Imperial College
London

Image, Shape & Multimedia Resource Discovery

Stefan Rüger
Multimedia & Information Systems
mmis.doc.ic.ac.uk

Frederic Fol Leymarie
Arts Computing
doc.gold.ac.uk/~ffl

Outline

Part 1: New paradigms for media access

Part 2: Shape representation

Background: Stefan Rüger

- Trained as theoretical physicist
- PhD in theory of neural networks
- Imperial College London
  - PostDoc in data mining and text mining (1997-1999)
  - EPSRC Advanced Research Fellow (1999-2004)
  - Research Lecturer (2000-2002), Lecturer (2002-2005), Reader in Multimedia and Information Systems (2005-2006)
- Professor of Knowledge Media, KMI, the Open University

Background: Frederic F. Leymarie

- Trained as an engineer (electrical and physics)
- PhD in 3D shape theory
- Brown University (USA)
  - Co-founder of the SHAPE Lab. (1999); NSF supported
  - SHAPE Lab manager (2002-2004)
- Professor of Arts Computing, Goldsmiths Digital Studios

Outline

- **New Paradigms in Media Access**
  - **Multimedia Management and Access**
    - Semantic gap and Polysemy
    - Content-based visual search
    - Lateral browsing
    - Commercial applications
- Shape representation
  - 2D
  - 2.5D, images, motion
  - 3D

Research: Multimedia Management and Access

What have you got?

Multimedia Management and Access

- Archive
  - text, video, images, speech, music, web content, …
- Query
  - text, stills, sketch, speech, humming, examples, …
  - **content-based**
- Value-added services
  - browsing, summaries, story boards
  - document clustering, cluster summaries
- Putting the user in the loop

New types of Search Engines

text	stills	sketch	speech	sound	humming	motion	query \ doc	Example
							text	conventional text retrieval
							video	get a wildlife documentary
							images	type “flowers” and get photos
							speech	radio news
							music	hum a tune and get a music piece
							sketches	
							multimedia	

Applications

- Medicine
- Personal collections
- Multimedia digital libraries
- Media archives
- Entertainment
- Tourism
- E-learning
- Retail

Outline

- **New Paradigms in Media Access**
  - Multimedia Management and Access
  - **Semantic gap and Polysemy**
  - Content-based visual search
  - Lateral browsing
  - Commercial applications
- Shape representation
  - 2D
  - 2.5D, images, motion
  - 3D

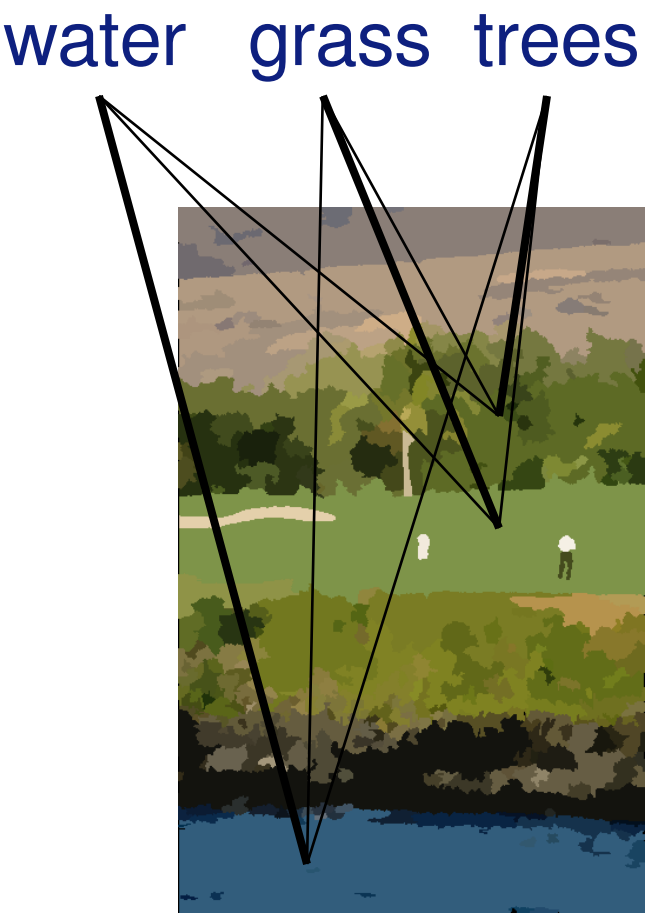
The semantic gap

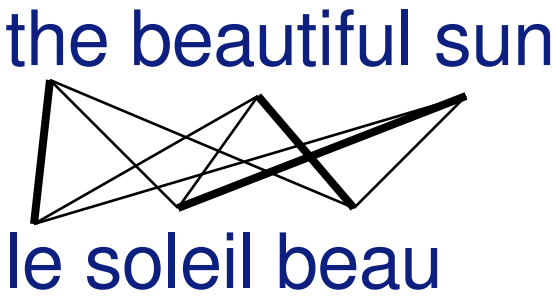
1. 120,000 pixels with a particular spatial colour distribution

2. faces and a vase-like object

3. victory, triumph, ...

Automated image annotation





Automated image annotation as machine learning problem

- Training set, e.g., 32,000 images such as

animal, creature, face, grass, nature, outdoors, outside, portrait, teeth, tiger, vertical, water, wild, wildlife

[with Magalhães, CIVR 2006]

[PhD work in progress: A Yavlinsky, Nov 2003-, partly ORS funded]

[PhD work in progress: J Magalhães, Oct 2004-, funded by FCT]

[with Magalhães, book chapter 2006]

[with Magalhães, SIGIR 2005]

[with Yavlinsky and Schofield, CIVR 2005]

[with Yavlinsky, Heesch and Pickering, ICASSP 2004]

Automated image annotation: an example

Automated: water buildings city sunset aerial

behold™

Visual (selected) Text

scenic panorama view Search in Universities

Previous Next

Results 1 to 16 of 52980


http://sub.domain.com/image.jpg


http://sub.domain.com/image.jpg


http://sub.domain.com/image.jpg


http://sub.domain.com/image.jpg


http://sub.domain.com/image.jpg


http://sub.domain.com/image.jpg


http://sub.domain.com/image.jpg


http://sub.domain.com/image.jpg

Bridging the semantic gap

- Region classification (grass, water, sky, ...) + simple vocabulary
- Infer complex concepts from simple models (**barbecue** = grass+plates+people), e.g., using ontologies

[collaboration with AT&T Research, Cambridge 2001]

Polysemy

Relevance feedback

- endow system with plasticity (parameters)
- **system needs to learn from user**
  = change the parameters

[with Heesch and Yavlinsky, CIVR 2003]
[with Heesch, ECIR 2003]

Applet Viewer: Core.Aetos.class

Applet

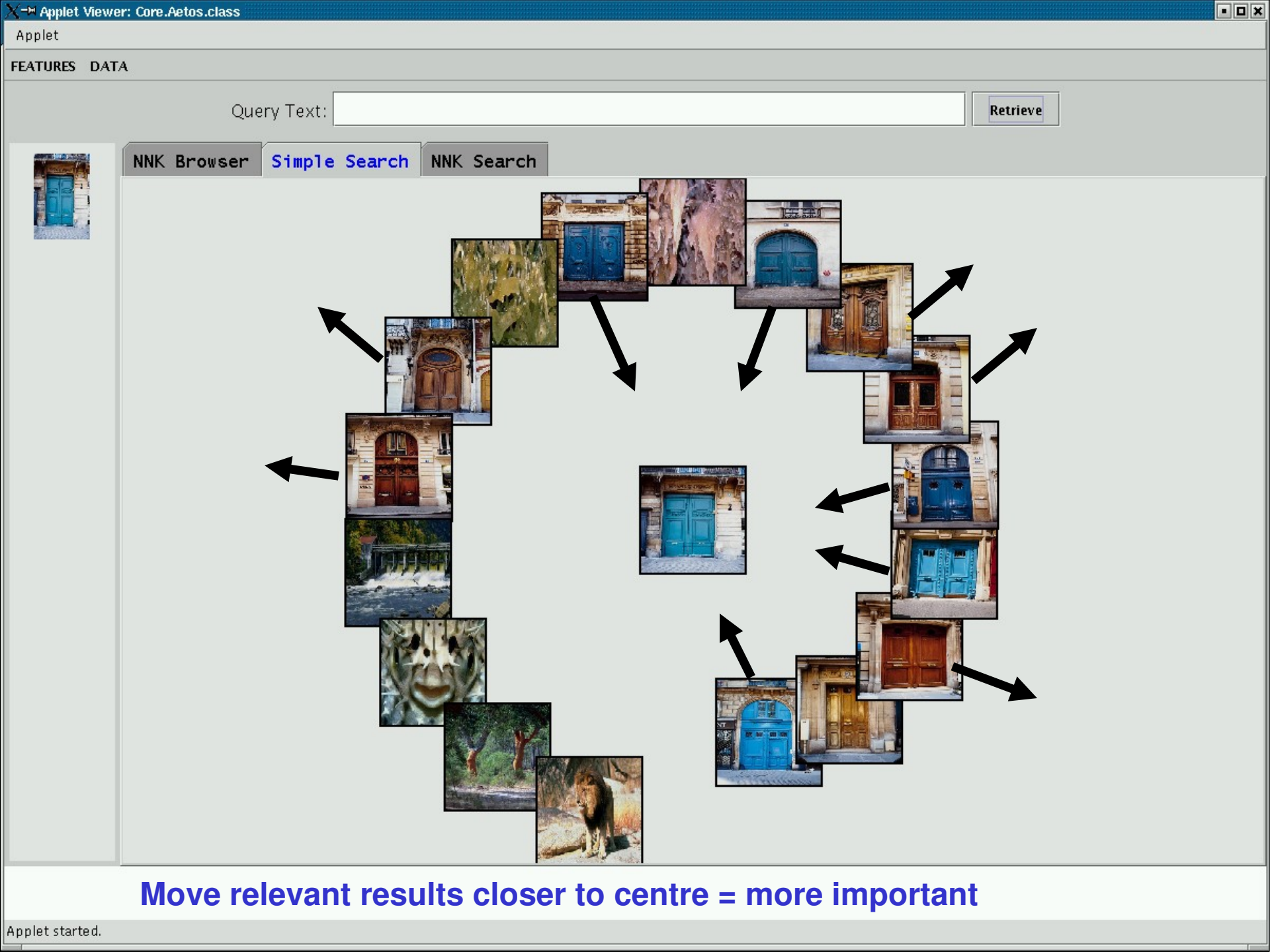
FEATURES DATA

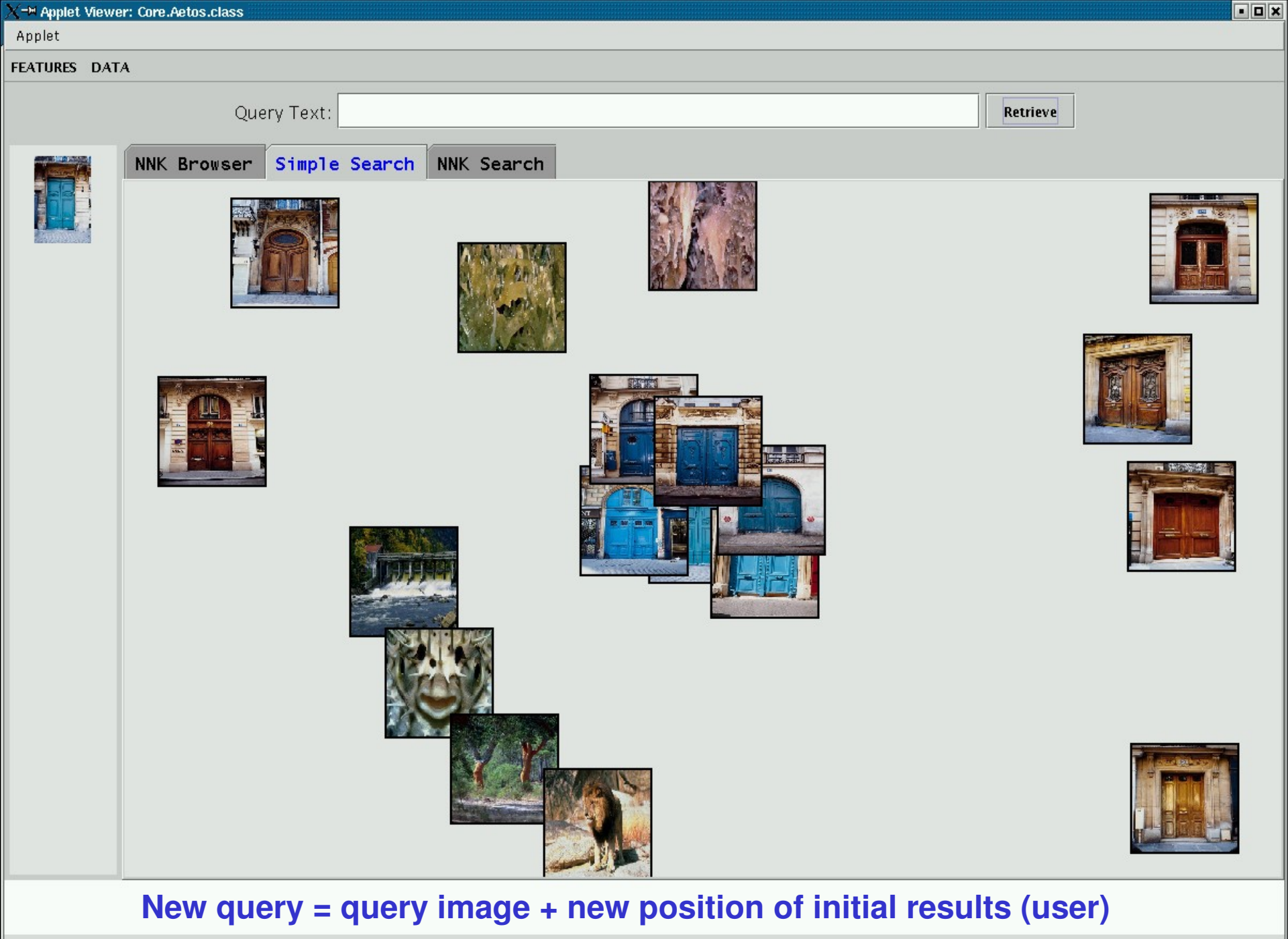
Query Text: Retrieve

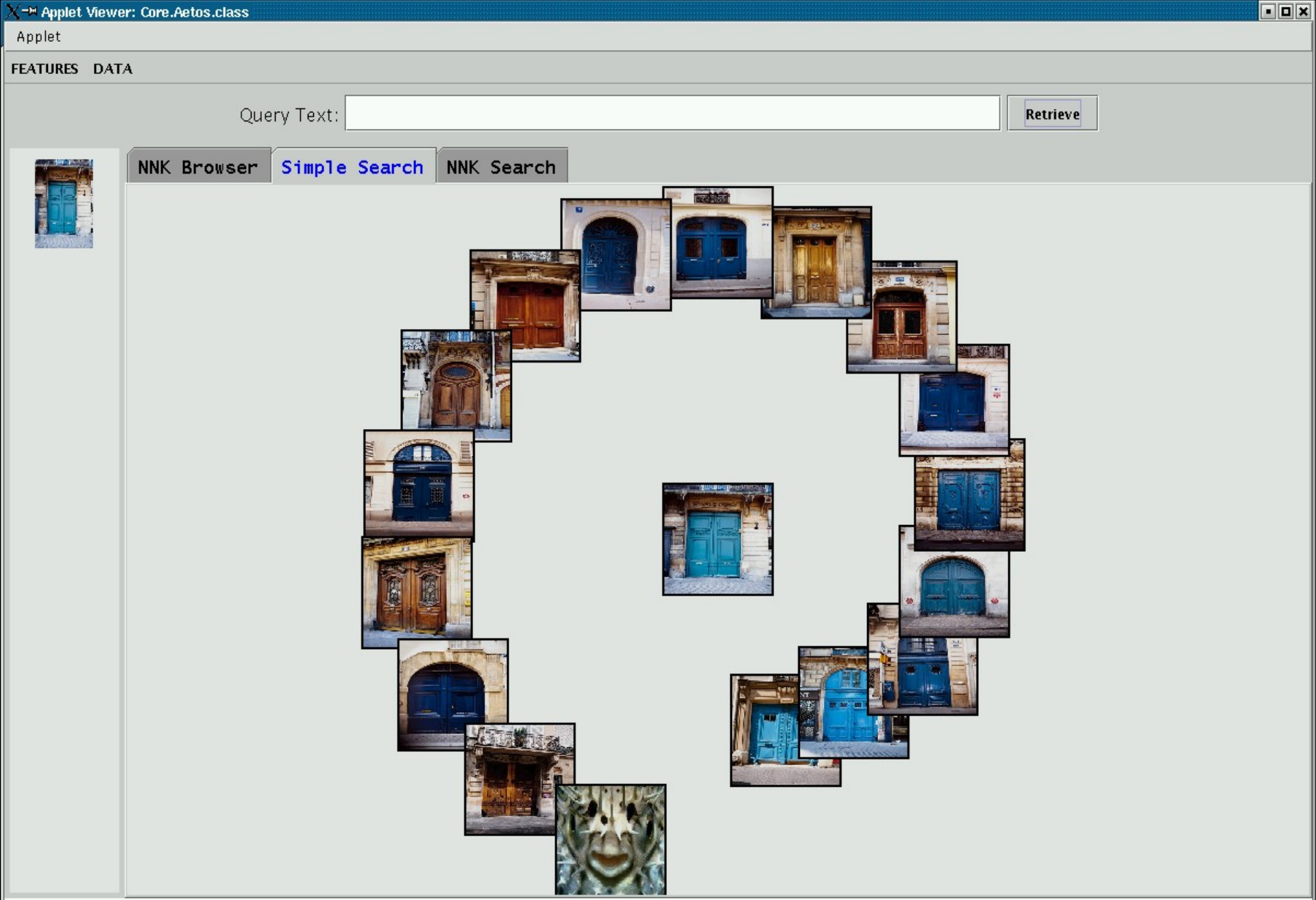
NNK Browser | Simple Search | NNK Search

Search for images of blue doors by example image

Applet started.


