VK Multimedia Information Systems

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

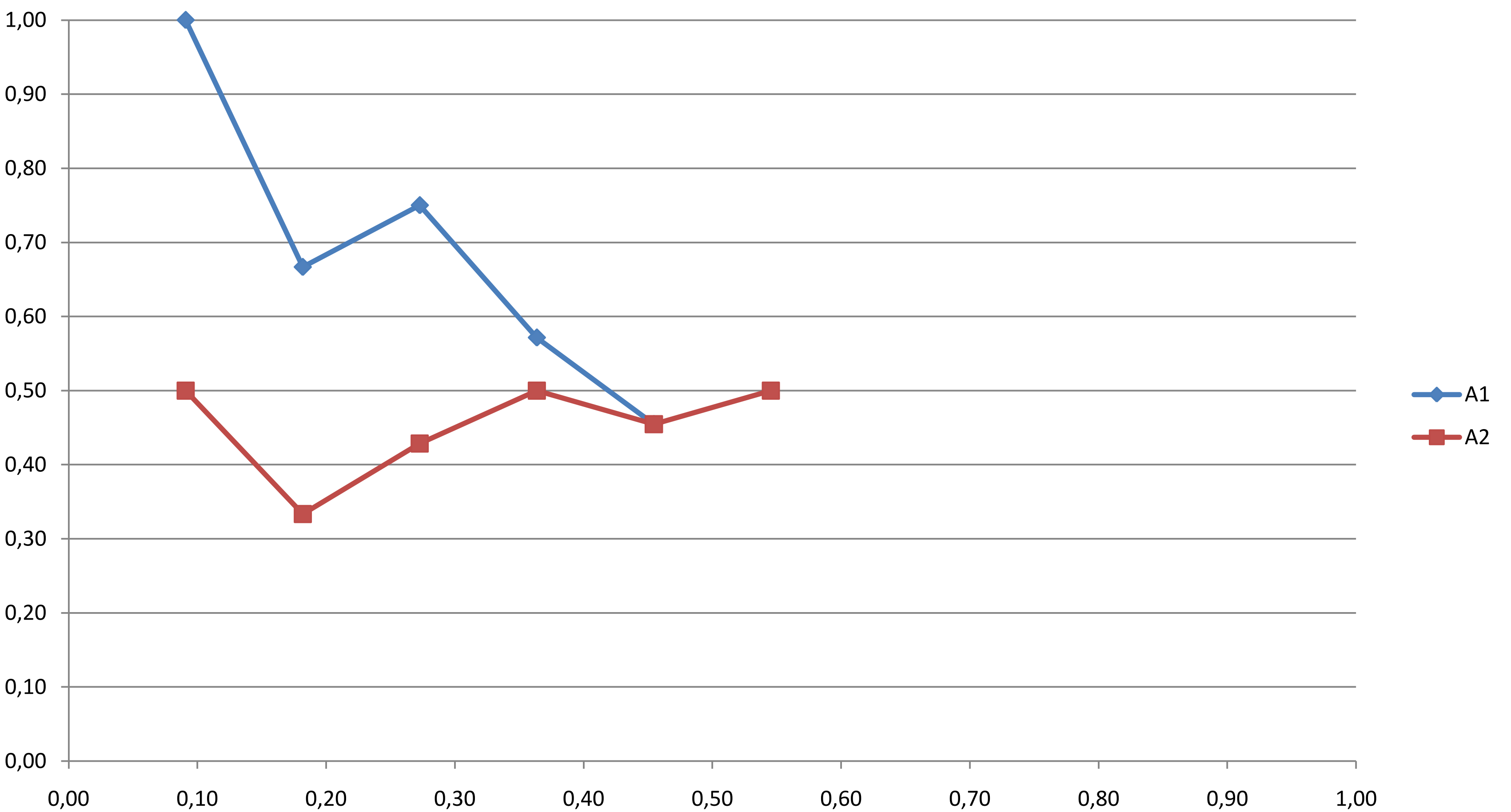
Dienstags, 16.oo Uhr c.t., E.1.42

Results

- A1:
  - 0,455 Precision
  - 0,455 Recall
- A2:
  - 0,500 Precision
  - 0,545 Recall

