VK Multimedia Information Systems

Mathias Lux, mlux@itec.uni-klu.ac.at

Dienstags, 16.oo Uhr c.t., E.1.42

Sample Projects (not to choose)

- LSA of MP3s ID3v2
  - Clustering & retrieval based on metadata
- Java based video summarization
  - Greedy keyframe distance maximization
- Image mosaics
  - Based on LIRe, now part of it

Proposed projects

- Flickr Photo Grabber
  - Cmd-line
  - File + metadata, incl. license
  - Feature: tag search
  - Metadata stored in IPTC

```
$> flickrgrab <numImages> [-t <tag>] [-d <directory>]
```

MSER Implementation

- Implement the Maximally stable extremal regions algorithm in Java

see http://en.wikipedia.org/wiki/Maximally_stable_extremal_regions

Feature Extraction with OpenCL

- Prepare Introduction to OpenCL
- Show exemplary work with OpenCL
- Compare to sequential programming
  - E.g. on ImageSeams, ROI, MSER, Pixel rank or Color Correlogram

LSA in LIReDemo

- Implement LSA in LIReDemo
  - Extend an existing program ...

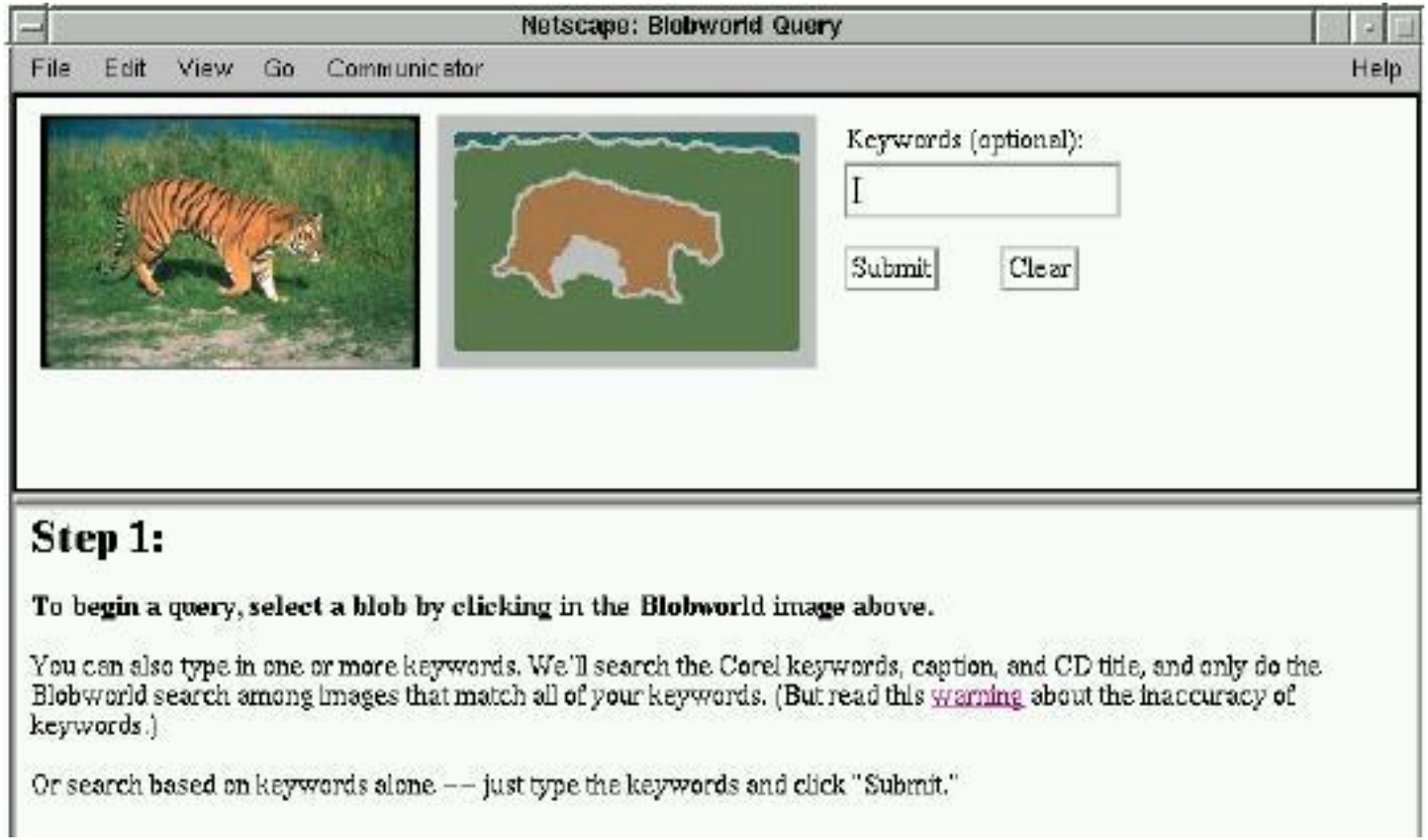
Content Based Image Retrieval II

- Segmentation
- Regions of Interest
- Object Recognition
- Vector Images
- Evaluation

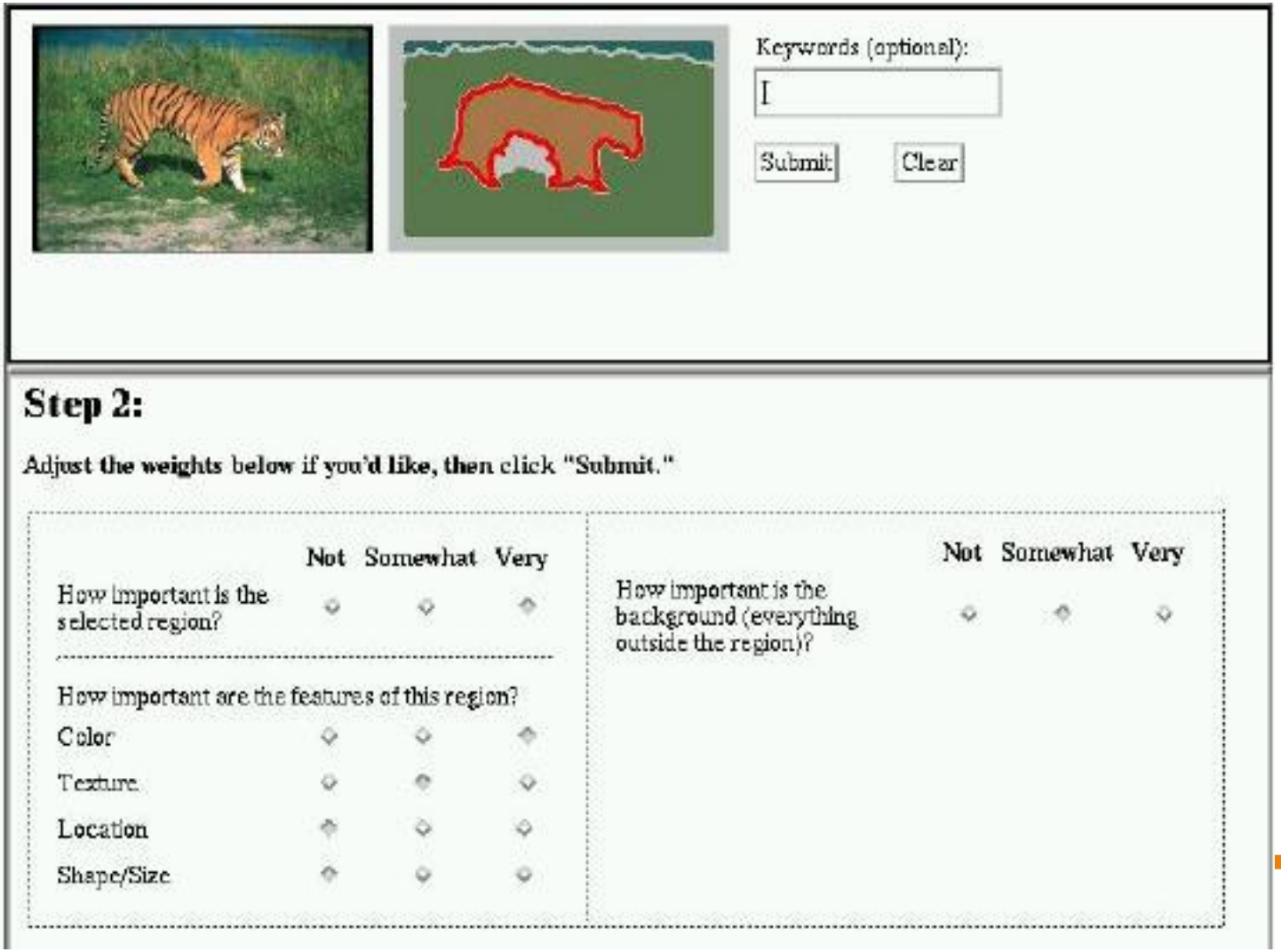
Segmentation

- Automatically decompose image
  - Foreground & background
  - Different objects & shapes
- Example applications
  - Computer vision in industry: Counting
  - Security: Erasing background

