Computer Vision I

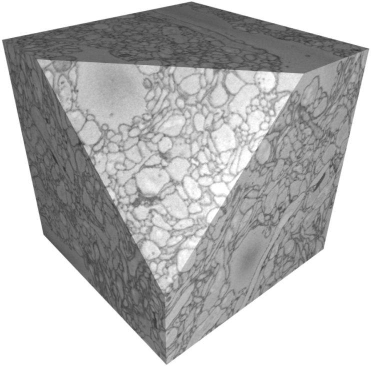
Bjoern Andres, Holger Heidrich

TU Dresden

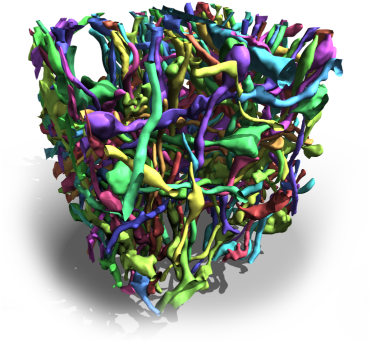
Winter Term 2021/2022

- So far, we have studied **pixel classification**, a problem whose feasible solutions define decisions at the pixels of an image
- Next, we will study **image decomposition**, a problem whose feasible solutions decide whether pairs of pixels are assigned to the same or distinct components of the image
- Image decomposition has applications where components of the image are indistinguishable by appearance (see next slide)

Image decomposition



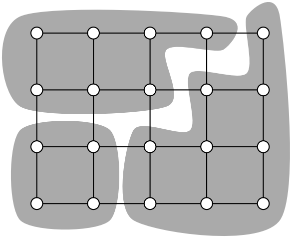
$\mapsto$



Volume Image (32 nm/voxel)
(Denk and Horstmann, 2004)

Decomposition
(Andres et al., 2012)

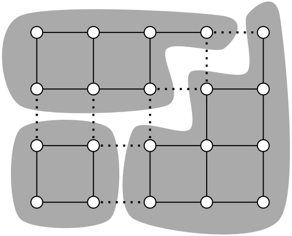
Image decomposition



Decomposition of a graph $G = (V, E)$

- A mathematical abstraction of a decomposition of an image is a decomposition of the pixel grid graph.
- A decomposition of a graph is a partition of the node set into connected subsets (one example is depicted above in gray).

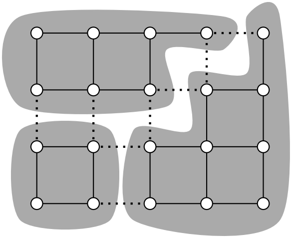
Image decomposition



Decomposition of a graph $G = (V, E)$

- A decomposition of a graph is characterized by the set of edges that straddle distinct components (depicted above as dotted lines)
- Those subsets of edges are called **multicuts** of the graph

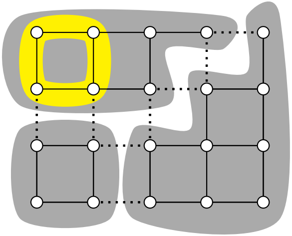
Image decomposition



Multicut of a graph $G = (V, E)$

- A decomposition of a graph is characterized by the set of edges that straddle distinct components (depicted above as dotted lines)
- Those subsets of edges are called **multicuts** of the graph

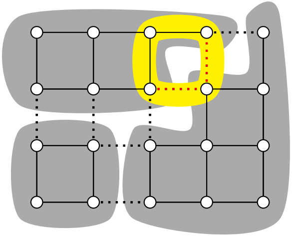
Image decomposition



Multicut of a graph $G = (V, E)$

- The defining property of multicuts is that no cycle in the graph intersects with the multicut in precisely one edge

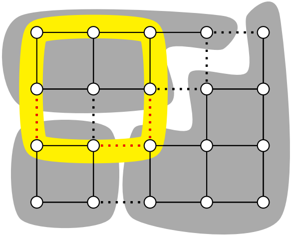
Image decomposition



Multicut of a graph $G = (V, E)$

- The defining property of multicuts is that no cycle in the graph intersects with the multicut in precisely one edge

