Action Recognition with HOG-OF Features

Florian Baumann  
Florian.Baumann@tnt.uni-hannover.de

Applications
- YouTube
- Video search, indexing, mining
- Medicine and Sociology
- Home-based rehabilitation
- Behavior analysis

Challenges of Human Action Recognition
- Moving backgrounds
- Different view points
- Moving camera
- Occlusions
- Intra-class variation
- Time-varying actions
- Few training examples
- Long training time
- Real-time capability

Contribution and Workflow
- Frame-by-Frame learning approach
- Create two processing streams
  - HOGs to gather static object appearances [1]
  - Optical Flow (OF) to incorporate motion information [2]
- Separately learning two Random Forest classifiers [3]
- Combining probability functions by product law
- Evaluation on the well-known KTH dataset [4]
  - 5-fold cross-validation
  - Original training/testing split

KTH Dataset
- Publicly available and well-known
- Single-view dataset
- 25 persons
- Four different scenarios
- 600 videos

Experiments
5-fold cross validation

<table>
<thead>
<tr>
<th></th>
<th>Schindler and Good</th>
<th>Zhang et al.</th>
<th>Proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>87.98 %</td>
<td>94.60 %</td>
<td>97.59 %</td>
</tr>
</tbody>
</table>

Original training/testing split

<table>
<thead>
<tr>
<th></th>
<th>Laptev et al.</th>
<th>Zhang et al.</th>
<th>Wang et al.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>91.80 %</td>
<td>94.00 %</td>
<td>94.20 %</td>
</tr>
</tbody>
</table>

Proposed method 94.44 %

O’Hara and Draper 97.90 %
Sadanand and Corso 98.20 %

Conclusion
- Frame-by-Frame learning approach
- Combination of HOG and OF in two parallel streams
- Simple and efficient
- State-of-the-art results

References