Computer Vision II

Bjoern Andres

Machine Learning for Computer Vision
TU Dresden

2020-05-08

Pixel classification

- So far, we have studied a local search algorithm for the smooth pixel classification problem.
- On the one hand, this algorithm is easy to implement and has straight-forward generalizations, e.g., to the case of more than two classes.
- On the other hand, it does not necessarily solve smooth pixel classification with two classes to optimality.
- Next, we will reduce the smooth pixel classification problem with two classes to the well-known **minimum st-cut problem** that can be solved exactly and efficiently.
- The notes are organized as follows
  - Definition of the minimum st-cut problem
  - Submodularity
  - Reduction of the smooth pixel classification problem

Definition 1

A 5-tuple $N = (V, E, s, t, \gamma)$ is called a **network** *iff (V, E) is a directed graph and $s \in V$ and $t \in V$ and $s \neq t$ and $\gamma : E \to \mathbb{R}_0^+$.*

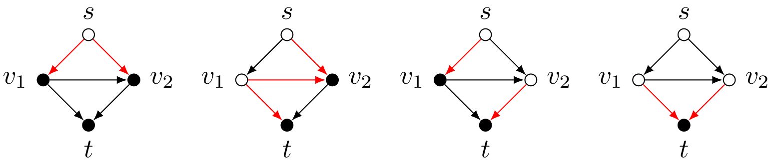
The nodes s and t are called the **source** *and the* **sink** *of N, respectively.*

For any edge $e \in E$, γ_e is called the **capacity** *of e in N.*

Definition 2

Let (V, E) *be a directed graph. Let* $s \in V$ *and* $t \in V$ *and* $s \neq t$.

- $X \subseteq V$ *is called an* st-**cutset** *of* (V, E) *iff* $s \in X$ *and* $t \notin X$.
- $Y \subseteq E$ *is called an* st-**cut** *of* (V, E) *iff there exists an* st*-cutset* X *such that* $Y = \{vw \in E \mid v \in X \wedge w \notin X\}$.



Definition 3

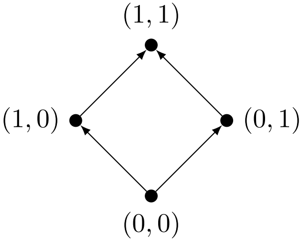
The instance of the **Minimum** st**-Cut Problem** *w.r.t. a network* $N = (V, E, s, t, \gamma)$ *is to*

$$\begin{aligned} \min_{x \in \{0,1\}^V} \quad & \sum_{vw \in E} x_v \left(1 - x_w\right) \gamma_{vw} && (1) \\ \text{subject to} \quad & x_s = 1 && (2) \\ & x_t = 0 && (3) \end{aligned}$$

Definition 4

A **lattice** $(S, \preceq)$ *is a set* S*, equipped with a partial order* $\preceq$*, such that any two elements of* S *have an infimum and a supremum w.r.t.* $\preceq$*.*

Example. $(\{0,1\}^2, \preceq)$ with $\preceq \, := \{(s,t) \in S \times S \mid s_1 \leq t_1 \wedge s_2 \leq t_2\}$.



For any $s, t \in \{0,1\}^2$,

$$\sup(s,t) = (\max\{s_1, t_1\}, \max\{s_2, t_2\})$$
$$\inf(s,t) = (\min\{s_1, t_1\}, \min\{s_2, t_2\})$$

Definition 5

A function $f : S \to \mathbb{R}$ *is called* **submodular** *w.r.t. a lattice* $(S, \preceq)$ *iff*

$$\forall s, t \in S \qquad f(\inf(s,t)) + f(\sup(s,t)) \leq f(s) + f(t) \ . \tag{4}$$

Lemma 1

For any $f : \{0,1\}^2 \to \mathbb{R}$, the following statements are equivalent.

