Semantic Image Interpretation

Ivan Donadello$^{1,2}$ Luciano Serafini (Advisor)$^1$

$^1$Fondazione Bruno Kessler, Via Sommarive, 18 I-38123, Trento, Italy

$^2$DISI University of Trento, Via Sommarive, 9 I-38123, Trento, Italy

November 26, 2015
Context

- Huge diffusion of digital images in recent years;
- Lack of semantic based retrieval systems for images, that is no complex queries: “a person riding a horse on a meadow”; 
- Semantic gap between numerical image features and human semantics;
- Need a method that automatically understands the semantic content of images.

Relevance:

- Semantic content based image retrieval via a query language;
- Semantic content enrichment with Semantic Web resource.
Case Study: Clustering-Based Cost Function

- Task: part-whole recognition, i.e., discovery complex objects from their parts;
- part-whole recognition can be seen as a clustering problem;
  - parts of the same object tend to be grouped together;
Case Study: Clustering-Based Cost Function

- Task: **part-whole recognition**, i.e., discovery complex objects from their parts;
- part-whole recognition can be seen as a **clustering problem**;
  - parts of the same object tend to be grouped together;
- cost function as a clustering optimisation function.
Clustering: grouping a set of input elements into groups (clusters) such that:

- Intra-cluster distance minimized
- Inter-cluster distance maximized
Case Study: Clustering-Based Cost Function

- Clustering: grouping a set of input elements into groups (clusters) such that:

- *clustering solution* of \((\mathcal{P}, \mathcal{I}_p, \mathcal{G})\) is \(\mathcal{C} = \{ C_d \mid d \in \Delta^{\mathcal{I}_p} \}\) where \(C_d = \{ \mathcal{G}(d') \mid d' \in \Delta^{\mathcal{I}_p}, \langle d, d' \rangle \in \text{hasPart}^{\mathcal{I}_p} \}\);

- \(d\) represents the composite object, the centroid of the cluster;
Case Study: Clustering-Based Cost Function

Mixing numeric and semantic features:

- **grounding distance** $\delta_G(d, d')$: the Euclidean distance between the centroids of $G(d)$ and $G(d')$;
- **semantic distance** $\delta_O(d, d')$ is the shortest path in $O$:

  - if $\text{Muzzle}(d')$, $\text{Tail}(d'')$ then $\delta_O(d', d'') = 2$;
  - if $\text{Muzzle}(d')$, $\text{Horse}(d)$ then $\delta_O(d', d) = 1$;
Case Study: Clustering-Based Cost Function

- **Inter-cluster distance** $\Gamma$:

- **Intra-cluster distance** $\Lambda$:
  \[ \delta = \delta_G + \delta_O \]

- **Cost function**:
  \[ S(\mathcal{P}, \mathcal{I}_P, \mathcal{G})_O = \alpha \cdot \Gamma + (1 - \alpha) \cdot \Lambda \]
Minimising the Cost Function

The Clustering Part-Whole Algorithm (CPWA) approximates the minimum of the cost function.
Minimising the Cost Function

The Clustering Part-Whole Algorithm (CPWA) approximates the minimum of the cost function.
Minimising the Cost Function

The Clustering Part-Whole Algorithm (CPWA) approximates the minimum of the cost function.
Minimising the Cost Function

The Clustering Part-Whole Algorithm (CPWA) approximates the minimum of the cost function.
Minimising the Cost Function

The Clustering Part-Whole Algorithm (CPWA) approximates the minimum of the cost function.
Minimising the Cost Function

The Clustering Part-Whole Algorithm (CPWA) approximates the minimum of the cost function.
The Clustering Part-Whole Algorithm (CPWA) approximates the minimum of the cost function.
Minimising the Cost Function

The Clustering Part-Whole Algorithm (CPWA) approximates the minimum of the cost function.

