# Image Processing and Computer Vision

## Assignment T6

Jan Hendrik Dithmar
2031259

Group T1: Tue, 14-16 (Sebastian Zimmer)

## 6.1 Mumford-Shah Cartoon Model

*Unfortunately no solution. :-(*

## 6.2 The Method of Bigün et al.

**a**. Based on
`http://mathworld.wolfram.com/GershgorinCircleTheorem.html`
I compute the Gershgorin circles.

(i)

$$\begin{aligned}\lambda_1 &= \{z \in \mathbb{R} \mid |z-5| \leq 5\} \Rightarrow [0,10]\\ \lambda_2 &= \{z \in \mathbb{R} \mid |z-1| \leq 3\} \Rightarrow [2,4]\\ \lambda_3 &= \{z \in \mathbb{R} \mid |z-1| \leq 2\} \Rightarrow [1,3]\end{aligned}$$

After a short calculation you can easily see that the eigenvalues are $-1, 1, 7$. This leads to $rank(J_\rho) = 3$.
Since $J_\rho$ is not symmetric where the spatiotemporal structure tensor is, the criterion is not fulfilled in this case.

(ii)

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

Since all eigenvalues lie in the interval $[1, 11]$, i.e. are $\geq 0$, we have $rank(J_\rho) = 3$.
Due to the fact that the matrix is symmetric and all eigenvalues are $\geq 0$, the requirements for a spatiotemporal structure tensor are fulfilled.

(iii)

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

Since one of the eigenvalues is negative (because of one negative element in the diagonal), this matrix cannot be a spatiotemporal structure tensor, because the matrix is not positive semidefinite.
All eigenvalues are $\neq 0$, so $rank(J_\rho) = 3$.

(iv) As one can easily see, $rank(J_\rho) = 0$. Since all eigenvalues are zero, this is not a semitemporal structure tensor.

**b**. Having three large eigenvalues means that either the grey value assumption is not fulfilled or the flow is not locally constant.
Here is the explanation for that:
If we have three large eigenvalues, we have a "jump" also in the temporal direction that does not allow the estimation of a locally constant flow.

## 6.3 Variational Optic Flow Methods

**a**.

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

Since we have one equation with three unknowns and the possible solutions in a point $(x, y, z)$ have to be orthogonal to $\nabla f = (f_x, f_y, f_z)^\top$, we have the aperture problem.

**b**.

$$E(u, v, w) = \int_\Omega \left((f_x u + f_y v + f_z w)^2 + \alpha \left(|\nabla u|^2 + |\nabla v|^2 + |\nabla w|^2\right)\right) dx\ dy\ dz$$

**c**. To compute the Euler-Lagrange equations, I only have to look at the integrad of

$$E(u, v, w) = \int_\Omega F(x, y, z, u, v, w, u_x, u_y, u_z, v_x, v_y, v_z, w_x, w_y, w_z) dx\ dy\ dz$$

which is

$$F = (f_x u + f_y v + f_z w)^2 + \alpha(u_x^2 + u_y^2 + u_z^2 + v_x^2 + v_y^2 + v_z^2 + w_x^2 + w_y^2 + w_z^2).$$

This leads to the partial derivatives

$$\begin{aligned}
F_u &= 2f_x(f_xu + f_yv + f_zw) \\
F_v &= 2f_y(f_xu + f_yv + f_zw) \\
F_w &= 2f_z(f_xu + f_yv + f_zw) \\
F_{u_x} &= 2\alpha u_x \\
F_{u_y} &= 2\alpha u_y \\
F_{u_z} &= 2\alpha u_z \\
F_{v_x} &= 2\alpha v_x \\
F_{v_y} &= 2\alpha v_y \\
F_{v_z} &= 2\alpha v_z \\
F_{w_x} &= 2\alpha w_x \\
F_{w_y} &= 2\alpha w_y \\
F_{w_z} &= 2\alpha w_z
\end{aligned}$$

This yields the Euler-Lagrange equations

$$\begin{aligned}
\Delta u - \frac{1}{\alpha} f_x(f_xu + f_yv + f_zw + f_t) &= 0 \\
\Delta v - \frac{1}{\alpha} f_y(f_xu + f_yv + f_zw + f_t) &= 0 \\
\Delta w - \frac{1}{\alpha} f_z(f_xu + f_yv + f_zw + f_t) &= 0
\end{aligned}$$

with the boundary conditions

$$\begin{aligned}
\mathbf{n}^\top \nabla u &= 0 \\
\mathbf{n}^\top \nabla v &= 0 \\
\mathbf{n}^\top \nabla w &= 0
\end{aligned}$$

where $\mathbf{n}$ is the unit normal vector and all equations have been divided by $2\alpha$.
Let's approximate the Laplacian similar to the lecture

$$\Delta u|_i \approx \sum_{j \in \mathcal{N}(i)} (u_j - u_i)$$

where $\mathcal{N}(i)$ are the neighbors of pixel $i$.

This yields the equations

$$
\begin{aligned}
0 &= \sum_{j \in \mathcal{N}(i)} (u_j - u_i) - \frac{1}{\alpha} f_{xi}(f_{xi}u_i + f_{yi}v_i + f_{zi}w_i + f_{ti}) \\
0 &= \sum_{j \in \mathcal{N}(i)} (v_j - v_i) - \frac{1}{\alpha} f_{yi}(f_{xi}u_i + f_{yi}v_i + f_{zi}w_i + f_{ti}) \\
0 &= \sum_{j \in \mathcal{N}(i)} (w_j - w_i) - \frac{1}{\alpha} f_{zi}(f_{xi}u_i + f_{yi}v_i + f_{zi}w_i + f_{ti})
\end{aligned}
$$

for all pixel ($i = 1, ..., N$).
$\Rightarrow$ This system of equations is linear.