1. *f is is submodular w.r.t. the the lattice $(\{0,1\}^2, \preceq)$*
2. $f(0,0) + f(1,1) \leq f(1,0) + f(0,1)$
3. *The unique form*

$$c_\emptyset + c_{\{1\}}x_1 + c_{\{2\}}x_2 + c_{\{1,2\}}x_1x_2$$

of f is such that $c_{\{1,2\}} \leq 0$.

Proof.

- $f(0,0) + f(1,1) \leq f(1,0) + f(0,1)$ is the only condition in

$$\forall s, t \in S \qquad f(\inf(s,t)) + f(\sup(s,t)) \leq f(s) + f(t)$$

  which is not generally true. Thus, (1.) is equivalent to (2.).

- We have

$$\begin{aligned} f(0,0) &= c_{\emptyset} \\ f(1,0) &= c_{\emptyset} + c_{\{1\}} \\ f(0,1) &= c_{\emptyset} \qquad\quad + c_{\{2\}} \\ f(1,1) &= c_{\emptyset} + c_{\{1\}} + c_{\{2\}} + c_{\{1,2\}} \enspace . \end{aligned}$$

  Therefore,

$$c_{\{1,2\}} = f(1,1) - f(1,0) - f(0,1) + f(0,0)$$

  and thus, (2.) is equivalent to (3.).

Lemma 2

The sum of finitely many submodular functions is submodular.

Lemma 3

For every $f : \{0,1\}^2 \to \mathbb{R}$, *there exist unique* $a_0 \in \mathbb{R}$ *and* $a_1, a_{\bar{1}}, a_2, a_{\bar{2}}, a_{12}, a_{\bar{1}2} \in \mathbb{R}_0^+$ *such that*

$$a_1 a_{\bar{1}} = a_2 a_{\bar{2}} = a_{12} a_{\bar{1}2} = 0 \tag{5}$$

and

$$\begin{aligned} \forall x \in \{0,1\}^2 \quad f(x) = \quad & a_0 \\ & + a_1 x_1 + a_{\bar{1}}(1 - x_1) \\ & + a_2 x_2 + a_{\bar{2}}(1 - x_2) \\ & + a_{12} x_1 x_2 + a_{\bar{1}2}(1 - x_1) x_2 \ . \end{aligned} \tag{6}$$

Proof.

- Comparison of (6) with the unique form

$$c_\emptyset + c_{\{1\}} x_1 + c_{\{2\}} x_2 + c_{\{1,2\}} x_1 x_2$$

  yields

$$\begin{aligned} a_0 + a_{\bar{1}} + a_{\bar{2}} &= c_\emptyset \\ a_1 - a_{\bar{1}} &= c_{\{1\}} \\ a_2 - a_{\bar{2}} + a_{\bar{1}2} &= c_{\{2\}} \\ a_{12} - a_{\bar{1}2} &= c_{\{1,2\}} \end{aligned} \tag{7}$$

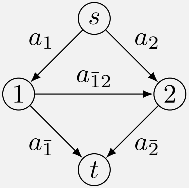
- By these equations (from bottom to top), (5) and c define a uniquely.

Lemma 4 (Kolmogorov and Zabih)

For every **submodular** $f : \{0,1\}^2 \to \mathbb{R}$ *and its unique coefficient* $a_0 \in \mathbb{R}$ *from Lemma 3,*

$$\min_{x \in \{0,1\}^2} f_x - a_0 \tag{8}$$

is equal to the weight of a **minimum** st-**cut** *in the graph below whose edge weights are the (unique, non-negative) coefficients from Lemma 3.*



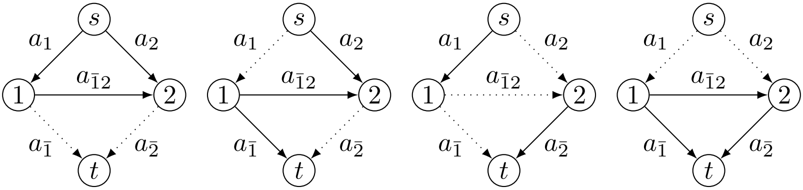
Moreover, f *is minimal at* $\hat{x} \in \{0,1\}^2$ *iff* $\{j \in \{1,2\} \mid \hat{x}_j = 0\}$ *is a* **minimum** st-**cutset** *of the above graph.*

Proof.

