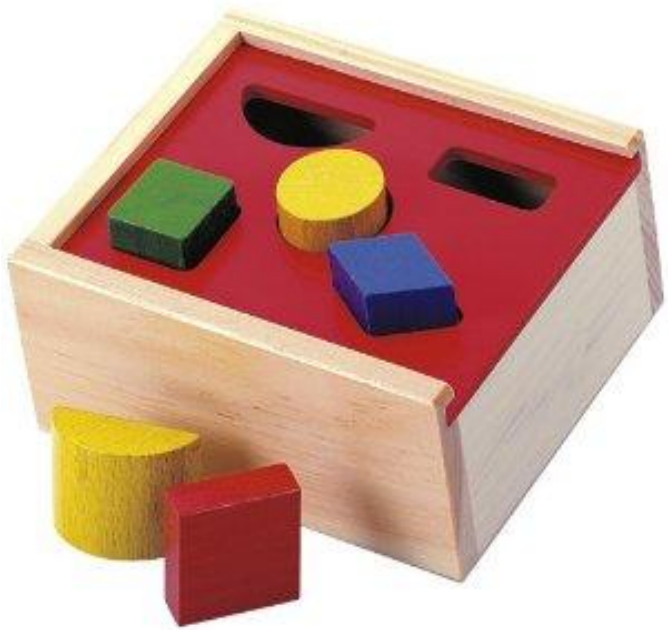
Similarity

What is a cube?

What is a cylinder?

How can you measure the similarity?



Computational Linguistics

Words and Images

Stefan Thater & Dietrich Klakow

FR 4.7 Allgemeine Linguistik (Computerlinguistik)

Universität des Saarlandes

Summer 2013

Example

Look at two examples

- Retrieving images using text queries
- Generating captions for images

Image Retrieval using Query Words

Tiger - Google Search - Mozilla Firefox

Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

Tiger - Google Search

www.google.de/search?num=10&hl=en&site=imghp&tbm=isch&source=hp&q=Tiger&oq=Tiger&gs_l=img.3..0l10.2192.2874.0.3056.5.4.0.1.1.0.99.343.4.4.0...0.0.1JbPE

Google

Meistbesuchte Seiten Erste Schritte Aktuelle Nachrichten Kostenlose Hotmail Links anpassen Windows Media Windows

+You Search Images Maps Play YouTube Gmail Documents Calendar Translate More

Google

Tiger

Sign in

Search

About 761,000,000 results (0.17 seconds)

SafeSearch moderate

Web
Images
Maps
Videos
News
Shopping
More

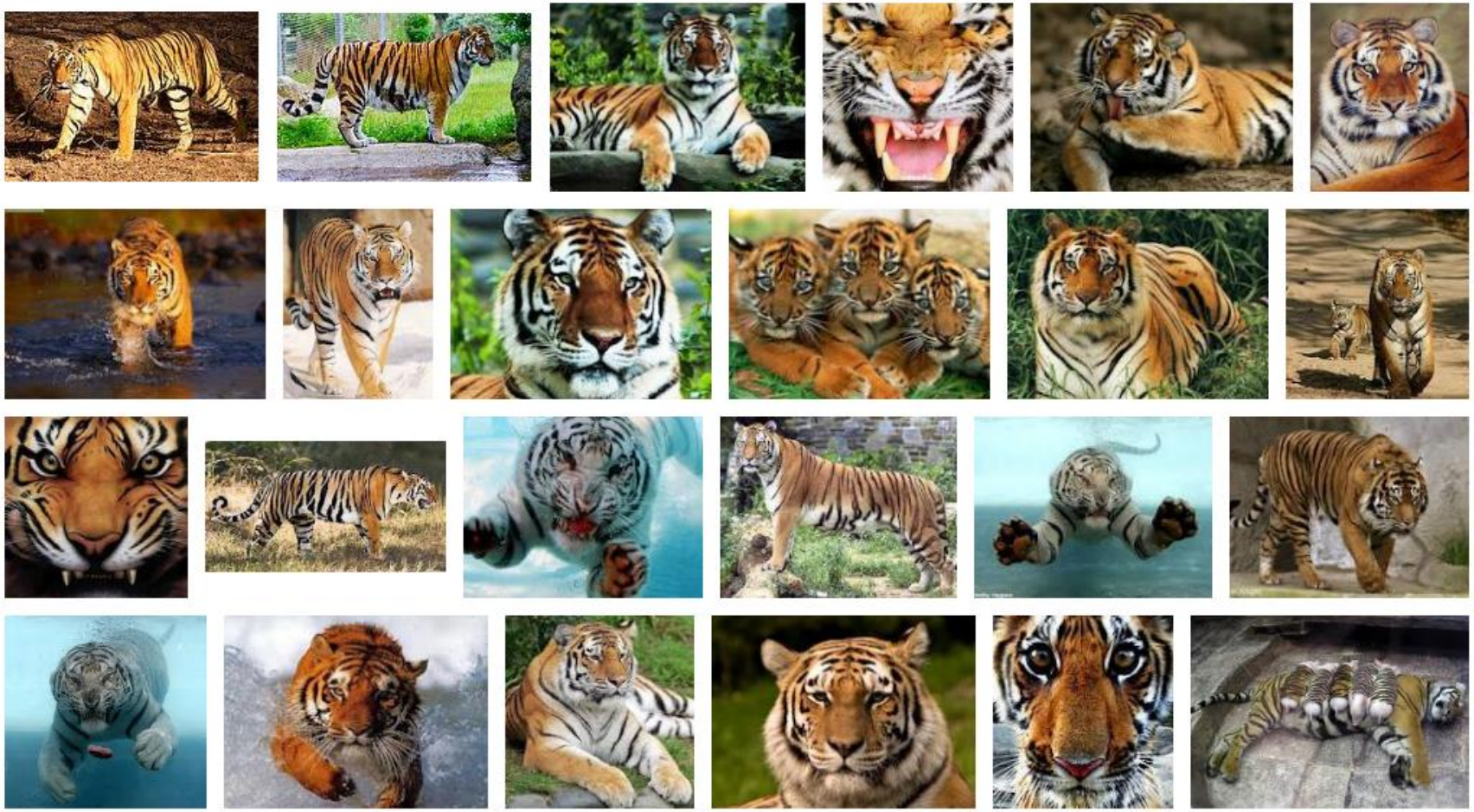
Any time
Past 24 hours
Past week
Custom range...

