Computer Vision II - Homework Assignment 1

Stefan Roth, Uwe Schmidt, Tobias Plötz, GRIS, TU Darmstadt

April 30, 2014

This homework is due on May 12, 2014 at 20:00.
Please read the instructions carefully!

General remarks

Your grade will of course depend on the correctness of your answer, but also on a clear presentation of your results and good writing style. It is your responsibility to find a way to *explain clearly how* you solved the problems. Note that you will get grades for the solution, not for the result. If you get stuck, try to explain why and describe the problems you encountered – you can get partial credit even if you did not complete the task. So please hand in enough information for us to understand what you did, what you tried, and how it worked!

We encourage interaction about class-related topics both within and outside of class. However, you should not share solutions with your classmates, and *everything you hand in must be your own work*. You are also not allowed to use material from the web. You are required to **acknowledge any source of information that you used to solve the homework** (i.e. books other than the course books, papers, web sites, etc.). Acknowledgements will *not* affect your grade. Thus, there is no reason why you would not acknowledge sources properly. Not acknowledging a source that you have used, on the other hand, is a clear violation of academic ethics. Note that the university as well as the department is very serious about plagiarism. For more details please see `http://www.informatik.tu-darmstadt.de/index.php?id=202` and `http://plagiarism.org`.

Groups

Please form groups of two people to work on the exercises and do not change groups during the semester. If you do not find a group member get in touch with us *early* and we will assign you to another person.

Programming exercises

For the programming exercises you will be asked to hand in Matlab code. If you used any other tool to write your code, say Octave, it is *your responsibility* to make sure that the code also works on Matlab, which is what we will use for grading. We have Matlab access available for everyone who needs it. In order for us to be able to grade the programming assignments properly, you need to comment your code in sufficient detail so that it will be easily clear to us what each part of your code does. Sufficient detail does not mean that you should comment every line of code (that defeats the purpose), nor does it mean that you should comment 20 lines of code using only a single sentence. Of course, all this is good coding practice anyway, so you would want to do this no matter what we expect from you.

Your Matlab code should display your results so that we can judge if your code works from the results alone. Of course, we will still look at the code. If your code displays results in multiple stages, please insert appropriate `pause` commands between the stages so that we can step through the code. Please be sure to name each file according to the naming scheme included with each problem. This also makes it easier for us to grade your submission. And finally, please be sure to include your name and email in the code.

Files you need

All the data you will need for the problems will be made available at the course web page `http://www.gris.tu-darmstadt.de/teaching/courses/ss14/cv2`, next to the link for this exercise sheet.

What to hand in

As mentioned, you need to show your solution and how you got there. Your handin should contain a PDF file (plain text is ok, too) with any textual answers that may be required. You do not have to include images of your results. Your code should show these instead.

For the programming parts, please hand in all documented `.m` scripts and functions that your solution requires. Make sure your code actually works (also in an empty workspace) and that all your results are displayed properly! Also include a separate text file which lists all `.m` files with a brief description for each one.

Handing in

Please upload your writeup and your code to the corresponding moodle area: `https://moodle.tu-darmstadt.de/course/view.php?id=3243`. Only one group member has to hand in your work. If *and only if* you experience problems with uploading your solution, you may also email it to `cv2staff@gris.tu-darmstadt.de`

You are supposed to send all your solution files as a single `.zip` or `.tar.gz` file. **Please note that we cannot accept file formats other than the ones specified!** These are widespread standards that are available on any platform.

Late Handins

We will accept late handins, but we will take 20% off for every day that you are late. Note that even 15 minutes late will be counted as being one day late! After the exercise has been discussed in class, you can no longer hand in.

If you, for some serious (say medical) reason, cannot make the deadline, you need to contact us *before* the deadline. We might waive the late penalty in such a case.

Problem 1 - Probabilities and Statistics - 14 points

1. Give three real-world examples of a joint distribution $p(x, y)$ where x and y are

x	y
discrete	discrete
continuous	discrete
continuous	continuous

2. What remains if we marginalize a joint distribution $p(v, w, x, y, z)$ over five variables with respect to variables w and z? What remains if we marginalize the resulting distribution with respect to x?

3. Show that the following relation is true:

$$p(w, x, y, z) = p(w|y) \cdot p(x|w, y, z) \cdot p(y) \cdot p(z|w, y)$$

4. In my pocket there are two coins. Coin 1 is unbiased, so the likelihood $p(h = 1|c = 1)$ of getting heads is 0.5 and the likelihood $p(h = 0|c = 1)$ of getting tails is also 0.5. Coin 2 is biased so that the likelihood $p(h = 1|c = 2)$ of getting heads is 0.7 and the likelihood $p(h = 0|c = 2)$ of getting tails is 0.3. I reach into my pocket and draw one of the coins at random. There is an equal prior probability I might have picked either coin. I flip the coin and observe a head. Use Bayes rule to compute the posterior probability that I chose coin 2.

5. The joint probability $p(w, x, y, z)$ over four variables factorizes as

$$p(w, x, y, z) = p(w)p(z|x, w)p(y|z)p(x).$$

   Show that x is independent of w by showing that $p(x, w) = p(x)p(w)$.

6. Consider two variables x and y with joint distribution $p(x, y)$. Show the following results:

   - $\mathbb{E}[x] = \mathbb{E}_y[\mathbb{E}_x[x|y]]$
   - $\mathrm{var}[x] = \mathbb{E}_y[\mathrm{var}_x[x|y]] + \mathrm{var}_y[\mathbb{E}_x[x|y]]$

   Here $\mathbb{E}_x[x|y]$ denotes the expectation of x under the conditional distribution $p(x|y)$, with a similar notation for the conditional variance.

