Machine Learning Research
2024, Vol. 9, No. 2, pp. 75-79
https://doi.org/10.11648/j.mlr.20240902.16

Research Article

An Union Method Combining the Stitching of Normal Images and the Unsupervised Semantic Segmentation of Stitched Image

Xing Hu[1, 2], Xinjian Li[2], Zhengguang Wang[2], Jie Ren[2], Yi An[1, 3, *], Cheng Shao[1], Hongsheng Tian[4], Qingru Guo[5]

[1]School of Control Science and Engineering, Dalian University of Technology, Dalian, China
[2]College of Information Technology, Nanchang Vocational University, Nanchang, China
[3]School of Electrical Engineering, Xinjiang University, Urumqi, China
[4]State Grid Tieling Electric Power Supply Company, Liaoning, China
[5]Product Development Department, Googosoft, Jinan, China

Abstract

The union combining the stitching of normal images and unsupervised semantic segmentation of the stitched image is an important region, which is crucial for autonomous driving, intelligent robots, and vehicle detection. This paper designs an union method combining the stitching of normal images and unsupervised semantic segmentation of the stitched image. The normal images are stitched by using the image stitching method designed by Ribeiro D. The semantic segmentation method for the stitched image uses the method opened in the github. The stitched image contains image distortion. The distortion of the stitched image will make the feature extraction unreasonable. The distortion form of the stitched image is different from the distortion form of the panoramic image combined by line images. Therefore, the DCM proposed by Xing Hu is useless to extract features of the stitched image reasonable. This paper improves the DCM as the improved distortion convolution module (IDCM) by using the deformable convolution, the clamp module, the type transformation module, and the gather module. The IDCM is added before the unsupervised semantic segmentation method opened in the github to extract features reasonable. The IDCM-NUSSM method and the ISM-IDCM-NUSSM method are proposed. The experimental results show the better performance of the designed methods.

Keywords

Panoramic Image, Image Stitching, Unsupervised Semantic Segmentation

1. Introduction

The union combining the stitching of normal images and unsupervised semantic segmentation of the stitched image is an important region, which is crucial for autonomous driving, intelligent robots, and vehicle detection. For the semantic

*Corresponding author: anyi@dlut.edu.cn (Yi An)

Received: 1 November 2024; **Accepted:** 19 November 2024; **Published:** 29 November 2024

segmentation methods of the panoramic image, Oztekin et al. [1] designed an automatic semantic segmentation method for the dental restorations using U-Net model. Yang et al. [2] proposed the PASS method to make the proposed method be used to the semantic segmentation of the panoramic annular image. Yang et al. [3] proposed an omnisupervised omnidirectional semantic segmentation method for the panoramic image. Zhang et al. [4] designed a distortion-aware transformer for adapting to the semantic segmentation of the panoramic image, which contains a deformable patch embedding, a deformable MLP, and a mutual prototypical adaptation method. Xu et al. [5] improved a dataset of the panoramic image using the dataset of the normal image and uses the improved dataset to segment the panoramic image. Orhan et al. [6] proposed a semantic segmentation method for the panoramic image using the equirectangular convolution to handle the distortion in the panoramic image. Yang et al. [7] improved a context-aware CNN to explore the wide context information of the panoramic image.

For the image stitching methods, the image stitching methods can be mainly divided into two categories: feature-based methods and deep-learning-based methods.

In deep-learning-based methods, Zhu et al. [8] improved a novel panorama generative model for synthesizing realistic and sharp-looking panorama, which does not require a large number of labeled ground-truth data. Sumantri S. J. et al. [9] designed a learning-based approach the reconstructs the scene in 360° ×180° from a sparse set of conventional images. Wu et al. [10] tackled the problem of synthesizing a ground-view panorama image conditioned on a top-view aerial image, which is a challenging problem in this domain.

In feature-based methods, Liu et al. [11] designed a novel image stitching method based on the scale-invariant feature transform (SIFT) and the color constancy theory, which can be used to complex environments and uneven lighting conditions. Zhang et al. [12] proposed a view synthesis approach based on optical flow to generate a high-quality omnidirectional panorama. Lin et al. [13] designed an image synthesis method which performs not only the image stitching process, but also a collaborative scheme that facilitates the image acquisition process via the use of multiple mobile phones. Kim et al. [14] presented a method to generate image mosaics of a panoramic scene to solve the problem that the relation between images cannot be expressed by a signal homography. Xiong et al. [15] proposed a spatially-varying deformation method that uses multiple local transformation matrices compared with the traditional panoramic image stitching algorithms that use only one global projection transformation matrix for the image alignment.

