



Suche nach Bildern und Videos mithilfe von Ontologien

Rainer Marquardt
Software- & Internettechnologie
6. Semester
Universität Mannheim

Übersicht



- 1 Motivation
- 2 Möglichkeiten nach Bildern und Videos zu suchen
- 3 Aufbau einer Ontologie zur Suche
- 4 Experiment: Vergleiche der Bildsuche mit Google & Co
- 5 Schlussfolgerung

1. Motivation



- 2004 wurden mehr Informationen als in der bisherigen Menschheitsgeschichte produziert
- Das waren im Durchschnitt 800 MB pro Mensch
- Das Erdbeobachtungssystem der NASA kann täglich ein Terrabyte an Bildmaterial produzieren
- Der Anteil an Multimediadaten wie Fotos und Videos hat dabei einen immer größeren Anteil (z.B. aufgrund von Mobiltelefonen)
- Die Suche in dieser Datenflut wird also immer wichtiger



2.1 Suche ohne Metadaten

(vereinfachtes) Bild als Anfrage



Zu durchsuchende Bestände



Suche

0xFFFFFFFF	0xFFFFFFFF	0xFFFFFFFF
0xFFFFFFFF	0x000000	0x123123
0x222212	0x2123AD	0xFF55FF
0xAAAAAA	0x324234



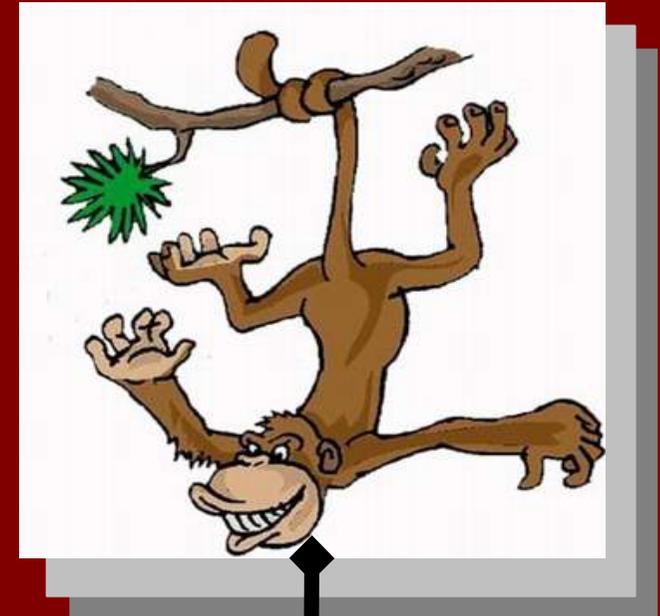
0xFFFFFFFF	0xFFFFFFFF	0xFFFFFFFF
0xFFFFFFFF	0x000000	0x123123
0x222245	0x2123AD	0xFF66FF
0xAAAAAA	0x324234

2.2 Suche mit Metadaten



[Web](#) [Bilder](#) [Groups](#) [Verzeichnis](#) [News](#) [Froogle](#) [Mehr »](#)

Suche



Annotation

```
SELECT `picture` FROM `annotations`  
WHERE `subject` LIKE '%affe%'
```

Subject = „Affe“,
Type = „Clipart“

2.2.1 Keywords als Metadaten



Bilder werden mit einfachen, den Inhalt beschreibenden Worten versehen. Die Strukturen sind meist flach.

Bsp.: „ein großer Affe sitzt neben einem kleinen Baum“

Mit welchen Keywords kann genau dies ausgedrückt werden?