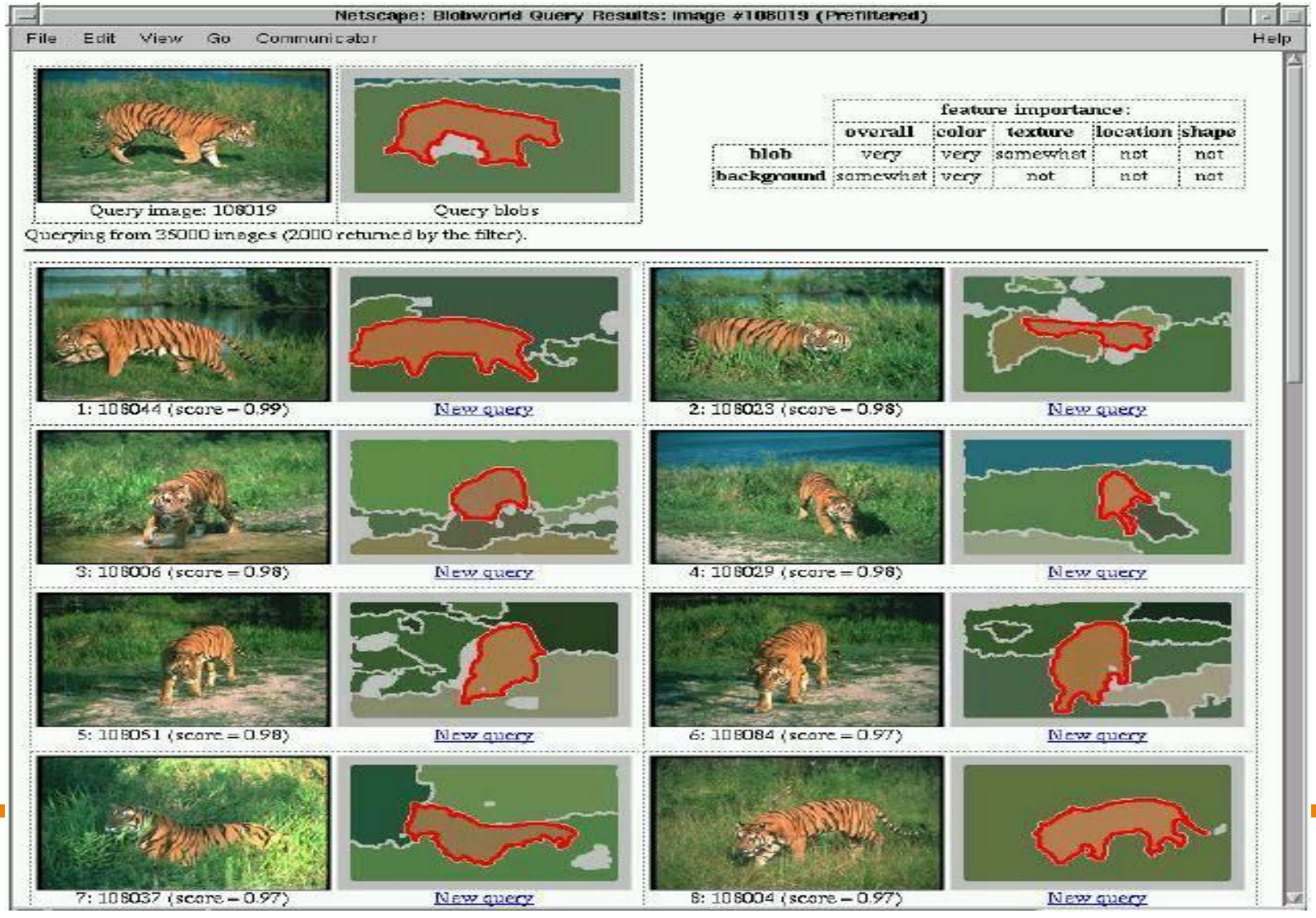
Application: Blobworld



Application: Blobworld



Application: Blobworld



Application: Blobworld

- Segmentation Method: Pixel clustering
  - Each pixel has feature vector
    - Color, texture in neighbourhood, etc.
  - Disjunctive clusters group similar pixels
- Method available in MatLab source code

