Image Processing and Computer Vision 2005/06
Example Solutions for Theoretical Assignments 6 (T6)

Problem 1.

We are looking for the cooccurrence matrix with displacement vector $d = (-1,1)^T$, i.e. we compare each pixel to its lower left neighbour.

$$
\begin{array}{ccccccccccc}
0 & & 2 & & 1 & & 0 & & 1 & & 0 \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & \\
0 & & 1 & & 3 & & 3 & & 2 & & 2 \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & \\
3 & & 2 & & 1 & & 2 & & 3 & & 0 \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & \\
3 & & 2 & & 0 & & 1 & & 0 & & 3 \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & \\
1 & & 1 & & 3 & & 2 & & 0 & & 3 \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & \\
2 & & 3 & & 1 & & 2 & & 0 & & 1
\end{array}
$$

which gives cooccurrence matrix

$$R = \frac{1}{25}\begin{pmatrix} 1 & 1 & 3 & 1 \\ 0 & 1 & 2 & 3 \\ 2 & 2 & 1 & 2 \\ 2 & 2 & 1 & 1 \end{pmatrix}$$

where $R_{i,j}$ is the relative frequency of pixels with grey value $i-1$ whose lower left neighbours have grey value $j-1$, i.e. "arrows starting at $i-1$ and ending at $j-1$". E.g. there are 3 pixels with grey value 1 and lower left neighbour 3, therefore $R_{2,4} = \frac{3}{25}$.

Problem 2.

The Mumford Shah cartoon model uses the variational ansatz

$$E(K) = \int_{\Omega - K} |u - f|^2 \, dx \, dy + \lambda l(K) \ ,$$

where $l(K)$ measures the length of K and $\lambda > 0$ is a scale parameter. We assume that $\Omega_1, \ldots, \Omega_n$ is a decomposition of Ω into disjoint regions with boundaries K. The segments are separated completely by K such that they can be treated independently for the minimisation of the energy. Thus the function u minimising the energy is constant in each segment and equal to the mean value of f in the segment:

$$u_m = \frac{1}{|\Omega_m|} \int_{\Omega_m} f \, dx \, dy \quad \text{for } m = 1, \ldots, n \ . \tag{1}$$

This can be seen by setting the derivative of the similarity term to zero. Let Ω_i, Ω_j denote two different segments. Merging the two segments Ω_i and Ω_j thus creates a new function v which is equal to u in all segments except in the new segment $\Omega_i \cup \Omega_j$ where it has the value

$$\tilde{u} := \frac{1}{|\Omega_i| + |\Omega_j|} \int_{\Omega_i \cup \Omega_j} f \; dx \; dy = \frac{|\Omega_i| u_i + |\Omega_j| u_j}{|\Omega_i| + |\Omega_j|} \quad . \tag{2}$$

We want to show that merging the two regions Ω_i and Ω_j results in the following change of energy:

$$E(K - \partial(\Omega_i, \Omega_j)) - E(K) = \frac{|\Omega_i| \cdot |\Omega_j|}{|\Omega_i| + |\Omega_j|} |u_i - u_j|^2 - \lambda l(\partial(\Omega_i, \Omega_j)) \;\; ,$$

where $\partial(\Omega_i, \Omega_j)$ denotes the common boundary between Ω_i and Ω_j.

We write down the difference between the energy values before and after merging:

$$\begin{aligned} &E(K - \partial(\Omega_i, \Omega_j)) - E(K) = \\ &\qquad \int_{K - \partial(\Omega_i, \Omega_j)} |v - f|^2 \; dx \; dy + \lambda l(K - \partial(\Omega_i, \Omega_j)) - \int_K |u - f|^2 \; dx \; dy - \lambda l(K). \end{aligned}$$

The length of the boundary is reduced by the length of the common boundary between Ω_i and Ω_j during the merging step. The difference between the length terms thus can be expressed as

$$l(K - \partial(\Omega_i, \Omega_j)) - l(K) = -l(\partial(\Omega_i, \Omega_j)) \;\; . \tag{3}$$