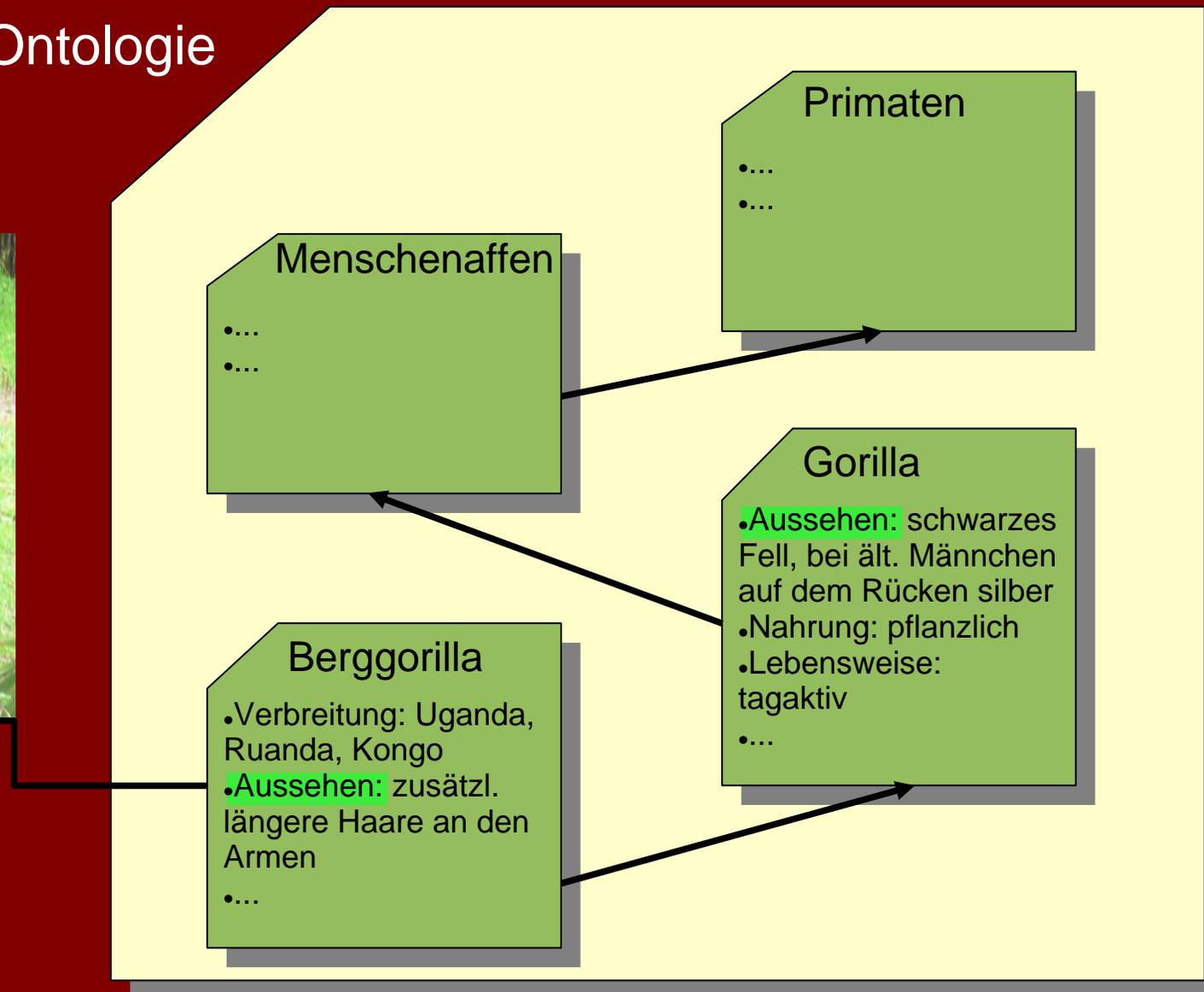
- „Affe“, „Baum“?
- „Affe“, „Baum“, „klein“, „groß“, „sitzen“?
- „großer Affe“, „sitzen“, „kleiner Baum“?



2.2.2 Ontologien



Beispiel für eine Ontologie



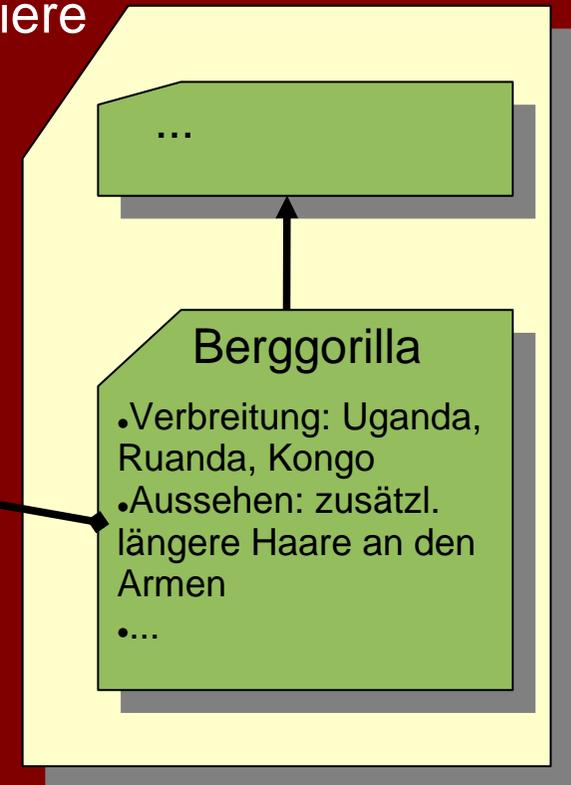
3.0 Aufbau einer Annotation



Annotation auf Basis einer Ontologie

- ...
- ...
- Bildinhalt

(Existierende) Ontologie mit Hintergrundwissen über Tiere



3.1 Aufbau einer Annotation



Was kann über ein (digitales) Bild ausgesagt werden?