Segmentation Example: Region Growing - Flood Fill

- Image is represented by its gradient map
- Some pixel is selected as seed
- Imaginary water is poured onto this seed
  - Flooding pixels with lower or equal height

- Problems:
  - Dam breach vs. amount of water
  - Good seed

Content Based Image Retrieval II

- Segmentation
- Regions of Interest
- Object Recognition
- Vector Images
- Evaluation

Regions of Interest

- Identify interesting patches in images
- Automatic extraction of ROIs
  - Top-down, based on a model
  - Bottom-up, e.g. stimulus-driven
- Applications
  - Image re-targeting
  - Image cropping

Src.: Borba, Gamba, Marques and Mayron , "Extraction of salient regions of interest using visual attention models", SPIE Conference on Multimedia Content Access: Algorithms and Systems III, 2009

Bottom-Up Visual Attention

- Attention Models
  - Find most interesting point in visual scene
  - Direct gaze towards this point
  - *Selective* or *focal* attention or *attention for perception*
- Metaphor of a spotlight
  - Sweeping the scene
  - Highlighting most important parts

Model of Itti, Koch & Niebur

- Biologically inspired
- Three low level dimensions of an image
  - Color, orientation and intensity
- Features are extracted in different scales
  - This results in feature maps
- Normalization -> conspicuity maps
- Normalization & summing -> saliency map
  - Peaks are salient points

Itti, Koch & Niebur, "A Model of Saliency-based Visual Attention for Rapid Scene Analysis", PAMI 1998

Model of Itti, Koch & Niebur

- In iterations
  - Preserve prominent peaks
  - Inhibit small peaks
- Number of iterations decides on the outcome

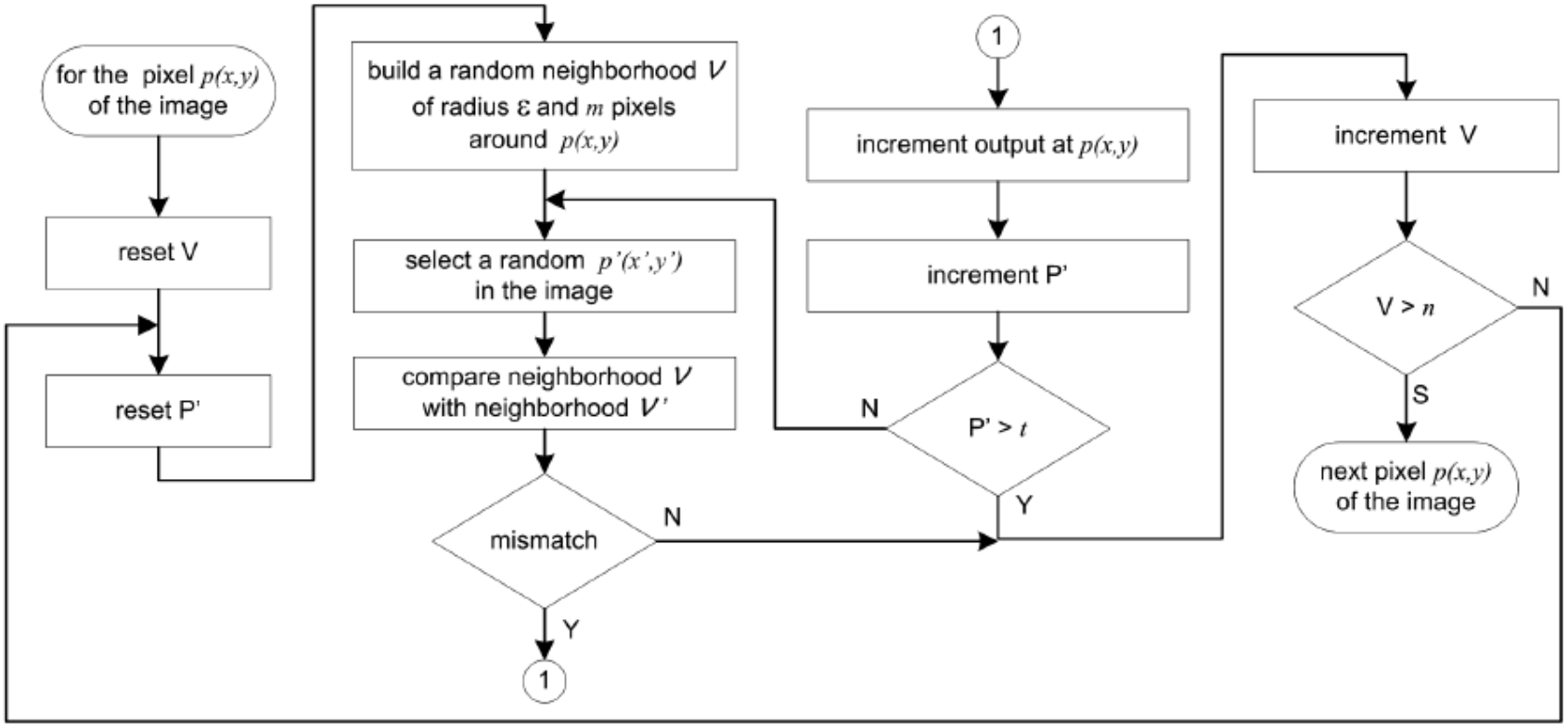
Src.: Borba, Gamba, Marques and Mayron , "Extraction of salient regions of interest using visual attention models", SPIE Conference on Multimedia Content Access: Algorithms and Systems III, 2009