Problem 2 - Modeling - 6 points

When we create models in computer vision, we are encoding prior assumptions about the real world in a mathematical framework. We will have a look at the simple Markov Random Field (MRF) model for binocular stereo that we encountered in lecture 2 and try to recap which assumptions we put in there.

$$\begin{aligned} p(\mathbf{d}|\mathbf{I}^0, \mathbf{I}^1) &\propto p(\mathbf{I}^0, \mathbf{I}^1|\mathbf{d})p(\mathbf{d}) \\ p(\mathbf{I}^0, \mathbf{I}^1|\mathbf{d}) &= \prod_{i,j} f(\mathbf{I}^0_{i,j} - \mathbf{I}^1_{(i-d_{i,j}),j}) \\ p(\mathbf{d}) &= \prod_{i,j} f_H(d_{i,j}, d_{(i+1),j}) f_V(d_{i,j}, d_{i,(j+1)}) \end{aligned}$$

Please (briefly) answer the following questions **in your own words**:

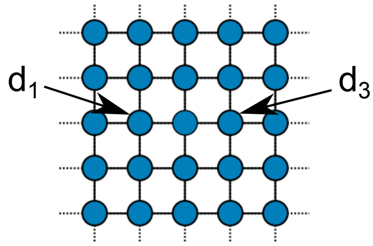

Figure 1: Markov Random Field prior over disparity values

1. How can we justify to factor the likelihood $p(\mathbf{I}^0, \mathbf{I}^1|\mathbf{d})$ into individual terms for each pixel?
2. What may be the reason to choose a robust likelihood function (e. g. Laplacian, Student-t) over the Gaussian likelihood?
3. What assumptions does the pairwise MRF prior over the disparity values encode? Are the disparity values d_1 and d_3 in figure 1 independent?
4. Can you think of different compatibility function (or "potential functions") than the delta function used in the Pott's model?
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?

Problem 3 - Getting to know Matlab - 1 point

In this first, very simple programming problem you will get to know Matlab if you don't know it already. If you do, this shouldn't take you more than a few minutes, so please do this problem anyway even if it is incredibly simple. You will likely be reusing this code piece many times anyway.

Commands useful for this problem include `imread, rgb2gray, imagesc, colormap, axis, min, max, mean, disp, title`. You should not have to use any `for` loop to solve this problem.

Tasks:

- Get and use the image `a1p3.png` from the ZIP file.
- Write a Matlab script `a1p3.m` that does the following:
  1. Load the image, convert it to grayscale, and then convert it into the double format we will use internally for all our problems. Make sure that the image is scaled to $[0, 255]$.
  2. Display the image with the correct aspect ratio, and with an appropriate color map (i.e. so that we see a graylevel image on screen). `Note:` Please always display your image results like this, unless otherwise specified.
  3. Compute and display (either at the Matlab prompt or as annotation of the image display) the minimum, maximum, and mean pixel value.

Problem 4 - Likelihood - 11 Points

In this problem you are going to implement a simple likelihood model for stereo. Commands useful for this problem include `normpdf` and `randperm`.

Tasks:

- Write a Matlab script `ml.m` that does the following:

  1. Load the images `i0.ppm` and `i1.ppm` from the Tsukuba dataset. Load the ground truth disparity map $\mathbf{d}_{\mathrm{GT}}$ contained in the file `gt.pgm`. You should divide all values from `gt.pgm` by 16 to obtain the actual disparity in pixels.
  2. Assuming conditional independence between pixels, compute and display the likelihood $p(\mathbf{I}^0, \mathbf{I}^1|\mathbf{d})$ of a Gaussian likelihood model with $\mu = 0$ and $\sigma = 1$ for the ground truth disparity $\mathbf{d}_{\mathrm{GT}}$. Note that the ground truth is not defined along the image borders.
  3. Now compute and display the value of the negative log-likelihood (note that you should not just take the negative log from the output of the previous part, but rather implement everything in the log-domain). What is the reason for computing the log instead of working with the actual probability densities?
  4. Take the second input image (corresponding to `i1.ppm`) and artificially generate pixels for which the brightness constancy is violated by replacing 10% (and also 30%) of all pixels (pixel position chosen uniformly at random across the entire image) with random brightness values (chosen uniformly at random from the brightness range $[0, 255]$). We will call this modified image $\mathbf{I}_o^1$ due to the outliers it contains. Compute and display the likelihood $p(\mathbf{I}^0, \mathbf{I}_o^1|\mathbf{d})$ and the negative log-likelihood $-\log p(\mathbf{I}^0, \mathbf{I}_o^1|\mathbf{d})$.
  5. Now implement (in the log-domain) a more robust likelihood model using a Laplacian distribution:
  $$\hat{p}(\mathbf{I}^0, \mathbf{I}^1|\mathbf{d}) = \prod_{i,j} \frac{1}{2s} \exp\left\{ -\frac{|I^0_{i,j} - I^1_{(i-d_{ij}),j}|}{s} \right\}.$$
  Choose $s = 1$. Compute, display and compare $\hat{p}(\mathbf{I}^0, \mathbf{I}^1|\mathbf{d})$ and $\hat{p}(\mathbf{I}^0, \mathbf{I}_o^1|\mathbf{d})$.
  6. Discuss your findings regarding outlier robustness based on the Gaussian and the Laplacian likelihood models.