Applet Viewer: Core.Aetos.class
Applet
FEATURES DATA
Query Text:
Retrieve
NNK Browser
Simple Search
NNK Search
Move relevant results closer to centre = more important
Applet started.


Applet Viewer: Core.Aetos.class
Applet
FEATURES DATA
Query Text:
Retrieve
NNK Browser
Simple Search
NNK Search
New query = query image + new position of initial results (user)
Applet started.



Next round: many more relevant results (blue doors)

Outline

- **New Paradigms in Media Access**
  - Multimedia Management and Access
  - Semantic gap and Polysemy
  - **Content-based visual search**
  - Lateral browsing
  - Commercial applications
- Shape representation
  - 2D
  - 2.5D, images, motion
  - 3D

Content-based Visual Search

[book chapter: 2006 to appear]
[Choplin 1.3 open-source software, 3 Q 2006]
[iBase 1.2 open-source software, 3 Q 2006]
[keynote: EU infoday for a call on audio-visual search, 19 Dec 2005]
[invited talks: AMR 2005, National e-Science workshop 2005]
[invited panels: EWIMT 2004, AMR 2005, EWIMT 2005]
[with Howarth, *Information Retrieval,* submitted 2005]
[with Howarth CIVR 2005]
[with Chawda, Craft, Cairns and Heesch, HCI 2005]
[with Howarth, ECIR 2005]
[with Heesch, Pickering, Howarth and Yavlinsky, JCDL 2004]
[with Yavlinsky, Heesch and Pickering, ICASSP 2004]
[with Howarth, CIVR 2004]
[with Pickering, *Computer Vision and Image Understanding*, 92(1), 2003]
[with Heesch, Pickering and Yavlinsky, TRECVID 2003]
[with Heesch and Yavlinsky, CIVR 2003]
[with Pickering, Heesch, O'Callaghan and Bull, TREC 2002]
[with Pickering and Sinclair, CIVR 2002]
[with Pickering, TREC 2001]

Evaluation conference TRECVID

- TRECVID topic 102: Find shots from behind the pitcher in a baseball game as he throws a ball that the batter swings at

Outline

- **New Paradigms in Media Access**
  - Multimedia Management and Access
  - Semantic gap and Polysemy
  - Content-based visual search
  - **Lateral browsing**
  - Commercial applications
- Shape representation
  - 2D
  - 2.5D, images, motion
  - 3D

Putting the user in the look

- Precompute browsable network within media connecting ``similar'' stuff

[with Heesch, CIVR 2005]
[with Heesch, CIVR 2004]
[with Heesch, ECIR 2004]

Nearest neighbours wrt features

query

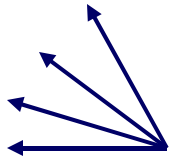
Structure

Colour

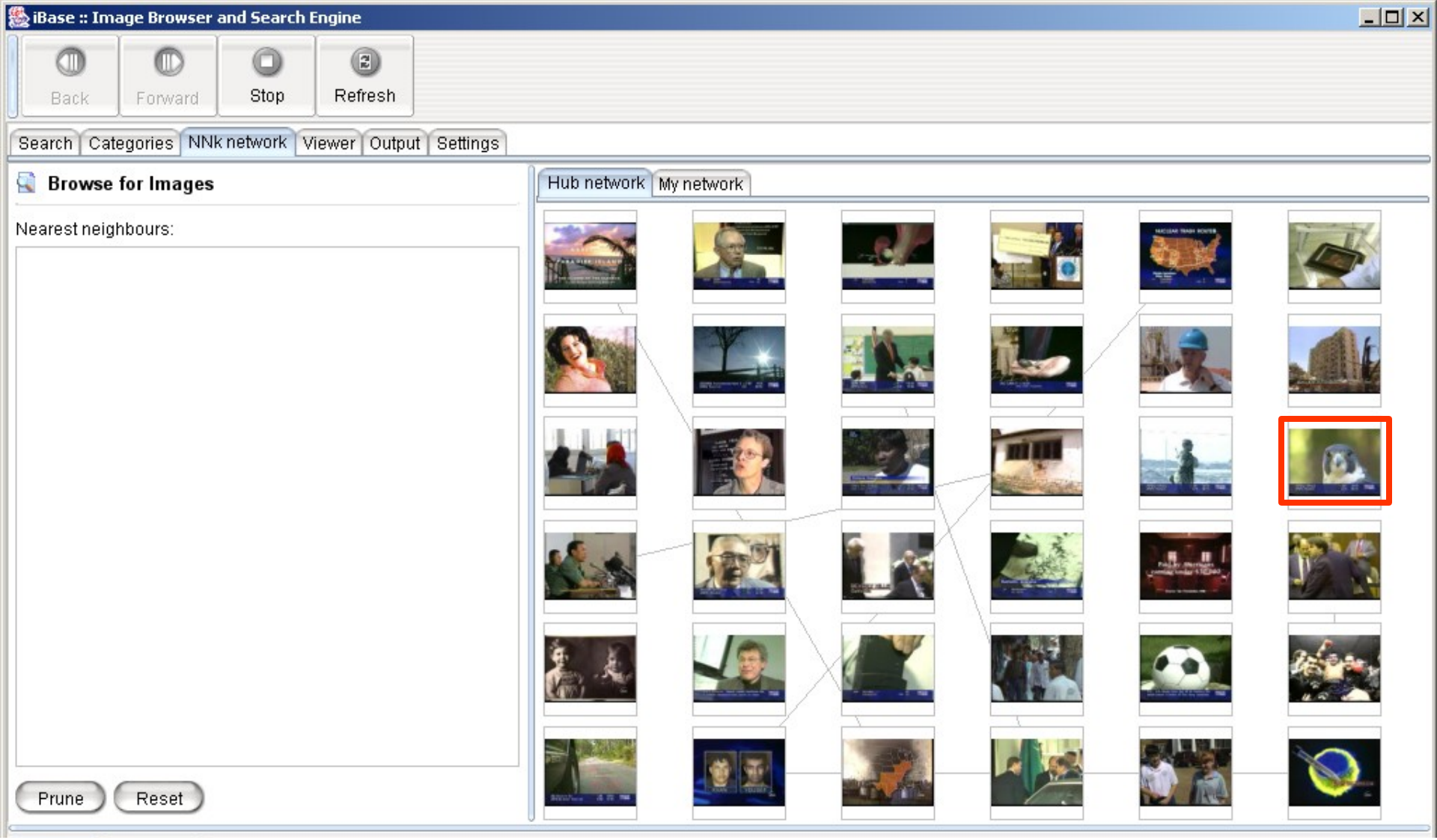
Text

Browsing example

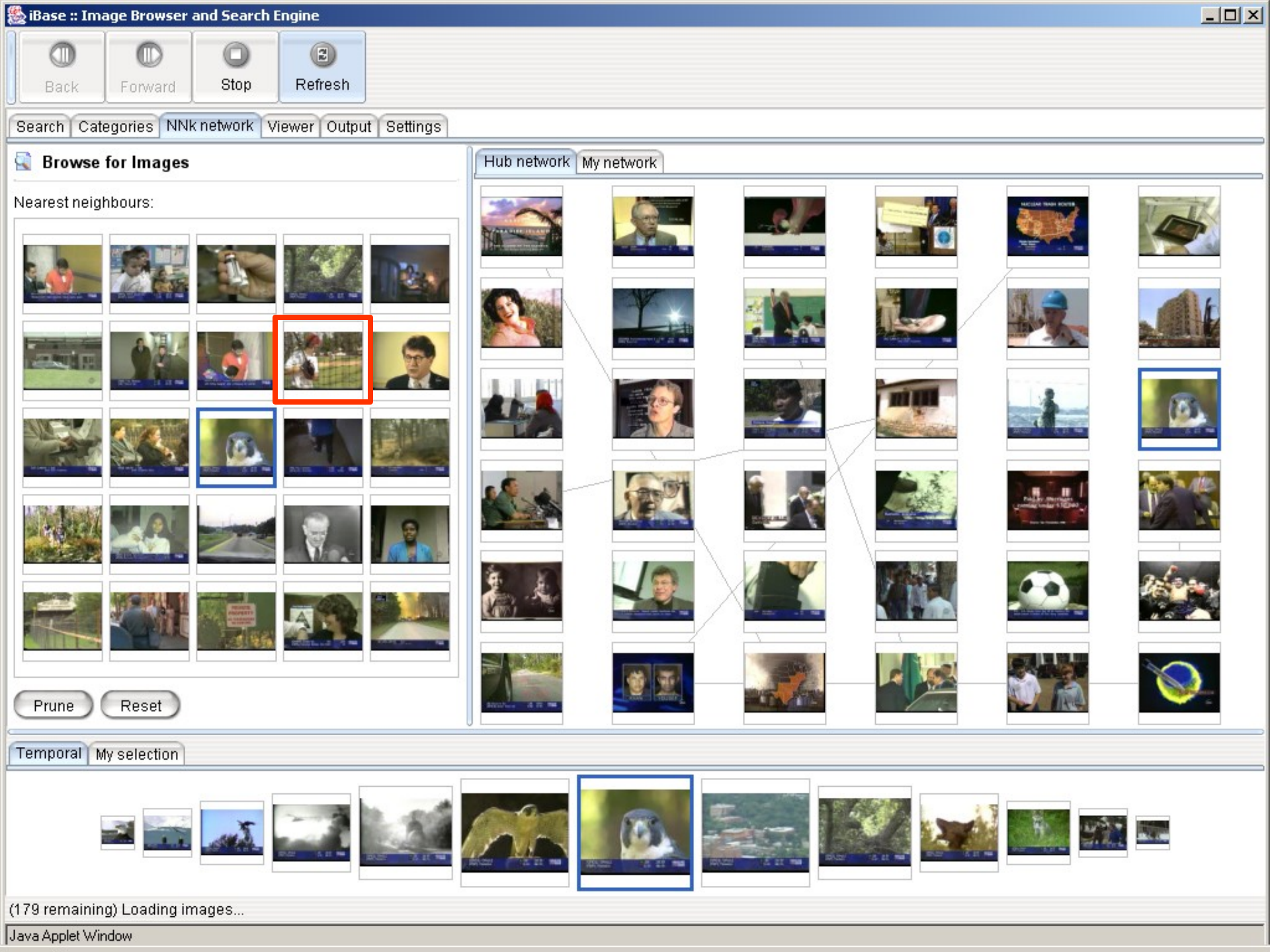
- TRECVID topic 102: Find shots from behind the pitcher in a baseball game as he throws a ball that the batter swings at



In your mind



[Alexander May: SET award 2004: best use of IT – **national prize**]


iBase :: Image Browser and Search Engine
Back
Forward
Stop
Refresh
Search
Categories
NNk network
Viewer
Output
Settings
Browse for Images
Nearest neighbours:
Prune
Reset
Hub network
My network
Temporal
My selection
(179 remaining) Loading images...
Java Applet Window

iBase :: Image Browser and Search Engine

Back | Forward | Stop | Refresh

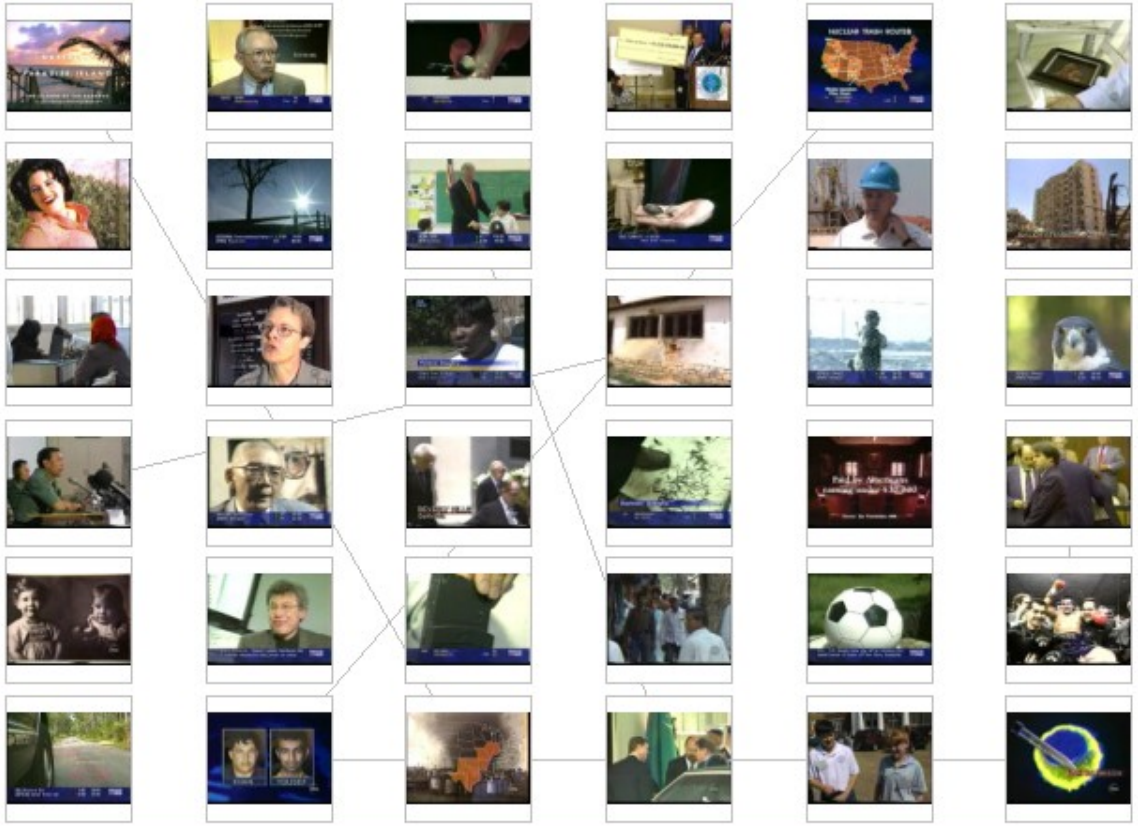
Search | Categories | NNk network | Viewer | Output | Settings

Browse for Images

Nearest neighbours:

Prune | Reset

Hub network | My network



Temporal | My selection

Image ID: 6635 Filename: TRECVID2003_156/shot156_308_RKF.jpg Keywords: THE UNDERDOG FIGHTING TO BE RECOGNIZED OR THE COMMITMENT THEY MADE TO EACH OTH...

