1 Introduction

### Challenges
- Visual and textual data significantly differ in representation and distribution characteristics
- Visual features: typically inaccurate and unreliable
- Tag labels: inherently sparse and incomplete

### Existing methods
- Tags are organised and used in a flat structure
- Visual and textual data significantly differ in representation and distribution characteristics

### Contributions
- A novel tag-based video clustering method capable of effectively fusing information from ambiguous visual features and sparse textual tags
- A unified tag correlation based algorithm to cope with tag sparseness/incompleteness

2 Methods

### Hierarchical-Multi-Label Random Forest (HML-RF)
- Conventional individual information gain
- Multi-tags information gain
- Adaptive hierarchical multi-label information gain

#### Advantages of the HML-RF model
- A hybrid classification and clustering forests
- Employing tags to constrain tree structure learning
- A new objective function for multi-tags
- Model abstract-to-specific text tag ontology and multiple tag correlations

3 Evaluations

<table>
<thead>
<tr>
<th>Input</th>
<th>Method</th>
<th>Purity</th>
<th>NMI</th>
<th>RI</th>
<th>F1</th>
<th>AIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>V&amp;T-bn</td>
<td>SpClust-bn[1]</td>
<td>0.71</td>
<td>0.73</td>
<td>0.91</td>
<td>0.50</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>CC-Forest[7]</td>
<td>0.23</td>
<td>0.23</td>
<td>0.86</td>
<td>0.13</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>AASC[8]</td>
<td>0.26</td>
<td>0.10</td>
<td>0.86</td>
<td>0.13</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>MMC[4]</td>
<td>0.24</td>
<td>0.11</td>
<td>0.88</td>
<td>0.13</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>HML-RF(Ours)</td>
<td>0.83</td>
<td>0.80</td>
<td>0.96</td>
<td>0.74</td>
<td>0.72</td>
</tr>
</tbody>
</table>

4 Take-home Messages

- An unsupervised text tag-based video semantic clustering framework for jointly exploiting heterogeneous visual and tag data
- A novel Hierarchical Multi-Label Random Forest (HML-RF) model incorporating tag abstractness hierarchy (ontology)
- Handling tag sparseness and incompleteness by exploiting multi-tag statistical correlations

References: