Chapter 1

Matlab

1.1 A matrix m in matlab is entered as `m = [1 2 3;4 5 6;7 8 9]` The semicolon denotes a new row, when entering a matrix. A semicolon after an expression tells matlab not to display the result. An element of a matrix can be accessed as `m(2,3)`. A submatrix can be accessed as `m(2:3,1:2)` and whole rows or colums using `m(1,:)` and `m(:,2)`. Try these commands. Also try

```
d = det(m);
[v,d]=eig(m);
[u,s,v]=svd(m);
```

Compare the matrix `u*s*v'` with `m`.

1.2 The matlab `help` command can be useful. Try `help ops` and find out what `*`, `.*` and `'/` means. Also try

```
help help
help type
help which
help whos
help det
help eig
help svd
demo
```

1.3 Write a matrix representation of the point $(2,3)$, the line $2x+3y+7=0$ and the conic $x^2+(y-2)^2=9$. What line does the matrix representation

$$l = \begin{bmatrix} 4 \\ 2 \\ -3 \end{bmatrix}$$

represent? What conic does the matrix representation

$$c = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -9 \end{bmatrix}$$

represent? Try `rital(l)` and `ritac(c)`.

1.4 Try `help pflat` and `help psphere`. Try these routines on the following matrix.

```
m = [1 2 10;3 6 1;1 2 5];
```

How do the routines work on a point at infinity

```
m = [1;2;0];
```

1.5 Try `rq` on the following matrix.

$$P = \begin{pmatrix} 5 & -14 & 2 & 17 \\ -10 & -5 & -10 & 50 \\ 10 & 2 & -11 & 19 \end{pmatrix}$$

Chapter 2

Introduction to Computer Vision

2.1 Let the projection be given by a camera matrix

$$P = \begin{pmatrix} 5 & -14 & 2 & 17 \\ -10 & -5 & -10 & 50 \\ 10 & 2 & -11 & 19 \end{pmatrix}$$

and a three-dimensional point in homogeneous coordinates as

$$X = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 1 \end{pmatrix} .$$

What image coordinates (x, y) does the projection of X have? Is the camera calibrated? What are the three-dimensional coordinates of the focal point? What are the three-dimensional coordiantes of the point X?

2.2 Study the following calibrated camera matrix

$$P_f = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & f \end{pmatrix} .$$

Let the matrix

$$K_f = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

represent internal calibration. Study the camera matrix $K_f P_f$ for $f = 1, 10, 100$ and 1000 pixels. Where is the focal point of the camera? What are the image coordinates of the projection of $(0, 0, 0)$? Is it fair to use a scaled orthographic model for object points

$$X = \begin{pmatrix} X & Y & Z & 1 \end{pmatrix}^T .$$

close to the origin?

2.3 Study the following calibrated camera matrix

$$P_f = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 10 \end{pmatrix} .$$

Let the matrix

$$K_f = \begin{pmatrix} 10 & 1 & 0 \\ 0 & 11 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

represent internal calibration. Study the camera matrix KP. Assume that the object points are located at $Z \approx Z_0 \pm 10$, where $Z_0 = 100$ or 1000. Approximate the camera with an affine and estimate the errors. Compare the errors for the two different Z_0. Conclude.

2.4 Two points represented in homogeneous coordinates as u and v are identical if they differ by scale only, i.e.

$$u \sim v \Longleftrightarrow \exists \lambda u = v$$

Show that $u \sim v$ if and only if $u \times v = 0$. Show that the operator

$$T_u(v) = u \times v$$

is linear (in v). Since T_u is linear in v it can be written using matrix represenation $T_u(v) = T_u v$. What is the matrix T_u? Each row of the matrix T_u can be interpreted as a line. That $v \sim u$ can thus be seen as the constraint $T_u v = 0$ that v lies on three lines. Interpret these three lines.

2.5 In the previous exercise it was shown that $u \times v$ if and only if $T_u v = 0$ for an antisymmetric matrix T_u. using this terminology the camera matrix equation

$$x \sim PX$$

can be rewritten as

$$T_x P X = 0$$

What is the size of the matrix $T_x P$. What is the rank of the matrix T_x. Assuming that the point x is not at infinity, can one choose two rows of T_x so that the submatrix always has rank 2.

