Image segmentation using graph based community detection algorithms

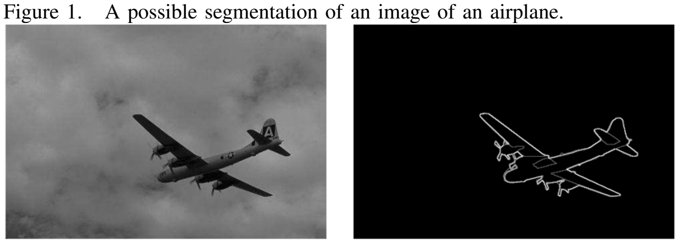
Pavlin Gregor Poličar
University of Ljubljana
Faculty of Computer and Information Science
Email: pavlin.g.p@gmail.com

***Abstract*—Image segmentation is a very broad field that combines methods from various different fields. Graph based techniques achieved some early successes in this field and are still in use today. However, community detection techniques have progressed significantly in the past few years, but none have been applied to this problem until quite recently, where the Louvain modularity optimization was applied and appeared to achieve reasonable results. However, we know that modularity optimization is simply one metric for evaluating community goodness, and others exists, most notably conductance. This raises the question, how well would other more recent community detection algorithms perform for image segmentation? In this work we examine more recent community detection algorithms and evaluate them on well known image segmentation benchmark datasets and provide a comparison to other state of the art computer vision segmentation algorithms.**

I. Introduction

The segmentation of an image into meaningful regions can be achieved through a number of methods. However, we have not found one single method that would outperform all else for different problems, therefore the landscape of image segmentation algorithms is vast.

Figure 1. A possible segmentation of an image of an airplane.



Graph-based techniques have proved to be effective in this domain, most notably – the normalized cut algorithm and its variants were proposed in 2000 [1], yet are still widely used today due to their simplicity and computation efficiency. Over the years, network based community detection methods have progressed, yet it is only quite recently that some of these methods have been tested for the task of image segmentation.

II. Related work

Image segmentation is a very broad field of study within computer vision with many different approaches and techniques, therefore we restrict ourselves mainly to the review graph-based procedures. These procedures generally build an undirected weighted graph G where each node represents a pixel in the input image. Each weighted edge of the graph G represents the similarity between a pair of pixels i and j and is stored in the weighted adjacency matrix W. Typically, we define w_{ij} as

$$w_{ij} = \begin{cases} e^{\frac{d(i,j)^2}{\sigma_x^2}} \, e^{\frac{|F(i)-F(j)|^2}{\sigma_i^2}} & \text{if } d(i,j) < d_{max}, \\ 0 & \text{otherwise} \end{cases}$$

where the term $d(i,j)$ represents spatial proximity using some distance metric between pixels i and j (most commonly the Euclidean distance) and $F(i)$ is some feature vector based on the properties of the pixel [2]. With d_{max} we specify an upper bound on pixel distances that we will consider. This parameter is necessary due to practical considerations, since the total number of connections between pixels is n^2, which is computationally unfeasible, therefore it is necessary to restrict ourselves to some field around individual pixels. In practice, this means that we will connect two pixels with a weighted edge if $d < d_{max}$.

In 2000, Shi & Malik proposed normalized cuts for image segmentation which attempt to find such a set of edges that creates disconnected components when removed from the graph, such that the sum of the weighted edges is minimal. These methods are typically referred to as *cuts* [1].

More recently, community detection algorithms have been tested for the image segmentation task. Browet et al. propose a variant of the Louvain method for greedy modularity optimization. The tests were performed on the Berkley segmentation benchmark dataset (BSDS) [3] and appears to work reasonably well for the images provided in the paper, however the method is never actually benchmarked to other image segmentation techniques and no score is provided [2]. An appropriate error metric may be *region covering* [4], [3] and is defined as

$$O(S_{seg}, S_{label}, p_i) = \frac{|C(S_{seg}, p_i) \setminus C(S_{hum}, p_i)|}{C(S_{seg}, p_i)}$$

where S_{seg} and S_{hum} are the machine segmentation and the human segmentation, respectively.

