# Computer Vision, Assignment 5
# Local Optimization and Structure from Motion

## 1 Instructions

In this assignment you will study model fitting using optimization. In particular you will compute the maximal likelihood solution for a couple of structure and motion problems. The data for the assignment is available in Ping Pong.

The assignment is due Thursday Dec 20, 2018. Make sure you answer all questions and provide complete solutions to the exercises. Write your name, the name of your collaborator (or that you have completed the exercises on your own) and the assignment number. After each exercise there is a gray box with instructions on what should be included in the report. In addition, all the code should be submitted as m-files. Make sure that your matlab scripts are well commented and can be executed directly (that is, without loading any data, setting parameters etc. Such things should be done in the script).

You will have time to work with the assignments during the computer laboratory sessions / exercise sessions. These sessions are intended to provide an opportunity for asking questions on things you have had problems with or just to work with the assignment. More specifically, during the laboratory sessions you should concentrate on the exercises marked "Computer Exercise". The rest of the exercises are intended to provide hints and prepare you for the computer exercises. You are expected to have solved these before you go to the laboratory sessions.

The report should be written individually, however you are encouraged to work together in pairs (in the lab session you might have to work in pairs). Note that it is NOT allowed to do the assignments in larger groups than two persons. Keep in mind that everyone is responsible for their own report and should be able to explain all the solutions.

For a passing grade (3) in the course, all exercises except for the ones marked as OPTIONAL need to be completed and submitted before the due date. For higher grades (4 or 5), sufficiently many of the optional exercises should be correctly completed and submitted before the due date.

# 2 Maximum Likelihood Estimation for Structure from Motion Problems

Exercise 1. Suppose the 2D-point $x_{ij} = (x^1_{ij}, x^2_{ij})$ is an observation of the 3D-point $\mathbf{X}_j$ in camera $P_i$. Also we assume that the observations are corrupted by Gaussian noise, that is,

$$(x^1_{ij}, x^2_{ij}) = \left( \frac{P^1_i \mathbf{X}_j}{P^3_i \mathbf{X}_j}, \frac{P^2_i \mathbf{X}_j}{P^3_i \mathbf{X}_j} \right) + \epsilon_{ij}, \tag{1}$$

where $P^1_i, P^2_i, P^3_i$ are the rows of the camera matrix $P_i$ and $\epsilon_{ij}$ is normally distributed with covariance $\sigma I$. The probability density function is then

$$p(\epsilon_{ij}) = \frac{1}{2\pi\sigma} e^{-\frac{1}{2\sigma^2}||\epsilon_{ij}||^2}. \tag{2}$$

Assuming that the $\epsilon_{ij}$ are independent, that is

$$p(\epsilon) = \prod_{i,j} p(\epsilon_{ij}), \tag{3}$$

show that the model configuration (points and cameras) that maximizes the likelihood of the obtaining the observations $x_{ij} = (x^1_{ij}, x^2_{ij})$ is obtained by solving

$$\min \sum_{i=1}^{n} \sum_{j=1}^{m} || \left( x^1_{ij} - \frac{P^1_i \mathbf{X}_j}{P^3_i \mathbf{X}_j},\ x^2_{ij} - \frac{P^2_i \mathbf{X}_j}{P^3_i \mathbf{X}_j} \right) ||^2. \tag{4}$$

Hint: Maximize the log-likelihood $\log p$.

For the report: Complete solution.

# 3 Calibrated Structure from Motion and Local Optimization

Exercise 2. (OPTIONAL.) Unfortunately there is no closed-form solution for computing the ML estimation when we use general pinhole cameras. The only way to find the ML estimate is to try to improve a starting solution using local optimization. Suppose that we want to minimize

$$\sum_i r_i(v)^2 = ||r(v)||^2, \tag{5}$$

where $r_i(v)$ are the error residuals and $r(v)$ is a vector containing all the $r_i(v)$. The first order Taylor expansion at a point $v_0$ is

$$r(v) \approx r(v_0) + J(v_0)\delta v, \tag{6}$$

where $\delta v = (v - v_0)$ and $J(v_0)$ is a matrix whose rows are the gradients of $r_i(v)$ at $v_0$. Show that the steepest descent direction (that is, the negative gradient) of the approximation

$$||r(v_0) + J(v_0)\delta v||^2 \tag{7}$$

at the point $v_0$ is

$$d = -2J(v_0)^T r(v_0). \tag{8}$$

For the report: Complete solution.

Exercise 3. (OPTIONAL.) A direction d is called a descent direction (of $f$ at $v_0$) if

$$\nabla f(v_0)^T d < 0, \tag{9}$$

since this means that the directional derivative in the direction $d$ is negative (see your multidimensional calculus book).

