Homework Assignment 2

The homework assignment should be solved *individually.* However, it is allowed to ask questions to fellow students. All solutions should be clearly written and well motivated. Regarding the language, you may use Swedish or English.

The assignment is due on Wednesday, January 21, 2009. Please email your solutions to schon@isy.liu.se.

The assignment is implementation oriented. The idea is to provide examples of the most important parts of the "surrounding infrastructure" of a sensor fusion framework including a camera.

1. *Harris corners*
   The task is to implement a Harris corner extractor in MATLAB.

   (a) *compute image gradients* Your task is to implement a function that computes the image gradient

   $$\nabla I(x_p, y_p) = \begin{pmatrix} I_x(x_p, y_p) \\ I_y(x_p, y_p) \end{pmatrix} \tag{1}$$

   The input to this function should be

   - The image
   - The standrad deviation for the smoothing step σ_s.
   - The standrad deviation for the differentiation σ_d.

   The output should be

   - The image gradient in the x-direction, I_x.
   - The image gradient in the y-direction, I_y.

   Verify that your result makes sense by providing plots of I_x and I_y. *Hint: Appendix 4.A in Ma et al. (2006) is probably useful. Furthermore, the commands conv2, imread, rgb2gray and imagesc can be useful.*

   (b) Implement a function that computes the Harris strength

   $$c(x_p, y_p) = \det(H(x_p, y_p)) + k\,\mathrm{tr}^2(H(x_p, y_p)), \tag{2}$$

   for each pixel in the image. Recall that

   $$H(x_p, y_p) = \sum_{x=-N/2}^{N/2} \sum_{y=-N/2}^{N/2} \begin{pmatrix} I_1 & I_{12} \\ I_{12} & I_2 \end{pmatrix} \in \mathbb{R}^{2\times 2} \tag{3a}$$

   where

   $$\begin{aligned} I_1 &= I_x^2(x_p + x, y_p + y), && \text{(3b)} \\ I_{12} &= I_x(x_p + x, y_p + y) I_y(x_p + x, y_p + y) && \text{(3c)} \\ I_2 &= I_y^2(x_p + x, y_p + y). && \text{(3d)} \end{aligned}$$

   The input to this function should be

- The constant k used in (2). A commonly used value for this constant is $k = 0.03$.
- The size N of the window used in computing H in (3).
- Image gradient in the x-direction, I_x.
- Image gradient in the y-direction, I_y.

The output should be

- The Harris strength c given by (2) for each pixel.

Verify your result by plotting the 50 most distinctive Harris corners overlayed on the actual image.

Investigate the strength $c(x_p, y_p)$ by using a 3D surf plot of the strength. Does this plot make sense?

(c) Discuss how the computational time be reduced. Hence, it is not required that you implement this.

2. *Feature tracking using block matching*
The task in this assignment is to implement data association and simple feature tracking using block matching. The idea in block matching is to cut out a patch around the most distinctive Harris corners. These patches are then searched for in the consecutive images using normalized cross-correlation. This approach is commonly used for visual odometry, see e.g., Cheng et al. (2006) and SLAM, see e.g., Davison et al. (2007). Data to be used in solving this assignment can be downloaded from the course web site. The data was collected using the combined camera and IMU sensor using during a trip from Linköping to Göteborg. The sensor was mounted close to the rear view mirror looking through the front window. Further description of how to use the data is also provided in the `README.TXT` file.

(a) *Data association* Implement a function that performs data association based on normalized cross-correlation (NCC). The input to the function should be

- The current image
- The patches to be associated in the current image
- Predicted patch positions in pixels
- The size of the search region to be used. The search region is the part of the current image that should be considered in performing the data association.

The output should be

- The positions of the patches in the new image. If a patch cannot be associated (could be outside the current FOV) indicate this using NaN.

Hint: The function normxcorr2 is useful here.

(b) Test your data association using the data available from the course web site by plotting the associated patch position in a sequence of consecutive images. More specifically, compare the performance for the data in `normal.mat` and `braking.mat`. The data in `normal.mat` was collected during normal driving and the data in `braking.mat` stems from a panic braking situation.

(c) What is the average life time of a feature? That is, on average, in how many frames can you successfully track a feature?

(d) In the sequence used above you will typically follow features belonging to moving objects, such as cars. This is of course dangerous if the features are to be used for instance for positioning the host vehicle. Discuss how this can be detected within a sensor fusion framework.

3. *Improved feature prediction for Visual Odometry*
Let us now assume that you had information about the motion (including position, velocity and orientation) of the host vehicle and the position of the features in the world. How could this information be used in order to produce better feature predictions? Can this information aid in finding suitable search regions? It can be assumed that the involved variables are approximately Gaussian.

References

Cheng, Y., Maimone, M. W., and Matthies, L. (2006). Visual odometry on the Mars exploration rovers. *IEEE Robotics & Automation Magazine*, 13(2):54–62.

Davison, A. J., Reid, I., Molton, N., and Strasse, O. (2007). MonoSLAM: Real-time single camera SLAM. *IEEE Transactions on Patterns Analysis and Machine Intelligence*, 29(6):1052–1067.

Ma, Y., Soatto, S., Kosecka, J., and Sastry, S. S. (2006). *An invitation to 3-D vision – from images to geometric models.* Interdisciplinary Applied Mathematics. Springer.