Another possible error metric could be *intersection over union* [4] which checks the overlap between the machine and human segmentation

$$IoU(S_{seg}, S_{hum}) = \frac{S_{seg} \cap S_{hum}}{S_{seg} \cup S_{hum}}$$

.

The method proposed by Browet et al. is a variant of the Louvain method for greedy modularity optimization. Modularity is a metric used to evaluate the quality of a community structure in a graph. The following is a generalization of the modularity metric adapted to weighted graphs.

$$Q = \frac{1}{2m} \sum_{i,j=1}^{n} (w_{ij} - N_{ij}) \, \delta_{C_i C_j}$$

where n is the number of nodes, w_{ij} is the edge weight, N_{ij} is a null model between nodes i and j and $\delta_{C_i C_j}$ is the Kronecker delta function. In this case, m is the sum of weights of all the edges in the graph and is a scaling parameter ensuring that $Q \in [-1, 1]$.

However, modularity is known to have problems with *resolution limit* [5], and other measures have been proposed, notably *conductance* by Leskovec et al. [6], so perhaps this was not the best community detection algorithm that could have been used. There have been many developments in community detection algorithms over the years that have not yet (to the best of our knowledge) been tested for image segmentation e.g. Infomap, stochastic block models or spectral methods.

III. Project proposal

Our work will focus on testing more recent developments in community detection algorithms for the task of image segmentation such as stochastic block models [7].

The stochastic block model is a generative model for learning community structure in unweighted networks. Since the graph-based approach to image segmentation requires weighted edges, we use a generalization of the stochastic block model presented by Aicher et al. [8]. Another problem we observe is that the stochastic block model is parameterized by K i.e. the number of blocks we expect to find must be provided in advance. This is clearly not suitable for our task. Recently, Peixoto proposed a nonparametric Bayesian formulation of weighted stochastic block models that can be used to infer the dimensions of the model from the data, requiring no prior knowledge about the data being observed.

Apart from these more recent community detection methods, we also consider Leighton-Rao [9] because it works very well for mesh-like graph structures [6]. It seems appropriate since the procedure that converts images to graphs produce mesh-like graphs with pixels being only locally connected.

We will benchmark our results with other state of the art image segmentation methods from other areas of computer vision e.g. affinity propagation [10], normalized graph cuts and fully convolutional networks [11] on the Berkley segmentation benchmark dataset [3] and the larger Pascal VOC 2012 dataset [12]. We will evaluate the examined methods using the union over intersection metric as described above.

References

[1] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[2] A. Browet, P.-A. Absil, and P. Van Dooren, "Community detection for hierarchical image segmentation.," in *IWCIA*, vol. 11, pp. 358–371, Springer, 2011.

[3] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th Int'l Conf. Computer Vision*, vol. 2, pp. 416–423, July 2001.

[4] H. Zhu, F. Meng, J. Cai, and S. Lu, "Beyond pixels: A comprehensive survey from bottom-up to semantic image segmentation and cosegmentation," *Journal of Visual Communication and Image Representation*, vol. 34, pp. 12–27, 2016.

[5] S. Fortunato and M. Barthélemy, "Resolution limit in community detection," *Proceedings of the National Academy of Sciences*, vol. 104, no. 1, pp. 36–41, 2007.

[6] J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney, "Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters," *Internet Mathematics*, vol. 6, no. 1, pp. 29–123, 2009.

[7] P. W. Holland, K. B. Laskey, and S. Leinhardt, "Stochastic blockmodels: First steps," *Social networks*, vol. 5, no. 2, pp. 109–137, 1983.

[8] C. Aicher, A. Z. Jacobs, and A. Clauset, "Learning latent block structure in weighted networks," *Journal of Complex Networks*, vol. 3, no. 2, pp. 221–248, 2014.

[9] T. Leighton and S. Rao, "An approximate max-flow min-cut theorem for uniform multicommodity flow problems with applications to approximation algorithms," in *Foundations of Computer Science, 1988., 29th Annual Symposium on*, pp. 422–431, IEEE, 1988.

[10] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *science*, vol. 315, no. 5814, pp. 972–976, 2007.

[11] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, 2015.

[12] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results." http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html.