Optical Flow Computation with Locally Quadratic Assumption

Tomoya Kato[1], Hayato Itoh[1], and Atsushi Imiya[2(✉)]

[1] School of Advanced Integration Science, Chiba University, Chiba, Japan
[2] Institute of Management and Information Technologies,
Chiba University, Yayoi-cho 1-33, Inage-ku, Chiba 263-8522, Japan
imiya@faculty.chiba-u.jp

Abstract. The purpose of this paper is twofold. First, we develop a quadratic tracker which computes a locally quadratic optical flow field by solving a model-fitting problem for each point in its local neighbourhood. This local method allows us to select a region of interest for the optical flow computation. Secondly, we propose a method to compute the transportation of a motion field in long-time image sequences using the Wasserstein distance for cyclic distributions. This measure evaluates the motion coherency in an image sequence and detects collapses of smoothness of the motion vector field in an image sequence.

1 Introduction

In this paper, we develop a method to compute a locally quadratic optical flow field. Furthermore, we propose a method to evaluate the global smoothness and continuity of motion fields and detect collapses of smoothness of the motion fields in long-time image sequences using the Wasserstein distance for cyclic distributions [2,7,10].

In ref. [13], using the local stationarity of visual motion, a linear method for motion tracking was introduced. The Lucas-Kanade (LK) method was proposed as an image-matching and -registration method assuming that the deformation field between images is locally constant. Subsequently, the method has been widely used for fast optical flow computation [11].

It is possible to extend the local assumption on the optical flow field to higher-order constraints on the motion field. In this paper, we assume that optical flow fields are locally quadratic. This local property of the optical flow field allows us to decompose an optical flow computation scheme to a collection of systems of linear equations defined in the neighbourhood of each point. The size of the system-matrix of each system of linear equations is 14×14, where 14 is the number of parameters used to describe a locally quadratic vector field on a plane. Furthermore, this decomposition property also allows us to select regions of interest for the optical flow computation and to construct a parallel method which computes the optical flow vectors of all points simultaneously.

For the computation of the three-dimensional scene flow from a stereo image sequence [4,5,6], we are required to solve four image-matching problems and

G. Azzopardi and N. Petkov (Eds.): CAIP 2015, Part I, LNCS 9256, pp. 223–234, 2015.
DOI: 10.1007/978-3-319-23192-1_19

their deformation fields. Two of them are optical flow computations for left and right image sequences. The other two of them are stereo matching for two successive stereo pairs. The displacements between stereo pairs are at most locally affine transformations caused by perspective projections based on the camera geometry. The displacement between a pair of successive images in left and right sequences, however, involves higher-order transformations caused by camera motion if a pair of cameras is mounted on a mobile vehicle. Therefore, since we are required to adopt different-order constraints on the optical flow computation and stereo matching for a stereo pair sequence, we develop a method with locally higher-order constraints for the fast computation of the optical flow field.

The Wasserstein distance defines a metric among probability measures [10]. In computer vision and pattern recognition, the 1-Wasserstein [2] distance is known as the earth movers' distance (EMD). We deal with the distribution of optical flow vectors as directional statistics [3]. Then, using the Wasserstein distance for cyclic distributions [7], we evaluate the temporal total transportation between a pair of successive optical flow fields. If the motion in an image sequence has a constant speed, this transportation measure between a pair of successive images is zero. Therefore, we use this temporal transportation of the flow fields as a measure to evaluate motion smoothness and continuity. Ustundag-Unel [15] and Chaudhry et. al. [16] analysed human motion using the orientation histogram of the optical flow field. There histogram is based on the histogram of oriented gradient (HoG) method [14] although we deal with the directional statistics to compute the temporal transportation of the optical flow fields.

Although classical optical flow computation methods [1,12] are based on least-squares optimisation, the total variation (TV) of the solution as a prior [20,17] and the L_1-constraint allow us to deal with the sparsity of images. The TV-L_1 minimisation for optical flow computation is solved by the primal-dual method [19]. There are a number of numerical schemes for (TV-L_1)-based image analysis [21,22]. These methods minimise a criterion defined over the whole image [19].

On the other hand, by dividing the region of interest into windowed areas and assuming that the optical flow field is locally quadratic in each region, our method solves a large system of diagonal linear equations. Furthermore, the local method allows us to select a region of interest for the optical flow computation, because the method computes the optical flow for each point in its local neighbourhood with a local constraint on the optical flow field [18].