Now we take a closer look at the similarity term. With $M = \{1, \dots, n\}$ we denote the index set of our regions. Considering the fact that u and v are constant in each region, we split up the integrals in sums over the several regions

$$\int_{\Omega - K} |u - f|^2 \quad = \quad \sum_{m \in M} \int_{\Omega_m} |u_m - f|^2 \; dx \; dy$$

In the second case we remember that v equals u in all regions except $\Omega_i \cup \Omega_j$:

$$\begin{aligned} \int_{\Omega - (K - \partial(\Omega_i, \Omega_j))} |v - f|^2 \; dx \; dy \quad &= \quad \sum_{m \in M - \{i,j\}} \int_{\Omega_m} |u_m - f|^2 \; dx \; dy \\ &\qquad + \int_{\Omega_i \cup \Omega_j} |\tilde{u} - f|^2 \; dx \; dy \;\; . \end{aligned}$$

The difference between the two similarity terms then can be written as

$$
\begin{aligned}
&\int_{\Omega-(K-\partial(\Omega_i,\Omega_j))} |v-f|^2 \; dx \; dy - \int_{\Omega-K} |u-f|^2 \; dx \; dy \\
= \;& \int_{\Omega_i\cup\Omega_j} |\tilde{u}-f|^2 \; dx \; dy - \int_{\Omega_i} |u_i-f|^2 \; dx \; dy - \int_{\Omega_j} |u_j-f|^2 \; dx \; dy \\
= \;& \int_{\Omega_i\cup\Omega_j} \left(\tilde{u}^2 - 2\tilde{u}f + f^2\right) \; dx \; dy - \int_{\Omega_i} \left(u_i^2 - 2u_i f + f^2\right) \; dx \; dy \\
& - \int_{\Omega_j} \left(u_j^2 - 2u_j f + f^2\right) \; dx \; dy \\
= \;& \int_{\Omega_i\cup\Omega_j} \left(\tilde{u}^2 - 2\tilde{u}f\right) \; dx \; dy - \int_{\Omega_i} \left(u_i^2 - 2u_i f\right) \; dx \; dy - \int_{\Omega_j} \left(u_j^2 - 2u_j f\right) \; dx \; dy \\
= \;& \left(|\Omega_i| + |\Omega_j|\right)\tilde{u}^2 - 2\tilde{u}\int_{\Omega_i\cup\Omega_j} f \; dx \; dy - |\Omega_i| u_i^2 + 2u_i \int_{\Omega_i} f \; dx \; dy \\
& - |\Omega_j| u_j^2 + 2u_j \int_{\Omega_j} f \; dx \; dy
\end{aligned}
$$

We use the definition of $\tilde{u}$, u_i and u_j as mean values in (1) and (2) to transform this expression to

$$
\left(|\Omega_i| + |\Omega_j|\right)\tilde{u}^2 - 2\left(|\Omega_i| + |\Omega_j|\right)\tilde{u}^2 - |\Omega_i| u_i^2 + 2|\Omega_i| u_i^2 - |\Omega_j| u_j^2 + 2|\Omega_j| u_j^2
$$

which can be immediately simplified to

$$
|\Omega_i| u_i^2 + |\Omega_j| u_j^2 - \left(|\Omega_i| + |\Omega_j|\right)\tilde{u}^2 \; . \tag{4}
$$

Using the relation between $\tilde{u}$, u_i and u_j given in (2), we can finally write this with a common denominator as

$$
\begin{aligned}
&|\Omega_i| u_i^2 + |\Omega_j| u_j^2 - \left(|\Omega_i| + |\Omega_j|\right)\tilde{u}^2 \\
= \;& \frac{1}{|\Omega_i| + |\Omega_j|} \Big(\left(|\Omega_i| + |\Omega_j|\right)\left(u_i^2 |\Omega_i| + u_j^2 |\Omega_j|\right) - \left(|\Omega_i| u_i + |\Omega_j| u_j\right)^2 \Big) \\
= \;& \frac{1}{|\Omega_i| + |\Omega_j|} \left(|\Omega_i| \cdot |\Omega_j| \left(u_i^2 - 2u_i u_j + u_j^2\right)\right) \\
= \;& \frac{|\Omega_i| \cdot |\Omega_j|}{|\Omega_i| + |\Omega_j|} |u_i - u_j|^2
\end{aligned} \tag{5}
$$

Together the equations (5) and (3) show the formula we wanted to prove:

$$
E(K - \partial(\Omega_i,\Omega_j)) - E(K) = \frac{|\Omega_i| \cdot |\Omega_j|}{|\Omega_i| + |\Omega_j|} |u_i - u_j|^2 - \lambda l(\partial(\Omega_i,\Omega_j)) \; .
$$

Problem 3.

