramic video. The smaller the stitching score, the better aligned the panoramic video.

B. Video Data

The video captured by ourselves in a very simple way: we let a participant hold two Go Pro HERO BLACK by two hands and capture videos while diving. Due to the fact of "diving", the captured videos are very shaky. All videos have 100 frames of images with 2704×1520 resolution.

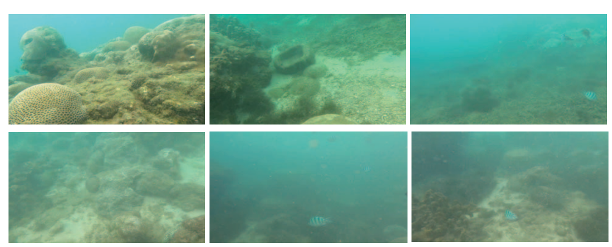

Fig. 4. The video data taken in Lingshui County, Haikou, China.

C. Experimental Information

a) Parameter: During video stitching, we extract and match approximately 1000 SIFT feature points per frame. In the video stitching component of our system, the most important parameter is the weight β of the stitching item, which is used to balance the stitching accuracy and stability of the stitched video. In our experiment, β is set between 10 and 25, depending on the parallax between the input videos. If there is significant disparity, it means that the weight of the stitching item should be set larger to avoid excessive stitching error. A larger stitching weight is good for the output stitching score, but it also sacrifices stability. In our experiments, by observing the image stitching effect corresponding to each stitching score value, we found that the stitching error is acceptable when the stitching score is lower than 3. In the case of different inputs, other parameters that are not specified are all constants.

b) Performance Analysis: We conduct experiments with un-optimized MATLAB implementation of our method on a PC with AMD Ryzen 7 5800H 3.20 GHz and 32GB memory. We use SIFT implementation provided by vlfeat library for feature correspondence matching and the MATLAB's built-in KLT tracker for trajectory extraction. These two procedures cost 0.5s and 15s for a frame of a video with 2704×1520 resolution. Underwater image enhancement, as the preprocessing part of the video, takes about 90 seconds per frame. The main component of our system is unified video stitching. For the background video stitching part, it takes about 36 seconds per frame. The background video stitching part is relatively slow, because there are SIFT feature extraction and matching, background identification and other parts. If the foreground stitching part is added, The entire stitching process takes an average of 45 seconds per frame. Therefore, we look forward to a more efficient feature extraction and background identification method to speed up the overall speed of video stitching.

D. Result and Discussion

(a) (b) (c) (d)

Fig. 5. (a) Original; (b) DCP; (c) UCM; (d) Our method.

a) Underwater Image Enhancement: The method we used, characterized by improved contrast, equalized saturation, and enhanced brightness, was compared qualitatively and quantitatively with the two traditional methods. Single image haze removal using dark channel prior by He [24] is a classic technique for image haze removal. Because underwater images are usually processed as haze images, we use this method as one of the comparison objects. The other comparison method is UCM [26] because it is an effective non-physical method, and the histogram is modified like the method we used. The resulting images obtained by different underwater image enhancement methods are shown in Fig 5. Obviously, the image enhancement effect of DCP is not good, and the image before and after processing does not change much in sharpness and visibility. UCM oversaturates the color of the image, which is more unnatural than our method. At the same time, the result image of UCM has obvious noise, and the sharpness of the image has decreased. We use SSIM [16], UCIQE [17], [18], UIQM [19] three indicators to quantitatively analyze the three methods. We use three sets of underwater images based on different cases to compare the three methods. Table I show the evaluation results of the three methods in Case 1, 2, and 3 respectively. Our method obtained the best results in the evaluation. It retains the most details, so that underwater images have excellent visibility and quality. The maximum UCIQE value means that our method can effectively balance the chroma, saturation, and contrast of underwater images. DCP is the worst method among the three methods. This means that the dark channel prior method cannot be directly used for underwater image enhancement. However, the underwater images produced by UCM have poor clarity and low image quality. The results show that our method can produce high-quality underwater images, and is superior to the current mainstream underwater image enhancement methods.

