Correlation Clustering

April 12, 2013

Abbreviation:	correlation-clustering
Number of instances:	715
Number of variables:	∼300
Number of labels:	∼300
Number of factors:	∼300
Order:	∼300
Function type:	Potts

Description Correlation Clustering (CC) is a graph-partitioning algorithm that infers the edge labels of the graph by simultaneously maximizing intra-cluster similarity and inter-cluster dissimilarity by optimization of a global objective function. In order to capture long-range dependencies of distant nodes in a global context, [4] proposes higher-order CC to incorporate higher-order relations over hyper-graph, and the CC was shown to be effective in image segmentation [4]. For this image segmentation, [4] defined the CC problem over a superpixel-based hyper-graph in which an edge referred to as hyperedge can connect to two or more nodes (superpixels). For example, as shown in Figure 1, one can introduce binary labels for each adjacent vertices forming a triplet such that $x_{ijk} = 1$ if all vertices in $\{i, j, k\}$ are in the same cluster; otherwise, $x_{ijk} = 0$. Define a hyper-graph $\mathcal{HG} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ is the set of all nodes (superpixels) and $\mathcal{E}$ is the set of all hyperedges (subsets of $\mathcal{V}$) such that $\bigcup_{e\in\mathcal{E}} = \mathcal{V}$. Here, a hyperedge e has at least two nodes, i.e. $|e| \geq 2$. Therefore, the hyperedge set $\mathcal{E}$ can be divided into two disjoint subsets: pairwise edge set $\mathcal{E}_p = \{e \in \mathcal{E} \mid |e| = 2\}$ and higher-order edge set $\mathcal{E}_h = \{e \in \mathcal{E} \mid |e| > 2\}$ such that $\mathcal{E}_p \bigcup \mathcal{E}_h = \mathcal{E}$.

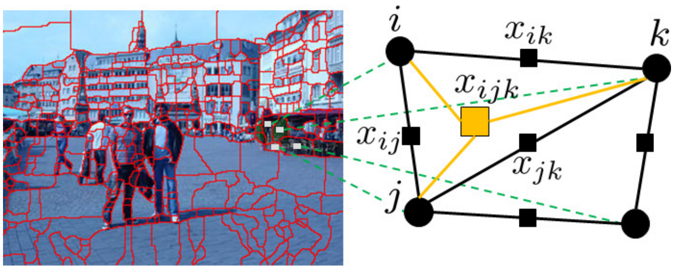

Figure 1: Illustration of a part of the triplet graph built on superpixels.

In this paper, we perform CC for image segmentation on the Stanford background dataset [3] (SBD). The SBD consists of 715 outdoor images with corresponding pixel-wise annotations such that each pixel is labeled with either one of 7 background classes or a generic foreground class. From the given pixel-wise ground-truth annotations, we obtain a ground-truth segmentation for each image. From a given image, a hyper-graph is constructed as follows. First, unsupervised multiple partitionings are obtained by merging not pixels but superpixels with different image quantizations using the ultrametric contour maps [1]. Then, the obtained regions are used to define hyperedges of the hyper-graph.

Objective / Learning The CC can be formulated as a linear objective function which allows for approximate polynomial-time inference by linear programming. An image segmentation is to infer the hyperedge label, $\hat{x}$, over the hyper-graph $\mathcal{HG}$ by minimizing J such that

$$\hat{x} = \operatorname*{argmin}_{x\in X(\mathcal{HG})} J(x) \quad (1)$$

$$= \operatorname*{argmin}_{x\in X(\mathcal{HG})} \sum_{e\in\mathcal{E}} \varphi_e(x_e) \quad (2)$$

$$= \operatorname*{argmin}_{x\in X(\mathcal{HG})} \sum_{e\in\mathcal{E}} \langle w, \phi_e\rangle x_e \quad (3)$$

$$= \operatorname*{argmin}_{x\in X(\mathcal{HG})} \sum_{e_p\in\mathcal{E}_p} \langle w_p, \phi_{e_p}\rangle x_{e_p} + \sum_{e_h\in\mathcal{E}_h} \langle w_h, \phi_{e_h}\rangle x_{e_h} \quad (4)$$

where $X(\mathcal{HG})$ is the set of $\{0, 1\}^{\mathcal{E}}$ that corresponds to a *valid segmentation* and the homogeneity measure among nodes in e, φ_e, is the inner product of the weight vector $w = [w_p; w_h]$ and the feature vector ϕ_e and takes values of both signs such that a large negative value indicates strong homogeneity while a large positive value indicates high degree of non-homogeneity. Here, we construct a 771-dimensional feature vector $\phi_e = [\phi_{e_p}; \phi_{e_h}]$ by concatenating several visual cues with different quantization levels and thresholds and estimate w by structured support vector machine [4]. The relaxed polytope to approximately solve (1) by linear programming is defined by the following three inequalities.

1. Cycle inequality: Let Path(j, k) be the set of paths between nodes j and k. The cycle inequality is a generalization of the triangle inequality [2] and is defined as

$$(1 - x_{jk}) \leq \sum_{(s,t)\in p} (1 - x_{st}), \quad p \in \text{Path}(j, k). \quad (5)$$

2. Odd-wheel inequality: Let a q-wheel be a connected subgraph $\mathcal{S} = (\mathcal{V}_s, \mathcal{E}_s)$ with a central vertex $j \in \mathcal{V}_s$ and a cycle of q vertices in $\mathcal{C} = \mathcal{V}_s \setminus \{j\}$. For every odd $q(\geq 3)$-wheel, a valid segmentation x satisfies

$$\sum_{(s,t)\in\mathcal{E}(\mathcal{C})} (1 - x_{st}) - \sum_{k\in\mathcal{C}} (1 - x_{jk}) \leq \lfloor \frac{1}{2} q \rfloor, \quad (6)$$

where $\mathcal{E}(\mathcal{C})$ denotes the set of all edges in the outer cycle $\mathcal{C}$.

3. Higher-order inequality:

$$x_{e_h} \leq x_{e_p}, \ \forall e_p \in \mathcal{E}_p | e_p \subset e_h, \qquad (7)$$
$$(1 - y_{e_h}) \leq \sum_{e_p \in \mathcal{E}_p | e_p \subset e_h} (1 - x_{e_p}).$$

The higher-order inequalities reflect valid relations between higher-order edge labels and pairwise edge labels.

References

[1] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33:898–916, 2011.

[2] S. Chopra and M. R. Rao. The partition problem. *Math. Program*, 59:87–115, 1993.

[3] Stephen Gould, Richard Fulton, and Daphne Koller. Decomposing a scene into geometric and semantically consistent regions. In *ICCV*, 2009.

[4] Sungwoong Kim, Sebastian Nowozin, Pushmeet Kohli, and Chang D. Yoo. Higher-order correlation clustering for image segmentation. In *NIPS*. 2011.