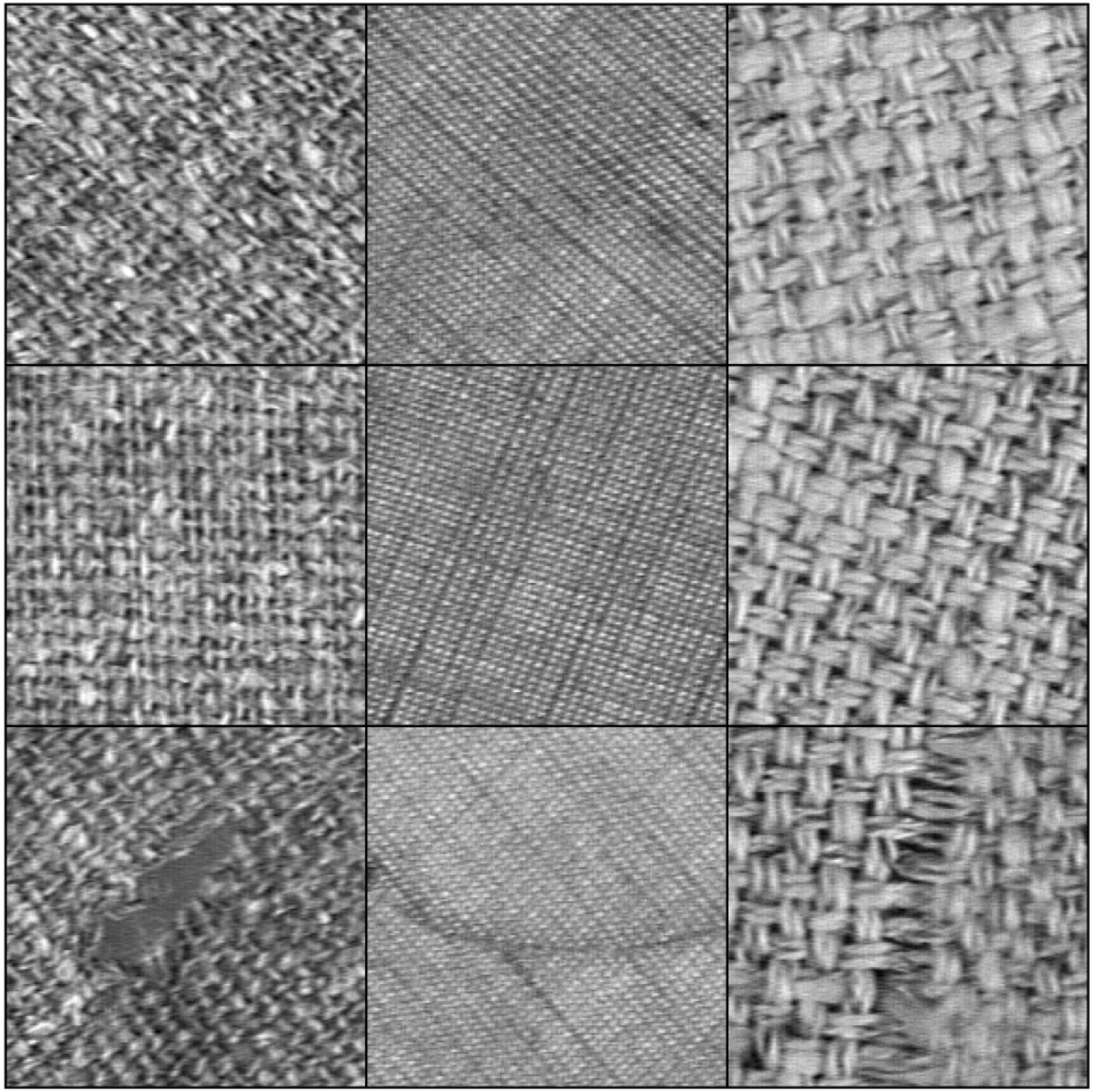
Kap. 5b
Anwendungen:
Invarianten über Gruppenmittel

- Sichtprüfungsaufgaben (Textilien)
- Bildsuchmaschinen (query by example) SIMBA und MICHELscope
- Automatische Klassifikation von Blütenpollen
- Suchen in Proteindatenbanken
- Suchen nach Wasserzeichen in alten Drucken

Complexity of Search tasks

- Finding an identical object (pixel by pixel) is trivial !
- Challenge for the future:
  Searching in Equivalence classes (up to a very high semantic level) can be very complex and demanding
- Data Bases without an intelligent access and retrieval mechanism is a graveyard for Bits and Bytes!

Visual inspection of textiles with anisotropic texture



Textildatenbank TILDA

http://www.informatik.uni-freiburg.de/~lmb/tilda (ca. 3200 images)

