Exam Preparation Questions, Part 2[1]

For Chs. 6–10 of

Reinhard Klette: Concise Computer Vision
Springer-Verlag, London, 2014

[1]See last slide for copyright information.

Agenda

6.1-6.5

6.1. List five properties of high-quality cameras which are of importance for their use in computer vision applications.

6.2. What is the purpose of the Bayer pattern?

6.3. How to use a "Color Checker" for testing color accuracy of a given color camera?

6.4. Characterize (by means of a graphical sketch or informally by words) the barrel transform and the pincushion transform.

6.5. How is "linearity" defined for a gray-level camera?

6.6–6.8

6.6. Provide a graphical sketch of the model of a pinhole camera (including 2D camera and 3D coordinates) and specify the central projection equations. How are those equations in homogeneous coordinates?

6.7. What is the principal point of a camera? How are pixel locations (x, y) and undistorted coordinates (x_u, y_u) related to each other?

6.8. Define canonical stereo geometry (also known as "standard stereo geometry'). Specify epipolar lines and epipoles for this geometry.

6.9–6.10

6.9. Specify the point at infinity on the line $31x + 5y - 12 = 0$. Determine the homogeneous equation of this line. What is the intersection point of this line with the line $31x + 5y - 14 = 0$ at infinity?
Generalize by studying lines $ax + by + c_1 = 0$ and $ax + by + c_2 = 0$, for $c_1 \neq c_2$.

6.10. The fundamental matrix **F** describes binocular stereo geometry (in the general case, not only for standard stereo geometry). Specify the fundamental matrix **F** for canonical stereo geometry.

6.11

6.11. Consider a camera defined by the following 3×4 camera matrix:

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Compute the projections of the following 3D points (in world coordinates) with this camera:

$$\mathbf{P}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{P}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{P}_3 = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \end{bmatrix}, \text{ and } \mathbf{P}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

6.12–6.14

6.12. When calibrating a camera, we map points (X_w, Y_w, Z_w) in 3D world coordinates into pixel coordinates (x, y). List and briefly describe the sequence of coordinate transforms involved in this process.

6.13. What are intrinsic and extrinsic camera parameters? Give a general definition and provide examples for both types of parameters.

6.14. Show that the matrix **F**, defined by the (ideal) relation $p_R^\top \cdot \mathbf{F} \cdot p_L = 0$ (in homogeneous coordinates) between corresponding points, specifies a line $\mathbf{F} \cdot p_L$ in the image plane of the right camera.

Agenda

7.1–7.3

7.1. What is the Euler characteristic $\chi(S)$ if S is a surface of a torus, a surface of a sphere, a surface of a cube, and a Moebius band? Which of those surfaces is orientable, and which not?

7.2. Consider a surface $Z = f(X, Y)$. How is the gradient and the surface normal defined for Z? What is the unit normal? How to represent the unit normal by tilt and slant on the Gaussian sphere?

7.3. Consider the *ab* gradient space. What is the dual straight line to a gradient (a, b) in this space? What is the meaning of the dual straight line when modeling surface reflectance in gradient space with respect to a given light source?

7.4–7.7

7.4. Describe how 2^n different light planes are generated by using n projected binary patterns which follow the Gray code.

7.5. By using a 2D sketch for illuminating a surface point $P = (X_s, Y_s)$ and recording this point in a (1D) camera, describe how the coordinates X_s and Y_s can be calculated based on known system parameters.

7.6. How is “disparity” and “base distance” defined for a general two-camera system?

7.7. Consider a 3D point $P = (X_s, Y_s, Z_s)$ mapped into two corresponding points p_{uL} and p_{uR} in the left and right image in canonical stereo geometry. Derive step by step the 3D reconstruction formulas for X_s, Y_s, and Z_s.

7.8–7.9

7.8. Discuss the meaning of the non-linear distribution of depth layers (defined by integer disparities) for possible accuracies in depth calculations when using stereo vision.

7.9. Describe the workflow when applying 3PSM for shape recovery.

7.10. Provide a graphical sketch for a surface point illuminated by an ideal point-light source and recorded by a camera, and specify Lambertian reflectance by a formula, and also explain informally what the formula means for surface reflectance.