Model of Stentiford

- Suppress areas of repetitive color patterns
- For each pixel:
  - Compare a number of randomly selected pixels
  - Based on color in neighbourhood
  - High value: low number of similar areas
  - Low value: lots of similar areas
- Result added up to saliency map

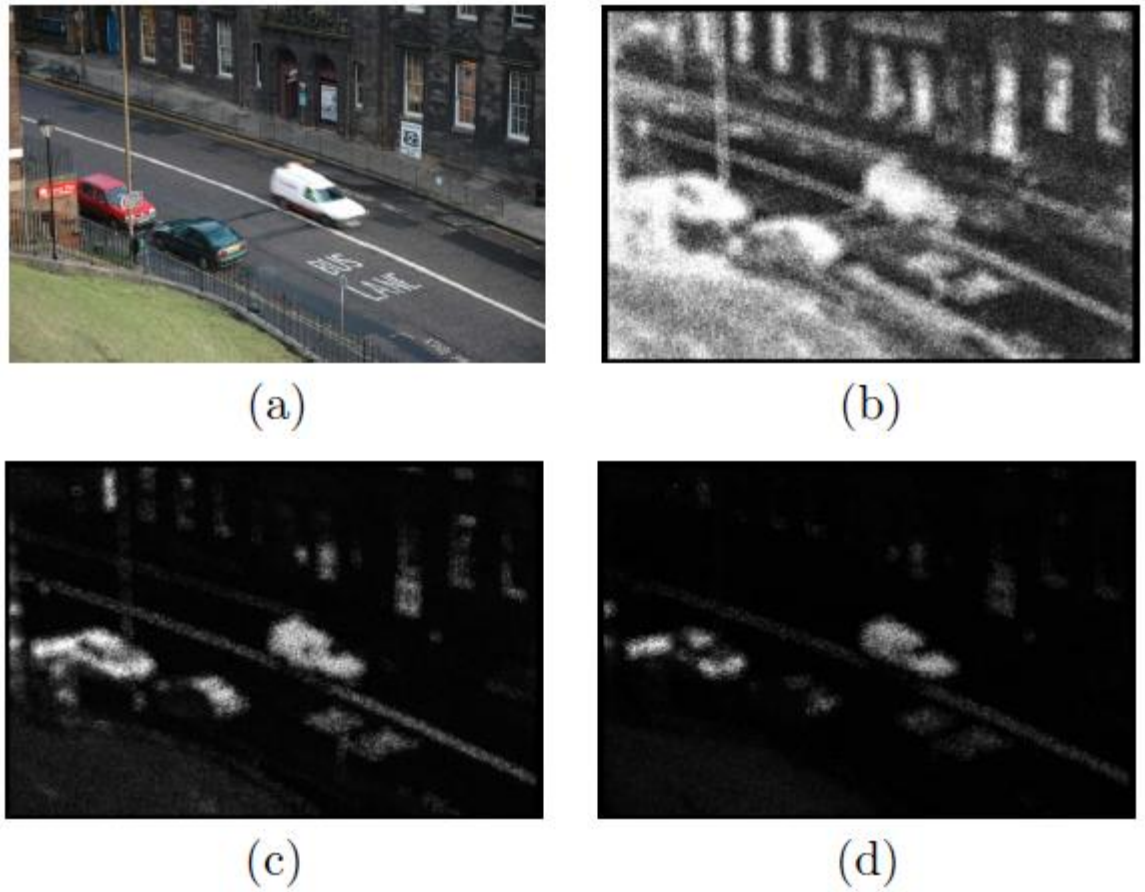
F. W. M. Stentiford, "An estimator for visual attention through competitive novelty with application to image compression," Proc. Picture Coding Symposium, pp 101-104, Seoul, 24-27 April, 2001.