RC-Lev	PR-A1	PR-A2
0,09	1,00	0,50
0,18	0,67	0,33
0,27	0,75	0,43
0,36	0,57	0,50
0,45	0,45	0,45
0,55		0,50
0,64		
0,73		
0,82		
0,91		
1,00		

Results Ex-04



Content Based Image Retrieval

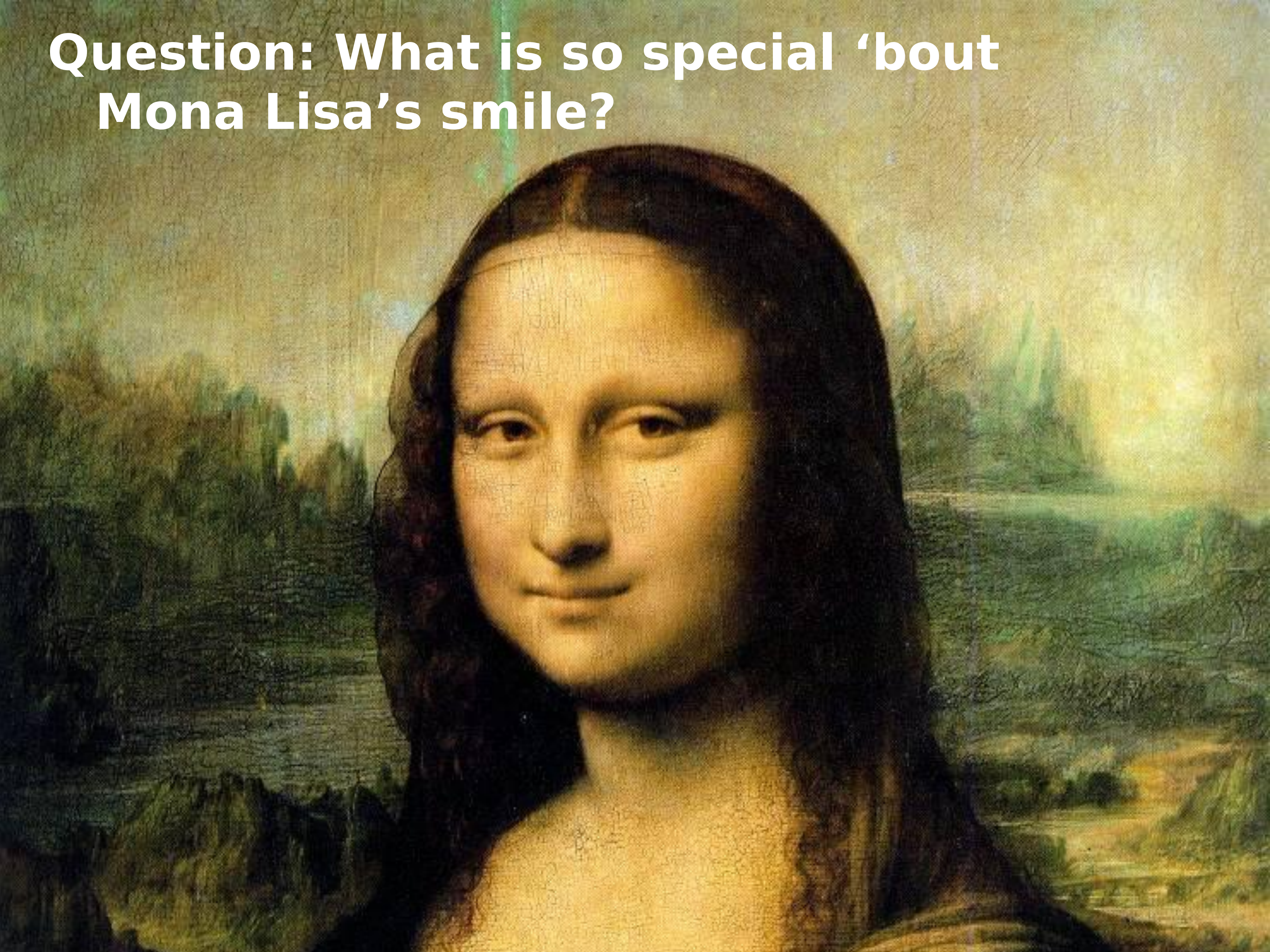
- Motivation & Semantic Gap
- Perception
- Color Based Features
- Texture Based Features
- Exercise 5

Motivation

Lots of good reasons …

- Visual information overload
  - – Devices (cameras, mobile phones, etc.)
  - – Communication (email, mo-blogs, etc.)
- Metadata not available
  - – Time consuming
  - – No automation

Question: What is so special ‘bout Mona Lisa’s smile?

