PROFESSUR
MEDIENINFORMATIK

Content Based Image Retrieval (CBIR)

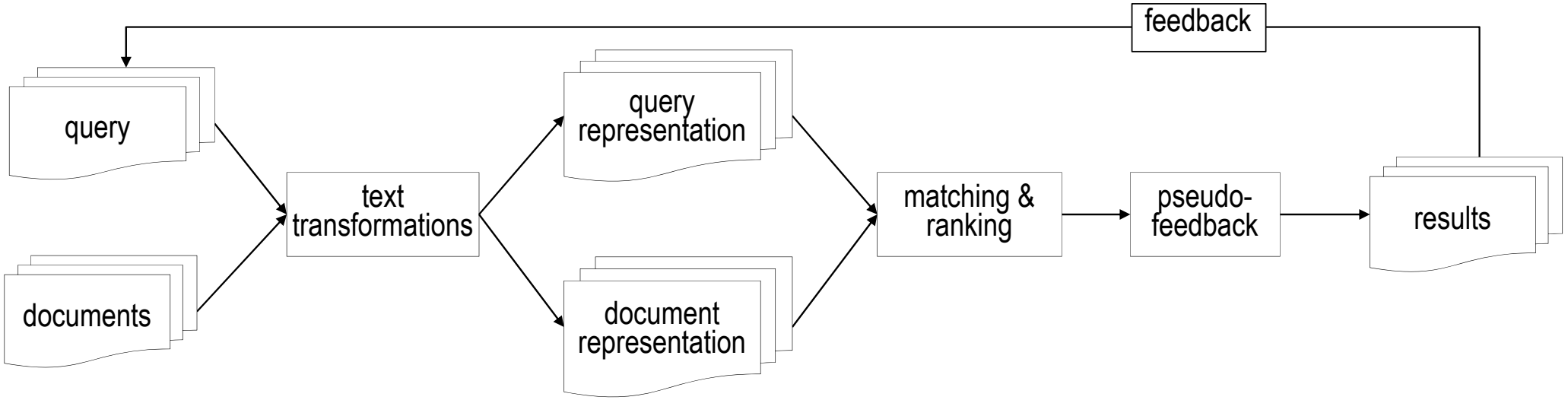
Lecture Media Retrieval

Maximilian Eibl, Medieninformatik, TU Chemnitz

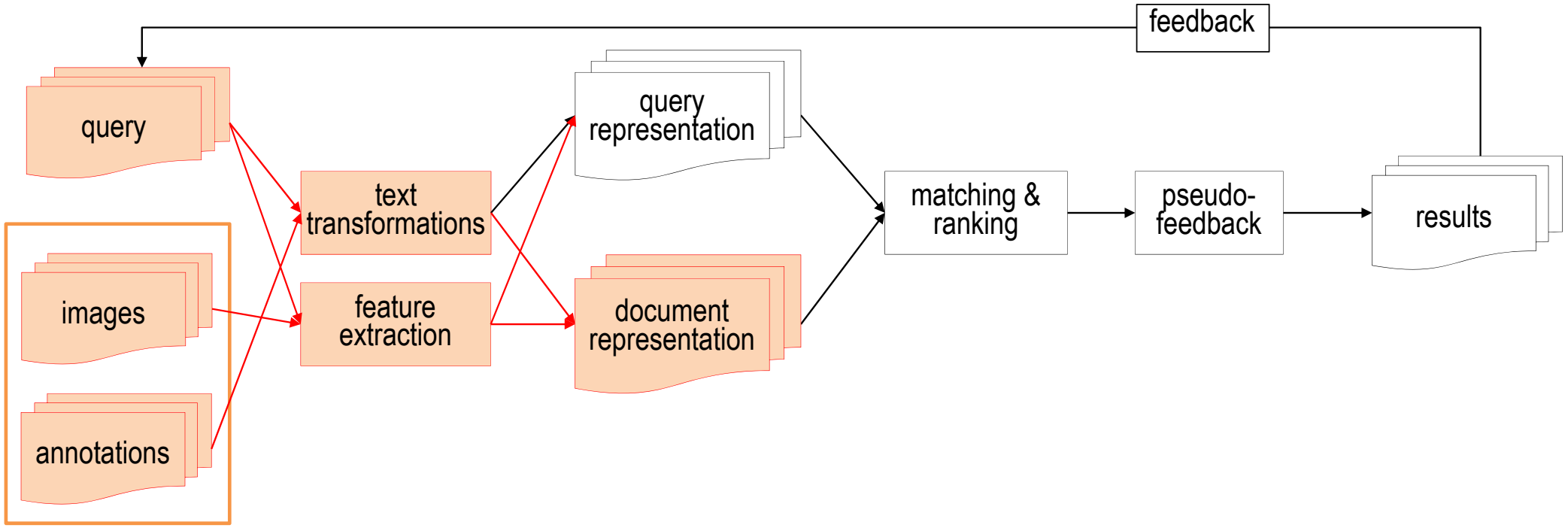
Image and Text: Different Modalities

Text	Image
Linear	2 dimensional
Syntactic structure	No implicit structure
Syntactic hierarchy	No implicit hierarchy
Semantic abstraction	Photorealism

IRS Components: Text Retrieval



IRS Components: Image Retrieval



Types of Queries in Image Retrieval

- Query-by-text
- Query-by-example
- Query-by-sketch
- Query-by-color
- Query-by-texture

} Content Based Image Retrieval CBIR

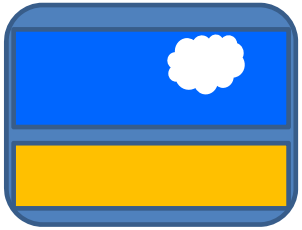
Van den Broek et al. 2002

Types of Queries in Image Retrieval

- Query-by-text: “beach”
- Query-by-example
- Query-by-sketch
- Query-by-color
- Query-by-texture

Types of Queries in Image Retrieval

- Query-by-text
- Query-by-example:
- Query-by-sketch:
- Query-by-color
- Query-by-texture

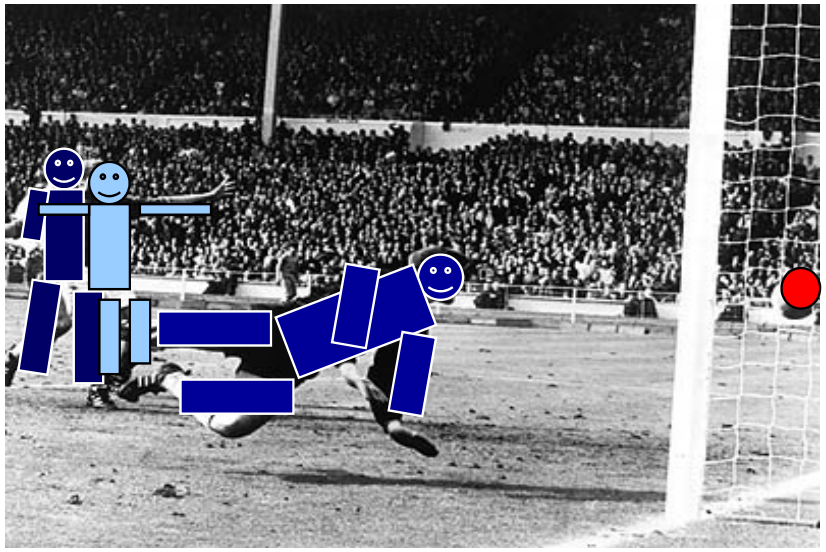


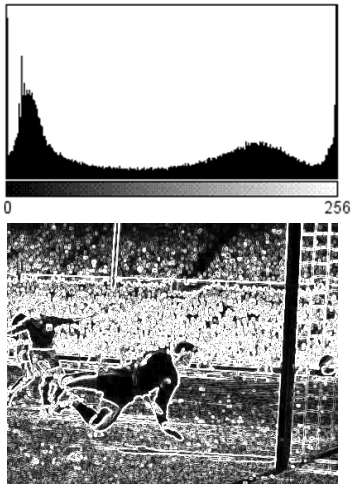
Feature Types

Level 1: Primitive features

Level 2: Logical features

Level 3: Abstract features





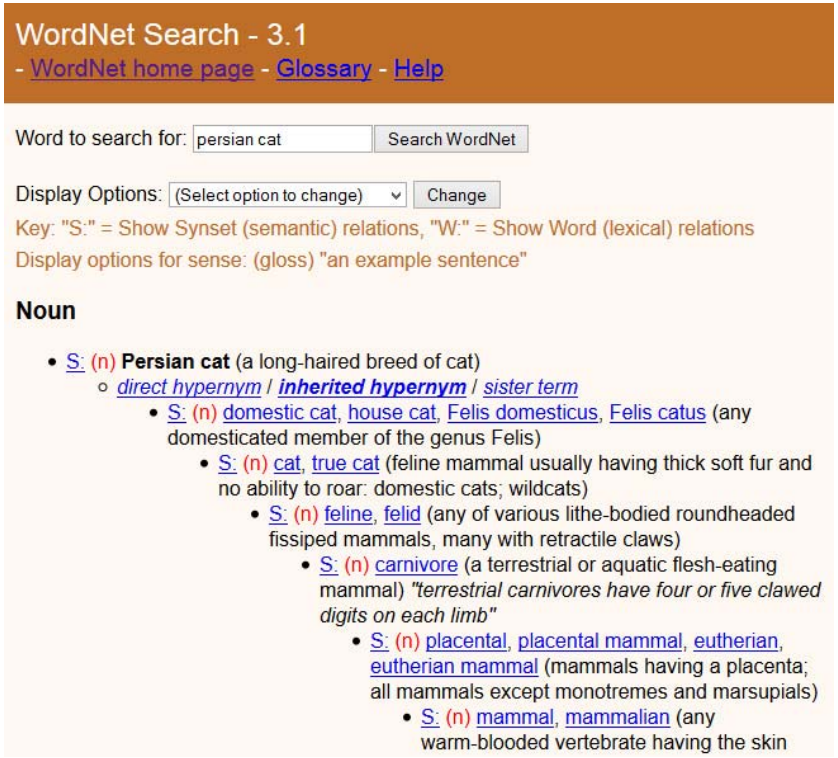
„Wembley goal“: World Cup final, July 30th 1966, England – BRD, minute 101, 3:2, end result: 4:2, by: Geoff Hurst; Goal Keeper: Hans Tilkowski; the swiss referee Gottfried Dienst decides after discussion with sowjet line referee Tofiq Bəhramov