- Submodularity of f implies $a_{12} = 0$ in (7), by Lemma 1 and (5).
- Comparison of the four possible minima of f,

$$
\begin{aligned}
f(0,0) &= a_0 + a_{\bar{1}} + a_{\bar{2}} \\
f(1,0) &= a_0 + a_1 + a_{\bar{2}} \\
f(0,1) &= a_0 + a_{\bar{1}} + a_2 + a_{\bar{1}2} \\
f(1,1) &= a_0 + a_1 + a_2 + a_{12} \quad ,
\end{aligned}
$$

with the four possible minimum cuts below proves the Lemma.



Definition 6

For any smooth pixel classification problem

$$\min_{y \in \{0,1\}^V} \quad \underbrace{\sum_{v \in V} c_v \, y_v + \sum_{\{v,w\} \in E} c'_{\{v,w\}} \, |y_v - y_w|}_{\varphi(y)} \tag{9}$$

the **induced minimum** st**-cut problem** *is defined by the network* (V', E', s, t, γ) *such that* $V' = V \cup \{s, t\}$*,*

$$\begin{aligned} E' = &\{(s,v) \in V'^2 \mid c_v > 0\} \cup \{(v,t) \in V'^2 \mid c_v < 0\} \\ &\cup \{(v,w) \in V'^2 \mid \{v,w\} \in E\} \end{aligned} \tag{10}$$

and $\gamma \colon E' \to \mathbb{R}_0^+$ *such that*

$$\forall (s,v) \in E' \colon \quad \gamma_{(s,v)} = c_v \tag{11}$$

$$\forall (v,t) \in E' \colon \quad \gamma_{(v,t)} = -c_v \tag{12}$$

$$\forall \{v,w\} \in E \colon \quad \gamma_{(v,w)} = \gamma_{(w,v)} = c'_{\{v,w\}} \quad . \tag{13}$$

Lemma 5

For any smooth pixel classification problem w.r.t. a pixel grid graph $G = (V, E)$ and the induced minimum st-cut problem with the network (V', E', s, t, γ), $\hat{y} : V \to \{0, 1\}$ is an optimal pixel classification iff $\{v \in V \mid \hat{y}_v = 0\}$ is an optimal st-cutset.

Proof (sketch). The function φ is submodular, by Lemma 2 and $c' > 0$. The statement holds by Lemma 3 and the fact that for all $y \in \{0, 1\}^V$:

$$\varphi(y) = \sum_{v \in V} c_v \, y_v + \sum_{\{v,w\} \in E} c'_{\{v,w\}} \left(y_v(1 - y_w) + (1 - y_v) y_w \right) \quad .$$

Suggested self-study:

- Solve the smooth pixel classification problems $-2y_1 + 3y_2 + c|y_1 - y_2|$ for $c \in \{1, 5\}$ via the induced minimum st-cut problem
- Implement a solver for the smooth pixel classification problem using any existing implementation of any algorithm for the minimum st-cut problem[1].
- Apply this algorithm to the pixel classification problem from the previous lecture
- Compare the classifications $y \in \{0,1\}^V$ and objective values $\varphi(y)$ found by local search with those found by minimum st-cuts
- Share your results using OPAL.

[1] Alternatively, use any algorithm for computing a maximum st-flow, e.g. `https://www.boost.org/doc/libs/1_48_0/libs/graph/doc/edmonds_karp_max_flow.html`, and consider the minimum st-cutset of all nodes reachable from the source node in the residual network.