2.6 Assume that points are projected according to

$$x \sim PX$$

The constraint that the image point x lies on a line l can be written as

$$l^T x = 0$$

Insert the camera projection equation into this constraint. What constraint do we get on the object point X that project to points on the line? Interpret $l^T P$?

2.7 Let the projection be given by a camera matrix

$$P = \begin{pmatrix} 5 & -14 & 2 & 17 \\ -10 & -5 & -10 & 50 \\ 10 & 2 & -11 & 19 \end{pmatrix}$$

and two three-dimensional point in homogeneous coordinates as

$$X_1 = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 1 \end{pmatrix}$$

and

$$X_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 1 \end{pmatrix} \quad .$$

Calculate the image points x_1 and x_2. Calculate the line l through x_1 and x_2. Calculate the plane Π of points X that project onto the line l. Verify that X_1 and X_2 lie on the plane Π.

Chapter 3

Mathematical tools

3.1 Show that the charts $t \mapsto t$ and $t \mapsto t^3$ gives two different but homeomophic differentiable structures on the line.

3.2 Find coordinate maps that makes the unit circle in $\mathbb{R}^2$ a differentiable manifold.

3.3 Show that the coordinate functions

$$p_n^{-1}(x,y) = \frac{4}{x^2+y^2+4}(x,y,x^2+y^2-4) \tag{3.1}$$

$$p_s^{-1}(x,y) = \frac{4}{x^2+y^2+4}(x,y,4-x^2-y^2) \tag{3.2}$$

defines a differentiable structure on S^2, i.e. that $p_s \circ p_n^{-1}$ and $p_s \circ p_n^{-1}$ are of class C^∞.

3.4 Show that if S is a skew-symmetric matrix, then

$$S \mapsto (I+S)(I-S)^{-1}$$

is an orthogonal matrix.

3.5 Show that the set of orthogonal matrix can be given a manifold structure as an implicitly defined manifold. Do the same for the orthogonal matrices with determinant equal to 1.

3.6 What is the intersection of the lines

$$2x+3y=2 \quad \text{and} \quad 4x+6y=2$$

in $\mathcal{P}^2$? Interpretation? (In matlab try `null([2 3 -2;4 6 -2])`

3.7 Show that the dual to a conic

$$\mathbf{x}^T C \mathbf{x} = 0$$

can be written as

$$\mathbf{n}^T C^{-1} \mathbf{n} = 0 \ .$$

3.8 Show that

$$(1,1,1), \quad (1,2,1), \quad (2,1,1), \quad (2,2,1)$$

constitute a projective basis in $\mathcal{P}^2$. Find a collineation, that maps these points to a standard prjective basis.

3.9 Show that any projective basis can be transferred to a standard projective basis.

3.10 Show that the subgroup of projective transformations that preserves the absolute conic is the similarity transformations.

3.11 Show that the cross-ratio is invariant under projective transformations.

3.12 Calculate the angles between the lines

$$\mathbf{l}_1 = (1,0,0) + t(1,1,1) \quad \mathbf{l}_2 = (0,1,0) + t(1,0,1)$$

using the Laguerre formula.

3.13 What is the dimension of the manifold of lines in the plane? Suggest some coordinate maps for this manifold.

3.14 What is the intersection of the two lines: $l_1 = \begin{pmatrix}1 & 2 & 3\end{pmatrix}^T$ $l_2 = \begin{pmatrix}1 & 2 & 4\end{pmatrix}^T$.

3.15 A partial differential equation is given by

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} = 0$$

Introduce new coordinates

$$u = 5y, \qquad v = 2x + y \ .$$

What is the partial differential equation in the new coordinates? How does the vector $(1,1)$ in the $x-y$-coordinate system transform? What is the connection with covariant and contravariant vectors.