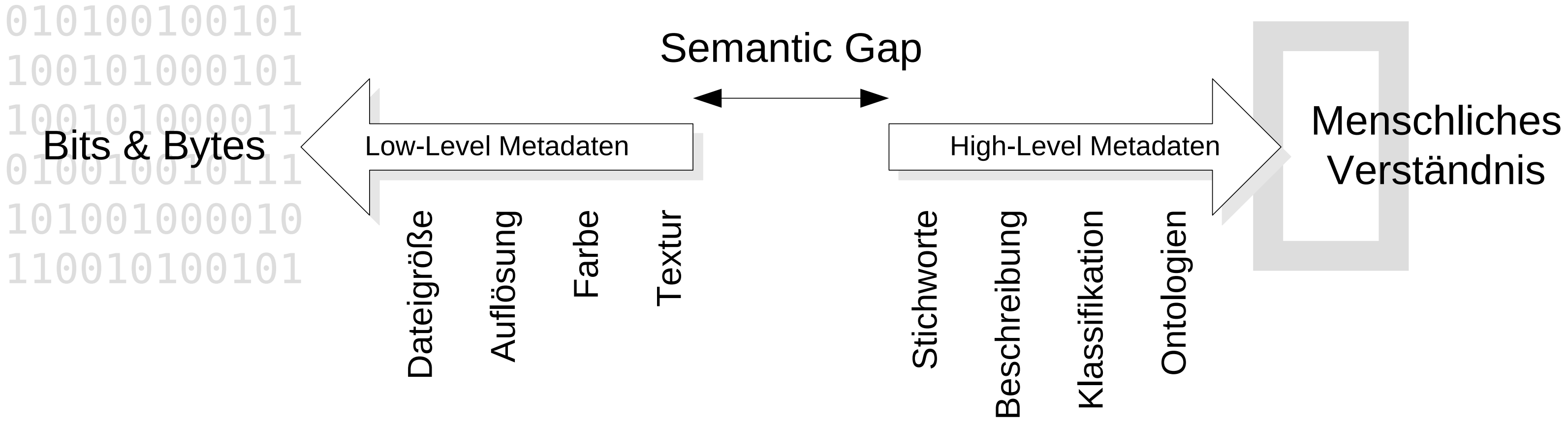


Semantic Gap

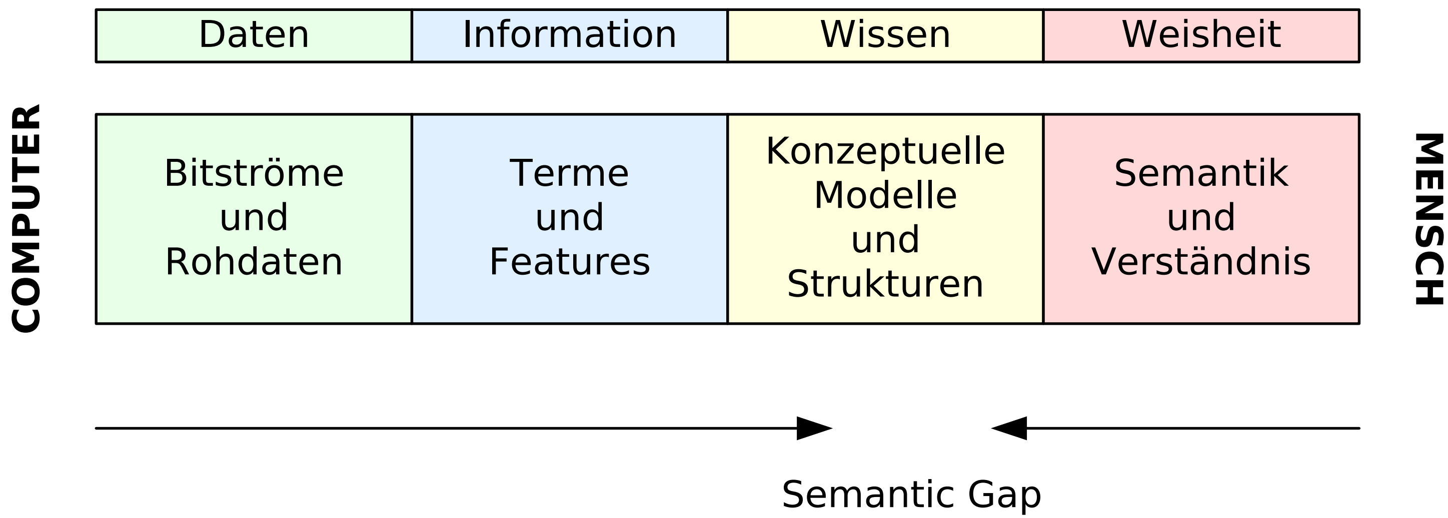
- Defined as
  - Inability of automatic understanding
  - Gap between high- and low-level features / metadata
- Actually hard task for humans also