<table>
<thead>
<tr>
<th>Labelled Picture</th>
<th>Singletons</th>
<th>Agglomerative Step</th>
<th>Cost Minimization</th>
<th>Agglomerative Step</th>
</tr>
</thead>
<tbody>
<tr>
<td>{face1}</td>
<td>{face1}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>{tail2}</td>
<td>{tail2}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>{leg3}</td>
<td>{leg3}</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The algorithm proceeds through steps, starting with labelled pictures and singletons, then moving to agglomerative steps, cost minimization, and finally more agglomerative steps.
Minimising the Cost Function

The Clustering Part-Whole Algorithm (CPWA) approximates the minimum of the cost function.
Evaluation

Comparing the predicted partial model with the ground truth, two measures:

- **grouping (GRP):**

![Diagram showing grouping (GRP)]
Evaluation

Comparing the predicted partial model with the ground truth, two measures:

- **grouping (GRP):**

  ![Diagram for grouping (GRP)]

- **complex-object type prediction (COP):**

  ![Diagram for complex-object type prediction (COP)]
Evaluation

Comparing the predicted partial model with the ground truth, two measures:

- **grouping (GRP):**
  ![Diagram of grouping (GRP)]

- **complex-object type prediction (COP):**
  ![Diagram of complex-object type prediction (COP)]

- precision, the fraction of predicted pairs that are correct;
- recall, the fraction of correct pairs that are predicted.
Experiments and Results

Experiments Setting

- **Ground truth** of 203 manually obtained labelled pictures on the urban scene domain;
- manually built **ontology** with basic formalism of meronymy of the domain;
- **task**: discovering complex objects from their parts in pictures.

Results

<table>
<thead>
<tr>
<th></th>
<th>prec&lt;sub&gt;GRP&lt;/sub&gt;</th>
<th>rec&lt;sub&gt;GRP&lt;/sub&gt;</th>
<th>F&lt;sub&gt;1&lt;/sub&gt;&lt;sub&gt;GRP&lt;/sub&gt;</th>
<th>prec&lt;sub&gt;COP&lt;/sub&gt;</th>
<th>rec&lt;sub&gt;COP&lt;/sub&gt;</th>
<th>F&lt;sub&gt;1&lt;/sub&gt;&lt;sub&gt;COP&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPWA</td>
<td>0.67</td>
<td>0.81</td>
<td>0.71</td>
<td>0.71</td>
<td>0.82</td>
<td>0.86</td>
</tr>
</tbody>
</table>
Experiments and Results

Experiments Setting

- **Ground truth** of 203 manually obtained labelled pictures on the urban scene domain;
- manually built **ontology** with basic formalism of meronymy of the domain;
- **task**: discovering complex objects from their parts in pictures.

Results

<table>
<thead>
<tr>
<th></th>
<th>prec\textsubscript{GRP}</th>
<th>rec\textsubscript{GRP}</th>
<th>F1\textsubscript{GRP}</th>
<th>prec\textsubscript{COP}</th>
<th>rec\textsubscript{COP}</th>
<th>F1\textsubscript{COP}</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPWA</td>
<td>0.67</td>
<td>0.81</td>
<td>0.71</td>
<td>0.71</td>
<td>0.82</td>
<td>0.86</td>
</tr>
<tr>
<td>Baseline</td>
<td>0.45</td>
<td>0.71</td>
<td>0.48</td>
<td>0.66</td>
<td>0.69</td>
<td>0.66</td>
</tr>
</tbody>
</table>

- **Baseline**: clustering without semantics;
Error Analysis (Low Precision)

Low precision: less clusters with respect to the dataset (≈ more false positives).
Error Analysis (Low Recall)

Low recall: more clusters with respect to the dataset (= more false negatives).
Conclusions and Future Work

- Theoretical framework for SII: partial model that minimizes a cost function;
- cost function as a clustering optimization function;
- clustering algorithm that approximates the cost function;
- explicitly using semantics improves the results;
- future work:
  - integrating of semantic segmentation algorithms;
  - generalizing to other relations;
  - extending the evaluation to a standard dataset;
  - using general purposes ontologies;
Conclusions and Future Work

- Theoretical framework for SII: partial model that minimizes a cost function;
- cost function as a clustering optimization function;
- clustering algorithm that approximates the cost function;
- explicitly using semantics improves the results;
- future work:
  - integrating of semantic segmentation algorithms;
  - generalizing to other relations;
  - extending the evaluation to a standard dataset;
  - using general purposes ontologies;
Thanks for listening

Questions?