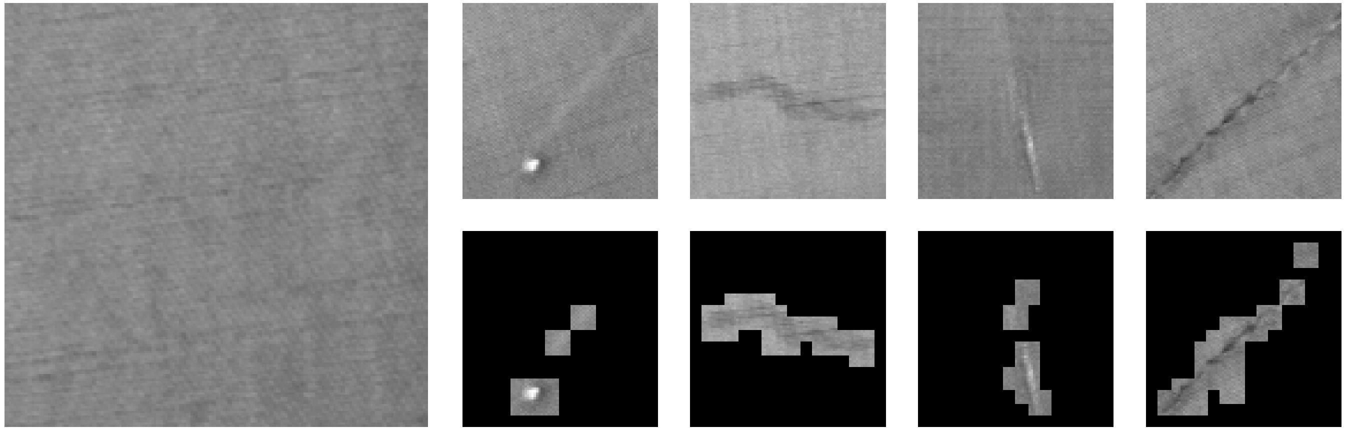
Texturdefektdetektion bei Textilstoffen

Segmentierung der Defekte

Texturklasse c1r1

Fehlerklassen

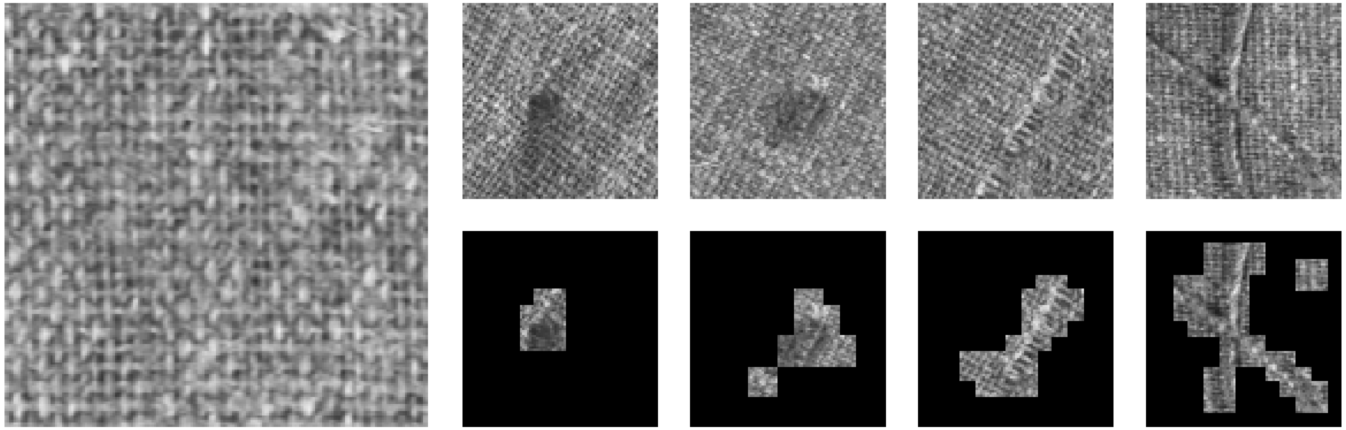
E1 E2 E3 E4



Texturklasse c2r2

Fehlerklassen

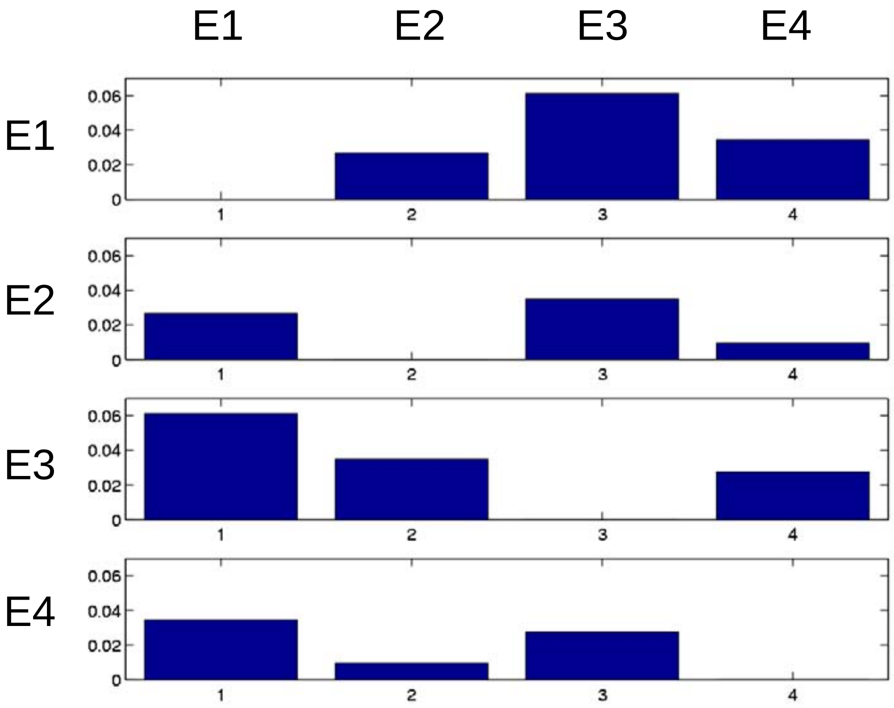
E1 E2 E3 E4



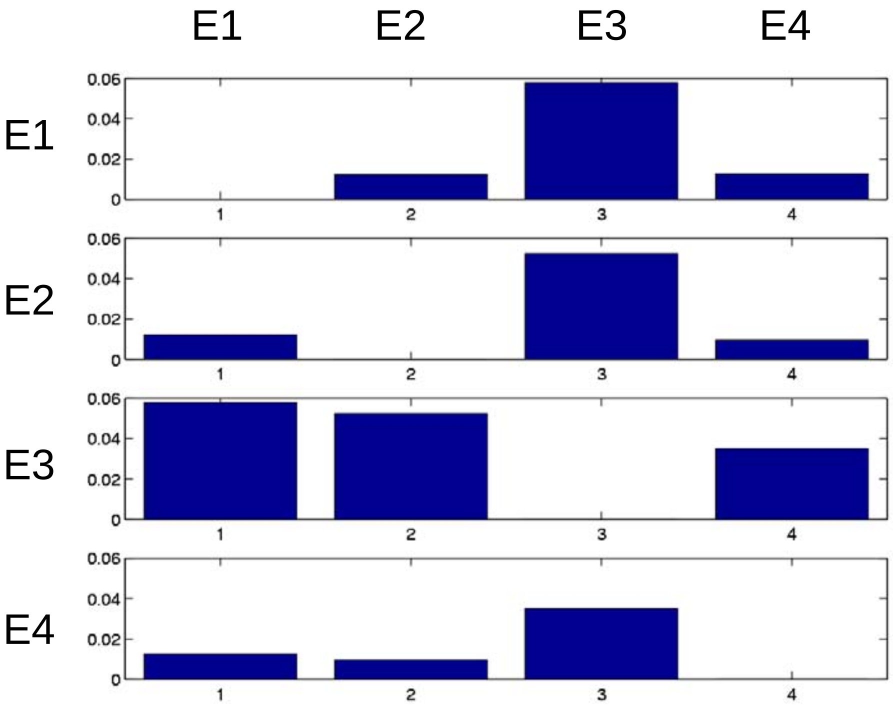
Texturdefektdetektion bei Textilstoffen

Diskriminierung der Defektklassen

Klassendistanzen Texturklasse c1r1

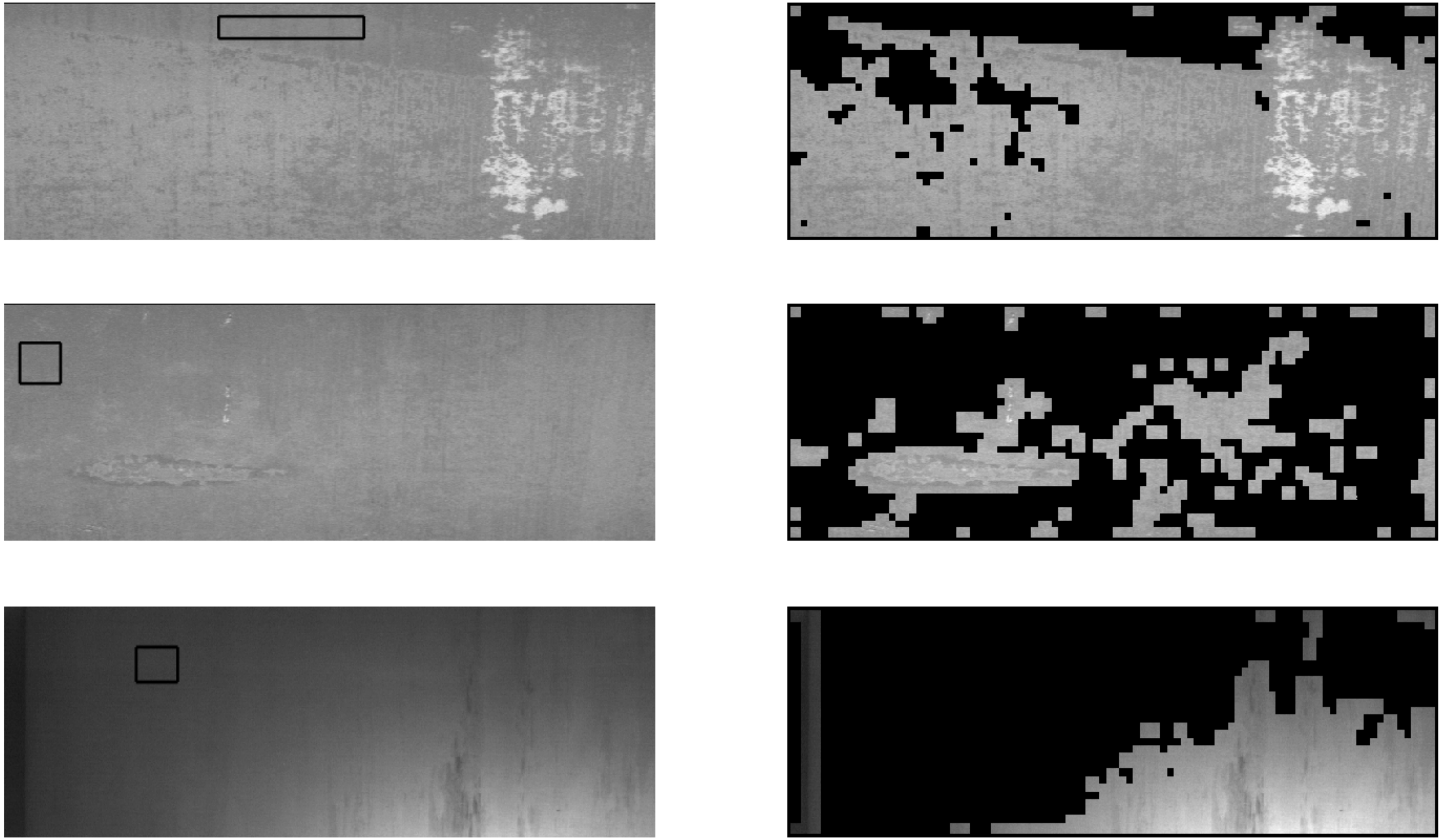


Klassendistanzen Texturklasse c2r2

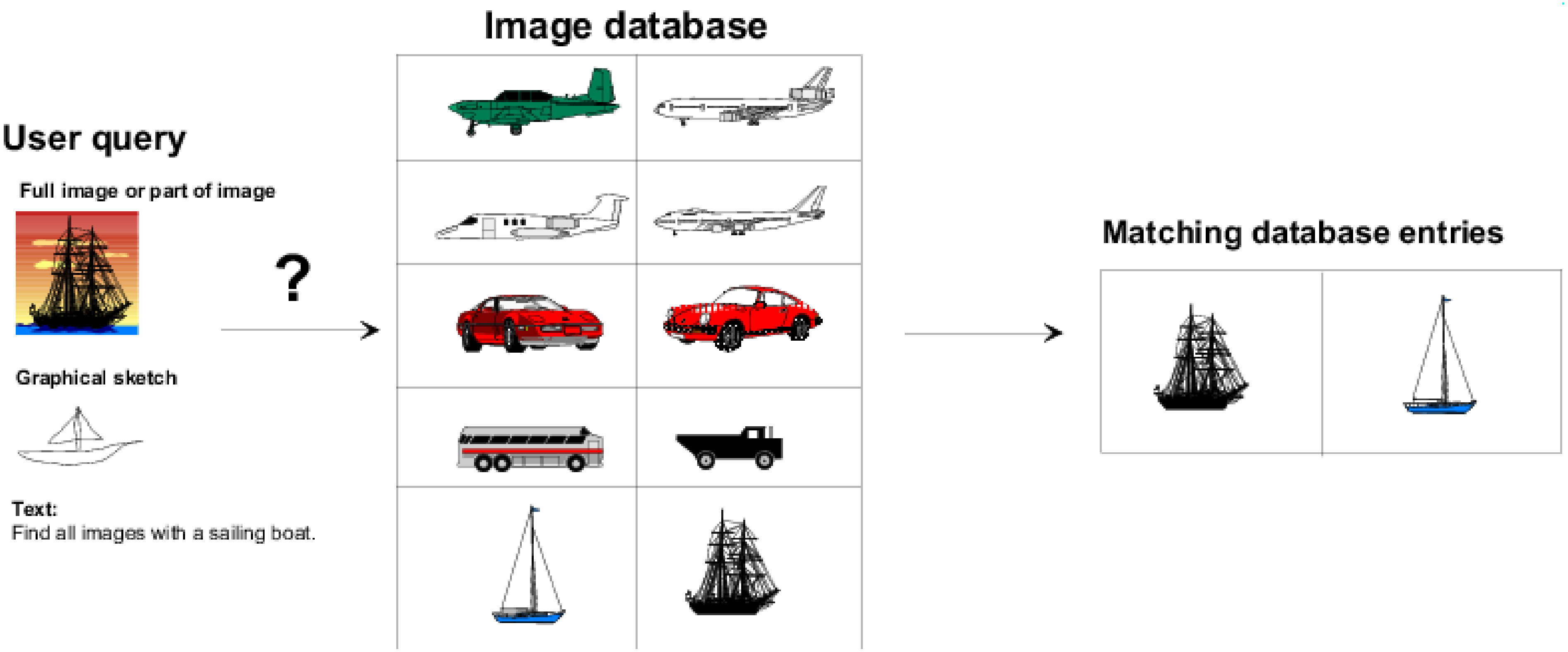


Texturdefektdetektion auf Stahloberflächen

Segmentierung der Defekte



Content based image retrieval



Find all images with a sailing boat.