This paper designs an union method combining the stitching of normal images and unsupervised semantic segmentation of the stitched image. The normal images are stitched by using the image stitching method designed by Ribeiro D. [16]. The semantic segmentation method for the stitched image uses the method opened in the github [17]. The stitched image contains image distortion. The distortion of the stitched image will make the feature extraction unreasonable. The distortion form of the stitched image is different from the distortion form of the panoramic image combined by line images. Therefore, the DCM proposed by Xing Hu [18] is useless to extract features of the stitched image reasonable. This paper improves the DCM as the improved distortion convolution module (IDCM) by using the deformable convolution, the clamp module, the type transformation module, and the gather module. The IDCM is added before the unsupervised semantic segmentation method opened in the github to extract features reasonable. The IDCM-NUSSM method and the ISM-IDCM-NUSSM method are proposed. The experimental results show the better performance of the designed methods.

2. Methods

2.1. The Image Stitching Method

The normal images are stitched by using the image stitching method designed by David Ribeiro.

2.2. The IDCM

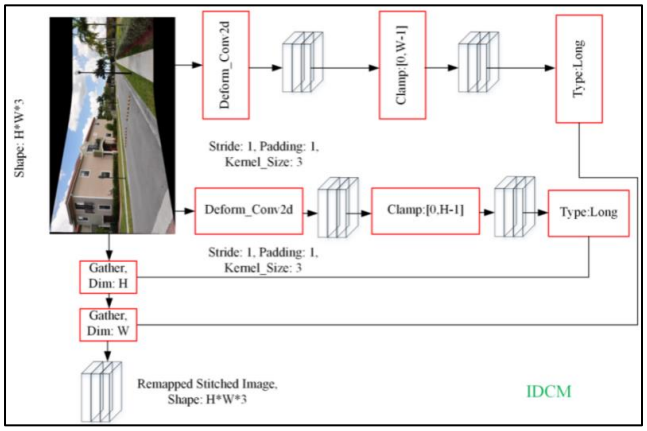

Figure 1. *The IDCM.*

The IDCM is as shown in Figure 1. The size of the stitched image is $H \times W \times 3$. The width of the stitched image is W, the height of the stitched image is H, and the number of the color channel of the stitched image is 3. The stitched image is first extracted features using the horizontal deformable convolution, which is implemented by the torchvision library. The parameters of the horizontal deformable convolution are the kernel size 3, stride 1, and padding 1. Then, the stitched image is extracted features using the vertical deformable convolution, which is also implemented by the torchvision library. The parameters of the vertical deformable convolution are the kernel size 3, stride 1, and pad-

ding 1.

The feature maps output by the horizontal deformable convolution module are clamped by $[0, W-1]$ and are transformed into the data type of long integer. The feature maps output by the vertical deformable convolution module are clamped by $[0, H-1]$ and are transformed into the data type of long integer.

Then, the clamped and transformed feature maps of the vertical deformable convolution module are used as the index to remap the stitched image by H-dimension and the clamped and transformed feature maps of the horizontal deformable convolution module are used as the index to secondly remap the remapped stitched image by W-dimension to obtain the output remapped stitched image. The output remapped stitched image is also $H \times W \times 3$.

The IDCM is added before the unsupervised semantic segmentation method designed by Asako Kanezaki to extract features reasonable and a new unsupervised semantic segmentation method called IDCM-NUSSM is proposed. Then, the stitched image is semantically segmented by the IDCM-NUSSM. The overall method of the image stitching method and the unsupervised semantic segmentation method is proposed, which is called the ISM-IDCM-NUSSM.

3. Results

3.1. The Experiment Settings

We select four images as the datasets of the image stitching method. The training epochs for the unsupervised semantic segmentation method is 40, the minimum number of labels in the unsupervised semantic segmentation method is 5, and the maximum number of labels in the unsupervised semantic segmentation method is 256.

3.2. The Ablation Experiments

The ablation experiments are as shown in Figure 2. In Figure 2, the left column is the images used in the image stitching method, the up image in the right column is the image stitching result used in this paper, the middle image in the right column is the semantic segmentation result using the method designed by Asako Kanezaki, and the down image in the right column is the semantic segmentation result using the method designed by Asako Kanezaki combined with the IDCM.

