FAKULTÄT
FÜR INFORMATIK
Faculty of Informatics

Domain Specific IR
Lecture 4 of 5: Beyond Text

Allan Hanbury

hanbury@ifs.tuwien.ac.at

Russian Summer School on Information Retrieval

August 6-11, 2012   Yaroslavl, Russian Federation

Image retrieval

- The most obvious step beyond text retrieval is usually image retrieval, due to the ubiquity of images
- But general images can be similar in so many different ways...

Contents

- Introduction to image retrieval
- Domain-specific image retrieval
  - Image search in the medical world
  - Image search in patents
- Challenges and conclusion

Search results from LTU technologies: http://corbis.ltutech.com/

Retrieval of images

- Text retrieval of images
  - Is there any text attached to the images?
  - Doing this manually is expensive, subjective, language dependent, …
  - Take text close to the images (such as captions)
  - Semantic concepts could help in some cases
- Visual retrieval of images
  - Using automatically extracted visual features
  - Content-based image retrieval (CBIR)
  - Query by Image Example(s) (QBE)
- Multimodal retrieval (text+images)

RuSSIR 2012, August 6-11   Domain Specific IR / Hanbury / Lupu   5

Components of image retrieval systems
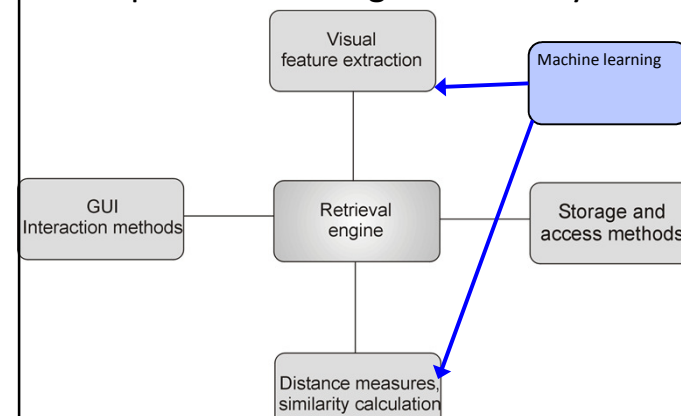
RuSSIR 2012, August 6-11   Domain Specific IR / Hanbury / Lupu   7

Visual information for retrieval

- Object detection
  - Then potentially mapping the objects to an ontology
  - Usually works well for a small number of objects
- Image classification
  - Training data, limited set of classes
  - Global classification of images vs. local classification of pixels, regions
- Similarity retrieval of images
  - Global image information, regions of interest (ROIs), small
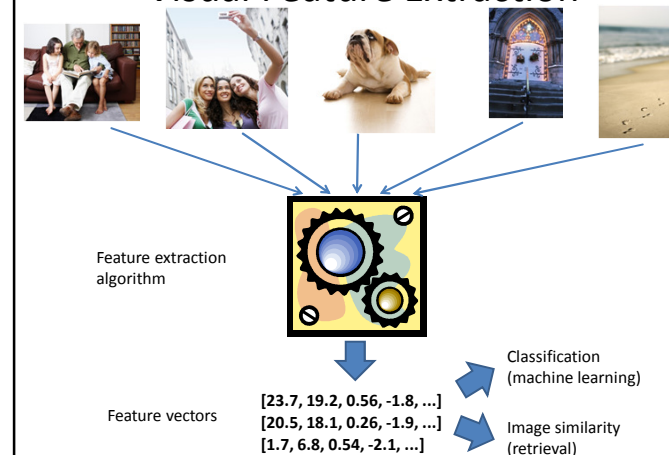 - No training data, relevance as criterion for quality
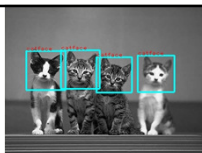
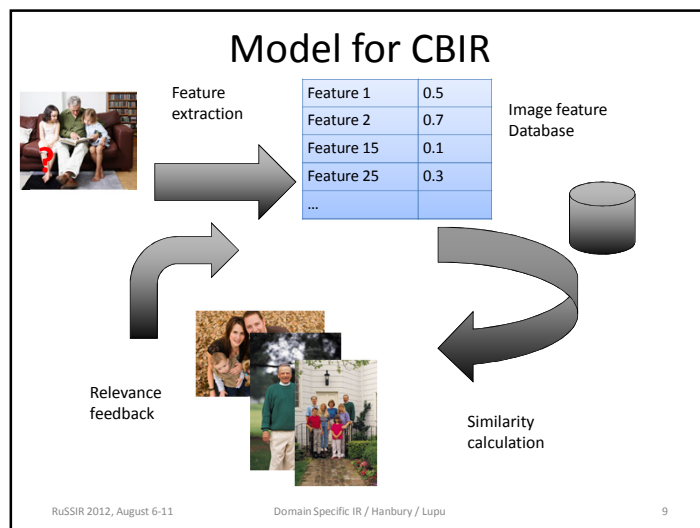RuSSIR 2012, August 6-11   Domain Specific IR / Hanbury / Lupu   6