Java Applet Window

iBase :: Image Browser and Search Engine

Back | Forward | Stop | Refresh

Search | Categories | NNk network | Viewer | Output | Settings

Image Properties

Image ID: 1263
Collection: Video
Resolution: 352 x 264
Filename: TRECVID2003_138/shot138_210_RKF.jpg

Keywords: FOR ONE WIN OVER THE SHARKS IN THE MAJORS THE BRAVES BEAT THE

Temporal | My selection

Image ID: 1263 Filename: TRECVID2003_138/shot138_210_RKF.jpg Keywords: FOR ONE WIN OVER THE SHARKS IN THE MAJORS THE BRAVES BEAT THE

Java Applet Window

Properties

- Exposure of semantic richness
- User decides which image meaning is correct
- Network precomputed -> fast and interactive
- Supports search without query formulation
- **3 degrees of separation** (for 32,000 images)
- **scale-free, small-world graph**

Outline

- **New Paradigms in Media Access**
  - Multimedia Management and Access
  - Semantic gap and Polysemy
  - Content-based visual search
  - Lateral browsing
  - **Commercial applications**
- **Shape representation**
  - 2D
  - 2.5D
  - 3D

Video summary

story-level segmentation

keyframe summary
videotext summary
named entities

full-text search

[technology being licensed by Imperial Innovations to industry]

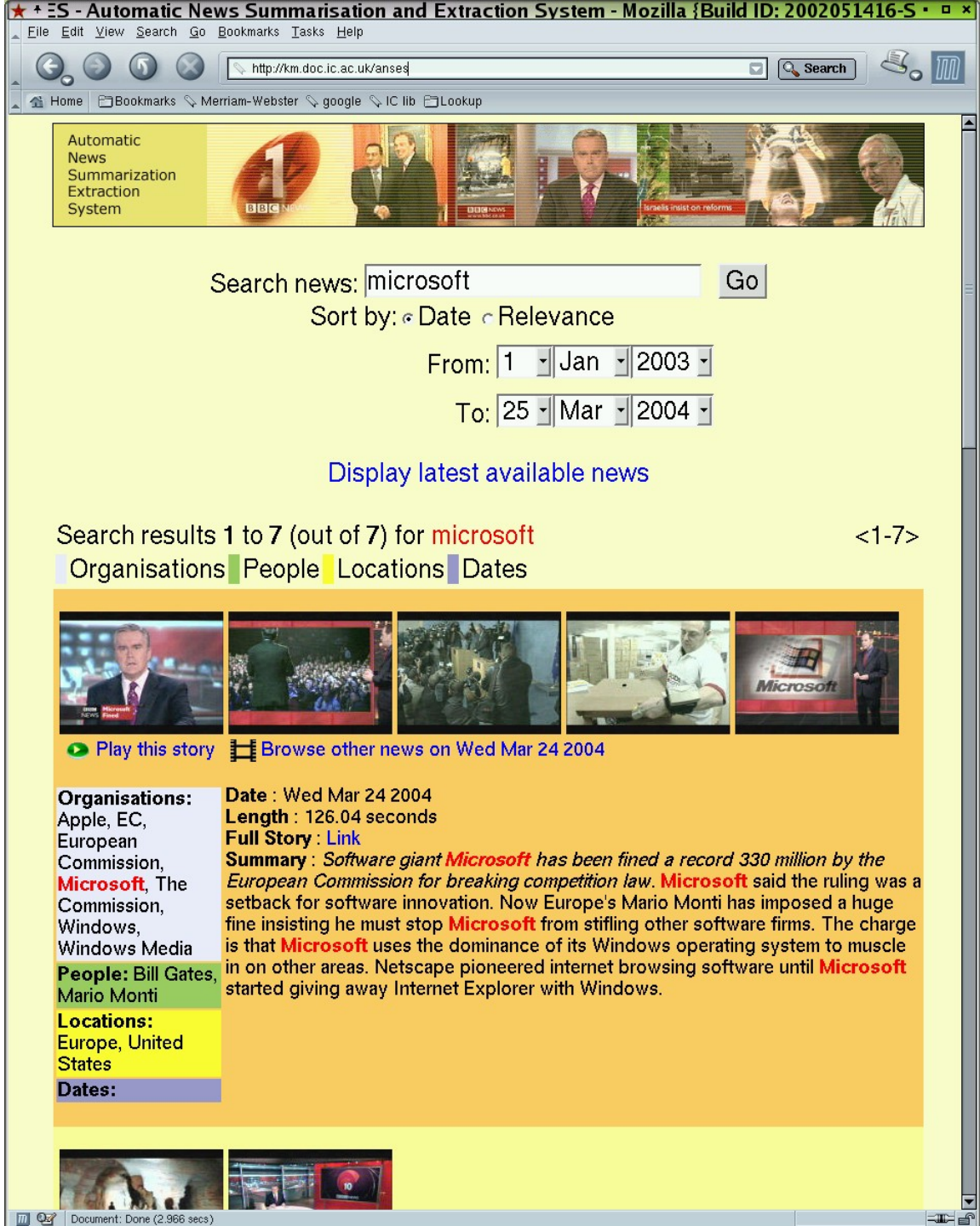
[patent application 2004]

[finished PhD project: Pickering, now at HTB, London]

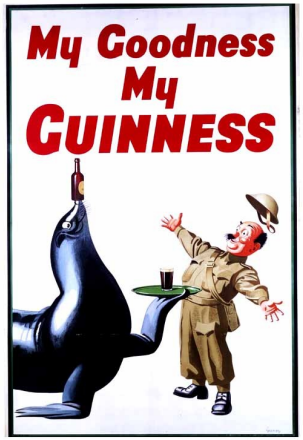
[with Wong and Pickering, CIVR 2003]

[with Lal, DUC 2002]

[Pickering: best UK CS student project 2000 – **national prize**]

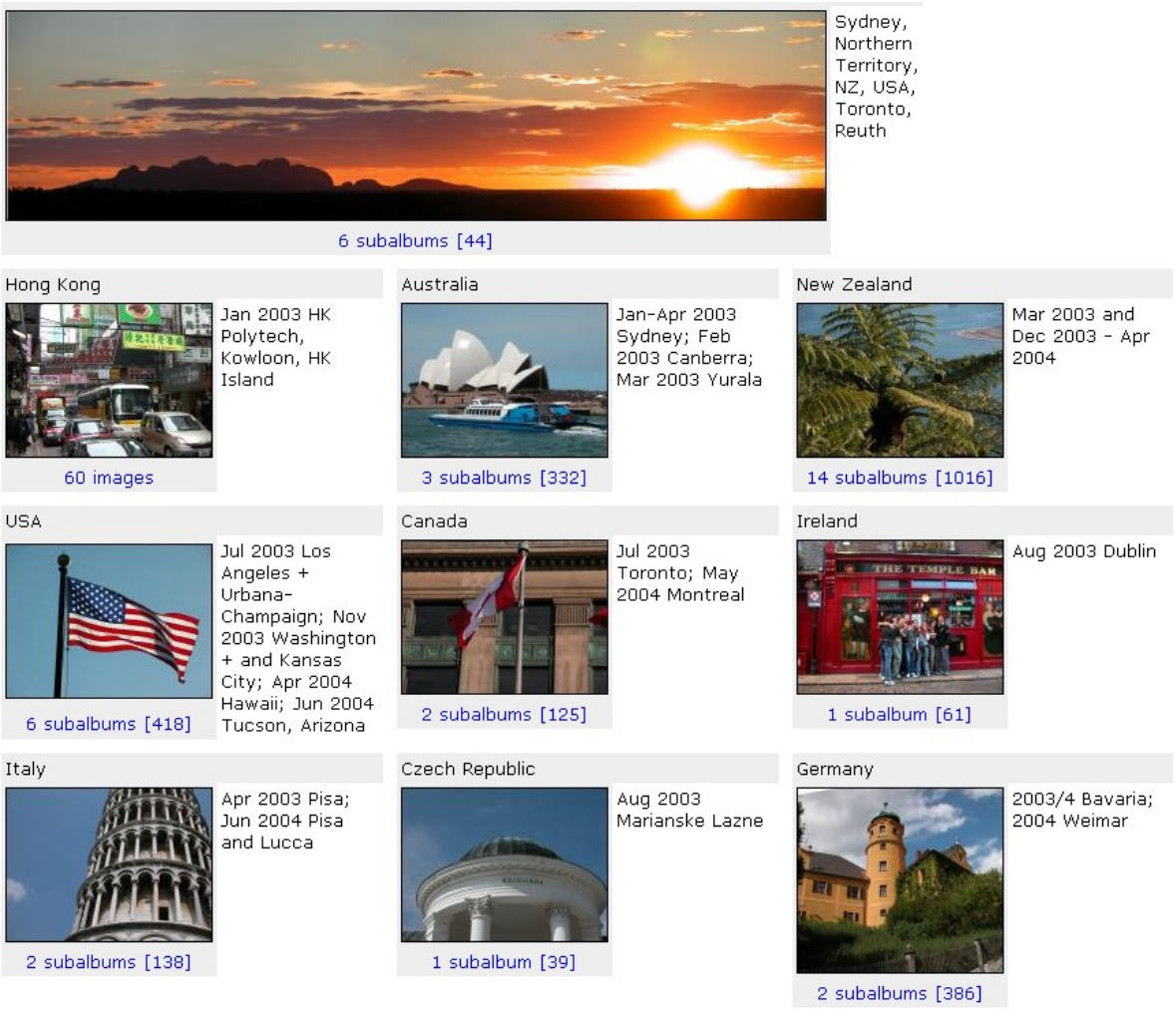


Multimedia Digital Libraries: Edutainment



Digital Shoebox

- Personal photo collection



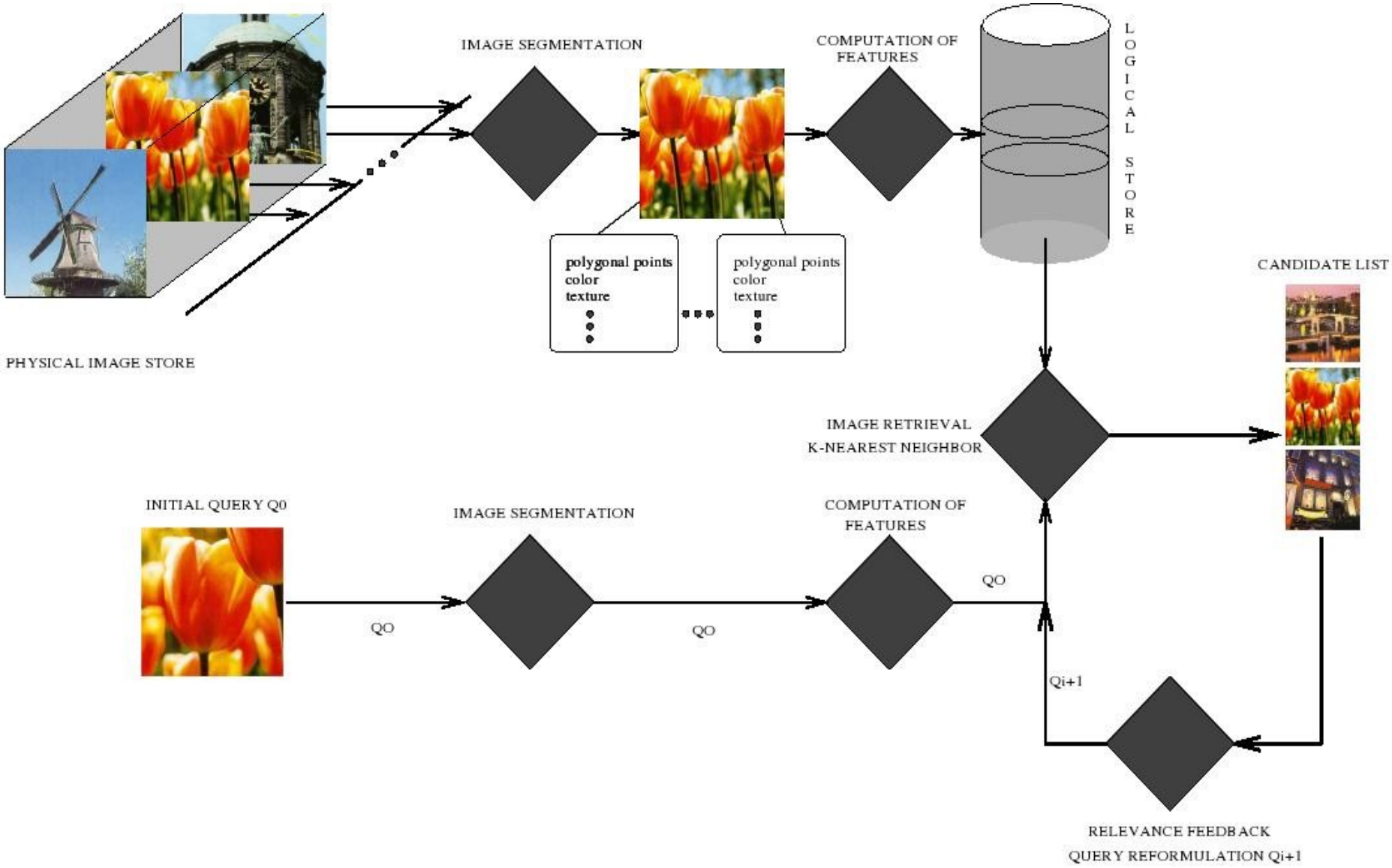
Multimedia Digital Libraries

- Visual, geo-temporal and text access to multimedia collections
- Novel visual browsing and search modes
- Information visualisation
- Collections from BBC, BL, V&A and 2 Unis

Outline

- New Paradigms in Media Access
  - Multimedia Management and Access
  - Semantic gap and Polysemy
  - Content-based visual search
  - Lateral browsing
  - Commercial applications
- **Shape representation**
  - 2D
  - 2.5D, images, motion
  - 3D

Content-Based Image Retrieval



From “Image search engines: an overview”, Gevers & Smeulders, 2003

Color

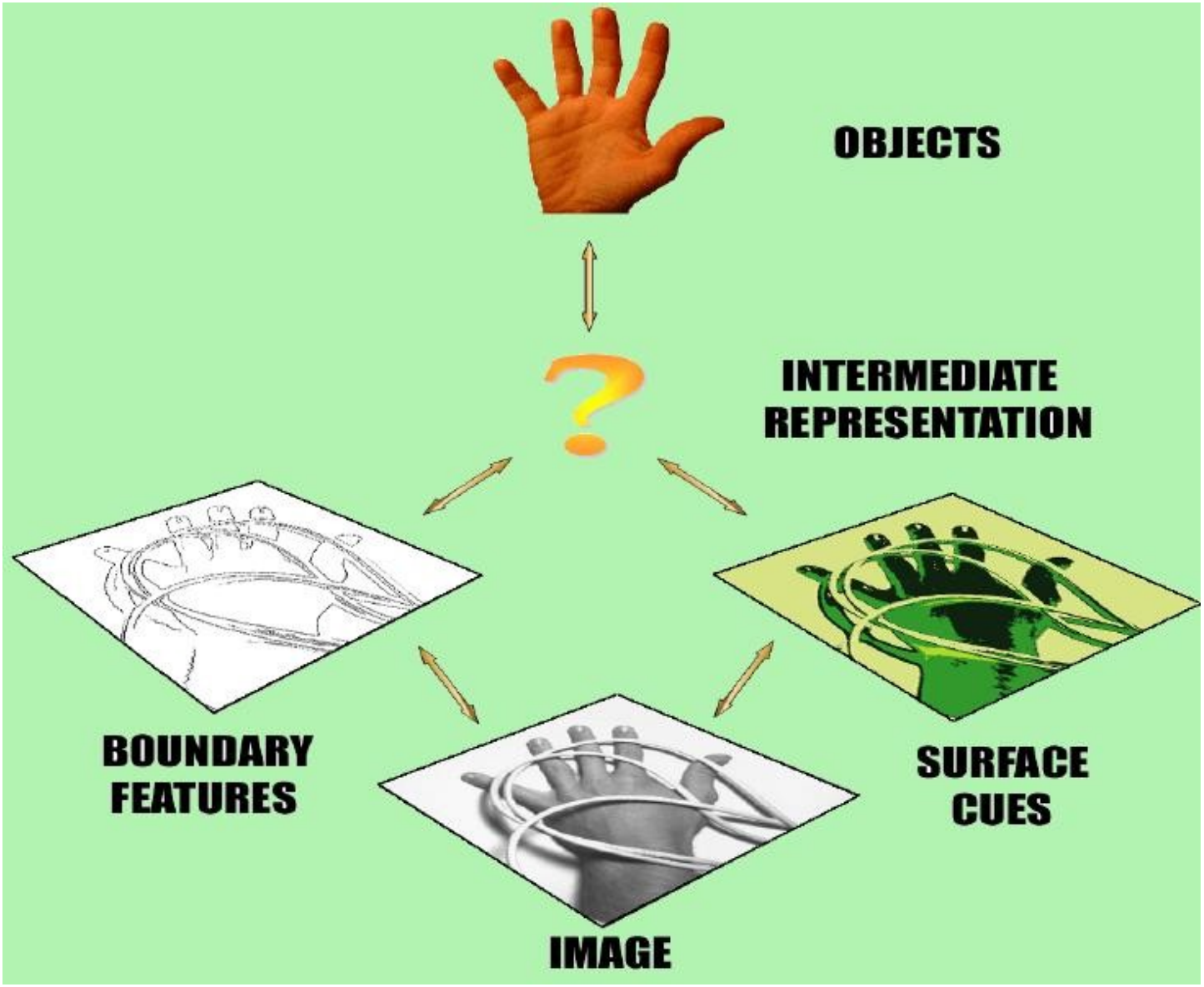
Shading

Shape

Texture

Motion

Shape as a Representation / Unification



From “On the Role of Medial Geometry in Human Vision”, Kimia, 2003.

