ClusterFuG: Clustering Fully connected Graphs by Multicut

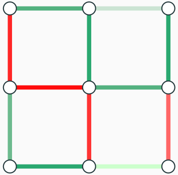
Ahmed Abbas[*], Paul Swoboda[*,†]

[*]Max Planck Institute for Informatics, Saarland Informatics Campus, Germany
[†]University of Mannheim, Germany

Multicut problem

- Given undirected graph $G = (V, E)$ with edge costs c
  - Attractive edge with $c > 0$
  - Repulsive edge with $c < 0$
- Find a clustering by

$$\min_{y \in \{0,1\} \cap \mathcal{M}_G} \sum_{ij \in E} c_{ij} y_{ij}$$

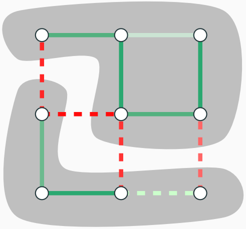


Multicut problem

- Given undirected graph $G = (V, E)$ with edge costs c
  - Attractive edge with $c > 0$
  - Repulsive edge with $c < 0$
- Find a clustering by

$$\min_{y \in \{0,1\} \cap \mathcal{M}_G} \sum_{ij \in E} c_{ij} y_{ij}$$

- Also known as weighted correlation clustering
- Many applications in computer vision such as neuron segmentation[1], instance segmentation[2]



[1] Andres 2012, [2] Gao 2019

Problem statement

- Graph structure needs to be known or hand designed
- Naively considering complete graph does not scale to large problems

Our contributions

- A novel formulation over complete graphs with costs for edge ij as

$$c_{ij} = \langle f_i, f_j \rangle$$

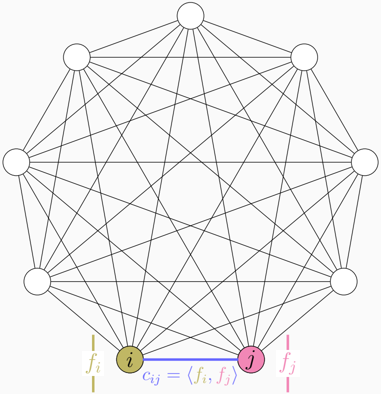
- Efficient and scalable algorithms for clustering of complete graphs
- Improved results on ImageNet clustering and panoptic segmentation benchmarks

Dense multicut problem

$$\min_{y \in \mathcal{M}_G} \sum_{i \in V} \sum_{j \in V \setminus i} \langle f_i, f_j \rangle y_{ij}$$

$\mathcal{M}_G$ constraints edge labels to be binary and consistent

- *Efficient representation:* Edge costs available on demand
- Node features f_i from machine learning model



Overall algorithm: Edge contraction for dense multicut

- Initialize nearest neighbour graph to find attractive node pairs
- Repeat until nearest neighbour graph has no attractive edges
    1. Contract most attractive edge
    2. Update nearest neighbour graph only at affected regions

Overall algorithm: Edge contraction for dense multicut

- Initialize nearest neighbour graph to find attractive node pairs
- Repeat until nearest neighbour graph has no attractive edges
  1. Contract most attractive edge
  2. Update nearest neighbour graph only at affected regions

Devise criteria for maximum reuse of nearest neighbour graph after contraction

Experiments

1. ImageNet clustering
   - ImageNet pretrained (unsupervised) features[1] to cluster validation set
   - Consider two problem splits containing $5k$, (100) and $50k$, (1000) images (classes)
   - *Baseline*: k-means with prespecified number of clusters

[1] Chen 2021, [2] Wang 2021, [3] Abbas 2021

Experiments

1. ImageNet clustering
   - ImageNet pretrained (unsupervised) features[1] to cluster validation set
   - Consider two problem splits containing $5k$, (100) and $50k$, (1000) images (classes)
   - *Baseline*: k-means with prespecified number of clusters
2. Panoptic segmentation on Cityscapes
   - Take pretrained Axial-ResNet50[2] features to cluster each instance class
   - *Baseline*: Conventional multicut on hand-designed graph structure[3]

[1] Chen 2021, [2] Wang 2021, [3] Abbas 2021

ImageNet clustering results

Method	t [s] ↓	NMI ↑	AMI ↑	# clusters
ImageNet-100 ($\|V\| = 5k$)				
k-means	16	0.42	0.27	100*
k-means	32	**0.53**	0.26	333*
Our	**3.2**	0.47	0.26	333†
ImageNet-1000 ($\|V\| = 50k$)				
k-means	701	0.54	0.2	1000*
k-means	1801	**0.61**	0.19	2440*
Our	**65**	0.56	**0.26**	2440†

$*$: Given as input, †: Determined automatically, NMI: Normalized mutual info., AMI: Adjusted mutual info.

Cityscapes panoptic segmentation

	Panoptic quality (%) ↑	
Category	Hand designed sparse graph	Dense multicut (Our)
Person	40.0	**46.9**
Rider	53.0	**54.4**
Car	50.7	**60.5**
Truck	**52.7**	52.3
Bus	**72.1**	71.1
Train	**65.6**	62.9
Motorcycle	**47.0**	46.8
Bicycle	45.7	**46.9**
All instance classes (PQ_{th})	53.3	**55.2**

Conclusion and outlook

1. Dense multicut formulation reducing human effort of graph structure design
2. Propose efficient algorithms
3. Improved results on two large scale clustering tasks from computer vision

Possible extensions

1. Methods for lower bound to the objective
2. Methods for backpropagation

References i

- Abbas, Ahmed (2021). “Combinatorial Optimization for Panoptic Segmentation: A Fully Differentiable Approach”. In: *Advances in Neural Information Processing Systems* 34, pp. 15635–15649.
- Andres, Bjoern (2012). “Globally Optimal Closed-Surface Segmentation for Connectomics”. In: *Computer Vision – ECCV 2012*. Ed. by Andrew Fitzgibbon. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 778–791. ISBN: 978-3-642-33712-3.
- Chen, Xinlei (2021). “An empirical study of training self-supervised vision transformers”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9640–9649.

Gao, Naiyu (2019). "SSAP: Single-shot instance segmentation with affinity pyramid". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 642–651.

Wang, Huiyu (2021). "Max-deeplab: End-to-end panoptic segmentation with mask transformers". In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5463–5474.