Image Processing and Computer Vision 2007/08
Example Solutions for Theoretical Assignments 6 (T6)

Problem 1 (Mumford-Shah Cartoon Model)

The Mumford Shah cartoon model uses the variational ansatz

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

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:

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

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

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

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:

$$E(K - \partial(\Omega_i, \Omega_j)) - E(K) =$$
$$\int\limits_{K - \partial(\Omega_i, \Omega_j)} |v - f|^2 \; dx \; dy + \lambda l(K - \partial(\Omega_i, \Omega_j)) - \int\limits_{K} |u - f|^2 \; dx \; dy - \lambda l(K).$$

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 $M = \{1, \ldots, n\}$ 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

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

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

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

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

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

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

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

which can be immediately simplified to

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

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

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

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 2 (The Method of Bigün *et al.*)

(a) The spatiotemporal stucture tensor looks as follows:

$$J_\rho := K_\rho * \left(\nabla_3 f \nabla_3 f^\top\right) = \begin{pmatrix} K_\rho * (f_x^2) & K_\rho * (f_x f_y) & K_\rho * (f_x f_z) \\ K_\rho * (f_x f_y) & K_\rho * (f_y^2) & K_\rho * (f_y f_z) \\ K_\rho * (f_x f_z) & K_\rho * (f_y f_z) & K_\rho * (f_z^2) \end{pmatrix}.$$

We can see that the matrix is symmetric. Moreover, one can show that it is positive semidefinite by construction (all eigenvalues ≥ 0). Vice versa, only symmetric positive semidefinite 3×3 matrices can be a spatiotemporal structure tensor.

We immediately see that (i) is not symmetric and thus cannot represent a spatiotemporal structure tensor.

We now must estimate the eigenvalues of the matrices (ii)-(iv), and check if the are ≥ 0. The fastest way to do it is using the theorem of Gershgorin. The circles of Gershgorin of the matrix (ii)

$$J_\rho = \begin{pmatrix} 3 & -1 & 1 \\ -1 & 6 & -3 \\ 1 & -3 & 7 \end{pmatrix}$$

are:

$$\begin{aligned} \lambda_1 &= \{z \in \mathbb{R} : |z-3| \leq 1+1 = 2 \Rightarrow [1,5]\}, \\ \lambda_2 &= \{z \in \mathbb{R} : |z-6| \leq 1+3 = 4 \Rightarrow [2,10]\}, \\ \lambda_3 &= \{z \in \mathbb{R} : |z-7| \leq 1+3 = 4 \Rightarrow [3,11]\}. \end{aligned}$$

This indicates that all eigenvalues are in the interval $[1, 11]$. Then, the matrix (ii) fulfills all requirements to be a spatiotemporal structure tensor. In fact, it corresponds to Bigün's second case, where the smallest eigenvalue is significantly larger than zero: Either the assumption of a locally constant flow or the assumption of a constant grey value over time is violated.

Let us now take a look at matrix (iii) given by

$$J_\rho = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -3 & -1 \\ 2 & -1 & 5 \end{pmatrix}$$

Here, the circles of Gershgorin read:

$$\begin{aligned}\lambda_1 &= \{z \in \mathbb{R} : |z-1| \leq 0+2 = 2 \Rightarrow [-1, 3]\}\,,\\ \lambda_2 &= \{z \in \mathbb{R} : |z+3| \leq 0+1 = 1 \Rightarrow [-4, -2]\}\,,\\ \lambda_3 &= \{z \in \mathbb{R} : |z-5| \leq 2+1 = 3 \Rightarrow [2, 8]\}\,.\end{aligned}$$

This shows that at least one eigenvalue is negative (λ_2), while another one is positive (λ_3). Thus the matrix is indefinite and cannot represent a spatiotemporal structure tensor.

In the case of the matrix (iv), 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 Bigün's first case, where nothing can be said about the local optical flow.