2 Local Optical Flow Computation

For $f(x, y, t)$, the optical flow vector [12] $\boldsymbol{u} = \dot{\boldsymbol{x}} = (\dot{x}, \dot{y})^\top$, where $\dot{x} = u = u(x, y)$ and $\dot{y} = v = v(x, y)$, of each point $\boldsymbol{x} = (x, y)^\top$ is the solution of the singular equation

$$f_x u + f_y v + f_t = \nabla f^\top \boldsymbol{u} + \partial_t f = \boldsymbol{J}^\top \boldsymbol{u} + f_t = 0. \tag{1}$$

Assuming $\boldsymbol{u}$ to be constant in the neighbourhood $\Omega(\boldsymbol{x})$ of point $\boldsymbol{x}$ [12], the optical flow vector is the minimiser of the criterion

$$E_0 = \frac{1}{2|\Omega(\boldsymbol{x})|}\int_{\Omega(\boldsymbol{x})} |\boldsymbol{J}^\top \boldsymbol{u} + \boldsymbol{f}_t|^2 d\boldsymbol{x} = \frac{1}{2}\boldsymbol{u}\boldsymbol{G}\boldsymbol{u} + \boldsymbol{a}^\top \boldsymbol{u} + \frac{1}{2}c, \tag{2}$$

where $\boldsymbol{J} = (f_x, f_y)^\top$, for

$$\boldsymbol{G} = \frac{1}{|\Omega(\boldsymbol{x})|}\int_{\Omega(\boldsymbol{x})} \boldsymbol{J}^\top \boldsymbol{J} d\boldsymbol{x}, \tag{3}$$

$$\boldsymbol{a} = \frac{1}{|\Omega(\boldsymbol{x})|}\int_{\Omega(\boldsymbol{x})} f_t \nabla f d\boldsymbol{x},\ c = \frac{1}{|\Omega(\boldsymbol{x})|}\int_{\Omega(\boldsymbol{x})} |f_t|^2 d\boldsymbol{x}. \tag{4}$$

If the displacement is locally affine such that $\boldsymbol{u} = \boldsymbol{D}\boldsymbol{x} + \boldsymbol{d}$, where $\boldsymbol{D}$ and $\boldsymbol{d}$ are a 2×2 matrix and a two-dimensional vector, we estimate $\boldsymbol{D}$ and $\boldsymbol{d}$ which minimise the criterion [1]

$$\begin{aligned} E_1 &= \frac{1}{2}\cdot\frac{1}{|\Omega(\boldsymbol{x})|}\int_{\Omega(\boldsymbol{x})} |\boldsymbol{J}^\top(\boldsymbol{D}\boldsymbol{x} + \boldsymbol{d}) + \boldsymbol{f}_t|^2 d\boldsymbol{y} \\ &= \frac{1}{2}\cdot\frac{1}{|\Omega(\boldsymbol{x})|}\int_{\Omega(\boldsymbol{x})} \left|\left(\boldsymbol{J}, (\boldsymbol{x}^\top \otimes \boldsymbol{J})\right) + \boldsymbol{f}_t\right|^2 d,\ \boldsymbol{v}_{(1)} = \begin{pmatrix} \boldsymbol{d} \\ vec\boldsymbol{D} \end{pmatrix} \end{aligned} \tag{5}$$

as an extension of eq. (2). From the relation $\frac{\partial E_1}{\partial \boldsymbol{v}_{(1)}} = 0$, we have the system of linear equations

$$\boldsymbol{A}_{(1)}\boldsymbol{v}_{(1)} + \boldsymbol{b}_{(1)} = 0 \tag{6}$$

for

$$\boldsymbol{A}_{(1)} = \begin{pmatrix} \boldsymbol{G}, & \boldsymbol{x}^\top \otimes \boldsymbol{G} \\ \boldsymbol{x} \otimes \boldsymbol{G}, & (\boldsymbol{x}\boldsymbol{x}^\top) \otimes \boldsymbol{G} \end{pmatrix},\ \boldsymbol{v}_{(1)} = \begin{pmatrix} \boldsymbol{d} \\ vec\boldsymbol{D} \end{pmatrix},\ \boldsymbol{b}_{(1)} = \begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{x} \otimes \boldsymbol{a} \end{pmatrix} \tag{7}$$

for the point $\boldsymbol{x}$ which is the centre point of the windowed area $\Omega(\boldsymbol{x})$.

The piecewise quadratic optical flow field is expressed as

$$\boldsymbol{u} = \begin{pmatrix} \boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{x} \\ \boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} \end{pmatrix} + \boldsymbol{D}\boldsymbol{x} + \boldsymbol{d} = (\boldsymbol{e} \otimes \boldsymbol{x})^\top Diag(\boldsymbol{P}, \boldsymbol{Q})(\boldsymbol{e} \otimes \boldsymbol{x}) + \boldsymbol{D}\boldsymbol{x} + \boldsymbol{d} \tag{8}$$

for 2×2 symmetric matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$, a 2×2 matrix $\boldsymbol{D}$ and a two-dimensional vector $\boldsymbol{d}$. Therefore, the minimiser of the criterion

$$E_2 = \frac{1}{2|\Omega(\boldsymbol{x})|}\int_{\Omega(\boldsymbol{x})} \left|\boldsymbol{J}^\top \left(\begin{pmatrix} \boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{x} \\ \boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} \end{pmatrix} + \boldsymbol{D}\boldsymbol{x} + \boldsymbol{d}\right) + \boldsymbol{f}_t\right|^2 d\boldsymbol{y} \tag{9}$$

is the the matrix equation

$$\boldsymbol{G}_{(2)}\boldsymbol{v}_{(2)} + \boldsymbol{b}_{(2)} = 0 \tag{10}$$

[1] The matrix equation $\boldsymbol{A}\boldsymbol{X}\boldsymbol{B} = \boldsymbol{C}$ is replaced to the linear system of equations $(\boldsymbol{B}^\top \otimes \boldsymbol{A})vec\boldsymbol{X} = vec\boldsymbol{C}$.