(a) The spatiotemporal structure tensor looks as follows:

$$J_p = (\nabla_3 f \nabla_3 f^T) = \begin{pmatrix} K_\rho * f_x^2 & K_\rho * (f_x f_y) & K_\rho * (f_x f_z) \\ K_\rho * (f_x f_y) & K_\rho * f_y^2 & K_\rho * (f_y f_z) \\ K_\rho * (f_x f_z) & K_\rho * (f_y f_z) & K_\rho * f_z^2 \end{pmatrix}$$

We can see that the matrix is symmetric. Moreover, one can show that it is positive semidefinite by construction (all eigenvalues ≥ 0). Vice versa, every symmetric positive semidefinite 3×3 matrix is possible spatiotemporal structure tensor. In the following it may also help to recall that positive semidefiniteness also implies that that all main diagonal elements are positive or zero.

We immediately see that (ii) is not symmetric. Moreover, we can see that one entry of the main diagonal of (iv) is negative. Consequently, it least one of the eigenvalues is negative. (This can be verified by using the theorem of Gershgorin). So (ii) and (iv) do not represent spatiotemporal structure tensors.

Now we must estimate the eigenvalues of the matrices (i) and (iii) and check if they are ≥ 0.
The fastest way is to use the theorem of Gershgorin

(i)

$$J_\rho = \begin{pmatrix} 4 & -1 & 1 \\ -1 & 5 & -3 \\ 1 & -3 & 6 \end{pmatrix}$$

circles of Gerschgorin:

$$\lambda_1 = \{z \in \mathbb{R}| \ \ |z-4| \leq 1+1 = 2 \Rightarrow [2,6]$$

$$\lambda_2 = \{z \in \mathbb{R}| \ \ |z-5| \leq 1+3 = 4 \Rightarrow [1,9]$$

$$\lambda_3 = \{z \in \mathbb{R}| \ \ |z-6| \leq 1+3 = 4 \Rightarrow [2,10]$$

Since this shows that all of the three eigenvalues are in the interval $[1, 10]$, they are obviously ≥ 0. This means that (ii) fulfills all requirements to be spatiotemporal structure tensor. In fact, it corresponds to Bigün's second case, where the smallest eigenvalue is significantly larger than zero: Thus either the assumption of a locally constant flow or the assumption of a const grey value over time is violated.

(iii) In this matrix it is easy to see that all eigenvalues are zero.
The matrix is symmetric and positive semidefinite $\Rightarrow$ spatiotemporal structure

tensor. This structure tensor corresponds to the first case: Nothing can be said about the local optic flow.

(b) The usual meaning of the eigenvalues:

- no large eigenvalue $\Rightarrow$ no information
- one large eigenvalue $\Rightarrow$ edge
- two large eigenvalues $\Rightarrow$ corner

If there are three large eigenvalues, we have a spatiotemporal corner (a corner in the image and a leap in time). This means that the structural information in temporal direction does not allow the estimation of a locally constant flow. This may have two reasons: Either the flow is really not locally constant or the grey value constancy assumption is not fulfilled.

Problem 4.

(a) For a 1-D signal $f(x, t)$, the grey value constancy assumption can be formulated as

$$f(x + u, t + 1) - f(x, t) = 0.$$

If u is small and f varies smoothly, we can perform a linearisation by means of first order Taylor expansion:

$$f(x + u, t + 1) \approx f(x, t) + f_x(x, t)u + f_t(x, t).$$

Thus, we obtain the following optic flow constraint:

$$f_x(x, t)u + f_t(x, t) = 0.$$

Since one equation is sufficient to determine one unknown uniquely, no aperture problem exists in the 1-D case. In fact, we can compute u via

$$u = -\frac{f_t}{f_x}.$$

However, there are cases where u cannot be computed: At locations, where $f_x = 0$.

