Event Extraction Using Transportation of Temporal Optical Flow Fields

Itaru Gotoh[1], Hiroki Hiraoka[1], and Atsushi Imiya[2(✉)]

[1] School of Science and Engineering, Chiba University,
Yayoi-cho 1-33, Inage-ku, Chiba 263-8522, 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. In this paper, we develop a method to transform a sequence of images to a sequence of events. Optical flow, which is the vector fields of pointwise motion computed from monocular image sequences, describes pointwise motion in an environment. The method extracts the global smoothness and continuity of motion fields and detects collapses of the smoothness of the motion fields in long-time image sequences using transportation of the temporal optical flow field.

1 Introduction

Optical flow is fundamentally pointwise local motion on an imaging plane (retina) [1–4]. This pointwise motion is low-level information for perception of global motion [5–8,34,35]. In this paper, we introduce a model for the extraction of queues for perception of global motion from the optical flow fields using the temporal transportation [9,10] of optical flow fields along times.

Flow vectors locally extract point correspondences between a pair of successive images on the retina [1]. These local correspondences are applied to motion tracking because temporal evolution of a correspondence describes the temporal trajectory of a point in a video stream of images.

In Field VI of the human brain for motion perception, independent components of optical flow field on the retina [11,36] are transmitted from the medial superior temporal area (MST) to the middle temporal area (MT) [11–16]. Then, pointwise local motion is transformed to intermediate-level information for motion cognition. Flying insects also control motion using optical flow. Honey bees navigate using optical flow [17–21]. The compound eyes [18,19,38,39] of insects perceive spherical optical flow fields [38,39]. The divergence on the spherical optical flow filed indicates the direction of flying in the global environment [19]. Disparities of optical flow fields between the fields on the left and right hemi-spheres control the direction of fling in the local environment. Therefore, temporal optical flow fields generated on the spherical retina of an omnidirectional camera system provide queue for navigation [39]. These geometric properties of optical flow fields on the spherical retina are the basis of insect-inspired

L. Leal-Taixé and S. Roth (Eds.): ECCV 2018 Workshops, LNCS 11134, pp. 692–705, 2019.
https://doi.org/10.1007/978-3-030-11024-6_52

visual navigation. Geometrical processing of optical flow fields on the spherical retina yields syntactical information for robot navigation [37–39].

Autonomous vehicles navigate using images captured by a planar retina [40, 41]. We have develop an algorithm for the generation of motion semantics from optical flow fields generated on a planar retina, which is a common imaging process for non-compound eye systems.

In the previous paper, we introduced a model for the extraction of queue for recognising global spatial motion from scene flow fields using the temporal transportation of the vector field [40]. As a comparative study with our previous results, we apply the same idea to the optical flow field on a planar retina. This comparative study implies that for global motion perception, the optical flow fields, which is computed from monocular image sequence, possess the similar properties with those of the scene flow fields.

2 Metric for Optical Flow Fields

Setting $\boldsymbol{u}(\boldsymbol{x}) = (u(\boldsymbol{x}), v(\boldsymbol{x}))^\top$ for $\boldsymbol{x} = (x, y)^\top \in \mathbf{R}^2$ to be the optical flow field on two-dimensional Euclidean space, the directional histogram [22] of $\boldsymbol{u}(\boldsymbol{x})$ is obtained by integration of the magnitude of $\boldsymbol{u}(\boldsymbol{x})$ in the region of interest (ROI), that is,

$$h_{\boldsymbol{x}}(\theta; \boldsymbol{u}) = \frac{1}{|\Omega(\boldsymbol{x})|} \int_{\Omega(\boldsymbol{x})} \frac{\boldsymbol{u}(\boldsymbol{y})}{|\boldsymbol{u}(\boldsymbol{y})|} = (\cos\theta, \sin\theta)^\top |\boldsymbol{u}(\boldsymbol{y})| d\boldsymbol{y}, \tag{1}$$

where $\Omega(\boldsymbol{x})$, $|\Omega(\boldsymbol{x})|$ and $\boldsymbol{x} \in \mathbf{R}^2$ are the ROI, the area measure of the ROI and the reference point of the ROI, respectively.

The distance between two optical flow fields $\boldsymbol{u}(\boldsymbol{x})$ and $\boldsymbol{v}(\boldsymbol{x})$ in the region Λ is defined as

$$D(\boldsymbol{u}, \boldsymbol{v}) = \left(\int_\Lambda \left(\min_\phi \min_{c(\theta,\theta')} \int_0^{2\pi} \int_0^{2\pi} |h_{\boldsymbol{x}}(\theta - \phi; \boldsymbol{u}) - h_{\boldsymbol{x}}(\theta'; \boldsymbol{v})|^2 c_{\boldsymbol{x}}(\theta, \theta') d\theta d\theta' \right) d\boldsymbol{x} \right)^{\frac{1}{2}}, \tag{2}$$

where

$$\int_0^{2\pi} c_{\boldsymbol{x}}(\theta, \theta') d\theta \le h_{\boldsymbol{x}}(\theta'; \boldsymbol{u}), \quad \int_0^{2\pi} c_{\boldsymbol{x}}(\theta, \theta') d\theta' \le h_{\boldsymbol{x}}(\theta; \boldsymbol{v}), \tag{3}$$

for $c_{\boldsymbol{x}}(\theta, \theta') \ge 0$, using the transportation [9] of the directional histograms [22] of the fields.

