Spatiotemporal Deformable Prototypes for Motion Anomaly Detection: Supplementary material

Robert Bensch
bensch@cs.uni-freiburg.de
Thomas Brox
brox@cs.uni-freiburg.de
Olaf Ronneberger
ronneber@cs.uni-freiburg.de

Department of Computer Science and BIOSS Centre for Biological Signalling Studies
University of Freiburg
Germany
http://lmb.informatik.uni-freiburg.de

1 Supertrajectory Representation

1.1 Supertrajectories

We denote the set of raw trajectories that form one supertrajectory by $\mathcal{X}_i \subset \{1,\dots,N_{\text{raw}}\}$, where $i \in \{1,\dots,N_{\text{super}}\}$. A supertrajectory i is computed by averaging the positions of all grouped raw trajectories at each time point

$$\mathbf{x}(i,t) = \begin{cases} \frac{\sum_{i_{\text{raw}} \in \mathcal{X}_i} w_{\text{raw}}(i_{\text{raw}},t) \cdot \mathbf{x}_{\text{raw}}(i_{\text{raw}},t)}{\sum_{i_{\text{raw}} \in \mathcal{X}_i} w_{\text{raw}}(i_{\text{raw}},t)} & \text{if } w(i,t) = 1 \\ \mathbf{0} & \text{else,} \end{cases} \tag{1}$$

where $w(i,t) = \max_{i_{\text{raw}} \in \mathcal{X}_i} w_{\text{raw}}(i_{\text{raw}},t)$. Additionally, $\mathbf{x}$ is low-pass filtered in temporal direction (average filter), to suppress noise and especially high frequencies that occur, when trajectories start and end frequently.

2 Detection and Elastic Registration of Motion Patterns

2.1 Detection Hypotheses via Efficient Hashing

To efficiently detect a prototype pattern $\mathbf{x}^a$ in a new test sequence $\mathbf{x}^b$ we modified [1] to deal with the spatiotemporal setting and point trajectories. Random point pairs of $\mathbf{x}^a$ and $\mathbf{x}^b$ are chosen and stored in a hash table using rotationally invariant features as table indices. Hash collisions provide transformation hypotheses with a time complexity of $O(n)$ for the first hypothesis, that converges to $O(1)$ for further hypotheses. The original algorithm [1] works on surface points and surface normals. To adapt it to our setting, we use points on supertrajectories and their velocity vectors instead: $\mathbf{q} = \big(\mathbf{x}(i,t_1);\dot{\mathbf{x}}(i,t_1);\mathbf{x}(j,t_2);\dot{\mathbf{x}}(j,t_2);t_1;t_2\big)$, where $\dot{\mathbf{x}} := \frac{d\mathbf{x}}{dt}$ (see Fig. 1). The rest of the algorithm (including hypothesis verification by

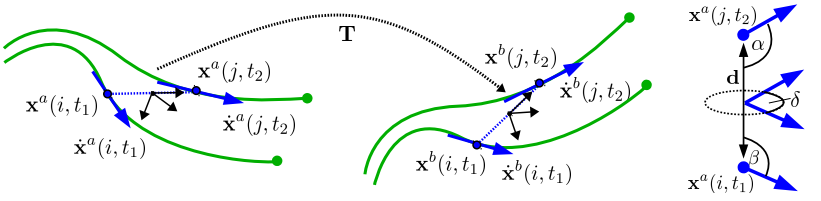

Figure 1: (a) Relative transformation $\mathbf{T}$ between matching two point constellations $\mathbf{q}_a$ and $\mathbf{q}_b$ on point trajectories. (b) Relation between $\mathbf{x}^a(i,t_1)$ and $\mathbf{x}^a(j,t_2)$. Transferred from [1].

computing "overlapping areas") was adapted accordingly. The final output of the algorithm is a number of rigid transformation hypotheses. We parameterize them as a temporal shift t_{shift} and a spatial rigid transformation $\mathbf{T}:\mathbb{R}^3\to\mathbb{R}^3:\mathbf{x}(i,t)\mapsto R\mathbf{x}(i,t)+\mathbf{b}$, with 3×3 rotation matrix R and translation $\mathbf{b}$, such that $\mathbf{x}'(i,t)=\mathbf{T}(\mathbf{x}(i,t-t_{\text{shift}}))$, $\forall i\in\{1,\ldots,N_{\text{super}}\}$ and $w'(i,t)=w(i,t-t_{\text{shift}})$ respectively.