Outline

- New Paradigms in Media Access
  - Multimedia Management and Access
  - Semantic gap and Polysemy
  - Content-based visual search
  - Lateral browsing
  - Commercial applications
- **Shape representation**
  - **2D**
  - 2.5D, images, motion
  - 3D

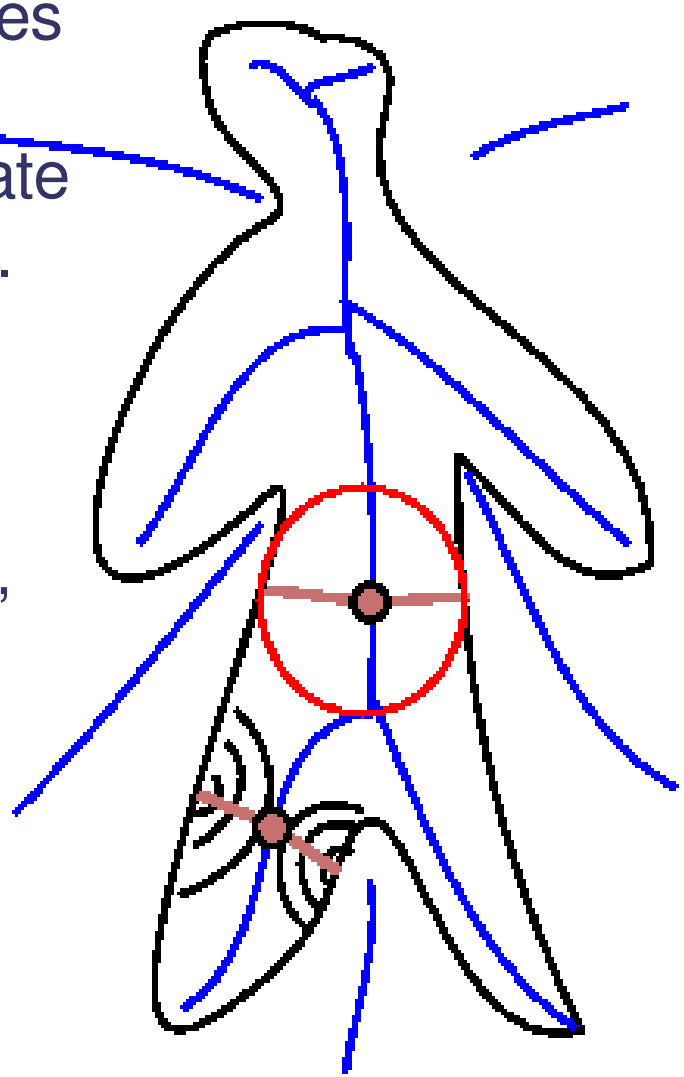
Medial representations of shape

1- Consider a set of inflatable primitives centered mid-way between 2 or more outline segments and grown to osculate (touch) the object on these segments.

Special case: **maximal contact spheres**.

Medial loci: let the outline "propagate" (i.e., continuously deform) until it self-intersects.

Special case: Euclidean **wave propagation** (optics).

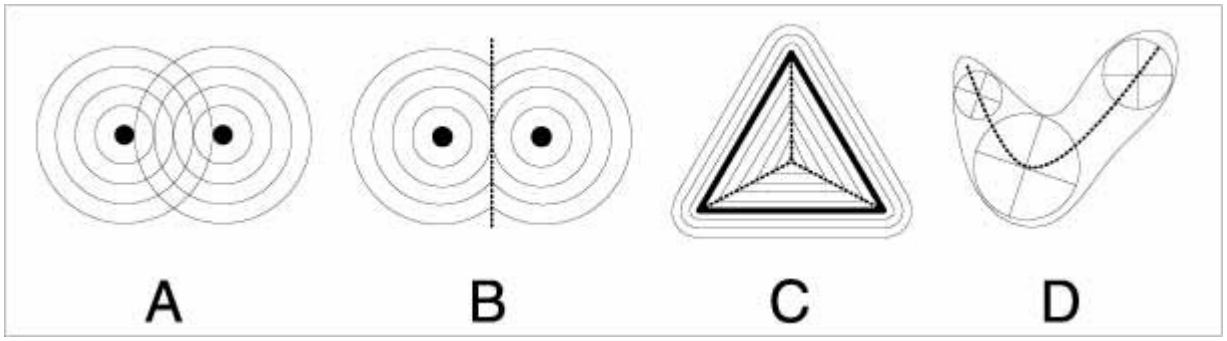


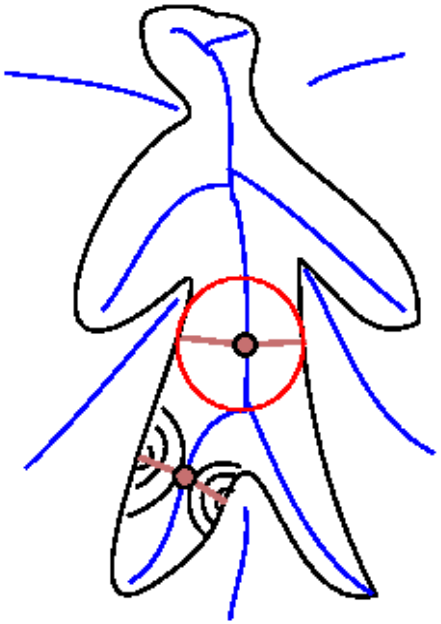
Medial representations of shape

The Medial Axis Transform: 1st singular solutions of the Eikonal equation of geometric optics [Blum].

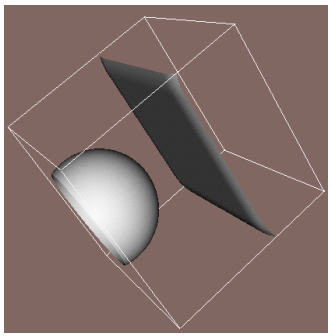
$$\sum_{i=1}^{n} \left(\frac{\partial u}{\partial x_i} \right)^2 = 1.$$

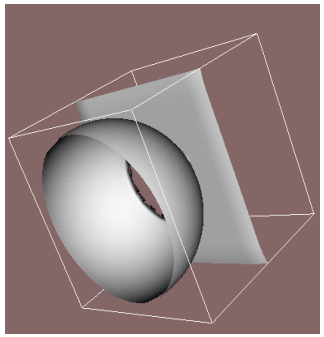
From Gert van Tonder's website:





3D cellular automata simulation
by Leymarie and Kimia:





Medial Axis representation of shape

Quantitative descriptions: encode the varying width of objects via the radius function (of contact spheres).

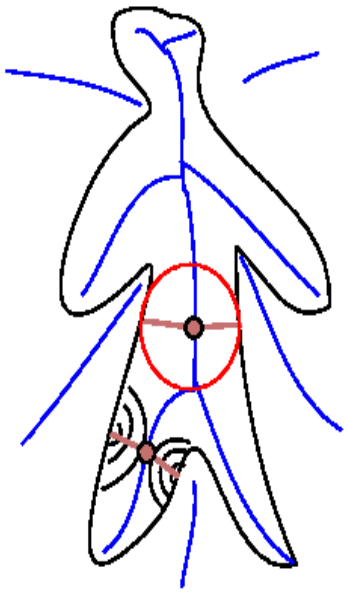
Qualitative descriptions: *e.g.*, branching structures of elongated objects.

Shape **features**: make explicit corners, bumps, surface ridges (sharp or smooth).

Completeness: exact **reconstruction** possible.

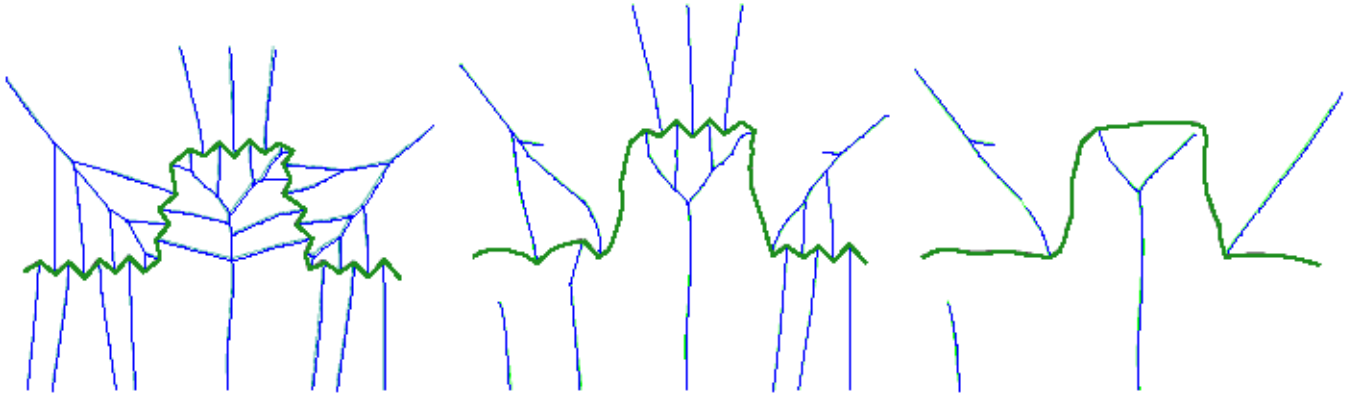
Built-in **hierarchy** of scales: mix of spatial and width properties.

Shape **dynamics**: modeling of deformations, kinematics (including animation, motion), genesis (growth).



2D graph “regularization”

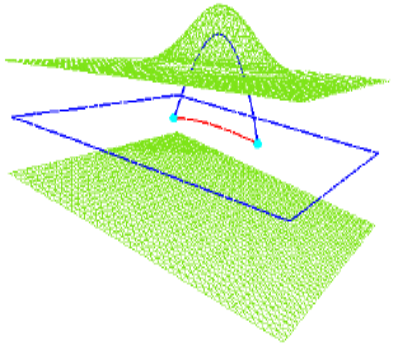
- Transition removal, *i.e.*, remove topological instability
- **2D Boundary Smoothing** ordered by “scale”

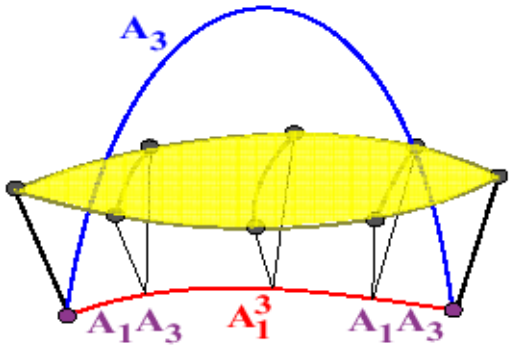


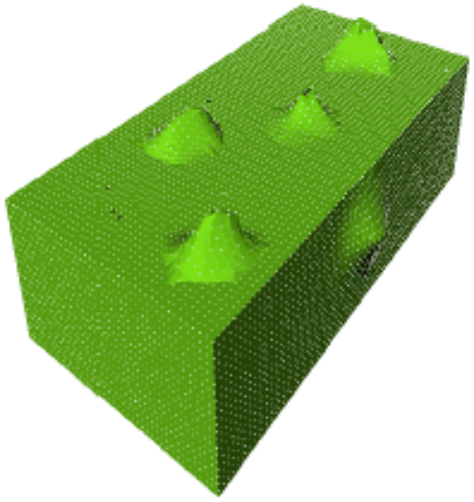
Iterative removal of MA branches, ordered by boundary support (*i.e.*, how much of the contour is represented), coupled with local boundary model adjustment, results in corner enhancement and small perturbations' smoothing.

(After Tek and Kimia, 2001)

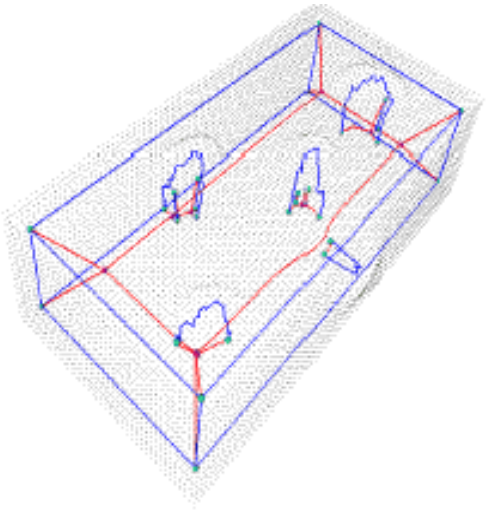
3D scaffold regularization



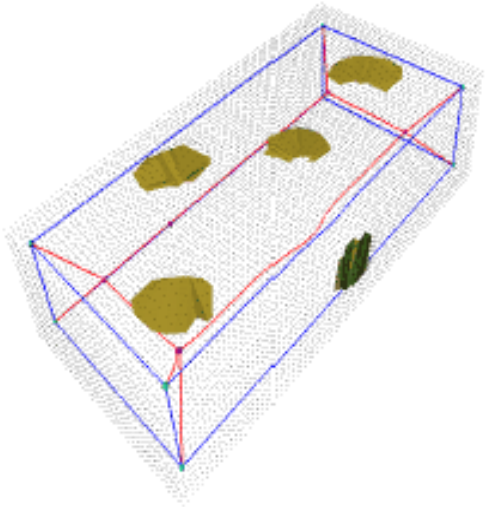




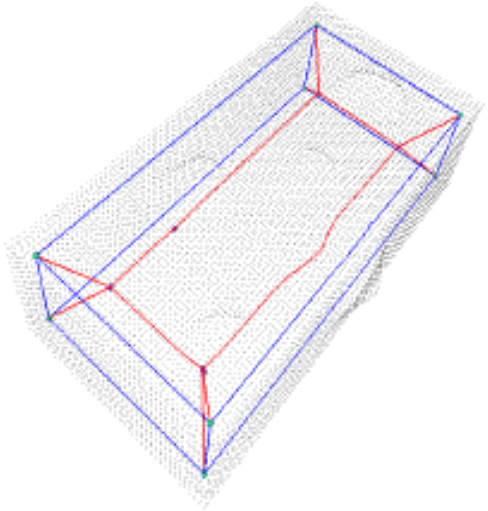
(f)



(g)

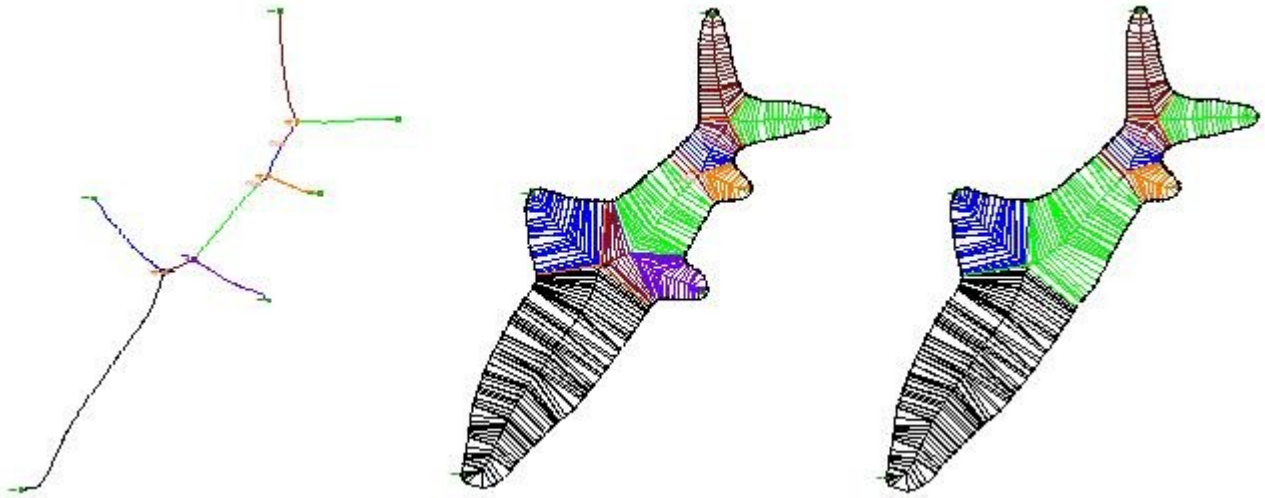


(h)