TABLE I
ASSESSMENT RESULTS OF THE THREE METHODS IN CASE 1 TO 3

CASE	*METHODS*	*SSIM*	*UCIQE*	*UIQM*
1	DCP	0.776	0.451	0.316
	UCM	0.832	0.508	0.551
	OUR	0.85	0.518	0.896
2	DCP	0.738	0.381	0.2
	UCM	0.743	0.445	0.409
	OUR	0.824	0.466	0.979
3	DCP	0.746	0.4	0.234
	UCM	0.641	0.465	0.296
	OUR	0.762	0.481	0.808

b) Unified Video Stitching: Our method is compared with three other existing methods: (1) The method of dynamic video stitching through shakiness removal proposed by Nie et al. [21]; (2) The method of combining DCP with our unified video stitching component; and (3) The method of combining UCM with our unified video stitching component. Fig 6. shows our Background Video Stitching process. The process is mainly composed of two parts: SIFT feature point extraction and matching and video stitching based on background feature points. Fig 7. shows the image changes before and after the processing of Foreground Video Stitching. We found that foreground video stitching eliminates the ghosting on the common background of the left and right images, making the stitching result more natural. In addition, we use Stitching Score [9] as the evaluation index of our video stitching (the smaller the index means the better the stitching quality). Table II, III shows the evaluation results of our video stitching. We not only compare with the method without image preprocessing [21], but also compare with DCP and UCM on the premise of only

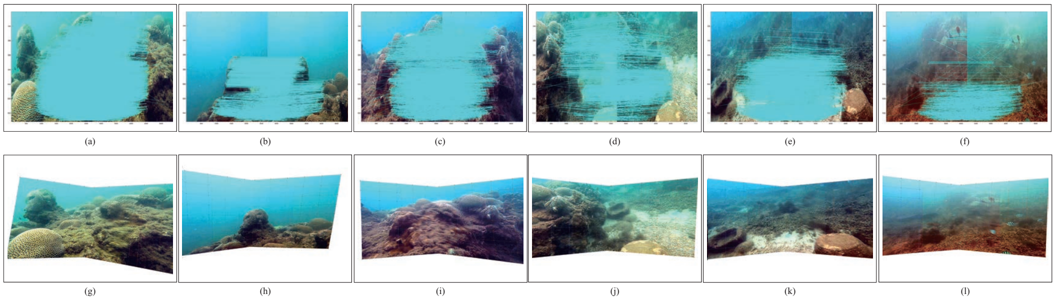

Fig. 6. The first row (a)(b)(c)(d)(e)(f) shows matching results of feature point. The second row (g)(h)(i)(j)(k)(l) shows stitching results based only on background feature points.

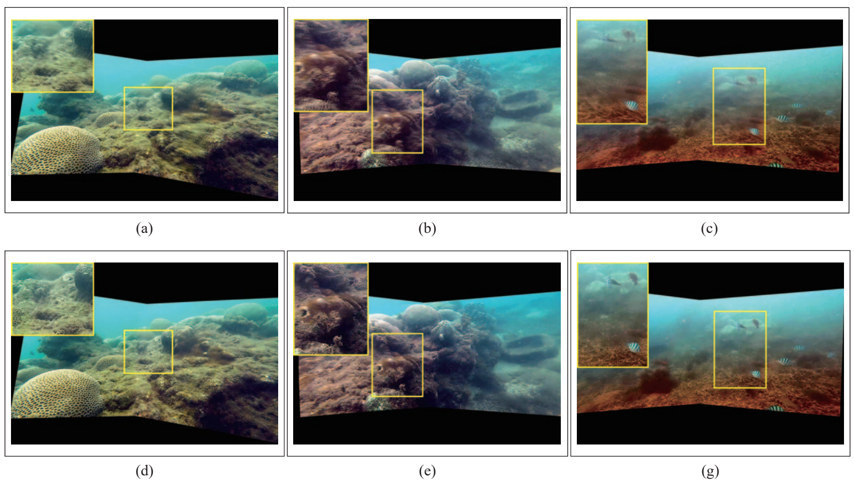

Fig. 7. The first row (a)(b)(c) shows stitching results without foreground stitching. The second row (d)(e)(f) shows stitching results with foreground stitching.