(b) The usual meaning of the eigenvalues:

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

If we have 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 3 (Variational Optic Flow Methods)

(a) First of all, please note that in this 3-D setting, the flow field we are going to estimate will be three dimensional as well, and for a given time t we will approximate the three functions $u, v, w : \mathbb{R}^3 \to \mathbb{R}$ as flow components. Assuming that corresponding pixels have the same grey values in successive frames, this constancy assumption reads in its original nonlinear form

$$f(x, y, z, t) - f(x + u, y + v, z + w, t + 1) = 0 \; . \tag{6}$$

If f varies slowly and if u, v and w are sufficiently small, we can perform a Taylor expansion around the point (x, y, z, t) and linearise the nonlinear term

$$\begin{aligned} 0 &= f(x + u, y + v, z + w, t + 1) - f(x, y, z, t) \\ &\approx f_x(x, y, z, t)\, u + f_y(x, y, z, t)\, v + f_z(x, y, z, t)\, w + f_t(x, y, z, t) \; . \end{aligned}$$

Omitting the arguments we obtain the 3-D Optic Flow Constraint for the grey value constancy assumption:

$$f_x\, u + f_y\, v + f_z\, w + f_t = 0 \; . \tag{7}$$

Concerning the aperture problem, we see that equation (7) contains three unknowns (u, v and w). Thus, in this 3-D scenario we have even one more unknown variable than in the 2-D case and just one equation. So, without any additional assumptions only the normal flow can be computed here:

$$\omega_n = -\frac{f_t}{|\nabla_3 f|} \frac{\nabla_3 f}{|\nabla_3 f|} \; .$$

In this context, $\nabla_3 := (\partial_x, \partial_y, \partial_z)^\top$ denotes the 3-D gradient operator.

(b) Squaring the OFC from equation (7) we obtain the following Horn-and-Schunck-like energy functional for 3-D:

$$\begin{aligned} E(u, v, w) \;\; = \;\; & \int_\Omega \Big((f_x\, u + f_y\, v + f_z\, w + f_t)^2 \\ & + \alpha \left(|\nabla_3\, u|^2 + |\nabla_3\, v|^2 + |\nabla_3\, w|^2 \right) \Big)\, dx\, dy\, dz \; . \end{aligned} \tag{8}$$

Note that the functional now depends on three functions, since we are searching for all three flow components.

(c) Following the usual procedure, we express the energy functional from equation (8) in terms of the *Lagrangian* $F(x, y, u, v, w, \nabla u, \nabla v, \nabla w)$

$$F(x, y, z, u, v, w, \nabla u, \nabla v, \nabla w) = \\ (f_x\, u + f_y\, v + f_z\, w + f_t)^2 + \alpha\left(|\nabla_3\, u|^2 + |\nabla_3\, v|^2 + |\nabla_3\, w|^2\right) \; . \quad (9)$$

This leads to the following energy functional for which we are searching a minimiser

$$E(u, v, w) = \int_\Omega F(x, y, z, u, v, w, \nabla u, \nabla v, \nabla w)\, dx\, dy\, dz \; . \quad (10)$$

In 3-D, the three Euler-Lagrange equations for u, v and w read

$$\begin{aligned} F_u - \partial_x\, F_{u_x} - \partial_y\, F_{u_y} - \partial_z\, F_{u_z} &= 0\;, \\ F_v - \partial_x\, F_{v_x} - \partial_y\, F_{v_y} - \partial_z\, F_{v_z} &= 0\;, \\ F_w - \partial_x\, F_{w_x} - \partial_y\, F_{w_y} - \partial_z\, F_{w_z} &= 0\;. \end{aligned} \quad (11)$$