For the discrete optical flow field $\boldsymbol{u}_{mn} = (u_{mn}, v_{mn})^\top$ at the point $(m, n)^\top$ on discrete plane $\mathbf{Z}^2$, let $\{f_{mn}(p)\}_{p=0}^{N-1}$ be the cyclic directional histogram for the directions $\boldsymbol{\omega}_N = (\cos 2\pi \frac{p}{N} \sin 2\pi \frac{p}{N})^\top$. For the discrete cyclic histograms $F_{mn} = \{f_{mn}(i)\}_{i=0}^{N-1}$ and $G_{mn} = \{g_{mn}(i)\}_{i=0}^{N-1}$, such that $f_{mn}(i + N) = f_{mn}(i)$

and $g_{mn}(i+N) = g_{mn}(i)$, we define the transportation between the histograms as

$$d_{mn}(F_{mn}, G_{mn}) = \left(\min_{k} \min_{c_{ij}^{mn}} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |f_{mn}(i) - g_{mn}(j-k)|^2 c(k)_{ij}^{mn} \right)^{\frac{1}{2}}, \quad (4)$$

$$\sum_{i=0}^{N-1} c_{ij}(k)^{mn} \leq g_{mn}(j-k), \quad \sum_{j=0}^{N-1} c_{ij}(k)^{mn} \leq f_{mn}(i), \quad c_{ij}^{mn} \geq 0. \quad (5)$$

Setting $A_{ij}^{mn}(k) = |f_{mn}(i) - g_{mn}(j-k)|^2$, the minimisation of $J_{mn}(k)$

$$J_{mn}(k) = \min_{c(k)_{ij}^{mn}} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} A_{ij}^{mn}(k) c(k)_{ij}^{mn}, \quad (6)$$

with the constraints of Eq. (5) is solved by linear programming for each $k = 0, 1, \cdots, N-1$. Then, we define the metric between discrete vector fields $\boldsymbol{u}_{mn}$ and $\boldsymbol{v}_{mn}$ in the ROI Λ on the two-dimensional discrete plane $\mathbf{Z}^2$ as

$$d(\boldsymbol{u}, \boldsymbol{v}) = \sqrt{\sum_{(m,n)^\top \in \Lambda} d_{mn}(F_{mn}, G_{mn})^2}, \quad d_{mn}(F_{mn}, G_{mn}) = \min_{k} \sqrt{J_{mn}(k)}. \quad (7)$$

Figure 1 shows the process of transportation of a pair of circular histograms f and g. (a) and (b) show two probabilistic distribution on a circle. and their samples on the circle. The top row in (c) shows the residual values after the maximum flows moved from each bin of P to bins of Q. The bottom row in (c) shows the flows that moved from P to Q as the maximum flow between histograms.

3 Symbolisation of Global Motion

The temporal trajectory of the distance between a successive pair of optical flow fields $\boldsymbol{u}(\boldsymbol{x}, t+1)$ and $\boldsymbol{u}(\boldsymbol{x}, t)$ of the spatiotemporal image $f(\boldsymbol{x}, t)$ is

$$H(t; f) = d(\boldsymbol{u}(\boldsymbol{x}, t+1), \boldsymbol{u}(\boldsymbol{x}, t)). \quad (8)$$

Setting $H_t(t; f)$ and $H_{tt}(t; f)$ to be the first and second derivatives, respectively, of the histogram $H(t : f)$, we define the interval $I_i = [t_i, t_{i+1}]$ along the time axis t using a pair of successive points for extremals $H_{tt}(t; f) = 0$. Using the l_1 linear approximation of $H(t; f)$ such that

$$\bar{H}(t; f) = a_i t + b_i, \quad (9)$$

which minimises the criterion

$$J(a_i, b_i) = \sum_{i=1}^{n} \sum_{j=1}^{n(i)} |H(t_{i(j)}; f) - (a_i t_{i(j)} + b_i)|, \quad (10)$$

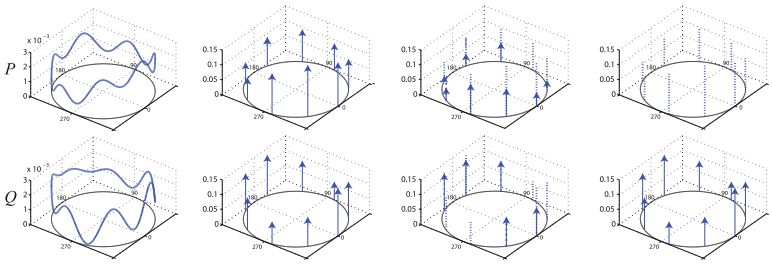


Fig. 1. Examples of the computation of the transportation distances. (a) and (b) show two probabilistic distributions on a circles and their samples on the circles. The top row in (c) shows the residual values after the maximum flows move from each bin of P to the bins of Q. The bottom row in (c) shows the flows that move from P to Q as the maximum flow between samples. (d) The state at the end of the computation. All sampled values of P are moved to bins of Q.

where $t_{i(j)} \in I_i$, we allocate signs for spatial motion.