changing the image enhancement method. We stitched together 6 scenes. It can be seen from Table II that for cases 1 and 2, there are no foreground objects in the case, so the final video stitching quality is not good, but the frames of underwater videos are greatly improved after image enhancement. For Case 3 with fish as the foreground object, the video stitching quality has also been significantly improved. The results of the three comparisons show that our proposed work of image enhancement before underwater video stitching can effectively improve the quality of video stitching. It can be seen from Table III that for cases 5 and 6 with fish as the foreground object, RGHS has a significant improvement in the quality of video stitching, DCP has a small improvement in the quality of video stitching, and UCM has a negative effect on video stitching. For Case 4 without any foreground objects, only RGHS improves the video stitching quality, while the other two schemes reduce the stitching quality. The results of these three sets of comparisons show that the RGHS underwater video stitching we used has the best improvement effect, while UCM has the worst. In addition, because RGHS has a lifting effect in all three scenarios, it can be seen that RGHS has better adaptability.

TABLE II
STITCHING SCORE OF [21] AND OUR METHODS ON CASE 1 TO 3

Case	***1***	***2***	***3***
[21]	7.6293	3.5101	3.3434
OUR	4.2509	1.9510	2.5762

TABLE III
STITCHING SCORE OF [21], DCP, UCM AND OUR METHODS ON CASE 4 TO 6

Case	***4***	***5***	***6***
[21]	1.4009	1.3825	1.3121
DCP	1.4862	0.9084	0.8018
UCM	1.8392	1.4877	2.0302
OUR	1.3035	0.6635	0.4291

c) ERP Projection: We conducted a separate experiment for ERP projection. With the VR glasses, we found the output shown in Fig. 8 is effective: a strong sense of visual difference between the two eyes, but without too much vertigo-inducing and excessive parallax.

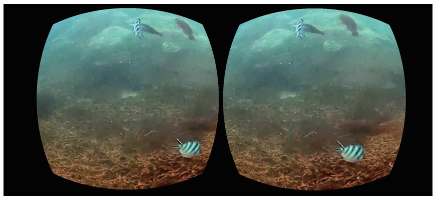

Fig. 8. The projection output of case 3.

E. Limitations

Although this method has achieved a good stable stitching effect, it is time-consuming both for unified video stitching and underwater image enhancement, which is a major limitation of its application expansion. And in the set of experimental parameter, the weight β of the stitching item is set by human factors. Thus, the the weight β of the stitching item requires further quantitative exploration. The different underwater environments in different waters will lead to different optical imaging. Therefore, we will optimize the existing methods to increase the computing speed and study more underwater scene materials as future work.

V. Conclusion And Future Work

We introduce a new system that can effectively stitch underwater videos. Our method consists of three components, underwater image enhancement, unified video stitching and 3D projection. Compared with previous methods, ours can produce better stitching results. We also compared different underwater image enhancement methods in terms of image quality and video stitching quality to prove that our method has better effectiveness and robustness. At the same time, we have also conducted the most suitable playback format of panoramic underwater video, so that the audience can get a better sense of immersion. The method proposed in this paper has potential applications in the fields of robotics [44-46]. In further research, we will work to improve the performance and efficiency of our current system.

VI. Acknowledgment

This work was supported by Key Research and Development Program of Hainan Province (Grant No.ZDYF2020033), Hainan Provincial Natural Science Foundation of China (Grant No.2019RC107, Grant No.2019RC098, Grant No.621RC612, Grant No.621RC1082), Young Talents'Science and Technology Innovation Project of Hainan Association for Science and Technology (Grant No.QCXM202007), Key Project of College Students' Innovation and Entrepreneurship of Hainan University (Grant No.20210110, Grant No.20210112), and Innovation and Entrepreneurship Training Program for College Students (Grant No.202110589017X).

References

[1] K. Zuiderveld, "Contrast limited adaptive histogram equalization," *Graphics Gems*, pp. 474–485, 1994.

[2] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," in *IEEE Conference on Computer Vision & Pattern Recognition*, 2012.

[3] K. He, S. Jian, Fellow, IEEE, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, 2011.

[4] R. Hummel, "Image enhancement by histogram transformation," *Computer Graphics and Image Processing*, 1977.

