Technische Universität Berlin

Wintersemester 2018/2019

Digital Image Processing

Exercise 01

Jochen Jacobs - 373163
Herbert Potechius - 353420
Sourabh Raj - 397371
Sangeetha Reji - 406043

November 1, 2018

1 Theory

1. **What is a digital image?**
   A digital image is a 2D-representation of a 3D-scene. During the *image formation* process, points of the 3D-scene will be projected on an image plane with the equation (1).

$$x = PX \tag{1}$$

   The variables mean the following:

   - X: 3D-point of the scene
   - x: 2D-point; projection of X on the image plane
   - P: projection matrix (2)

$$P = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \tag{2}$$

   Interior orientation variables (principle distance α_x, principle point$(x_0, y_0)^T$, skew facor s) and exterior orientation variables (rotation R, translation t) determine how X will be projected.
   The continuous values x has to be discretized (e.g. to an image with the pixel size of 320x240 px) and to be quantised (e.g to a 24-bit RGB-image).
   The resulting image can be represented as a *Digital Image Function* f which maps spatial positions *(x, y)* to intensity values:

$$f : \mathbb{N}^2 \to \mathbb{N}^c \tag{3}$$

   In most cases we want to describe greyscale images (*c=1*) or color images (*c=3*). c describes the number of color channels. The more specific representation for *c=1* looks like this:

$$f : [1, N_x] \times [1, N_y] \to [0, 255] \tag{4}$$

   In this case, we described the image as *set of matrices*. Other representations are the following:

   - chain codes for object boundaries
   - topological data structures like graphs
   - hierarchical representations like pyramids and trees

2. **What does the paradigm "bottom-up processing" mean?**
   Bottom-up processing is a data-driven method (in contrast to the experience-driven method of the *Top-down processing*). During the image analysis, object hypothesis will be constructed by linking evidence in the image. These evidences are patterns of light and dark areas. No object patterns are known, i.e., information are only available in the stimulus itself. A pure bottom-up processing can not tell that there is a cube in fig. 1, because the concept of a cube is not given.

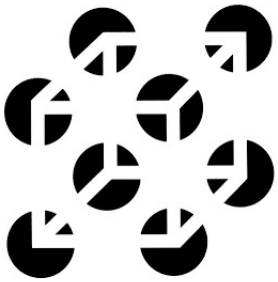

Figure 1: Top-down processing example

3. **State at least three fundamentally different image sources!**

   (a) *digital-camera*: The light of the 3D-scene passes through an IR-blocking filter and a color filter. The color filter makes sure, that only one wavelength (either red, green or blue) reaches the color blind sensor. This sensor converts the light into electricity. This gives us the raw image. Each pixel describes the intensity of one color (R, G or B). To get a 3-channel image, i.e, to get a green and a red value for a pixel that contains only a blue value, the values will be interpolated with the help of the neighbouring pixels.

   (b) *Sonar*: It stands for *sound navigation and ranging* and is a method that uses the reflected wave of a sound to calculate the distance of objects. The resulting image maps a 2D-coordinate to a distance value.

   (c) *MRI*: It stands for *Magnetic resonance imaging* and is a medical imaging technique for the representation of the structure of human tissue and organs. This method uses electromagnetic fields.

Image Sources

- *Figure 1*: http://f10323jdelapena.blogspot.com/2010/09/top-down-processing.html