

# Semantic and Visual Image Clustering

## Retrieving Search Term Related Pictures in Structured Clusters

Seminar paper

SEMANTIC MULTIMEDIA

Summer Term 2013

Hasso-Plattner-Institut für Softwaresystemtechnik GmbH

Universität Potsdam

written by

Mandy Roick  
Claudia Exeler  
Tino Junge  
Nicolas Fricke

30. August 2013

## **Abstract**

Abstract goes here.

Write it at the end.

# Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Retrieving Images in Clusters</b>                  | <b>3</b>  |
| 1.1      | Problem Statement . . . . .                           | 3         |
| 1.2      | Motivation . . . . .                                  | 3         |
| <b>2</b> | <b>Related Work</b>                                   | <b>4</b>  |
| 2.1      | Semantics . . . . .                                   | 4         |
| 2.2      | Visuals . . . . .                                     | 4         |
| 2.3      | Approaches to Combine Semantics and Visuals . . . . . | 4         |
| <b>3</b> | <b>Image Tree Based on Wordnet</b>                    | <b>5</b>  |
| 3.1      | Wordnet . . . . .                                     | 5         |
| 3.2      | Constructing a Searchtree . . . . .                   | 5         |
| 3.3      | Synset Detection . . . . .                            | 5         |
| 3.4      | Assigning Synsets to Pictures . . . . .               | 5         |
| <b>4</b> | <b>Context-related and Visual Clustering</b>          | <b>6</b>  |
| 4.1      | General Approach . . . . .                            | 6         |
| 4.2      | Keyword Clusters . . . . .                            | 6         |
| 4.2.1    | Keyword Clustering . . . . .                          | 6         |
| 4.2.2    | Assigning Images to Keyword Clusters . . . . .        | 6         |
| 4.3      | Visual Clusters . . . . .                             | 6         |
| 4.3.1    | Features . . . . .                                    | 6         |
| 4.3.2    | Clustering . . . . .                                  | 6         |
| <b>5</b> | <b>Evaluation</b>                                     | <b>7</b>  |
| 5.1      | Method . . . . .                                      | 7         |
| 5.2      | Results . . . . .                                     | 7         |
| <b>6</b> | <b>Results Discussion</b>                             | <b>8</b>  |
| 6.1      | MCL-based Clusters . . . . .                          | 8         |
| <b>7</b> | <b>Future Work</b>                                    | <b>9</b>  |
| <b>A</b> | <b>Glossary</b>                                       | <b>10</b> |
| <b>B</b> | <b>Abbreviations and Acronyms</b>                     | <b>11</b> |
|          | <b>References</b>                                     | <b>12</b> |

# **1 Retrieving Images in Clusters**

## **1.1 Problem Statement**

What do we do?

challenge: cluster 1M pictures of the MIR1M flickr file set

Why do we do that?

## **1.2 Motivation**

## **2 Related Work**

Im vorangegangenen Kapitel hatten wir bereits die Gliederung einer wissenschaftlichen Arbeit im Fachgebiet der Informatik kurz vorgestellt und erläutert, welche inhaltlichen Punkte in der „Einleitung“ behandelt werden sollten. Die folgenden Abschnitte skizzieren inhaltlich die übrigen der bereits genannten Gliederungspunkte.

### **2.1 Semantics**

### **2.2 Visuals**

### **2.3 Approaches to Combine Semantics and Visuals**

## **3 Image Tree Based on Wordnet**

### **3.1 Wordnet**

### **3.2 Constructing a Searchtree**

two typical types use hyponym relation to span tree of specializations  
if no hyponyms (usually the case for geographic terms), use part-meronyms

### **3.3 Synset Detection**

for each tag and word in title, try to find synset (limiting ourselves to nouns)  
problem: multiple possible synsets for a word, how to find correct meaning?  
use best-first search with limited queue

### **3.4 Assigning Synsets to Pictures**

for higher recall: find strongly co-occurring tags that could not be mapped to synset  
take all pictures that are annotated with at least

## **4 Context-related and Visual Clustering**

Why do we do subclusters for nodes?