From the sign of a_i, we define the symbols of motion of $f(\boldsymbol{x}, t)$ in the interval $I_i = [t_i, t_{i+1}]$ as $\{\nearrow, \rightarrow, \searrow\}$, where

$$S(H(t;f)) = \begin{cases} \nearrow \text{ if } a_i > 0 \text{ if } t \in I_i, \\ \rightarrow \text{ if } a_i = 0 \text{ if } t \in I_i, \\ \searrow \text{ if } a_i < 0 \text{ if } t \in I_i, \end{cases} \quad S(\log H(t;f)) = \begin{cases} \nearrow \text{ if } a_i > 0 \text{ if } t \in I_i, \\ \rightarrow \text{ if } a_i = 0 \text{ if } t \in I_i, \\ \searrow \text{ if } a_i < 0 \text{ if } t \in I_i. \end{cases} \tag{11}$$

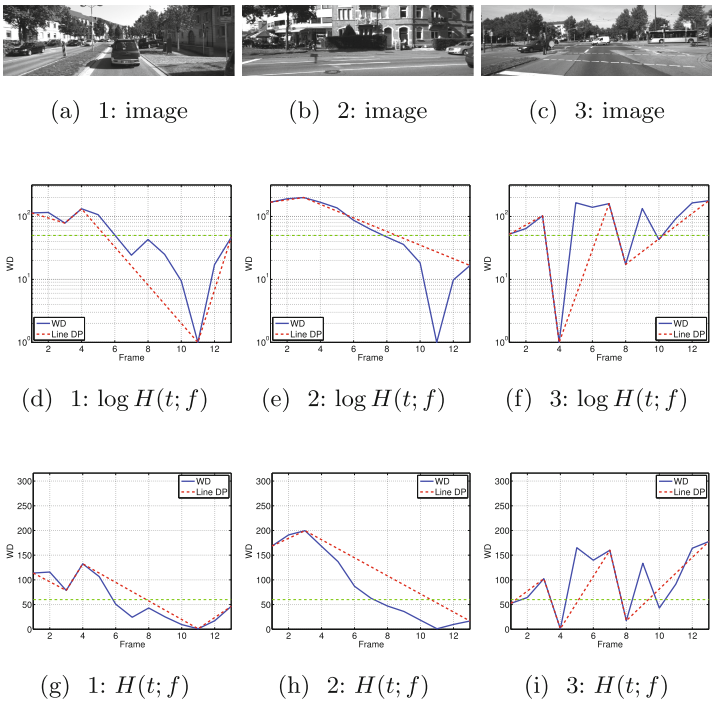
4 Numerical Examples

Table 1. Statuses of three sequences from KITTI sceneflowDataset2015[44].

No.	Motion of car	Front car	Inter-vehicle dist.	Oncoming car	Additional conditions
1	accelerating	with	increasing	without	
2	turning right	with	outside of screen	with in the first frame	
3	stopping after braking	with		without	approaching to road crossing

For numerical experiments, three image sequences from left images of KITTI-Scene Flow Dataset2015 [44] are selected.

For event extraction using Eq. (11), we employ $S(H(t;f))$ and $S(\log H(t;f))$, since $S(\log H(t;f))$ allows us to detect symbols from small perturbations of $H(t;f)$.



(a) 1: image (b) 2: image (c) 3: image

(d) 1: $\log H(t;f)$ (e) 2: $\log H(t;f)$ (f) 3: $\log H(t;f)$

(g) 1: $H(t;f)$ (h) 2: $H(t;f)$ (i) 3: $H(t;f)$

Fig. 2. Examples of the Wasserstein distances

Table 2. Extracted events

No.	Status	Car in front	Oncoming car	WD graph	
				$S(\log H(t;f))$	$S(H(t;f))$
1	acceleration	exists	no	→↘↘↘↘↗	↘↗↘↘↘↗
2	turn right	out of the view	exists	→↘↘↘↘↗	↗↗↘↘↘↘
3	deceleration and stop	exists	no	↗↘↗↘↗↗	↗↗↘↗↗↗

Table 1 lists statuses of the images. Figure 2 shows the temporal trajectories of the transportation of the vector fields. Table 2 shows the event strings extracted by linear approximation by using Eq. (10) and symbolisation by using Eq. (11). These experiments show that the algorithm extracts symbol strings, which describe the states in front of driving cars in various environments.

5 Dictionary Generation

Tables 4 and 3 show status and speed of objects in synthetic video sequences. Figure 3 shows top views of geometric configurations of objects in synthetic video sequences. Table 5 shows combinations of events as symbol strings captured by vehicle mounted camera in a synthetic world. In Figs. 4, 5 and 6, (a) and (b) show a frame view of the image sequence and its optical flow field, respectively.

In Fig. Figures 4, 5 and 6, (c), $\log H(t;f)$ and $\bar{H}(t;f)$ are the blue curve and red polygonal curve, respectively.