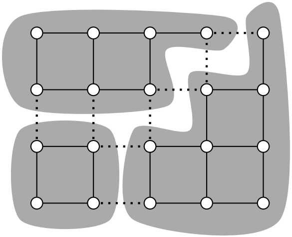
Image decomposition



Multicut of a graph $G = (V, E)$

- The defining property of multicuts is that no cycle in the graph intersects with the multicut in precisely one edge

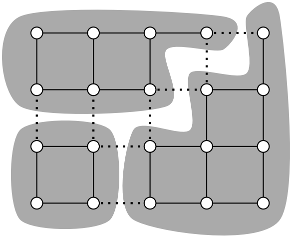
Image decomposition



Multicut of a graph $G = (V, E)$

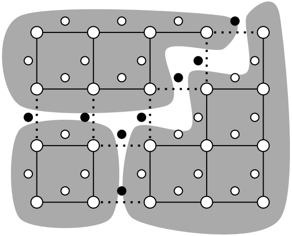
$$\text{multicuts}(G) := \{M \subseteq E \,|\, \forall C \in \text{cycles}(G) :\ |M \cap C| \neq 1\}$$

Image decomposition



Multicut of a graph $G = (V, E)$

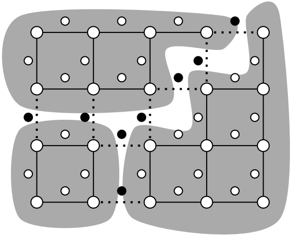
Image decomposition



Multicut of a graph $G = (V, E)$

- The characteristic function $y \colon E \to \{0, 1\}$ of a multicut $y^{-1}(1)$ can be used to encode the decomposition induced by the multicut in an $|E|$-dimensional 01-vector
- For any $e \in E$, $y_e = 1$ indicates that an edge is cut, straddling distinct components

Image decomposition

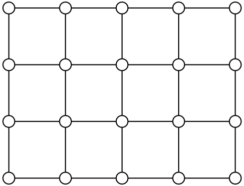


Multicut of a graph $G = (V, E)$

- The set of the characteristic functions of all multicuts of G:

$$Y_G := \left\{ y : E \to \{0,1\} \;\middle|\; \forall C \in \mathsf{cycles}(G)\, \forall e \in C : \; y_e \leq \sum_{f \in C \setminus \{e\}} y_f \right\}$$

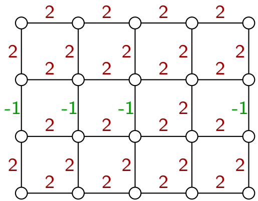
Image decomposition



Graph $G = (V, E)$

- An instance of the image decomposition problem is given by a graph $G = (V, E)$ and, for every edge $e = \{v, w\} \in E$, a (positive or negative) cost $c_e \in \mathbb{R}$ that is payed iff the incident pixels v and w are put in distinct components
- Such costs are often estimated from examples using machine learning technqiues

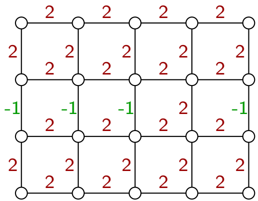
Image decomposition



Graph $G = (V, E)$. Edge costs $c : E \to \mathbb{R}$

- An instance of the image decomposition problem is given by a graph $G = (V, E)$ and, for every edge $e = \{v, w\} \in E$, a (positive or negative) cost $c_e \in \mathbb{R}$ that is payed iff the incident pixels v and w are put in distinct components
- Such costs are often estimated from examples using machine learning technqiues

Image decomposition



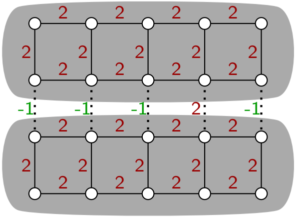
Graph $G = (V, E)$. Edge costs $c : E \to \mathbb{R}$

- Image decomposition problem:

$$\min_{y \in Y_G} \sum_{e \in E} c_e \, y_e$$

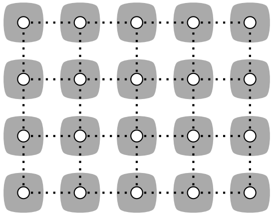
- The optimal solution is shown in the next slide