Visual Feature Extraction

Feature extraction algorithm

Feature vectors

[23.7, 19.2, 0.56, -1.8, …]
[20.5, 18.1, 0.26, -1.9, …]
[1.7, 6.8, 0.54, -2.1, …]

Classification (machine learning)

Image similarity (retrieval)

RuSSIR 2012, August 6-11   Domain Specific IR / Hanbury / Lupu   8

Model for CBIR

Feature extraction

Feature 1	0.5
Feature 2	0.7
Feature 15	0.1
Feature 25	0.3
...	

Image feature Database

Relevance feedback

Similarity calculation

Types of visual feature

- Image thumbnails
- Colors
  - Different color spaces, human perception
- Textures
  - Many different models, no clear definition
- Shapes
  - Automatic segmentation is an ill defined problem
- Visual words
- Invariance of the extracted features
  - Rotation, shift, size, illumination, …

Steps of a retrieval system

- Image pre-processing
  - Normalization, background removal, …
- Salient region or point detection
- Visual feature extraction
  - Feature selection
  - Then feature modeling in the case of visual words
- Distance calculation or similarity measurement
- Results fusion (text, visual, …) for ranked list
- Results filtering (i.e. by modality)
- …

Similarity by colour

Interest point-based approaches

- A set of interest points is extracted from the image by an algorithm.
- Many algorithms are available to do this.
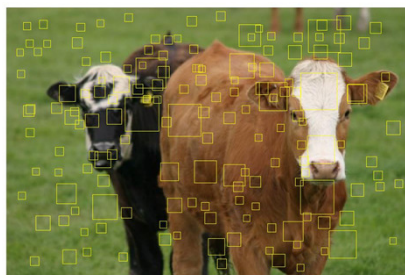
From Bishop ICPR talk

Bags of keypoints (words)

- **Bags of keypoints** approaches are widely used to recognise objects based on interest points.
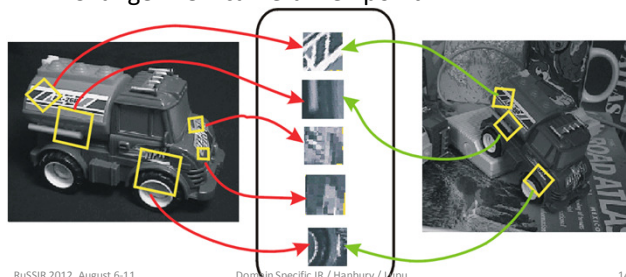
Object → **Bag of 'words'**

Slide from Li Fei-Fei, CVPR 2007 Tutorial

- To be useful, these features should be invariant to:
  - Image scaling
  - Image rotation
  - Change in illumination
  - Change in 3D camera viewpoint.