Tables 6 and 7 show the strings $S(\log H(t;f))$ and $S(H(t;f))$ detected by the algorithm using $\log H(t;f)$ and $H(t;f)$, respectively. Since both $\log H(t;f)$ and $H(t;f)$ are approximated by polygonal curves for the extraction of symbol strings, events are described by using $\vee$, $\wedge$ and M based on the semi-local shapes of the curves.

Table 3. Speed of objects in synthetic image sequences.

Object	ac-/de-celeration	The first frame spd. [km/h]	The last frame spd. [km/h]
ego-vehicle	no	30	30
oncoming vehicle	no	40	40
ego-vehicle	acceleration	5	30
ego-vehicle	deceleration	30	5
vehicle	no	30	30
vehicle in front	no	45	45
vehicle in front	no	15	15

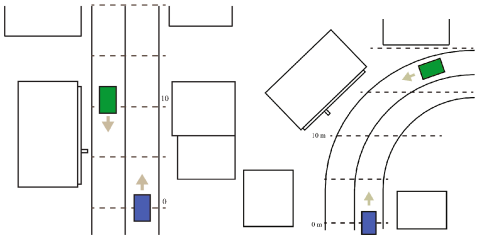
Table 4. Geometry in synthetic image sequences.

Objects	Position [m]	Width [m]	Length [m]
camera	1.5	base line width 0.3	
white line	0.05	0.4	
lane		2.8	
pavement		1.0–1.5	
oncoming vehicle			$2.0 \times 4.5 \times 1.75$

Five pairs 1 and 2, 5 and 7, 6 and 8, 13 and 14, and 15 and 16 provide same environments with and without oncoming vehicles. These examples show that pairs 1 and 2, 5 and 7, 6 and 8, 13 and 14, possess same properties for symbol string. Pairs 1 and 2, and 13 and 14 imply that the temporal transportation of optical flow vector fields achieves recognition of oncoming vehicles. The algorithm detects acceleration and deceleration of the ego-vehicle.

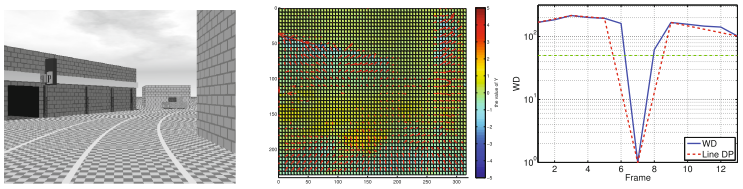
The results observed in a pair 7 and 8 show that for the detection of the directions of turning additional information is required, since the optical flow fields for left and right turning possess the same statistical properties.

The difference of the results observed in a pair 15 and 16 depends on the background properties caused by trees, since the correspondences between a pair of natural scene contains ambiguities. Moreover, for the detection of oncoming vehicle, the pointwise optical flow vectors are required.



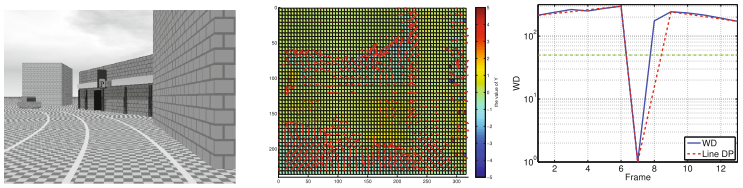
(a) Straight lane (b) Curve lane

Fig. 3. Top views of geometric configurations of objects in synthetic video sequences simulating city environments. The blue car is the ego-vehicle for experiments. The green car is the object-vehicle for experiments. The lane width is 2.8 m. The pavement width is 1 m–1.5 m.



(a) Input (b) Optical flow field (c) $\log H(t; f)$

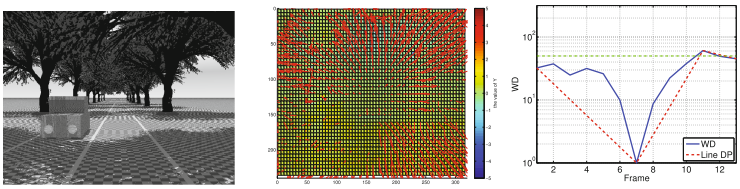
Fig. 4. Motion of synthetic image No. 7. (a) A frame view of the image seqence. (b) Optical flow field. (c) $\log H(t; f)$ and $\log \bar{H}(t; f)$ are drawn in the blue curve and red polygonal curve, respectively. (Color figure online)



(a) Input (b) Optical flow field (c) $\log H(t; f)$

Fig. 5. Motion of synthetic image No. 8. (a) A frame view of the image sequence. (b) Optical flow field. (c) $\log H(t; f)$ and $\log \bar{H}(t; f)$ are drawn in the blue curve and red polygonal curve, respectively. (Color figure online)

The algorithm does not distinguish left and right turns, since the time trajectory of the distance between to field possess the shape profiles. However, it is possible to detect the stating frame of the turns, since the symbol $\wedge$ is detected on the frame. For detection of the straight motion from real sequences, symbol strings both $S(H(t; f))$ and $S(\log H(t; f))$ are necessary, since in real sequences of the straight motion temporal local-perturbation of the optical flow vectors are detected. This local-perturbation derives perturbations on $H(t; f)$ and $\log H(t; f)$.