at each point $\boldsymbol{x}$, for

$$\boldsymbol{G}_{(2)} = \begin{pmatrix} \boldsymbol{G}, & \boldsymbol{x}^\top \otimes \boldsymbol{G}, & \boldsymbol{x}_\otimes^\top \otimes \boldsymbol{G}\boldsymbol{X}_\otimes^\top \\ \boldsymbol{x} \otimes \bar{\boldsymbol{G}}, & (\boldsymbol{x}\boldsymbol{x}^\top) \otimes \boldsymbol{G}, & (\boldsymbol{x}\boldsymbol{x}_\otimes^\top) \otimes \boldsymbol{G}\boldsymbol{X}_\otimes^\top \\ \boldsymbol{x}_\otimes \otimes \boldsymbol{X}_\otimes \boldsymbol{G}, & (\boldsymbol{x}_\otimes \boldsymbol{x}^\top) \otimes \boldsymbol{X}_\otimes \boldsymbol{G}, & (\boldsymbol{x}_\otimes \boldsymbol{x}_\otimes^\top) \otimes \boldsymbol{X}_\otimes \boldsymbol{G}\boldsymbol{X}_\otimes^\top \end{pmatrix}$$

$$\boldsymbol{v}_{(2)} = \begin{pmatrix} \boldsymbol{d} \\ vec\boldsymbol{D} \\ vec\boldsymbol{C} \end{pmatrix}, \quad \boldsymbol{b}_{(2)} = \begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{x} \otimes \boldsymbol{a} \\ ((\boldsymbol{x}\boldsymbol{x}_\otimes^\top) \otimes \boldsymbol{X})\boldsymbol{a} \end{pmatrix}, \tag{11}$$

where $\boldsymbol{e} = (1,1)^\top$, $\boldsymbol{C} = Diag(\boldsymbol{P}, \boldsymbol{Q})$, $\boldsymbol{x}_\otimes = \boldsymbol{e} \otimes \boldsymbol{x}$, $\boldsymbol{X}_\otimes = \boldsymbol{I} \otimes \boldsymbol{x}$ and $\boldsymbol{I}$ is the 2×2 identity matrix. This matrix equation is used for the computation of the piecewise quadratic field as an extension of the LK method. For numerical experiment, we set $\Omega(\boldsymbol{c}) = \{\boldsymbol{x}|\ |\boldsymbol{x} - \boldsymbol{c}|_\infty \leq k\}$, for a positive integer k, where $|\boldsymbol{x}|\infty$ is the l_∞ norm on the plane $\mathbf{R}^2$. Therefore, in eq. (11), we set $\boldsymbol{x} := \boldsymbol{x} - \boldsymbol{c}$ for the computation of the optical flow vector of point $\boldsymbol{c}$.

3 l_2^2-l_2 Optimisation

For the computation of the local optical flow field, we deal with the minimisation of the functional

$$J_{2221}(\boldsymbol{x}) = \frac{1}{2}|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}|_2^2 + \lambda|\boldsymbol{x}|_2, \tag{12}$$

where we set $\boldsymbol{A} := \boldsymbol{G}\backslash\boldsymbol{G}_{(2)}$, $\boldsymbol{x} := \boldsymbol{u}\backslash\boldsymbol{v}_{(2)}$ and $\boldsymbol{b} := \boldsymbol{b}\backslash\boldsymbol{v}_{(2)}$. Since the functional derivative of J_{2221} with respect to $\boldsymbol{x}$ is

$$\frac{\delta J_{2221}(\boldsymbol{x})}{\delta \boldsymbol{x}} = \boldsymbol{A}^\top(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}) + \lambda\frac{\boldsymbol{x}}{|\boldsymbol{x}|_2}, \tag{13}$$

the minimiser of eq. (12) is the solution of

$$(\boldsymbol{A}^\top\boldsymbol{A} + \frac{\lambda}{|\boldsymbol{x}|_2}\boldsymbol{I})\boldsymbol{x} = \boldsymbol{A}^\top\boldsymbol{b}. \tag{14}$$

We compute the solution of eq. (14) using the iteration form

$$\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{x}^{(n)} = \boldsymbol{b}^{(n)}, \quad \boldsymbol{b}^{(n)} = \boldsymbol{A}^\top\boldsymbol{b} - \frac{\lambda}{|\boldsymbol{x}^{(n-1)}|_2}\boldsymbol{x}^{(n-1)} \tag{15}$$

until $|\boldsymbol{x}^{(n+1)} - \boldsymbol{x}^{(n)}|_2 < \epsilon$, where

$$\boldsymbol{x}^{(n)} = (\boldsymbol{A}^\top\boldsymbol{A} - \lambda^{(n)}\boldsymbol{I})^{-1}\boldsymbol{A}^\top\boldsymbol{b}, \ \lambda^{(n)} = \frac{\lambda}{|\boldsymbol{x}^{(n-1)}|} \tag{16}$$

This procedure is performed in Algorithm 1. In this algorithm, $\boldsymbol{x}(i)$ expresses the ith element of vector $\boldsymbol{x}$.