Show that the direction $d$ in equation (8) is a descent direction for the original function (5) at the point $v_0$.

A matrix $M$ is called positive definite if $v^T M v > 0$ for any $v$ such that $||v|| \neq 0$. Show that the direction

$$d = -M\nabla f(v) \tag{10}$$

is a descent direction. In the Levenberg-Marquardt method we chose the update step

$$\delta v = -(J(v_0)^T J(v_0) + \lambda I)^{-1} J(v_0)^T r(v_0) \tag{11}$$

Is this a step in a descent direction?

For the report: Complete solution.

Computer Exercise 1. In Computer Exercise 3 and 4, Assignment 3, you computed a solution to the two-view structure form motion problem for the two images of Figure 1 using the 8-point algorithm. In this exercise the goal is to use the solution from Assignment3 as a starting solution and locally improve it using the Levenberg-Marquardt method.

Figure 1: kronan1.jpg and kronan2.jpg.

The file `LinearizeReprojErr.m` contains a function that for a given set of cameras, 3D points and imagepoints, computes the linearization (6). The file `update_solution.m` contains a function that computes a new set of cameras and 3D points from an update $\delta v$ computed by any method. The file `ComputeReprojectionError.m` computes the reprojection error for a given set of cameras, 3D points and image points. It also returns the values of all the individual residuals as a second output.

In the Levenberg-Maquardt method the update is given by

$$\delta v = -(J(v_k)^T J(v_k) + \lambda I)^{-1} J(v_k)^T r(v_k). \tag{12}$$

Using this scheme and starting from the solution that you got in Assignment 3, plot the reprojection error versus the iteration number for $\lambda = 1$. Also plot histograms of all the residual values before and after running the Levenberg-Maquardt method.

Try varying $\lambda$. What happens if $\lambda$ is very large/small?

```
Useful matlab commands:
%Takes two camera matrices and puts them in a cell.
P = {P1,P2}

%Computes the reprejection error and the values of all the residuals
%for the current solution P,U,u.
[err,res] = ComputeReprojectionError(P,U,u);

%Computes the r and J matrices for the appoximate linear least squares problem.
[r,J] = LinearizeReprojErr(P,U,u)

% Computes the LM update.
C = J'*J+lambda*speye(size(J,2));
```

```
c = J'*r;
deltav = -C\c;

%Updates the variabels
[Pnew,Unew] = update_solution(deltav,P,U);
```

For the report: Submit plots of the function value versus the number of iterations and the two histograms.

Computer Exercise 2. (OPTIONAL) In this exercise the goal is to compute a dense textured model of kronan, see Figure 2, using the two images from the previous exercise. The file `compex2data.mat`

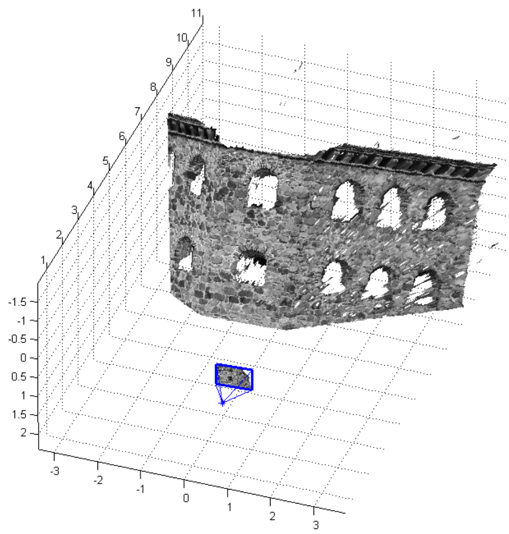

Figure 2: Dense reconstruction of kronan.

contains foreground segmentations of the two images. The file `compute_ncc.m` contains a function that the normalized cross correlation for each pixel and depth using the two images and the segmentations. The file `disp_result.m` plots the resulting model. Use the commands listed below to compute the textured model.

```
Useful matlab commands:
%Selects suitable depths for the planesweep algorithm
d = linspace(5,11,200);

%rescales the images to incease speed.
sc = 0.25;
%Compute normalized corss correlations for all the depths
[ncc,outside_image] = compute_ncc(d,im2,P{2},im1,segm_kronan1,P{1},3,sc);

%Select the best depth for each pixel
[maxval,maxpos] = max(ncc,[],3);

%Print the result
disp_result(im2,P{2},segm_kronan2,d(maxpos),0.25,sc)
```

For the report: Submit the m-file and a plot of the final 3D reconstruction.