

Computer Vision II

Bjoern Andres

Machine Learning for Computer Vision
TU Dresden

2020-05-29

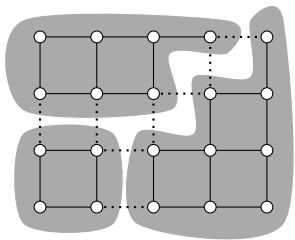
Joint pixel classification and image decomposition

- ▶ So far, we have studied
 - ▶ **pixel classification**, a problem whose feasible solutions define decisions at the pixels of an image
 - ▶ **image decomposition**, a problem whose feasible solutions decide whether pairs of pixels are assigned to the same or distinct components of the image.
- ▶ Applications exists (as we will see) for which both problems are too restrictive:
 - ▶ In **pixel classification**, there is no way of assigning neighboring pixels of the same class to distinct components of the image.
 - ▶ In **image decomposition**, there is no way of expressing that a unique decision shall be made for pixels that belong to the same component of the image.

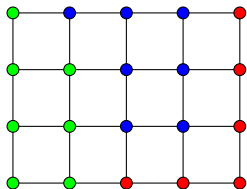


M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele. The Cityscapes Dataset for Semantic Urban Scene Understanding. CVPR 2016. See also: <https://www.cityscapes-dataset.com/>

- ▶ One application where a joint generalization of pixel classification and image decomposition is useful is called **semantic image segmentation**.
- ▶ In the above image, thin boundaries are left between pixels of the same class (e.g. pedestrian) that belong to different instances of the class (e.g. distinct pedestrians).
- ▶ Next, we are going to introduce a strict generalization of both, pixel classification and image decomposition that does not require these boundaries.

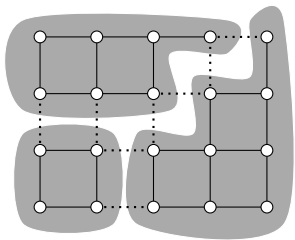


Graph Decomposition

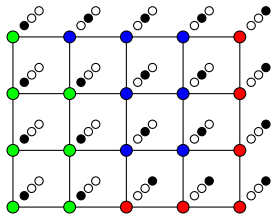


Node Labeling

We state an optimization problem whose feasible solutions define both, a decomposition of a graph $G = (V, E)$ and a labeling $l: V \rightarrow L$ of its nodes.



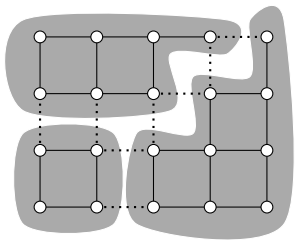
Graph Decomposition



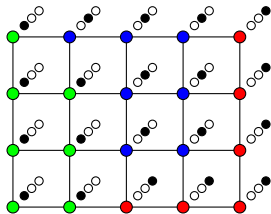
Node Labeling

We encode every feasible node labeling in a binary vector from the set

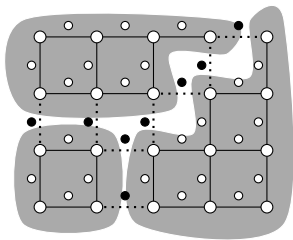
$$Y_{VL} := \left\{ y : V \times L \rightarrow \{0, 1\} \mid \forall v \in V : \sum_{l \in L} y_{vl} = 1 \right\}$$



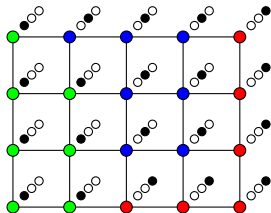
Graph Decomposition



Node Labeling



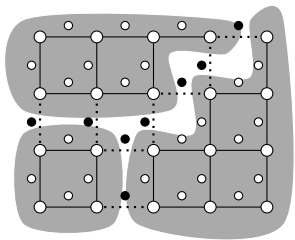
Graph Decomposition



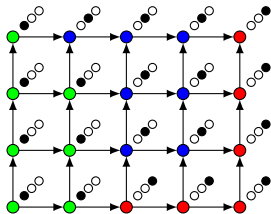
Node Labeling

We encode every feasible graph decomposition by the characteristic function of the multicut it induces:

$$X_G := \left\{ x : E \rightarrow \{0, 1\} \mid \forall C \in \text{cycles}(G) \forall e \in C : x_e \leq \sum_{f \in C \setminus \{e\}} x_f \right\}$$

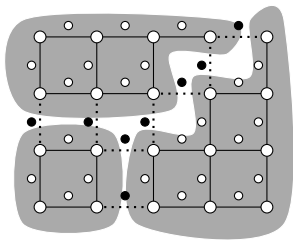


Graph Decomposition

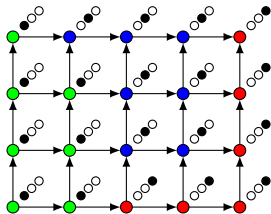


Node Labeling

We choose an arbitrary orientation (V, A) of the edges E , i.e., for each $v, w \in V$, we have $\{v, w\} \in E$ if and only if either $(v, w) \in A$ or $(w, v) \in A$.



Graph Decomposition



Node Labeling

W.r.t. the orientation (V, A) of the graph $G = (V, E)$, the set L of labels, any (costs) $c: V \times L \rightarrow \mathbb{R}$ and any (costs) $c', c'': A \times L^2 \rightarrow \mathbb{R}$, the instance of the joint graph decomposition and node labeling problem has the form

$$\min_{(x,y) \in X_G \times Y_{VL}} \sum_{v \in V} \sum_{l \in L} c_{vl} y_{vl} + \sum_{(v,w) \in A} \sum_{(l,l') \in L^2} c'_{vwl'l'} y_{vl} y_{wl'} x_{\{v,w\}} + \sum_{(v,w) \in A} \sum_{(l,l') \in L^2} c''_{vwl'l'} y_{vl} y_{wl'} (1 - x_{\{v,w\}})$$

Suggested self-study:

- ▶ Generalize your implementation of a local search algorithm for the image decomposition problem such that it becomes applicable to the joint pixel classification and image decomposition problem
 1. by alternating between transformations of the labeling and transformations of the decomposition
 2. by searching, whenever a node is moved from one component to another, all possible labels