From D. Lowe, CVPR03 tutorial

Analogy to documents

Of all the sensory impressions proceeding to the brain, the visual experiences are the dominant ones. Our perception of the world around us is based essentially ... reach the brain from o ... thought th ... point by ... cerebra ... upon w ... Throug ... now kne ... perceptio ... more comp ... the visual imp ... various cell layer ... Wiesel have been abl ... message about the image falling on the ... undergoes a step-wise analysis in a syste ... nerve cells stored in columns. In this syste ... cell has its specific function and is responsi ... a specific detail in the pattern of the retinal image.

sensory, brain, visual, perception, retinal, cerebral cortex, eye, cell, optical nerve, image Hubel, Wiesel

China is forecasting a trade surplus of $90bn (£51bn) to $100bn this year, a threefold increase on 2004's $32bn. The Commerce Ministry said the surplus would ... dicted 30% jump in expo ... a 18% rise in imp ... ly to further a ... hat China's ... deliber ... the sur ... one fac ... Xiaochua ... more to bo ... stayed within ... value of the yuan ... in July and permitted it ... ow band, but the US wants the yuan to be ... ed to trade freely. However, Beijing has made ... that it will take its time and tread carefull ... e allowing the yuan to rise further in value.

China, trade, surplus, commerce, exports, imports, US, yuan, bank, domestic, foreign, increase, trade, value
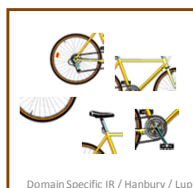
Slide from Li Fei-Fei, CVPR 2007 Tutorial

A clarification: definition of "BoW"

- Looser definition
  - Independent features
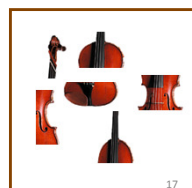
Slide from Li Fei-Fei, CVPR 2007 Tutorial

Bag of Visual Words in Practice

Obtaining visual words:

1. Extract keypoints in a (large) set of images
2. Extract features at each keypoint
3. Cluster the features, producing the visual words
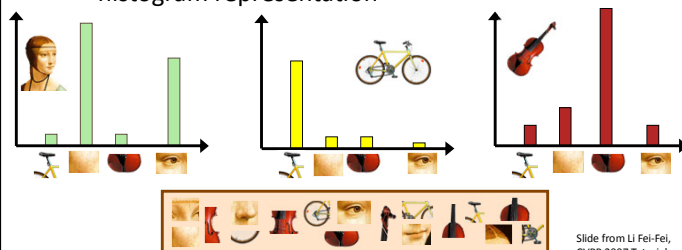4. Assign each keypoint to a visual word (cluster)

A clarification: definition of "BoW"

- Looser definition
  - Independent features
- Stricter definition
  - Independent features
  - histogram representation

Slide from Li Fei-Fei, CVPR 2007 Tutorial

For a new (unseen) image:

1. Extract keypoints
2. Extract features at keypoints
3. Assign each feature to the closest visual word
4. (Text) classification or similarity measures can now be applied

Distance measures

- Histogram intersection
  - For visual words, color features
- TF/IDF weighting
  - For binary/sparsely populated feature spaces/visual words
- Euclidean distance (L2), city block distance (L1)
  - Very frequently used, simple to employ
- Mahalanobis, Bhattacharyya, …
  - Depending on the feature spaces this can lead to much better results
  - See article by Vasconcelos

Nuno Vasconcelos, Andrew Lippman, "A Unifying View of Image Similarity," Pattern Recognition, International Conference on, p. 1038, 15th International Conference on Pattern Recognition (ICPR'00) - Volume 1, 2000