Semantic Gap (1)

- General Definition: Santini & Jain (1998)



Semantic Gap (2)



Where actually is the Semantic Gap?

- Classification based on Conceps
- Segmentation & Object Recognition
- ...

Applications

- Home User & Entertainment
  - Find picture of / from / at ....
  - Search & browse personal digital library
- Graphics & Desgin
  - Find picture representing something (Color in CD/CI, feeling, etc.)
- Medical Applications
  - Diagnosis, segmentation & classification
  - X-Ray images, patient monitoring

Applications

- Accessibility
  - ‘Explain’ image to visually disabled people
- Industrial application
  - Select / Sort out products (chips, buns)
  - Monitor processes (e.g. sensors unavailable)
- Security
  - Match fingerprints
  - Search face database

Applications

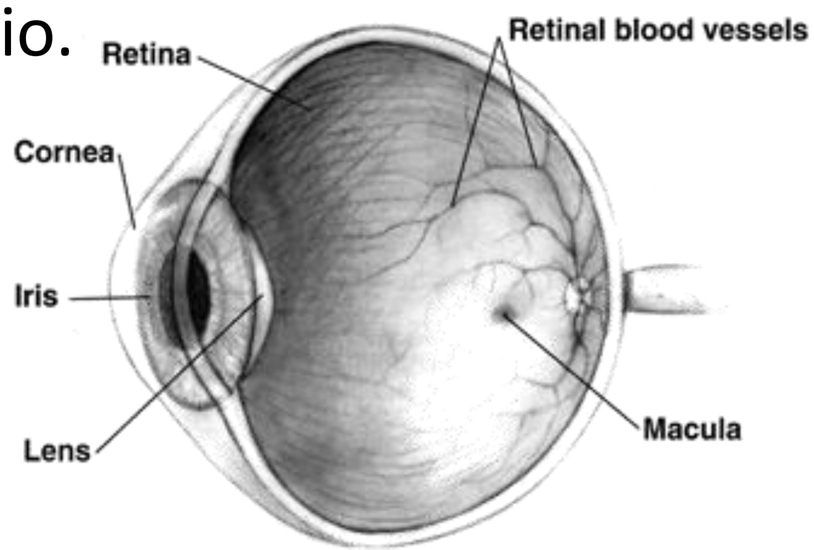
- Biology
  - Analysing cell samples
  - Recognizing animals, insects & plants
- Astronomy
  - Classifying stars & events
- Weather forecasting
  - Satellite images, clouds
- Cartography
  - Mapping (e.g. aerial photo - earth model)

Content Based Image Retrieval

- Motivation & Semantic Gap
- Perception
- Color Based Features
- Texture Based Features
- Exercise 5

