OBSIR: Object-Based Stereo Image Retrieval

Xiangyang Xu, Wenjing Geng, Ran Ju, Yang Yang, Tongwei Ren and Gangshan Wu

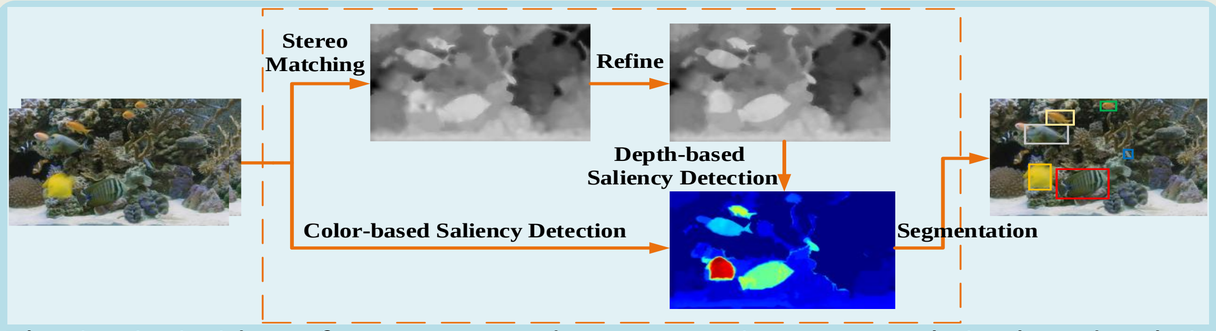
State Key Laboratory for Novel Software Technology, Department of Computer Science and Technology, Nanjing University, China

Introduction

Recent years, the stereo image has become an emerging media in the field of 3D technology and numerous 3D devices have emerged. For example, 3D smartphone, stereo camera, DV, 3D TV, 3D projector and so on. And this made the stereo media more and more popular. When searching stereo images on Flickr, we get a large number of results and the number is still increasing today. Consequently, there is an urgent demand of stereo image retrieval.

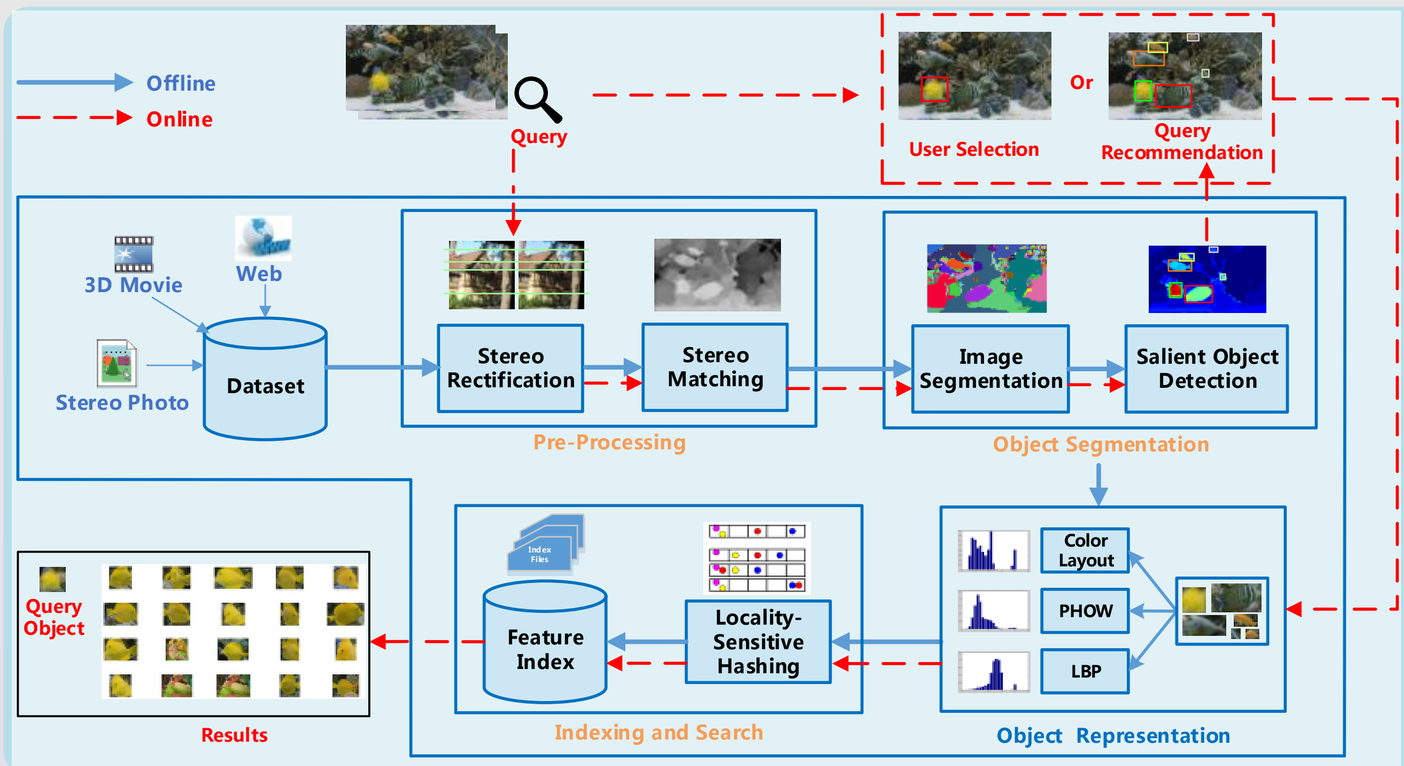
To the best of our knowledge, we believe our work is the first attempt to explicitly establish a systematically framework for object-based stereo image retrieval, and also the first one to build up a stereo image dataset for the evaluation of stereo image retrieval task.