http://simba.informatik.uni-freiburg.de/

Query image

0.978

0.828

0.792

0.686

0.674

0.649

0.630

0.595

Query image

0.98

0.82

0.81

0.79

0.78

0.78

0.77

0.76

Query image

0.99

0.79

0.77

0.68

0.68

0.66

0.66

0.66

Query image

0.98

0.82

0.79

0.77

0.74

0.73

0.71

0.71

Query image

0.98

0.79

0.77

0.68

0.68

0.65

0.65

0.63

Query image

0.98

0.80

0.78

0.74

0.70

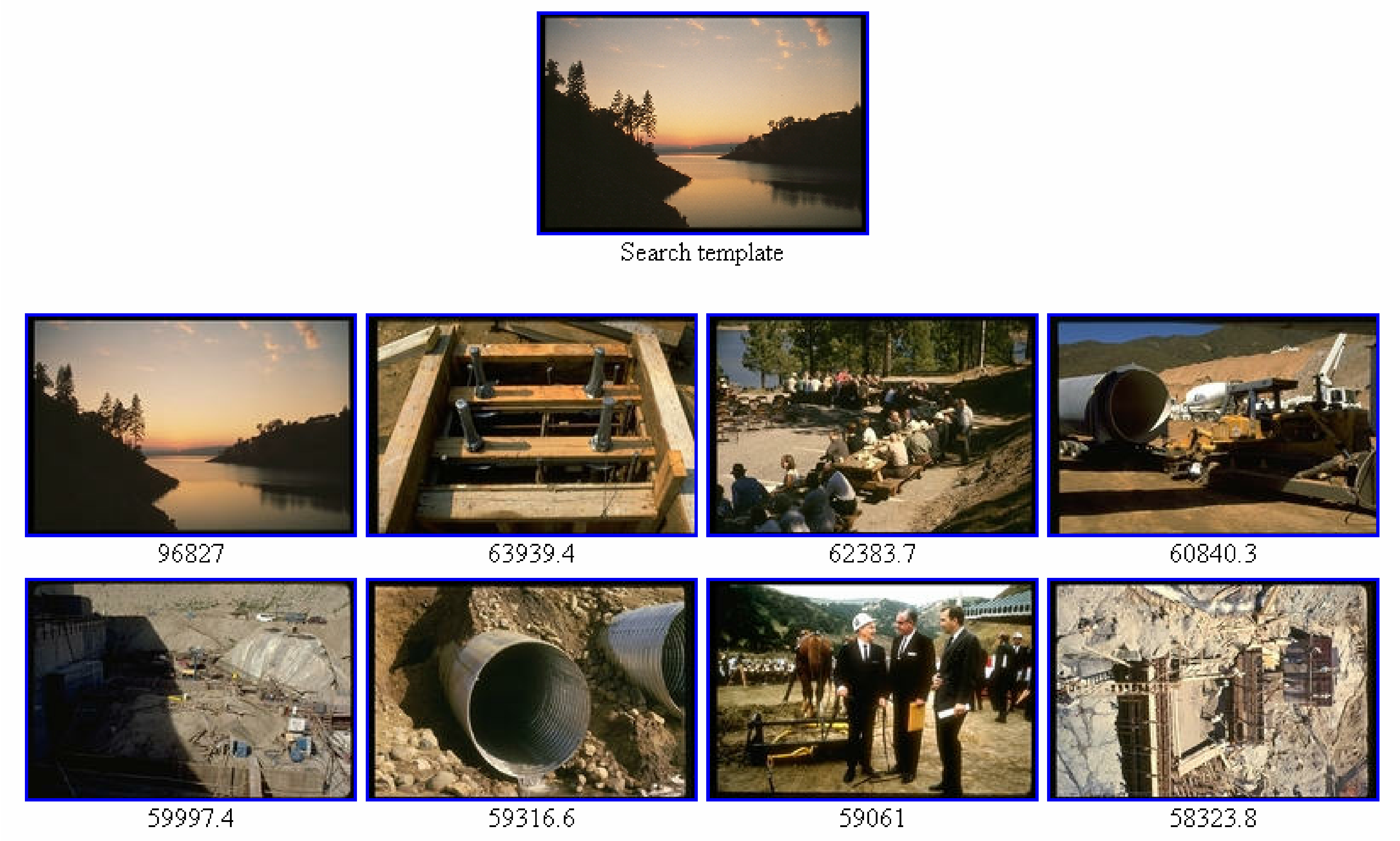
0.60

0.59

0.57

SIMBA – Search IMages By Appearance

COLOR only



SIMBA – Search IMages By Appearance

RELATIONAL only

Search template

87948.7

85491.7

85163

84896.4

84603.2

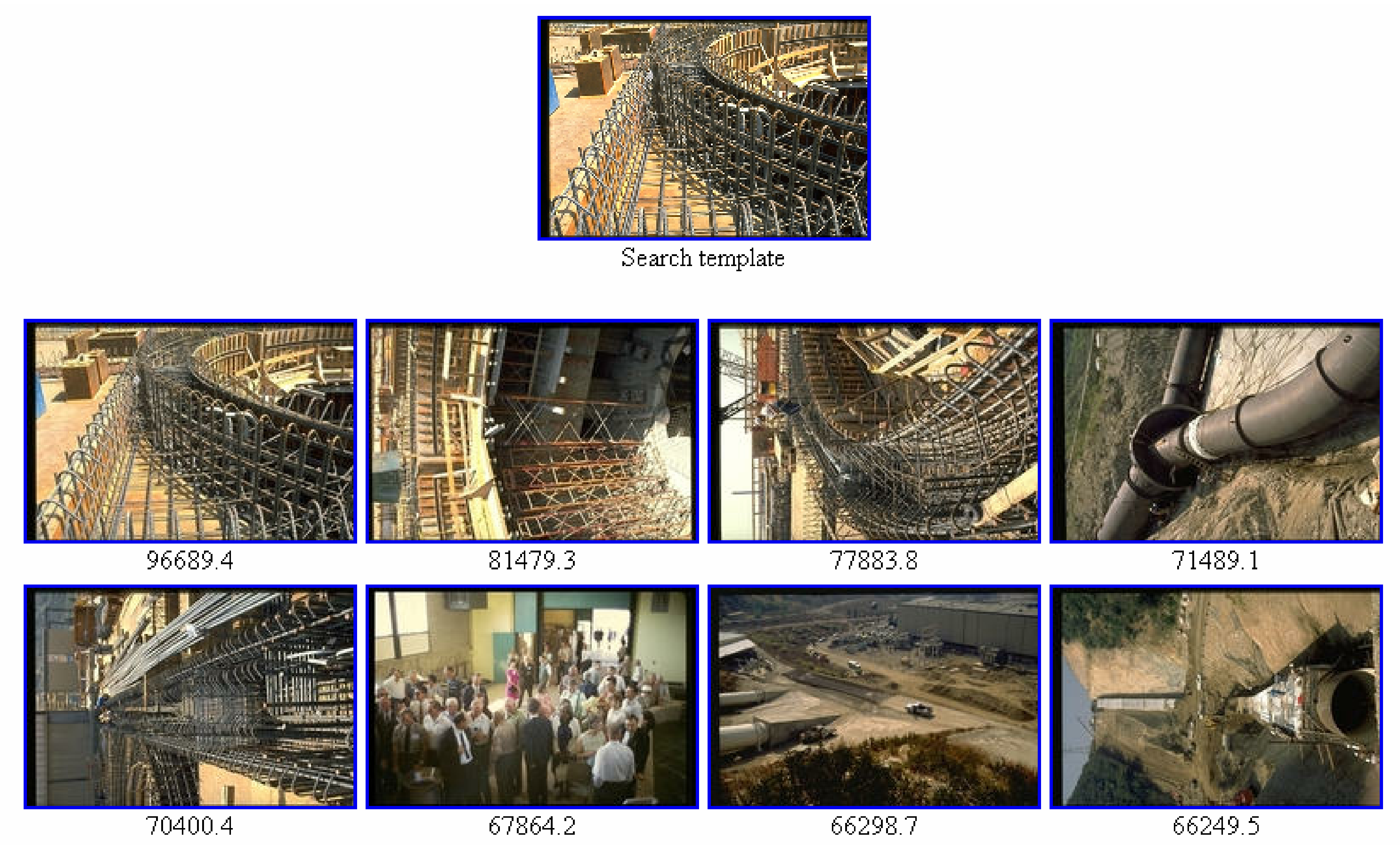
84379.5

84065.4

83475.7

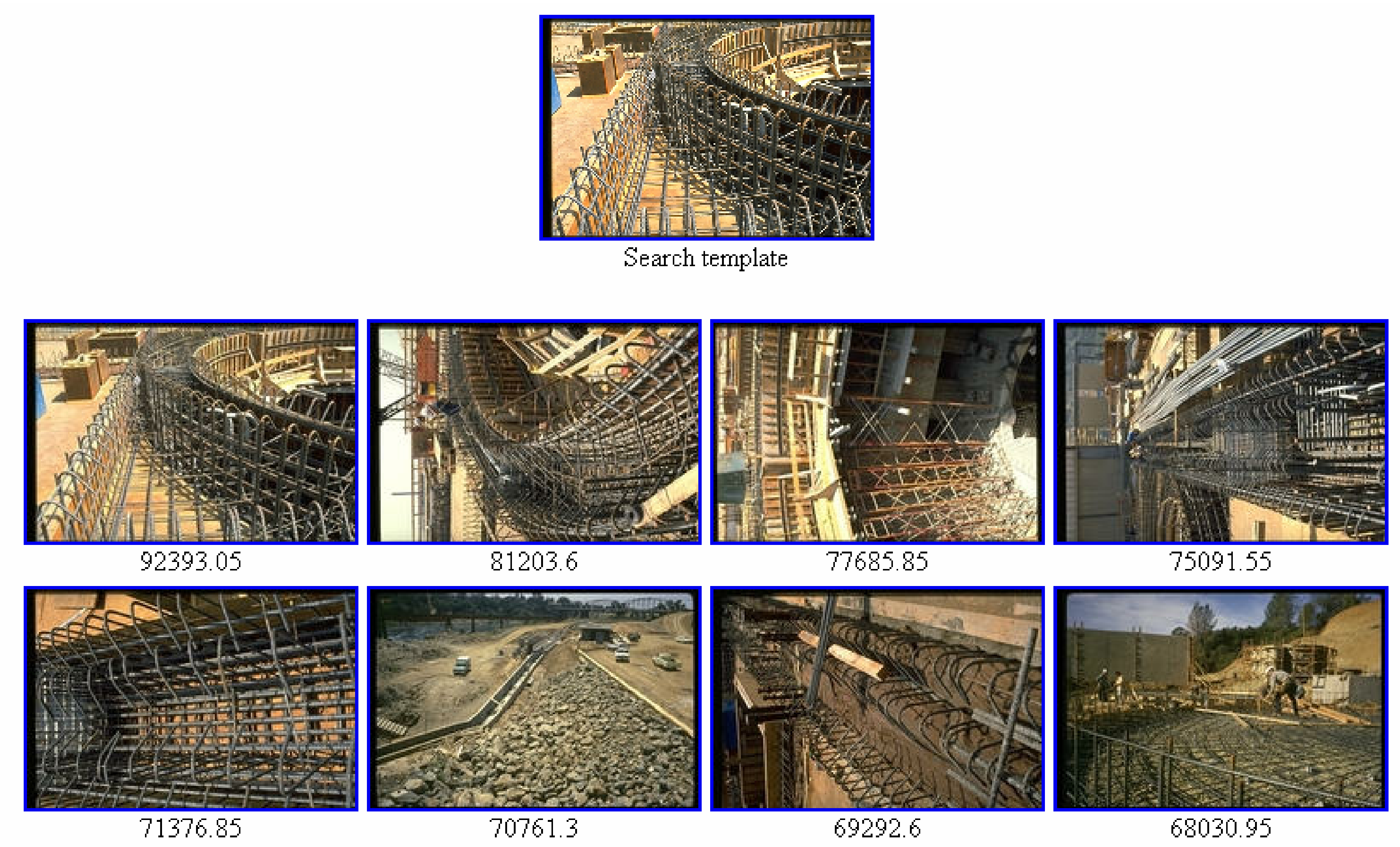
SIMBA – Search IMages By Appearance

COLOR only



SIMBA – Search IMages By Appearance

COLOR+RELATIONAL



Search template

92393.05

81203.6

77685.85

75091.55

71376.85

70761.3

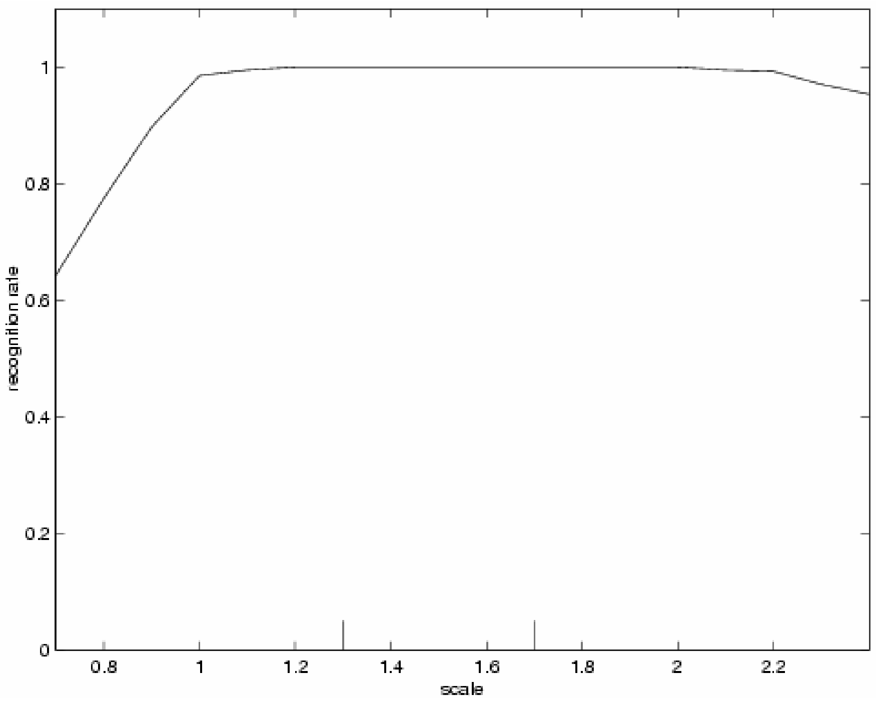
69292.6

68030.95

Experiments

Robustness to scaling

- Database as before, grayvalue features only
- Kernels **M**(1,0)**M**(0,2), **M**(2,0)**M**(0,4)
- Images reduced/zoomed (bilinear interpolation) to scales from 0.7 to 2.4
- Database contains scales 1.3 and 1.7 only



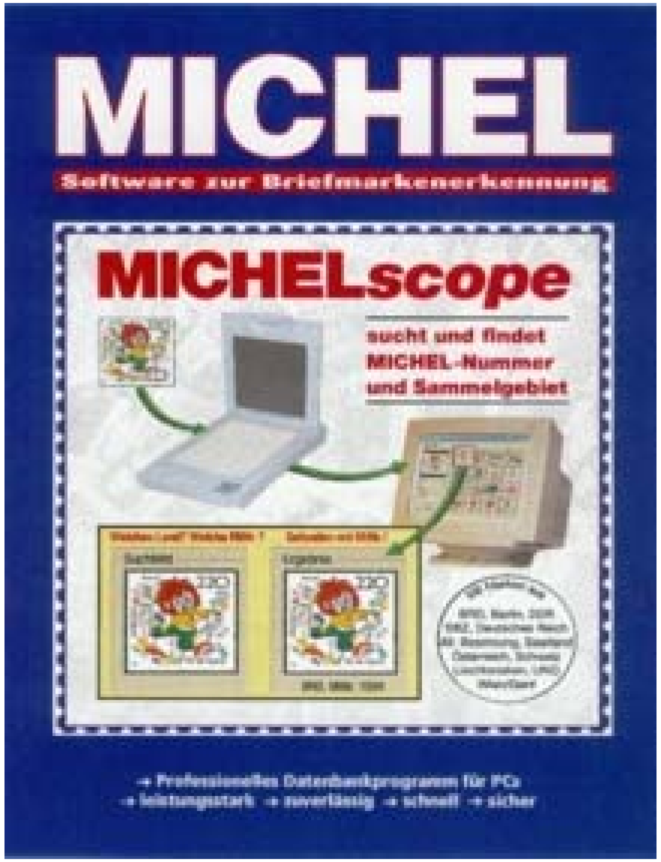
⟶ Good recognition from scale 1 to 2.2, i.e. scale 4.8 in area

