Vision 1 - Intro + Camera models + Calibration

1. Projection
   1.1. What size if the projection matrix P?
   1.2. If we have a 3D point M and a projection matrix P, how do we compute the corresponding pixel location where the camera will see the 3D point? Give steps.
   1.3. Given the projection matrix (P) and a 3D point compute where that point will be seen in the camera.
      1.3.1. $P = \begin{pmatrix} 1000 & 0 & 500 & 0 \\ 0 & 1000 & 500 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}; M = \begin{bmatrix} -400 \\ 300 \\ 1000 \end{bmatrix}$
      1.3.2. $P = \begin{pmatrix} 0 & -500 & -1000 & -190000 \\ 1000 & -500 & 0 & -120000 \\ 0 & 0 & 0 & -300 \end{pmatrix}; M = \begin{bmatrix} 60 \\ -900 \\ -200 \end{bmatrix}$
   1.4. If we have m (2D point on image plane), can we compute M (3D point)? (If yes: How?, if no: Why not?)
   1.5. What is the meaning/use of the individual intrinsic parameters $(f, \alpha_u, \alpha_v, s, u_0, v_0)$?
      K*A = $\begin{pmatrix} f*\alpha_u & f*s & u_0 & 0 \\ 0 & f*\alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

   1.6. What are a camera's extrinsic parameters? What do they model?
2. Distortion
   2.1. What is the name of the two distortion models we discussed?
   2.2. Why can the distortion models not be integrated into the projection matrix? (Why do these two components (distortion, projection) need to be handled separately?)
3. Calibration
   3.1. How does one of the three discussed camera calibration methods work (choose which one you want to describe)?
4. Projective space
   4.1. What is the inhomogeneous version of

$$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 6 \\ 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ 3 \\ 0.5 \end{bmatrix}, \begin{bmatrix} 1 \\ 10 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -4 \\ 1.1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \\ 10 \end{bmatrix}, \begin{bmatrix} -15 \\ 3 \\ 6 \\ 3 \end{bmatrix}$$

   4.2. Which of the points (p) lie on the line (l) (points and line are described in projective 2 space (P2))?

$$l = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}, p = \left\{ \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 110 \\ -40 \\ 10 \end{bmatrix} \right\}$$

   4.3. What is the geometric interpretation of a projective space point with a zero as last entry? (e.g., $[a \quad b \quad 0]^T, [c \quad d \quad e \quad 0]^T$)

Vision 2 - Stereo + Triangulation + Epipolar Geometry

1. What are the two main steps in the stereopsis process?
2. What is the correspondence problem?
3. What is the reconstruction problem?
4. What two general types of stereo did we talk about? What are the advantages disadvantages of both?
5. What do we know about the location of a 3D point when we have the image location where we see the point? (You can assume you know P.)
6. What is the epipole? How can you construct it?
7. What is an epipolar line? In which part, and how, can it help speed up the stereopsis process?
8. How can you construct the epipolar line (not only a direction)?
9. Compute ray:
   9.1. How do you compute the ray on which a 3D point corresponding to a location in the image (m) will lie (given P or K, A and H)? Represent the ray as point + direction or as a Plücker line.
   9.2. Compute point on line + direction or Plücker line given:

$$K * A = \begin{pmatrix} 1000 & 0 & 500 & 0 \\ 0 & 1000 & 500 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}; H = \begin{pmatrix} 0 & 0 & -1 & -40 \\ 1 & 0 & 0 & 30 \\ 0 & -1 & 0 & -300 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$P = \begin{pmatrix} 0 & -500 & -1000 & -190000 \\ 1000 & -500 & 0 & -120000 \\ 0 & -1 & 0 & -300 \end{pmatrix}; \ m = \begin{bmatrix} 767 \\ 650 \end{bmatrix}$$

