Guide to answers for written examination in TSBB09 Image Sensors, 2019-04-23

Maria Magnusson, maria.magnusson@liu.se
Robert Forchheimer, robert.forchheimer@liu.se,

PART I: STANDARD CAMERAS & IR SENSORS

Exercise 1 Blue light gives a sharper image than red light.
The width of the point-spread function is given by $\Delta x \approx 1.22 \cdot \lambda \cdot f_L/D$, where λ is the light wavelength. Consequently, a short wavelength gives a narrower point-spread function and a sharper image. Blue light has a shorter wavelength than red light.

Exercise 2 The refraction index of matter (e.g. lenses) is wavelength dependent. Therefore, a ray of white light is decomposed into rays of different colors that intersect the image plane at different points.

Exercise 3 The fill factor is the percentage of the total sensor area which is light sensitive. The larger sensor area, the more photons can be collected, and the less photon noise.

Exercise 4 Short-wave infrared, Mid-wave infrared, Long-wave infrared.

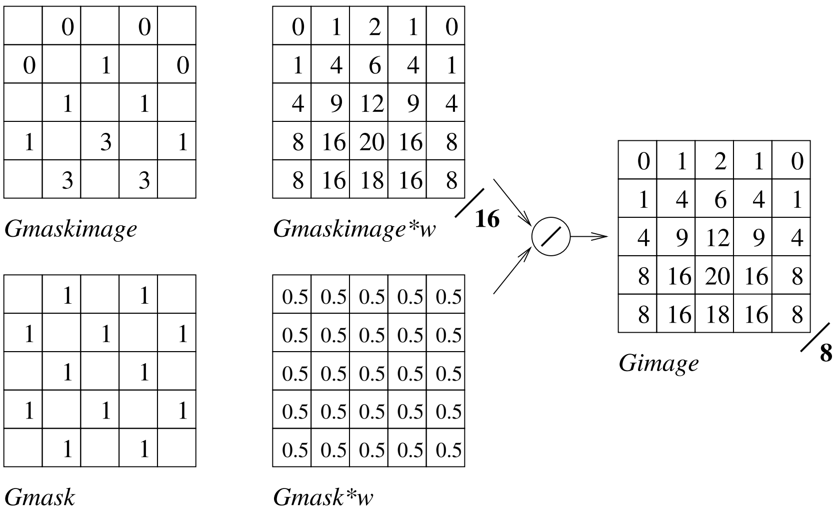
Exercise 5 At each pixel it measures the light energy in multiple wavelength bands. An RGB color camera is a trivial example, but more commonly the term is used for more than three wavelength bands that are both in the visual and IR range of the spectrum.

Exercise 6 The two curves show how much energy a surface radiates (vertical axis) in different wavelengths, for two different temperatures. The higher curve, which peaks at a shorter wavelength, is for a higher temperature.

Exercise 7

$$\begin{cases} \hat{f}_A = c(b_A + d) \\ \hat{f}_B = c(b_B + d) \end{cases} \Rightarrow \begin{cases} c = (\hat{f}_B - \hat{f}_A)/(b_B - b_A) \\ d = (\hat{f}_A \cdot b_B - \hat{f}_B \cdot b_A)/(\hat{f}_B - \hat{f}_A) \end{cases}$$

Exercise 8



PART II: GEOMETRY AND MULTIPLE VIEWS

Exercise 9 The two (straight) parallel lines in B show up as:
- Two straight parallel lines in A.
- Two straight, but not necessary parallel, lines in H.

Exercise 10 Otherwise, image points in the two images are not related by a global geometric transformation that only depends on the rotation between the two views. Instead points in the two images will move also depending on their distance to the camera. It may happen that a point that is visible in one image will be occluded in the other image.

Exercise 11 A pinhole camera without any lens distortion maps 3D lines to 2D lines in the image. The calibration for lens distortion can therefore be based on adjusting the lens distortion parameters to make the lines in the image as straight as possible.

Exercise 12 **C** is a 3×4-matrix that can be determined up to a scale factor, i.e. 11 unknowns in **C** need to be determined. One corresponding point-pair on the 3D calibration object and the image, contributes with 2 equations to solve for **C**. Consequently, the minimum number of 3D points on the calibration object is 6.

Exercise 13 The parameters α and β denotes the scaling in the u- and v-direction, respectively. If $\alpha \neq \beta$, the proportions between height and width

will not be preserved in the real image plane (measured in pixels). For example, a quadrat in the world, orthogonal to the optical axis, will become a rectangle with different height and width in the real image.
γ is the skewing parameter. If $\gamma \neq 0$, the real image plane (measured in pixels) will be skewed. For example, a quadrat in the world, orthogonal to the optical axis, will become a skewed rhombus in the real image.