Name:	monkey2.jpg
Typ:	JPEG-Bild
Größe:	43,5 KB (44 577)
Geändert:	22.05. 16:48
Eigentümer:	rainer - users
Rechte:	-rw-r--r--
<hr/>	
Abmessungen:	640 x 480 Pixel
Farbmodus:	Farbe
Blitz benutzt?:	Nein
JPEG-Prozess:	Baseline

• **Inhaltliche Betrachtungsgegenstände** (*subject matter*) z.B.: ein Affe im Wald auf einem Ast

• **Eigenschaften des digitalen Mediums** (*medium feature*) z.B.: Format, Auflösung, Farben

• **Umstände unter denen das Foto gemacht wurde** (*photo feature*) z.B.: Zeit, Ort, Perspektive



3.2 Aufbau einer Annotation



Die Struktur von *subject matter*:



- *agent & agent-modifier* z.B. „Affe“ mit „silber braunem Fell“
- *action* z.B. „essen“
- *object* z.B. „Banane“
- *setting* z.B. „steiniger Untergrund“

3.3 Annotieren

- Auswahl von Werten aus vorgegebenen Wertebereichen (*ranges*)

Der Wertebereich für *agent* & *agent-modifiers* wird von der **Ontologie für Hintergrundwissen** vorgegeben

The screenshot shows an 'Ontology browser' window. On the left is a tree view of the ontology hierarchy. On the right is a table of properties. A yellow box highlights the 'agent' and 'agent-modifiers' properties in the table, and a line points from the text above to this box.

Class	Passive agent	
Meta Class	Class	
Properties		
Name	Class	Range
setting	Subject matter description	Setting description
agent_modifiers	Subject matter description	Animal characteristics
agent	Subject matter description	Species

3.4 Suchen



- Eingabe von Werten für Bilder mit den gesuchten Eigenschaften
- Gesucht werden Annotation (bzw. Bilder) mit gleichen oder spezielleren Eigenschaften



3.5 Annotation & Suche bei Videos



- Ein Video ist eine Abfolge von Bildern, oder besser kurzen Sequenzen mit jeweils der selben Kameraeinstellung (Shots)
- Zur Annotation müssen also alle, oder die wichtigsten Shots annotiert werden
- Hierzu wird aus jedem Shot ein Keyframe ermittelt und dessen Inhalt annotiert
- Die Suche verläuft wieder ähnlich: zunächst werden zu suchende Werte angegeben und anhand dieser werden Shots ermittelt



4.0 Vergleiche der Bildsuche



In einem Experiment wurden Bilder von Affen auf die beschriebene Weise annotiert und dann durchsucht. Die erste Suche galt „großen Affen“ („great ape“)

Vergleich mit bekannten Suchmaschinen mit automatisch generierten Suchindizes

	Angelegte Ontologie	Alta Vista <great + ape> 2001	Alta Vista <great ape> 2001	Google <„great ape“> 2006	Alta Vista <„great ape“> 2006
Recall / Treffer	100%	13	45000	1900	648
Präzision	100%	50%	32%	32%	60%

Vergleich mit der Suchmaschine gettyone.com, die manuell, auf Basis von Keywords annotierte Bildkollektionen durchsucht

	Angelegte Ontologie	gettyone.com <great ape> 2001	gettyone.com <ape> 2001	gettyone.com <great ape> 2006	gettyone.com <ape> 2006
Recall / Treffer	100%	0	521	2	1754
Präzision	100%	-	85%	100%	94%

4.0 Vergleiche der Bildsuche



Ein weiterer Vergleich wurde mit einer komplexeren Suchanfrage durchgeführt: „suche einen Schimpanse der sich mit der Hand am Kopf kratzt“

Mit der erzeugten Annotation ist eine solche Abfrage leicht möglich (siehe auch 3.3), mit Keywords ist die Anfrage aber schwer zu formulieren

	Ontologie	gettyone.com <chimpanzee scratching> 2006	gettyone.com <chimpanzee AND hand AND head> 2006
Recall / Treffer	100%	11	4
Präzision im Vgl. zur Anfrage („irgendwo kratzen“, „hand irgendwie am kopf“)	100%	8 (73%)	2 (50%)
Präzision im Vgl. zu „mit der Hand am Kopf kratzen“	100%	6 (55%)	0 (0%)

5.0 Schlussfolgerung



- Manuelle Annotationen haben bei der Suche eine deutlich höhere Präzision als automatisch erzeugte
- Ontologien sind ein geeignetes Mittel Hintergrundwissen und speziellere Eigenschaften *wie die action eines agent in eine Annotation einfließen zu lassen (Klassifizierung, Beziehungen, Wertebereiche)*
- *Ontologien können (halb-) automatische Verfahren unterstützen*
- Manuelle Annotation sind aufwendig (kostspielig)
- Volle Nutzung der Möglichkeiten wohl nur im professionellen/kommerziellen Bereich
- Zur Verlinkung mit Hintergrundwissen muss dieses in geeigneter Form vorhanden sein
- Ein ontologie-basiertes Suchtool müsste die typische Erwartung „gesucht wird über ein Textfeld“ befrieden



Bildquellen:

- Abb. 1: www.vskrems-lerchenfeld.ac.at
- Abb. 1: www.game-reserve.com
- Abb. 2: www.freenaturephotos.com
- Abb. 3: www.magiccarpetrides.com
- Abb. 4: de.wikipedia.org
- Abb. 4: whalersdiveclub.org
- Abb. 4,5: A. Th. Schreiber, B. Dubbeldam, J. Wielemaker: *Ontology-Based Photo Annotation*, University of Amsterdam, 2001
- Abb. 6: Stephan Kopf, Thomas Haenselmann, Dirk Farin, Wolfgang Effelsberg: *Automatic Generation of Summaries for the Web*, University of Mannheim, 2004