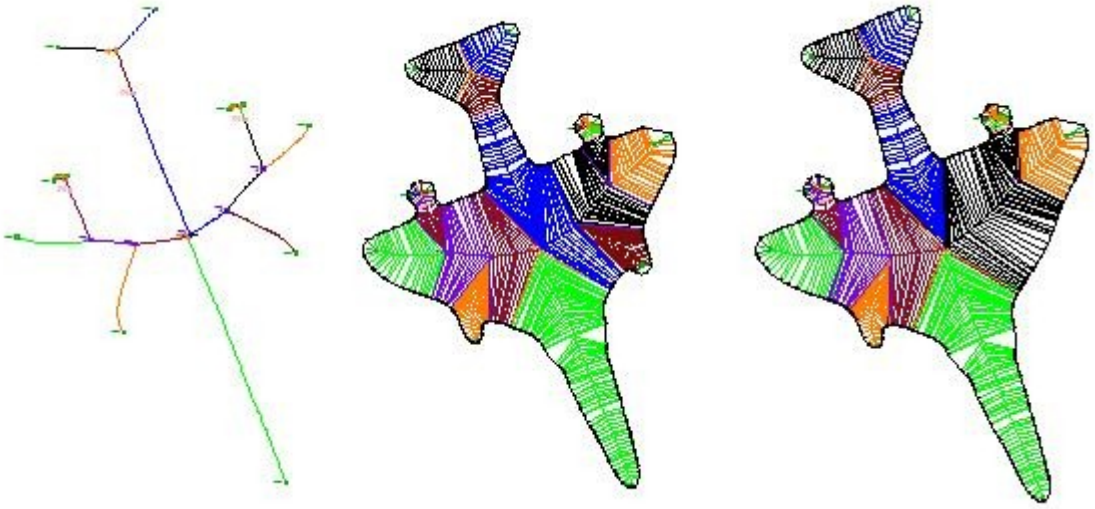


(i)

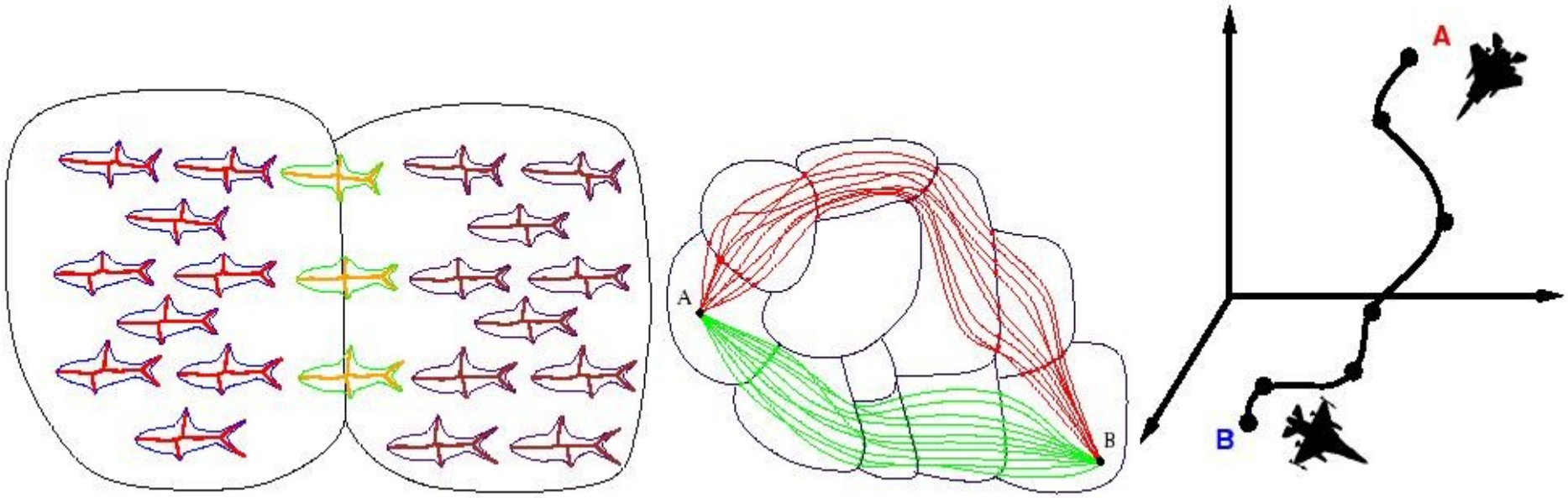
Examples: parts, deformation



Examples: parts, deformation

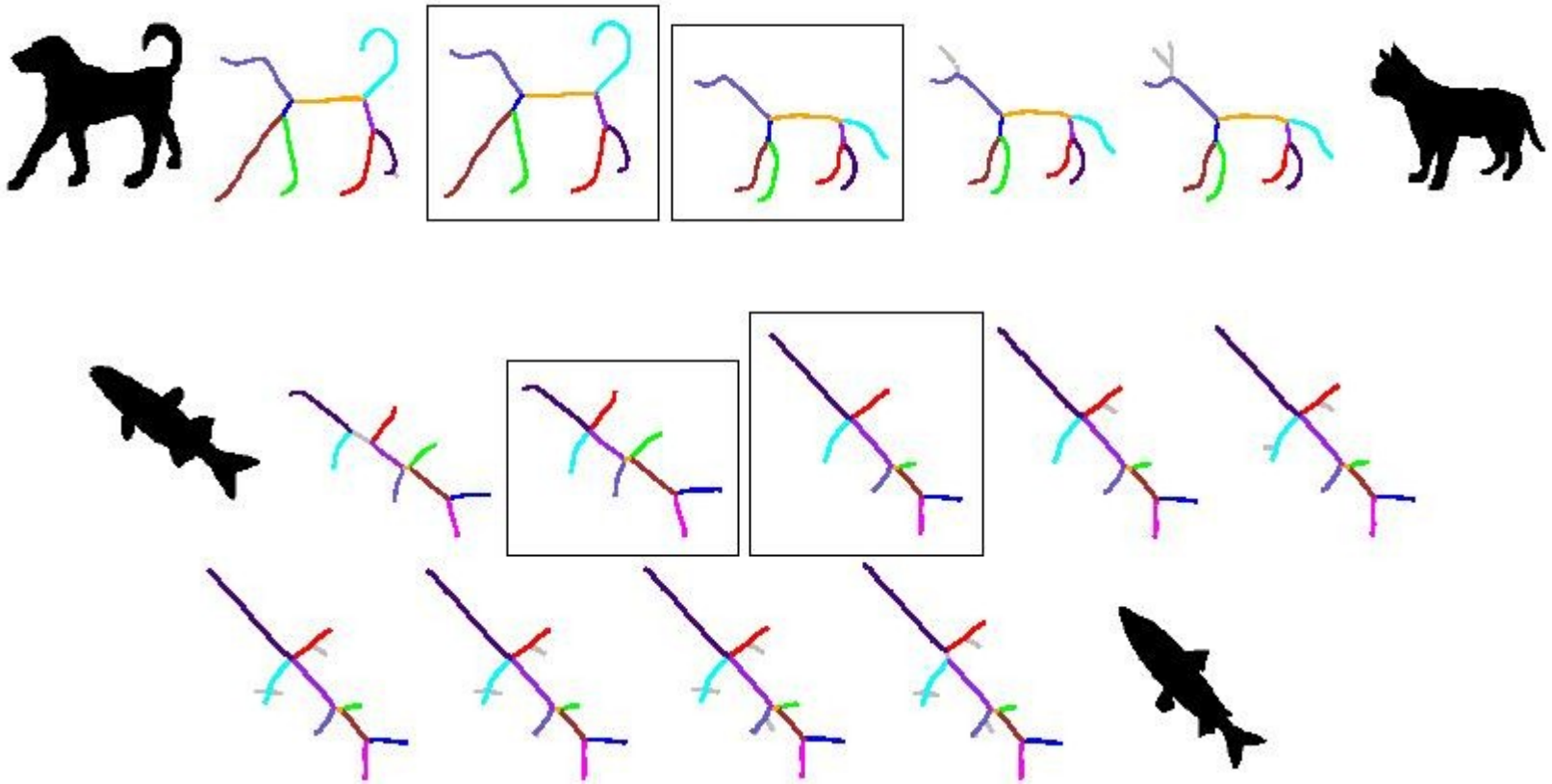


Challenge: How to navigate the “shape-space”



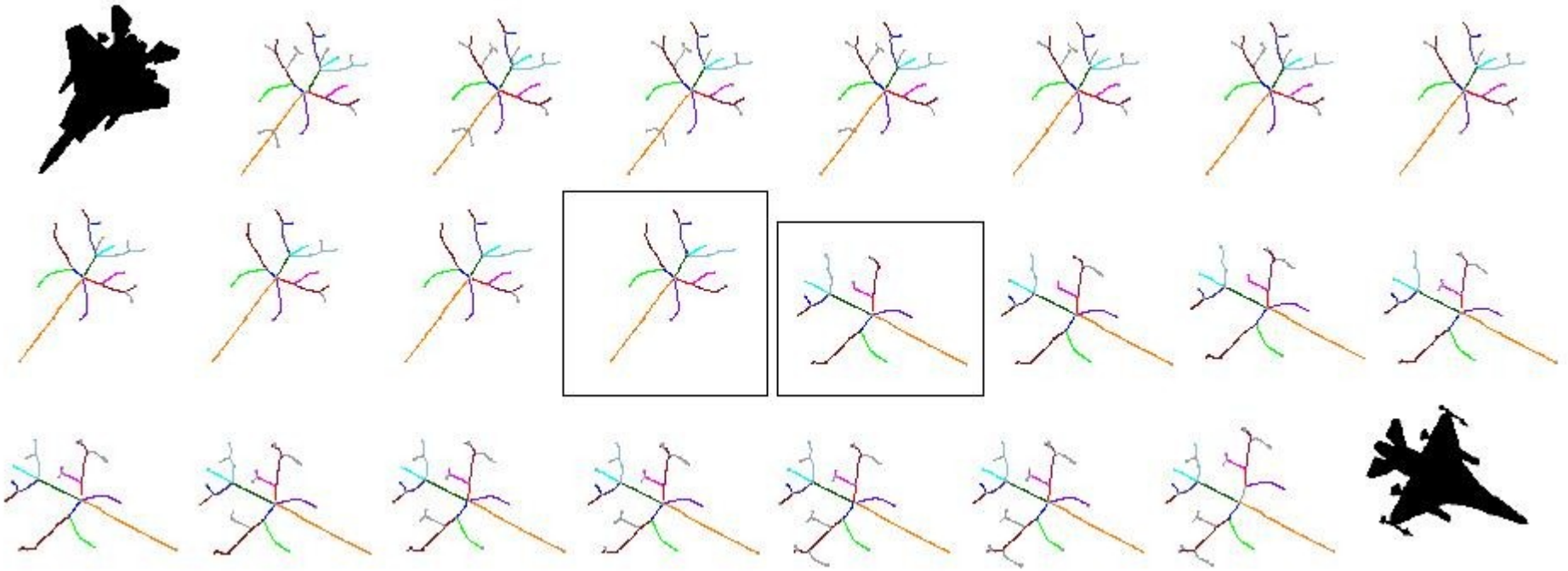
Challenge: How to navigate the “shape-space”

Build metrics on (graph) deformations

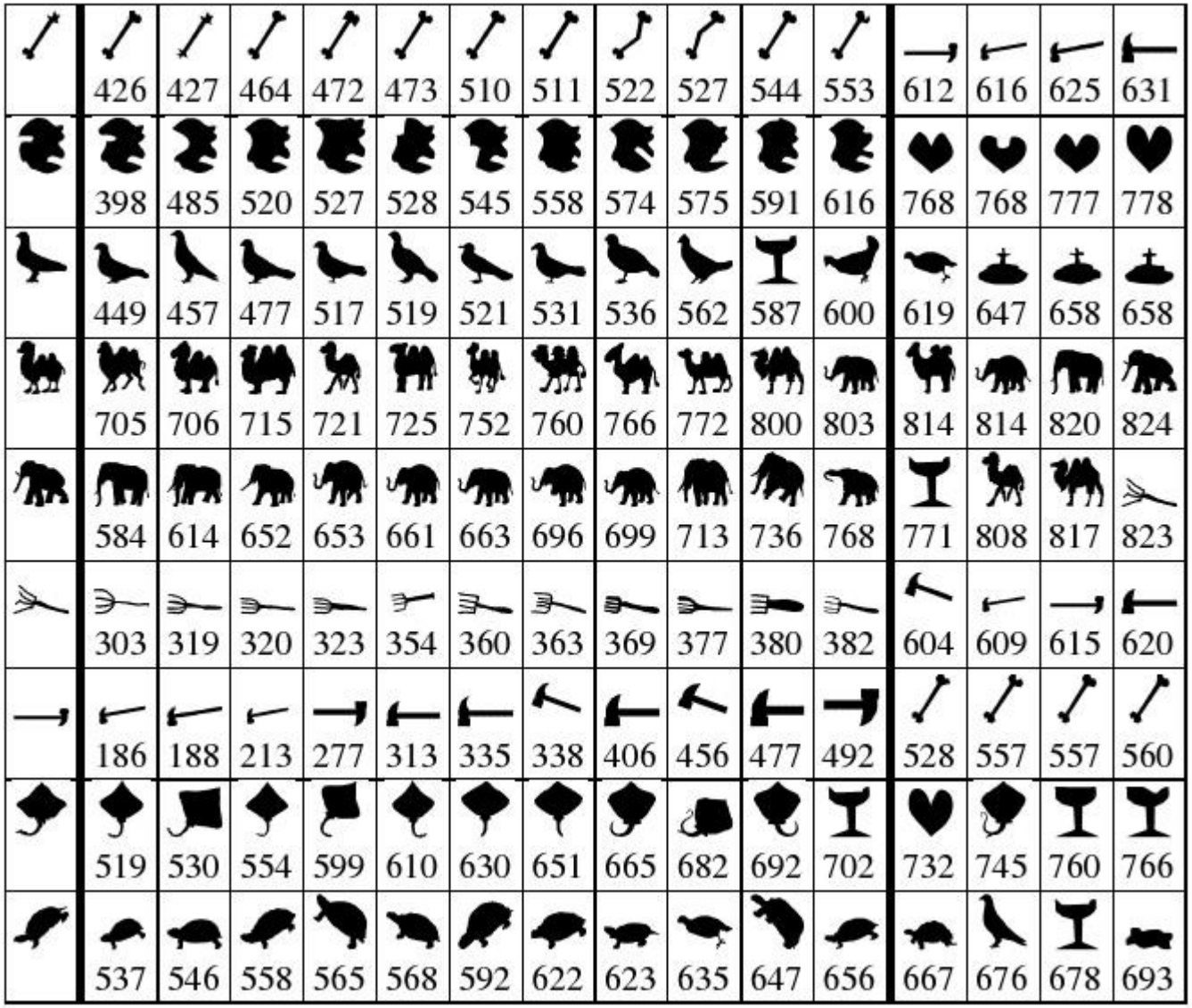


Challenge: How to navigate the “shape-space”

Build metrics on (graph) deformations



DB indexing by (2D) shape



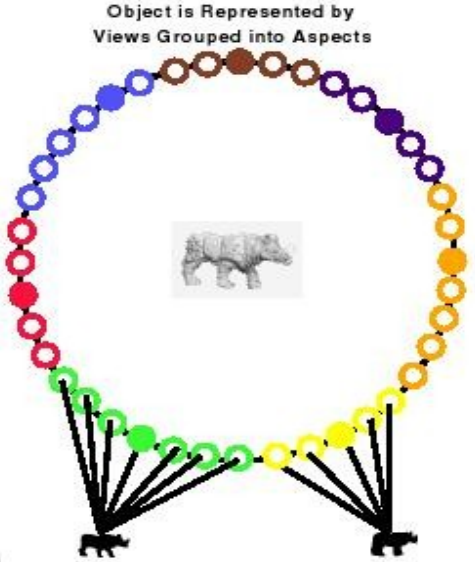
After Sebastian, Kimia, *et al.*, IEEE-PAMI, 2004.