21

Example: Face Image

RuSSIR 2012, August 6-11   Domain Specific IR / Hanbury / Lupu   23

Example of contemporary image search

RuSSIR 2012, August 6-11   Domain Specific IR / Hanbury / Lupu   22
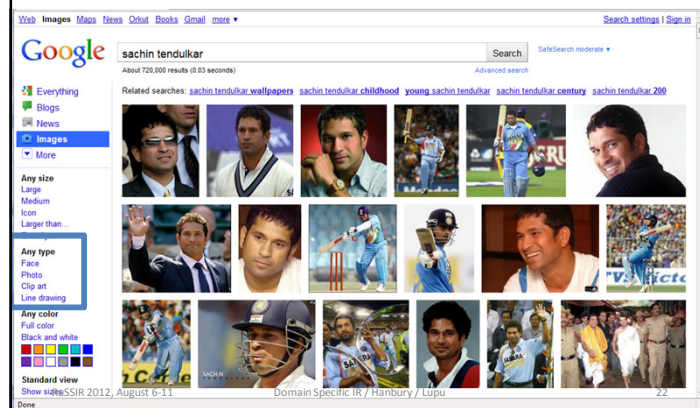
Example: Requesting a Similar Image

RuSSIR 2012, August 6-11   Domain Specific IR / Hanbury / Lupu   24

6
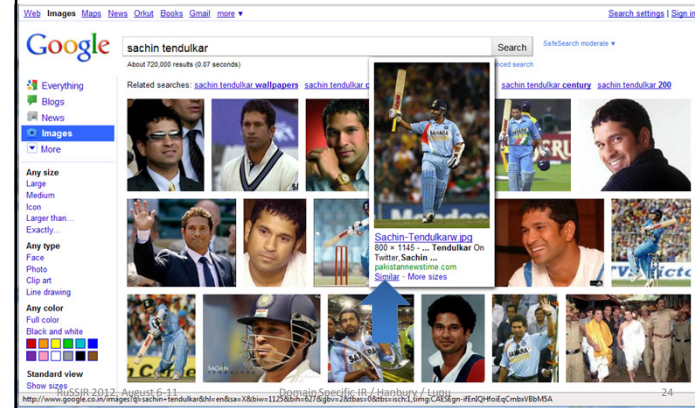
Example: Similar Image Result
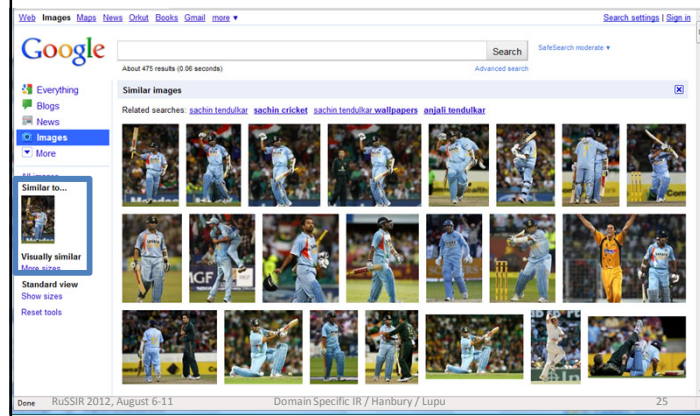
Domain-specific image retrieval

- So far, the same methods as for general image retrieval have generally been used for domain-specific image retrieval
- For domain-specific retrieval, domain knowledge is generally available to guide the choice of:
  - Features
  - Similarity functions

Contents

- Introduction to image retrieval
- Domain-specific image retrieval
  - Image search in the medical world
  - Image search in patents
- Challenges and conclusion

Medical images

Number of images captured per day in a large hospital (each slice of a volume is counted as an image)
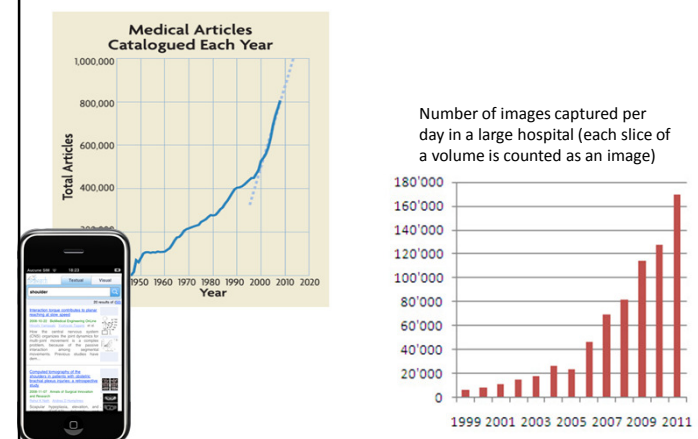
Medical imaging data production