Algorithm 1. $l_2^2 - l_2$ Minimisation

Data: $\boldsymbol{x}^0 := \mathbf{1}$, $k := 0$, $0 \leq \delta \ll 1$, $0 < \epsilon$
Result: minimiser of $\frac{1}{2}|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}|_2^2 + \lambda|\boldsymbol{x}|_2$
while $|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}|_2 > \delta$ **do**
  $\lambda^{(k-1)} := \frac{\lambda}{|\boldsymbol{x}^{(k-1)}|_2}$;
  solve $(\boldsymbol{A}^\top \boldsymbol{A} + \lambda^{(k)} \boldsymbol{I})\boldsymbol{x}^{(k)} = \boldsymbol{A}^\top \boldsymbol{b}$;
  if $\boldsymbol{x}^{(k)}(i) = 0$ **then**
    $\boldsymbol{x}^{(k)}(i) := \boldsymbol{x}^{(k)}(i) + \epsilon$;
  end
  $k := k + 1$
end

We call this method of the optical flow computation the $l_2^2 - l_2$ quadratic Kanade-Lucas-Tomasi tracker (2221QKLT tracker). Moreover, to ensure stable and robust computation, we use the pyramid-transform-based [8,9] multiple resolution method described in Algorithm 2. We call the method based on Algorithm 2 the pyramid-based 2221QKLT (2221PQKLT) tracker.

Algorithm 2. Quadratic-Optical-Flow Computation with Gaussian Pyramid

Data: $\boldsymbol{u}^{L+1} := 0$, $L \geq 0$, $l := L$
Data: $f_k^L \cdots f_k^0$
Data: $f_{k+1}^L \cdots f_{k+1}^0$
Result: optical flow u_k^0
while $l \geq 0$ **do**
  $f_{k+1}^l := f_{k+1}^l(\cdot + E(\boldsymbol{u}_k^{l+1}), k+1)$;
  compute $\boldsymbol{C}_k^l$, $\boldsymbol{D}_k^l$ and $\boldsymbol{d}_k^l$;
  $\boldsymbol{u}_k^l := \boldsymbol{x}_1^{l\top} \boldsymbol{C}_k^l \boldsymbol{x}_1^l + \boldsymbol{D}_k^l \boldsymbol{x}^l + \boldsymbol{d}_k^l$;
  $l := l - 1$
end

4 Transportation of Motion Direction

We define the coherency of motion along the time axis and in a scene. Then, we introduce a measure for the evaluation of the coherency of motion in an image sequence.

Definition 1. *If a vector field on an image generated by the motion of a scene and moving objects is spatially and temporally smooth, we call this property of the field motion coherency.*

Therefore, rapid changes in the spatial direction of motion causes collapses of motion coherency on the imaging plane, even if the spatial motion of the object is smooth. Moreover, the sudden halting of a moving object destroys motion smoothness and causes the collapses of the motion coherency.

The p-Wasserstein distance between a pair of distributions $f(x)$ and $g(y)$ for $x \in X$ and $y \in Y$ is

$$W_p(f,g) = \left(\min_c \int_X \int_Y |f(x) - g(y)|^p c(x,y) dx dy \right)^{\frac{1}{p}}. \tag{17}$$

For discrete probabilistic distributions $F = \{f_i\}_{i=1}^n$ and $G = \{g_j\}_{j=1}^n$ such that $\sum_{i=1}^n f_i = 1$ and $\sum_{i=1}^n g_i = 1$, setting $d_{ij} = |f_i - g_j|^p$, the distance between distributions F and G is computed as

$$DW_p(F,G) = \min_{x_{ij}} \left(\sum_{i=1}^n \sum_{j=1}^n d_{ij} x_{ij} \right) \tag{18}$$

subject to the conditions $\sum_{i=1}^n x_{ij} = f_i$, $\sum_{j=1}^n x_{ij} = g_j$ and $x_{ij} \geq 0$. The minimisation is achieved by solving the linear programming for transportation problem. DW_1 is called the earth mover's distance in computer vision and pattern recognition.

