Assignment 1

June 29, 2017

Task 1. **Maximum Likelihood (difficulty 1).** Given a set of noisy observations of a template J, estimate the mean template and noise variance.

- Input: a sequence of images of the object, array JJ.

Task 2. **Bayesian Decision Theory (difficulty 2).** Consider the *Template Matching Example* from the lecture, compute optimal Bayesian decision for

a) Box loss:

$$W(k, k^*) = \begin{cases} 0, & \text{if } |k - k^*| < 10; \\ 1, & \text{otherwise.} \end{cases}$$

b) Grid loss: for a set of quadrants forming a partition of all positions $(K_1, \ldots, K_l)$, the goal is to find i such that $k \in K_i$. What is the loss function?

Input:

- Observed image I;
- Template image J;
- Detected vertical position ky;
- Noise level σ;
- The matrix K of size $l \times N$, in which row j contains the mask of the set K_j of the grid loss.
- The data consists of two images `background.png`, `template.png`, a matrix $\mathtt{K} \in \mathbb{N}^{(l \times 2)}$ describing the partition (for grid loss) and a variable `vpos` indicating the vertical position of the template.
- Compute and plot posterior probability of tank at position k for all k.
- Plot loss to be paid for making decision d for all positions d.
- Compute the optimal decision.

Task 3. **Tracking (difficulty 4).** The following information is available on the input:

- A sequence of images as in Task 1, array II
- Tank template J
- Noise level σ
- Detected vertical position ky
- Detected speed Ss
- Motion model: probability q to keep direction
- Ground truth positions kk for verification

Follow these steps to construct Kalman filter. Compute:

- Position p.d.f. of the object for every frame, p_t. Visualize the result.

- State: the state will be discrete, formed by position $k \in 0 \dots N$ and velocity $v \in \{-S, S\}$;
- Motion model: With probability q the new state is $(k+v, v)$ and with probability $1-q$ it is $(k+v, -v)$;
- Prediction: Given a state p.d.f. for frame t ($2 \times N$ numbers), compute predicted state p.d.f. in frame $t+1$ using the motion model. Visualize the result. Assume that in time $t=0$, the velocity has probability 0.5 for each direction.
- Update: Infer the posterior state estimate from frame $t+1$ given the predicted state estimate and the evidence p_t.
- Plot recovered position p.d.f. over time as an image. Overlay ground truth true trajectory. Overlay maximum a posterior estimate of the position $\hat{k}_t$ and uncertainty bounds given by $\pm$ standard deviation of the position.

Task 4. **EM-Reconstruction (difficulty 4).**

Legend You built a telescope to observe night sky. You decided to take many images with not so long exposure time to integrate them later on using computer. Usually, simply averaging many images produced a good result. But another day the targeting mechanism of you telescope seem to have short-circuiting problems and each image taken seem to have a random offset. This was a very suspicious breakdown and it was so unlucky to happen just when you hoped to get a good view of a planet. You felt determined to reconstruct the picture with a more clever algorithm.

Problem Formulation and Model The true unknown image is denoted by J and $I = (I_k)_{k=1}^N$ are the observed images. Each I_j is a noisy and shifted observation of J. The shifts are denoted by $d = (d_k)_{k=1}^N$ where shift d_j corresponds to image I_j. Furthermore it is known that for all shifts it holds that $d_k \in \{-D, \cdots, D\}^2$ where D is a known constant.

To model the problem we make the following assumptions:

- The noise of the image is Gaussian i.i.d. with known variance σ.
- The shifts d are independent and uniformly distributed, *i.e.* $p(d_k) = {}^1/_{(2D+1)^2}$
- Given the shifts d (and J), the observations I are independent

$$p(I|d;J) = \prod_{k=1}^{N} p(I_k|d_k;J) \tag{1}$$

- given J and d_k, the pixels in the image I_k are independent:

$$p(I_k|d_k;J) = \prod_{x\in\Omega} p_{\mathcal{N}}(I_k(x) - J(x-d_k), \sigma) \tag{2}$$

  where Ω is the set of all pixel and $p_{\mathcal{N}}(y,\sigma)$ is the normal distribution with zero mean and variance σ, *i.e.*:

$$p_{\mathcal{N}}(y,\sigma) = {}^1/_{\sqrt{2\pi}\sigma} \exp\left(-{}^{y^2}/_{2\sigma^2}\right)$$

By applying the rules of conditional probability to (2) and combining with (1), one gets:

$$p(I,d;J) = \prod_{k=1}^{N} \left(\prod_{x\in\Omega} p_{\mathcal{N}}(I_k(x) - J(x-d_k), \sigma) \right) p(d_k).$$

Since we are not interested in the shifts, we marginalize over them:

$$p(I;J) = \sum_{d\in\{-D,\cdots,D\}^2} p(I,d;J)$$

The maximum likelihood for J is now expressed as

$$\arg\max_J \sum_{d \in \{-D, \cdots, D\}^2} p(I, d; J)$$

The input contains:

- `II` – array of images, shape $H \times W \times 3 \times N$
- `D` – maximum absolute displacement
- `sigma` – the noise variance

Your task is to:

- Describe the estimation step
- Describe the maximization step
- Implement the EM Algorithm in order to reconstruct the image.

Task 5. **Tracking* (bonus task, difficulty 4).** In the setting of Task 3:

- Repeat the estimation process backwards in time starting from the last frame (can be easily done by reverting the sequence).
- You will obtain somewhat different posterior estimate of the trajectory.
- Can you combine the two directions and derive the Bayesian estimate for position at time t given all preceding and all subsequent frames?
- Can you recover a better template from the estimated trajectory?
- Try to formulate the problem of template estimation and trajectory recovery as a joint semi-supervised problem.