SIMBA
Searching Images by Appearance

http://simba.informatik.uni-freiburg.de/

MICHELscope

A search engine for philatelists



Suchbild

Suchmethode

Farbe — Textur

☑ Motiv

☑ Größe

☑ Seitenverhältnis

Datenbank: BRD

Ergebnisse

Brd0644 Brd0638 Brd0732 Brd0689

Brd0729 Brd0636 Brd0727 Brd0637

Brd0645 Brd0640 Brd0642 Brd0635

Brd0691 Brd0730 Brd0639 Brd0643

Suchbild

Suchmethode

Farbe – Textur

☑ Motiv

☑ Größe

☑ Seitenverhältnis

Datenbank: BRD

☑ Motiv

☑ Größe

☑ Seitenverhältnis

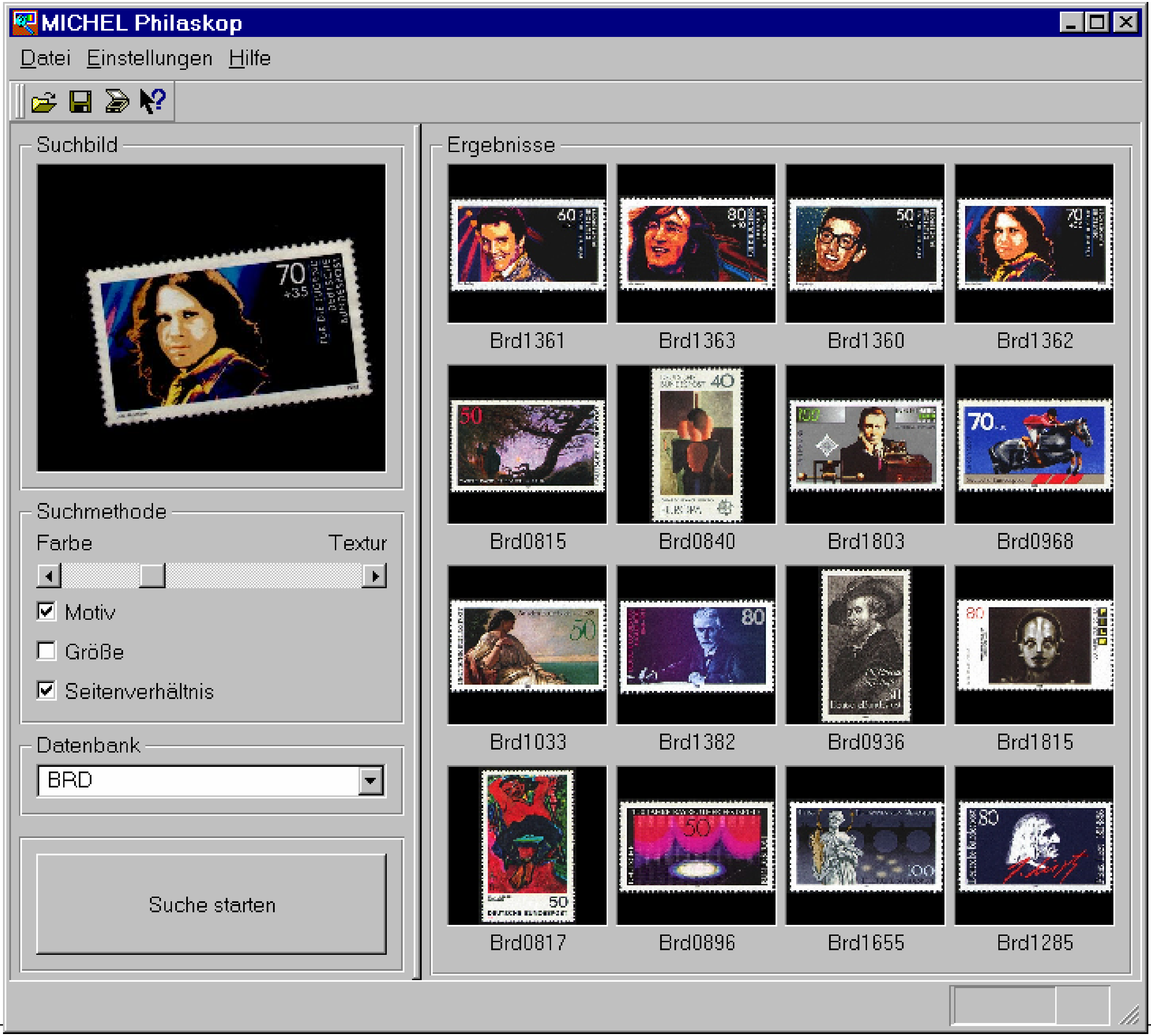
Datenbank: BRD

Brd0523 Brd0485 Brd0538 Brd0232

Brd0408 Brd0409 Brd0486 Brd0447

Brd0540 Brd0164 Brd0449 Brd0266

Brd0386 Brd0173 Brd1610 Brd0487



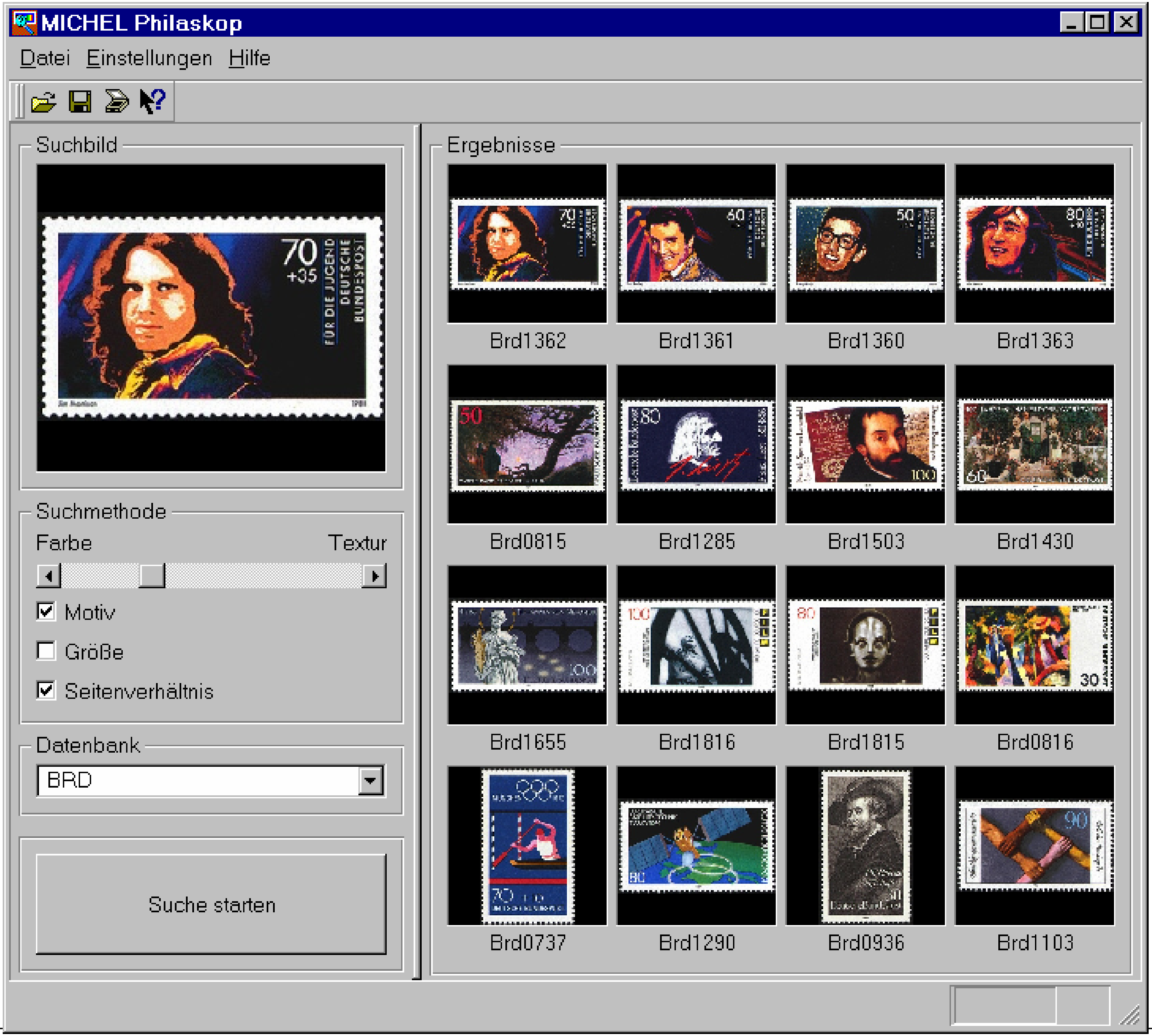
☑ Motiv

☐ Größe

☑ Seitenverhältnis

Datenbank: BRD

Brd1361	Brd1363	Brd1360	Brd1362
Brd0815	Brd0840	Brd1803	Brd0968