### **4.1 General Approach**

### **4.2 Keyword Clusters**

#### **4.2.1 Keyword Clustering**

MANDY

#### **4.2.2 Assigning Images to Keyword Clusters**

### **4.3 Visual Clusters**

#### **4.3.1 Features**

color histogram, edge length and angle histogram

#### **4.3.2 Clustering**

k-means, late fusion by intersecting

## 5 Evaluation

### 5.1 Method

Picture annotation tool, 200 pictures

compare two pictures: for each, does it display food? semantically not similar / same object / same object and same context? visually similar / not similar?

receive testset from users

how are quality measures calculated?

search food: precision / recall of picture inclusion (compare synset detection mechanisms?)

evaluate tree nodes based on same object annotations

evaluate mcl clusters based on same object and on same context annotations (compare both, what does mcl actually do?)

evaluate visuals with large minimal node size

vary parameters given by frontend

### 5.2 Results



## **6 Results Discussion**

Are our results good? Are they biased by something?

All depends on annotations - inappropriate tagging leads to bad results, as well as limitation to nouns (adjectives, adverbs and verbs are wrongly matched nouns)

### **6.1 MCL-based Clusters**

highly depend on quality of keyword clusters

## 7 Future Work

How to improve, what other approaches to take

## A Glossary

**Synset:** Ein spezielles Programm, mit dem man über das WWW Zugang zu WWW-Servern erlangen und von diesem angeforderte Dokumente anzeigen kann.

**Wordnet:** Bezeichnet ein Programm, dass einen Server kontaktiert und von diesem Informationen anfordert. Der im WWW eingesetzte Browser ist in diesem Sinne ein Client. Aber es gibt auch andere Clients im WWW, die WWW-Server kontaktieren und Informationen von diesen herunterladen, wie z.B. Suchmaschinen oder Agenten.

**HTML:** Hypertext Markup Language; das einheitliche Dokumentenformat für Hypermedia-Dokumente im WWW. Dokumente, die im WWW übertragen und vom Browser dargestellt werden sollen, sind in HTML kodiert.

**HTTP:** Hypertext Transfer Protocol; das Protokoll, das die Kommunikation von Browsern und WWW-Servern im WWW regelt. Fordert ein Browser ein Dokument vom WWW-Server an oder beantwortet der WWW-Server eine Anfrage, muss diese Anfrage den Konventionen des HTTP-Protokolls gehorchen.

**Netzanwendung:** Ein Anwendungsprogramm, dessen Ablauf den Zugriff auf Ressourcen einschließt, die nicht lokal auf dem ausführenden Rechner liegen, sondern auf einem entfernten Rechner über das Netzwerk zugegriffen werden.

**Server:** Bezeichnet einen Prozess, der von Clients kontaktiert wird, um diesen Informationen zurück zu liefern. Oft wird auch der Rechner, auf dem ein Server-Prozess abläuft, als Server bezeichnet.

## B Abbreviations and Acronyms

|      |                                       |
|------|---------------------------------------|
| 4CIF | 4 fach Common Intermediate Format     |
| AAC  | Advanced Audio Coding                 |
| AAL  | ATM Adaption Layer                    |
| ABR  | Available Bit Rate                    |
| AC   | Audio Code                            |
| ACK  | Acknowledgement                       |
| ADM  | Add Drop Multiplexer                  |
| ADSL | Asymmetric Digital Subscriber Line    |
| AH   | Authentication Header                 |
| AIFF | Audio Interchange File Format         |
| AM   | Amplituden-Modulation                 |
| ANSI | American National Standards Institute |
| API  | Application Programming Interface     |
| ARP  | Address Resolution Protocol           |
| W3C  | World Wide Web Community              |
| WWW  | World Wide Web                        |

## References