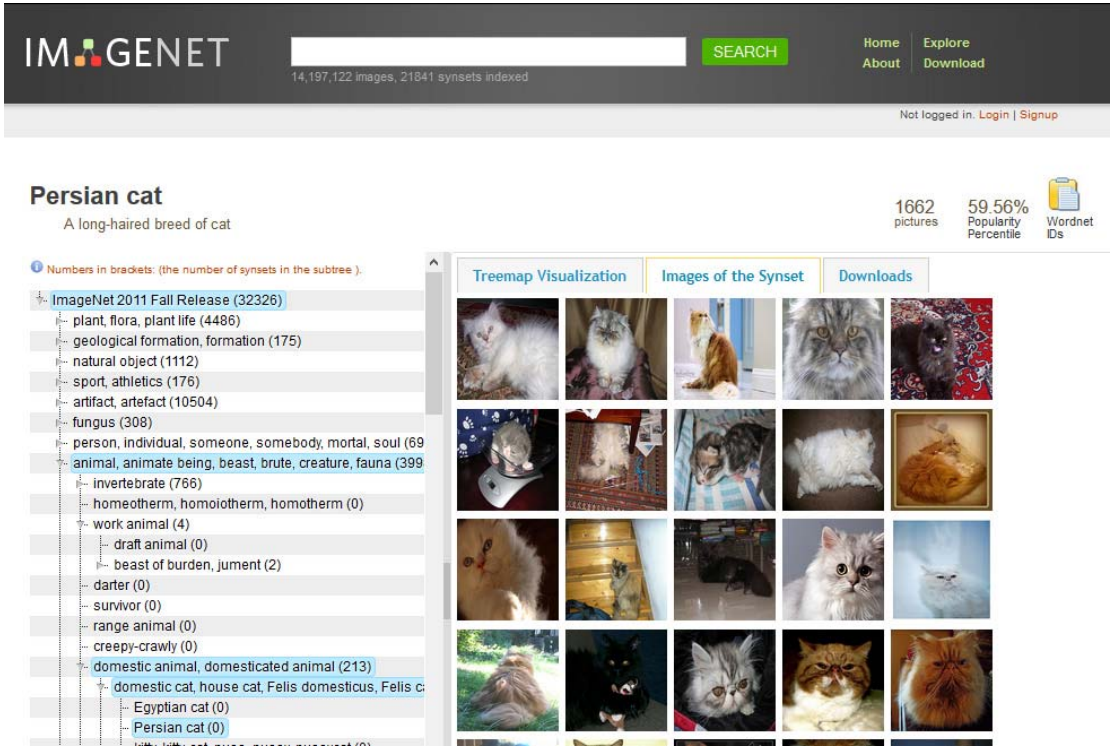
Semantic Gap

- Annotation
- Gamification

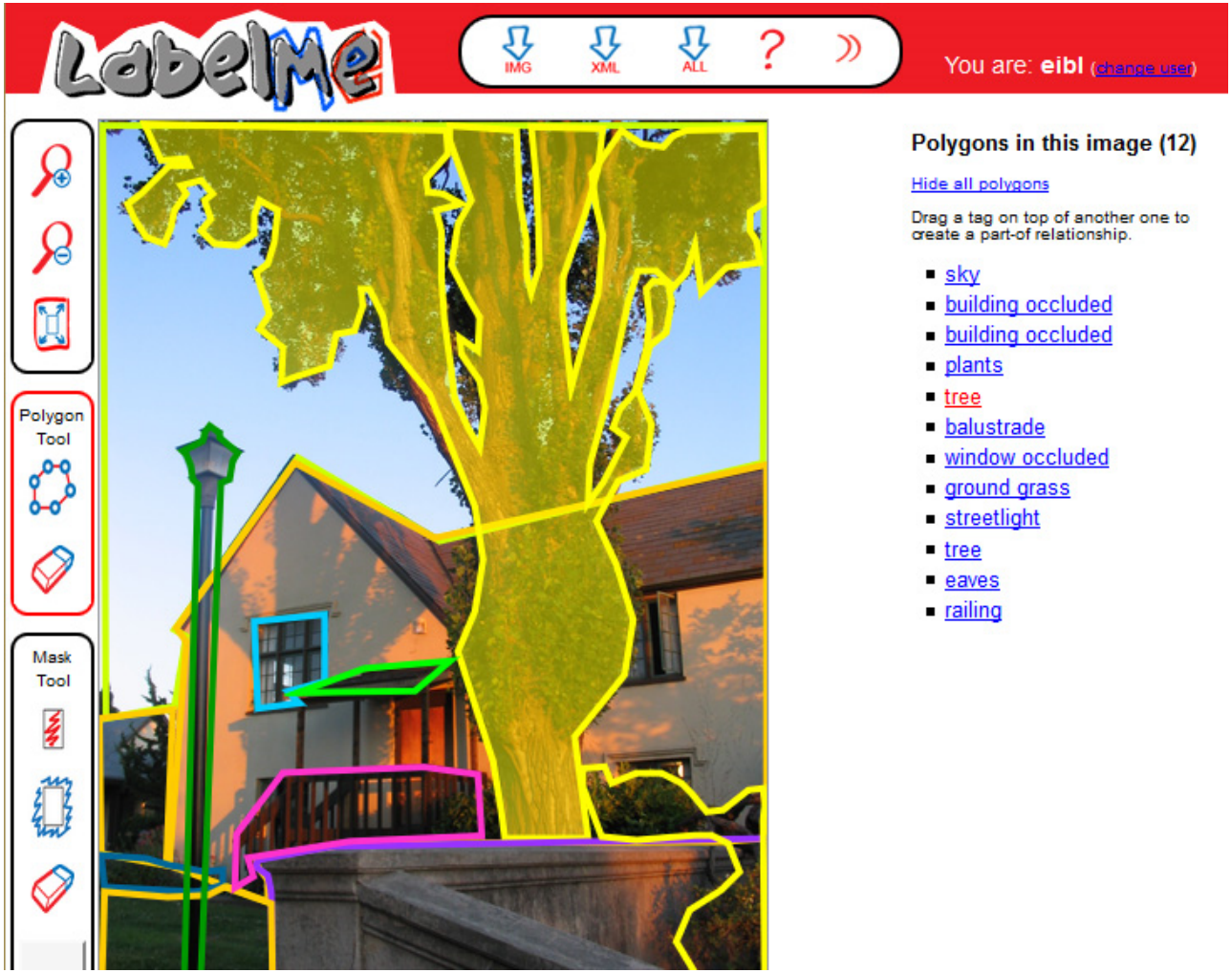
ImageNet

- Basis: WordNet (http://wordnet.princeton.edu/)
- Homepage: http://www.image-net.org/





MIT Label Me


LabelMe
IMG
XML
ALL
You are: eibl (change user)
Polygon Tool
Mask Tool
Polygons in this image (12)
Hide all polygons
Drag a tag on top of another one to create a part-of relationship.
sky
building occluded
building occluded
plants
tree
balustrade
window occluded
ground grass
streetlight
tree
eaves
railing

MIT Label Me - data

```
<?xml version="1.0"?>
- <annotation>
    <filename>123_2334.jpg</filename>
    <folder>static_outdoor_bay_area_submitted_alyosha_efros</folder>
  - <source>
      <sourceImage>The MIT-CSAIL database of objects and scenes</sourceImage>
      <sourceAnnotation>LabelMe Webtool</sourceAnnotation>
    </source>
  + <object>
  + <object>
  + <object>
  + <object>
  + <object>
  + <object>
  + <object>
  + <object>
  - <object>
      <name>streetlight</name>
      <deleted>0</deleted>
      <verified>0</verified>
      <date>07-May-2007 16:53:13</date>
      <id>8</id>
    - <polygon>
        <username>anonymous</username>
      - <pt>
          <x>196</x>
          <y>1956</y>
        </pt>
      - <pt>
          <x>227</x>
          <y>838</y>
        </pt>
```

Semantic Gap: Image Labeling – ESP Game

a image

b entered descriptions

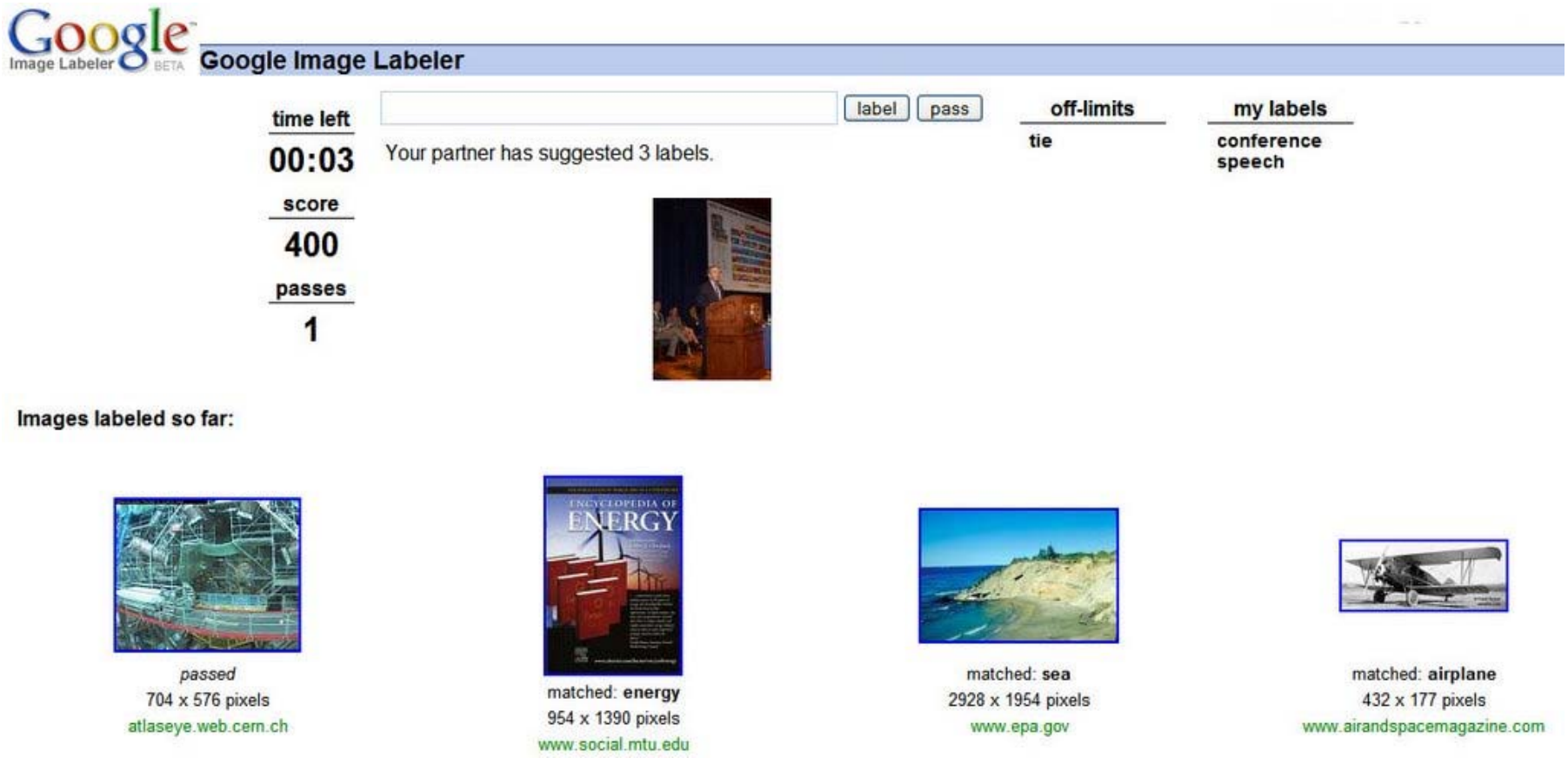
c taboo words

d match

e pass

https://www.interaction-design.org/images/encyclopedia/social_computing/fig3_social_computing_research_social_media_plawindow_illustration.jpg