Setting $f(t) \geq 0$ and $g(t) \geq 0$ to be cyclic distributions on $[0, 2\pi]$, the Wasserstein distance for the cyclic distributions $f(s + 2\pi) = f(x)$ and $g(t + 2\pi) = g(t)$ is

$$CW_p(f,g) = \left(\min_{c,\theta} \int_0^{2\pi} \int_0^{2\pi} |f(t) - g(s-\theta)|^p c(t,s) dt ds \right)^{\frac{1}{p}}, \tag{19}$$

where $\int_0^{2\pi} f(t)dt = 1$ and $\int_0^{2\pi} g(t)dt = 1$. For the discrete cyclic distributions $F_c = \{f_i\}_{i=0}^{N-1}$ and $G_c = \{g_i\}_{i=0}^{N-1}$, such that $f_{i+N} = f_i$ and $g_{j+N} = g_j$, eq. (19) becomes

$$DCW_p(F_c, G_c) = \left(\min_{c_{ij},k} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |p_i - q_{j-k}|^p c_{ij} \right)^{\frac{1}{p}}. \tag{20}$$

Therefore, setting

$$D_k = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} d_{ijk} c_{ij}, \;\; d_{ijk} = |f_i - g_{j-k}|^p, \tag{21}$$

we have the relation

$$DCW_p(F_c, G_c) = (\min_k D_k)^{\frac{1}{p}}. \tag{22}$$

We apply $CDW_p(F,G)$ to compute the transportation distance between the directional statistics F and G.

For a vector-valued function $\boldsymbol{f}(\boldsymbol{x})$ such that $\boldsymbol{x} \in \mathbf{R}^n$, setting $d = \hat{\boldsymbol{f}}(\omega; \boldsymbol{x})$ $\omega \in S^{n-1}$, to be the spherical expression of $\boldsymbol{f}$ at point $\boldsymbol{x}$, we construct the directional statistics of $\boldsymbol{f}$ at the point $\boldsymbol{a} \in \mathbf{R}^n$ as

$$h(\omega; \boldsymbol{a}, \alpha) = d, \ \boldsymbol{x} \in \Omega_{\alpha,r}(\boldsymbol{a}) \tag{23}$$

for $\Omega_{\alpha,r}(\boldsymbol{a}) = \{\boldsymbol{x}| \ |\boldsymbol{x} - \boldsymbol{a}|_\alpha \leq r\}$, where $|\boldsymbol{x}|_\alpha$ is the l_α-norm.

From the temporal optical flow field $\boldsymbol{u}(\boldsymbol{x}, t)$, we define

$$u^d(\theta; \boldsymbol{a}, t, \alpha) = \sqrt{u(x,y)^2 + v^2(x,y)}, \ \theta = \tan^{-1} \frac{v(x,y)}{u(x,y)} \tag{24}$$

for $\boldsymbol{x} = (x, y)^\top \in \Omega_{\alpha,r}(\boldsymbol{a})$. For the evaluation of motion coherency, we define

$$W(t,k) = \frac{1}{|\mathbf{A}|} \int_{\boldsymbol{y} \in \mathbf{A} \subset \mathbf{R}^2} \int_{\Omega_{r,\alpha}(\boldsymbol{y})} CW_p(u^d(\theta; \boldsymbol{x}, t+1, \alpha), u^d(\theta; \boldsymbol{x}, t, \alpha)) d\boldsymbol{x} d\boldsymbol{y}. \tag{25}$$

If the motion along an image sequence has a constant speed, the transportation of the optical flow field between a pair of successive images is zero. We use this transportation computed as the Wasserstein distance along flow fields as a measure to evaluate the motion smoothness along an image sequence.

In our numerical examples, we set $\alpha = \infty$ on $\mathbf{R}^2$ and 2. Furthermore, for the sampled optical flow fields, $\Omega_{1,r}(\boldsymbol{a})$ is the 7×7 neighbourhood of each point. Figure 1 shows the procedure for the construction of the directional statistics. Moreover, we set $N = 16$ for cyclic histograms, that is, we devide the $0 \leq \theta < 2\pi$ to $\frac{\pi}{8} i \leq \theta < \frac{\pi}{8}(i+1)$ for $i = 0, 1, \cdots, 15$.

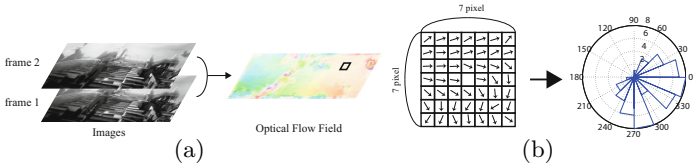


Fig. 1. Construction of directional statistics from optical flow field. (a) The optical flow field is computed from a pair of successive images from an image sequence. (b) For each point, directional statistics is constructed from flow vectors in the 7×7 neighbourhood of the point.

5 Numerical Examples

5.1 Effects of Window Size and Pyramid Hierarchy Level

Before the frame-wise evaluation, we evaluated the effects of the regularisation parameter, the size of the windows and the level of the pyramid hierarchy using the temporal continuity of the optical flow fields. For the evaluation of temporal continuity, we use the warp error (RMS error) and the temporal derivative. For the flow vector $\boldsymbol{u}(x, y, t) = (u, v)^\top$, setting

$$f'(x, y, t) = f(x - u, y - v, t + 1), \tag{26}$$

Table 1. RMS error of 2221PQKLT tracker for Motorway sequence. These experiments imply that the preferred window size and the pyramid hierarchy level for accurate and stable computation of the optical flow are 7×7 and 6. The preferred regularisation parameter λ is 0.5.