All results
By subject

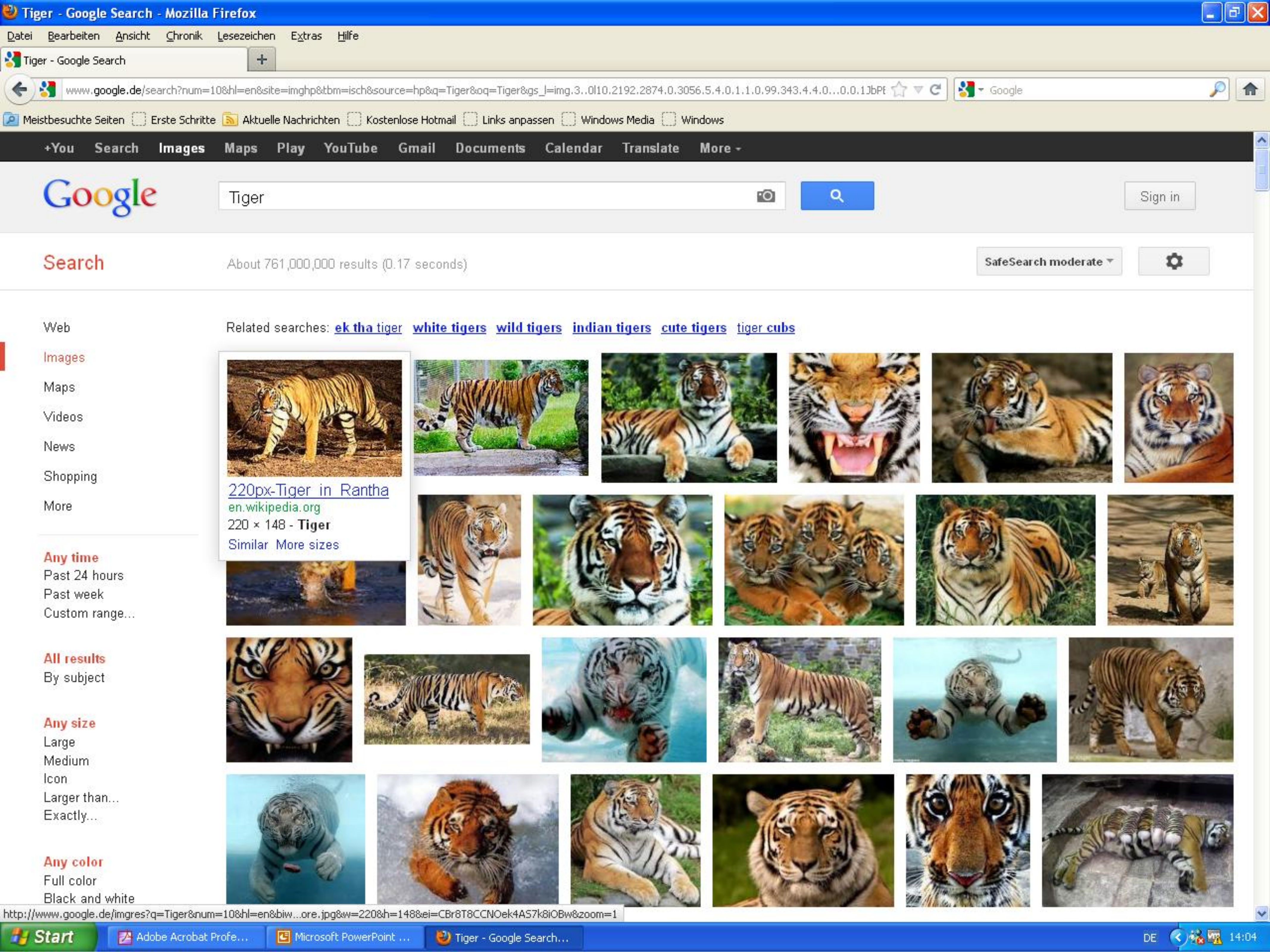
Any size
Large
Medium
Icon
Larger than...
Exactly...

