Guide to answers for written examination in TSBB09 Image Sensors, 2016-08-19

Klas Nordberg, klas.nordberg@liu.se,
Maria Magnusson, maria.magnusson@liu.se

PART I: STANDARD CAMERAS & IR SENSORS

Exercise 1 A photon of wavelength λ_1 has higher energy than a photon of wavelength λ_2 when $\lambda_1 < \lambda_2$. This means that there must be fewer detected photons from the first light source than from the second source to give the same total energy; $n_2 > n_1$. See lecture A, slide 18.

Exercise 2 See lecture B, slide 42.

Exercise 3 See lecture A, page 73.

Exercise 4 See lecture A, slides 56.

Exercise 5 See lecture B, slide 52.

Exercise 6 The visible spectral range.

Exercise 7 CCD: charges are shifted out row by row and column by column from the sensor chip, i.e., all of the image needs to be transferred to circuits where charges are converted to voltage and then to a binary representation. In a CMOS camera each individual detector can be connected to a row and column readout line, similar to a RAM memory, from which conversion to a binary representation is done. This makes is possible to read out smaller parts of the image, but at a higher rate in time compared to a whole image. Due to the long readout line, from the detector to the conversion stage, however, the basic CMOS has higher noise levels than CCD. By using additional transistors at each detector element (APS), the CMOS design can be improved in terms of noise.

Exercise 8 See lecture B, slide 14-19.

PART II: GEOMETRY AND MULTIPLE VIEWS

Exercise 9 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.

Exercise 10 See lecture F, slide 25.

Exercise 11 γ represents the skewness of the optical transformation from the 3D space to the image plane, which in the ideal case should be $= 0$. See lecture E, slide 16.

Exercise 12 In order to avoid large geometric distortion when one image is transformed to the coordinate system of another. See lecture X, slide 17-.

Exercise 13 The camera coordinate system is transformed with

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 10 \\ 0 & 1 & 0 & 20 \\ 0 & 0 & 1 & 30 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Because transformation of a coordinate system with $\mathbf{M}$, corresponds to transform its coordinates with

$$\mathbf{M^{-1}} = \begin{pmatrix} 1 & 0 & 0 & -10 \\ 0 & 1 & 0 & -20 \\ 0 & 0 & 1 & -30 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \text{ the desired matrix is } [\mathbf{R}\ \mathbf{t}] = \begin{pmatrix} 1 & 0 & 0 & -10 \\ 0 & 1 & 0 & -20 \\ 0 & 0 & 1 & -30 \end{pmatrix}.$$

Exercise 14 See lecture F, slides $59 - 60$.

Exercise 15 See lecture F, slides $7 - 12$.

Exercise 16 The 3×3 calibration matrix $[\,\mathbf{h}_1\,\mathbf{h}_2\,\mathbf{h}_3\,]$ is determined through calibration with one planar calibration pattern.
The relation between the camera coordinate system and the world coordinate system at the actual planar pattern is described by $[\mathbf{Rt}] = [\,\mathbf{r}_1\,\mathbf{r}_2\,\mathbf{t}\,]$, were $\mathbf{r}_1$ and $\mathbf{r}_2$ are orthonormal rotation vectors and $\mathbf{t}$ is the translation vector.
$\mathbf{A}$ is the camera calibration matrix, describing the inner parameters of the camera.
The two constrains are: $\mathbf{r}_1 \cdot \mathbf{r}_2 = 0$ and $||\mathbf{r}_1|| = ||\mathbf{r}_2|| = 1$.
They give two constraints for solving the parameters in $\mathbf{A}$. See lecture E, slides $29 - 31$.

PART III: NON-STANDARD IMAGE SENSORS

Exercise 17 A line camera can have a very high resolution (number of pixels) along the line. A high resolution across the line can be accomplished by reducing the motion of the push broom camera or by increasing the number of line images that are taken per time unit. Finally, by splitting the light beam through a prims in to a spectrum, the line image can converted into a hyperspectral 2D image.

Exercise 18 If the reflected laser pulse comes from a border area which is dark on one side and bright on the other side, the detected pulse shape becomes distorted. Then there will be a slight change in detected max position which will show up as a height difference in the range image. Consequently, the range deviations will occur along the border around the M letter as well as on both sides of the surrounding rectangle. See lecture I, slide 32.

Exercise 19 The X-ray source and the detector rotates continuously along a circle around the patient. The X-ray source and the detector are mounted 180 degree apart with the patient in between. The patient is translated slowly through the circle.

Exercise 20

$$\cos\varphi = \frac{\mathbf{L}\cdot\mathbf{N}}{\|\mathbf{L}\|\ \|\mathbf{N}\|}$$
$$\cos\theta = \frac{\mathbf{R}\cdot\mathbf{V}}{\|\mathbf{R}\|\ \|\mathbf{V}\|}$$

Exercise 21 A diffuse surface reflects the light in all directions and one of these light rays has a direction straight into the camera. A specular surface, on the other hand, reflects the light mainly in one direction, and then there is a risk that the reflected light does not reach the camera.

Exercise 22 The laser sheet gives a curved line on the sensor $f(s,t)$. For every column t, the maximum intensity $f_{max}(t)$ gives the position of the laser line $s_L(t)$. If $f_{max}(t)$ is close to zero, we have occlusion due to one of the cases described in the previous exercise.

Exercise 23 A position (x, y) in the image corresponds to
$r = x\cos(\phi)+y\sin(\phi)$ in the detector. The position $(x, y) = (2, -1)$ corresponds to $r = 2\cos(-26.565°) + (-1)\sin(-26.565°) \approx 2.236$.
Linear interpolation then gives $q(2.236) = 3\cdot(1 - 0.236) + 2\cdot 0.236 = 2.764$, i.e. the backprojected value is 2.764.

Exercise 24 Call the 3D volume $f(x, y, z)$. Calculate derivative volumes

$$\frac{\partial f(x,y,z)}{\partial x} = f(x,y,z) * \mathrm{sobel}_{\mathrm{x}},$$
$$\frac{\partial f(x,y,z)}{\partial y} = f(x,y,z) * \mathrm{sobel}_{\mathrm{y}},$$
$$\frac{\partial f(x,y,z)}{\partial z} = f(x,y,z) * \mathrm{sobel}_{\mathrm{z}},$$

which gives the gradient in every voxel as

$$\nabla(x,y,z) = (\frac{\partial f(x,y,z)}{\partial x}, \frac{\partial f(x,y,z)}{\partial y}, \frac{\partial f(x,y,z)}{\partial z})$$

and the surface normal in every voxel as

$$\mathbf{n}(x,y,z) = -(\nabla(x,y,z)/|\nabla(x,y,z)|).$$

Assume that the light source is located in the direction $\mathbf{L}$ and the camera is located in the direction $\mathbf{V}$. Suppose that the desired image size is $N \times N$. From the camera position, step into the 3D volume $f(x, y, z)$ along $N \times N$ straight lines, one for each pixel in the 2D image. When a voxel value larger than a predefined threshold is reached, stop there and calculate the shading value, which is a function of $\mathbf{n}(x, y, z)$, $\mathbf{L}$ and $\mathbf{V}$. The calculated shading value gives the pixel intensity.