3.16 Calculate the projection of the line that goes through the two points

$$X_1 = \begin{pmatrix}0\\2\\2\\1\end{pmatrix} \qquad X_2 = \begin{pmatrix}1\\0\\2\\1\end{pmatrix} \ .$$

The camera matrix is given by

$$P = \begin{pmatrix} 5 & -14 & 2 & 17 \\ -10 & -5 & -10 & 50 \\ 10 & 2 & -11 & 19 \end{pmatrix} \ .$$

3.17 A quadric is given in homogeneous coordinates as $X^TQX = 0$ with

$$Q = \begin{pmatrix} 9 & 0 & 0 & -18 \\ 0 & 4 & 0 & -16 \\ 0 & 0 & 1 & -6 \\ -18 & -16 & -6 & -1 \end{pmatrix} \ .$$

Calculate the image conic $x^TCx = 0$ with camera matrix given by

$$P = \begin{pmatrix} 2 & -1 & 2 & 0 \\ 2 & 2 & -1 & 0 \\ -1 & 2 & 2 & 100 \end{pmatrix} \ .$$

3.18 A simplified image of a railway track and the horizon is given by three lines represented in homogeneous dual coordinates as

$$l_1 = \begin{pmatrix} -100 \\ 1 \\ 100 \end{pmatrix} \quad l_2 = \begin{pmatrix} 100 \\ 1 \\ 100 \end{pmatrix}, \quad l_H = \begin{pmatrix} 0 \\ 1 \\ -100 \end{pmatrix} .$$

Sketch the image. What does the intersection of l_1 and l_2 correspond to in reality? Such points are called vanishing points. Which object points does the horizon correpond to in reality?

Chapter 4

Feature extraction

4.1 Use the result from the previous exercise. Construct a filter that is an approximation of the laplacian,

```
laplace = [0 1 0;1 -4 1;0 1 0];
```

Use this as a corner detector

```
cornerim = conv2(bild,laplace,'same'); imagesc(abs(cornerim));
```

Is this a good cornerdetector? Try thresholding the image with different thresholds,

```
imagesc(abs(cornerim)>5);
```

4.2 Use the result from the first exercise. Construct a filter that is an approximation of derivative in the x direction,

```
dx = [1 0 -1;2 0 -2;1 0 -1];
```

Use this as a edge detector (for vertical edges)

```
dbilddx = conv2(bild,dx,'same'); imagesc(abs(dbilddx));
```

Is this a good edgedetector? Try thresholding the image with different thresholds,

```
imagesc(abs(dbilddx)>5);
```

Do the same thing for horizontal edges (derivative in y direction). One could try to find corners by looking at points where both the x and y derivatives are large.

4.3 The Harris corner detector has already been implemented. Try

```
bild = pgmlas(filnamn,1);
[Y,X,f]=cordata(eightmax(harris(bild,10),1000))
u=[X';Y';ones(size(X'))];
hold off
imagesc(bild);
hold on;
rita(u,'*');
```

What happens?

4.4 In low-level vision it is common to interpolate sub-pixel intensity values. To find the sub-pixel intensity at position x at scale b one calculates

$$w_b(x) = \sum_{i \in \mathbf{Z}^2} f_b(i - x) w(i)$$

where $f = \operatorname{sinc} * g_b$ with G_b Show that

$$|| \operatorname{sinc} * G_b - G_b||_2^2 \leq \frac{1}{b\sqrt{\pi}} \Phi(-\pi b \sqrt{2}) \tag{4.1}$$

where Φ is the primitive function of

$$\frac{2}{\sqrt{\pi}} e^{-x^2} \quad .$$

Chapter 5

Tracking and fitting

5.1 The terms Hotelling transform, Principal Component Analysis, Karhunen-Loeve transform and Singular Value Decomposition are in principal the same thing. In matlab the singular value decomposition is implemented as a routine `svd`:

```
[u,s,v]=svd(M);
```

The input is a $m \times n$ matrix M. The result are three matrices u, s and v, such that u and v are orthogonal, s is diagonal with non-negative decreasing scalars on the diagonal. Show that

$$\min_{w|w|=1} |Mw| = s_n$$

where s_n is the smallest singular value. The corresponding vector w that minimizes $|Mw$ is column n in the matrix v. Show that

$$\min_{rank\tilde{A}\leq k} ||A - \tilde{A}||_F$$

is obtained by

$$\tilde{A} = u\tilde{s}v^T$$

where $\tilde{s}$ is formed from s by zeroing all but the k largest eigenvalues.