- Medical imaging is estimated to occupy 30% of world storage capacity in 2010!
- Mammography data in the US in 2009 amounts to 2.5 Petabytes

Riding the wave – how Europe can gain from the rising tide of scientific data, *report of the European Commission*, 10/2010.
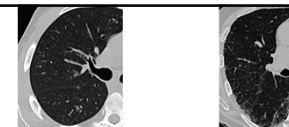
Content vs. context

- Most often text around images does not describe the image content itself
  - Unless specifically annotated for retrieval
  - Text often gives the context in which the images were taken (private, also medical)
- Image content is rarely described precisely and completely with text
  - Visual features describe the pure content
  - Low level of semantics
- Content and context are complementary for search

Text used for image search

- Image captions (in longer texts, scientific articles)
  - Some journals require high quality of this
- Link names to the images (in html)
- URL names (sometimes cryptic)
- Full text
  - Not specific as often many images in a text
- Text close to the images
  - Often better than the full text
- Mapping of free text to semantics
  - MeSH, UMLS, WordNet

Modality classification

Goldminer.arrs.org (249,000 images)

www.yottalook.com (70,000 images)

Medgift demo (300,000 images)

Visual classification and retrieval

http://www.youtube.com/watch?v=cMoONC0Tz2c

Regions of interest

37
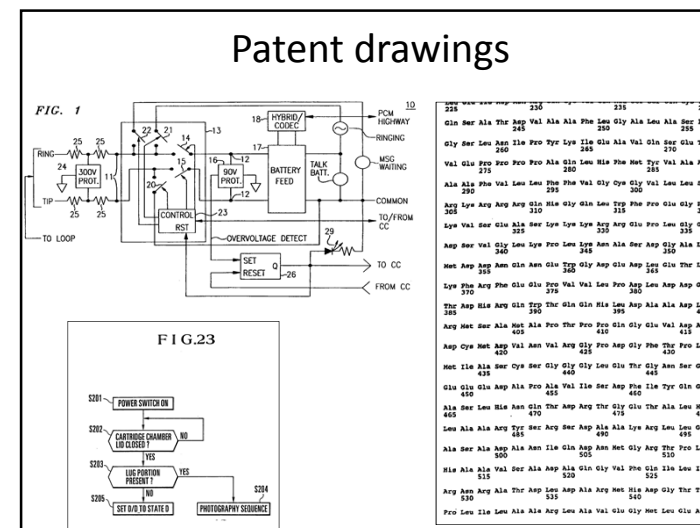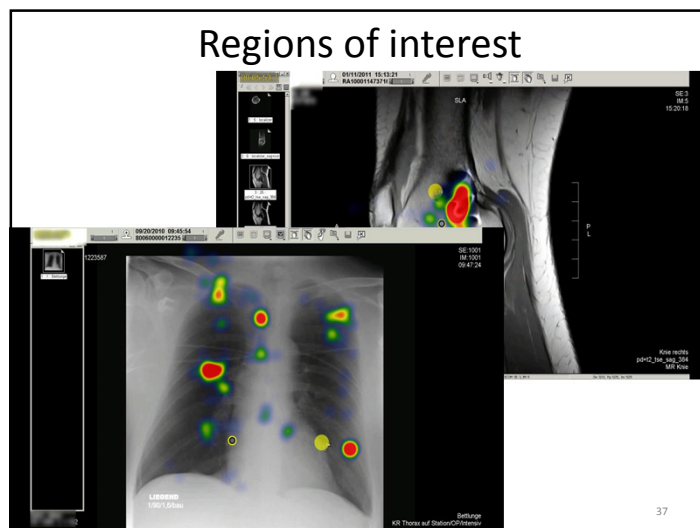
Patent drawings