Brd1033	Brd1382	Brd0936	Brd1815
Brd0817	Brd0896	Brd1655	Brd1285



Suchbild

Suchmethode

Farbe — Textur

☑ Motiv

☐ Größe

☑ Seitenverhältnis

Datenbank: BRD

Suchmethode

Farbe — Textur

☑ Motiv

☑ Größe

☑ Seitenverhältnis

Datenbank: BRD

☑ Motiv

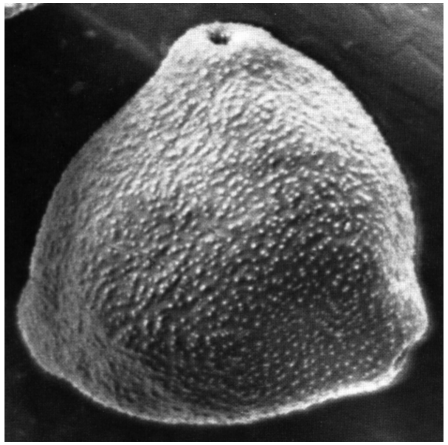
☑ Größe

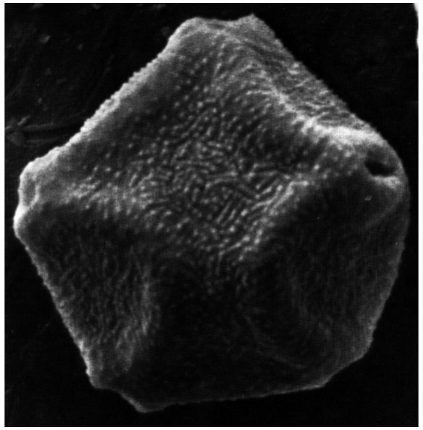
☑ Seitenverhältnis

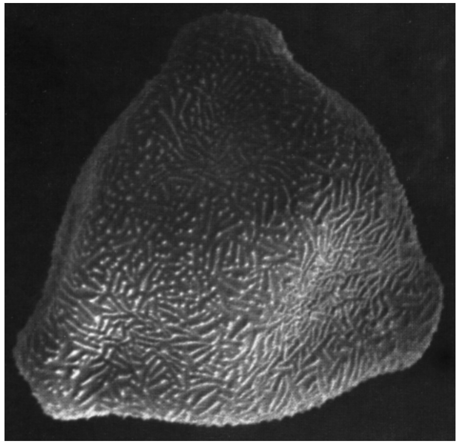
Datenbank: Deutschland

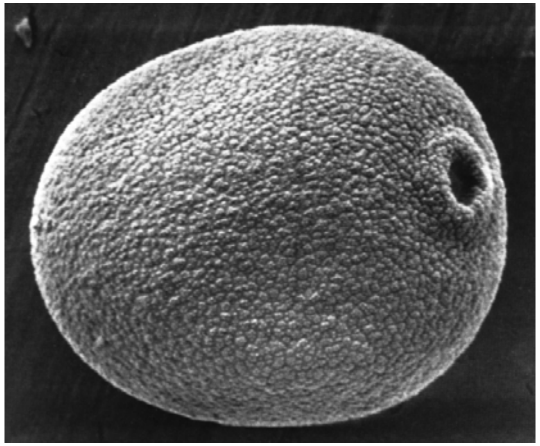
Deutschland Berlin (West) 0042

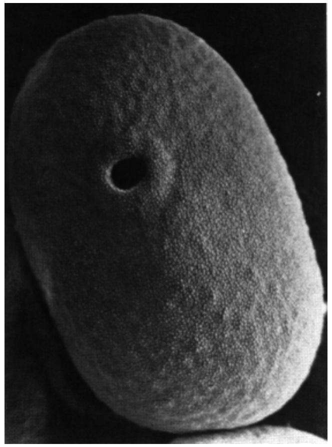
Projekt zur automatischen Erkennung von Blütenpollen - Elektronenmikroskopische Aufnahmen von Pollen

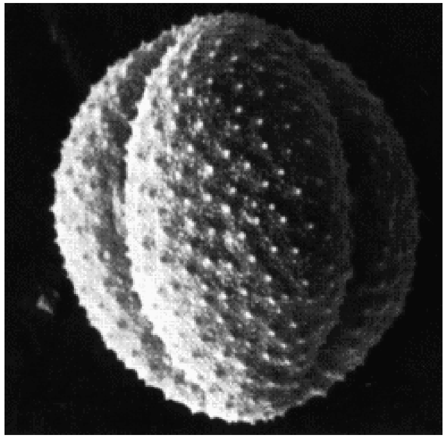

Hasel


Erle


Birke

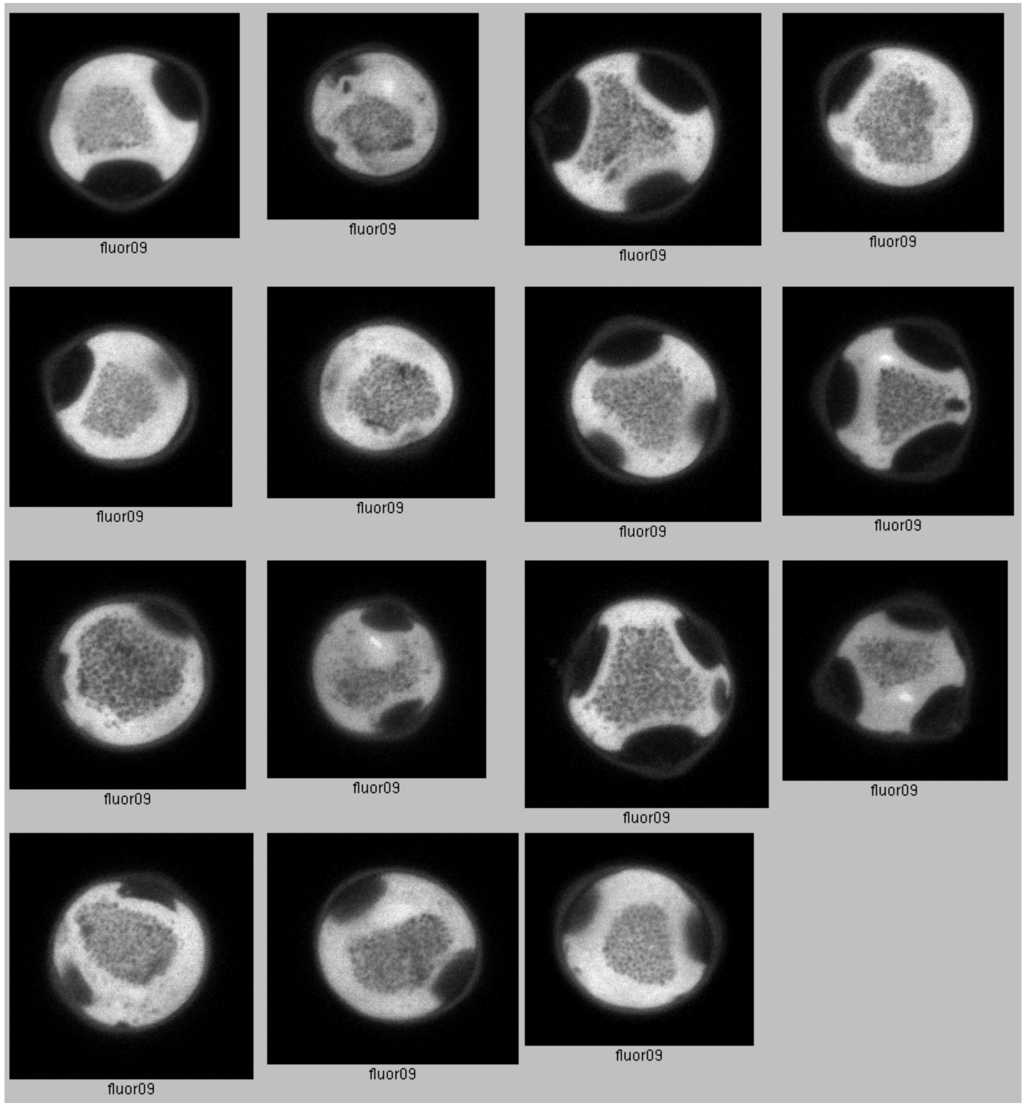

Gräser


Roggen

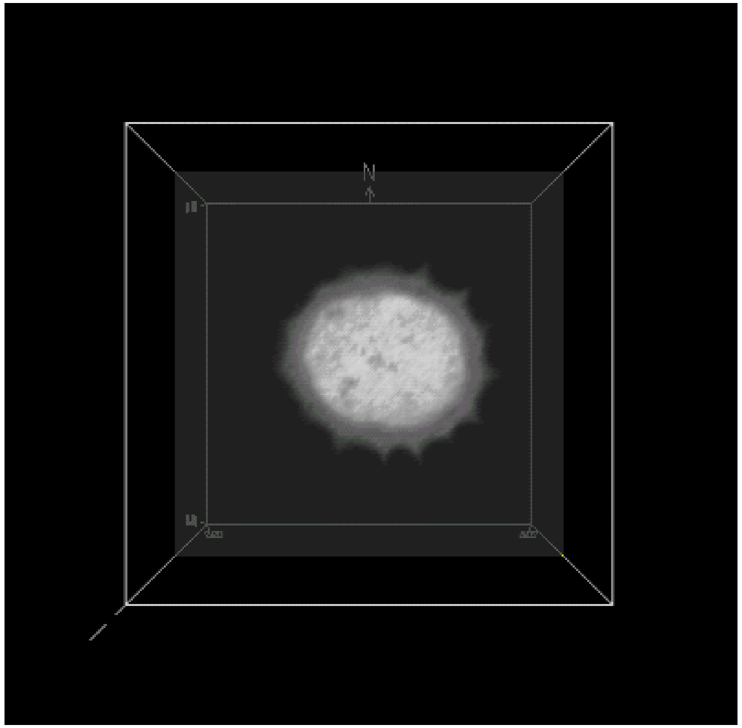

Beifuß

+ 33 further species (not relevant for allergies)