10. Compute the two endpoints of the epipolar line (they do not need to be inside the image boundaries)

$$(K * A)_{left} = (K * A)_{right} = \begin{pmatrix} 1000 & 0 & 500 & 0 \\ 0 & 1000 & 500 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix};$$

$$H_{left} = \begin{pmatrix} 0 & 0 & -1 & -40 \\ 1 & 0 & 0 & 30 \\ 0 & -1 & 0 & -300 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

$$P_{left} = \begin{pmatrix} 0 & -500 & -1000 & -190000 \\ 1000 & -500 & 0 & -120000 \\ 0 & 0 & 0 & -300 \end{pmatrix};$$

$$H_{right} = \begin{pmatrix} 0 & 0.052336 & -0.99863 & -80 \\ 1 & 0 & 0 & 30 \\ 0 & -0.99863 & -0.052336 & -300 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

$$P_{left} = \begin{pmatrix} 0 & -446.98 & -1024.8 & -230000 \\ 1000 & -499.31 & -26.168 & -120000 \\ 0 & -0.99863 & -0.052336 & -300 \end{pmatrix};$$

$$m = \begin{bmatrix} 767 \\ 650 \end{bmatrix}$$

11. Draw a non-rectified stereo camera system. [SCANNING] Highlight:

- The epipole
- A 2D location in the left image of your choice
- The corresponding generated ray
- The corresponding epipolar line

12. Point reconstruction:
    - 12.1. How does the geometric method for point reconstruction work? What problem do we have to deal with here (influence of noise)?
    - 12.2. How does the algebraic method (linear alternative) work?
    - 12.3. In which situation is the algebraic method the better choice?
    - 12.4. Compute the 3D position of the point M causing m_left and m_right (using the method of your choice):

$$(K * A)_{left} = (K * A)_{right} = \begin{pmatrix} 1000 & 0 & 500 & 0 \\ 0 & 1000 & 500 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix};$$

$$H_{left} = \begin{pmatrix} 0 & 0 & -1 & -40 \\ 1 & 0 & 0 & 30 \\ 0 & -1 & 0 & -300 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

$$P_{left} = \begin{pmatrix} 0 & -500 & -1000 & -190000 \\ 1000 & -500 & 0 & -120000 \\ 0 & 0 & 0 & -300 \end{pmatrix};$$

$$H_{right} = \begin{pmatrix} 0 & 0.052336 & -0.99863 & -80 \\ 1 & 0 & 0 & 30 \\ 0 & -0.99863 & -0.052336 & -300 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

$$P_{left} = \begin{pmatrix} 0 & -446.98 & -1024.8 & -230000 \\ 1000 & -499.31 & -26.168 & -120000 \\ 0 & -0.99863 & -0.052336 & -300 \end{pmatrix};$$

$$m_{left} = \begin{bmatrix} 767 \\ 650 \end{bmatrix};\ m_{right} = \begin{bmatrix} 620 \\ 648 \end{bmatrix}$$

13. Fundamental matrix
    - 13.1. What does the fundamental matrix tell us about the relationship of a point in the left and one in the right image? How does the corresponding equation look?
    - 13.2. How can we estimate the fundamental matrix from image correspondences?
    - 13.3. If the correspondences are noisy what approach can we use to filter out the good correspondences? How does the process look if we use that algorithm (steps)?
14. What is rectification?

Vision 3 - Matching + Dense Stereo

1. Make a drawing of a stereo setup in case of a rectified system. Indicate the camera centres, a 3D point, the projection onto time image planes, the image planes, the information in the images required to compute the 3D point. [SCANNING]
2. What is disparity? How does one compute it?
3. How can you compute the depth given disparity (d), baseline (B) and focal length (f)?
4. Given the pixel location (measured relative to the image center) in the left image (p), disparity (d), baseline (B), and focal length (f) compute the 3D position:
$$p = \begin{bmatrix} -222 \\ 333 \end{bmatrix}; d = 77px; f = 1234px; B = 0.1m$$
5. What are the three fundamental problems that occur during stereo matching?
6. Stereo matching relies heavily on the reflectance function. What assumption is being made here? Is that in reality always given? What kind of objects/materials do not follow that assumption?
7. What is the assumption in the disparity gradient criteria?
8. Describe a basic dense stereo algorithm? How can you compare pixel locations? How can you compare regions around pixel locations? (We talked about four different simple metrics, give one.)
9. What are the four steps in the discussed stereo matching taxonomy?
10. What is the trade-off given by the size of the window in window based stereo matching algorithms?
11. What are the steps you need to take to compute a point cloud using dense stereo (given an initially unknown stereo camera system)?

