Solutions for Exam Computer Vision

November 7, 2003, 9:00 hrs

Problem 1. (**2.5 pt**) Let X be a binary image X and A a structuring element as in Fig. 1.

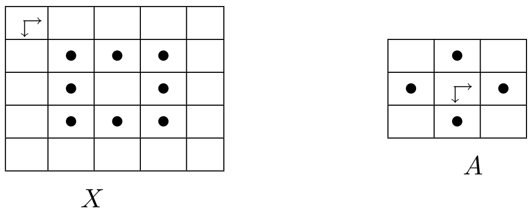
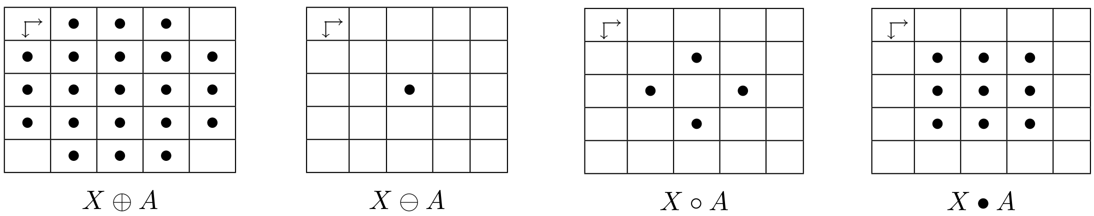
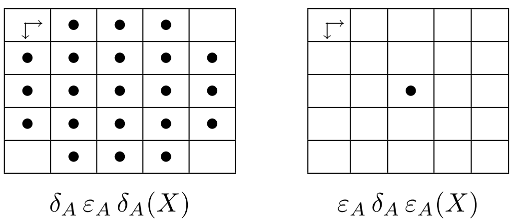


Figure 1: *Binary image X and structuring element A.*

a. (1 pt) In a similar way to Fig. 1 draw: the dilation $\delta_A(X) = X \oplus A$, the erosion $\varepsilon_A(X) = X \ominus A$, the opening $\gamma_A(X) = X \circ A$ and the closing $\phi_A(X) = X \bullet A$.



b. (0.5 pt) Furthermore, draw $\delta_A\, \varepsilon_A\, \delta_A(X)$ and $\varepsilon_A\, \delta_A\, \varepsilon_A(X)$.



c. (1 pt) Prove that for any X, A: $\delta_A\, \varepsilon_A\, \delta_A(X) = \delta_A(X)$.
Hint: Prove that $\delta_A\, \varepsilon_A\, \delta_A(X) \subseteq \delta_A(X)$ *and* dat $\delta_A\, \varepsilon_A\, \delta_A(X) \supseteq \delta_A(X)$.

Proof: By definition

$$\delta_A\, \varepsilon_A\, \delta_A(X) = \delta_A(\phi_A(X)) = \gamma_A(\delta_A(X)), , \tag{1}$$

furthermore,

$$X \subseteq \phi_A(X) \tag{2}$$

because a closing is extensive. Because dilation is an increasing operator, i.e. $X \subseteq Y \Rightarrow \delta_A(X) \subseteq \delta_A(Y)$, we have

$$\delta_A(X) \subseteq \delta_A(\phi_A(X)) = \delta_A\, \varepsilon_A\, \delta_A(X). \tag{3}$$

Conversely, because γ_A is anti-extensive

$$\delta_A\, \varepsilon_A\, \delta_A(X) = \gamma_A(\delta_A(X)) \subseteq \delta_A(X). \tag{4}$$

Equations (3) and (4) can only be true simultateously if

$$\delta_A\, \varepsilon_A\, \delta_A(X) = \delta_A(X) \tag{5}$$

Problem 2. (**2.5 pt**) Given a camera with unknown camera constant f, which images a parallellogram $ABCD$ via perspective projection on the plane $z = f$ (**ccc**-system). The sides AB and AD are at a known angle α, see Fig. 2. Furthermore, one corner of the parallellogram is known: $A = (0, 0, 3)$.

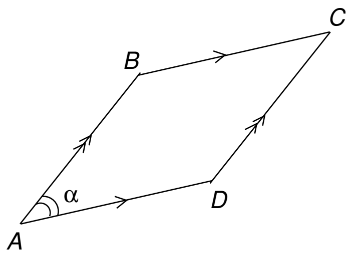


Figure 2: *Four line segments forming a parallellogram.*

The vanishing point of the parallel sides AB and DC is $(u_\infty, v_\infty) = (2, 1)$.
The vanishing point of the parallel sides AD and BC is $(u'_\infty, v'_\infty) = (-1, -2)$.

a. (1.5 pt) Compute the camera constant f as a function of α.