Images in patents

- *A patent application is required to contain drawings, if drawings are necessary to understand the subject matter to be patented.*
- *Most patent applications contain drawings.*
- *The drawings must show every feature of the invention as specified in the claims.*
- *Omission of drawings may cause an application to be considered incomplete and no application filing date will be granted by the USPTO.*

The USPTO "Nonprovisional (Utility) Patent Application Filing Guide"

Characteristics

- Sometimes information is in the drawings that is not in the text (e.g. mechanical gearing)
- Drawings in older patents are often done by hand
- Drawings are usually in bitmap format from scanning, and not in vector format
- Drawings can be of poor quality in older patents due to poor scanning technology

Chemical Structure

Abstract Drawing

Flow Chart

FIG. 7

popup * pop = [popup alloc]; /* allocate memory space to the pop up object */

[popup autorelease];

/* have the autorelease pool hold a pointer (i.e., retain) the pop up object */

return pop;

Program Code

$$ex(i) = 2 \sum_{j=1}^{(M+1)} a_{2j-1}[ex(i-2j+1)+ex(i+2j-1)],$$

Mathematical Formula

Graph

Cys Cys Val Pro Gln Ala Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val

Gly Arg Lys Pro Lys Val Glu Gln Leu Ser Asn Met Ile Val Arg Ser

Cys Lys Cys Ser

Gene Sequence

Character

CLEF-IP 2011 Patent Image Classes   Table

Types of Patent Image

From: S. Vrochidis, S. Papadopoulos, A. Moumtzidou, P. Sidiropoulos, E. Pianta, and I. Kompatsiaris. Towards content-based patent image retrieval: A framework perspective. World Patent Information, 32(2):94-106, 2010.

Variability

Links between drawings and text

FIG. 5 illustrates, in schematic, cross-sectional view […] rotatable drive shaft or spindle 66 […]

FIG 5
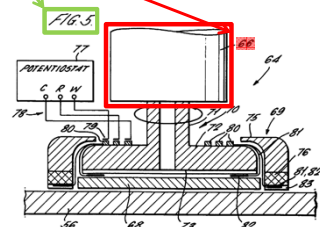
- Patent applicants discouraged from using words in drawing labels

Visual Feature Extraction

Feature extraction algorithm

Feature vectors
[3.7, 29.3, -0.6, -2.3, …]
[2.5, 8.1, -0.2, 1.8, …]
[0.7, 26.9, -0.4, 2.5, …]

Classification (machine learning)

Image similarity (retrieval)
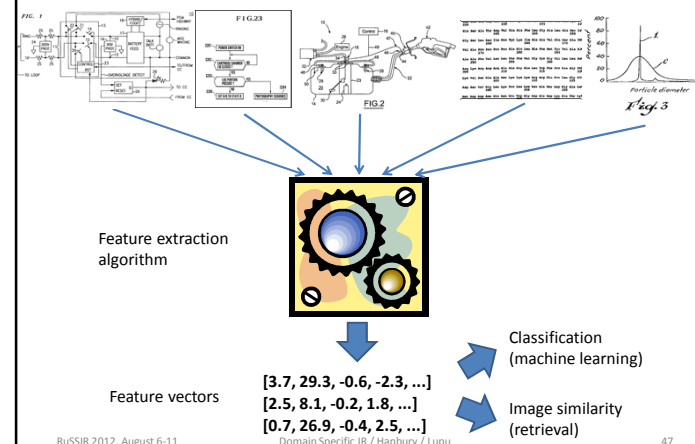
- Patents also contain brief descriptions of each drawing in the text (but not as a caption)
- Usually in a section headed *Brief Description of the Drawings*
- E.g. *FIG. 1. shows a longitudinal section through a first embodiment of an actuator*
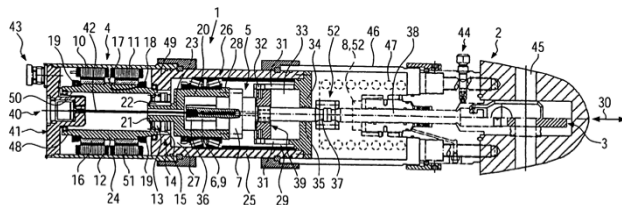
FIG. 1