Perception

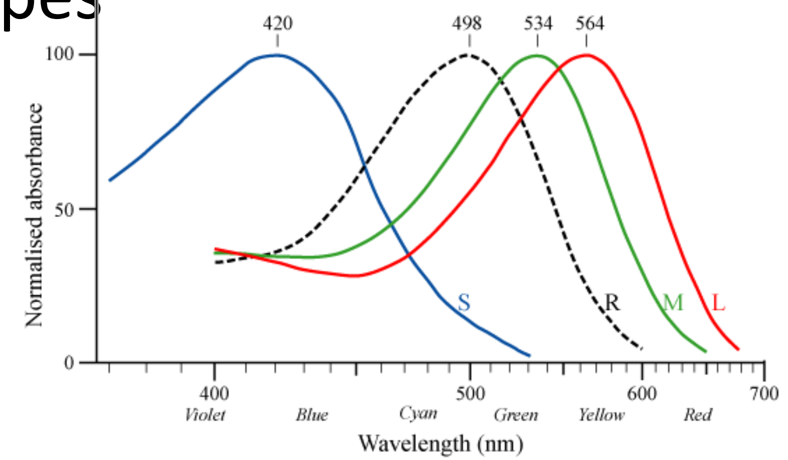
- The **eye** as instrument of perception
- Sensory capabilities
  - Cones (bright light): 6-7 Mio.
  - Rods (dim-light): 75-150 Mio.
  - Brain ‘corrects’ vision
    - e.g. blind spot



Color & Color Spaces

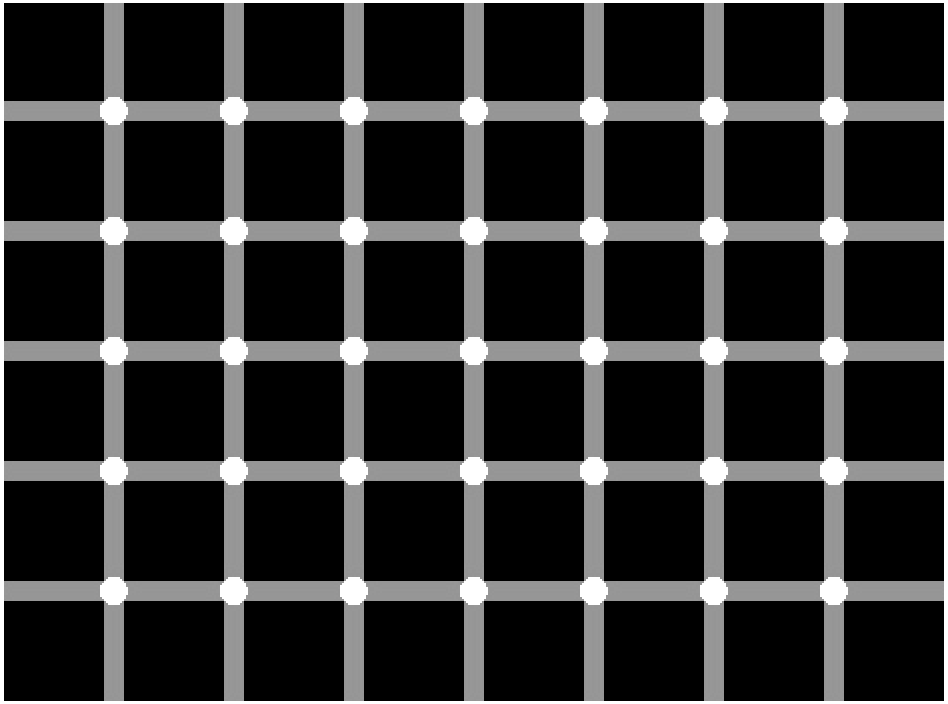
S-, M- and L-cones: Blue, green and red

- RGB based on these three colors
- CIE models perception better
  - Responsiveness of cone types
  - Number of cones / types
  - etc.



The human eye ...

- Count the black dots on the image:



The human eye …

- Rabbit or duck?