Vision 4 - Other depth sensors + approaches

1. How can you make stereo be able to deal with homogenous structures in the scene? (What do you need to add?)
2. Describe how time of flight works (pulsed modulation and/or continuous wave modulation)?
3. Mention (and give a short description) of three methods for acquiring depth data that we discussed in class.
4. What two different scene coding methods (for structured light systems) did we discuss in class?

Vision 5 - Point cloud processing

1. What is a point cloud?
2. What is the difference between ordered and un-ordered pointclouds?
3. What are typical problems in/with point clouds? (That we then address with the algorithms discussed in class?)
4. Local features typically depend on a surround region. What is in important parameter there? What are the trade-offs?
5. How does the outlier removal method we discussed in class work? (Statistically motivated.)
6. What is the approach (discussed in class) to estimate normals?
7. What is the approach (discussed in class) to estimate curvature?
8. How does a voxelgrid filter work? (steps)
9. It makes sense to remove the table plane from the point cloud if you are only interested in objects on the table. How can you remove the table plane? What assumption does your described method make?
10. Euclidian clustering:
    10.1. How does Euclidian clustering work? What are the steps?
    10.2. What is this useful for?
11. What are the steps in the pose estimation pre-processing pipeline we discussed in class?

Vision 6 - Pose estimation, 3D -> 3D

1. What is a pose of an object? What is the mathematical structure?
2. How does object pose relate to camera pose?
3. If you have (at least three) point correspondence between model and scene, which algorithm can you use to compute the transformation? What are the steps in that algorithm?
4. Estimate the transformation T that aligns the five points in P with the corresponding points in Q:
   P = [1 3 1 4 5
   2 2 1 4 5
   3 1 1 4 2];
   Q = [2.6213 2.2071 1.7071 3.8284 2.4142
   2.6464 5.3536 3.5000 5.0000 7.0000
   5.3536 4.6464 3.5000 8.0000 8.0000];
   P = [1 3 1 4 5; 2 2 1 4 5; 3 1 1 4 2]
   Matlab formatting:
   Q = [[2.621, 2.207, 1.707, 3.828, 2.414];[2.646, 5.354, 3.5, 5.0, 7.0];[5.354, 4.646, 3.5, 8.0, 8.0];]
5. What are the two alignment algorithms/types discussed in class?
6. What is an algorithm for local alignment? What are the steps in the algorithm? When do you stop?
7. Which are the steps in the algorithm for global alignment discussed in class?
8. What are the trade-offs between using few (3) and many (>3) correspondences during RANSAC?
9. What are the names of the three shape features discussed in class?
10. What can you say in general about shape features used for global pose estimation? (What are their properties?)
11. Describe one of the shape features discussed in class (how do they work)?
12. How are these local (typically, as described in class) descriptors compared to each other?

Vision 7 - Pose estimation, 2D -> 3D

1. What are the three general approaches for single image-based pose estimation we discussed in class?
2. Homography
   2.1. What kind of objects is the Homography based pose estimation applicable to?
   2.2. What assumption do we need to make about the plane coordinate system for the Homography to be applicable?
   2.3. Once we have estimated the Homography how can we get the pose of the plane relative to the camera? What equation do we use here as a basis? What are the steps?
   2.4. Compute the rotation and translation that aligns the plane coordinate system with the camera coordinate system given the Homography H that maps from the plane to the image.

$$K * \bar{A} = \begin{pmatrix} 1000 & 0 & 500 \\ 0 & 1000 & 500 \\ 0 & 0 & 1 \end{pmatrix};$$
$$H_{left} = \begin{pmatrix} 0 & 2500 & -4500 \\ 2500 & 0 & 8000 \\ 0 & 0 & 1 \end{pmatrix};$$