5.2 Fit a line to the measured points

$$(1, 1.1), (2, 1.9), (3, 3.3), (4, 4.5), (5., 4.9), (6, 5.7), (7, 7.3)$$

a) using algebraic fitting.

b) using non-linear fitting.

Compare the results. Do the same for the edge points that you saved in `edgepos` in an earlier exercise.

5.3 Experiment with b-spline representations of curves in MATLAB.

5.4 Show the matrix inversion lemma,

$$(B + C^T DC)^{-1} = B^{-1} - B^{-1}C^T(CB^{-1}C^T + D^{-1})^{-1}CB^{-1} \tag{5.1}$$

by direct substitution.

5.5 Discuss what state variables that could be used when tracking a line in a sequence of images.

Chapter 6

Multiple view geometry

6.1 This problem is about structure and motion estimation of points in two images. Run the script `kdemo6`. It loads two images called `bild1` and `bild4`. We have used sub-pixel methods to find 15 corresponding points. These are stored in matrices `x1` and `x4`.

a) First estimate the fundamental matrix F using the eight-point algorithm. The numerics is better if we condition the points by dividing all pixel values by 100. This has already been done. The result is stored in matrices `x1c` and `x4c`.

b) Calculate the epipoles from the fundamental matrix.

c) Does the matrix you have obtained have rank 2. If not exchange F with its closest singular matrix. Use `svd`. Now extract two projection matrices P_1 and P_4 from F.

d) Now that you know the motion (the camera matrices P_1 and P_4). Use triangulation to estimate the projective coordinates of the 15 points.

e) Finally reproject the 15 points and see whether they are close to the measured ones or not.

6.2 Try the script `kdemo7`. It loads two images and a fundamental matrix. It then waits for the user to click in figure 1. This point is the plotted. Then we use this point and the fundamental matrix to find the line in figure 2 of possible correspondences (transfer).

6.3 Is the matrix

$$E = \begin{pmatrix} 10 & -11 & -52 \\ 25 & 40 & 5 \\ -20 & -23 & 14 \end{pmatrix}$$

an essential matrix. If so calculate a rotation matrix R and a skew symmetric matrix T so that $E = RT$.

6.4 A digital image is taken of ten points with known coordinates

$$X = \begin{pmatrix} -2 & 1 & 6 & 1 & -3 & 1 & 0 & -7 & -3 & -8 \\ -8 & -6 & 0 & -1 & 11 & 5 & -4 & 4 & 4 & -7 \\ 1 & 6 & 2 & 4 & -1 & 0 & 1 & 8 & 6 & 3 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

The image points are

$$x = \begin{pmatrix} 2/23 & 1/4 & 16/79 & 11/86 & -19/104 & -1/30 & 2/25 & -1/56 & 1/52 & -1/27 \\ -7/23 & -1/5 & 10/79 & -2/43 & 17/104 & 2/15 & -3/25 & -1/8 & -1/26 & -11/27 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

In an earlier exercise it was shown that the camera equation $x \sim PX$ can be rewritten as $T_x P X = 0$. This equation is linear in the 12 elements of the camera matrix P. Show that the equation $T_x P X = 0$ can be written as Mp where p is the row-stacked vector and M is a 3×12 matrix. Express M in terms of x and X. Thus each image-object point pair gives a linear constraint on p. Construct a big matrix

$$A = \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ M_1 0 \end{pmatrix}$$

by stacking these constraints (for the ten image-object point pairs above. Calculate the camera matrix P (using svd). Are the image coordinates corrected for internal calibration? Where is the focal point of the camera?

6.5 inversion of fundamental matrix.

6.6 Visa att P1 p2 till F är oberoende av coordinatsystem.