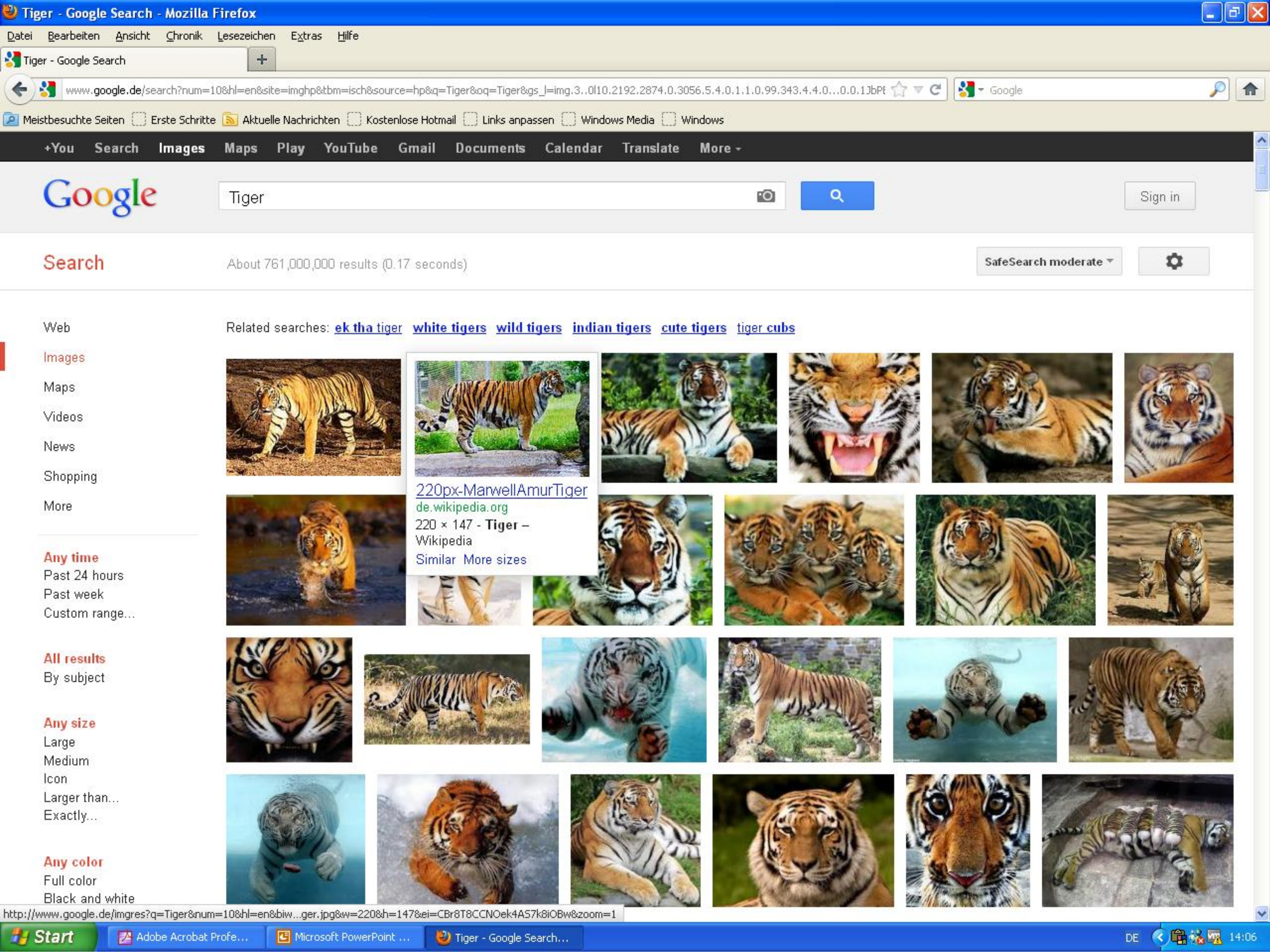
Any color
Full color
Black and white

Related searches: ek tha tiger white tigers wild tigers indian tigers cute tigers tiger cubs



Start Adobe Acrobat Profe... Microsoft PowerPoint ... Tiger - Google Search... DE 14:05


