**Cross-view Activity Recognition using Hankelets**

Binlong Li, Octavia I. Camps and Mario Sznaier
Dept. of Electrical and Computer Engineering

---

**Abstract**

We introduce a new feature for cross-view activity recognition: the "Hankelet". This type of feature captures dynamic properties of short tracklets that are invariant to viewpoint changes and time shifts. Experiments using Hankelets on the IXMAS database show a 20% improvement over the state of the art.

**GOAL**

To recognize an activity from a different viewpoint than the one used for training.

**EXISTING APPROACHES**

- Geometric constraints [32]
- Track body joints [21,22]
- 3D Models [8,15,30,31]
- Quasi-invariant geometric features [10,11]
- Transfer features across views [7,18]
  
Best performance is far below the state of art performance for single view activity recognition.

**A Dynamics-based Feature: Hankelet**

Tracklets Detection (shown in green)

*HANKELETS* Codebook: Clustering

Hankelet Codebook: Clustering

Dissimilarity Score:

\[
\text{Dissimilarity Score} = \frac{1}{n} \sum_{i=1}^{n} \text{Dissimilarity}_{\text{ij}}
\]

\[
\text{Dissimilarity}_{\text{ij}} = \left( 1 - \frac{\text{Distance}}{\text{Max Distance}} \right)
\]

\[
\text{Distance} = \sqrt{\sum_{i=1}^{n} (\text{Feature}_i - \text{Feature}_{ij})^2}
\]

**BAGS of HANKELETS**

Bi-lingual Hankelets: subset visible from different viewpoints

**Experimental Results**

**Single View: KTH Dataset**

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Perf</th>
<th>Act Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ours</td>
<td>95.89</td>
<td>Walking</td>
</tr>
<tr>
<td>Cao et al. [3]</td>
<td>95.02</td>
<td>Walking</td>
</tr>
<tr>
<td>Wang et al. [29]</td>
<td>94.7</td>
<td>Jogging</td>
</tr>
<tr>
<td>Le et al. [13]</td>
<td>93.9</td>
<td>Jogging</td>
</tr>
<tr>
<td>Li et al. [14]</td>
<td>93.6</td>
<td>Jogging</td>
</tr>
</tbody>
</table>

**Cross-View: IXMAS Dataset**

Cross-view activity recognition.

**Comparing against state-of-the-art cross-view techniques**

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Perf</th>
<th>Act Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ours</td>
<td>95.79</td>
<td>Walking</td>
</tr>
<tr>
<td>Cao et al. [3]</td>
<td>95.02</td>
<td>Walking</td>
</tr>
<tr>
<td>Wang et al. [29]</td>
<td>94.7</td>
<td>Jogging</td>
</tr>
<tr>
<td>Le et al. [13]</td>
<td>93.9</td>
<td>Jogging</td>
</tr>
<tr>
<td>Li et al. [14]</td>
<td>93.6</td>
<td>Jogging</td>
</tr>
</tbody>
</table>

---

This work is supported in part by NSF grants 08-31753 and 0833982/DMS, AFOSR grant FA9550-09-1-0253, and the MC2 Center of Excellence under award number 09-ST-061-ED0001.

---

*See Table 5 in paper for detailed comparisons*