Outline

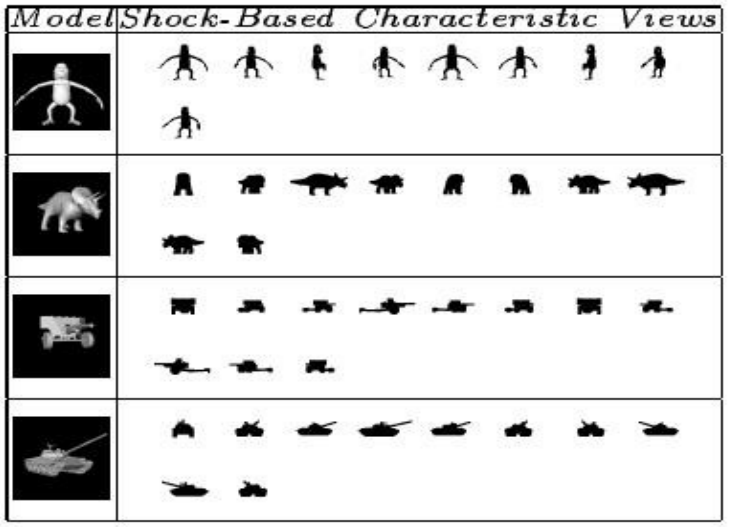
- New Paradigms in Media Access
  - Multimedia Management and Access
  - Semantic gap and Polysemy
  - Content-based visual search
  - Lateral browsing
  - Commercial applications
- **Shape representation**
  - 2D
  - **2.5D, images, motion**
  - 3D

DB indexing by (“2.5D”) shape

(a)



(b)



(c)

	Shock-Based Results		
Unknown	Match 1	Match 2	Match 3
Cost:	0.359	0.366	0.374
Cost:	0.279	0.286	0.288
Cost:	0.120	0.129	0.189
Cost:	0.158	0.279	0.279

(d)

After Kimia, *et al.*, 2000.

Examples: multiple views, animation

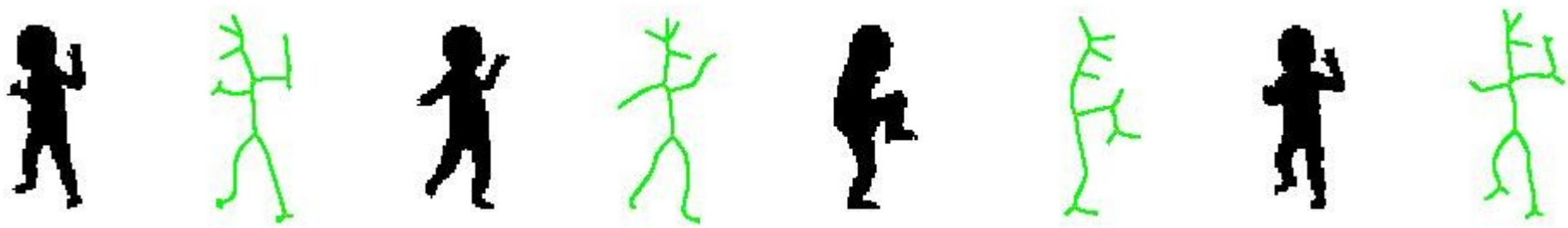
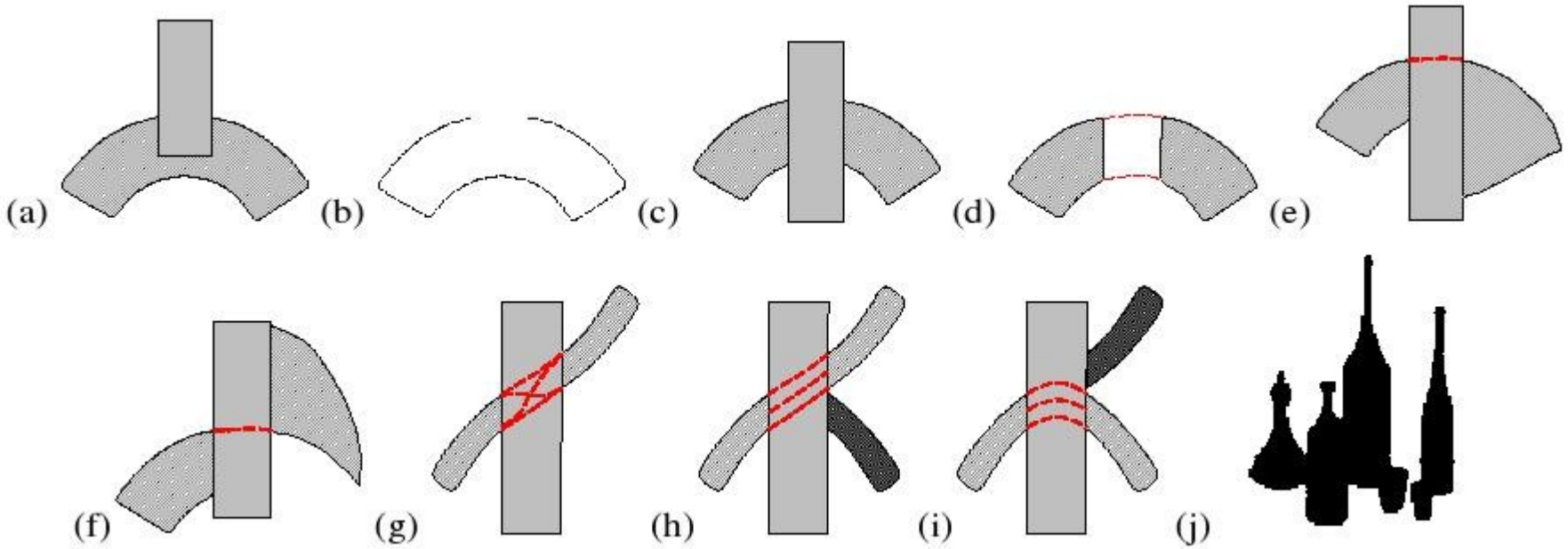


Image grouping / segmentation

Figure 12.15. Using the edge map associated with an image, visual fragments are defined and are used as canonical elements for perceptual grouping Tamrakar and Kimia (2004).

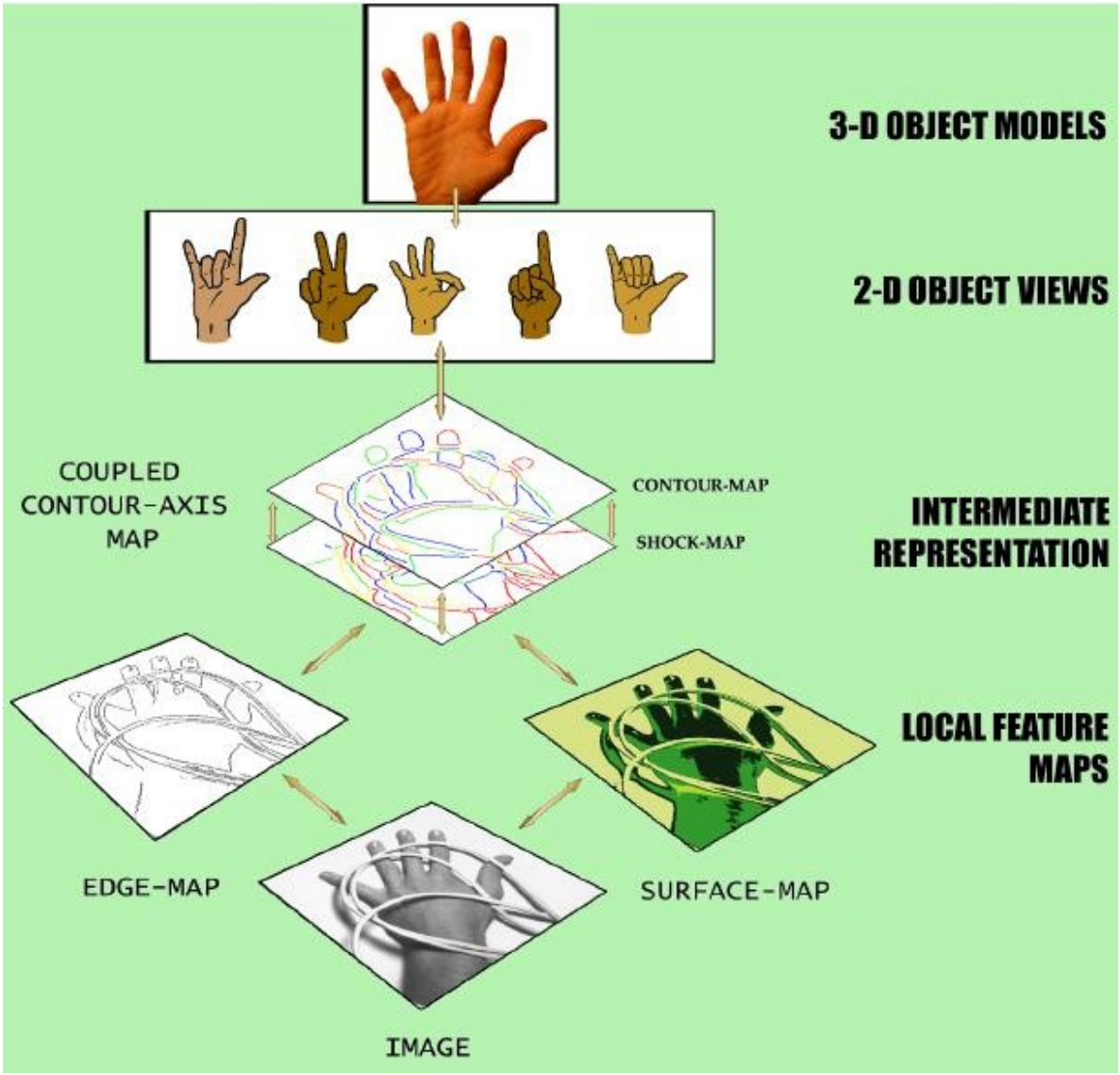
After Kimia *et al.*, 2004.

Image grouping / segmentation



After Kimia *et al.*, 2004.

Shape as a Representation / Unification



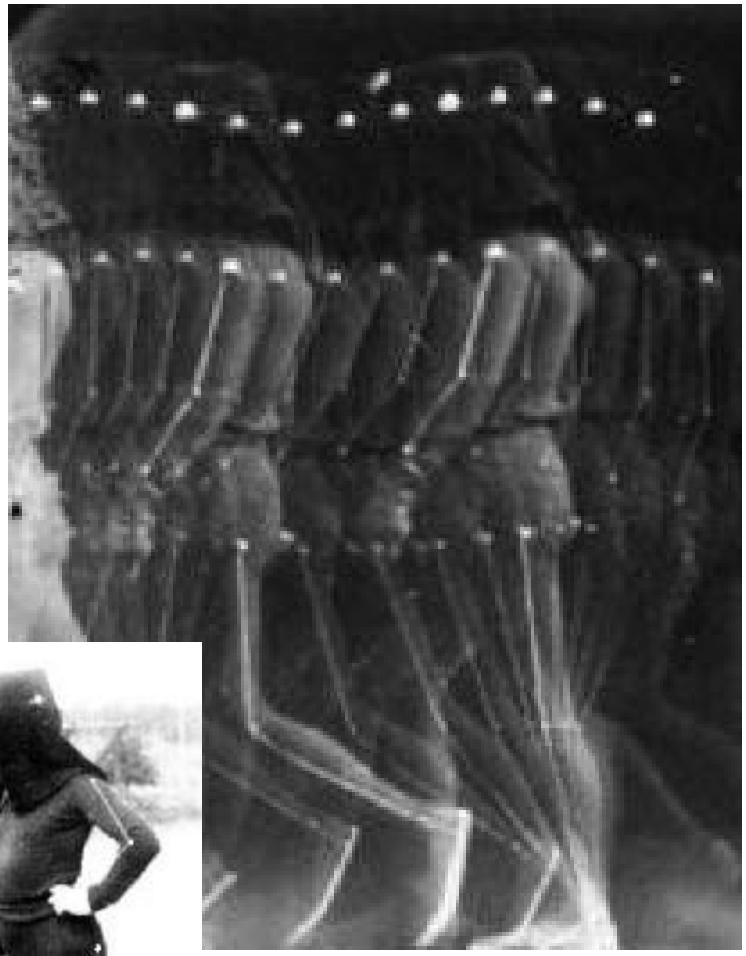
From “On the Role of Medial Geometry in Human Vision”, Kimia, 2003

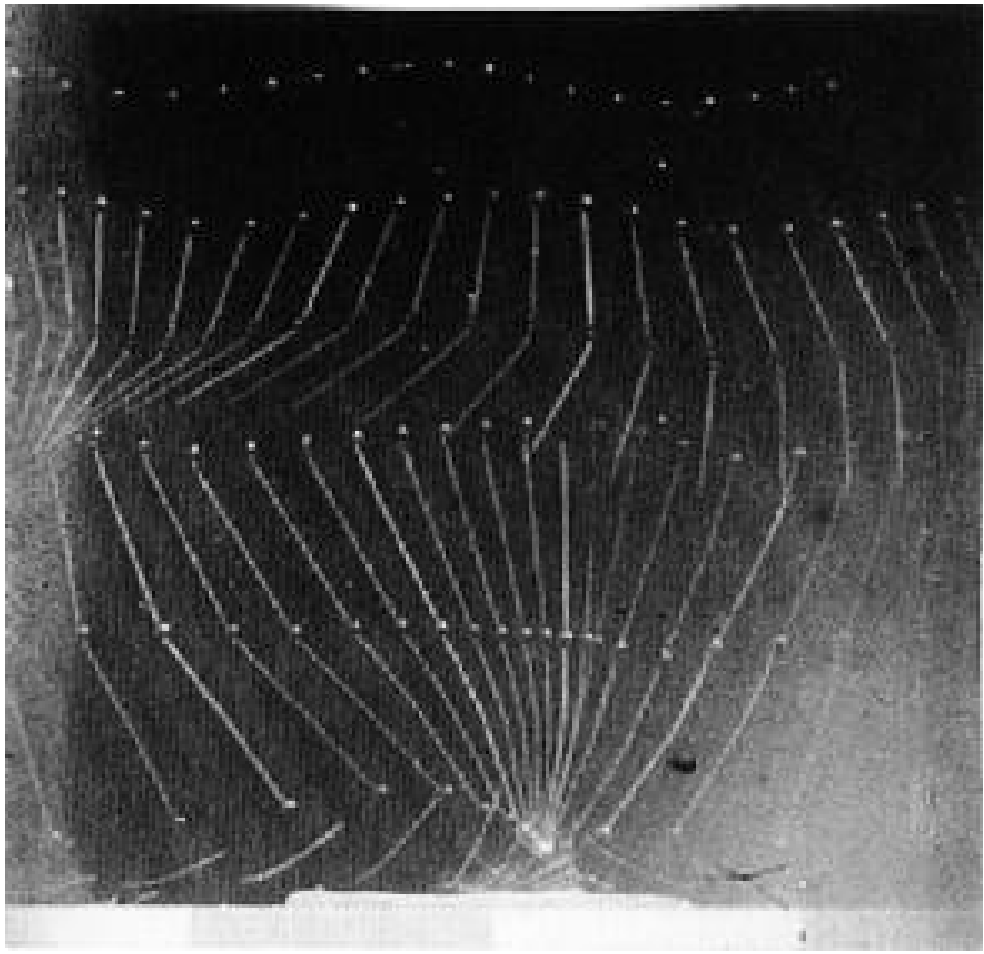
How to represent complex moving forms

Animation by Charles Lucassen, after Muybridge (1887)

