Exam Preparation Questions, Part 1[1]

For Chs. 1–5 of

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

[1] See last slide for copyright information.

Agenda

1.1–1.3: Basics

1.1. Show the correctness of

$$\sigma_I^2 = \left[\frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} I(x,y)^2 \right] \quad - \quad \mu_I^2$$

1.2. How was the contrast $C(I)$ defined, for image I. How does the contrast change when applying a local mean or local median operator?

1.3. What are phase and magnitude values of the complex numbers $z_1 = 5 - 4i$, $z_2 = 0$, and $z_3 = 7i$, for $i = \sqrt{-1}$?

1.4–1.6: Fourier Transform

1.4. How is the 2D DFT defined for an $N \times N$ image I? What is the value $\mathbf{I}(0,0)$ in frequency space?

1.5. Why are powers of the Mth root of unity located on the unit circle in the complex plane?

1.6. Assume an $N \times N$ image as input for the 2D DFT, for even N. Prove that a multiplication with $(-1)^{x+y}$ in the spatial domain causes a shift by $N/2$ (both in row and column direction) in the frequency domain.

1.7–1.9: Color

1.7. How is the *xy* color space of CIE defined? Start with *explaining* the meaning of tristimulus values X, Y, and Z, and end with a graphical sketch of the *xy*-diagram.

1.8. Define the RGB color space. Where are the gray-levels located? How is saturation and intensity defined for one of those colors? What are the saturation values for gray-levels?

1.9. How is *hue* defined for a color in RGB space? You do not need to provide a formula; just give an informal explanation how to define a hue value for RGB colors.

Agenda

2.1–2.3

2.1. Linear local operators are those which can be defined by a convolution. Classify the following whether they are linear operators or not: box, median, histogram equalization, sigma-filter, Gauss filter, and LoG.

2.2. Equalization of color pictures is an interesting area of research. Discuss why the following approach is expected to be imperfect: do histogram equalization for all three color (e.g., RGB) channels separately; use the resulting scalar pictures as color channels for the resulting image.

2.3. Derive the formula for linear scaling from given values u_{min} and u_{max}.

2.4–2.6

2.4. Specify exactly how the integral image can be used for minimizing run time for a box filter of large kernel size.
2.5. Following the derivation of a Laplace mask in the lecture, what could be a filter kernel for the quadratic variation (instead of the one derived for the Laplace operator)?

2.6. The sigma-filter replaces $I(p)$ by $J(p)$ as defined in the lecture. The procedure uses the histogram $H(u)$ computed for values u in the window $W_p(I)$ which belong to the interval $[I(p) - \sigma, I(p) + \sigma]$. Alternatively, a direct computation can be applied:

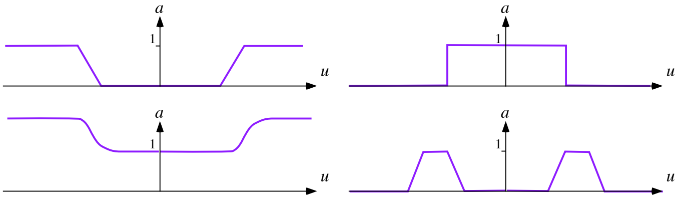
$$J(p) = \frac{\sum_{q \in Z_{p,\sigma}} I(q)}{|Z_{p,\sigma}|}$$

where $Z_{p,\sigma} = \{q \in W_p(I) : \; I(p) - \sigma \leq I(q) \leq I(p) + \sigma\}$. Analyze possible advantages of this approach for small windows.

2.7–2.8

2.7. What is the "symmetry property" in the frequency domain, and why do we have to ensure the validity of this property when designing a Fourier filter?

2.8. Sketch, as in the figure



filter curves in the frequency domain which might be called "exponential low-emphasis filter" and "ideal band-pass filter".

2.9-2.11

2.9. Specify the 2D Gauss function (the zero-mean version is sufficient). Describe how a discrete filter mask (for smoothing) can be derived from this function.

2.10. How is a LoG scale space defined? How is a DoG scale space defined?

2.11. How is the gradient for an image I defined, and how is the output of the Sobel operator related to the gradient? Demonstrate values of the Sobel operator for examples of 3×3 image windows.

Agenda

3.1

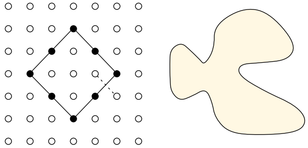
3.1. Consider *6-adjacency* in images, for pixel location $p = (x, y)$: In an even row (i.e. y is even) use

$$A_6^{even} = A_4 \cup \{(x+1, y+1), (x+1, y-1)\}$$

and in an odd row use

$$A_6^{odd} = A_4 \cup \{(x-1, y+1), (x-1, y-1)\}$$

Consider the case on the left in the figure. Discuss the result. Consider a chess-board type of binary image. What is the result?



3.2–3.6

3.2. Explain why using 4-adjacency alone for a binary image leads to topological problems. Explain why using 8-adjacency alone is also not a topologically sound solution.

3.3. What is a component or region in an image?

3.4. Specify a way for defining K-adjacency for a gray-level image? How is K-adjacency motivated by Euclidean topology in the plane?

3.5. What is the stop criterion in the border tracing algorithm of Voss?

3.6. Why is the number of pixels on a border not a multi-grid convergent estimator for the perimeter of a digitized set?

3.7–3.9

3.7. Specify two ways how to define curvature at a point on a continuous arc in the Euclidean plane.

3.8. Select a simple (e.g. 4×4) gray-level image, a pixel adjacency set A, and explain the concept of a co-occurrence matrix by using the chosen example.