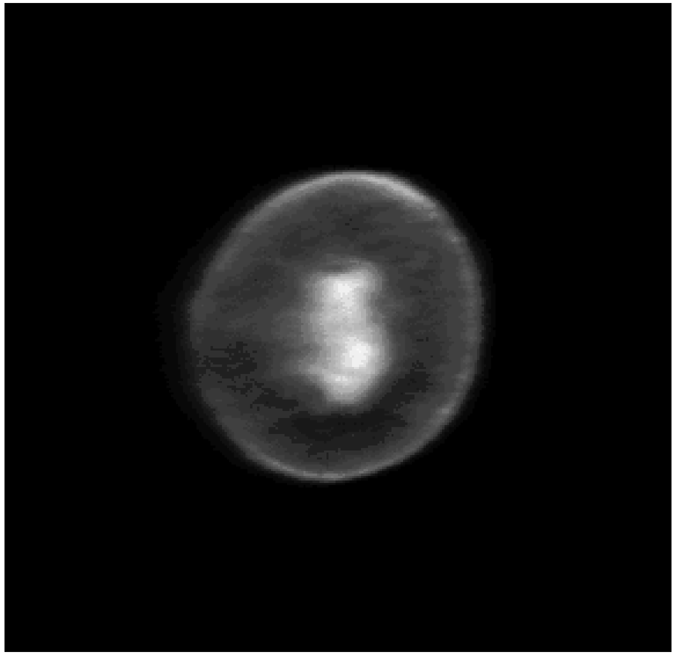
Haselpollen

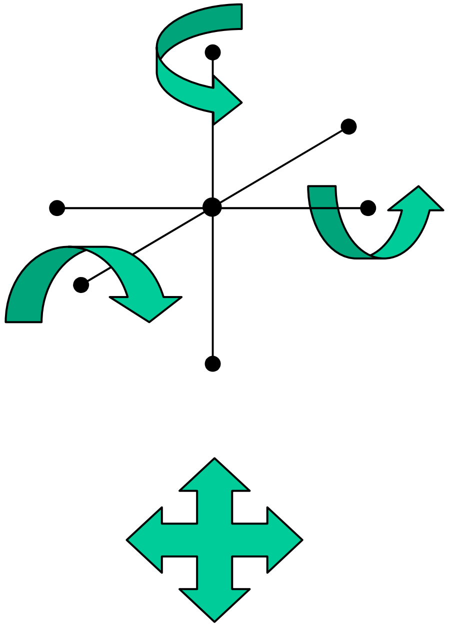


Gänseblümchen/daisy pollen grain



Eibe/Taxus



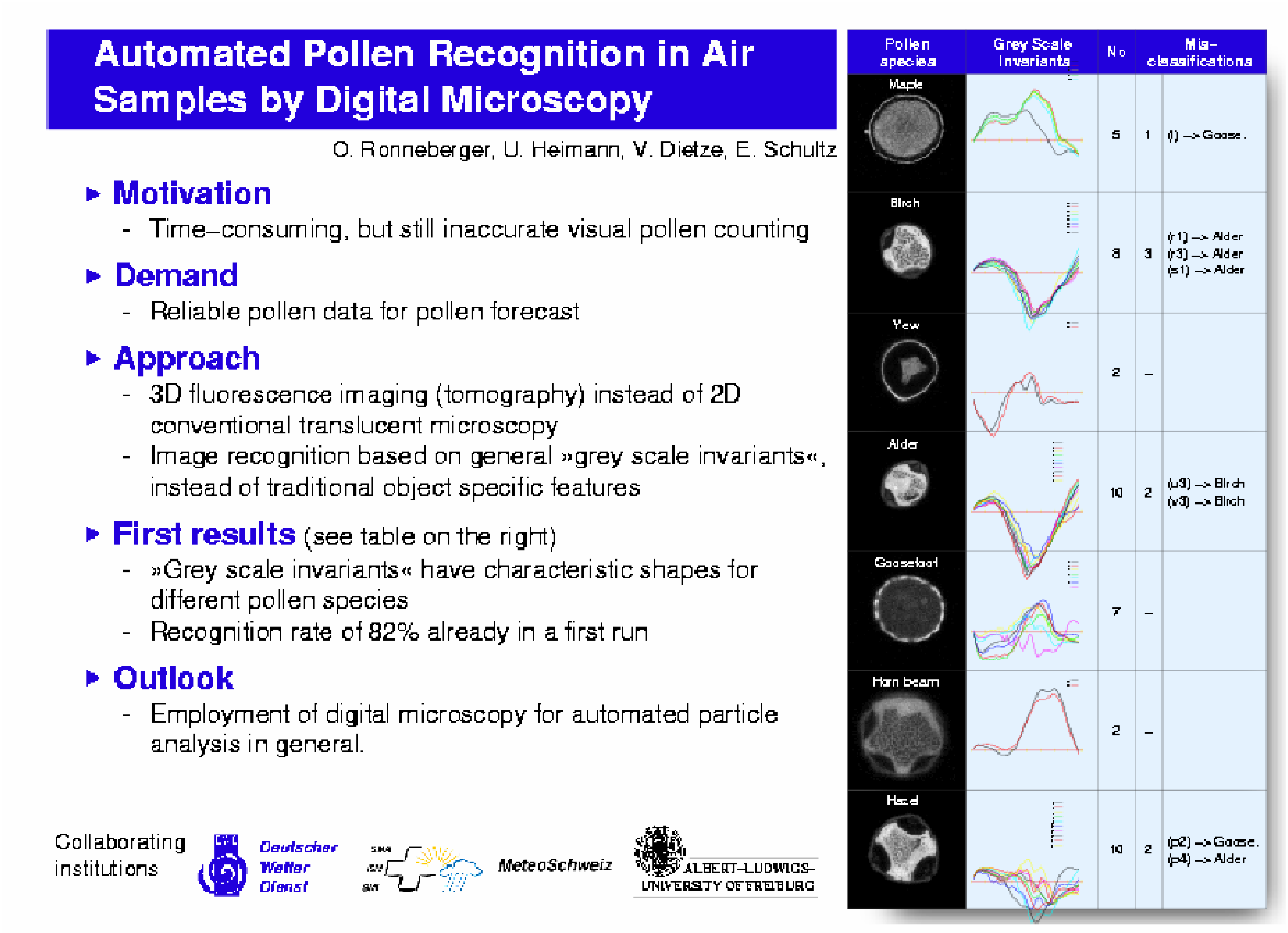


Integriere über die 3D Euklidsche Bewegung

Reference Data Base Description

- 26 pollen species directly sampled from the plants
  - Correctly assigned species
  - Maybe less variation in size and shape compared to airborne pollen (no different sub-species, no regional variation)
  - No deformation or contamination or agglomeration

- Recording 3D volumes of about 15 samples of each species with a Confocal Laser Scanning Microscope
  - No distortions of the data due to imperfect imaging of a usual microscope



Automated Pollen Recognition in Air Samples by Digital Microscopy

O. Ronneberger, U. Heimann, V. Dietze, E. Schultz

- **Motivation**
  - Time–consuming, but still inaccurate visual pollen counting
- **Demand**
  - Reliable pollen data for pollen forecast
- **Approach**
  - 3D fluorescence imaging (tomography) instead of 2D conventional translucent microscopy
  - Image recognition based on general »grey scale invariants«, instead of traditional object specific features
- **First results** (see table on the right)
  - »Grey scale invariants« have characteristic shapes for different pollen species
  - Recognition rate of 82% already in a first run
- **Outlook**
  - Employment of digital microscopy for automated particle analysis in general.

Collaborating institutions

Pollen species	Grey Scale Invariants	No	Mis-classifications	
Maple		5	1	(l) –> Goose.
Birch		8	3	(r1) –> Alder (r3) –> Alder (s1) –> Alder
Yew		2	–	
Alder		10	2	(u3) –> Birch (v3) –> Birch
Goosefoot		7	–	
Horn beam		2	–	
Hazel		10	2	(p2) –> Goose. (p4) –> Alder

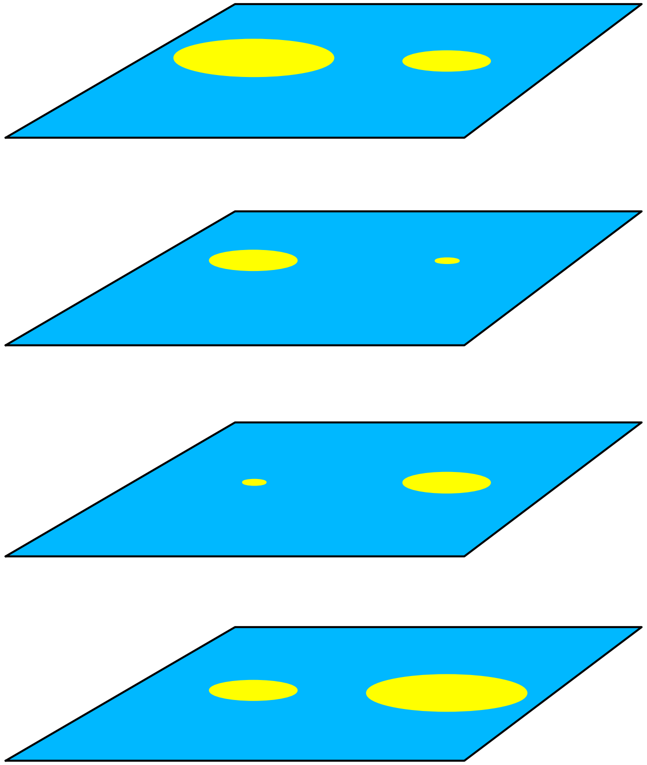
Classification Results using 3D LSM Data