From “Capturing time”, Ilona Kovacks, 2006

How to represent complex moving forms

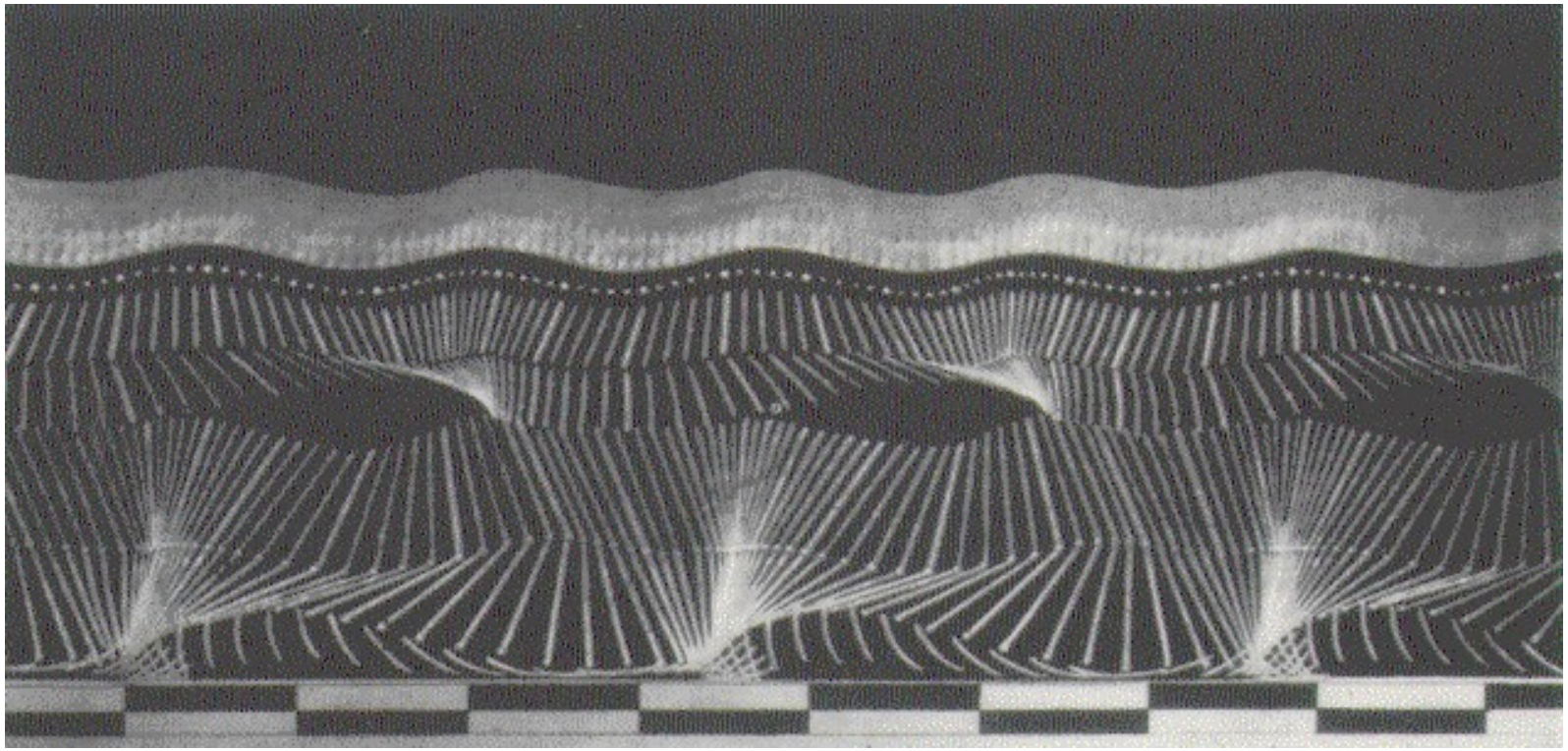




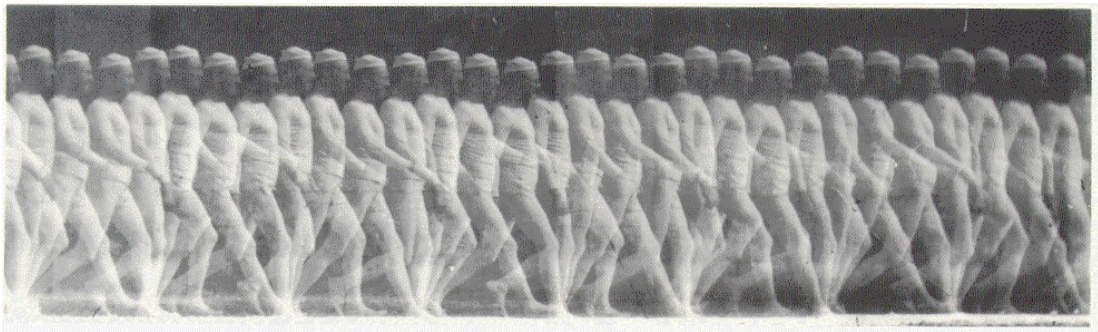
Chronophotographs, Marey, 1884, College de France

From “Capturing time”, Ilona Kovacks, 2006

How to represent complex moving forms



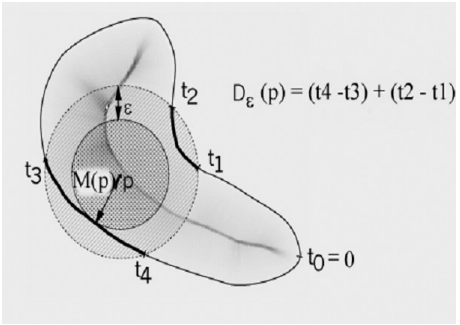
Marey, 1883, College de France

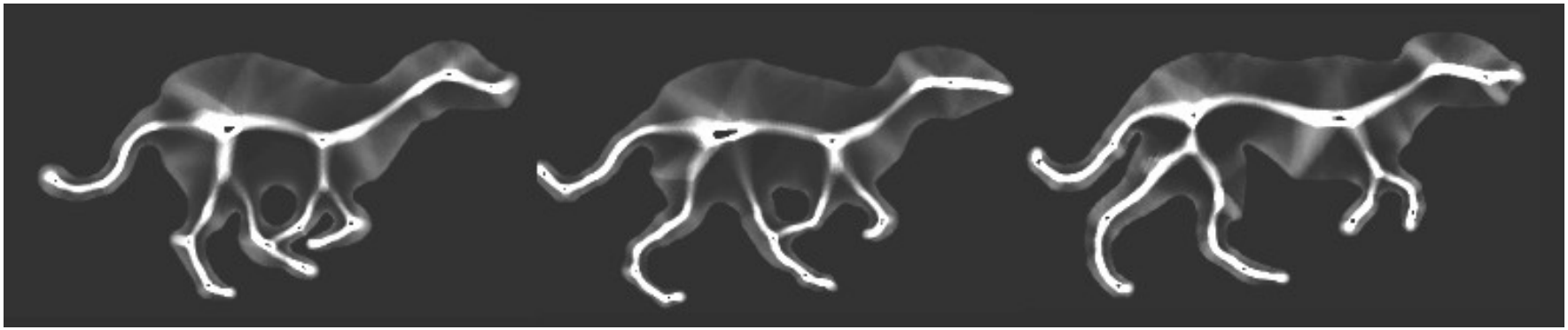


From “Capturing time”, Ilona Kovacks, 2006

How to represent complex moving forms

- human contrast sensitivity
- neuronal responses in monkey occipital cortex



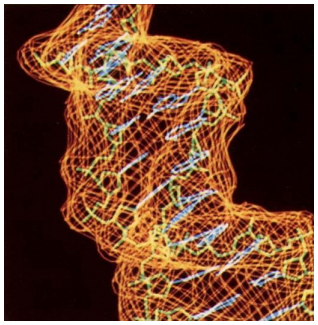
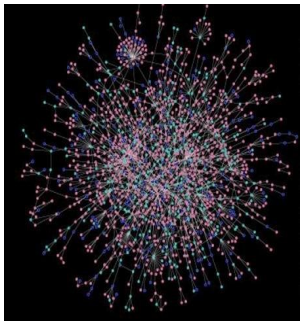
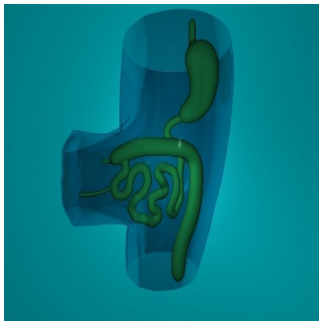
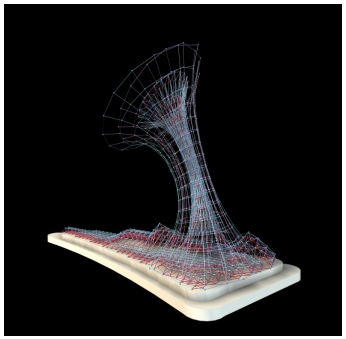
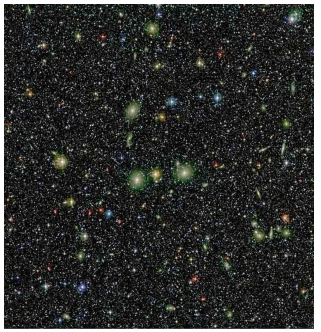


From “Capturing time”, Ilona Kovacks, 2006

Outline

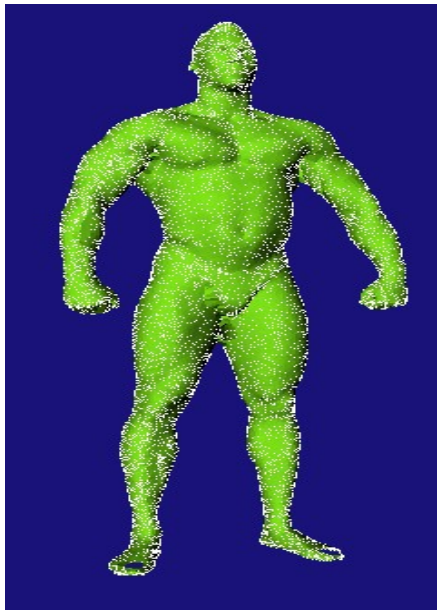
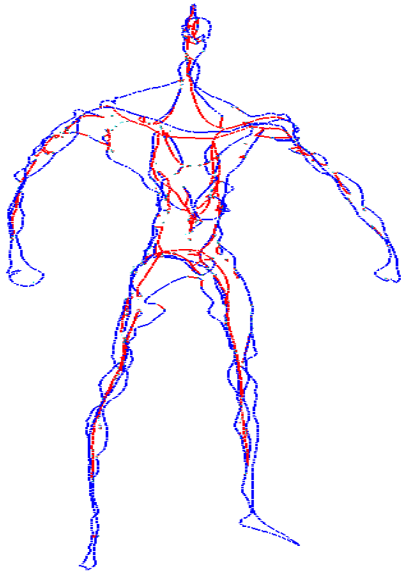
- New Paradigms in Media Access
  - Multimedia Management and Access
  - Semantic gap and Polysemy
  - Content-based visual search
  - Lateral browsing
  - Commercial applications
- **Shape representation**
  - 2D
  - 2.5D, images, motion
  - **3D**

3D

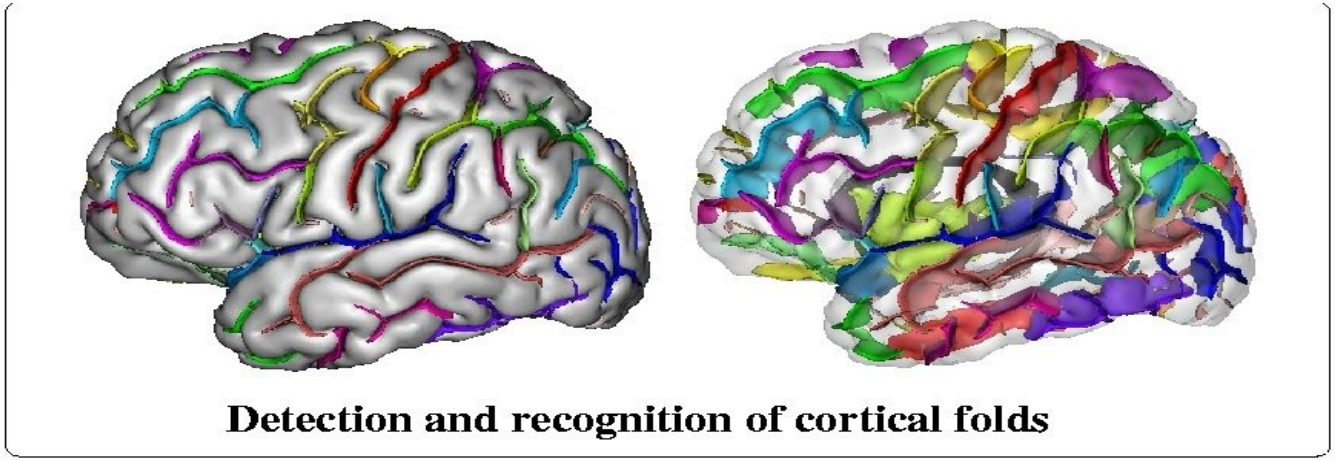






www.doc.gold.ac.uk/morpholingua/

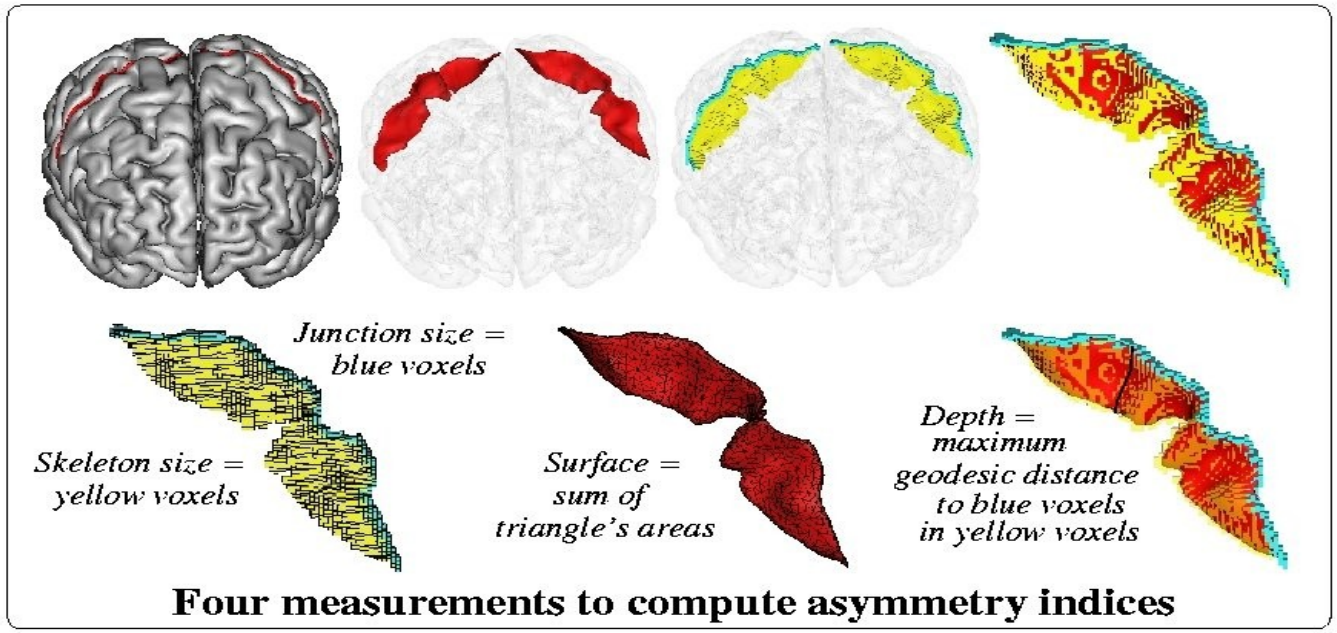
Leymarie-Kimia:
Keep only **singular** points of the flow (radius) to build a **graph** (2001).