Model of Stentiford



Src.: Borba, Gamba, Marques and Mayron , "Extraction of salient regions of interest using visual attention models", SPIE Conference on Multimedia Content Access: Algorithms and Systems III, 2009

Model of Stentiford



Content Based Image Retrieval II

- Segmentation
- Regions of Interest
- Object Recognition
- Vector Images
- Evaluation

Object Recognition

- Motivated by practical applications
  - License plate detection & recognition
  - Logo detection & classification (e.g. in TV)
  - Face detection & recognition
  - Identifying spatial regions of interest
    - possible tumors in medical imaging

Object Recognition

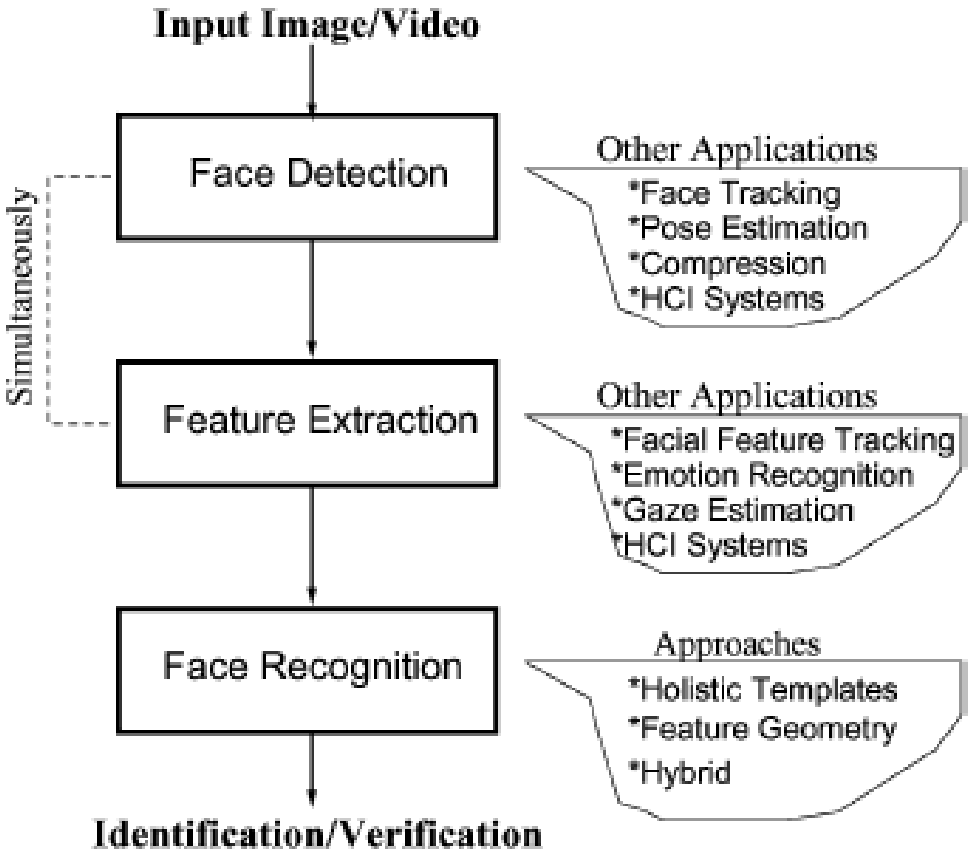
Weight of problem depends on domain ...