From Figure 2, we can see the image stitching method used in this paper can obtain better the stitched image, the semantic segmentation result using the method designed by Asako Kanezaki combined with the IDCM is better compared with the semantic segmentation result using the method designed by Asako Kanezak. Therefore, the performances of our IDCM and the image stitching method and the unsupervised semantic segmentation method are validated.

Figure 2. *The images, stitched Image, the semantic segmentation result of stitched image using the method designed by Asako Kanezaki, and the semantic segmentation result of stitched image using the method designed by Asako Kanezaki combined with the IDCM.*

4. Discussion

This paper designs an union method combining the stitching of normal images and unsupervised semantic segmentation of the stitched image. The normal images are stitched by using the image stitching method designed by Ribeiro D.. The semantic segmentation method for the stitched image uses the method opened in the github. The stitched image contains image distortion. The distortion of the stitched image will make the feature extraction unreasonable. The distortion form of the stitched image is different from the distortion form of the panoramic image combined by line images. Therefore, the DCM proposed by Xing Hu is useless to extract features of the stitched image reasonable. This paper improves the DCM as the improved distortion convolution module (IDCM) by using the deformable convolution, the clamp module, the type transformation module, and the gather module. The IDCM is added before the unsupervised semantic segmentation method opened in the github to extract features reasonable. The IDCM-NUSSM method and the ISM-IDCM-NUSSM method are proposed. The experimental results show the better performance of the proposed methods.

5. Conclusion

This paper designs an union method combining the stitching of normal images and unsupervised semantic segmentation of the stitched image. The stitched image contains image

distortion. The distortion of the stitched image will make the feature extraction unreasonable. The distortion form of the stitched image is different from the distortion form of the panoramic image combined by line images. Therefore, the DCM proposed by Xing Hu is useless to extract features of the stitched image reasonable.

1) this paper improves the DCM as the improved distortion convolution module (IDCM) by using the deformable convolution, the clamp module, the type transformation module, and the gather module.
2) the IDCM is added before the unsupervised semantic segmentation method opened in the github to extract features reasonable. The IDCM-NUSSM method and the ISM-IDCM-NUSSM method are proposed.

The experimental results show the better performance of the designed methods.

Abbreviations

CNN	Convolution Neural Network
DCM	Distortion Convolution Module
IDCM	Improved Distortion Convolution Module
IDCM-NUSSM	Improved Distortion Convolution Module-New Unsupervised Semantic Segmentation Method
ISM-IDCM-NUSSM	Image Stitching Method-Improved Distortion Convolution Module-New Unsupervised Semantic Segmentation Method

Acknowledgments

Lingyue Zeng, Employment and Entrepreneurship Service Office, Nanchang Vocational University, Nanchang, China.

Shunan Zhao, School of Control Science and Engineering, Dalian University of Technology, Dalian, China.

Lin Wu, School of Control Science and Engineering, Dalian University of Technology, Dalian, China.

Hongsheng Tian, State Grid Tieling Electric Power Supply Company, Liaoning, China.

Author Contributions

Xing Hu: Data curation, Software, Validation, Writing – original draft

Xinjian Li: Conceptualization, Formal Analysis, Methodology, Project administration, Resources

Zhengguang Wang: Conceptualization, Data curation, Resources

Jie Ren: Data curation, Methodology, Software, Project administration, Visualization

Yi An: Formal Analysis, Funding acquisition, Validation, Writing – review & editing

Cheng Shao: Data curation, Funding acquisition, Resources, Visualization

Hongsheng Tian: Conceptualization, Funding acquisition, Investigation, Supervision

Qingru Guo: Investigation, Project administration, Validation, Visualization

Code Availability

https://download.csdn.net/download/accdgh/89867700

Data Availability Statement

https://download.csdn.net/download/accdgh/89867696

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Oztekin F., Katar O., Sadak F., Aydogan M., Yildirim T. T., Plawiak P., Yildirim O., Talo M., and Karabatak M., "Automatic Semantic Segmentation for Dental Restorations in Panoramic Radiography Images Using U-Net Model", *Int. J. Imaging Syst. Technol.*, 2022, 32(6): 1990-2001, https://doi.org/abs/10.1002/ima.22803

[2] Yang K., Hu X., Bergasa L. M., Romera E., Huang X., Sun D., and Wang K., "Can we PASS beyond the Field of View? Panoramic Annular Semantic Segmentation for Real-World Surrounding Perception", in *Proc. IEEE Intell. Veh. Symp. Proc.*, Paris, France, 2019: 446-453, https://doi.org/10.1109/IVS.2019.8814042