3D



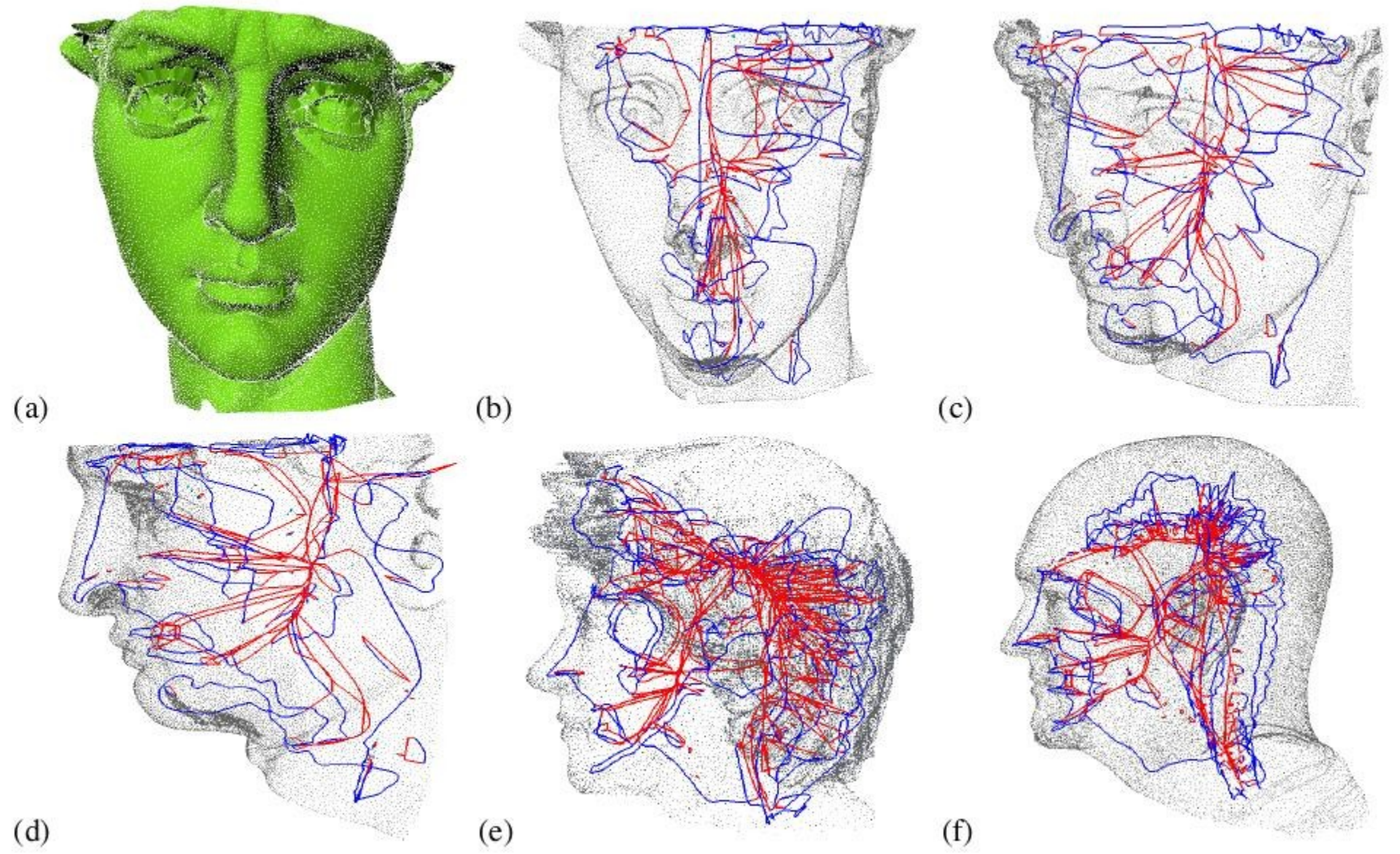
Detection and recognition of cortical folds



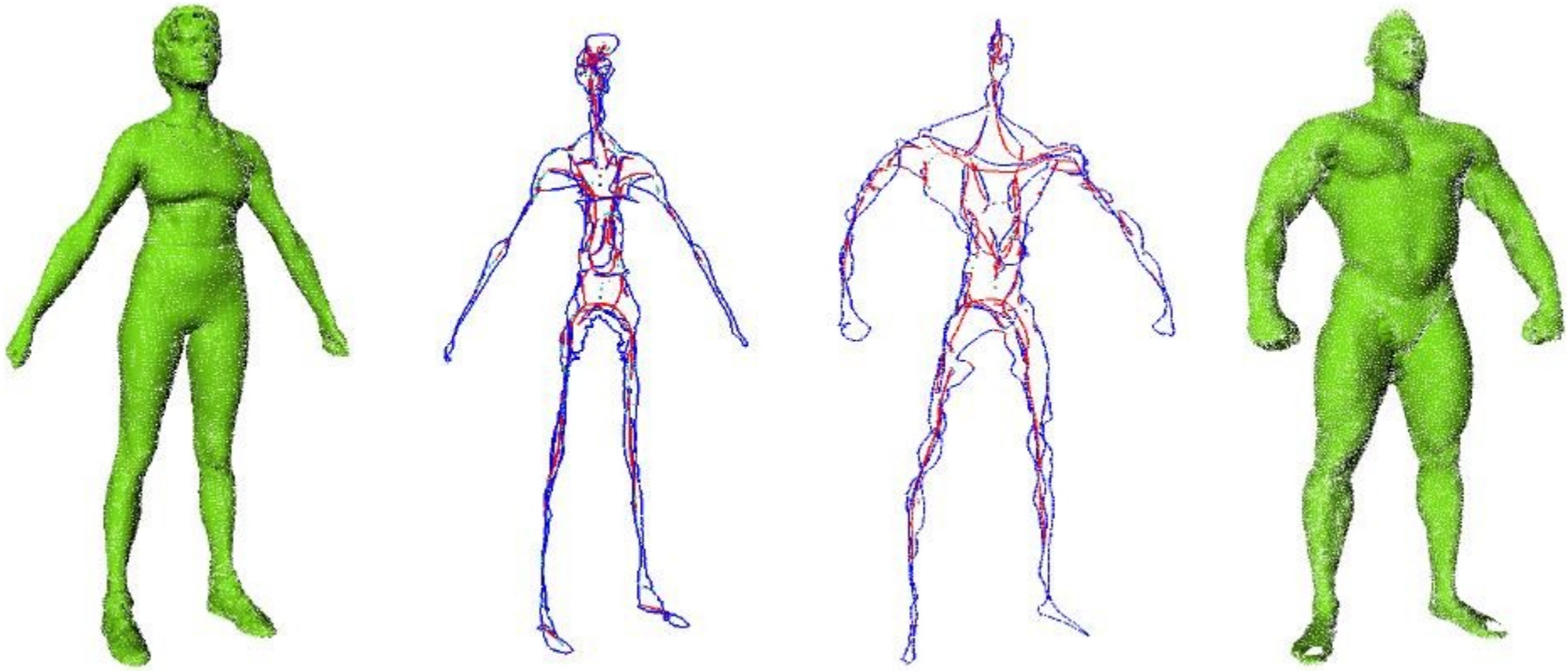
Four measurements to compute asymmetry indices

Mangin *et al.*, 2004; Cortical folds' recognition.

3D : Towards Object Recognition



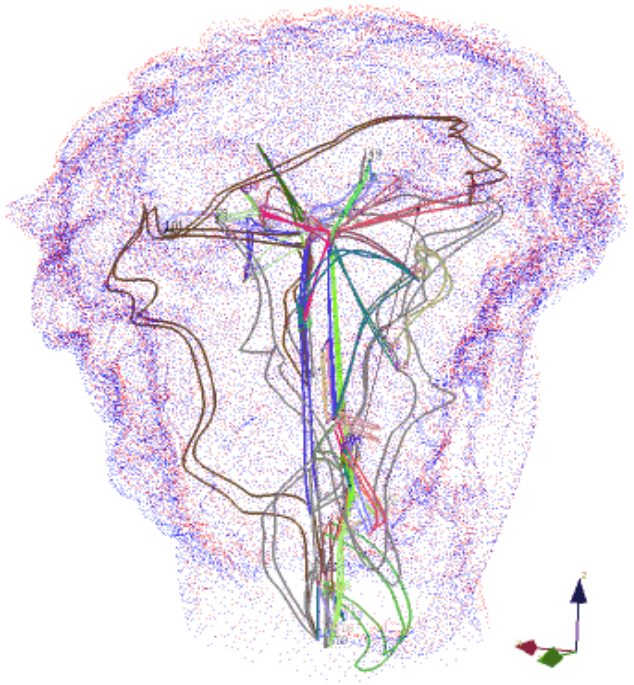
3D : Towards Object Recognition



Visual Search in large 3D DB

Graph matching via graduated assignement
(presented at 3DPVT, Greece, Sept. 2004, Chang, Leymarie & Kimia)
a solution to the **Global Registration** problem.

Digital Michelangelo
Stanford, Firenze, NRC



Challenge: robust automatic extraction of graphs,
dealing with topological instabilities/events.

Outline

- New Paradigms in Media Access
  - Multimedia Management and Access
  - Semantic gap and Polysemy
  - Content-based visual search
  - Lateral browsing
  - Commercial applications
- Shape representation
  - 2D
  - 2.5D, images, motion
  - 3D

What next ...?

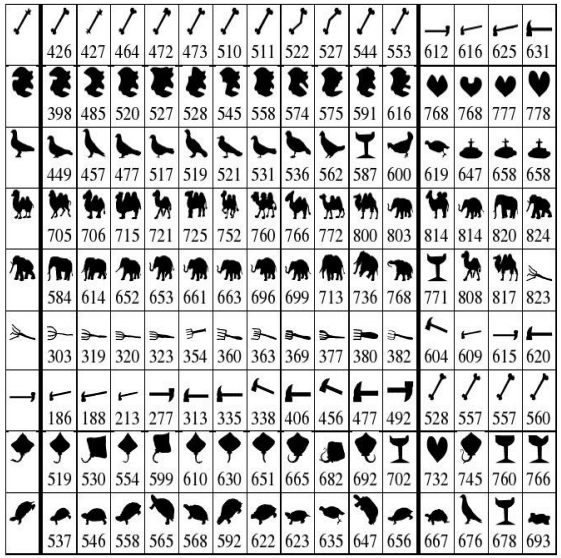
iBase :: Image Browser and Search Engine

Back | Forward | Stop | Refresh

Search | Categories | NNk network | Viewer | Output | Settings

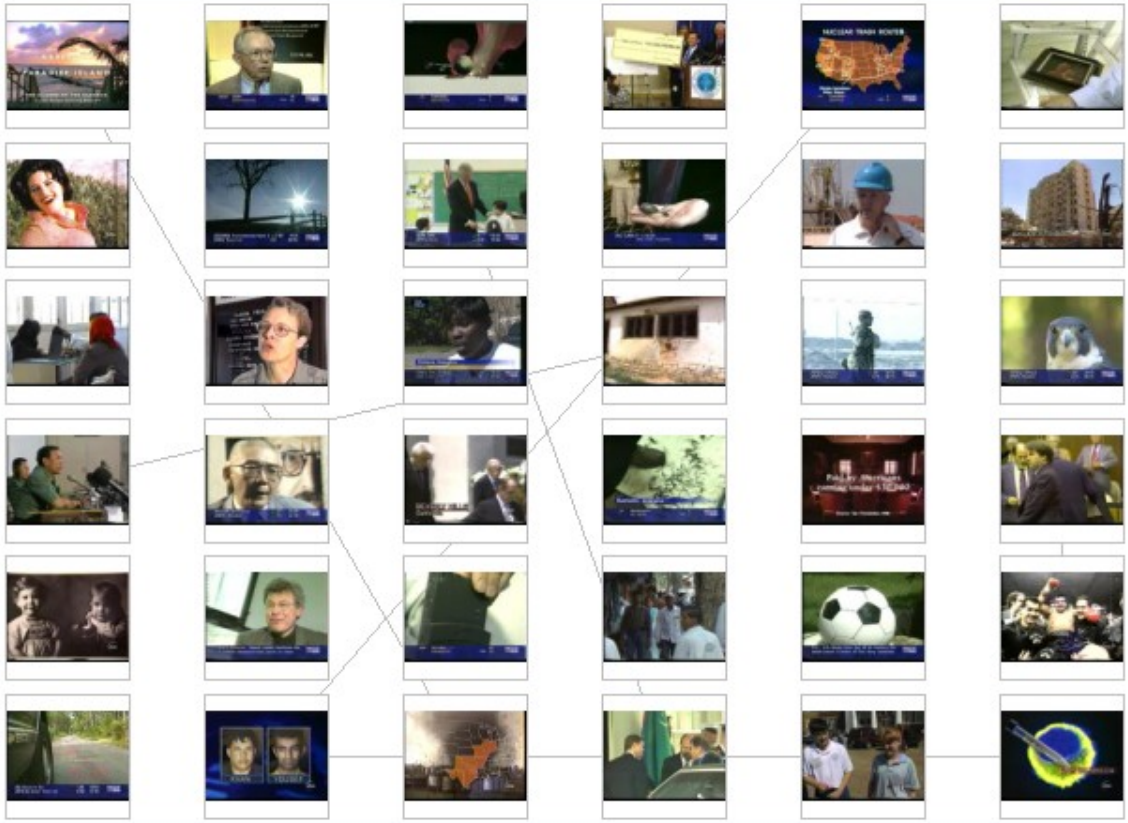
Browse for Images

Nearest neighbours:



Prune | Reset

Hub network | My network



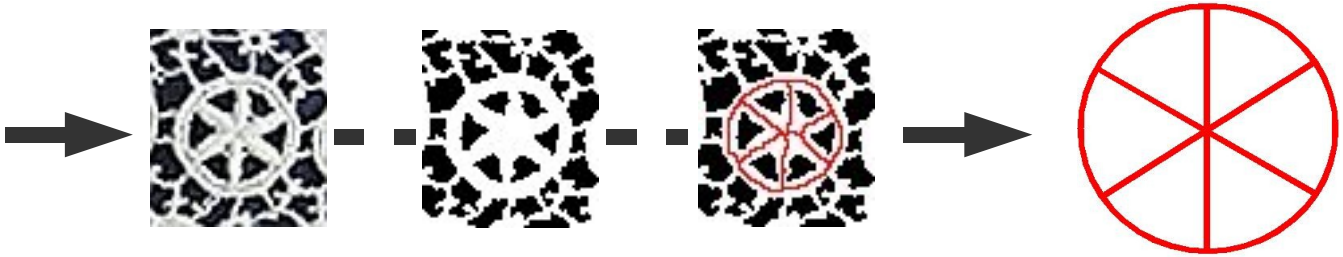
Temporal | My selection

Done

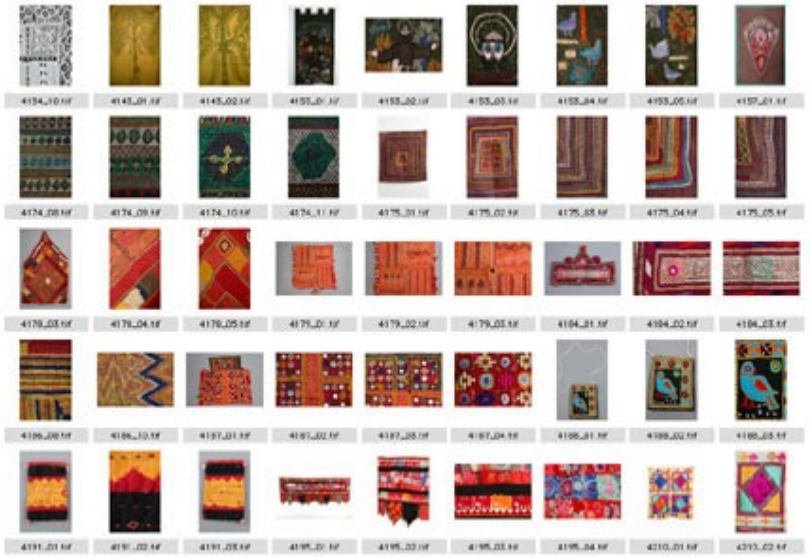
Start | C.. | M.. | C.. | c... | n... | M.. | M.. | h... | i... | W.. | EN | 15:51

Visual Search in large 2D DB

Can the shape representation lead to functional descriptors and help in **automatically creating meta-data?**



Challenge: robust automatic extraction of graphs, dealing with topological instabilities/events.



4227_03.tif

Imperial College
London

Image, Shape & Multimedia Resource Discovery

Stefan Rüger
Multimedia & Information Systems
mmis.doc.ic.ac.uk

Frederic Fol Leymarie
Arts Computing
doc.gold.ac.uk/~ffl