Google Image Labeler



Models of CBIR

- Color-Based Modeling
- Texture-Based Modeling
- Shape-Based Modeling
- Spatial-Based Modeling

→ Level 1&2

→ Level 3: Semantic Gap

Color Space

- Different existing color spaces:
  - Physiological: CIE-RGB, CIE-XYZ
  - Psychological: CIE L*a*b*, CIE L*u*v*
  - Hardware oriented: RGB, CMY(K), YIQ
  - User oriented: HLS, HCV, HSV, HSB, MTM
- For Image Retrieval we need perceptual uniformity (*einheitliche Wahrnehmung*).

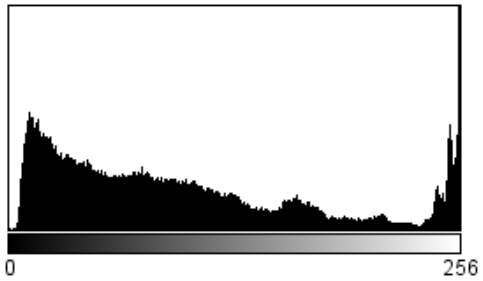
→CIE L*a*b*, CIE L*u*v* are hardware independend and well suited

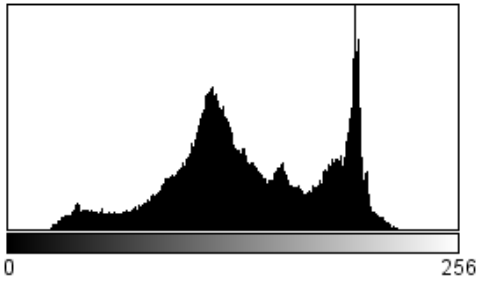
Color Moments

- First order: mean (*Mittelwert*) $\mu_i = \frac{1}{N}\sum_{j=1}^{N} f_{ij}$
- Second order: variance (*Varianz*) $\sigma_i = \left(\frac{1}{N}\sum_{j=1}^{N}\left(f_{ij} - \mu_i\right)^2\right)^{\frac{1}{2}}$
- Third order: skewness (*Verzerrung*) $s_i = \left(\frac{1}{N}\sum_{j=1}^{N}\left(f_{ij} - \mu_i\right)^3\right)^{\frac{1}{3}}$
- where:
  - *fij* : value of the *i* color component in pixel *j*
  - *N* : number of pixels in image

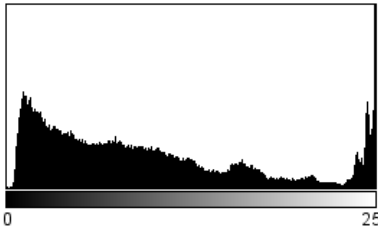
Color Histogram

- Effective representation of colors of an image when color distribution is unique
- Robust concerning translation and rotation
- Usefull for global and local color distribution



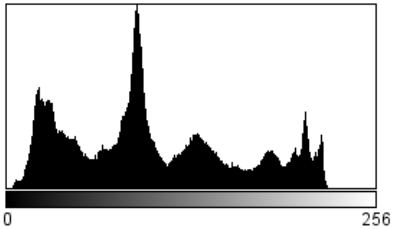


Color Histogram



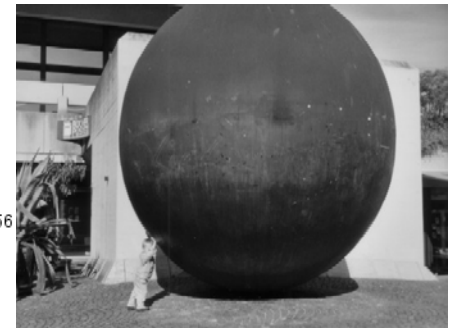
Count: 196608
Mean: 97.403
StdDev: 74.819

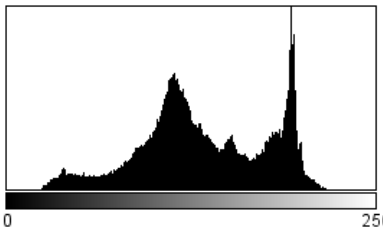
Min: 0
Max: 255
Mode: 255 (4182)



Count: 196608
Mean: 100.894
StdDev: 58.918

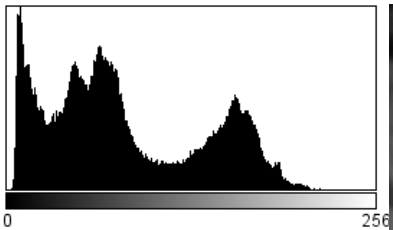
Min: 2
Max: 227
Mode: 90 (3607)





Count: 196608
Mean: 135.993
StdDev: 44.652

Min: 3
Max: 252
Mode: 197 (4179)



Count: 196608
Mean: 81.490
StdDev: 54.534

Min: 1
Max: 229
Mode: 9 (2644)

Color Histogram: Process

- Select color space
- Quantisize the three components of the color space to „bins“ (*Farbeimer*)
- Count absolute color distribution → fill the bins
- Calculate relative color distribution

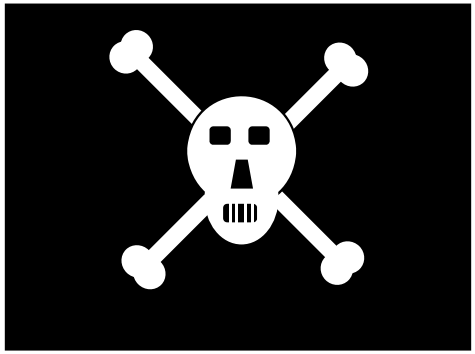
Color Histogram: Challenges

- Hughe amount of images in a data base

  → Simple comparison of histograms is ineffective

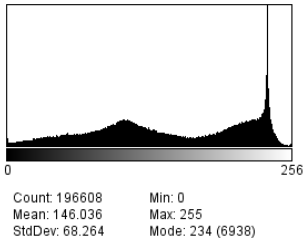
- Spatial distibution of pixels is not considered

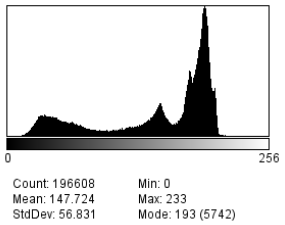
  → Different solutions for spatial distribution



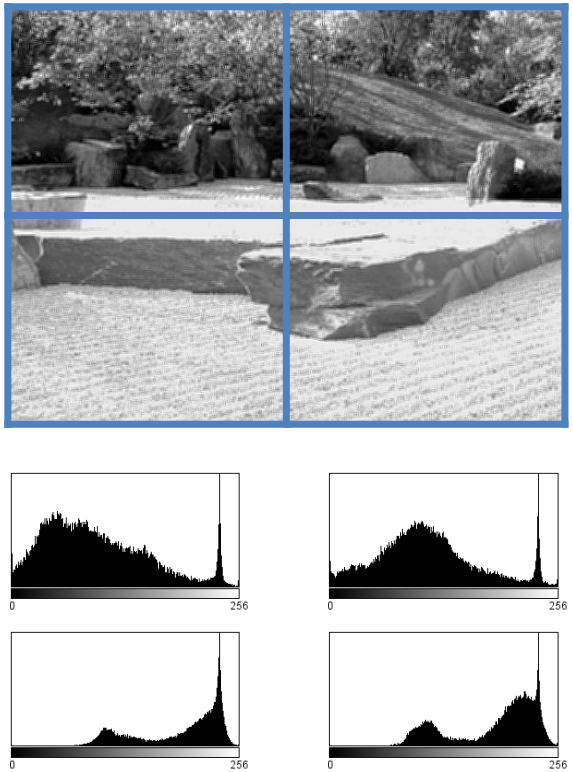
Color Histogram: Spatial Distribution

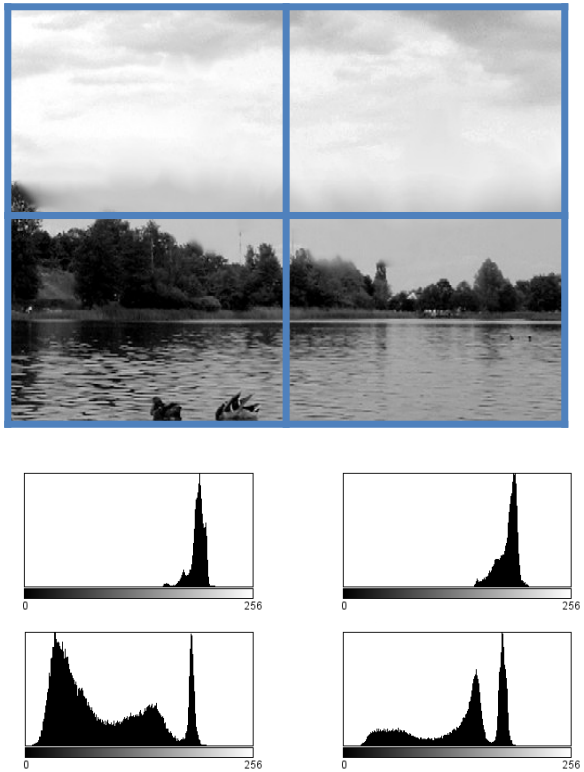
- Grid
- Annular Histogram
- Angular Histogram
- Color Coherence Vector (CVV)
- Color Correlogram
- Segmenting
- …





Color Histogram: Grid





Color Coherence Vector (CVV)

- Devide the histograms in coherent and incoherent pixels:

  CVV: <(α1,β1), (α1,β1),…, (αN,βN)>
  where
  αi: number of coherent pixels of color i of an image
  βi: number of incoherent pixels of color i of an image

  (simple hstogram: <(α1+β1), (α1+β1),…, (αN+βN)>)

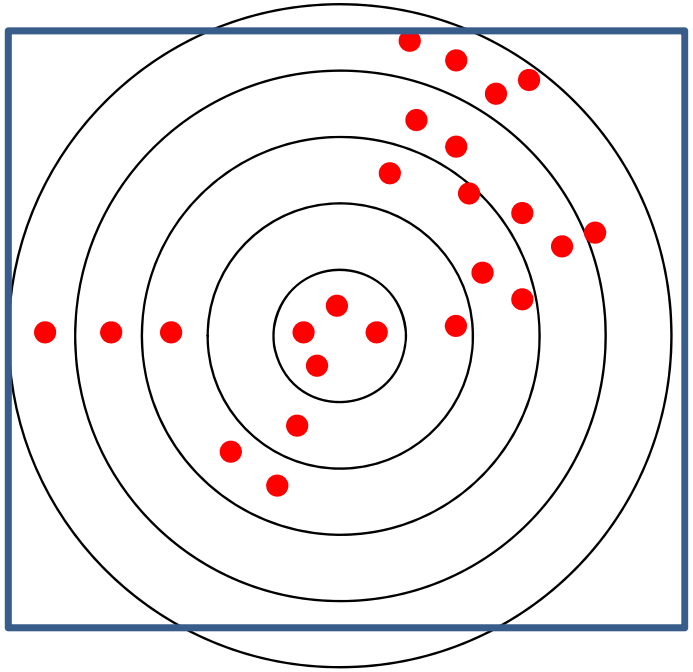
- Especially for images with large cohrent color regions.
- HSV performs better than CIE L*u*v* oder CIE L*a*b*

Color Histogramm: annular

1. Divide the image in circular regions
2. Count colors from the center to the outer regions

→ annular distribution density vector :

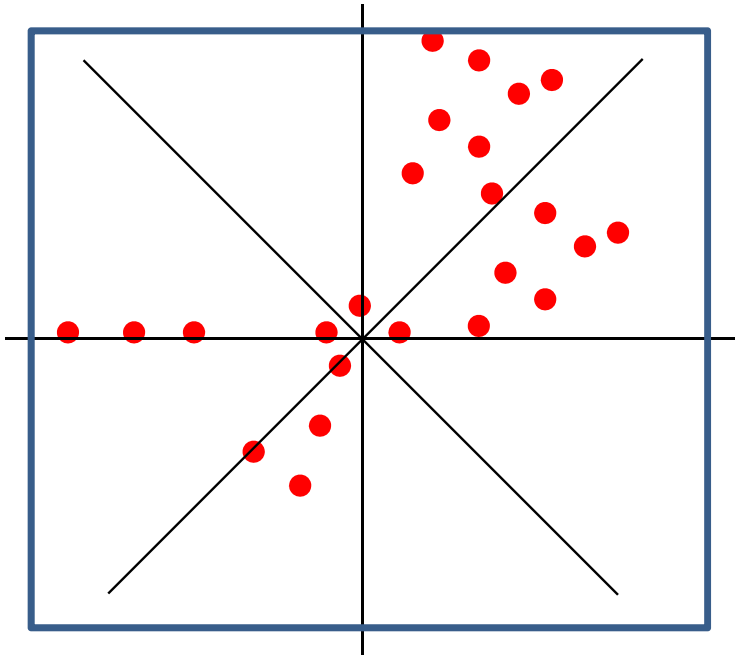
<4,2,7,5,6>



siehe: Rao et al. 1999

Color Histogramm: angular

1. Divide the image in angles
2. Count the colors counter clockwise
3. → *angular distribution density vector* :
   <7,8,1,4,0,4,0,0>

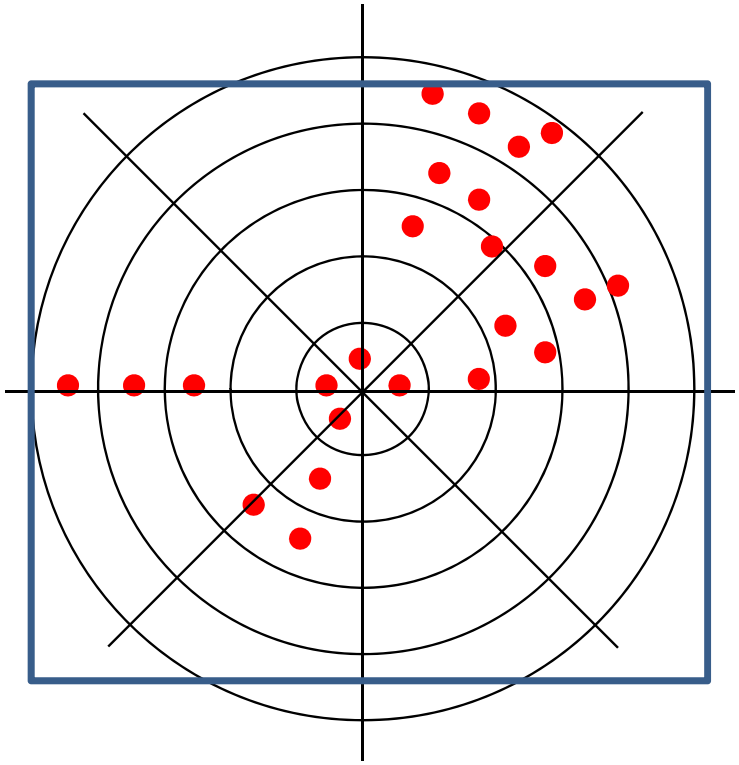


siehe: Rao et al. 1999

Color Histogramm: hybrid

1. Divide the image in angles and circles
2. Count the colors counter clockwise from the center to the outer regions
3. → *hybrid distribution density vector* :

   <1,0,1,1,0,1,0,0,1,0,0,0,0,1,0,0,2,2,0,1,0,
   2,0,0,2,2,0,1,0,0,0,0,1,4,0,1,0,0,0,0>



siehe: Rao et al. 1999

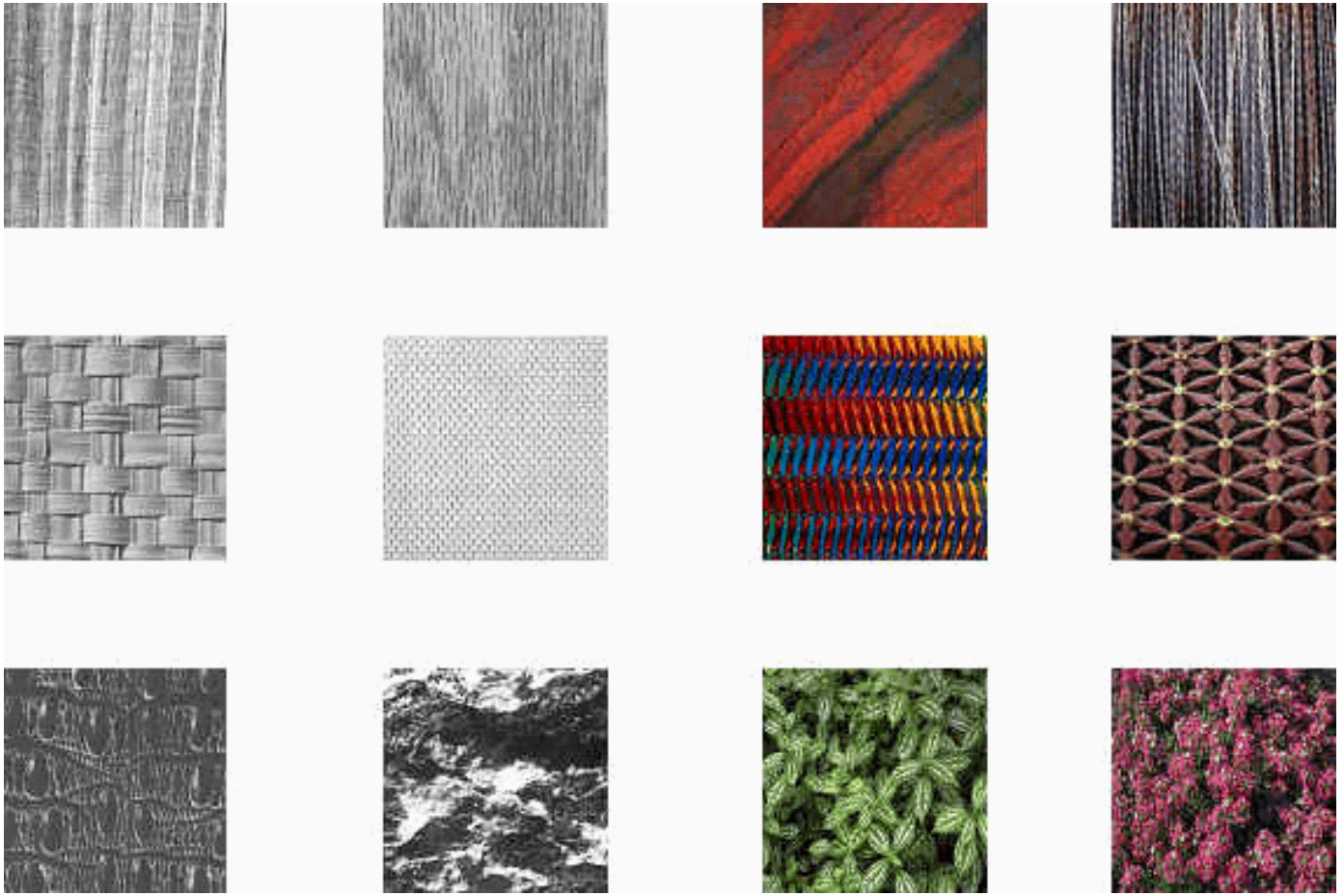
Color Histogram: Test

- Test von Rao, Srihari&Zhang (1999)
- Test setting: 500 images in 41 similarity groups
- 5 types of histograms: „normal“, CCV, annular, angular, hybrid
- Results - performance:
  1. Annular
  2. Angular
  3. Hybrid
  4. CCV
  5. „normal“

Textur

- „…the way a fabric or other substance feels to the touch, its thickness or firmness or solidity.“ (Oxford American Dictionary)
- „…something composed of closely interwoven elements. Frequent repetition of similar elements (*texemes*).“ (*Webster’s Dictionary*)
- „In very general terms, it is reasonable to say that a texture is a combination of a relatively large number of objects in which the individuality of each of object is not regarded as important, and in which the structural properties of the set of objects is predominant.“ (Santini 2001)