[3] Yang K., Hu X., Fang Y., Wang K., and Stiefelhagen R., "Omnisupervised Omnidirectional Semantic Segmentation", *IEEE Trans. Intell. Transp. Syst.*, 2020, 23(2): 1184-1199, https://doi.org/10.1109/TITS.2020.3023331

[4] Zhang J., Yang K., Shi H., Reiβ S., Peng K., Ma C., Fu H., Torr P. H. S., Wang K., and Stiefelhagen R., "Behind Every Domain There is a Shift: Adapting Distortion-aware Vision Transformers for Panoramic Semantic Segmentation", *IEEE Trans. Pattern Anal. Mach. Intell.*, 2024, https://doi.org/10.1109/TPAMI.2024.3408642

[5] Xu Y., Wang K., Yang K., Sun D., and Fu J., "Semantic Segmentation of Panoramic Images Using a Synthetic Dataset", in *Proc. SPIE Int. Soc. Opt. Eng.*, Strasbourg, France, 2019, Art. No. 111690B, https://doi.org/10.1117/12.2532494

[6] Orhan S. and Bastanlar Y., "Semantic Segmentation of Outdoor Panoramic Images", *Signal Image Video P.*, 2022, 16(3): 643-650, https://doi.org/10.1007/s11760-021-02003-3

[7] Yang K., Hu X., and Stiefelhagen R., "Is Context-aware CNN Ready for the Surroundings? Panoramic Semantic Segmentation in the Wild", *IEEE Trans. Image Process.*, 2021, 30: 1866-1881, https://doi.org/10.1109/TIP.2020.3048682

[8] Zhu D., Zhou Q., Han T., and Chen Y., "360 Degree Panorama Synthesis from Sequential Views Based on Improved FC-densenets", *IEEE Access*, 2019, 7: 180503-180511, https://doi.org/10.1109/ACCESS.2019.2958111

[9] Sumantri J. S. and Park I. K., "360 Panorama Synthesis from a Sparse Set of Images on a Low-Power Device", *IEEE Trans. Comput. Imag.*, 2020, 6: 1179-1193, https://doi.org/10.1109/TCI.2020.3011854

[10] Wu S., Tang H., Jing Xiao-Y., Zhao H., Qian J., Sebe N., and Yan Y., "Cross-View Panorama Image Synthesis", *IEEE Trans. Multimedia*, 2023, 25: 3546-3559, https://doi.org/10.1109/TMM.2022.3162474

[11] Liu Y. and Long Y., "Image Stitching Method of Non-Ferrous Smelting Scene Based on SIFT Algorithm and Color Constancy", *Open Journal of Applied Sciences*, 2023, 13(11): 2055-2067, https://doi.org/10.4236/ojapps.2023.1311161

[12] Zhang W., Wang Y., and Liu Y., "Generating High-quality Panorama by View Synthesis Based on Optical Flow Estimation", *Sensors*, 2022, 22(2), Art. No. 470, https://doi.org/10.3390/s22020470

[13] Lin J. and Yang Chuan-K., "Collaborative Panoramic Image Generation from Multiple Mobile Phones", in *Proc. IEEE Int. Conf. Multimed. and Expo Workshops*, Hong Kong, Hong Kong, 2017: 339-344, https://doi.org/10.1109/ICMEW.2017.8026222

[14] Kim S., Uh Y. J., and Byun H., "Generating Panorama Image by Synthesizing Multiple Homography", in *Proc. Int. Conf. Image Process.*, 2012: 2981-2984, https://doi.org/10.1109/ICIP.2012.6467526

[15] Xiong J., Li F., Long F., Xu Y., Wang S., Xu J., Ling Q., "Spatially-varying Warping for Panoramic Image Stitching", in *Proc. Chin. Control Decis. Conf.*, 2022: 575-580, https://doi.org/10.1109/CCDC55256.2022.10032845

[16] Ribeiro D., CustÓdio P., and Balasubramaniam L., "Image Stitching and 3D Point Cloud Registration", *Image Processing and Vision MEEC*, 2021.

[17] Kanezaki A., "Unsupervised Image Segmentation by Backpropagation", in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process Proc.*, 2018: 1543-1547, https://doi.org/10.1109/ICASSP.2018.8462533 [Accessed 19 June 2019]

[18] Hu X., An Y., Shao C., and Hu H., "Distortion Convolution Module for Semantic Segmentation of Panoramic Images Based on the Image-forming Principle", *IEEE Trans. Instrum. Meas.*, 2022, 71: 1-12, https://doi.org/10.1109/TIM.2021.3139710