We have two direction vectors: $\vec{V} = (2, 1, f)^T$ and $\vec{V}' = (-1, -2, f)^T$, with angle α. The inner product is therefore

$$f^2 - 4 = \|\vec{V}\| \|\vec{V}'\| \cos\alpha = (f^2 + 5)\cos\alpha. \tag{6}$$

This means

$$(1 - \cos\alpha) f^2 = (4 + 5\cos\alpha) \tag{7}$$

which yields

$$f^2 = \frac{4 + 5\cos\alpha}{1 - \cos\alpha}, \tag{8}$$

and therefore

$$f = \sqrt{\frac{4 + 5\cos\alpha}{1 - \cos\alpha}}, \tag{9}$$

b. (1 pt) The equation of the plane V containing the parallellogram, has the form:

$$V: \quad a\,x + b\,y + c\,z + d = 0.$$

Determine the constants a, b, c, d, assuming $\alpha = \pi/2$.

This means $\cos\alpha = 0$, i.e., $f = 2$. The normal to the plane is defined by (a, b, c), which is obtained from the cross-product of $\vec{V} = (2, 1, 2)^T$ and $\vec{V}' = (-1, -2, 2)^T$:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} -1 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ -6 \\ -3 \end{pmatrix} \tag{10}$$

This can be simplified to $(2, -2, -1)^T$. For d we fill in the known location of A:

$$0a + 0b + 3c + d = 0 \quad \Rightarrow \quad d = 3. \tag{11}$$

Problem 3. (**2 pt**) Consider a cylinder with an axis parallel to the x-axis, with equation

$$z = d - \sqrt{r^2 - y^2}, \qquad -r \leq y \leq r$$

The cylinder has a Lambertian surface with constant albedo $\rho_S = 1$, and is illuminated from below by a light source at a very large distance, from a direction defined by the unit vector (a, b, c). The camera is on the negative z-as.

Show that the image intensity under orthografic projection is given by

$$E(x, y) = \frac{by - c\sqrt{r^2 - y^2}}{r}$$

Solution. Note that the gradient-space representation (p, q) of each point (x, y) in the image plane is given by

$$(p, q) = \left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) = \left(0, \frac{y}{\sqrt{r^2 - y^2}}\right), \qquad -r \leq y \leq r \tag{12}$$

The surface normals $\vec{n}$ are given by $(p, q, -1)$. The irradiance function $R(p, q)$ is given by

$$R(p, q) = \frac{\vec{n} \cdot \vec{s}}{\|\vec{n}\|\|\vec{s}\|} \tag{13}$$

with $\vec{s} = (a, b, c)^T$, and $\|\vec{s}\| = 1$. Inserting (12) we have

$$R(p, q) = \frac{\frac{by}{\sqrt{r^2-y^2}} - c}{\sqrt{\frac{y^2}{r^2-y^2} + 1}} = E(x, y). \tag{14}$$

Multiplying the numerator and denominator in the middle term of (14) with $\sqrt{r^2 - y^2}$ we have

$$R(p, q) = \frac{by - c\sqrt{r^2 - y^2}}{\sqrt{y^2 + r^2 - y^2}} = \frac{by - c\sqrt{r^2 - y^2}}{r} = E(x, y). \tag{15}$$

Problem 4. (**2 pt**) Methods to segment grey-value images into foreground and background can be divided into two categories (1) region-based and (2) edge-based.

a. (1 pt) Assign each of the following methods to the abovementioned classes: (i) thresholding; (ii) watershed algorithm (iii) snakes.

(i) Thresholding is clearly region-based, because it assigns pixels due to regional properties (intensity) rather than looking at edges.

(ii) Watersheds are also region-based, because regions are grown either from local extrema or from markers. The locations where two growth fronts meet are labeled as watershed lines, but these need not correspond to edges. (Note that it could be argued that watershed segmentation is edge based IF restricted to segmentation using watersheds of the *gradient image*, given good arguments, this choice could be accepted).

(iii) Snakes are clearly edge based, because snakes try to approximate the boundaries of objects by being attracted to edges in images.

b. (1 pt) For each method, give one or more advantages and one or more disadvantages.

(i) Advantages: Easy to implement, very fast. Disdavantages: Selection of threshold not always easily automated, does not deal well with texture or very noisy images

(ii) Advantages: Flexible, due to option of marker selection. Disadvantage: Prone to oversegmentation, also marker selection itself can cause problems: e.g. how to automate?

(iii) Advantages: can deal well with missing bits of edges, flexible because constraints (such as prior knowledge of the shape) can be included). Disadvantages: Very sensitive to initialization, many parameters to set.

Note: these are the just some of the pros and cons, an extensive list would be too long.