7.11. Provide a graphical sketch of a Lambertian reflectance map, showing the direction to the light source, the dual straight line to this direction (note: it has to follow defined geometric constraints), and a sketch of a few more isolines.

7.12–7.14

7.12. Describe the Lambertian reflectance map in spherical coordinates on the Gaussian sphere (hint: isointensity curves are circles in this case). Use this model to answer the question why two light sources are not yet sufficient for identifying a surface normal uniquely.

7.13. What is a possible way to calibrate the direction to a light source by using inverse PSM?

7.14. How to obtain a simply-connected surface patch (in 3D coordinates) based on a given dense gradient field for this surface patch and by applying local integration along a defined integration path?

7.15–7.16

7.15. Specify a data term and a smoothness term for specifying global integration (of a given dense gradient field) by following the Markov random-field model, i.e. for minimizing the error or energy $E_{total}(Z) = E_{data}(Z) + E_{smooth}(Z)$.

7.16. Why the statement "We also use Parseval's theorem..." when introducing the gradient-field integration algorithms (with deriving solutions in frequency space)?

Agenda

8.1–8.4

8.1. Why are SSD, SAD, and AD not recommended data cost functions for stereo matching? Give three reasons.

8.2. What is the "3D data cost matrix"? Why is there a region in this matrix identified by the comment "Constrained disparity range"?

8.3. Describe the census data cost function as an application of the Hamming distance between two vectors.

8.4. How is the stereo matching strategy "The winner takes all" defined? Refer to the 3D data cost matrix when describing this strategy.

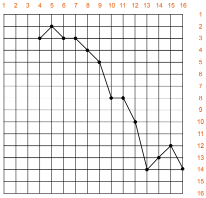
8.5–8.6

8.5. What is a “confidence measure” for stereo analysis? How does this approach differ from performance evaluation based on comparing with given disparity (or depth) ground truth? Specify two different confidence measures for stereo analysis results.

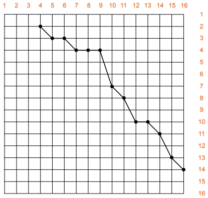
8.6. Outline the general dynamic programming methodology. Why does it make no sense to ask for a dynamic programming solution for ”The winner takes all”?

8.7

8.7. Are the shown graphs G_1 or G_2 epipolar profiles which represent disparity vectors?



G_1



G_2

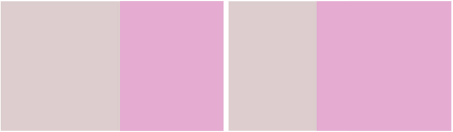
If so, which disparity vector? In reverse: Given are the disparity vectors

$$\mathbf{d}_1 = [f_4, f_5, \ldots, f_{16}]^\top = [1,0,1,2,3,0,1,1,0,1,2,3,2]^\top$$
$$\mathbf{d}_2 = [f_4, f_5, \ldots, f_{16}]^\top = [4,3,2,1,0,1,2,4,1,2,3,2,2]^\top$$

Draw the epipolar profiles defined by those two disparity vectors. Which profiles and which vectors satisfy the ordering constraint?

8.8–8.10

8.8. Consider the following rectified stereo pair of images:



Assume the AD data-cost function and "The winner takes all" or DPM with ordering constraint. Describe the expected outcome of those two stereo matchers.

8.9. What was the motivation for suggesting the penalty

$$c_2(p, q) = \frac{c}{|B(p) - B(q)|}$$

for DPM with smoothness term?

8.10. How is the 3D data cost matrix used (repeatedly) for multi-scanline DPM with smoothness constraint (also known as *semi-global matching*)?

Agenda

9.1–9.3

9.1. Explain the concept of "invariance w.r.t. changes in the scene", at first briefly in general words, and then by means of a positive and a negative example.

9.2. How are keypoints defined in LoG or DoG scale space? What is the "disk of influence" of a keypoint detected this way? How can this disk of influence be used for calculating sparse 3D flow?

9.3. Explain the general outline of RANSAC, and illustrate the general concept for detecting a straight segment in a given "noisy image" of a straight segment. Structure your presentations into "initialization", "test", "consensus set", "cardinality check", "stop criterion", and "refined model".