(a) Image (b) Optical flow field (c) $\log H(t; f)$

Fig. 6. Motion of synthetic image No. 16. (a) A frame view of the image sequence. (b) Optical flow field. (c) $\log H(t; f)$ and $\log \bar{H}(t; f)$ are drawn in the blue curve and red polygonal curve, respectively. (Color figure online)

Table 5. Events in synthetic data

No.	Motion	Car in front	Oncoming car	Event	Event label
1	straight	no	no	straight	const. spd
2			exist	straight	const. spd+obstacle
3			no	accelerate → straight	accelerate
4			no	straight → deceleration	deceleration
5	turn		no	straight → right curve	turn
6			no	straight → left curve	turn
7			exist	straight → right curve	turn +obstacle
8			exist	straight → left curve	turn +obstacle
9	straight	exist	no	inter-vehicle dist.	car in front
10			no	inter-vehicle dist. increase	car in front
11			no	inter-vehicle dist. decrease	car in front
12			no	passing in passing lane	car in front
13			no	passing in the opposite lane	car in front
14			exist	passing in the opposite lane	with obstacle
15		no	no	straight	const. spd
16			exist	straight	with obstacle

6 Discussions

For the function $f(\boldsymbol{x}, t)$ defined in $\mathbf{R}^n$, the total derivative with respect to the variable t is

$$\frac{df}{dt} = \nabla f^\top \frac{d\boldsymbol{x}}{dt} + f_t. \tag{12}$$

Mathematically, optical flow is the solution of the linear equation $\frac{df}{dt} = 0$. This inconsistent linear equation is solved by regularisation [3,23] and using local geometric constraints [1,33].

In the medical volumetric-image analysis, for instance, the motion analysis of the moving organs, we are required to deal with volumetric images defined in three-dimensional Euclidean space $\mathbf{R}^3$. In computer vision, optical flow is usually computed from planar images.

Table 6. Symbol strings extracted from $\log H(t; f)$

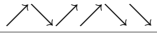
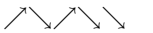
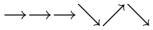
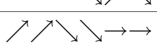
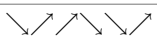
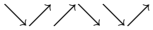
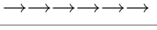
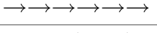
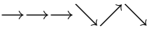
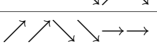
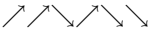
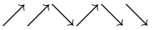
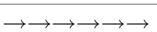
No	Signal shape	Symbols
1	M	
2	M	
3	$\vee\wedge$	
4	$\wedge$	
5	$\wedge\vee$	
6	$\wedge\vee$	
7	$\wedge\vee$	
8	$\wedge\vee$	
9	$(\vee\wedge)^n$	
10	$(\vee\wedge)^n$	
11	$(\vee\wedge)^n$	
12	polygonal curve	
13	$M2$	
14	$M2$	
15	$(\wedge\vee)^n$	
16	$\vee$ 2	

Table 7. Symbol strings extracted from $H(t; f)$

No	Signal shape	Symbols
1	flat	
2	flat	
3	$\vee\wedge$	
4	$\wedge$	
5	$\wedge\vee$	
6	$\wedge\vee$	
7	$\wedge\vee$	
8	$\wedge\vee$	
9	Flat	
10	flat	
11	flat	
12	polygonal curve	
13	$M2$	
14	$M2$	
15	flat	
16	flat	

For motion analysis with range data, setting $f(x, y, t)$ to be a grey-label image, we deal with the following system of equations

$$\begin{aligned} \frac{df}{dt} &= f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_t = 0, \\ \frac{dg}{dt} &= h_x \frac{dx}{dt} + h_y \frac{dy}{dt} - \frac{dz}{dt} + h_t = 0, \end{aligned} \tag{13}$$

where $g(x, y) = h(x, y, t) - z$ for depth z of the temporal range image $h(x, y, t)$ [24].

For colour and multi-channel images, the system of equations

$$\frac{df^\alpha}{dt} = f_x^\alpha \frac{dx}{dt} + f_y^\alpha \frac{dy}{dt} + f_t^\alpha = 0, \quad \alpha = 1, 2, \cdots, k \tag{14}$$

is derived from the k-channel images [25,26].

For the left image $f(x_l, y_l, t)$ and the right image $g(x_r, y_r, t)$ of temporal stereo-pair images, the system of equations

$$\frac{df}{dt} = f_x u_l + f_y v_l + f_t = 0 \tag{15}$$

$$\frac{dg}{dt} = g_x u_r + g_y v_r + g_t = 0 \tag{16}$$

derive the optical flow vectors $\boldsymbol{u}_l = (u_l, v_l)^\top$ and $\boldsymbol{u}_r = (u_r, v_r)^\top$ on the left and right images, respectively. After establishing correspondences between $\boldsymbol{x}_l$ and $\boldsymbol{x}_r$ and between $\boldsymbol{x}_l + \boldsymbol{u}_l$ and $\boldsymbol{x}_r + \boldsymbol{u}_r$, the stereo reconstruction algorithm computes scene flow $\dot{\boldsymbol{X}}$ in the space using disparities between temporal stereo-pair images. Estimation of correspondences is established by solving system of equations

$$f(x+d, y, t) = g(x, y, t), \;\; f(x+u_l+d_1', y+v_l+d_2', t) = g(x+u_r, y+v_r, t) \quad (17)$$

for the displacement $\boldsymbol{d} = (d, 0)^\top$ and $\boldsymbol{d}' = (d_1', d_2')^\top$.