Examples of Textures



Texture Representation Methods

Tamura Features
- *coarseness (*Granularität)
- *contrast (*Kontrast)
- *directionality (*Richtung)
- *linelikeness (*Linienhaftigkeit)
- *regularity (*Regelmäßigkeit)
- *roughness (*Rauhigkeit)

Rao & Lohse Features
- *repetitivness (*Wiederholung)
- *directionality (*Richtung)
- *granularity & complexity (*Granulatität & Kompexität)

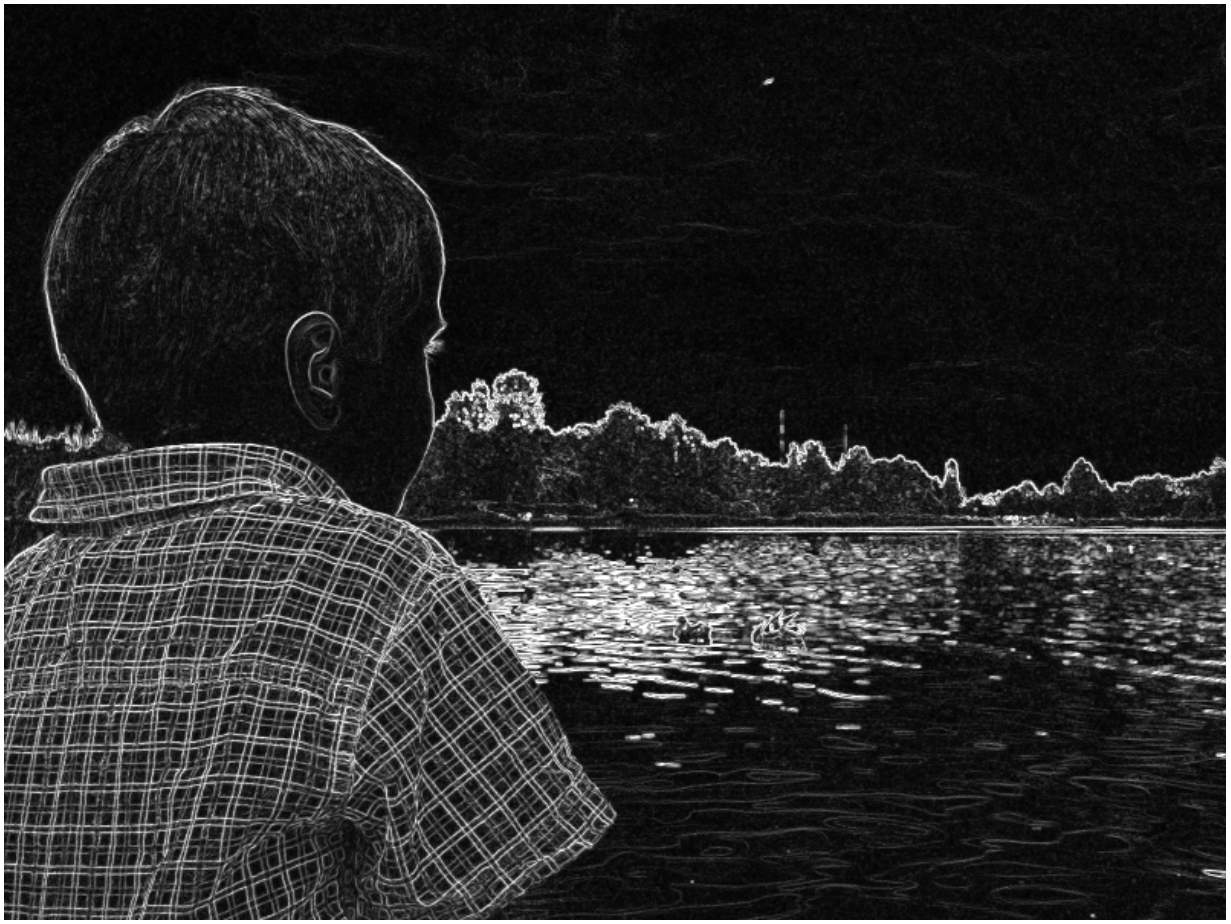
Wold Decomposition
- *harmonic (*harmonisch)
- *evanescent (*flüchtig)
- *indeterministic (*unbestimmt)

Amundson & King Features
- *coarseness (*Einfachheit, Granulatität)
- *contrast (*Kontrast)
- *busyness (*Hektik)
- *complexity (*Komplexität)
- *texture strength (*Wirksamkeit)

Shape-Based Modeling: simple measures

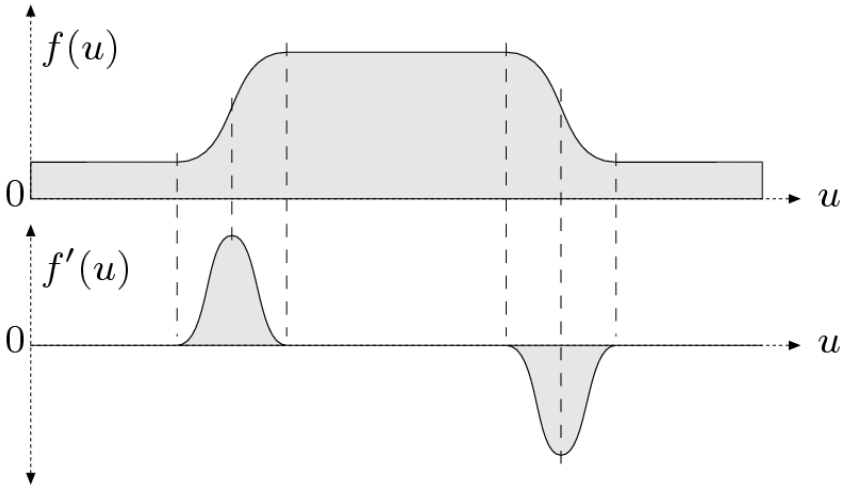
- Circumference (Umfang)
- Area (Fläche)
- Compactness
- Center of gravity (Schwerpunkt)

Edge Detection



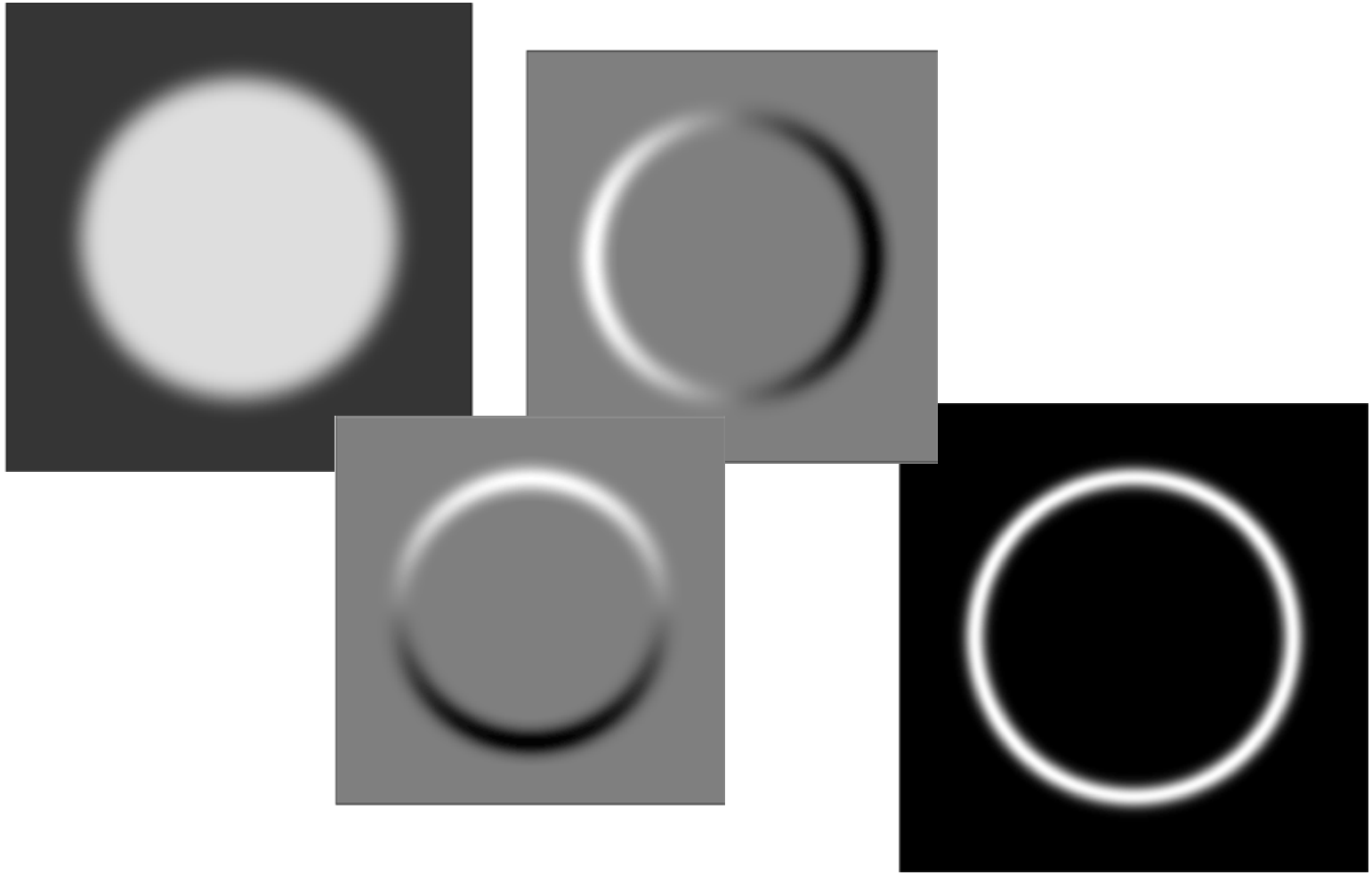
Gradient-based Edge Detection

(a)

