Weekly Exercise 6

Dr. Csaba Domokos and Lingni Ma
Technische Universität München, Computer Vision Group
May 24, 2016 (submission deadline: May 31, 2016)

Metrics (3 points)

Exercise 1 (**Metric, semi-metric**, 3 points)**.** Show that the followings hold:

1. The *truncated absolute distance*, defined as $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^+$

$$d(x, y) = \min(K, |x - y|), \quad \text{for some } K \in \mathbb{R}^+ ,$$

is a metric.

2. The *truncated quadratic function*, defined as $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^+$

$$d(x, y) = \min(K, |x - y|^2), \quad \text{for some } K \in \mathbb{R}^+ ,$$

is a semi-metric.

3. The *weighted Potts-model*, defined as $d : \mathbb{N} \times \mathbb{N} \to \mathbb{R}^+$

$$d(\ell_1, \ell_2) = w[\![\ell_1 \neq \ell_2]\!], \quad \text{for some } w \in \mathbb{R}^+ ,$$

is a metric.

Solution.

1. Let $d(x, y) = \min(K, |x - y|)$, the following holds,

$$d(x, x) = \min(K, |x - x|) = \min(K, 0) = 0, \forall K \in \mathbb{R}^+ \tag{1a}$$

$$d(x, y) = \min(K, |x - y|) = \min(K, |y - x|) = d(y, x) \tag{1b}$$

first, we have $\min(K, |x - y|) + \min(K, |y - z|) \geq \min(K, |x - y| + |y - z|)$
if $y \leq x \leq z, |x - y| + |y - z| = 2|x - y| + |z - x| \geq |x - z|$
if $x \leq y \leq z, |x - y| + |y - z| = |x - z|$
if $x \leq z \leq y, |x - y| + |y - z| = 2|y - z| + |z - x| \geq |x - z|$
in all cases $|x - y| + |y - z| \geq |x - z|$
therefore, $\min(K, |x - y| + |y - z|) \geq \min(K, |x - z|)$
therefore, $\min(K, |x - y|) + \min(K, |y - z|) \geq \min(K, |x - z|)$ (1c)

2. Let $d = (x, y) = \min(K, |x - y|^2)$, $K \in \mathbb{R}^+$, the following holds,

$$d(x, x) = \min(K, |x - x|^2) = \min(K, 0) = 0 \tag{2a}$$

$$d(x, y) = \min(K, |x - y|^2) = \min(K, |y - x|^2) = d(y, x) \tag{2b}$$

   Therefore, $d(x, y)$ is a semi-metric.

3. Let $d(\ell_1, \ell_2) = w[\![\ell_1 \neq \ell_2]\!]$, the following holds,

$$d(\ell_1, \ell_1) = 0 \tag{3a}$$

$$d(\ell_1, \ell_2) = w[\![\ell_1 \neq \ell_2]\!] = w[\![\ell_2 \neq \ell_1]\!] = d(\ell_2, \ell_1) \tag{3b}$$

   a) if $\ell_1 = \ell_2 = \ell_3 \Rightarrow d(\ell_1, \ell_2) + d(\ell_2, \ell_3) = 0 = d(\ell_1, \ell_3)$
   b) if otherwise $\Rightarrow d(\ell_1, \ell_2) + d(\ell_2, \ell_3) \geq w$ and $d(\ell_1, \ell_3) \leq w$
   in both cases, $d(\ell_1, \ell_2) + d(\ell_2, \ell_3) \geq d(\ell_1, \ell_3)$ (3c)

   Therefore, $d(\ell_1, \ell_2)$ is a metric.

Programming (10 points)

Exercise 2 (**Binary image segmentation with maxflow**, 5 points)**.** Solve the binary image segmentation problem for figure 1 by applying the maximum flow algorithm. For binary segmentation, we define $y_i \in \{0, 1\}$, where 0 denotes the background and 1 denotes the foreground. Let us consider the following energy function,

$$E(\mathbf{y}, \mathbf{x}) = \sum_{i \in \mathcal{V}} E_i(y_i; x_i) + w \sum_{i,j \in \mathcal{E}} E_{ij}(y_i, y_j; x_i, x_j) \quad , \tag{4}$$

where $w \in \mathbb{R}^+$ is a parameter, and $\mathcal{V}$ stand for the set of pixels and $\mathcal{E}$ includes 4-neighboring pixels.