For images on a manifold $\mathcal{M}$, the optical flow vector filed is the solution of the equation

$$\frac{df}{dt} = \nabla_{\mathcal{M}} f^\top \frac{d\boldsymbol{\nu}}{dt} + f_t = 0 \quad (18)$$

where $\nabla_{\mathcal{M}}$ is the gradient operation on the manifold. For example, if $\mathcal{M}$ is the unit sphere S^2 in three-dimensional Euclidean space $\mathbf{R}^3$, the gradient operation is

$$\nabla_{\mathcal{M}} f = \left(\frac{\partial}{\partial \theta} f, \frac{1}{\cos\theta} \frac{\partial}{\partial \phi} f \right)^\top. \quad (19)$$

Equation (18) allows us to compute the optical flow vectors on a spherical retina, which is the mathematical model of compound eyes.

In this paper, we have shown a method to extract intermediate queues for motion perception from optical flow on flow fields on the plane [34–36,41]. It is possible to apply the event extraction method based on the transportation of optical flow fields for scene flow [40] and the optical flow field on non-planar retina [38]. In reference [38], we have shown a method to extract intermediate queue for motion perception from optical flow fields on a sphere.

Moreover, we have developed a method to decompose the optical flow fields [27,28] on the surface of the moving organs [42] employing three-dimensional optical flow computation.

The optical flow fields between a pair of successive images in a sequence provide queues for image alignment. Aligning images along the time axis achieves the tracking of images in a video sequence [2]. Therefore, tracking is a sequential alignment. Multiple alignment in a space by deformation fields derives the deformation-based average of images.

For a collection of images $\{f_i(\boldsymbol{x})\}_{i=1}^m$, setting $\boldsymbol{u}_i(\boldsymbol{x})$ to be the deformation fields, the minimiser f of the energy functional

$$J = \sum_{i=1}^{m} \int_{\mathbf{R}^n} (f(\boldsymbol{x} + \boldsymbol{u}_i(\boldsymbol{x})) - f_i(\boldsymbol{x}))^2 d\boldsymbol{x} \quad (20)$$

with appropriate constraints derives the deformation-based average of the collection of images $\{f_i(\boldsymbol{x})\}_{i=1}^m$ [29,43]. The deformation-based average was applied for motion analysis of a volumetric beating-heart sequence.

The directional gradient of an image $f(\boldsymbol{x})$ at the point $\boldsymbol{x} = (x, y)^\top$ in the direction of $\boldsymbol{\omega} = (\cos\theta, \sin\theta)^\top$ is computed as $\boldsymbol{\omega}^\top \nabla f$. The directional gradient evaluates the steepness, smoothness and flatness of $f(\boldsymbol{x})$ along the direction of

vector $\boldsymbol{\omega}$. Setting F to be a injective mapping for gradient, the gradient-based feature constructed by F satisfies the relations $F(\nabla f) = 0$ and $F(\nabla f) = F(\nabla g)$ if $f = 0$ and $f = g + a$ for constant a, respectively.

The census transform is computed by

$$s(\boldsymbol{x}) = \frac{1}{2\pi} \int_0^{2\pi} u(\boldsymbol{\omega}^\top \nabla f) d\theta \tag{21}$$

where u is the Heaviside function. The directional histogram (DH) is computed by

$$h_{\boldsymbol{x}}(\theta) = \frac{G_f(\theta, \boldsymbol{x})}{\int_0^{2\pi} G_f(\theta, \boldsymbol{x}) d\theta}, \quad G_f(\theta, \boldsymbol{x}) = \int_{\Omega(\boldsymbol{x})} \boldsymbol{\omega}^\top \nabla f(\boldsymbol{y}) d\boldsymbol{y}, \tag{22}$$

such that $h_{\boldsymbol{x}}(\theta + 2\pi) = h_{\boldsymbol{x}}(\theta)$, where $\boldsymbol{x} \in \mathbf{R}^2$ is the centre of the region $\Omega(\boldsymbol{x})$. The vector $\boldsymbol{x}$ is used as the index of the DH. We call $h_{\boldsymbol{x}}(\theta)$ the HoG signature of f.

The census transform encodes local geometric property of the gradient vector fields as scalar function. The HoG signature encodes semi-global geometric properties of the gradient vector field as a scalar function. These encoded features are used for matching of images and motion detection [30]. Our transform in Eq. (1) encodes the global geometric properties of motions on the retina as a scalar function using optical flow vector fields. Then, using this encoded motion vector field, we define a metric between a pair of motion fields for the extraction of events on video streams.