	Pyramid Level	2			4			6		
		Window Size			Window Size			Window Size		
KLT	λ	3×3	7×7	11×11	3×3	7×7	11×11	3×3	7×7	11×11
	0.2	37.8	34.5	33.2	32.1	31.7	31.7	33.8	33.2	33.1
	0.5	37.8	34.5	33.2	32.1	31.7	31.7	33.8	33.2	33.1
	0.7	37.8	34.5	33.2	32.1	31.7	31.7	33.8	33.2	33.1
	1.0	37.8	34.5	33.2	32.1	31.7	31.7	33.8	33.2	33.1
	2.0	37.8	34.5	33.2	32.1	31.7	31.7	33.8	33.2	33.1
	10.0	37.8	34.5	33.2	32.1	31.7	31.7	33.8	33.2	33.1
	Pyramid Level	2			4			6		
		Window Size			Window Size			Window Size		
QKLT	λ	3×3	7×7	11×11	3×3	7×7	11×11	3×3	7×7	11×11
	0.2	35.5	33.7	33.3	31.8	31.4	31.3	32.2	31.8	31.8
	0.5	35.5	33.7	33.3	31.8	31.4	31.3	32.2	31.8	31.8
	0.7	35.5	33.7	33.3	31.8	31.4	31.3	32.2	31.8	31.8
	1.0	35.5	33.7	33.3	31.8	31.4	31.3	32.2	31.8	31.8
	2.0	35.5	33.7	33.3	31.8	31.4	31.3	32.2	31.8	31.8
	10.0	35.5	33.7	33.3	31.8	31.4	31.3	32.2	31.8	31.8

we define the RMS error as

$$RMS\ error = \sqrt{\frac{1}{|A|} \int \int_{\boldsymbol{x} \in A} (f(x, y, t) - f'(x, y, t))^2 dxdy} \tag{27}$$

in the region of interest A at time t, where $|A|$ is the area measure of region A.

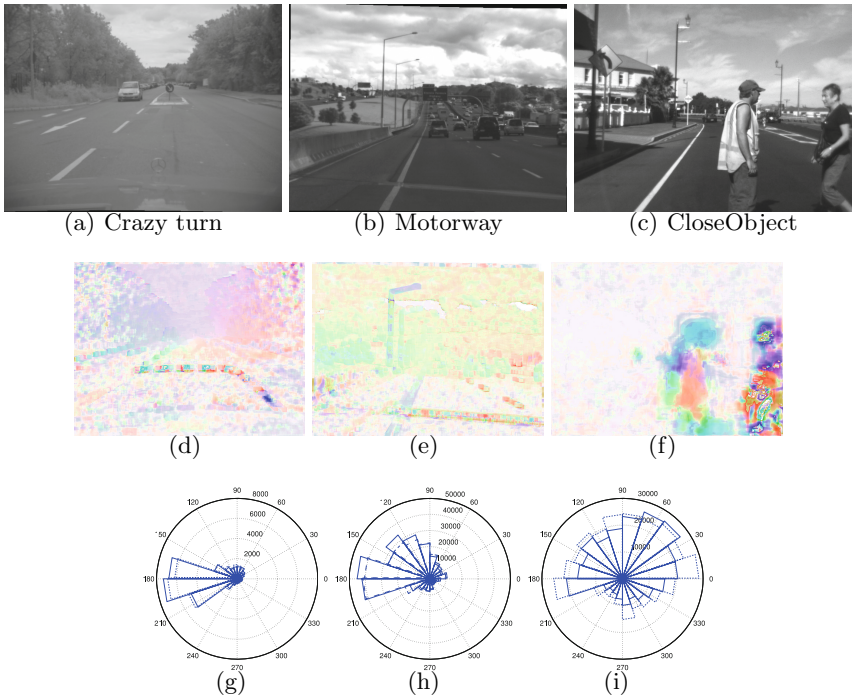
Tables 1 and 2 list the least mean errors and angle errors, respectively, for several window sizes and pyramid levels for the Motorway sequence. These results indicate that for all window sizes and pyramid levels, the performance of the 2221PQKLT tracker is superior to that of KLT tracker. These results imply the the preferred window size and pyramid hierarchy level for accurate and stable computation of the optical flow field are 7×7 and 6, respectively. Moreover, the preferred regularisation parameter λ is 0.5.

5.2 Performance of Motion Recognition

In the top row of Fig. 2, from left ot right, single images from the Crazy turn, Motorway and CloseObject sequences, respectively, are shown. The second and third rows of Fig. 2, colour charts of the optical flow fields and radar charts of the directional statistics, respectivey, are shown. Figure 3 shows the result for the

Table 2. Motorway Angle error of 2221QKLT tracker ($\lambda = 0.5$). These experiments imply that the preferred window size and the pyramid hierarchy for accurate and stable computation of the optical flow are 7×7 and 6, respectively.