3. P6P/P3P
   3.1. Which parameters will the P6P approach estimate?
   3.2. How does P6P work?
   3.3. What is the disadvantage of P6P over P3P? How can we employ the Iterative Approach to correct for this?
   3.4. What geometric property is the basis of the P3P approach (as discussed in class)? Draw the underlying geometry. [SCANNING]
   3.5. What is the basic equation for the Iterative Approach? How is the error defined?
4. Template based
   4.1. What is the basic approach of template/view-based approaches for pose estimation? (Basic steps without streamlining speeding up the process. Think simple.)
   4.2. How do we compare the images?
      4.2.1. What images do we typically use here for the comparison (colour, greyscale, intensity, …)?
      4.2.2. How do we then compare the actual images? Mention/describe one of the two methods discussed in class.
   4.3. How do we speed up the matching? Mention the two discussed approaches.
   4.4. How do aspect graphs work?
   4.5. Why is a pose correction needed at the end of discussed pose estimation process?
   4.6. What components are you correcting for?

Vision 8 - Uncertainty in Reconstruction

1. How are uncertainties commonly modelled in computer vision? Which model parameter quantifies the magnitude of the uncertainty?
2. In the lecture we made an approximation to the propagating function in order to let us propagate uncertainties using $\Lambda_y = \nabla f \, \Lambda_x \, \nabla f^T$ from input variable to output variables. Which approximation was used?
3. Propagate the uncertainties from uncertainties on vectors a and b to the uncertainty of $(a - b)^T(a - b)$. $a = [1,2]$ and $b = [3,3]$ The uncertainty on a and b are assumed independent. The covariance matrices corresponding to the vectors a and b are $cov_a = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ and $cov_b = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$ respectively.
4. When do we need to use the implicit function theorem to propagate uncertainties? Why do we need it? What does it let us do?
5. Describe the Monte-Carlo uncertainty propagation method in your own words. Use e.g., pseudo code to describe it
6. Name pros and cons of using the analytic method versus using the Monte Carlo method.
7. How can you approximate the uncertainty in the input variable?

Vision 9 - Kalman filter

1. What are the three elements of tracking?
2. What is the difference between the state of a system and the observations? Give an example.
3. What are the two typical steps in the filters we discussed? What do they do? (relation to time and to measurements)
4. What is the idea of the Bayesian filter? How do the equations look?
5. What is the implementation problem with the Bayesian filter?
6. What choices are made in the Kalman filter? (The Kalman filter is a specific version, with some restrictions, of the Bayesian filter)?
7. What are the two models used in the Kalman filter?
8. How does the Kalman filter deal with the fact that state space and measurement space are different?
9. There are two "external" uncertainties used in the Kalman filter. What do they represent/model? (What are they good for?)
10. What are the equations for the 1D Kalman filter?
11. What happens if you are very certain about your measurements? What does that mean for σ_m^2 (the measurement variance)? What does follow from that (trace through the Kalman equations)?
12. If your measurements are 3D positions coming out of a stereo process, how can you estimate good measurement uncertainties that are adaptive to the specific situation?
13. Extended Kalman filter:
    - 13.1. Which of the Kalman restrictions does the Extended Kalman filter remove?
    - 13.2. How does the Extended Kalman filter do that? (What is different between the Kalman filter and the Extended Kalman filter?)
14. What are the steps of the Prediction step of the Unscented Kalman Filter?

Vision 10 - Particle filter

1. How are probability density functions represented in particle filters?
2. What are the two (particle related) factors that influence the value of the pdf in a certain area?
3. What are the steps in a particle filter?
4. What probabilistic formulation is used in the update step? What is the meaning of that equation?
5. Why do we need to re-sample?
6. What is trade-off when it comes to the number of particles?
7. What kind of problems can the Particle Filter deal with that none of the discussed Kalman filters can deal with?