Tiger - Google Search - Mozilla Firefox
Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe
Tiger - Google Search
www.google.de/search?num=10&hl=en&site=imghp&tbm=isch&source=hp&q=Tiger&oq=Tiger&gs_l=img.3..0l10.2192.2874.0.3056.5.4.0.1.1.0.99.343.4.4.0...0.0.1JbPt
Google
Meistbesuchte Seiten
Erste Schritte
Aktuelle Nachrichten
Kostenlose Hotmail
Links anpassen
Windows Media
Windows
+You Search Images Maps Play YouTube Gmail Documents Calendar Translate More
Tiger
Sign in
Search
About 761,000,000 results (0.17 seconds)
SafeSearch moderate
Web
Images
Maps
Videos
News
Shopping
More
Any time
Past 24 hours
Past week
Custom range...
All results
By subject
Any size
Large
Medium
Icon
Larger than...
Exactly...
Any color
Full color
Black and white
Related searches: ek tha tiger white tigers wild tigers indian tigers cute tigers tiger cubs
220px-Tiger_in_Rantha
en.wikipedia.org
220 × 148 - Tiger
Similar More sizes
http://www.google.de/imgres?q=Tiger&num=10&hl=en&biw...ore.jpg&w=220&h=148&ei=CBr8T8CCNOek4AS7k8iOBw&zoom=1
Start
Adobe Acrobat Profe...
Microsoft PowerPoint ...
Tiger - Google Search...
DE
14:04


Tiger - Google Search - Mozilla Firefox
Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe
Tiger - Google Search
www.google.de/search?num=10&hl=en&site=imghp&tbm=isch&source=hp&q=Tiger&oq=Tiger&gs_l=img.3..0l10.2192.2874.0.3056.5.4.0.1.1.0.99.343.4.4.0...0.0.1JbPl
Google
Meistbesuchte Seiten Erste Schritte Aktuelle Nachrichten Kostenlose Hotmail Links anpassen Windows Media Windows
+You Search Images Maps Play YouTube Gmail Documents Calendar Translate More
Google
Tiger
Sign in
Search
About 761,000,000 results (0.17 seconds)
SafeSearch moderate
Web
Images
Maps
Videos
News
Shopping
More
Any time
Past 24 hours
Past week
Custom range...
All results
By subject
Any size
Large
Medium
Icon
Larger than...
Exactly...
Any color
Full color
Black and white
Related searches: ek tha tiger white tigers wild tigers indian tigers cute tigers tiger cubs
220px-MarwellAmurTiger
de.wikipedia.org
220 × 147 - Tiger – Wikipedia
Similar More sizes
http://www.google.de/imgres?q=Tiger&num=10&hl=en&biw...ger.jpg&w=220&h=147&ei=CBr8T8CCNOek4AS7k8iOBw&zoom=1
Start
Adobe Acrobat Profe...
Microsoft PowerPoint ...
Tiger - Google Search...
DE
14:06

Tiger - Google Search - Mozilla Firefox

Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

Tiger - Google Search

www.google.de/search?num=10&hl=en&site=imghp&tbm=isch&source=hp&q=Tiger&oq=Tiger&gs_l=img.3..0l10.2192.2874.0.3056.5.4.0.1.1.0.99.343.4.4.0...0.0.1JbPE

Meistbesuchte Seiten Erste Schritte Aktuelle Nachrichten Kostenlose Hotmail Links anpassen Windows Media Windows

+You Search Images Maps Play YouTube Gmail Documents Calendar Translate More

Google

Tiger

Sign in

Search

About 761,000,000 results (0.17 seconds)

SafeSearch moderate

Web
Images
Maps
Videos
News
Shopping
More

Any time
Past 24 hours
Past week
Custom range...

All results
By subject

Any size
Large
Medium
Icon
Larger than...
Exactly...

Any color
Full color
Black and white

Related searches: ek tha tiger white tigers wild tig

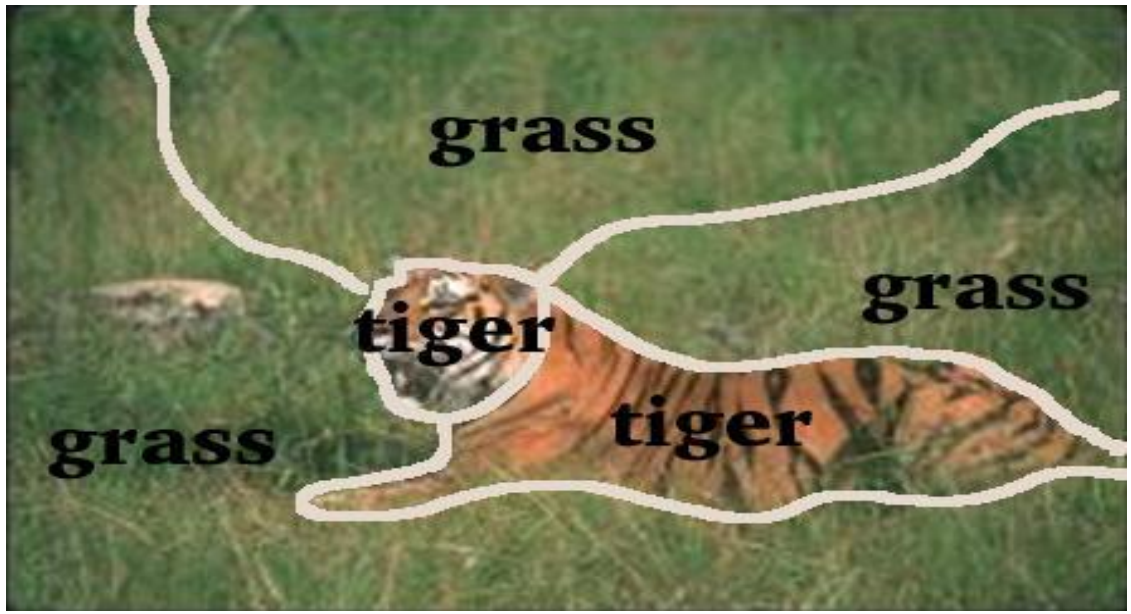
tiger-regal.jpg
fohn.net
690 × 461 - A Siberian **tiger** in all his majesty--and doesn't he know it!
Similar More sizes

