An evaluation of 3D motion flow and 3D pose estimation for human action recognition

Matteo Munaro, Stefano Michieletto, Emanuele Menegatti
University of Padova

Presenter: Stefano Michieletto

3D extension of optical flow


Skeleton features

IAS-Lab Action Dataset
http://robotics.dei.unipd.it/actions
IAS-Lab Action Dataset

(a) Check watch    (b) Cross arms    (c) Get up    (d) Kick    (e) Pick up

(f) Point    (g) Punch    (h) Scratch head    (i) Sit down    (j) Standing

(k) Throw from bottom up    (l) Throw over head    (m) Turn around    (n) Walk    (o) Wave
<table>
<thead>
<tr>
<th></th>
<th>#actions</th>
<th>#people</th>
<th>#samples</th>
<th>RGB</th>
<th>skel</th>
</tr>
</thead>
<tbody>
<tr>
<td>[32]</td>
<td>6</td>
<td>1</td>
<td>198</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>[25]</td>
<td>12</td>
<td>4</td>
<td>48</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>[14]</td>
<td>20</td>
<td>10</td>
<td>567</td>
<td>no²</td>
<td>yes</td>
</tr>
<tr>
<td>[27]</td>
<td>16</td>
<td>10</td>
<td>320</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>[28]³</td>
<td>10</td>
<td>21</td>
<td>461</td>
<td>yes⁴</td>
<td>no</td>
</tr>
<tr>
<td>[22]</td>
<td>11</td>
<td>12</td>
<td>660</td>
<td>yes</td>
<td>yes³</td>
</tr>
<tr>
<td>Ours</td>
<td>15</td>
<td>12</td>
<td>540</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

3D Motion Flow
Our approach estimates **motion vectors** of single 3D points

- No **viewpoint** and no **scale** dependence  
  (typical of 2D approaches)
- Computed directly on **point clouds**
- **No** need for **projectable** point clouds  
  (obtained from a dense matrix of depth data)
3D Motion Flow: Points Matching
Matching in XYZHSV space

SOURCE (CIRCLES) AND TARGET (TRIANGLES) POINT CLOUDS
3D Motion Flow: Points Matching
Matching in XYZHSV space

KNN IN XYZ SPACE
3D Motion Flow: Points Matching
Matching in XYZHSV space

NN IN HSV SPACE
1. **Correspondence** finding
   - K-NN in XYZ space
   - NN in HSV space
2. Outlier rejection by means of **reciprocal** correspondences
3. Computation of **3D velocity** vectors
   \[ \mathbf{v}_i = \left( \mathbf{p}_i^{\text{target}} - \mathbf{p}_i^{\text{source}} \right) / \left( t_i^{\text{target}} - t_i^{\text{source}} \right) \]
4. Outlier rejection based on velocity **magnitude**
   - Points with \( \| \mathbf{v}_i \| \) below threshold are discarded
   - Isolated moving points are discarded.
3D Motion Flow: Outlier Rejection

Velocity Magnitude
Descriptors
ô The space around a person is divided into a grid of cubes
ô The sum of the velocity vectors is computed for every cube
ô **Single-frame** descriptor: concatenation of sum vectors from every cube
ô **Sequence** descriptor: concatenation of single-frame descriptors (evenly spaced in time)
**Single-frame** descriptors:
- concatenation of joint *positions*
- concatenation of link *orientations*
- concatenation of joint *positions* and link *orientations*

\[
d_P = [x_1 \ y_1 \ z_1 \ \ldots \ x_N \ y_N \ z_N],
\]

\[
d_O = [q_1^1 \ q_1^2 \ q_1^3 \ q_1^4 \ \ldots \ q_N^1 \ q_N^2 \ q_N^3 \ q_N^4],
\]

\[
d_{TOT} = [d_P^1 \ d_O^1 \ \ldots \ d_P^N \ d_O^N].
\]

**Sequence** descriptor: concatenation of single-frame descriptors (evenly spaced in time)
Results
Results with 3D Motion Flow
IAS-Lab Action Dataset
Results with 3D Motion Flow
IAS-Lab Action Dataset

**MEANFLOW**

58%

**SUMFLOW**

NO outlier rejection

81.9%

<table>
<thead>
<tr>
<th>Action</th>
<th>MEANFLOW</th>
<th>SUMFLOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Check Watch-1</td>
<td>18.9</td>
<td>97.2</td>
</tr>
<tr>
<td>Cross Arms-2</td>
<td>28.0</td>
<td>28.0</td>
</tr>
<tr>
<td>Get Up-3</td>
<td>25.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Kick-4</td>
<td>4.8</td>
<td>100.0</td>
</tr>
<tr>
<td>Pick Up-5</td>
<td>7.8</td>
<td>100.0</td>
</tr>
<tr>
<td>Point-6</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Punch-7</td>
<td>3.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Scratch Head-8</td>
<td>55.8</td>
<td>38.9</td>
</tr>
<tr>
<td>Sit Down-9</td>
<td>25.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Standing-10</td>
<td>22.2</td>
<td>100.0</td>
</tr>
<tr>
<td>Throw From Bottom Up-11</td>
<td>22.2</td>
<td>100.0</td>
</tr>
<tr>
<td>Throw Over Head-12</td>
<td>13.9</td>
<td>100.0</td>
</tr>
<tr>
<td>Turn Around-13</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Walk-14</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Wave-15</td>
<td>2.8</td>
<td>2.8</td>
</tr>
</tbody>
</table>

Mean recognition accuracy: 58%

Mean recognition accuracy: 81.9%
SUMFLOW NO outlier rejection 81.9%

SUMFLOW with outlier rejection 85.2%

Results with 3D Motion Flow IAS-Lab Action Dataset
Results with Motion Flow
IAS-Lab Action Dataset

SUMFLOW with outlier rejection 85.2%

SUMFLOW after PCA 87.4%

Mean recognition accuracy: 85.2%

Mean recognition accuracy: 87.4%
Mean Recognition Accuracy

when varying the number of frames in the sequence descriptor
Link Orientations

55.9%

Joint Positions and Links Orientations

66.9%
Joint Positions and Links Orientations
66.9%

Mean recognition accuracy: 66.9%

Joint Positions
76.7%

Mean recognition accuracy: 76.7%
SUMFLOW 85.2%

Joint Positions 76.7%

Mean recognition accuracy: 85.2%

Mean recognition accuracy: 76.7%
The **skeleton** descriptors are very **fast** to compute (30 fps)

The **motion flow** computation and description takes **0.25s** (4 fps)

<table>
<thead>
<tr>
<th>Step</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Octree initialization</td>
<td>0.0150s</td>
</tr>
<tr>
<td>Point clouds matching</td>
<td>0.2300s</td>
</tr>
<tr>
<td>Flow vectors computation</td>
<td>0.0001s</td>
</tr>
<tr>
<td>Outlier removal</td>
<td>0.0004s</td>
</tr>
<tr>
<td>Descriptor computation</td>
<td>0.0003s</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>0.2458s</strong></td>
</tr>
</tbody>
</table>
Direct 3D motion flow estimation from colored point clouds

- Novel 3D Motion descriptor
- Action recognition (near) real time performance (4 fps) and 87% accuracy
- Better accuracy (8.5% more) than a skeleton based approach

Future works:

- 3D Flow estimation from mobile robots
- Test histogram-based flow descriptors
Thanks for the attention!

Matteo Munaro
munaro@dei.unipd.it

Stefano Michieletto
michieletto@dei.unipd.it

Emanuele Menegatti
emg@dei.unipd.it