(leave-one-out Classification)

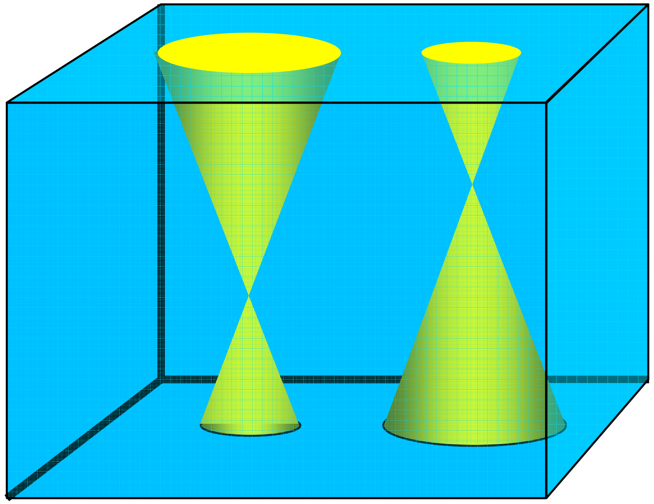
	Correct	Wrong classifications
Artemisia:	13	1 -> *Compositae*, 1 -> *Platanus*
Alnus:	15	-
Alnus viridis:	12	-
Betula:	13	2 -> *Plantago*
Corylus:	13	1 -> *Alnus*
Gramineae/Poaceae:	15	-
Secale:	11	3 -> *Fagus*, 1 -> *Tilia*
Allergolocial irrelevant*:	282	2 -> *Gramineae*
Total:	97.4%	2.6%

***** *Acer, Carpinus, Chenopodium, Compositae, Cruciferae, Fagus, Quercus, Aesculus, Juglans, Fraxinus, Plantago, Platanus, Rumex, Populus, Salix, Taxus, Tilia, Ulmus, Urtica*

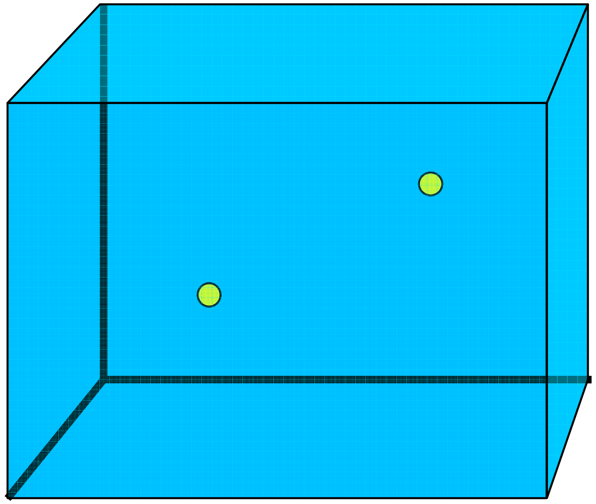
Gewinnung von tomographischen 3D-Daten mit einem normalen Fluoreszenz-Mikroskop



Aufnahme eines Bildstapels mit einem normalen Mikroskop



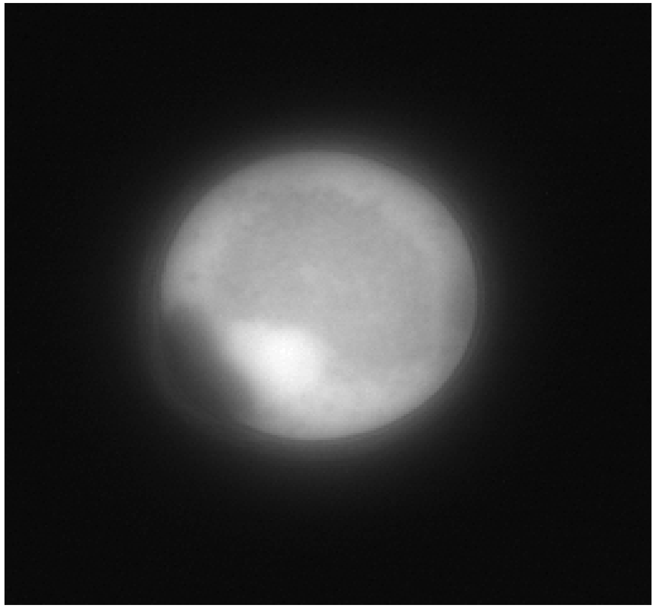
Zusammensetzen der Bilder im Computer zu einem Volumen



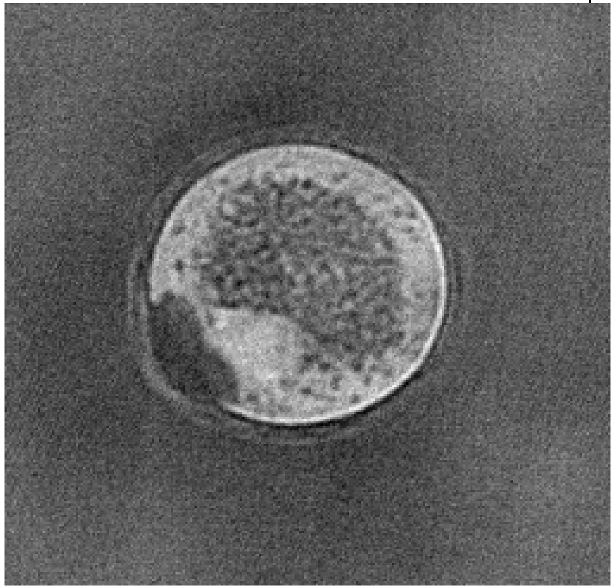
Nach Entfaltung mit der Punktbildfunktion sind alle unscharfen Bereiche entfernt

Deconvolution of a Pollen (Hasel/Corylus)

point-spread function



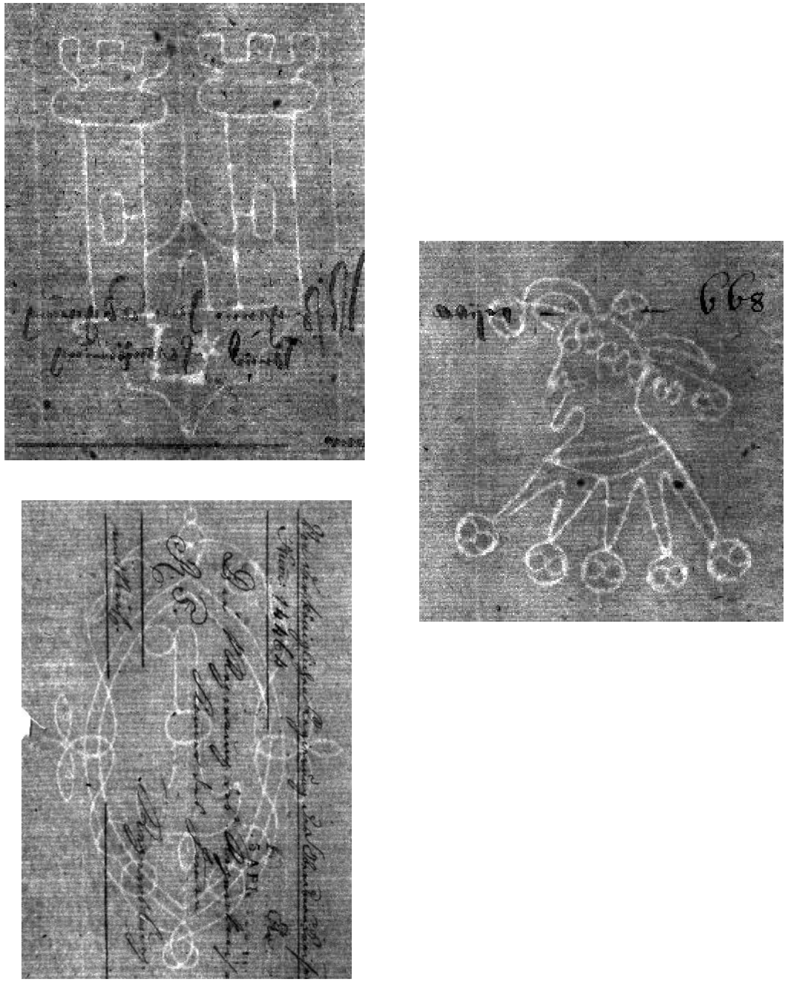
volume



deconvolved volume

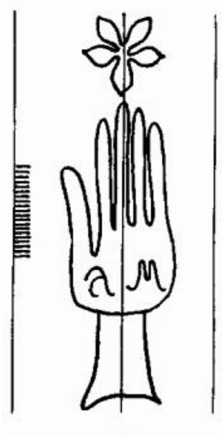
Search Engine for Watermarks in old paper prints
(Joint project with Basel Paper Mill, Swiss Museum for Paper, Writing and Printing)

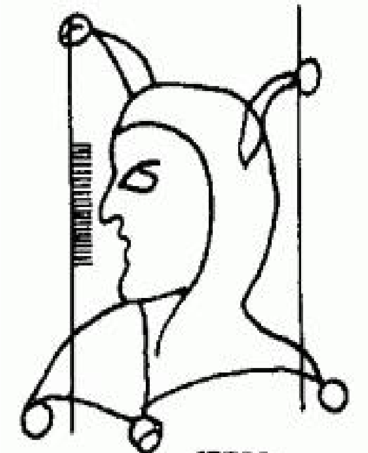
Scanned watermark images
(samples)

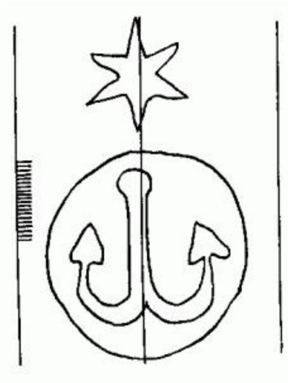


After preprocessing

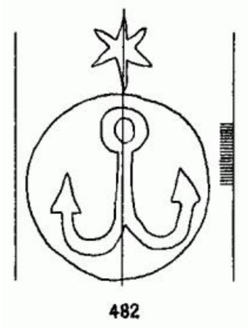
(sample images)