Furthermore, also the boundary conditions

$$\mathbf{n}^\top \begin{pmatrix} F_{u_x} \\ F_{u_y} \\ F_{u_z} \end{pmatrix} = 0\;, \qquad \mathbf{n}^\top \begin{pmatrix} F_{v_x} \\ F_{v_y} \\ F_{v_z} \end{pmatrix} = 0\;, \qquad \mathbf{n}^\top \begin{pmatrix} F_{w_x} \\ F_{w_y} \\ F_{w_z} \end{pmatrix} = 0\;. \quad (12)$$

have to be satisfied on the boundary of the integration domain $\partial\Omega$. Calculating the partial derivatives occurring in equations (12) and (12) we have

$$\begin{aligned} \partial_x\left(F_{u_x}\right) &= \partial_x\left(2\,\alpha\, u_x\right) = 2\,\alpha\, u_{xx}\;, \\ \partial_y\left(F_{u_y}\right) &= \partial_y\left(2\,\alpha\, u_y\right) = 2\,\alpha\, u_{yy}\;, \\ \partial_z\left(F_{u_z}\right) &= \partial_z\left(2\,\alpha\, u_z\right) = 2\,\alpha\, u_{zz}\;. \end{aligned} \quad (13)$$

Analogously we obtain

$$\begin{aligned} \partial_x\left(F_{v_x}\right) &= 2\,\alpha\, u_{xx} & \partial_x\left(F_{w_x}\right) &= 2\,\alpha\, w_{xx}\;, \\ \partial_y\left(F_{v_y}\right) &= 2\,\alpha\, v_{yy} & \partial_y\left(F_{w_y}\right) &= 2\,\alpha\, w_{yy}\;, \\ \partial_z\left(F_{v_z}\right) &= 2\,\alpha\, v_{zz} & \partial_z\left(F_{w_z}\right) &= 2\,\alpha\, w_{zz}\;, \end{aligned} \quad (14)$$

and finally we can compute

$$\begin{aligned} F_u &= 2\,\left(f_x\, u + f_y\, v + f_z\, w + f_t\right)\, f_x\;, \\ F_v &= 2\,\left(f_x\, u + f_y\, v + f_z\, w + f_t\right)\, f_y\;, \\ F_w &= 2\,\left(f_x\, u + f_y\, v + f_z\, w + f_t\right)\, f_z\;. \end{aligned} \quad (15)$$

Plugging all parts together, we obtain the following Euler-Lagrange equations:

$$0 = \alpha \underbrace{(u_{xx} + u_{yy} + u_{zz})}_{\triangle u} - (f_x\, u + f_y\, v + f_z\, w + f_t)\; f_x \;, \tag{16}$$

$$0 = \alpha \underbrace{(v_{xx} + v_{yy} + v_{zz})}_{\triangle v} - (f_x\, u + f_y\, v + f_z\, w + f_t)\; f_y \;, \tag{17}$$

$$0 = \alpha \underbrace{(w_{xx} + w_{yy} + w_{zz})}_{\triangle w} - (f_x\, u + f_y\, v + f_z\, w + f_t)\; f_z \;. \tag{18}$$

The associated boundary conditions from equation (12) simplify to

$$0 = \mathbf{n}^\top \begin{pmatrix} F_{u_x} \\ F_{u_y} \\ F_{u_z} \end{pmatrix} = \mathbf{n}^\top \begin{pmatrix} 2\,\alpha\, u_x \\ 2\,\alpha\, u_y \\ 2\,\alpha\, u_z \end{pmatrix} = \mathbf{n}^\top \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} = \mathbf{n}^\top \nabla u \;. \tag{19}$$

Analogously we obtain the equations

$$\mathbf{n}^\top \nabla u = 0 \;, \qquad \mathbf{n}^\top \nabla v = 0 \;, \qquad \mathbf{n}^\top \nabla w = 0 \;. \tag{20}$$

Evidently, the Euler-Lagrange equations (16)-(19) are linear, since they depend linearly on the unknowns u, v and w. In this context one should note that differentiation is a linear operation. This explains the linearity of Laplace operators Δu, Δv and Δw.

The fact that we obtain linear Euler-Lagrange equations, however, is not surprising: Since we used quadratic penalisation in the data term as well as in the smoothness term, the differentiation of the functional via the Euler-Lagrange framework must yields only linear dependencies. As in the case of ordinary polynomials, the differentiation reduces the degree of the expression by one.