How to retrieve images
that are not called *tiger*.jpg?

http://www.google.de/imgres?q=Tiger&num=10&hl=en&biw...gal.jpg&w=690&h=461&ei=CBr8T8CCNOek4AS7k8iOBw&zoom=1

Start Adobe Acrobat Profe... Microsoft PowerPoint ... Tiger - Google Search... DE 14:07

Are Tigers striped?



How important is it to detect “tiger color” and “tiger texture” together?

How important is it to detect “tiger color” and in the middle of the image?

- Image from [Barnard2003]
- 5000 images plus captions

Correlation concept/region is stronger than color/texture!

Overview

- Measuring regional mutual information
- Retrieval and annotation model:

  Regional language models
- Experiments

The TRECVID 2003 Data

Training : 9413 shots

Validation: 4085 shots

Test : 4787 shots

Num concepts = 75

On average 5 concepts per shot

Features extracted from grids

5x7 = 35 grids per image

Horse, outdoors, snow

- Feature extraction:

  Use existing tools (e.g. SIFT features using http://www.vlfeat.org/~vedaldi/code/sift.html)

- Features are quantized using k-Means (visterms)

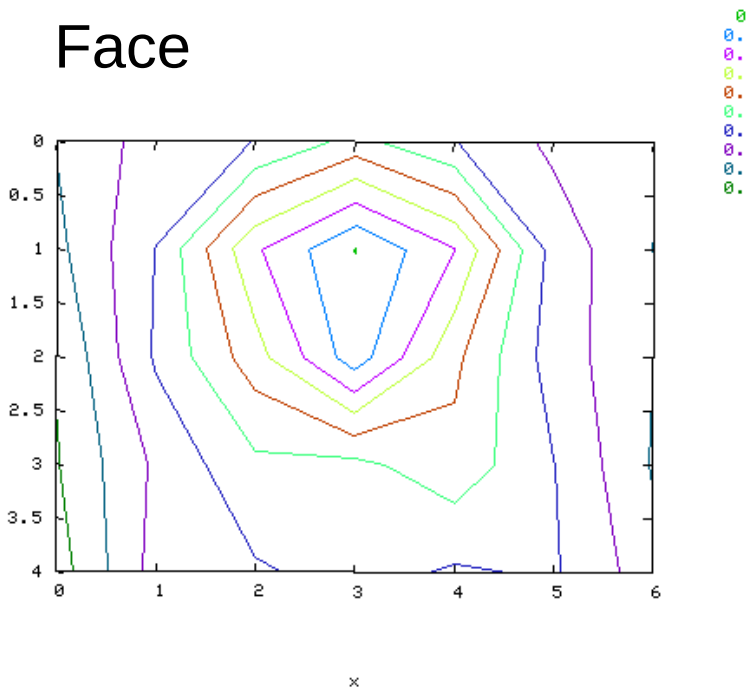
Regional Mutual Information

$$MI(r,c) = \sum_{v} N(v,r,c) \log\left(\frac{N(v,r,c)}{N(v,r)N(c)} \right)$$

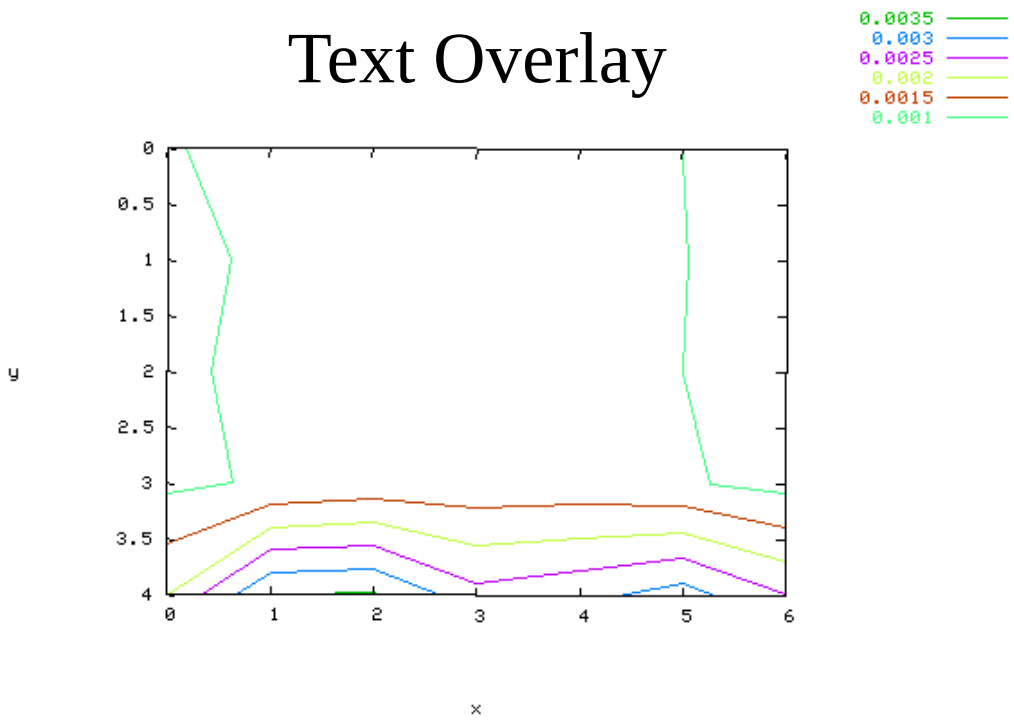
r: label of region (5x7 grid on the image)
c: concept label (e.g. face, vehicle, text_overlay)
v: visterm (discrete representation of color, texture and edge information)