(b) In the case of a 3-D signal $f(x, y, z, t)$, the grey value constancy assumption is given by

$$f(x + u, y + v, z + w, t + 1) - f(x, y, z, t) = 0.$$

As in the 1-D case, we can perform a linearisation of this assumptions by means of first order Taylor expansion here. To this end, we assume again that u is small and f only varies smoothly. This give us

$$\begin{aligned} f(x+u, y+v, z+w, t+1) \approx f(x,y,z,t) &+ f_x(x,y,z,t)u \\ &+ f_y(x,y,z,t)v \\ &+ f_z(x,y,z,t)w \\ &+ f_t(x,y,z,t). \end{aligned}$$

The resulting optic flow constraint then reads

$$f_x(x,y,z,t)u + f_y(x,y,z,t)v + f_z(x,y,z,t)w + f_t(x,y,z,t) = 0.$$

This time the aperture problem is present: We have three unknowns but only one equation. Taking a closer look at the previous constraint, one can see that all possible solutions at a point (x,y,z) have to lie on a plane through $f_t(x,y,z,t)$ with normal vector $(f_x, f_y, f_t)^\top$.

Let us now embed this optic flow constraints in a Bigün–like approach. Then, we obtain

$$E(\mathbf{a}) = \int\limits_{B_\rho(x_0,y_0,z_0,t_0)} \Big(f_x a_1 + f_y a_2 + f_z a_3 + f_t a_4\Big)^2 \, dx\, dy\, dz\, dt,$$

where the last component of the minimising vector has to be normalised after the computation to 1. This yields the 3-D optic flow

$$u = \frac{a_1}{a_4}, \qquad v = \frac{a_2}{a_4}, \qquad w = \frac{a_3}{a_4}.$$

As in the original method of Bigün *et al.* the approach can be formulated as a quadratic form. In the 3-D case this form is given by

$$E(\mathbf{a}) = \mathbf{a}^\top \hat{J}_\rho \, \mathbf{a},$$

with the 3-D spatiotemporal structure tensor

$$\hat{J}_\rho = \begin{pmatrix} K_\rho * (f_x^2) & K_\rho * (f_x f_y) & K_\rho * (f_x f_z) & K_\rho * (f_x f_t) \\ K_\rho * (f_x f_y) & K_\rho * (f_y^2) & K_\rho * (f_y f_z) & K_\rho * (f_y f_t) \\ K_\rho * (f_x f_z) & K_\rho * (f_y f_z) & K_\rho * (f_z^2) & K_\rho * (f_z f_t) \\ K_\rho * (f_x f_t) & K_\rho * (f_y f_t) & K_\rho * (f_z f_t) & K_\rho * (f_t^2) \end{pmatrix}.$$

Let $\mu_1 \geq \mu_2 \geq \mu_3 \geq \mu_4 \geq 0$ be the eigenvalues of $\hat{J}_\rho$. Then, we have five different cases that are processed in the following order

- **No eigenvalue is large** ($\mathrm{tr}\hat{J}_\rho = \hat{j}_{11} + \hat{j}_{22} + \hat{j}_{33} + \hat{j}_{44} \leq \tau_1$): There is a *homogeneous spatiotemporal volume* and thus not sufficient local information to compute the optic flow.
- **All eigenvalues are large** ($\mu_4 \geq \tau_2$): Either the grey value constancy assumption or the assumption of a constant flow is violated.
- **One eigenvalue is large** ($\mu_2 \leq \tau_3$): There is an *edge*, the solution lies on a plane.
- **Two eigenvalues are large** ($\mu_3 \leq \tau_4$): There is a *2-D corner*, the solution lies on a line.
- **Three eigenvalues are large** (the remaining case): There is a *3-D corner*. Then, we have to compute the eigenvector to the smallest eigenvalue of $\hat{J}_\rho$ and normalise it so that its last component becomes 1.