Key idea



The intrinsic idea of our proposed segmentation approach is that depth is usually discontinuous across the “object” edge. And this will be served as a complementary to color cues to facilitate the object segmentation.

Proposed approach



Given a stereo image, and the object region that the user specifies, we aim to search for the relevant images containing the similar object. The framework of our approach is shown in left figure, which is composed of two pipelines: offline and online processing. For the offline processing, we first collect a large dataset of stereo images. Then a preprocessing including image resizing, duplicate removal, stereo image rectification and stereo matching is taken. In this step we get the disparity maps which encode the depth information. After that, we extract salient objects from the image. For each object, we extract visual features and use the BoVW model for representation. At last, the feature vectors are indexed by a cluster-based LSH. For the online processing, a stereo image is uploaded by the user first and then passed through the preprocessing and object segmentation module sequentially. A few objects are generated then and displayed in the object box to give a recommendation to the user. The user may directly pick one object or drag a rectangular region of interest as the final query. Then the query region is passed to the object representation module and converted to a fixed BoVW vector. The search is performed by the cluster-based LSH and at last, a list of objects is returned and displayed in the user interface. Object results are highlighted in their owner images to show the entire context of the objects.

Performance Evaluation

Salient Object segmentation

Table 1. Salient object segmentation performance

	Ours	COMP [9]	RC [19]	SS [10]
Precision	0.703	0.304	0.592	0.614
Recall	0.571	0.527	0.259	0.268

