Faculty of Engineering and Information Technology
University of Technology Sydney

Image Co-saliency Detection and Co-segmentation from The Perspective of Commonalities

A thesis submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

by

Lina Li

March 2020

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This thesis is the result of me conducted jointly with Shanghai University as part of a collaborative Doctoral degree.

This research is supported by the Australian Government Research Training Program.

Production Note:
Signature removed prior to publication.

Signature of Candidate: _____

Date: 2020.03.09

Acknowledgments

Foremost, I would like to express the deepest appreciation to my supervisor, Associate Prof. Jian Zhang, for the constant support and guidance for my Ph.D work. As a collaborative degree candidate, I would also like to thank Prof. Zhi Liu, my co-supervisor, for his patient guidance, and scientific advice. Without their generous support, this dissertation would not have been possible.

I would like to thank my labmates in Global Big Data Technology Center: Junjie Zhang, Yifan Zuo, Xiaoshui Huang, Muming Zhao, Hao Cheng, Lu Zhang, Yongshun Gong, and Zhibin Li and labmates in Shanghai University, for the help and discussions and for all the fun we have had in my PhD study period.

Last but not the least, many thanks would go to my parents and my sisters, for their support and confidence in me through all these years.

Contents

List of Figures

List of Tables

List of Publications

Papers Published

- **Lina Li**, Zhi Liu, Wenbin Zou, Xiang Zhang, Olivier Le Meur(2014), Co-saliency detection based on region-level fusion and pixel-level refinement. *in* 'Proceedings of the International conference on Multimedia and Expo (ICME'14)', IEEE, pp. 1-6.

- **Lina Li**, Zhi Liu, Jian Zhang, (2018), Unsupervised image co-segmentation via guidance of simple images. *Neurocomputing*, vol. 275, pp. 1650-1661.

- **Lina Li**, Zhi Liu, Jian Zhang, Xiaofei Zhou, Learn Image Object Co-segmentation with Multi-scale Feature Fusion. *in* 'Visual Communications and Image Processing (VCIP'19)', IEEE, Accepted.

- Shuhua Luo, Zhi Liu, **Lina Li**, Xuemei Zou, Olivier Le Meur(2013), Efficient saliency detection using regional color and spatial information, *in* '4th European Workshop on Visual Information Processing (EUVIP'13)', pp. 184-189.

- Zhi Liu, Wenbin Zou, **Lina Li**, Liquan Shen, Olivier Le Meur(2014),Co-saliency detection based on hierarchical segmentation, *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 88-92.

- Linwei Ye, Zhi Liu, **Lina Li**(2015), Evaluation on fusion of saliency and objectness for salient object segmentation, *in* 'Proceedings of the 7th International Conference on Internet Multimedia Computing and Service(ICIMCS'15)', ACM, p.18.

- Linwei Ye, Zhi Liu, **Lina Li**, Liquan Shen, Cong Bai, Yang Wang(2017), Salient object segmentation via effective integration of saliency and objectness. *IEEE Transactions on Multimedia*, vol. 19, no. 8. pp. 1742-1756.

Abstract

Image co-saliency detection and image co-segmentation aim to identify the common salient objects and extract them in a group of images.

Image co-saliency detection and image co-segmentation are important for many content-based applications such as image retrieval, image editing, and content aware image/video compression. The image co-saliency detection and image co-segmentation are very close works. The most important part in these two works is the definition of the commonality of the common objects. Usually, common objects share similar low-level features, such as appearances, including colours, textures shapes, etc. as well as the high-level semantic features.

In this thesis, we explore the commonalities of the common objects in a group of images from low-level features and high-level features, and the way to achieve the commonalities and finally segment the common objects. Three main works are introduced, including an image co-saliency detection model and two image co-segmentation methods.

Firstly, an image co-saliency detection model based on region-level fusion and pixel-level refinement is proposed. The commonalities between the common objects are defined by the appearance similarities on the regions from all the images. It discovers the regions that are salient in each individual image as well as salient in the whole image group. Extensive experiments

on two benchmark datasets demonstrate that the proposed co-saliency model consistently outperforms the state-of-the-art co-saliency models in both subjective and objective evaluation.

Secondly, an unsupervised images co-segmentation method via guidance of simple images is proposed. The commonalities are still defined by handcrafted features on regions, colours and textures, but not calculated among regions from all the images. It takes advantages of the reliability of simple images, and successfully improves the performance. The experiments on the dataset demonstrate the outperformance and robustness of the proposed method.

Thirdly, a learned image co-segmentation model based on convolutional neural network with multi-scale feature fusion is proposed. The commonalities between objects are not defined by handcraft features but learned from the training data. When training a neural network with multiple input images simultaneously, the resource cost will increase rapidly with the inputs. To reduce the resource cost, reduced input size, less downsampling and dilation convolution are adopted in the proposed model. Experimental results on the public dataset demonstrate that the proposed model achieves a comparable performance to the state-of-the-art methods while the network has successfully gotten simplified and the resources cost is reduced.