		Window Size		
KLT	Pyramid Level	3×3	7×7	11×11
	2	1.72×10^{-1}	1.67×10^{-1}	1.67×10^{-1}
	4	1.58×10^{-1}	1.52×10^{-1}	1.52×10^{-1}
	6	1.47×10^{-1}	1.44×10^{-1}	1.43×10^{-1}
		Window Size		
QKLT	Pyramid Level	3×3	7×7	11×11
	2	1.66×10^{-1}	1.52×10^{-1}	1.50×10^{-1}
	4	1.53×10^{-1}	1.45×10^{-1}	1.44×10^{-1}
	6	1.42×10^{-1}	1.41×10^{-1}	1.41×10^{-1}



(a) Crazy turn (b) Motorway (c) CloseObject

(d) (e) (f)

(g) (h) (i)

Fig. 2. Images from Crazy turn, Motorway and CloseObject data sets. In the first row, From left ot right, single images from the Crazy turn, Motorway and CloseObject sequences, respectively. In the second and third rows, colour charts of the optical flow fields and radar charts of the directional statistics, respectivey.

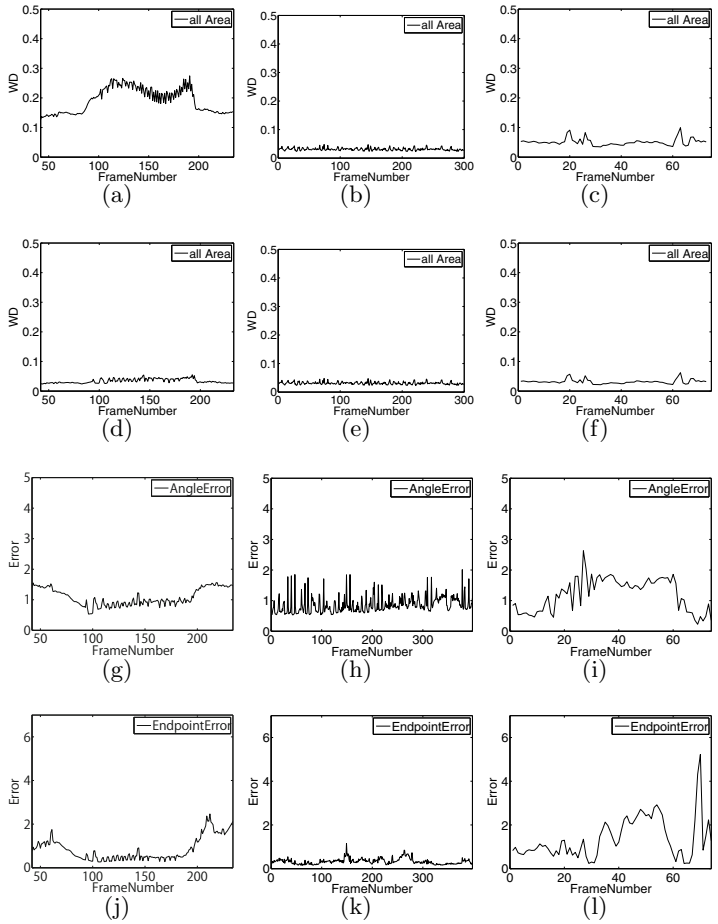


Fig. 3. Motion recognition using 2221PQKLT. From left to results, the results for the Crazy turn, Motorway and CloseObject sequences are shown. From top to bottom, temporal trajectories of the 2-Wasserstain distance, temporal trajectories of the cyclic 2-Wasserstain distance, the end point errors and the angle errors are shown. The angle errors and endpoint errors in the third and fourth rows do not detect events which brake motion coherency. The Wasserstein distance evaluates the transportation of the allows to detect events from transition of the optical flow field.

detection of motion coherency. From left to results, the results for the Crazy turn, Motorway and CloseObject sequences are shown. From top to bottom, temporal trajectories of the 2-Wasserstain distance, temporal trajectories of the cyclic 2-Wasserstain distance, the end point errors and the angle errors are shown. The window size and hierarchy of pyramid transform are 7×7 and 6, respectively.

If the motion displacement in a sequence of images is coherent, the transportation of directions of flow vectors is small, and the Wasserstein distance of the flow fields between a pair of successive images is small.

For the Crazy turn, Motorway and CloseObject sequences the optical flow fields and the radar charts of 184th, 127th and 35th frames, respectively are shown on the first in Fig. 2. In the Crazy turn and CloseObject sequences turn of a car and motion of men are captured around 184th and 34th frames, respectively.