Since $\boldsymbol{v} = \frac{-f_t}{|\nabla f|^2} \nabla f$ is a solution of $\frac{df}{dt} = 0$, the optical flow vector is expressed as $\boldsymbol{u} = \frac{-f_t}{|\nabla f|^2} \nabla f + \alpha \nabla f^\perp$ for an appropriate scalar α, where $\nabla f^\top (\nabla f^\perp) = 0$. If the motion perpendicular to the gradient of the edges of the segments is small, that is, α is small, $\boldsymbol{u} \sim \mu \nabla f$ for an appropriate real number μ. This relation between the optical flow filed and the gradient field implies that events in the image stream detected by the features encoded by Eq. (1) are those caused by the temporal fluctuations of the gradient of the foreground.

In ref. [32] the on-line algorithm for detection of a polygonal curve from a time signal of a string of conversation dialogs. was proposed based on the randomized Hough transform. This algorithm is pre-processing for the construction of the syntactic trees of conversation dialogs. The event detection from video sequence is an extension of syntactic analysis of dialog signals to image sequences.

In pedestrian detection, annotated data for designing classifier is generated using artificially generated virtual world [31]. It is possible to extend the idea for event detection from image observed by vehicle mounted camera system. We generated symbol sequences from events in virtual world. Events detected from generated symbol strings coincide with the events detected from real world test data sequences.

7 Conclusions

We proposed a method for the symbolisation of the temporal transition of environments using statistical analysis of the flow field. The algorithm allows us to interpret a sequence of images as a string of events.

A machine can control a car to avoid incidents by detecting abnormalities using event strings stored in a dictionary. The symbolisation of temporal optical fields is suitable for the generation of entries in such a dictionary.

We have introduced a framework for syntactical interpretation of dynamic scenes using the temporal transportation of the optical flow fields. The future work for us is to derive semantics of the motion fields from strings of symbols. Multiscale image analysis of the dynamic scenes provides hierarchies of the motions [32] in the scenes from temporal local deformation to global fluctuations. Therefore, these hierarchies of motions would define the syntactic structure and semantic meaning of dynamic scene. The optical flow fields are the important queries for linguistic analysis of the dynamic scene.

References

1. 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)
2. Tomasi, C., Kanade, T.: Detection and tracking of point features. Int. J. Comput. Vis. **9**, 137–154 (1991)
3. Horn, B.K.P., Schunck, B.G.: Determining optical flow. Artif. Intell. **17**, 185–203 (1981)
4. Hwang, S.-H., Lee, U.-K.: A hierarchical optical flow estimation algorithm based on the interlevel motion smoothness constraint. Pattern Recogn. **26**, 939–952 (1993)
5. Vaina, L.M., Beardsley, S.A., Rushton, S.K. (eds.): Optic Flow and Beyond. SL, vol. 324. Springer, Dordrecht (2004). https://doi.org/10.1007/978-1-4020-2092-6
6. Duffy, C.J.: Optic flow analysis for self-movement perception. Int. Rev. Neurobiol. **44**, 199–218 (2000)
7. Lappe, M., Bremmer, F., van den Berg, A.V.: Perception of self-motion from visual flow. Trends Cogn. Sci. **3**, 329–336 (1999)
8. Calow, D., Krüger, N., Wörgötter, F., Lappe, M.: Statistics of optic flow for self-motion through natural scenes. In: Ilg, U., Bülthoff, H.H., Mallot, A.H., et al. (eds.) Dynamic Perception, pp. 133–138. IOS Press (2004)
9. Villani, C.: Optimal Transport. Old and New. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-540-71050-9
10. Rabin, J., Delon, J., Gousseau, Y.: Transportation distances on the circle. JMIV **41**, 147–167 (2011)
11. Sabatini, S.P.: A physicalist approach to first-order analysis of optic flow fields in extrastriate cortical areas. In: ICANN 1999 (1999)
12. Park, K.-Y., Jabri, M, Lee, S.-Y., Sejnowski, T.J.: Independent components of optical flows have MSTd-like receptive fields. In: Proceedings of the 2nd International Workshop on ICA and Blind Signal Separation, pp. 597–601 (2000)
13. Wurtz, R.: Optic flow: a brain region devoted to optic flow analysis? Curr. Biol. **8**, R554–R556 (1998)