2.2 Elastic Registration

The elastic transformation is parameterized by a spatial deformation function $\mathbf{u}(i,t):\Omega_a\to\mathbb{R}^3$ and a temporal warping $\tau(i,t):\Omega_a\to\mathbb{R}$, such that $\mathbf{x}'(i,t)=\mathbf{x}(i,t-\tau(i,t))+\mathbf{u}(i,t-\tau(i,t))$, and $w'(i,t)=w(i,t-\tau(i,t))$ accordingly.

Data term The data term is the sum of squared distances of all points of test pattern $\mathbf{x}^b$ to the associated points in pattern $\mathbf{x}^a$, subject to spatial deformation and temporal warping

$$E_{\text{data}}(\mathbf{u},\tau,\sigma)=\sum_{\substack{(i_b,t_b)\in\Omega_b\\ w(i_b,t_b)=1\\ \sigma(i_b,t_b)\neq 0}}\Psi\left(\|\mathbf{x}^a(A[\tau,\sigma](i_b,t_b))+\mathbf{u}(A[\tau,\sigma](i_b,t_b))-\mathbf{x}^b(i_b,t_b)\|^2\right)+\sum_{\substack{(i_b,t_b)\in\Omega_b\\ w(i_b,t_b)=1\\ \sigma(i_b,t_b)=0}}d^2_{\text{undef}}. \quad (2)$$

Analogous to Eq. 2 (in the main paper), we define an association function $A[\tau,\sigma]:\Omega_b\to\Omega_a:(i_b,t_b)\to(i_a,t_a)$, which now depends on the temporal warping τ instead.

Spatial smoothness We define spatial smoothness across trajectories, such that the elastic coupling between trajectories $C(i,j)=\exp(-d(i,j)^2/2r^2)$ depends on the pairwise distances $d(i,j)$. The elastic coupling within the prototype pattern is described by a smoothness energy on the spatiotemporal deformation functions for each pair of prototype trajectories

$$E_{\text{spatial}}(\mathbf{u},\tau)=\sum_{\substack{i,j,t\\ (i,t)\in\Omega_a\wedge(j,t)\in\Omega_a\\ w(i,t)=1\wedge w(j,t)=1}}C(i,j)\cdot\left(\|\mathbf{u}(i,t)-\mathbf{u}(j,t)\|^2+\beta_{\text{temp}}(\tau(i,t)-\tau(j,t))^2\right).$$

Temporal smoothness A temporal smooth transformation along trajectories is enforced by the energy term

$$E_{\text{temp}}(\mathbf{u},\tau) = \sum_{\substack{(i,t)\in\Omega_a \\ w(i,t)=1}} \|\dot{\mathbf{u}}(i,t)\|^2 + \beta_{\text{temp}}(\dot{\tau}(i,t))^2, \tag{3}$$

where $\dot{\mathbf{u}} := \frac{d\mathbf{u}}{dt}$ and $\dot{\tau} := \frac{d\tau}{dt}$ respectively. The weighting between spatial deformation and temporal warping is determined by β_{temp}.

Assignment smoothness Moreover, a temporally smooth assignment is preferred by the smoothness term

$$E_{\text{assign}}(\sigma) = \sum_{\substack{i,t \\ (i,t)\in\Omega_b \wedge (i,t-\Delta t)\in\Omega_b \\ w(i,t)=1 \wedge w(i,t-\Delta t)=1 \\ \sigma(i,t)\neq 0 \wedge \sigma(i,t-\Delta t)\neq 0}} d^2\left(\sigma(i,t),\sigma(i,t-\Delta t)\right),$$

that penalizes temporal assignment changes from one to another trajectory by their pairwise distance $d(i,j)$. In this way, assignment changes between dissimilar trajectories get strong penalization. Let Δt_{min} be the shortest interval for assignment changes of σ, then $\Delta t \in (0,\Delta t_{min}] \subset \mathbb{R}$ must be chosen.