- License plate detection is rather easy
  - Plates always look the same
  - Small number of possible positions
  - Sensor fixed, background is separated easily

Face Recognition

Areas	Specific applications
Entertainment	Video game, virtual reality, training programs
	Human-robot-interaction, human-computer-interaction
Smart cards	Drivers' licenses, entitlement programs
	Immigration, national ID, passports, voter registration
	Welfare fraud
Information security	TV Parental control, personal device logon, desktop logon
	Application security, database security, file encryption
	Intranet security, internet access, medical records
	Secure trading terminals
Law enforcement and surveillance	Advanced video surveillance, CCTV control
	Portal control, postevent analysis
	Shoplifting, suspect tracking and investigation

Face Recognition



Face Detection

- Task:
  - Isolate face from background / other faces
- Problems:
  - Partially hidden
  - Point of view
- Methods:
  - Templates (whole face & feature based)
  - Skin color
  - Neural networks & machine learning

Feature Extraction

- Extraction of characteristics
  - Different approaches (e.g. *Eigenfaces* and *Fisherfaces*)
  - Psychological background
  - Index key facial features like position of nose, eyes and mouth
  - Also needed to normalize the holistic face

Face Recognition

- Holistic methods
  - Face as a whole is indexed (mostly PCA, then classification)
- Feature based
  - Eyes, mouth, nose, etc.
- Hybrid methods
  - Combination of both

Content Based Image Retrieval

- Segmentation
- Regions of Interest
- Object Recognition
- Vector Images
- Evaluation

Vector Images

- In contrast to raster images
  - Description of boundaries, regions & effects
  - Raster is rendered at view time
  - Aspects: Size, scaling, modification
- Applications
  - Graphs & charts
  - Cliparts & illustrations
  - Logos & Fonts
  - 3D models and scenes

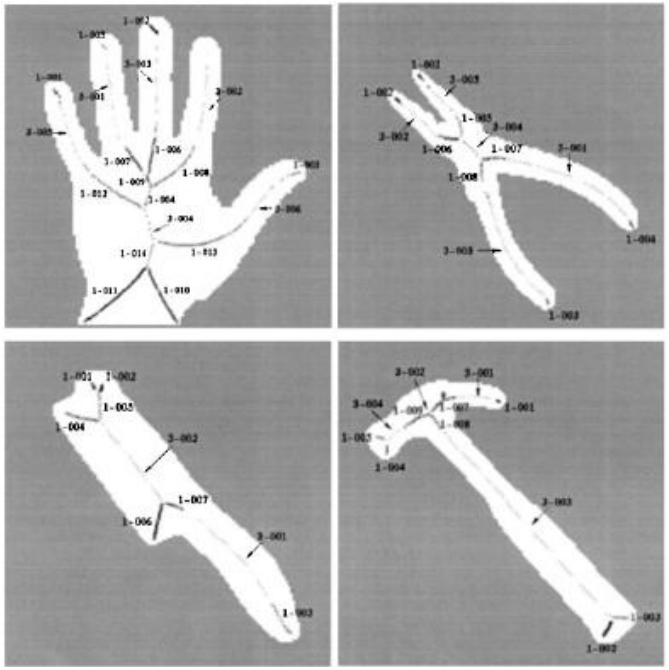
Vector Image Retrieval

- Same aspects as for raster images
  - Color, shape & texture
- Selection of aspects depends on domain
- Extraction might be easier
  - Shape, edge and texture features
  - Objects are not necessarily defined

```
<g> <polygon fill-rule="evenodd" clip-rule="evenodd" fill="#E58325" points="2.711,36.656 40.199,36.656 51.75,2.64
63.337,36.656 100.789,36.656 70.494,57.667 82.081,91.648 51.75,70.636 21.455,91.648 33.042,57.667 "/> <path
fill="#FED6AD" stroke="#FED6AD" stroke-width="0.0354" stroke-linejoin="round" stroke-miterlimit="10" d="M0.018,35.805
h40.181l-0.815,0.567L51.75,0.018l12.401,36.354l-0.814-0.567h40.181L70.99,58.376l10.319-0.992l12.401,36.461-32.421-
22.5h0.957 l-32.421,22.5l12.402-36.461l0.319,0.992L0.018,35.805L0.018,35.805zM34.034,57.348L22.27,91.931l-1.311-
0.957L51.75,69.608 l30.791,21.366l-1.275,0.957L69.502,57.348l30.827-
21.401l0.46,1.559H62.734L50.971,2.923h1.595L40.801,37.506H2.711l0.496-1.559 L34.034,57.348L34.034,57.348z"/> </g>
```

