Computer Vision II - Homework Assignment 1

Sample solution

Stefan Roth, Nikita Araslanov, Tobias Plötz
Visual Inference Lab, TU Darmstadt

May 16, 2018

Problem 1

1. - x discrete: "amount of work hours put into CV2" $\in \mathbb{Z}_+$,
     y discrete: "grade obtained in CV2 exam" $\in \{1.0, 1.3, 1.7, ..., 3.7, 4.0, 5.0\}$.
   - x continuous: *temperature* $\in \Re$,
     y discrete: $outlook = \{sunny, rainy, overcast\}$.
   - x continuous: "width of a lemon" $\in \Re$,
     y continuous: "height of a lemon" $\in \Re$.

2. We obtain the (marginal) distribution of y. Its density function $f(\cdot)$ can be computed as $f(y) = \int p(x, y)\, dx$.

3. Set A as a random variable describing the number of the box $i = 1, \ldots, 3$ which is chosen. $B = 1$ corresponds to the event that we find a gold coin in the randomly chosen drawer. We assume that the gold and silver coins are distributed as shown in Figure 1. Given the described setup, we have the following probabilities:

$$p(B = 1 \,|\, A = 1) = 1 \qquad p(A = 1) = \frac{1}{3}$$
$$p(B = 1 \,|\, A = 2) = \frac{1}{2} \qquad p(A = 2) = \frac{1}{3}$$
$$p(B = 1 \,|\, A = 3) = 0 \qquad p(A = 3) = \frac{1}{3}$$

   A gold coin is included in the second drawer if and only if we have chosen box $i = 1$, i. e. $A = 1$. With Bayes rule we therefore obtain

$$\begin{aligned} p(A = 1 \,|\, B = 1) &= \frac{p(B = 1 \,|\, A = 1) \cdot p(A = 1)}{p(B = 1)} \\ &= \frac{p(B = 1 \,|\, A = 1) \cdot p(A = 1)}{\sum_i p(B = 1 \,|\, A = i) \cdot p(A = i)} \\ &= \frac{1 \cdot \frac{1}{3}}{1 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}} \\ &= \frac{2}{3}. \end{aligned}$$

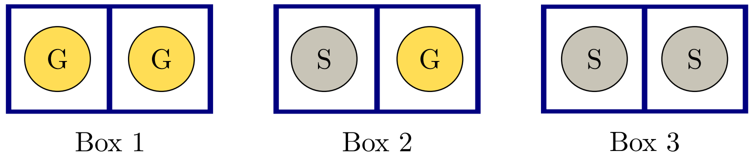

Figure 1: Setup of the boxes with each box containing two coins. Given that the first drawer contains a gold coin, the second drawer can only contain a gold coin if box 1 was chosen.

4. Using the definition of conditional probabilities, it holds that

$$\begin{aligned} p(x, y \mid z) &= \frac{p(x, y, z)}{p(z)} \\ &= \frac{p(x \mid y, z) \cdot p(y, z)}{p(z)} \\ &\overset{(\star)}{=} p(x \mid z) \cdot \frac{p(y, z)}{p(z)} \\ &= p(x \mid z) \cdot p(y \mid z) \end{aligned}$$

where $p(x \mid y, z) = p(x \mid z)$ is applied in $(\star)$.

Problem 2 - Modeling - 6 points

1. *How can we justify to factor the likelihood* $p(\mathbf{I}^1 \mid \mathbf{I}^0, \mathbf{d})$ *into individual terms for each pixel?*
   We assume that each pixel p in $\mathbf{I}^1$ is conditionally independent from all other pixels in $\mathbf{I}^1$ given the value of the corresponding pixel in $\mathbf{I}^0$ and the disparity value at that pixel. That is a reasonable simplification if we assume that the assigned pixel in $\mathbf{I}^1$ conveys all the information we can get about p up to some inherent noise.

2. *Why is it unfavorable to model the likelihood function as a Gaussian?*
   Gaussians have flat tails, i. e. they assign a very low probability to values that are far away from the mean. In the stereo setting, corresponding pixels in $\mathbf{I}^0$ and $\mathbf{I}^1$ can take drastically different intensity values, e. g. due to specular surfaces. If we want to allow such 'outliers' in our model, we need to give them a reasonable probability of occurrence. Hence, a Gaussian is unfavorable to model the likelihood. Instead, we choose a likelihood function that has heavy tails, e. g. the Student-t distribution.

3. *What assumptions does the pairwise MRF prior over the disparity values encode? Given the disparity values* d_1, d_2 *and* d_3 *in Figure 2, are the values* d_1 *and* d_3 *independent given* d_2*?*
   From a modeling point of view, using pairwise potentials implies that our model only expresses which pairwise relations are favorable and which are not, i. e. we can not directly express arbitrary relations between three or more disparity values. Regarding independence assumptions, the pairwise model tells us that a disparity value is independent of any other disparity value given *all* of its four direct neighbors (local Markov property). That means, that we can not get more information about each disparity value than by knowing the value

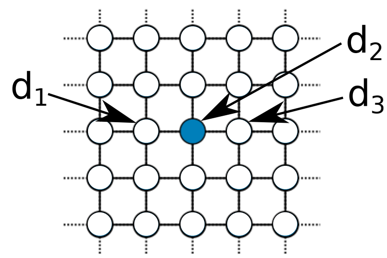

Figure 2: Markov Random Field prior over disparity values

of its neighbors. However, the disparity values d_1 and d_3 are not independent, as we have only observed the disparity value d_2 and hence knowing d_1 provides information about d_3 and vice versa.

4. *Can you think of different compatibility function (or "potential functions") than the delta function used in the Pott's model?*
   The delta functions are 'all or nothing', i. e. either the disparity values match and no penalty is assigned or they do not match and a penalty is inclined no matter how big the difference between disparity values is. As an alternative we may choose a potential function which grows with the absolute difference of disparity values.

5. *In window-based stereo we always compare a whole region around a single pixel with the corresponding region in the second image in order to measure, how good a certain disparity value for that pixel fits to the images. Why can we get away in our model with comparing only single pixel values in the factors of the likelihood function?*
   Although the likelihood function compares only single pixel values, the dependencies between nearby pixels are still modeled through the MRF prior on the disparity values. The information that a certain pixel p provides about another pixel q can 'flow' from p over its disparity value to the disparity value of q and finally to q. That means dependencies between pixels are not constrained to a bounded region! In theory the pixel in the upper-left corner influences the pixel in the lower-right corner, hence we call this a global model.

Problem 4 - Stereo Likelihood - 10 Points

Discussion

Why Log likeilhood?

First, note that computing the log probabilities does not change the argmax of the function because the logarithm is a strictly monotonically increasing function. Second, note that multiplying the actual probabilities equals to a product of a large number of very small values (smaller than one) which underflows the numerical precision of float and even double precision. This is avoided by computing the likelihood in the log-domain instead.

Overall observation

The Gaussian negative log-likelihood rises more than 7 times from no noise to 10% noise. However the Laplacian negative log-likelihood rises less the than 3 times. Therefore the Laplacian likelihood is much more robust regarding outliers.