In the Crazy turn sequence, a car with a mounted camera system turns to the left. This event is detected as a mode stated around 100th frame in trajectory of the Wasserstein distance. However, trajectory of the cyclic Wasserstein distance processes any modes. This statistical property implies that in a sequence object turns with a constant speed. The cyclic Wasserstein distance removes the appearance motion on an image screen. as shown in Fig. 3. Therefore, the transportation of the direction statics of optical flow fields allows us to detect the collapses of motion coherency on a screen in a sequence of images. In the Motorway sequence, cars in the opposite lane periodically moves toward the car-mounted camera then disappear, and cars in the same lane periodically move in front of the car-mounted camera. The temporal trajectory of the Wasserstein distance allows us to detect the temporal transportation of the optical flow fields as small periodic peaks in the temporal trajectory. In the CloseObject sequence, two men cross in front of a car. This event is detected as a peak of the temporal trajectory around 20th frame as shown in Fig. 3(c).

The angle errors and endpoint errors in the third and fourth rows in Fig. 3 do not detect events which brake motion coherency, since these errors evaluate the pointwise accuracy of the computed optical flow vectors. On the other hand, since the Wasserstein distance evaluates the transportation of the optical flow field, the distance allows to detect events from transition of the optical flow field. These results show that the Wasserstein distance of the directional statistics of an optical flow field sequence is an effective measure for the evaluation of motion coherency and its collapses in a sequence of images.

6 Conclusions

In this paper, by extending the KLT tracker, we develop a locally quadratic tracker for the motion analysis of long-time image sequences. The method computes locally quadratic optical flow fields using a model-fitting scheme. Furthermore, the method can select a region of interest for the optical flow computation.

References

1. Horn, B.K.P., Schunck, B.G.: Determining optical flow. Artificial Intelligence **17**, 185–204 (1981)
2. Rubner, Y., Tomasi, C., Guibas, L.J.: A metric for distributions with applications to image databases. In: Proceedings of ICCV 1998, pp. 59–66 (1998)

3. Fisher, N.I.: Statistical Analysis of Circular Data. Cambridge University Press (1993)
4. Wedel, A., Cremers, D.: Stereo Scene Flow for 3D Motion Analysis. Springer (2011)
5. Vogel, Ch., Schindler, K., Roth, S.L.: Piecewise rigid scene flow. In: Proceedings of ICCV 2013, pp. 1377–1384 (2013)
6. Vogel, Ch., Schindler, K., Roth, S.: 3D scene flow estimation with a rigid motion prior. In: Proceedings of ICCV 2011, pp. 1291–1298 (2011)
7. Rabin, J., Delon, J., Gousseau, Y.: Transportation distances on the circle. JMIV **41**, 147–167 (2011)
8. Hwang, S.-H., Lee, U.-K.: A hierarchical optical flow estimation algorithm based on the interlevel motion smoothness constraint. Pattern Recognition **26**, 939–952 (1993)
9. Amiaz, T., Lubetzky, E., Kiryati, N.: Coarse to over-fine optical flow estimation. Pattern Recognition **40**, 2496–2503 (2007)
10. Villani, C.: Optimal Transport, Old and New. Springer (2009)
11. Baker, S., Matthews, I.: Lucas-Kanade 20 years On: A unifying framework. IJCV **56**, 221–255 (2004)
12. Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: Proceedings of IJCAI 1981, pp. 674–679 (1981)
13. Shi, J., Tomasi, C.: Good features to track. In: Proceedings of CVPR 1994, pp. 593–600 (1994)
14. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proceedings of CVPR 2005 (2005)
15. Ustundag, B.C., Unel, M.: Human action recognition using histograms of oriented optical flows from depth. In: Bebis, G., et al. (eds.) ISVC 2014, Part I. LNCS, vol. 8887, pp. 629–638. Springer, Heidelberg (2014)
16. Chaudhry, R., Ravichandran, A., Hager, G.D., Vidal, R.: Histograms of oriented optical flow and binet-cauchy kernels on nonlinear dynamical systems for the recognition of human actions. In: Proceedings of CVPR 2009, pp. 1932–1939 (2009)
17. Mileva, Y., Bruhn, A., Weickert, J.: Illumination-robust variational optical flow with photometric invariants. In: Hamprecht, F.A., Schnörr, C., Jähne, B. (eds.) DAGM 2007. LNCS, vol. 4713, pp. 152–162. Springer, Heidelberg (2007)
18. Bruhn, A., Weickert, J., Schnörr, C.: Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. IJCV **61**, 211–231 (2005)
19. Zach, C., Pock, T., Bischof, H.: A duality based approach for realtime TV-L^1. In: Hamprecht, F.A., Schnörr, C., Jähne, B. (eds.) DAGM 2007. LNCS, vol. 4713, pp. 214–223. Springer, Heidelberg (2007)
20. Papenberg, N., Bruhn, A., Brox, T., Didas, S., Weickert, J.: Highly accurate optic flow computation with theoretically justified warping. IJCV **67**, 141–158 (2006)
21. Shin, Y.-Y., Chang, O.-S., Xu, J.: Convergence of fixed point iteration for deblurring and denoising problem. Applied Mathematics and Computation **189**, 1178–1185 (2007)
22. Chambolle, A.: An algorithm for total variation minimization and applications. JMIV **20**, 89–97 (2004)