[5] I. Kashif, R. A. Salam, O. Azam, and A. Z. Talib, "Underwater image enhancement using an integrated colour model," *Iaeng International Journal of Computer Science*, vol. 34, no. 2, pp. 239–244, 2007.

[6] K. Iqbal, M. O. Odetayo, A. E. James, R. A. Salam, and A. Z. Talib, "Enhancing the low quality images using unsupervised colour correction method," in *IEEE*, 2010.

[7] A. Ghani and N. Isa, "Underwater image quality enhancement through composition of dual-intensity images and rayleigh-stretching," *IEEE*, 2015.

[8] M. Zheng, X. Chen, and L. Guo, "Stitching video from webcams," in *International Symposium on Visual Computing*. Springer, 2008, pp. 420–429.

[9] B. He, G. Zhao, and Q. Liu, "Panoramic video stitching in multi-camera surveillance system," in *2010 25th International Conference of Image and Vision Computing New Zealand*. IEEE, 2010, pp. 1–6.

[10] A. Kaheel, M. El-Saban, M. Refaat, and M. Ezz, "Mobicast: a system for collaborative event casting using mobile phones," in *Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia*, 2009, pp. 1–8.

[11] W. Xu and J. Mulligan, "Panoramic video stitching from commodity hdtv cameras," *Multimedia systems*, vol. 19, no. 5, pp. 407–426, 2013.

[12] W. Jiang and J. Gu, "Video stitching with spatial-temporal content-preserving warping," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2015, pp. 42–48.

[13] F. Perazzi, A. Sorkine-Hornung, H. Zimmer, P. Kaufmann, O. Wang, S. Watson, and M. Gross, "Panoramic video from unstructured camera arrays," in *Computer Graphics Forum*, vol. 34, no. 2. Wiley Online Library, 2015, pp. 57–68.

[14] J. Li, W. Xu, J. Zhang, M. Zhang, Z. Wang, and X. Li, "Efficient video stitching based on fast structure deformation," *IEEE Transactions on Cybernetics*, vol. 45, no. 12, pp. 2707–2719, 2015.

[15] D. Huang, Y. Wang, W. Song, J. Sequeira, and S. Mavromatis, "Shallow-water image enhancement using relative global histogram stretching based on adaptive parameter acquisition," in *International conference on multimedia modeling*. Springer, 2018, pp. 453–465.

[16] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[17] M. Yang and A. Sowmya, "An underwater color image quality evaluation metric," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 6062–6071, 2015.

[18] A. Horé and D. Ziou, "Image quality metrics: Psnr vs. ssim," in *2010 20th International Conference on Pattern Recognition*, 2010, pp. 2366–2369.

[19] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures," *IEEE Journal of Oceanic Engineering*, vol. 41, no. 3, pp. 541–551, 2016.

[20] K. Lin, S. Liu, L.-F. Cheong, and B. Zeng, "Seamless video stitching from hand-held camera inputs," in *Computer Graphics Forum*, vol. 35, no. 2. Wiley Online Library, 2016, pp. 479–487.

[21] Y. Nie, T. Su, Z. Zhang, H. Sun, and G. Li, "Dynamic video stitching via shakiness removing," *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 164–178, 2017.

[22] T. Su, Y. Nie, Z. Zhang, H. Sun, and G. Li, "Video stitching for handheld inputs via combined video stabilization," in *SIGGRAPH ASIA 2016 Technical Briefs*, 2016, pp. 1–4.

[23] R. Schettini and S. Corchs, "Underwater image processing: state of the art of restoration and image enhancement methods," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, pp. 1–14, 2010.

[24] A. Galdran, D. Pardo, A. Picón, and A. Alvarez-Gila, "Automatic red-channel underwater image restoration," *Journal of Visual Communication and Image Representation*, vol. 26, pp. 132–145, 2015.

[25] P. Drews, E. Nascimento, F. Moraes, S. Botelho, and M. Campos, "Transmission estimation in underwater single images," in *Proceedings of the IEEE international conference on computer vision workshops*, 2013, pp. 825–830.