Regional Mutual Information for two Concepts

Face



Text Overlay



Same visterm in different regions can trigger different concepts

Regional Bayes Classifier

Bayes classifier $\bar{c} = \arg\max_{c} P(v_1 \ldots v_R \mid c) P(c)$

Independence assumption (Naïve Bayes)

$$P(v_1 \ldots v_R \mid c) = \prod_{r=1}^{R} P(v_r \mid c)$$

One model per region

$$P(v_1 \ldots v_R \mid c) = \prod_{r=1}^{R} P_r(v_r \mid c)$$

Absolute Discounting

$$P_r(v \mid c) = \max\left(\frac{N(v,r,c) - d}{N(r,c)}, 0 \right) + \frac{d\ B}{N(r,c)} P_{BG}(v \mid c)$$

Most widely used in speech recognition

d: discounting parameter

$P_{BG}(v|c)$: backing-off distribution

B: number of v for which N(v,r,c)>d

Smoothing Methods

Linear Interpolation

$$P_r(v \mid c) = (1-\lambda)\frac{N(v,r,c)}{N(r,c)} + \lambda P_{BG}(v \mid c)$$

λ: interpolation weight

Also known as Jelinek-Mercer-Smoothing

Dirichlet Prior

$$P_r(v \mid c) = \frac{N(v,r,c) + \mu P_{BG}(v \mid c)}{N(r,c) + \mu}$$

μ: weight for prior

Background Distributions

Uniform Distribution ("zerogram")

$$P_{BG}^{Zero}(v \mid c) = \frac{1}{|V|}$$

|V|: number of visterms)

Frequency of visterms ("unigram")

$$P_{BG}^{Unigram}(v \mid c) = \frac{N(v)}{\sum_{v} N(v)}$$

Results
(mean average precision)

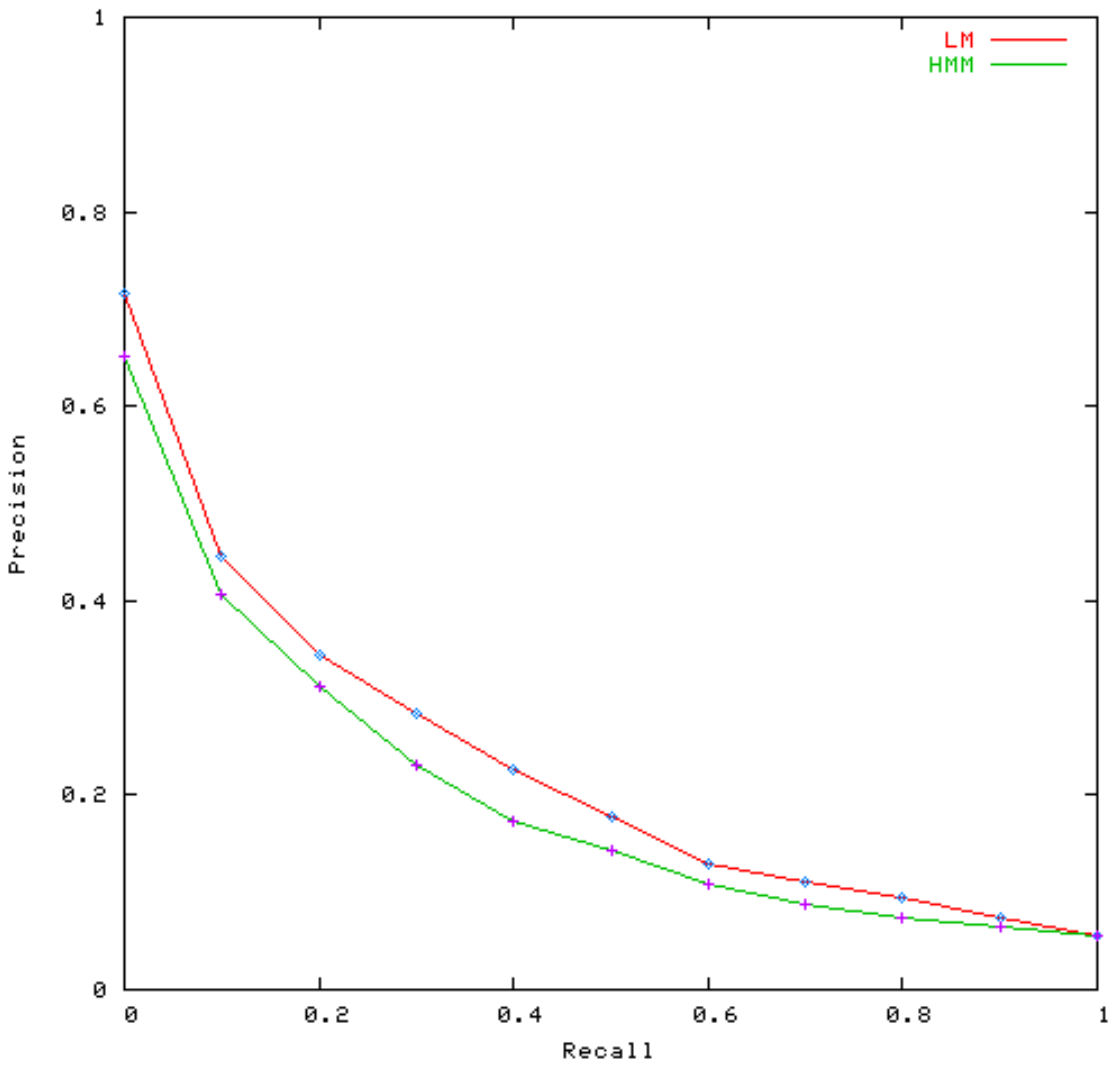
	Baseline	
Model	**BG Zero**	**BG Unigram**
Absolute Discounting	0.149	0.149
Dirichlet Prior	0.149	0.150
Linear Interpolation	0.145	0.148