Exercise 14 $||\mathbf{r}_1|| = ||\mathbf{r}_2|| = 1$.

Exercise 15

$$\begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} = \mathbf{K^{-1}} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} , \quad \begin{pmatrix} x_s \\ y_s \\ z_s \end{pmatrix} = \begin{pmatrix} x_n/\sqrt{x_n^2 + y_n^2 + 1} \\ y_n/\sqrt{x_n^2 + y_n^2 + 1} \\ z_n/\sqrt{x_n^2 + y_n^2 + 1} \end{pmatrix} ,$$

where $(x_n, y_n, 1)$ are coordinates on the normalized image plane and (x_n, y_s, z_s) are coordinates on the unit sphere.
The transformation is performed for two images and their respective coordinates on the unit sphere are send to Procrustes algorithm, which determines the rotation between the two images.

Exercise 16 The camera is rotated $60°(= \arccos(0.5))$ around the z-axis of the world coordinate system. Consequently the optical axis is directed in the same direction as the z-axis of the world coordinate system. The origin of the world is at $(20, 35, 75)$ in the camera coordinate system.

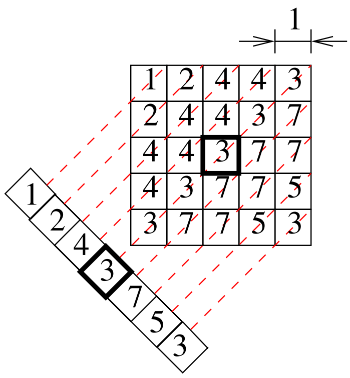
PART III: NON-STANDARD IMAGE SENSORS

Exercise 17 distance $= t \cdot c/2 = 11 \cdot 10^{-9} \cdot 3 \cdot 10^8/2 = 1.65$ m.

Exercise 18 When $a0 \tan\beta = b0 \tan\alpha$ (Scheimpflugs condition) is fulfilled, the whole sensor is in perfect focus.

Exercise 19 The ramp-filter is an infinitely long filter. In CT, it is convolved with projection data. Suppose that the size of the projection data is N. Then the ramp-filter can be truncated to the size 2N-1 (or larger), without disturbing the convolution result for the N points.
A convolution can be performed via multiplication in the DFT domain. Before DFT, the size of the projection data and the rampfilter must be the same. Therefore the projection data must be zero-padded to at least the size 2N-1.

Exercise 20



Exercise 21 In figure b) and c), only interpolation in the horizontal direction is needed. Figure a) demands interpolation in the vertical direction, too.

Exercise 22

- Household paper only gives diffuse reflection, therefore choose $k_d = 1$ and $k_s = 0$.
- A perfect mirror provides only specular reflection, therefore choose $k_d = 0$ and $k_s = 1$. A perfect mirror reflects light in a very small loob. Therefore, choose a very high value for n.
- Polished steel gives a little diffuse and much specular reflection, therefore choose $k_d = 1 - a$ and $k_s = a$, where e.g. $a = 0.8$. Polished steel reflects light in a much broader loob than the perfect mirror. Therefore, choose e.g. $n = 2,\ 3,\ 4$, or 5.

Exercise 23 A coded aperture replaces the (approximately) circular camera aperture with a more complex aperture pattern. In practice this is achieved by putting a mask onto the camera lens where the mask contains the aperture pattern. When taking a picture, defocused objects will then be convolved with a scaled image fk of the aperture pattern. Reconstructing a sharp image x from the unsharp image y is done by trying out various values for x and scales k (at different local positions) and minimizing the error $|f_k * x - y|^2$.
ExoticCameras2018 Slides 45-52

Exercise 24

$$
\begin{aligned}
(r, y, 1)^T &= k\mathbf{C}^{-1}(s, t, 1)^T \\
(r_1, y_1, 1)^T &= k\mathbf{C}^{-1}(s_1, t_1, 1)^T = k\mathbf{C}^{-1}(122, 116, 1)^T \\
(r_1, y_1, 1)^T &= k(38.5018, 59.4935, 1.1154)^T \Rightarrow (r_1, y_1) = (34.5184, 53.3383) \\
(r_2, y_2, 1)^T &= k\mathbf{C}^{-1}(s_2, t_2, 1)^T = k\mathbf{C}^{-1}(132, 416, 1)^T \\
(r_2, y_2, 1)^T &= k(-19.2492, -224.8005, 0.5464)^T \Rightarrow (r_2, y_2) = (-35.2291, -411.4211)
\end{aligned}
$$

$$\text{Answer : width} = \sqrt{(r_1 - r_2)^2 + (y_1 - y_2)^2} = 469.9638 \text{ mm} \approx 470 \text{ mm}.$$