Retrieval of MPEG-7 based Semantic Descriptions

Mathias Lux, Michael Granitzer
Know-Center

{mlux, mgrani}@know-center.at
Content

- MPEG-7 and Semantic Descriptions
- Indexing & Search of Semantic Descriptions
- Implementation Details
- Evaluation
- Demonstration
- Future Work
What is MPEG-7?

"Multimedia Content Description Interface"

ISO/IEC Standard for MuMe Meta Data Representation

- XML as well as Compressed Binary

Organized in Descriptors (D) and Descriptor Schemes (DS)
MPEG-7 Semantic Descriptions
(1/2)

*Semantic DS* allows Semantic Descriptions:

- **Base Descriptor** „SemanticBase“
- Inherited are Agents, Places, Times, Events, Concepts, Objects, ...
- D Instances are interpreted as Nodes

<table>
<thead>
<tr>
<th>Instanz</th>
<th>Descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathias Lux</td>
<td>Semantic Agent</td>
</tr>
<tr>
<td>Orient Express</td>
<td>Semantic Object</td>
</tr>
<tr>
<td>Traveling</td>
<td>Semantic Event</td>
</tr>
<tr>
<td>Railway München-Salzburg</td>
<td>Semantic Location</td>
</tr>
</tbody>
</table>
MPEG-7 Semantic Descriptions (2/2)

- Semantic Relations interconnect Semantic Objects.
- 45 different Relations, 44 inverse.
- Relations are directed Edges interconnecting Nodes.
Properties of Semantic Descriptions

- Semantic Descriptions are „Labeled Graphs“.
- Node Labels are unique.
- Number of possible Edge Labels is bounded.
- One Semantic Object Instance can be found in multiple Graphs (Domain specific).
Indexing

Indexing is needed for

- Semantic Objects (SO), which are the Nodes
  - Fulltext Index
  - Node IDs
- Semantic Descriptions (SD), which are the Graphs
  - Based on the Paths in the Graphs
  - Paths as Strings
Paths in the Index

<table>
<thead>
<tr>
<th>Term</th>
<th>Path length</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>_agentOf_1_location</td>
<td>1</td>
</tr>
<tr>
<td>_locationOf_4_2</td>
<td>1</td>
</tr>
<tr>
<td>_patientOf_3_2</td>
<td>1</td>
</tr>
<tr>
<td>_1_agentOf_2_patient_3</td>
<td>2</td>
</tr>
<tr>
<td>_1_agentOf_2_location_4</td>
<td>2</td>
</tr>
<tr>
<td>_3_patientOf_2_location_4</td>
<td>2</td>
</tr>
</tbody>
</table>
Constructing a Query (1/2)

Query: *Mathias Lux* is doing *something* at the *I-Know*

- *Mathias Lux* is identified as node with ID 1
- *I-Know* is identified as node with ID 2
- „Something“ is a wildcard

![Diagram]

1. agentOf

2. location
Constructing a Query (2/2)

The Query is based on the paths:

- Paths with Length 0: "_1" and "_2"
- Paths with Length 1: "_agentOf_1_*" and "_locationOf_2_*"
- Paths with Length 2: "_2_agentOf_*_location_2"
Query Interface

Creating Queries is like drawing Graphs.

- Possible Candidates for Nodes are identified.
- With Query Expansion Query Graphs are constructed.
- From each Query Graph a Query String is constructed.
- Support for Wildcards
Implementation (1/2)

Open Source Applications for Annotation and Retrieval of Digital Photos: Caliph & Emir

Caliph: Common and Light Weight Photo Annotation
Implementation (2/2)

Emir: **Experimental Metadata Based Image Retrieval**

- Similar Images (CBIR)
- Keywords in Full Text Index
- Semantic Descriptions
- Visualization based on:
  - CBIR (Color, Edges)
  - Similarity of Semantic Graphs
Evaluation

Definition of a Test Set

- No Standardized Test Set available

Evaluation of Retrieval Performance compared to

- Full Text Search
- Different Ranking (Scoring) Algorithms

Assumption: Maximum Common Subgraph Metric yields Optimal Results

\[
similarity(G_1,G_2) = \frac{|mcs(G_1,G_2)|}{\max(|G_1|,|G_2|)}
\]
Test Set

- 85 different Semantic Descriptions
- Photos of I-Know Conferences in 2002 and 2004
- Each Graph from the database was taken to query the whole database, precision and recall were averaged

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodes</td>
<td>3</td>
<td>11</td>
<td>5.5</td>
</tr>
<tr>
<td>Relations</td>
<td>2</td>
<td>12</td>
<td>5.6</td>
</tr>
</tbody>
</table>
Auswertung

![Graph showing precision and recall for different index types: Full text index, 2-Path Index TF, 2-Path Index TF-IDF, and 2-Path Index Lucene Scorer. The graph compares the performance of these index types across varying recall values.]
Lucene Scoring Function

\[
\text{score}(q,d) = \sum_{t \in q} \text{TF}(t,d) \cdot \text{IDF}(t) \cdot b(t.\text{field},d) \cdot \text{lNorm}(t.\text{field},d) \cdot \text{coord}(q,d) \cdot \text{qNorm}(q)
\]

Lucene Scorer:
- TF ... Term Frequency
- IDF ... Inverse Document Frequency
- b ... Boost Value
- lNorm ... Normalization based on Field Value Length
- qNorm ... Normalization based on Query
- coord ... Term Frequency in Query and Document
Evaluation Results

The Path Index based Retrieval outperforms the Full Text Retrieval on this test case.

Between classical TF*IDF implementation and the term frequency scoring function only slight differences in retrieval performance can be identified.

We assume that the coord($q, d$) factor is the reason for the different performance of the classical TF*IDF and the Lucene score function by reflecting the denominator of the maximum common distance metric.
Demonstration

Emir: Experimental Metadata based Image Retrieval
Future Work

- Creating an appropriate Test Set
- Advancing Clustering and MDS Algorithms
- Integration of different Similarity Metrics for Evaluation
  - Path based ST Model
  - Error Correcting Maximum Common Subgraph Metric
  - Different Path Lengths, Selection with TF*IDF
- Implementation for RDF Graphs / OWL
Thank you …

... for your attention!

Visit Caliph & Emir:
http://caliph-emir.sourceforge.net