9.4–9.7

9.4. How can RANSAC be used for estimating an affine transform between two sets of points in the plane?

9.5. Describe the selection procedure for keypoints for SIFT, and which descriptor is used for this image feature. Mention the motivations for the individual steps.

9.6. What are similarities, and what are differences between local binary patterns (LBP) and the BRIEF descriptor?

9.7. How to define a "repeatability measure" for a given feature detector and a given image sequence?

9.8–9.9

9.8. Explain the Newton-Raphson iteration scheme for detecting a zero of a smooth unary function. How has the basic idea of this scheme be followed in the design of the Lucas-Kanade tracker? Describe in general, no formulas needed.

9.9. Explain the general outline of a particle filter. Structure your presentation into “definition of particle space”, “initialization”, “update model”, “dynamic model (generation of clouds of particles)”, “observation model (weights of particles)”, “condensation”, and “decision for winning particles”.

9.10–9.11

9.10. For lane-border detection it was suggested in the lectures to generate a bird's-eye view by using a homography, defined by four marked points being corners of a trapezoid in the image, but actually corners of a rectangular region on the road. Specify this homography.

9.11. In which way did the discussed lane-border detection technique specify the general steps of a particle filter? No formulas needed, just an informal description is fine; but using formulas for supporting your discussion is also a possibility (but then used parameters need to be specified).

9.12–9.15

9.12. What is the equation of a continuous linear dynamic system, and what is the equation of a discrete linear dynamic system (not considering noise or control). How are both equations related to each other.

9.13. A discrete linear system with control and noise is defined by two equations for $\mathbf{x}_t$ and $\mathbf{y}_t$. Name all the symbols in those two equations.

9.14. When predicting the system state at time t (for time t), which knowledge is available at that moment for performing this prediction?

9.15. Which error is minimised by the optimal Kalman gain? Describe in words rather than giving just a formal expression.

Agenda

10.1–10.4

10.1. Explain the terms “classifier”, “weak classifier”, and “strong classifier”. What does it mean “to train” a classifier? What is “supervised learning”?

10.2. When is PR $>$ RC for an object detector? When is MR $>$ FPPI?

10.3. Which HoG descriptor results for a bounding box in which all pixels have constantly the same value?

10.4. Explain (in general, and by means of examples) how integral images can be used for a time-efficient calculation of sums S_W as defined in Equation (10.7) in the book.

10.5–10.7

10.5. Provide a pseudo-code for a procedure for sliding masks of stepwise increased sizes through a placed window in a given image. Now embed this procedure into a main program describing the move of the placed window through the whole image (with calling the procedure for every position of the placed window).

10.6. Explain how AdaBoost is used to derive a strong classifier from a set of given week classifiers.

10.7. Explain the meaning of the multiplications in the sum in Equation (10.16) in the book. Why do we start the iteration with weights $\omega_1(i) = 1/m$, and not, for example, with weights $\omega_1(1) = 0.1$, $\omega_1(2) = 0.2$, $\omega_1(3) = 0.3$, and so forth? Why do we stop with the iteration when arriving at $W_{a(T)} \geq 0.5$? Could we also already stop when arriving at $W_{a(T)} \geq 0.45$? Or when arriving at $W_{a(T)} \geq 0.55$?

10.8–10.11

10.8. Why is there a minus sign in Equation (10.40) in the book?

10.9. Describe how to train a random forest. Why is it called "random"? Do the trees of the forest grow during training, or are they already of defined structure when beginning with the training of a forest? Which constraints in the training process define leaf nodes?

10.10. In the given description of a Hough forest, objects are parametrised by the object's centroid. Modify the definition of a Hough forest for the case where objects are parametrised by a vertical line segment (say, their vertical "main axis", with given endpoints for this axis).

10.11. Explain the meaning of entries in Table 10.2 in the book.

Copyright Information

This slide show was prepared by Reinhard Klette
with kind permission from Springer Science+Business Media B.V.

The slide show can be used freely for presentations.
However, *all the material* is copyrighted.

In case of citation: just cite the book, that's fine.