The human eye …

- Anamorphic illusions

See e.g. http://users.skynet.be/J.Beever/pave.htm

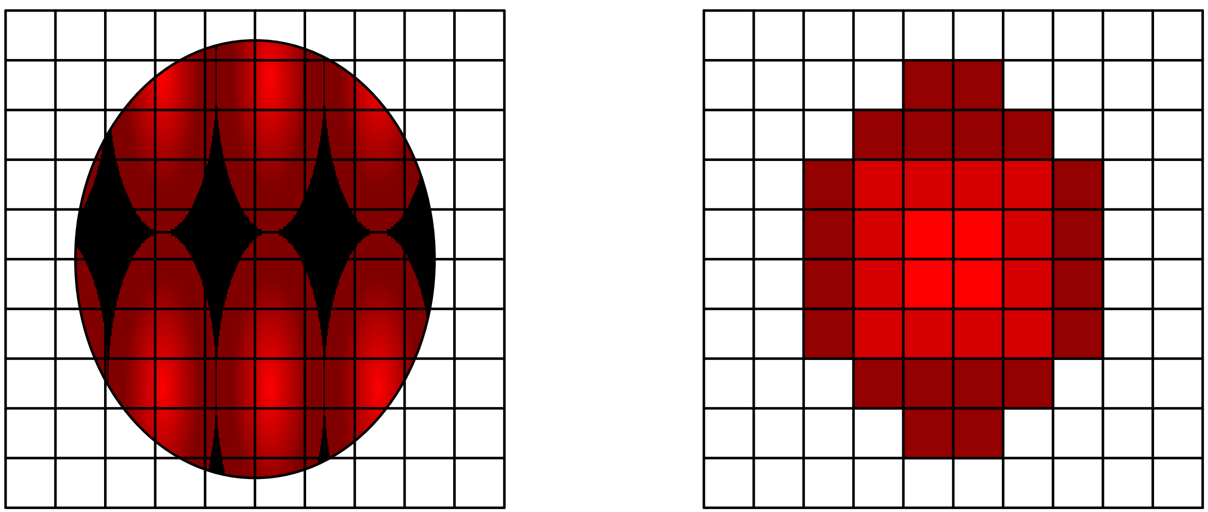
Anamorphic Illusions (Julian Beever)

What are (digital) images?

- An Image is
  - Created by a set of photons
    - With different frequency
    - Moving from different sources
    - Along different vectors
  - A representation of sensor unit activation
    - Activated by the set of photons
- Storing an image
  - Based on the set of photons ???

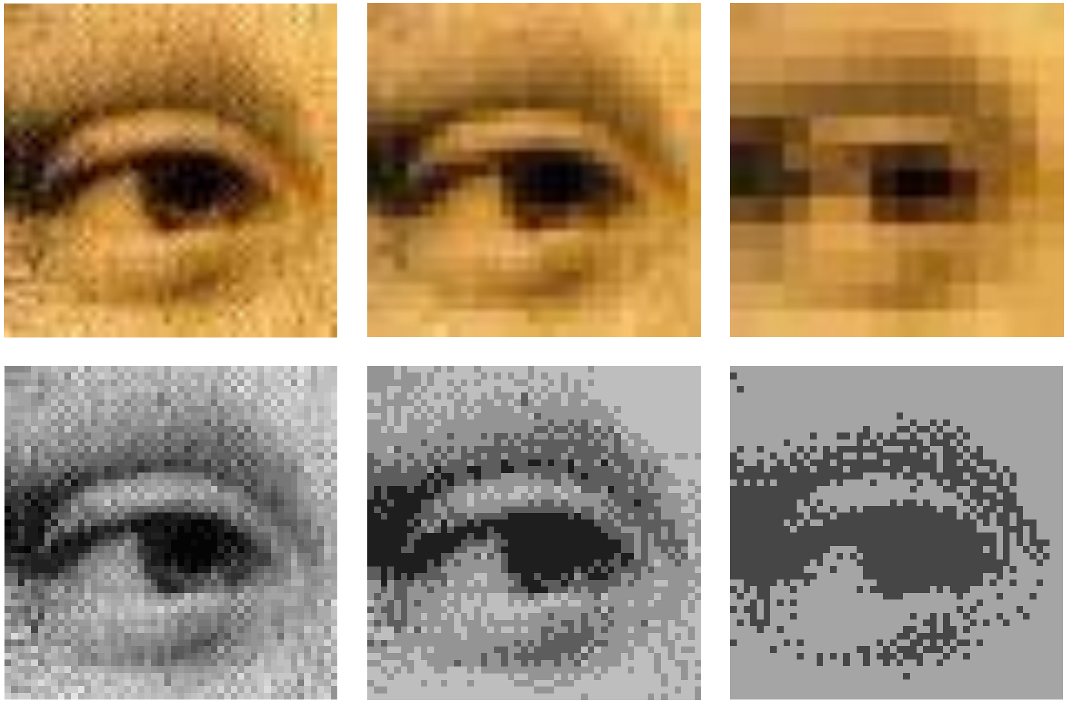
Sampling & Quantization

- Capturing continuous images on sensors
  - Sampling: Continuous to matrix
  - Quantization: Continuous color to value