14. Greenlee, M.: Human cortical areas underlying the perception of optic flow: brain imaging studies. Int. Rev. Neurobiol. **44**, 269–292 (2000)
15. Andersen, R.A.: Neural mechanisms of visual motion perception in primates. Neuron **18**, 865–872 (1997)
16. Newsome, W.T., Baré, E.B.: A selective impariment of motion perception following lesions of the middle temporal visual area (MT). J. Neurosci. **8**, 2201–2211 (1988)
17. Pan, C., Deng, H., Yin, X.-F., Liu, J.-G.: An optical flow-based integrated navigation system inspired by insect vision. Biol. Cybern. **105**, 239–252 (2011)
18. Franceschini, N.: Visual guidance based on optic flow: a biorobotic approach. J. Physiol. Paris **98**, 281–292 (2004)
19. Srinivasan, M.V.: Honeybees as a model for the study of visually guided flight, navigation, and biologically inspired robotics. Physiol. Rev. **91**, 413–460 (2011)
20. Serres, J.R., Ruffier, F.: Optic flow-based collision-free strategies: from insects to robots. Arthropod Struct. Dev. **46**, 703–717 (2017)
21. Sobey, P.J.: Active navigation with a monocular robot. Biol. Cybern. **71**, 433–440 (1994)
22. Fisher, N.I.: Statistical Analysis of Circular Data. Cambridge University Press, Cambridge (1993)
23. Weickert, J., Schnörr, C.: Variational optic flow computation with a spatio-temporal smoothness constraint. J. Math. Imaging Vis. **14**, 245–255 (2001)
24. Spies, H., Jähne, B., Barron, J.L.: Range flow estimation. Comput. Vis. Image Underst. **85**, 209–231 (2002)
25. Barron, J.L., Klette, R.: Quantitative color optical flow. In: Proceedings of ICPR 2002, vol. 4, pp. 251–255 (2002)
26. Golland, P., Bruckstein, A.M.: Motion from color. Comput. Vis. Image Underst. **68**, 346–362 (1997)
27. Kirisits, C., Lang, L.F., Scherzer, O.: Decomposition of optical flow on the sphere. GEM Int. J. Geomathematics **5**, 17–141 (2014)
28. Lukas, F., Lang, L.K., Scherzer, O.: Optical flow on evolving sphere-like surfaces. Inverse Probl. Imaging **11**, 305–338 (2017)
29. Rumpf, M., Wirth, B.: Variational methods in shape analysis. In: Scherzer, O. (ed.) Handbook of Mathematical Methods in Imaging, pp. 1819–1858. Springer, New York (2015). https://doi.org/10.1007/978-1-4939-0790-8_56
30. Hafner, D., Demetz, O., Weickert, J.: Why is the census transform good for robust optic flow computation? In: Kuijper, A., Bredies, K., Pock, T., Bischof, H. (eds.) SSVM 2013. LNCS, vol. 7893, pp. 210–221. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38267-3_18
31. Vázquez, D., López, A.M., Marín, J., Ponsa, D., Gómez, D.: Virtual and real world Adaptation for pedestrian detection. IEEE PAMI **36**, 797–809 (2014)
32. Imiya, A.: Detection of piecewise-linear signals by the randomized Hough transform. Pattern Recogn. Lett. **17**, 771–776 (1996)
33. Imiya, A., Iwawaki, K.: Voting method for the detection of subpixel flow field. Pattern Recognit. Lett. **24**, 197–214 (2003)
34. Ohnishi, N., Imiya, A.: Featureless robot navigation using optical flow. Connect. Sci. **17**, 23–46 (2005)
35. Ohnishi, N., Imiya, A.: Appearance-based navigation and homing for autonomous mobile robot. Image Vis. Comput. **31**, 511–532 (2013)
36. Ohnishi, N., Imiya, A.: Independent component analysis of optical flow for robot navigation. Neurocomputing **71**, 2140–2163 (2008)

37. Alibouch, B., Radgui, A., Rziza, M., Aboutajdine, D.: Optical flow estimation on omnidirectional images: an adapted phase based method. In: Elmoataz, A., Mammass, D., Lezoray, O., Nouboud, F., Aboutajdine, D. (eds.) ICISP 2012. LNCS, vol. 7340, pp. 468–475. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31254-0_53
38. Torii, A., Imiya, A., Sugaya, H., Mochizuki, Y.: Optical flow computation for compound eyes: variational analysis of omni-directional views. In: De Gregorio, M., Di Maio, V., Frucci, M., Musio, C. (eds.) BVAI 2005. LNCS, vol. 3704, pp. 527–536. Springer, Heidelberg (2005). https://doi.org/10.1007/11565123_51
39. Mochizuki, Y., Imiya, A.: Pyramid transform and scale-space analysis in image analysis. In: Dellaert, F., Frahm, J.-M., Pollefeys, M., Leal-Taixé, L., Rosenhahn, B. (eds.) Outdoor and Large-Scale Real-World Scene Analysis. LNCS, vol. 7474, pp. 78–109. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34091-8_4
40. Kato, T., Itoh, H., Imiya, A.: Motion language of stereo image sequence. In: CVPR Workshops, pp. 1211–1218 (2017)
41. Ohnishi, N., Mochizuki, Y., Imiya, A., Sakai, T.: On-line planar area segmentation from sequence of monocular monochrome images for visual navigation of autonomous robot. In: VISAPP 2010, pp. 435–442 (2010)
42. Kameda, Y., Imiya, A.: The William Harvey code: mathematical analysis of optical flow computation for cardiac motion. In: Rosenhahn, B., Klette, R., Metaxas, D.N. (eds.) Human Motion, Understanding, Modelling, Capture, and Animation, Computational Imaging and Vision, vol. 36, pp. 81–104. Springer, Dordrecht (2006). https://doi.org/10.1007/978-1-4020-6693-1_4
43. Inagaki, S., Itoh, H., Imiya, A.: Multiple alignment of spatiotemporal deformable objects for the average-organ computation. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) ECCV 2014. LNCS, vol. 8928, pp. 353–366. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16220-1_25
44. http://www.cvlibs.net/datasets/kitti/eval_scene_flow.php