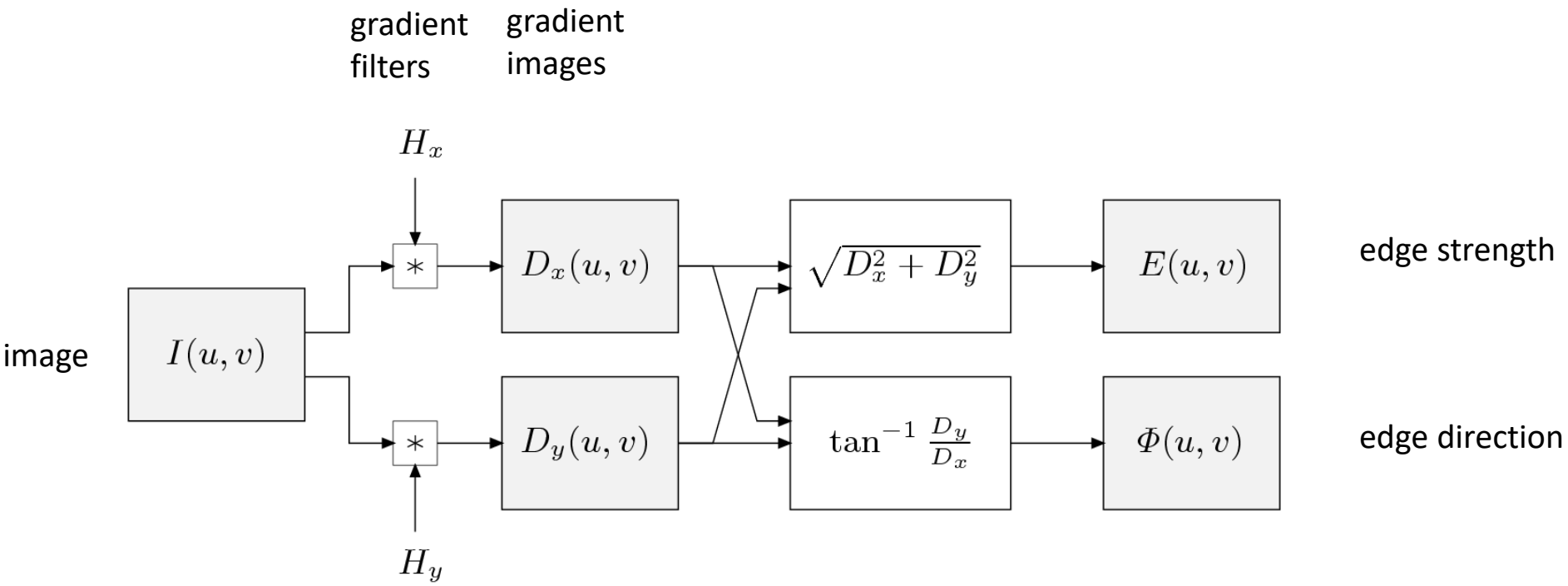


(b)