Image decomposition



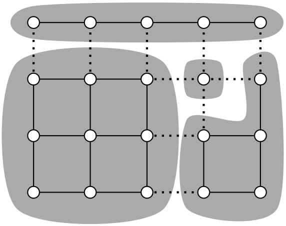
Graph $G = (V, E)$. Edge costs $c : E \to \mathbb{R}$

Image decomposition



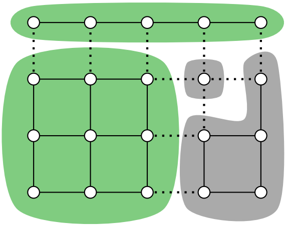
- One technique for finding feasible solutions to an image decomposition problem is **local search**.
- Starting from the finest decomposition into singleton components (depicted above), we greedily join neighboring components as long as this improves the cost (see next slide).

Image decomposition



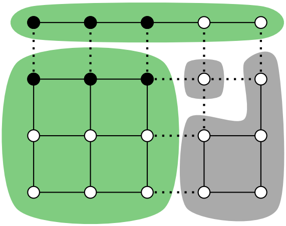
- Once no joining of neighboring components further reduces the cost, we consider all pairs of neighboring components (depicted in green) and all nodes at the shared boundary (depicted in black) and all possibilities of moving nodes from one component to the other.
- The procedure is iterated until no such transformation further reduces the cost

Image decomposition



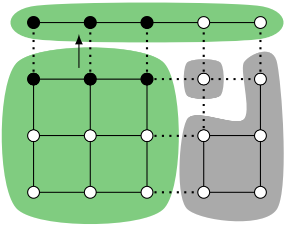
- Once no joining of neighboring components further reduces the cost, we consider all pairs of neighboring components (depicted in green) and all nodes at the shared boundary (depicted in black) and all possibilities of moving nodes from one component to the other.
- The procedure is iterated until no such transformation further reduces the cost

Image decomposition



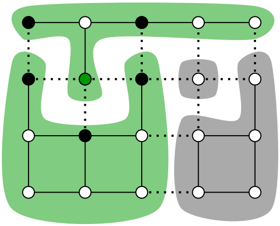
- Once no joining of neighboring components further reduces the cost, we consider all pairs of neighboring components (depicted in green) and all nodes at the shared boundary (depicted in black) and all possibilities of moving nodes from one component to the other.
- The procedure is iterated until no such transformation further reduces the cost

Image decomposition



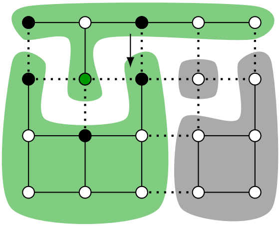
- Once no joining of neighboring components further reduces the cost, we consider all pairs of neighboring components (depicted in green) and all nodes at the shared boundary (depicted in black) and all possibilities of moving nodes from one component to the other.
- The procedure is iterated until no such transformation further reduces the cost

Image decomposition



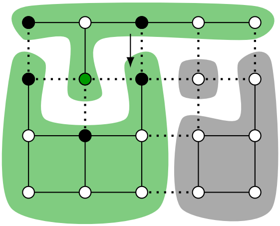
- Once no joining of neighboring components further reduces the cost, we consider all pairs of neighboring components (depicted in green) and all nodes at the shared boundary (depicted in black) and all possibilities of moving nodes from one component to the other.
- The procedure is iterated until no such transformation further reduces the cost

Image decomposition



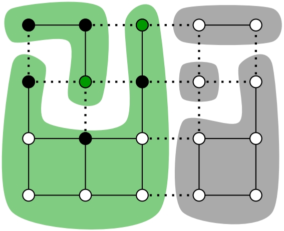
- Once no joining of neighboring components further reduces the cost, we consider all pairs of neighboring components (depicted in green) and all nodes at the shared boundary (depicted in black) and all possibilities of moving nodes from one component to the other.
- The procedure is iterated until no such transformation further reduces the cost

Image decomposition



- Once no joining of neighboring components further reduces the cost, we consider all pairs of neighboring components (depicted in green) and all nodes at the shared boundary (depicted in black) and all possibilities of moving nodes from one component to the other.
- The procedure is iterated until no such transformation further reduces the cost