3 Motion Anomaly Detection

3.1 Learning a Spatiotemporal Deformable Prototype Model

We build a statistical model that captures:

Global spatiotemporal deformations and data fitting costs after registration For both, we define bounds for validating prototype registrations. Global deformation parameters are computed by standard PCA on the concatenation of deformation parameters $\mathbf{u}'_d(i,t)$ and $\tau_d(i,t)$, where $(i,t) \in \Omega_a \wedge w(i,t) = 1$, $\forall d \in \{1,\cdots,D\}$. Let $\mathbf{y}_d$ be the representation of deformations for detection d in PCA space, with elements $y_d(k)$ and $k \in \{1,\cdots,K\}$. We define a bounding hyper-cuboid in PCA space based on the lower and upper bounds $y_{\min}(k) = \min_d(y_d(k))$ and $y_{\max}(k) = \max_d(y_d(k))$, for each k. Prototype data fitting costs c_d are computed analogous to the data term in Eq. 2, but averaging over valid points on the prototype domain Ω_a and applying association functions mapping from Ω_a to Ω_b instead. We define an upper bound based on the maximum $c_{\max} = \max_d(c_d)$.

Residual distances remaining after elastic registration For each prototype pattern point $\mathbf{x}^a(i,t)$ the obtained residual distances with associated training pattern points are locally aggregated and learned using a Gaussian residual model.

3.2 Reconstruction by Prototype Placements

For reconstructing a whole test pattern by prototype placements we apply a greedy search algorithm. It iteratively finds best placements of prototype patterns into the test pattern. Candidate placements are obtained from an over-complete set of prototype detections. A

priority list defines the order of placements. It is sorted by the score $s' = r \cdot s$, where $s = d^2_{\max} - c$, with data fitting cost c, described in Sec. 3.1, and maximum value $d^2_{\max}$. r is the ratio of the candidate pattern overlapping with so far unreconstructed test pattern. Thus, a high score is achieved only with good data fitting combined with good overlapping with so far unreconstructed test pattern. In each step, the k-best candidates are elastically registered. Registered candidates are accepted only if all deformation and data fitting parameters are within the learned bounds. Rejected candidates are removed from the priority list. The scores of accepted candidates are updated (temporarily) and the best accepted candidate is selected for reconstruction and removed from the priority list. The unreconstructed test pattern is updated accordingly. The algorithm stops, if the test pattern is reconstructed completely, or if no candidates remain that can reconstruct parts of the unreconstructed test pattern.

4 Experiments

4.1 Anomaly Detection in Juggling Patterns

Table 1 gives an overview of the test set of the juggling dataset. It contains 29 sequences with juggling patterns from five different persons including anomalies, recorded in different viewpoint settings (distance to camera, view angle).

Seq.	Person	Frames	Distance to camera	View angle	Seq.	Person	Frames	Distance to camera	View angle
1	1	141	1.5m	0°	15	5	460	1.5m	0°
2	1	460	1.5m	±20°	16	5	420	1.5m	±20°
3	1	220	2.5m	0°	17	5	440	2.5m	0°
4	1	460	2.5m	±20°	18	5	410	2.5m	±20°
5	2	250	1.5m	0°	19	3+5	370	2.5m	0°
6	2	440	2.5m	0°	20	3+5	480	2.5m	0°
7	3	410	1.5m	0°	21	1	280	1.5m	0°...±20°
8	3	430	1.5m	±20°	22	1	410	1.5m...2.5m	0°...±20°
9	3	420	2.5m	0°	23	1	350	1.5m...2.5m	0°...±20°
10	3	420	2.5m	±20°	24	1	320	1.5m	0°
11	4	430	1.5m	0°	25	1	250	1.5m	0°
12	4	360	1.5m	±20°	26	1	390	1.5m	0°
13	4	440	2.5m	0°	27	2	360	1.5m...2.5m	0°...±20°
14	4	440	2.5m	±20°	28	2	400	1.5m	0°...±20°
					29	2	450	1.5m	±20°
Total #sequences: 29, #frames: 11.111, #persons: 5									

Table 1: Juggling dataset test set overview.

Additional results are shown in Video 1, complementary to the results presented in Sec. 6.1 of the main paper.

Video 1: Additional anomaly detection results. The video shows the prototype and training sequences and anomaly detection results for five test sequences. Anomaly detection results are rendered into the test sequences and give a good impression of the anomaly localization capability. They include the scenario of a moving juggler and two persons juggling side-by-side.

References

[1] S. Winkelbach, S. Molkenstruck, and F. M. Wahl. Low-cost laser range scanner and fast surface registration approach. In *Pattern Recognition (Proc. DAGM)*, pages 718–728. Springer, LNCS, 2006.