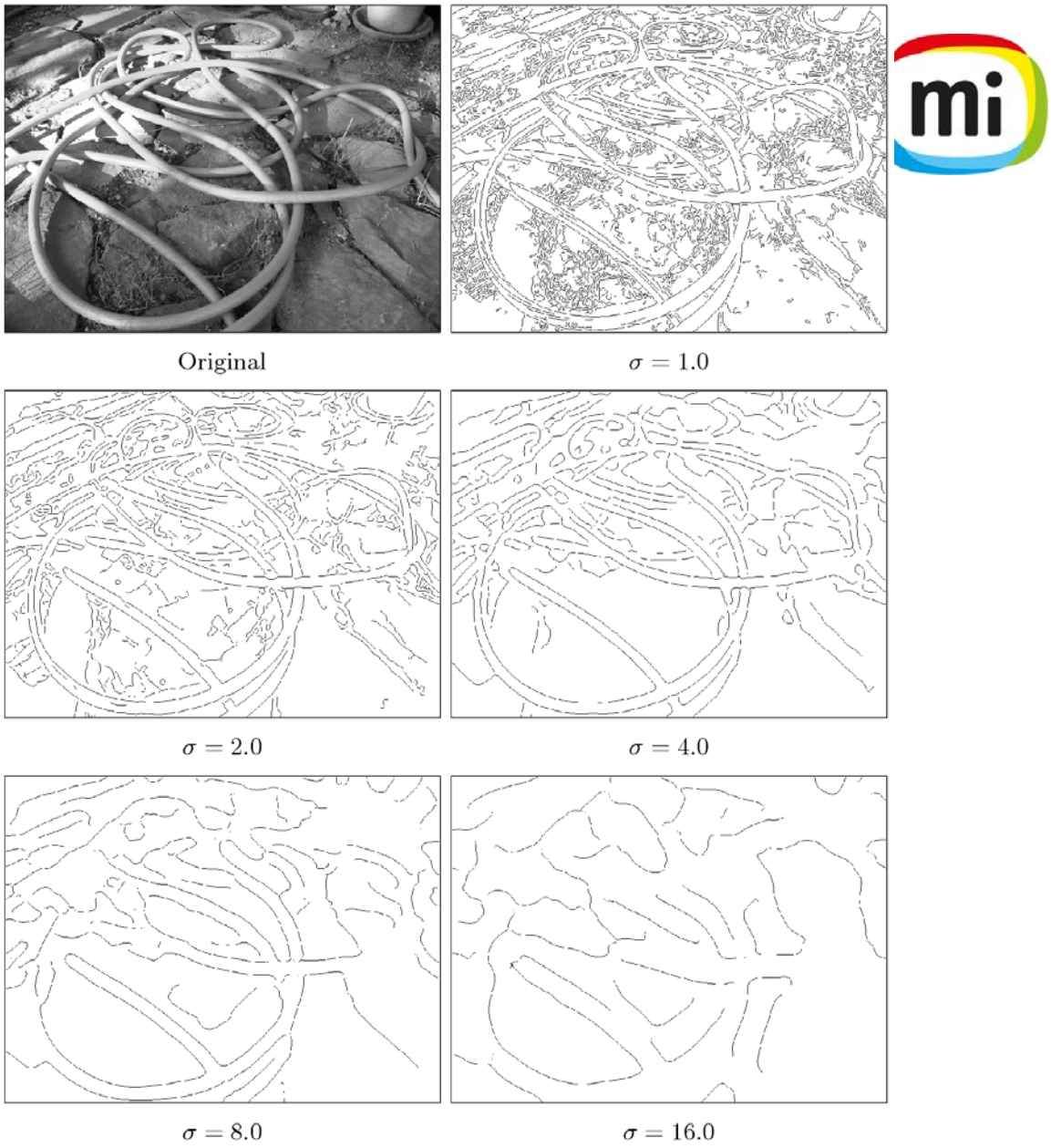
Gradienten-basierte Kantendetektion IV



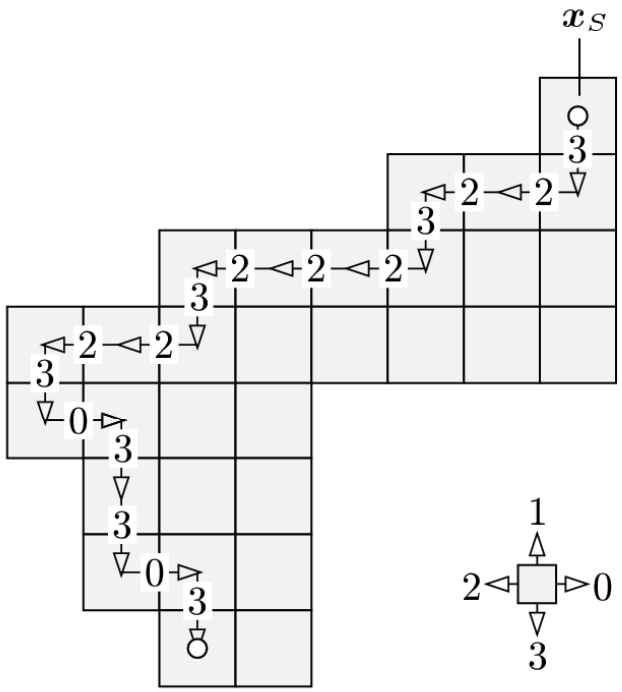
Process Edge Detection



Canny-Filter



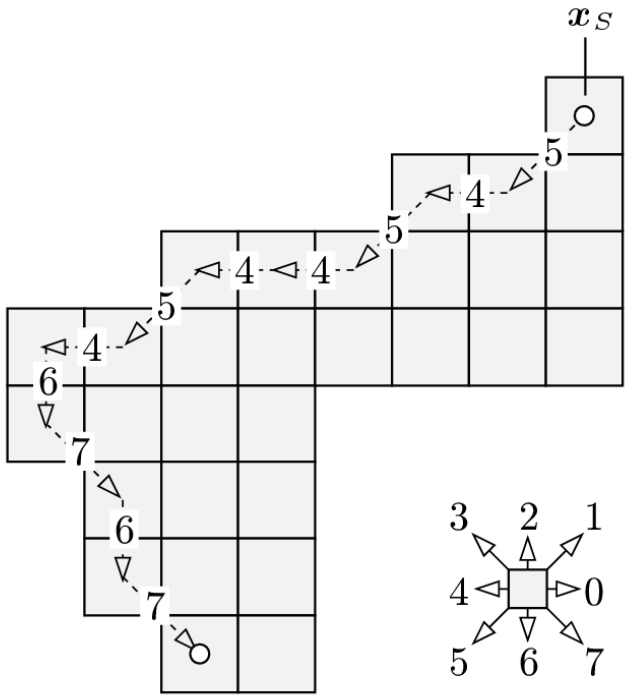
Codierung: Chain Codes



4-Chain Code

3223222322303303...111

$length = 28$



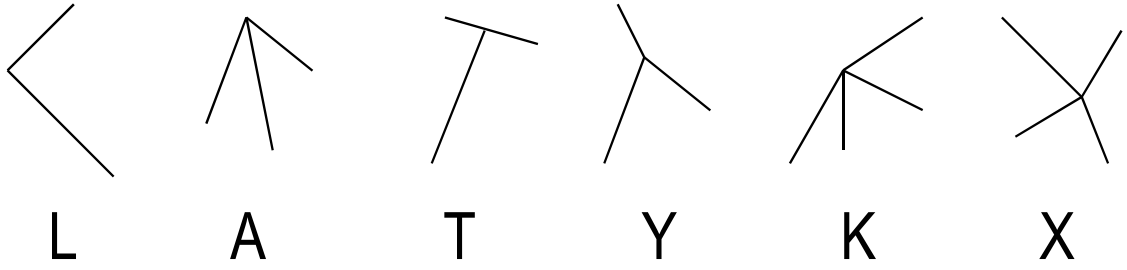
8-Chain Code

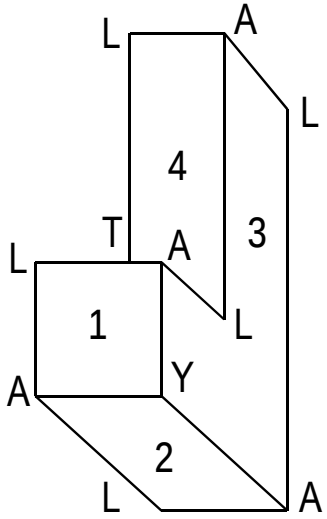
54544546767...222

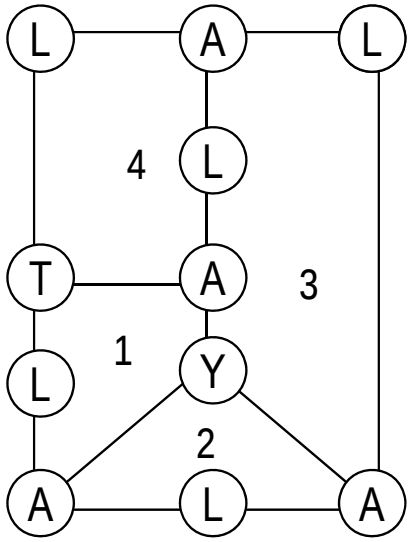
$length = 18 + 5\sqrt{2} \approx 25$

Codierung: LISG

- Labeled Image Structure Graph

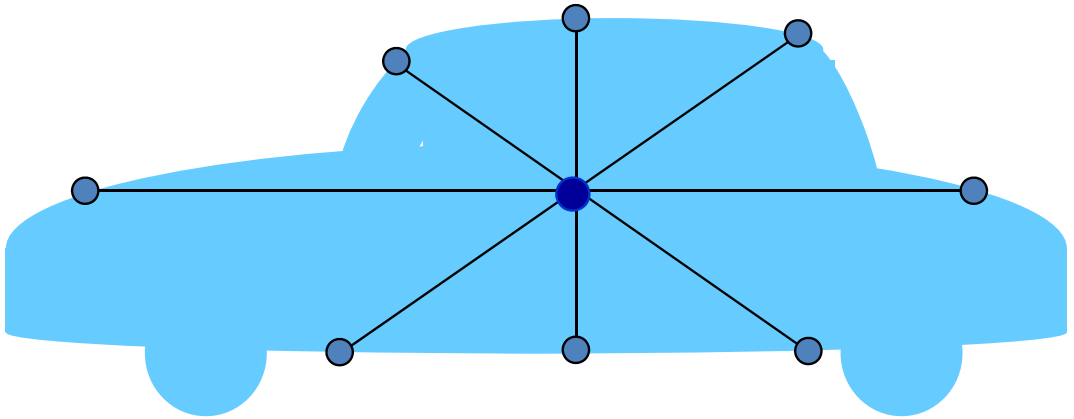




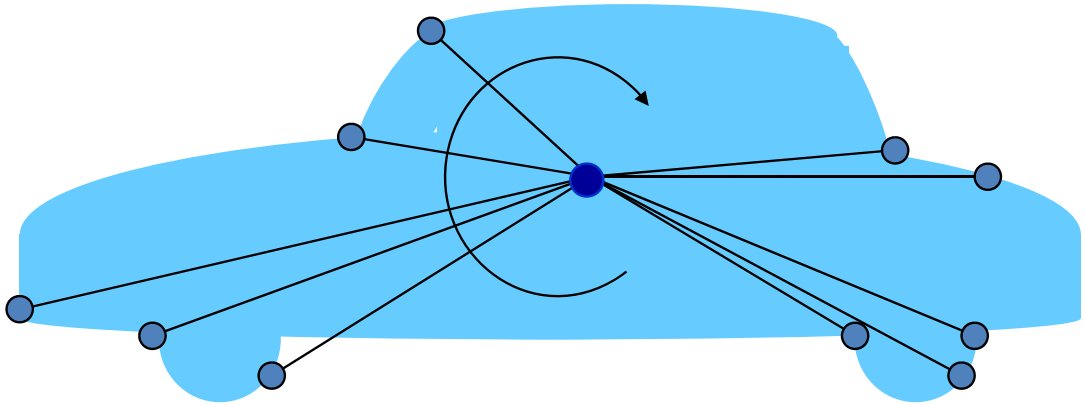


Outer Shape:
LALALTLA
Inner Shapes:
1: AYALT
2: YALA
3: LAYALA
4: ALATL

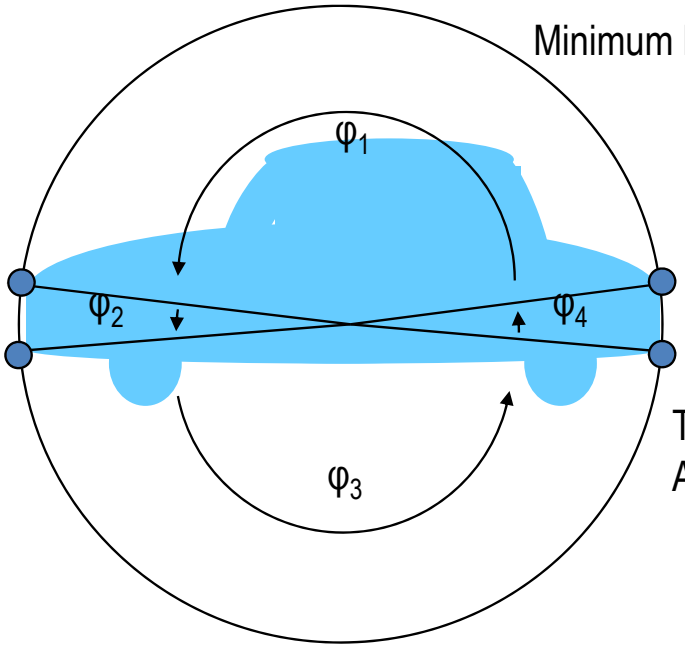
Shape Radii Representation

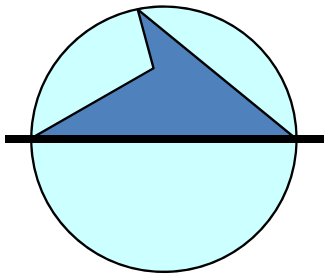


Turning Point Representation

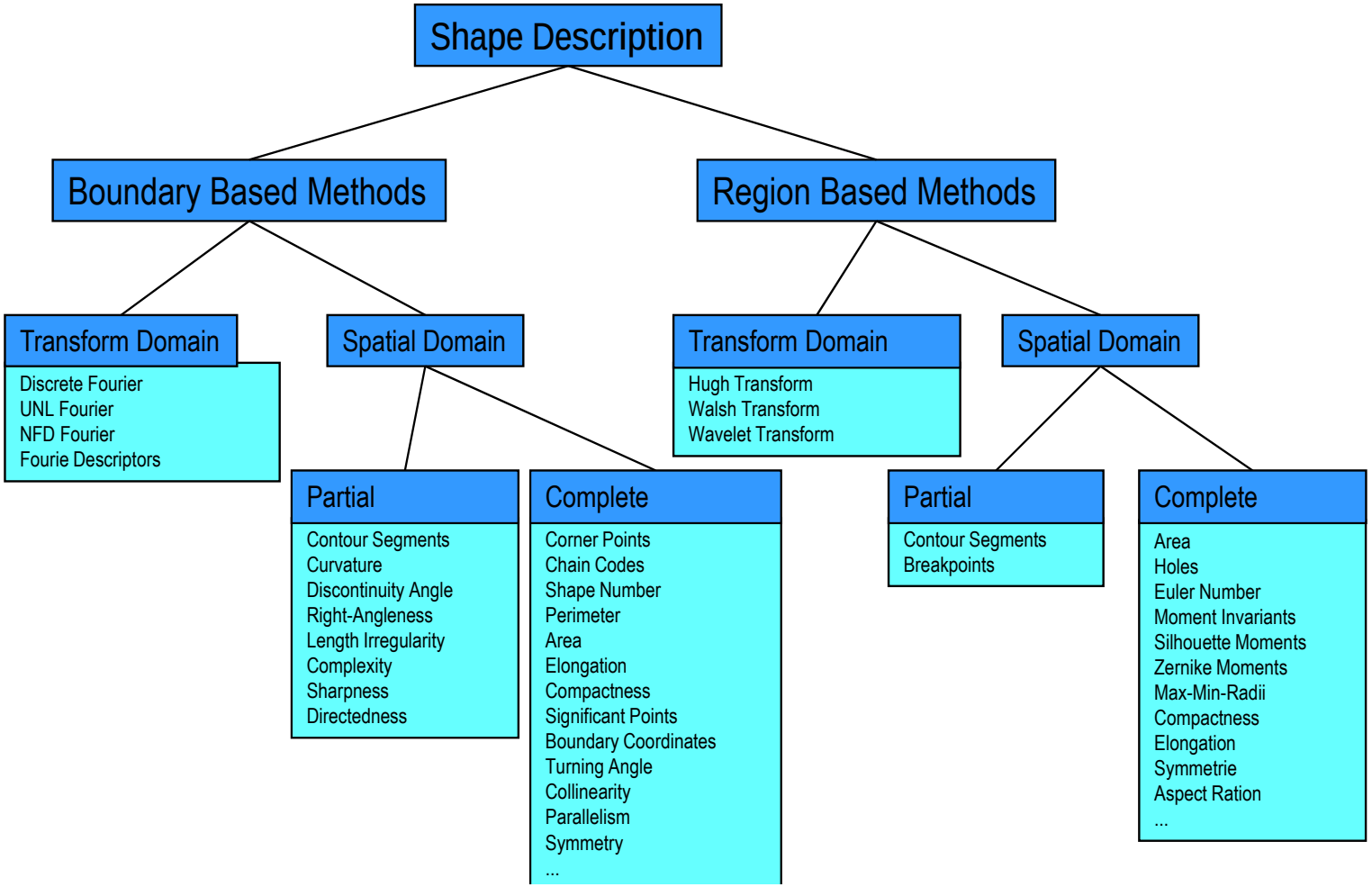


Minimum Bounding Circle (MBC)