Results
(mean average precision)

	Baseline		+ regional modeling	
Model	**BG Zero**	**BG Unigram**	**BG Zero**	**BG Unigram**
Absolute Discounting	0.149	0.149	0.209	0.215
Dirichlet Prior	0.149	0.150	0.207	0.218
Linear Interpolation	0.145	0.148	0.215	**0.221**

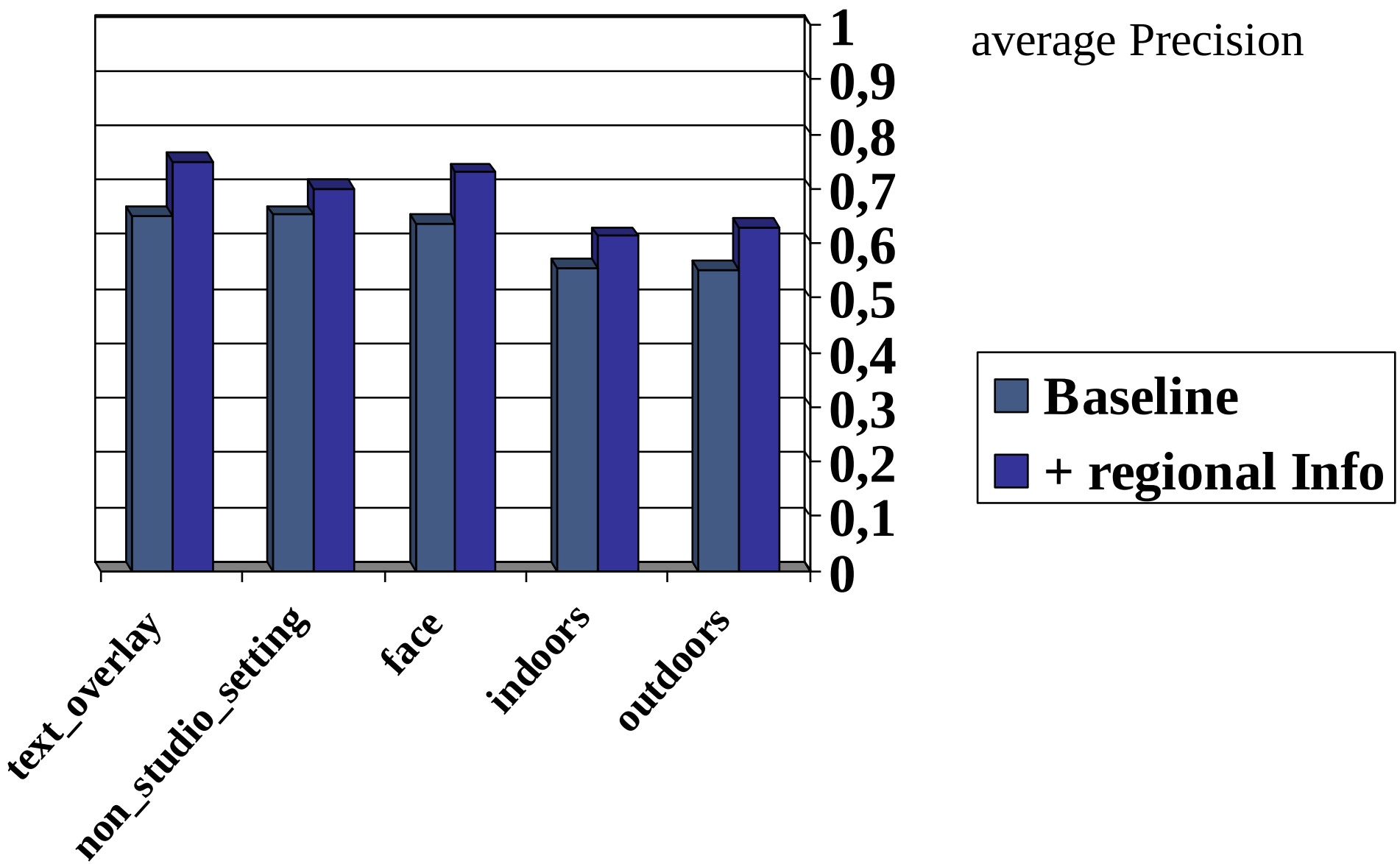
$\mapsto$ regional information gives significant improvement (32%)