US2012/0153198 A1

Image similarity in patents

- Patmedia
  - Informatics and Telematics Institute, Greece
  - Adaptive Hierarchical Density Histogram
    - Focus on geometry
    - Sidiropoulos et al., 2011

PatMedia

Similarity

- A difference in patents
  - Markush Structures
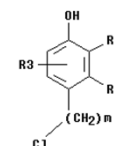
Substituent Variation:	R1 is methyl or ethyl
Homology Variation:	R2 is alkyl
Position Variation:	R3 is amino
Frequency Variation:	m is 1-3

Example from Barnard and Downs:1997

  - Potentially infinite combinations
  - Extremely difficult to properly match image, table and free text

RuSSIR 2012, August 6-11   Domain Specific IR / Hanbury / Lupu   51

Image similarity

- When images are not images (1)

- These images must be converted to a textual representation of chemical structure for search

RuSSIR 2012, August 6-11   Domain Specific IR / Hanbury / Lupu   50

Image similarity

- When images are not images (2)

52

Similarity

- Flowchart representation
  - MT Title "Fig.7"
  - MT Nodes 10
  - MT Edges 9
  - CO ====== Now comes the list of nodes ======
  - CO === identifier type text ============
  - NO 1 oval "BEGIN"
  - NO 2 rectangle "RECEIVE AND DIGITIZE IMAGE"
  - NO 3 rectangle "DISPLAY IMAGE AND SELECT CHART"
  - NO 4 rectangle "DESIGNATE APPROXIMATE POSITIONS"
  - NO 5 rectangle "RECOGNIZE GRAPHICAL OBJECT AND OUTPUT"
  - NO 6 oval "BEGIN"
  - NO 7 no-box "80"
  - NO 8 no-box "82"
  - NO 9 no-box "84"
  - NO 10 no-box "86"
  - CO ========= Here come the edges ===========
  - CO === start-node end-node type text ========
  - EG 1 2 plain ""
  - EG 2 3 plain ""
  - EG 2 7 wiggly ""
  - EG 3 4 plain ""
  - EG 3 8 wiggly ""
  - EG 4 5 plain ""
  - EG 4 9 wiggly ""
  - EG 5 6 plain ""
  - EG 5 10 wiggly ""
  - CO ======== THIS IS IT =====================

Fig. 7

BEGIN

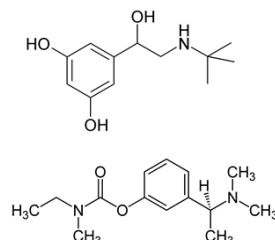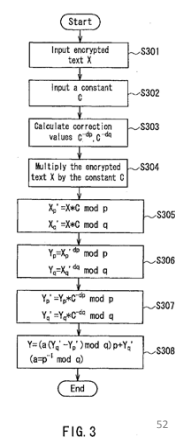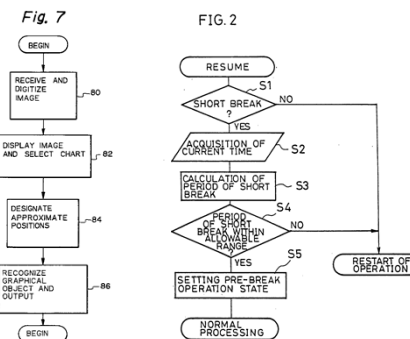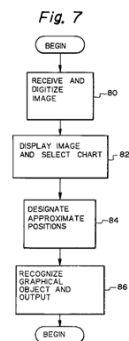RECEIVE AND DIGITIZE IMAGE — 80