Figure 2. Salient object extraction examples

Object-based stereo image retrieval performance

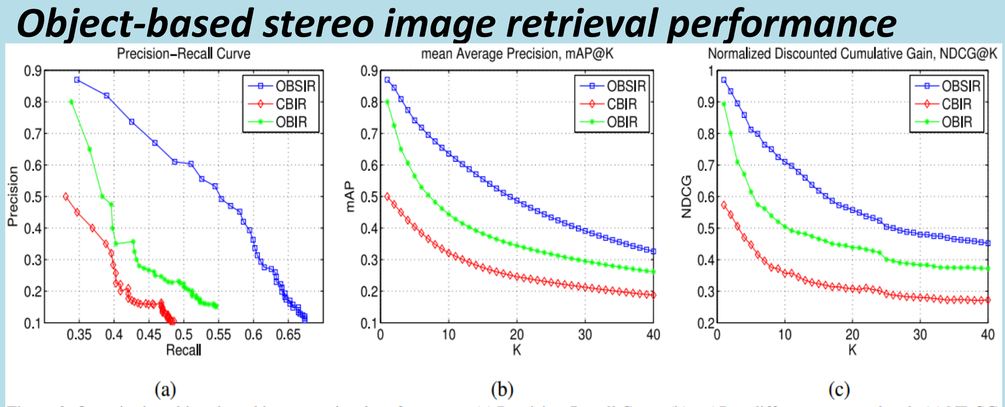


Figure 3. Quantitative object-based image retrieval performance: (a) Precision-Recall Curve (b) mAP at different truncate levels (c) NDCG at different truncate levels

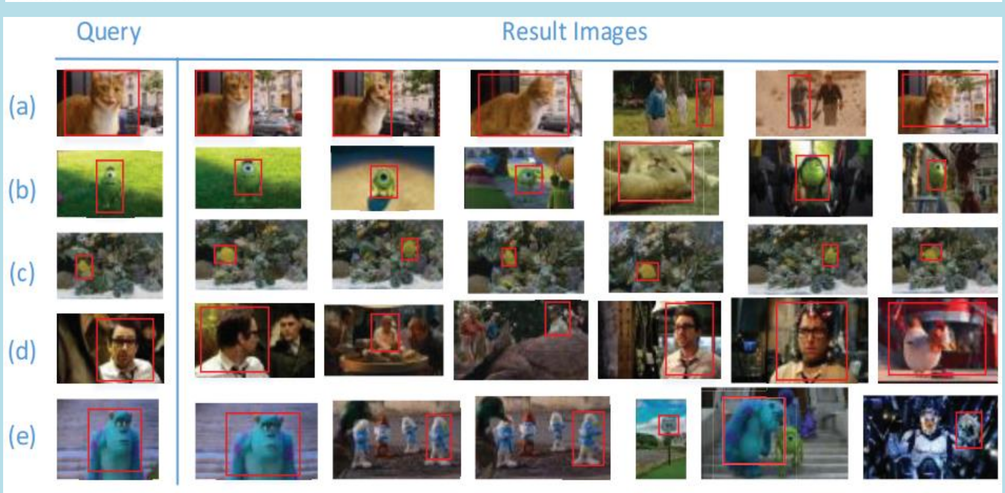


Figure 4. Top returned examples of some queries from OBSIR

Some References

[1] Shih-Fu Chang. How far we’ve come: Impact of 20 years of multimedia information retrieval. TOMCCAP, 2013.

[2] James Philbin, et al. Object retrieval with large vocabularies and fast spatial matching. CVPR, 2007.

[6] Andreas Geiger, et al. Efficient large-scale stereo matching. ACCV, 2011.

[9] Jie Feng, et al. Salient object detection by composition. ICCV, 2011.

[10] Yuzhen Niu, et al. Leveraging stereopsis for saliency analysis. CVPR, 2012.

[18] Radhakrishna Achanta, et al. Slic superpixels compared to state-of-the-art superpixel methods. TPAMI, 2012.

[19] Ming-Ming Cheng, et al. Global contrast based salient region detection. CVPR, 2011.

Conclusion

This paper presented an object-based retrieval framework for stereo images. In the experiment of salient object segmentation, our approach captured more accurate images than other state-of-the-art methods, and in the evaluation of retrieval, benefited from the proposed salient object segmentation approach, the proposed framework outperformed other retrieval frameworks which were based on monoscopic images.

Research Team

Our research team, Multimedia Computing Group (MCG), belongs to the State Key Laboratory for Novel Software Technology, Nanjing University, China, and the leader is Professor Gangshan Wu. For more information, please contact Prof. Gangshan Wu (gswu@nju.edu.cn).