3.9. Define moments of order $a + b$ for a set $S \subset \Omega$ and a given image I defined on Ω. How is the centroid of S defined? Assume that I is equal to 0 everywhere in S. What is the centroid of S in this case?

3.10 - 3.11

3.10. Why is the straight-line representation $y = ax + b$ not suitable for defining a parameter space for detecting line segments in an image? Which representation has been suggested by Duda and Hart instead?

3.11. Assume that we only map two pixel locations from an image into the $d\alpha$-Hough space. Sketch the resulting Hough space and explain how the straight line is detected defined by those two pixel locations. – Now assume that we have a "noisy line segment" in the image. How can this line segment be detected in Hough space?

3.12. Explain how image gradient information can be used to enhance accuracy and speed of circle detection using the concept of the Hough transform.

Agenda

4.1–4.4

4.1. What is the aperture problem for detecting local displacements (i.e. 2D motion) in an image?

4.2. Describe cases where 2D motion (projected into an image) does not correspond to the visible optical flow.

4.3. Calculate the coefficients of the Taylor expansion of the function $I(x, y, t) = 2x^2 + xy^2 + xyt + 5t^3$ up to derivatives of the 3^{rd} order.

4.4. What is the optical flow equation (also known as Horn-Schunck constraint)? Specify the defined straight line in uv-velocity space.

4.5–4.7

4.5. When presenting the optical flow problem as a labeling problem, defined by data-cost and smoothness-cost terms, what are those two terms in case of the Horn-Schunck algorithm?

4.6. Explain as a general concept how LSE optimization is used for solving a minimization problem such as defined by the Horn-Schunck approach towards optical flow? What is LSE standing for?

4.7. Provide a pseudo-code for the the Horn-Schunck algorithm based on the mathematical solution

$$
\begin{aligned}
u_{xy}^{n+1} &= \bar{u}_{xy}^{n} - I_x(x,y) \cdot \frac{I_x(x,y)\bar{u}_{xy}^{n} + I_y(x,y)\bar{v}_{xy}^{n} + I_t(x,y)}{\lambda^2 + I_x^2(x,y) + I_y^2(x,y)} \\
v_{xy}^{n+1} &= \bar{v}_{xy}^{n} - I_y(x,y) \cdot \frac{I_x(x,y)\bar{u}_{xy}^{n} + I_y(x,y)\bar{v}_{xy}^{n} + I_t(x,y)}{\lambda^2 + I_x^2(x,y) + I_y^2(x,y)}
\end{aligned}
$$

4.8–4.10

4.8. For a given pixel location p, let Q be the point in uv-velocity space where the straight line define by the optic-flow equation intersects with the line in direction $[I_x, I_y]^\top$. Derive (step by step) a solution for Q (i.e an expression in terms of I_x, I_y, and I_t).

4.9. What is the basic idea behind the Lucas-Kanade algorithm? How does this work for two adjacent pixels, and how for a whole neigborhood of k pixels?

4.10. Show for the Lucas-Kanade algorithm that the matrix inversion in the equation

$$\mathbf{u} = (\mathbf{G}^\top \mathbf{G})^{-1} \mathbf{G}^\top \mathbf{B}$$

fails if image gradients in the selected neighborhood are parallel to each other. Is this possible to happen for real image data?

Agenda

5.1–5.4

5.1. Specify one technique for image binarization which is more advanced than just using one given global threshold T.

5.2. Assume that segments are generated by seed growing. How to ensure independence of generated segments from the choice of a selected seed in a segment?

5.3. Specify a pseudo-code for the recursive labeling of all pixels in a defined region (defined by an equivalence relation on the given image).

5.4. Describe the basic ideas behind mean-shift image segmentation. An informal description is sufficient.

5.5–5.7

5.5. Assume that we only use gray-levels in feature space. Specify a variant of a mean-shift segmentation algorithm for this particular case.

5.6. Define the image segmentation problem as a labeling problem, specified by a data-cost term and a smoothness-cost term. Provide one example of a relevant data term.

5.7. Describe the message-update process (either just by an informal explanation, or by providing the update equation and explaining the symbols involved in your equation) of belief-propagation.

5.8

5.8. The lectures suggested to replace the simple Potts smoothness term by an alternative smoothness term in BP segmentation. For example, if μ_1 and μ_2 are the intensity means in adjacent segments, then constant c in equation

$$E_{smooth}(l - h) = E_{smooth}(a) = \begin{cases} 0 & \text{if } a = 0 \\ c & \text{otherwise} \end{cases}$$

can be replaced by a term where c is scaled in dependency of difference $|\mu_1 - \mu_2|$.
Specify and discuss modified smoothness functions based on this equation which include data characteristics into the smoothness term.

5.9–5.10

5.9. Consider the dissimilarity measure $\mathcal{D}$ defined by equation

$$\mathcal{D}(A, B) = \frac{|(A \cup B) \setminus (A \cap B)|}{|A \cup B|}$$

Show that this measure satisfies the following two properties of a metric d on a base set S:

1. $f = g$ *iff* $d(f, g) = 0$
2. $d(f, g) = d(g, f)$

for all $f, g, h \in S$. Here, S is the family of sets of pixels.
Optional: A proof of the triangularity property would give you *additional* marks.

5.10. What is "temporal consistency" when segmenting a sequence of images, and give one example of a procedure which might help to ensure this type of temporal consistency.

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.