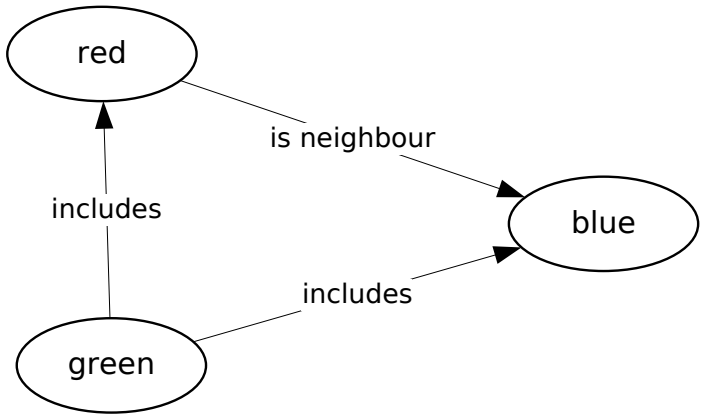
Example: Shock Graphs

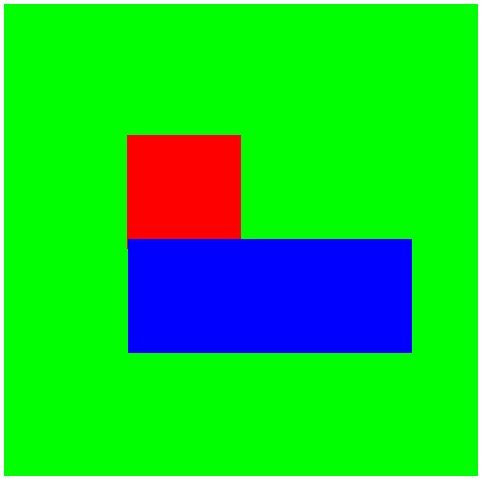
- Actual shape is reduced to a graph (tree)
  - Mostly preserving a characteristic
- Graphs are indexed
  - Using invariants
  - Low dimensional vector
- Retrieval is done by
  - Spatial access



Example: Graph Isomorphism

- Clipart is reduced to graph
  - Describing color & relation between regions
    - includes, is-neighbour





Example: Graph Isomorphism

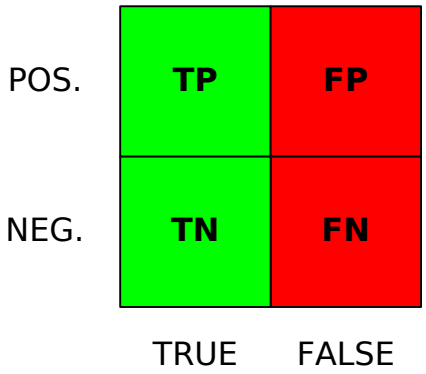
- Graphs are compared pairwise
  - Based on the similarity of nodes (colors)
  - And structure (edges and types)
- Corpus is clustered hierarchically
  - Generates a “tree of images”
  - That’s called “metric index”
- Retrieval is based on pairwise check with cluster representatives
  - Efficiency depends on the index.

Content Based Image Retrieval

- Segmentation
- Object Recognition
- Regions of Interest
- Vector Images
- Evaluation

Evaluation

- Evaluation methods are similar to information retrieval tasks:
  - Comparison of methods based on precision & recall, error classification rate, etc.
  - Evaluation of application user centered (subjective evaluation)
- For classification tasks
  - False/True Posititves/Negatives
- Test data sets
  - E.g. Wang-1000



Test data sets

- Wang-1000
  - 10 classes with 100 images each
  - Taken from the Corel stock photos
- University of Washington data set
  - 1,109 images, partially annotated with keywords
  - 18 categories, image relevant if at least 2 keywords match
  - http://www.cs.washington.edu/research/imagedatabase/groundtruth/

Test data sets

- IRMA database
  - 10,000 images, annotated, 57 classes
  - Greyscale, RWTH Aachen University Hospital
  - Used for imageCLEF2005
- Zurich Buildings Database for Image Based Recognition (ZuBuD)
  - 1,005 images of 201 buildings
  - 105 query images
- UCID database
  - 1,300 images & ground truth

Src.: Deselaers, Keysers & Ney, "Features for Image Retrieval: An Experimental Comparison", 2007
http://thomas.deselaers.de/system/files/cbir.pdf

User Centered Evaluation

- Sample Evaluation on the use of content based organization of images
  - User were presented layouts of images
  - Compare the 3 types and a list representation

User Centered Evaluation

- Results:
  - Content based method good for “graphical tasks”
  - Metadata based method depends on annotation
  - Random organization also helps user
    - Identification of “strong” images
    - Stick out of surroundings

Homework ?!?

- Think about a project
- Write a proposal
  - Title
  - Management summary (100 words)

Thank you ...

... for your attention