Precision-Recall Graph



On TREC-VID 03 Consistent improvement over HMM model by Ghoshal et al. (SigIR 2005)

Improvement on most frequent concepts



Retrieved Images: selected examples

Weather News

1 2 3 4 5

Football

Horse

Summary

- Analysis by regional mutual information
- Region specific modeling is important
- Best variant uses
  - Linear interpolation
  - Backing off to background unigram
- Extremely simple method

Generating Captions

How Many Words Is a Picture Worth?
Automatic Caption Generation for News Images

Yansong Feng and **Mirella Lapata**
School of Informatics, University of Edinburgh
10 Crichton Street, Edinburgh EH8 9AB, UK
Y.Feng-4@sms.ed.ac.uk, mlap@inf.ed.ac.uk

Task

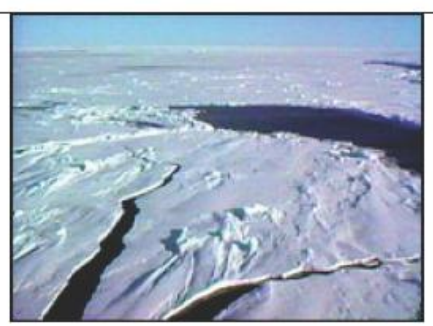
Thousands of Tongans have attended the funeral of King Taufa'ahau Tupou IV, who died last week at the age of 88. Representatives from 30 foreign countries watched as the king's coffin was carried by 1,000 men to the official royal burial ground.

King Tupou, who was 88, died a week ago.

Contaminated Cadbury's chocolate was the most likely cause of an outbreak of salmonella poisoning, the Health Protection Agency has said. About 36 out of a total of 56 cases of the illness reported between March and July could be linked to the product.

Cadbury will increase its contamination testing levels.

A Nasa satellite has documented startling changes in Arctic sea ice cover between 2004 and 2005. The extent of "perennial" ice declined by 14%, losing an area the size of Pakistan or Turkey. The last few decades have seen ice cover shrink by about 0.7% per year.



Satellite instruments can distinguish "old" Arctic ice from "new".

A third of children in the UK use blogs and social network websites but two thirds of parents do not even know what they are, a survey suggests. The children's charity NCH said there was "an alarming gap" in technological knowledge between generations.

Children were found to be far more internet-wise than parents.

Generate based on text and image

First Step

Annotate each image with key words

(e.g. using method from the previous section)

Second step: generate caption
Extractive approaches

Idea: pick a suitable sentence from the text to be the caption

Approach:

Measure similarity between sentence and key words describing image using

Word overlap

Cosine

KL-divergence

Second step: generate caption abstractive approaches

Idea: pick the most likely word sequence given the key words

Estimate N-gram based on

Surrounding text

Key words

Combine the two using LM adaptation techniques

Optional: include a phrase based constraint

Results

G:	King Tupou, who was 88, died a week ago.
KL:	Last year, thousands of Tongans took part in unprecedented demonstrations to demand greater democracy and public ownership of key national assets.
A_W:	King Toupou IV died at the age of Tongans last week.
A_P:	King Toupou IV died at the age of 88 last week.
G:	Cadbury will increase its contamination testing levels.
KL:	Contaminated Cadbury's chocolate was the most likely cause of an outbreak of salmonella poisoning, the Health Protection Agency has said.
A_W:	Purely dairy milk buttons Easter had agreed to work has caused.
A_P:	The 105g dairy milk buttons Easter egg affected by the recall.
G:	Satellite instruments can distinguish "old" Arctic ice from "new".
KL:	So a planet with less ice warms faster, potentially turning the projected impacts of global warming into reality sooner than anticipated.
A_W:	Dr less winds through ice cover all over long time when.
A_P:	The area of the Arctic covered in Arctic sea ice cover.
G:	Children were found to be far more internet-wise than parents.
KL:	That's where parents come in.
A_W:	The survey found a third of children are about mobile phones.
A_P:	The survey found a third of children in the driving seat.

Summary

Image retrieval using text queries

Caption generation