Taxonomie Shape Based Retrieval



Normalisierung

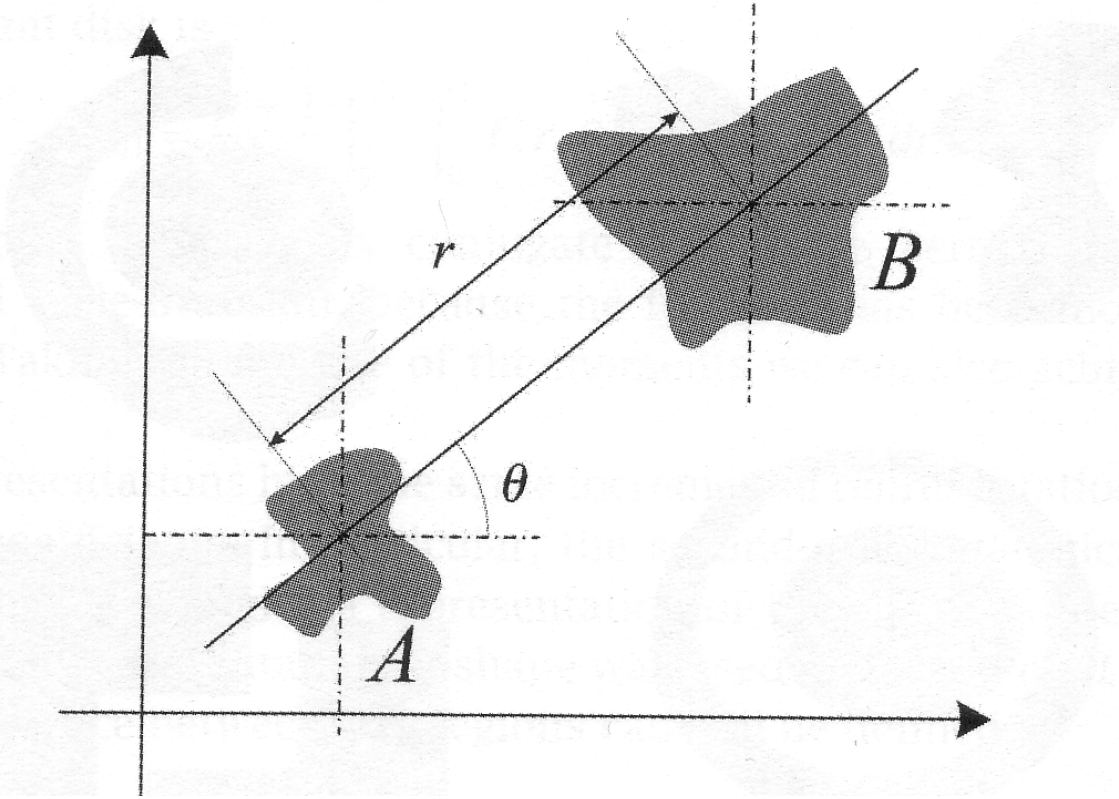
- Rotation

- Scale

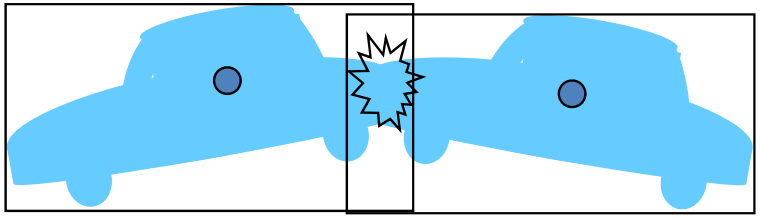
Spatial Relations

Type of Relation	Invariant to...	Example
Abolute	None	at position x,y
Distance	Translation	close to
Directional	Scale	left of
Topological	Continuous	Overlapping

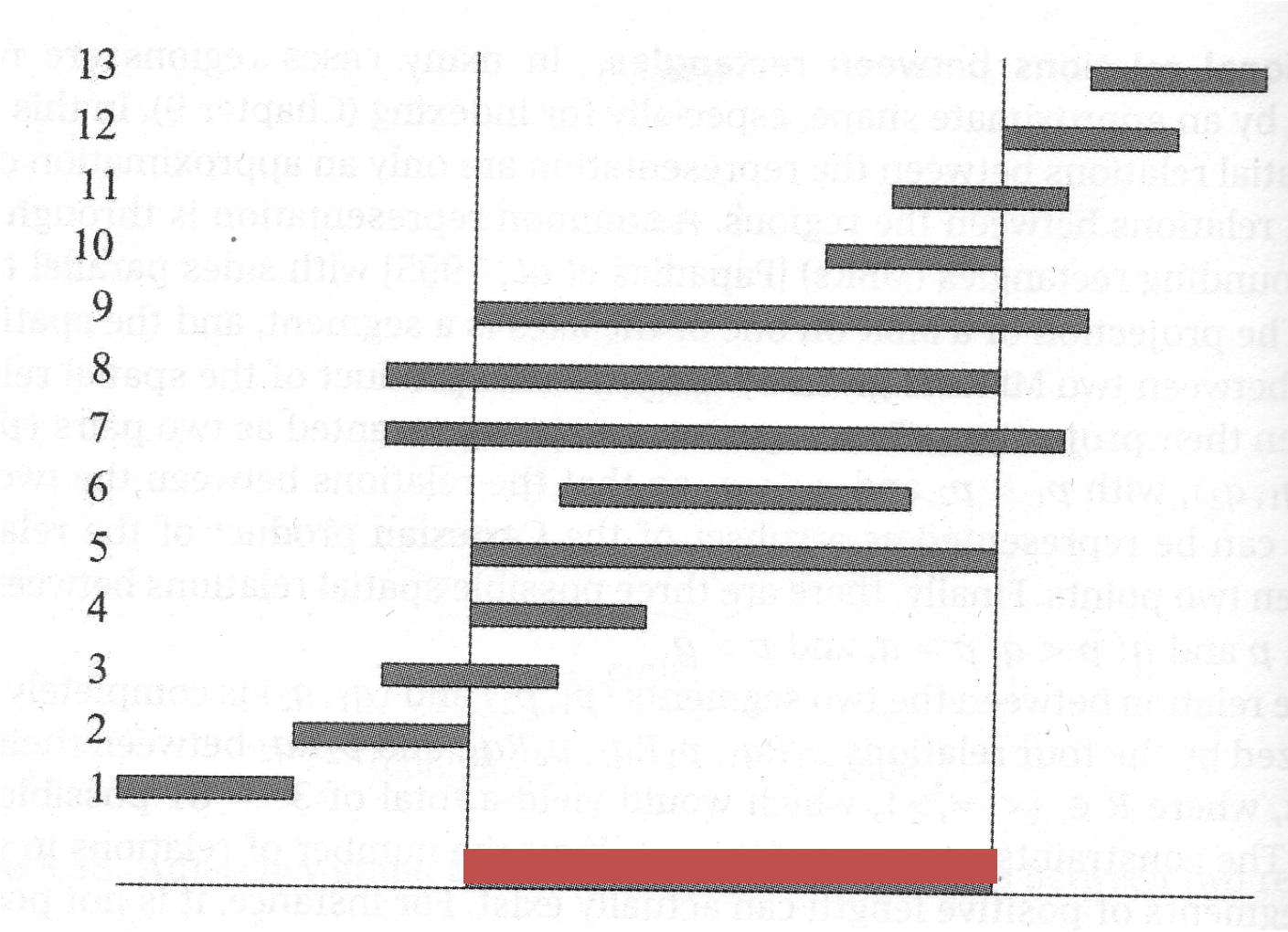
Comparison of centroids (*Schwerpunkte*)



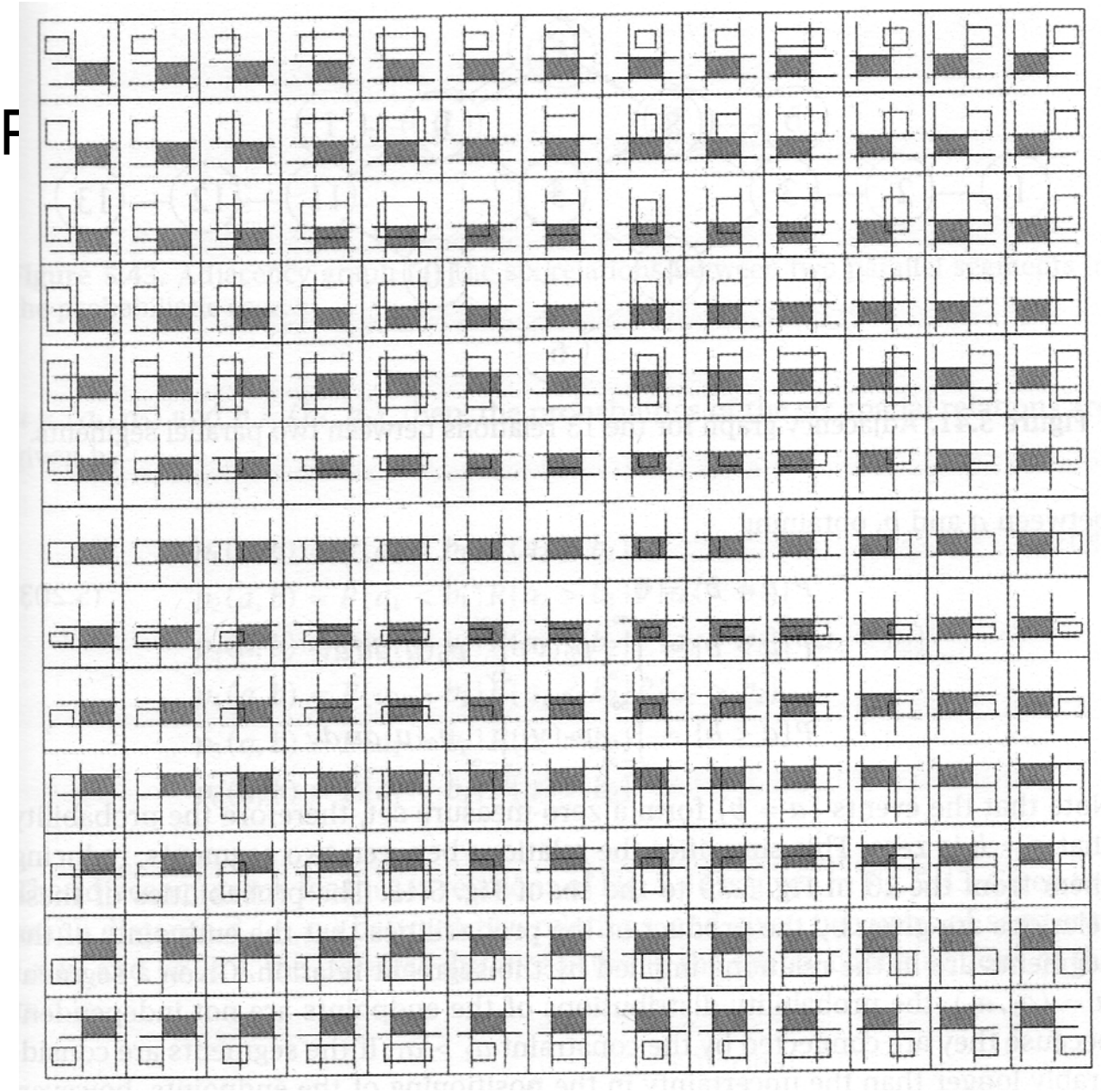
Topological Relations



Spatial Relations



Spatial R



169 possible spatial relations betwee two rectangles

Topological Relations