Image decomposition



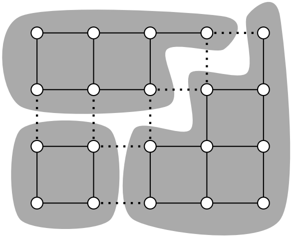
- Once no joining of neighboring components further reduces the cost, we consider all pairs of neighboring components (depicted in green) and all nodes at the shared boundary (depicted in black) and all possibilities of moving nodes from one component to the other.
- The procedure is iterated until no such transformation further reduces the cost

Joint pixel classification and image decomposition

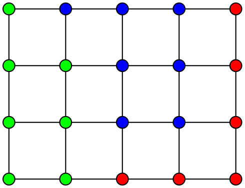
- So far, we have studied
  - **pixel classification**, a problem whose feasible solutions define decisions at the pixels of an image
  - **image decomposition**, a problem whose feasible solutions decide whether pairs of pixels are assigned to the same or distinct components of the image.
- Applications 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.

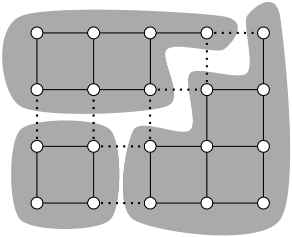


Graph Decomposition

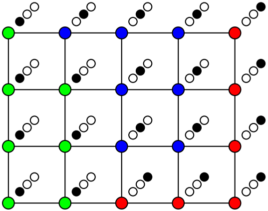


Node Labeling

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



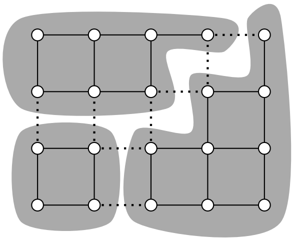
Graph Decomposition



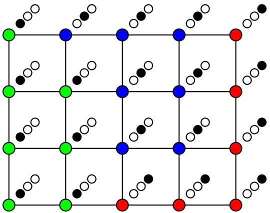
Node Labeling

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

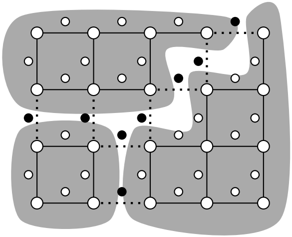
$$Y_{VL} := \left\{ y : V \times L \to \{0,1\} \;\middle|\; \forall v \in V : \sum_{l \in L} y_{vl} = 1 \right\}$$



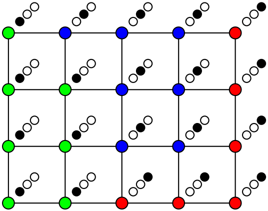
Graph Decomposition



Node Labeling



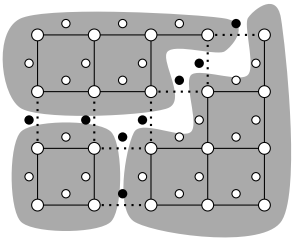
Graph Decomposition



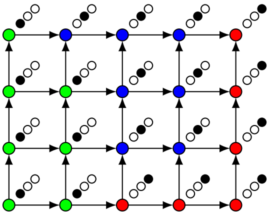
Node Labeling

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

$$X_G := \left\{ x : E \to \{0,1\} \;\middle|\; \forall C \in \mathsf{cycles}(G)\, \forall e \in C\colon\; x_e \leq \sum_{f \in C \setminus \{e\}} x_f \right\}$$

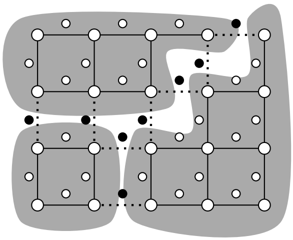


Graph Decomposition

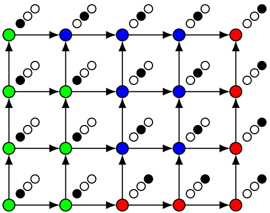


Node Labeling

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) $c\colon V \times L \to \mathbb{R}$ and any (costs) $c', c''\colon A \times L^2 \to \mathbb{R}$, the instance of the joint graph decomposition and node labeling problem has the form

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