[26] L. Sun, X. Wang, X. Liu, P. Ren, P. Lei, J. He, S. Fan, Y. Zhou, and Y. Liu, "Lower-upper-threshold correlation for underwater range-gated imaging self-adaptive enhancement," *Applied optics*, vol. 55, no. 29, pp. 8248–8255, 2016.

[27] R. Szeliski and H.-Y. Shum, "Creating full view panoramic image mosaics and environment maps," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, 1997, pp. 251–258.

[28] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *International journal of computer vision*, vol. 74, no. 1, pp. 59–73, 2007.

[29] J. Gao, S. J. Kim, and M. S. Brown, "Constructing image panoramas using dual-homography warping," in *CVPR 2011*. IEEE, 2011, pp. 49–56.

[30] W.-Y. Lin, S. Liu, Y. Matsushita, T.-T. Ng, and L.-F. Cheong, "Smoothly varying affine stitching," in *CVPR 2011*. IEEE, 2011, pp. 345–352.

[31] J. Zaragoza, T.-J. Chin, M. S. Brown, and D. Suter, "As-projective-as-possible image stitching with moving dlt," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2013, pp. 2339–2346.

[32] C.-H. Chang, Y. Sato, and Y.-Y. Chuang, "Shape-preserving half-projective warps for image stitching," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3254–3261.

[33] J. Gao, Y. Li, T.-J. Chin, and M. S. Brown, "Seam-driven image stitching." in *Eurographics (Short Papers)*, 2013, pp. 45–48.

[34] F. Zhang and F. Liu, "Parallax-tolerant image stitching," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3262–3269.
bibitemagarwala2004interactive A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen, "Interactive digital photomontage," in *ACM SIGGRAPH 2004 Papers*, 2004, pp. 294–302.

[35] H. Guo, S. Liu, T. He, S. Zhu, B. Zeng, and M. Gabbouj, "Joint video stitching and stabilization from moving cameras," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5491–5503, 2016.

[36] M. El-Saban, M. Izz, and A. Kaheel, "Fast stitching of videos captured from freely moving devices by exploiting temporal redundancy," in *2010 IEEE International Conference on Image Processing*. IEEE, 2010, pp. 1193–1196.

[37] M. El-Saban, M. Izz, A. Kaheel, and M. Refaat, "Improved optimal seam selection blending for fast video stitching of videos captured from freely moving devices," in *2011 18th IEEE International Conference on Image Processing*. IEEE, 2011, pp. 1481–1484.

[38] F.-L. Zhang, J. Wang, H. Zhao, R. R. Martin, and S.-M. Hu, "Simultaneous camera path optimization and distraction removal for improving amateur video," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5982–5994, 2015.

[39] J. Bai, A. Agarwala, M. Agrawala, and R. Ramamoorthi, "User-assisted video stabilization," in *Computer Graphics Forum*, vol. 33, no. 4. Wiley Online Library, 2014, pp. 61–70.

[40] J. Sauer, J. Schneider, and M. Wien, "Improved motion compensation for 360° video projected to polytopes," in *2017 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE,2017, pp. 61–66.

[41] L. Li, Z. Li, X. Ma, H. Yang, and H. Li, "Co-projection-plane based 3-d padding for polyhedron projection for 360-degree video," in *2017 IEEE International conference on Multimedia and Expo (ICME)*. IEEE, 2017, pp. 55–60.

[42] L. Li, Z. Li, M. Budagavi, and H. Li, "Projection based advanced motion model for cubic mapping for 360-degree video," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 1427–1431.

[43] Y. Ma, H. Derksen, W. Hong, and J. Wright, "Segmentation of multivariate mixed data via lossy data coding and compression," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 9, pp. 1546–1562, 2007.

[44] B. Liu, L. Wang, and M. Liu, "Lifelong federated reinforcement learning: a learning architecture for navigation in cloud robotic systems," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4555–4562, 2019.

[45] B. Liu, L. Wang, M. Liu, and C.-Z. Xu, "Federated imitation learning: A novel framework for cloud robotic systems with heterogeneous sensor data," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3509–3516, 2019.

[46] B. Liu, L. Wang, X. Chen, L. Huang, D. Han, and C.-Z. Xu, "Peer-assisted robotic learning: a data-driven collaborative learning approach for cloud robotic systems," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 4062–4070.