DISPLAY IMAGE AND SELECT CHART — 82

DESIGNATE APPROXIMATE POSITIONS — 84

RECOGNIZE GRAPHICAL OBJECT AND OUTPUT — 86

BEGIN

CLEF-IP 2011 Image-Based Prior Art Search

- 3 IPC sub-classes (of 631) used:

A43B	CHARACTERISTIC FEATURES OF FOOTWEAR; PARTS OF FOOTWEAR
A61B	DIAGNOSIS; SURGERY; IDENTIFICATION
H01L	SEMICONDUCTOR DEVICES; ELECTRIC SOLID STATE DEVICES NOT OTHERWISE PROVIDED FOR

- 47.000 xml files (patent documents), 290.000 tiff files – 5.5 GB
- 211 query patents
- Retrieval of patents within the same IPC class (i.e. group of images)

Patent image search evaluation

- Patent image retrieval evaluations have been done on small datasets (max. 2000 images)
- Changed in 2011 with 3 patent image evaluations:
  - CLEF-IP Image-based Prior Art
  - CLEF-IP Image Classification
  - TREC-CHEM Chemical Structure Recognition
- Even more data is available for such tasks...

Results

- Only 1 participant (Xerox)
- Best MAP for 3 types of runs are shown below:

Run Type	MAP
Image only	0.035
Text only	0.203
Text + Image	0.212

- Combining image and text search can improve the retrieval results

CLEF-IP 2011 Patent Image Classification

Training images	
5565	**Abstract drawing**
5957	**Chemical structure**
5573	**Program listing (code)**
5983	**Gene sequence (dna)**
310	**Flow chart**
1663	**Graph**
5950	**Mathematical formula**
5502	**Table**
1578	**Character (symbol)**
38081	

1000 test images

Patent drawing search: present and future

- Currently, much drawing search in practice is done by manual visual comparison
- There are initiatives to change the rules for patent drawing in patent submissions:
  - Why not use colour?
  - Why submit 2D cross-sectional black-and-white drawings when a 3D CAD model would be much more useful?
- For automated search, more effort is being put on specific search techniques for specific drawing classes

---

Results

run	TPR
joanneum.alphacentauri	0.66
joanneum.arcturus	0.50
joanneum.betelgeuse	0.62
joanneum.canopus	0.65
joanneum.procyon	0.27
joanneum.rigel	0.63
joanneum.sirius	0.64
joanneum.vega	0.28
xerox-sas.RUNORH	0.85
xerox-sas.RUNORH_ROTRAIN	0.91
xerox-sas.FV_ORH_SP	0.85
xerox-sas.MEAN_ALL	0.85

True Positive Rate

Contents

- Introduction to image retrieval
- Domain-specific image retrieval
  - Image search in the medical world
  - Image search in patents
- Challenges and conclusion

Some challenges

- Combining text retrieval, visual retrieval and visual classification based on needs, queries
  - Analysis of the query to estimate goal
  - Quick learning from small training data, changing images
  - Work on robustness of fusion techniques
    - Rank, not score based
- Domain-specific retrieval
  - Medical, expert search, …
- …

Conclusion

- Many techniques for general image search are also used for domain-specific image search
- For domain-specific image search, it is usually possible to use domain knowledge to make informed choices on
  - Image features
  - Similarity functions
- Still much research to do, especially on combining information from different modalities

Combinations of techniques

- Classify modalities, then search in sub space
  - Basically remove noise
  - Allow for tabbed browsing
  - Separate compound figures, separate them and reclassify the sub figures
- Detect important regions
  - Segment areas, visualize the results
- Retrieve similar cases/documents based on structured data, free text and the visual information
  - Case/document-based retrieval is much more realistic than image-based retrieval