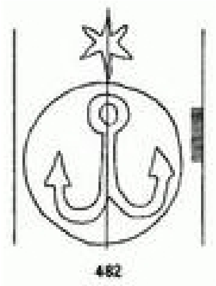

10730


15722


492

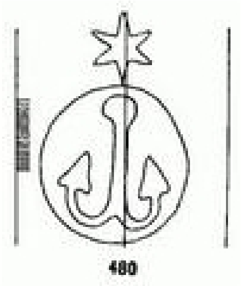
Query Image



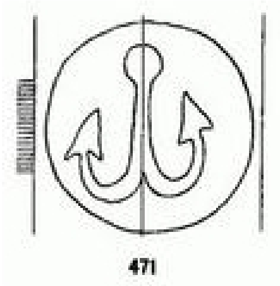


1.00000

0.98446

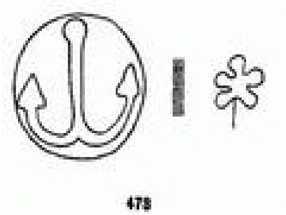


0.98229

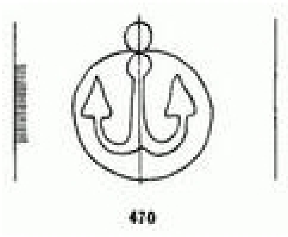


0.98045

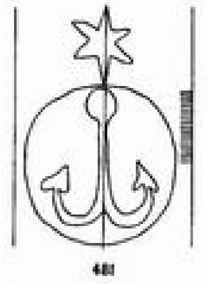
0.97986



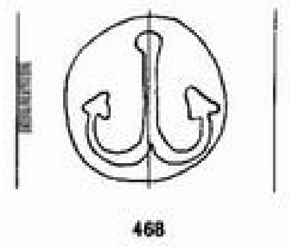
0.97949



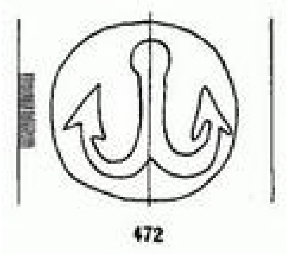
0.97924



0.97890

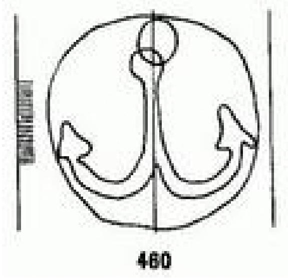


0.97787



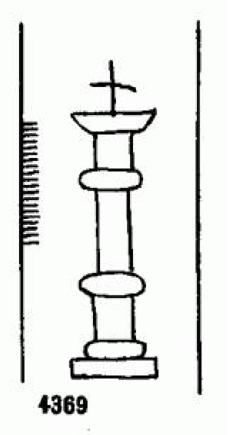
0.97580

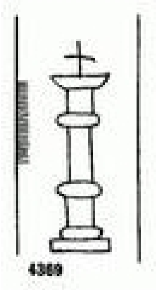
0.97565



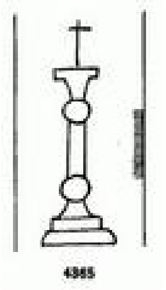
0.97326

Query Image

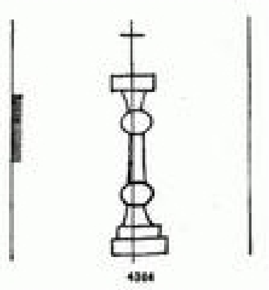


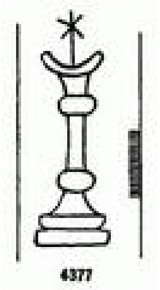


1.00000



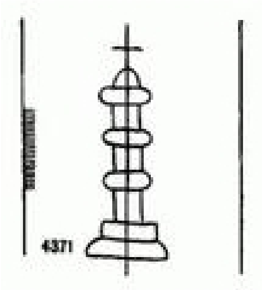
0.80769


0.80128

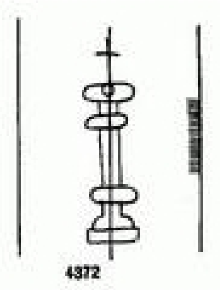


0.78125

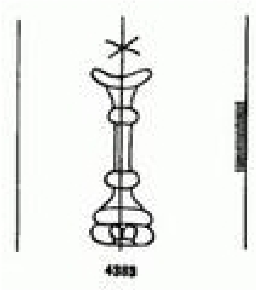
0.77926

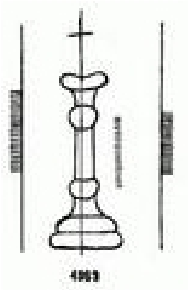


0.77350

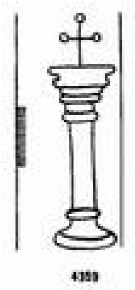


0.76923

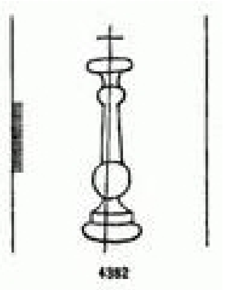

0.72436



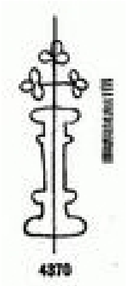
0.72192



0.71923

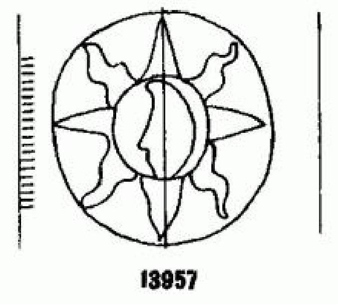


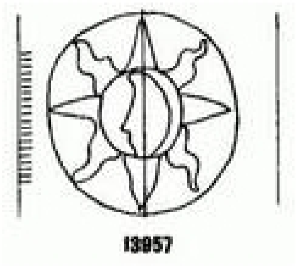
0.71795



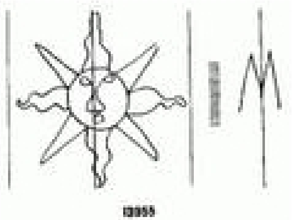
0.70964

Query Image

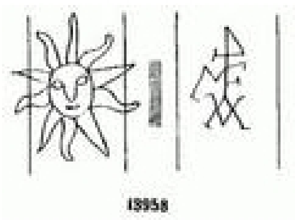




1.00000

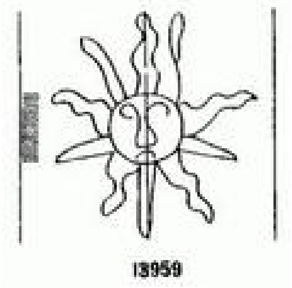


0.94530



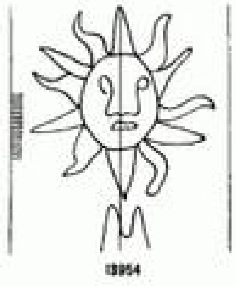
0.93585

0.91158

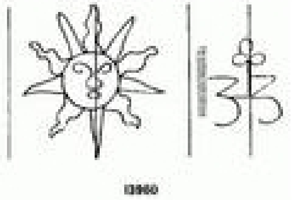


0.87267

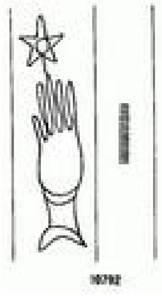
0.86996



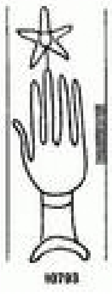
0.86776



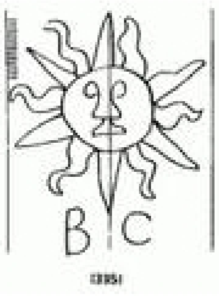
0.83026



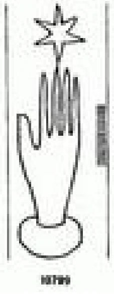
0.76324



0.76137



0.75978



0.75896

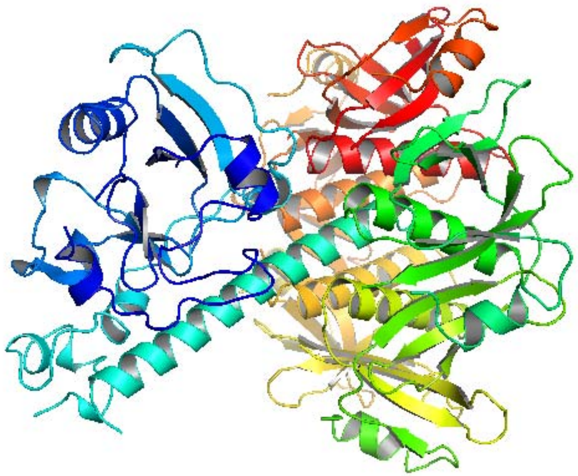
USED FEATURES:

- We encode inter-relational structure knowledge from feature vectors.
- Line segments are grouped by hierarchical clustering.
- Relative spatial arrangements of the grouped line segments are captured in histograms.
- The resulting histogram is robust against illumination changes, rotation, translation and scaling.

Development of a Search Engine for Protein Databases

- There is no canonical way of describing the folding structure of proteins
- This search engines develops space-invariant descriptors for the folding structure of proteins which can be used as index for searching in databases

A Search Engine for Protein Databases based on 3D Folding-Structure



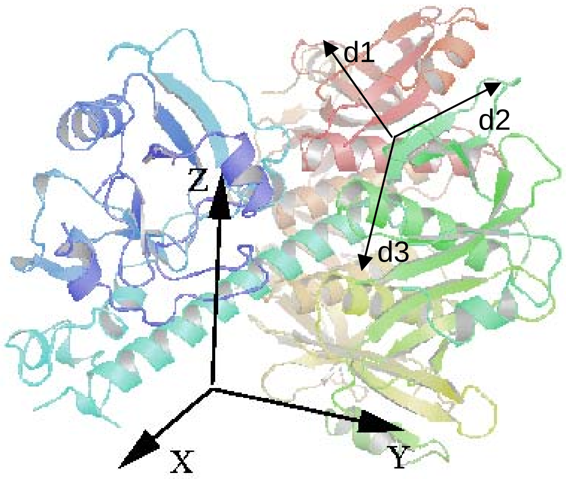
- A Protein is a highly structured 3D object consisting of amino acid chains
- Typical number of amino acids: 100-1000
  → several thousands of atoms
- In the past proteins were compared solely by its amino acid sequence, but
  different sequence → same 3D-structure
  3D-structure → functional behavior

- Growing amount of 3D-macromolecular structural data (currently > 32000 in Brookhaven‘s PDB)
- need for efficent search and retrieval techniques
  - drug design
  - evolutionary link detection
- current structural search techniques rely on alignment algorithms
  → slow (retrieval times >hour or >day)

A Search Engine for Protein Databases based on 3D Folding-Structure

- Methods known from Content-Based 3D-Shape/Image Retrieval offer numerous fast algorithms
- Search and Classification is feature-based. The features are invariant to the 3D-pose and the numbering of the amino acids.
- *Invariant Protein Features* are numerical quantities that describe global properties of the Protein, for example

F1 = „How often occur two atoms of the Protein within some distance d“
F2 = „Percentage of atoms which are part of a alpha-Helix structure“



Preliminary Results:

Classification Accuracy:
- for SCOP-‘class‘ granularity: **99.8%**
- for SCOP-‘fold‘ (class alpha): **97.6%**

Retrieval times: less than a second in DB with
more than 32000 proteins (>65000 chains)

Video
“Licht ins Dunkel”

Licht ins Dunkel\VTS_01_1.avi