Use the GMM models you have trained in Exercise 4 to define the unary energy functions for all $i \in \mathcal{V}$:

$$E_i(y_i = 0) = -\log(p_B(x_i))$$
$$E_i(y_i = 1) = -\log(p_F(x_i)) \ .$$

Use the contrast-sensitive Potts-model to define the pairwise energy functions for all $(i, j) \in \mathcal{E}$:

$$E_{ij}(y_i, y_j; x_i, x_j) = \exp(-\lambda \|x_i - x_j\|^2)[\![y_i \neq y_j]\!] \quad , \tag{5}$$

where x_i is the intensity vector for pixel i, and you may choose $\lambda = 0.5$.

To solve the maximum flow problem, you can use the provided maxflow package `maxflow-v3.04.src.zip`.

- Choose a set of different values for w, and report what you observe.

Figure 1: The test image for binary image segmentation.

- How are the segmentation results compared to the results you obtained in Exercise 4?

Exercise 3 (**multi-class image segmentation with α-expansion**, 5 points)**.** Consider the energy function

$$E(\mathbf{y}, \mathbf{x}) = \sum_{i \in \mathcal{V}} E_i(y_i; x_i) + w \sum_{i,j \in \mathcal{E}} E_{ij}(y_i, y_j; x_i, x_j) \quad , \tag{6}$$

for the multi-class labeling problem, i.e. $\mathbf{y} \in \mathcal{L}^{\mathcal{V}}$, where $\mathcal{L}$ stands for the label set. Implement the α-expansion algorithm to solve the image segmentation for the images shown in figure 2.

The test images come from a subset of the MSRC image understanding dataset, which contains 21 classes, i.e. $\mathcal{L} = \{1, 2, \ldots, 21\}$.

To define the unary energy functions E_i, use the provided `*.c_unary` files. Each test image has its own unary file, specified by the same filename. In each unary file, you can read out a $K \times H \times W$ array of float numbers. The H and W are the image height and width, and $K = 21$ is the number of possible classes. This array contains the 21-class probability distribution for each pixel. We provide the `multilabel_demo.cpp` to demonstrate how to load the unary file and read out the corresponding probability values. The unary energy functions E_i for all $i \in \mathcal{V}$ are then defined as

$$E_i(y_i = l) = -\log(p_l) \quad .$$

The pairwise energy is again defined by the contrast sensitive Potts-model (see Equation (5)).

- Choose different w for Equation (6) and compare the segmentation results.
- The meaning of the classes are specified in the `21class.txt` file. Use it to check if your results make sense.

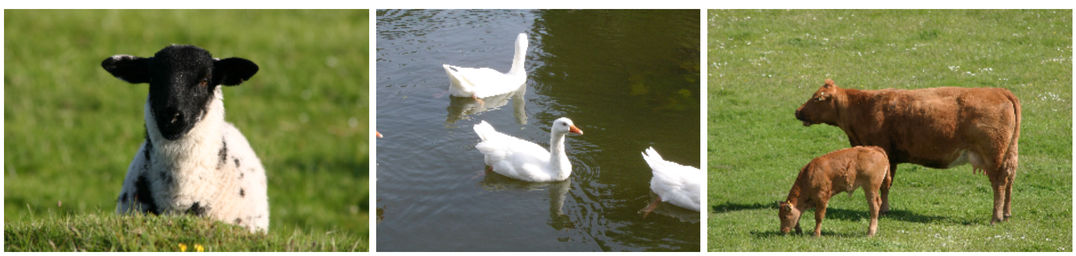

Figure 2: The target images for multi-class image segmentation.