Sampling & Quantization

- Size of a captured image:
  - # of samples (width*heigth) * # of colors

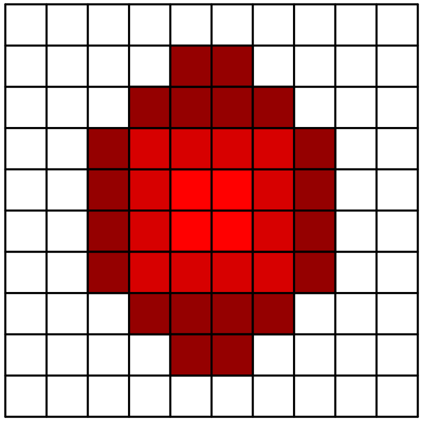


Content Based Image Retrieval

- Motivation & Semantic Gap
- Perception
- Color Based Features
- Texture Based Features
- Exercise 5

Color Histogram

- Count how often which color is used
- Algorithm:
  - Allocate int array h with dim = # of colors
  - Visit next pixel -> it has color with index i
  - Increment h[i]
  - IF pixels left THEN goto line 2
- Example: 4 colors, 10*10 pixels
  - histogram: [4, 12, 20, 64]



Color Histogram

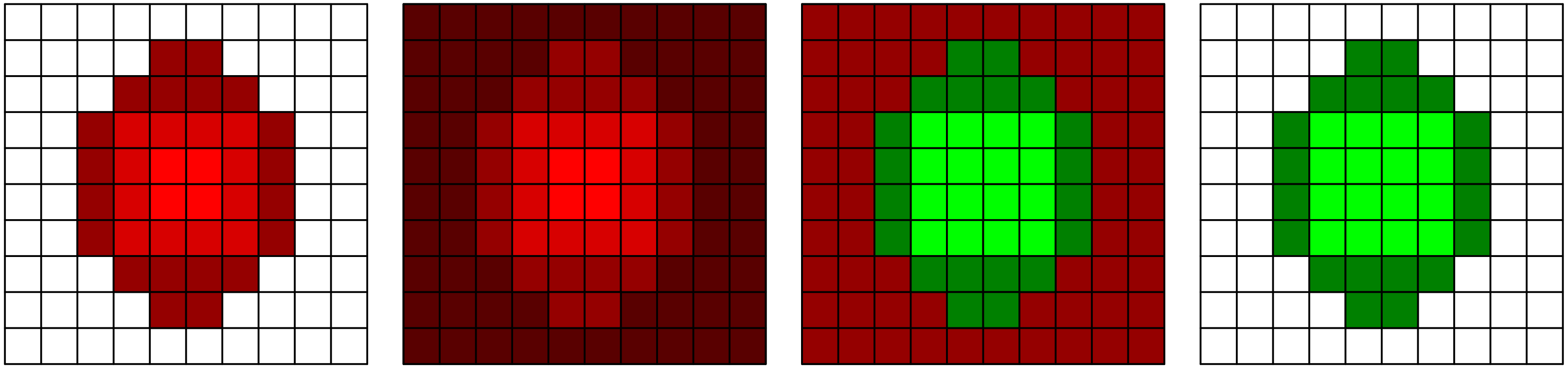
- Strategies:
  - Quantize if too many colors
  - Normalize histogram (different image sizes)
  - Weight colors according to use case
  - Use (part of) color space according to domain
- Distance / Similarity
  - Assumption: All images have the same colors
  - L_1 or L_2 is quite common

Color Histogram

- Benefits
  - Easy to compute, not depending on pixel order
  - Matches human perception quite good
  - Quantization allows to scale size of histogram
  - Invariant rotation & reflection
- Disadvantages
  - Distribution of colors not taken into account
  - Colors might not represent semantics
  - Find quantization fitting to domain / perception
  - Image scaling might be a problem

Color Histogram

- Example: 4 images, 7 colors
  - 1: [0, 4, 12, 20, 64, 0, 0]
  - 2: [66, 4, 12, 20, 0, 0, 0]
  - 3: [0, 0, 0, 64, 0, 20, 16]
  - 4: [0, 0, 0, 0, 64, 20, 16]



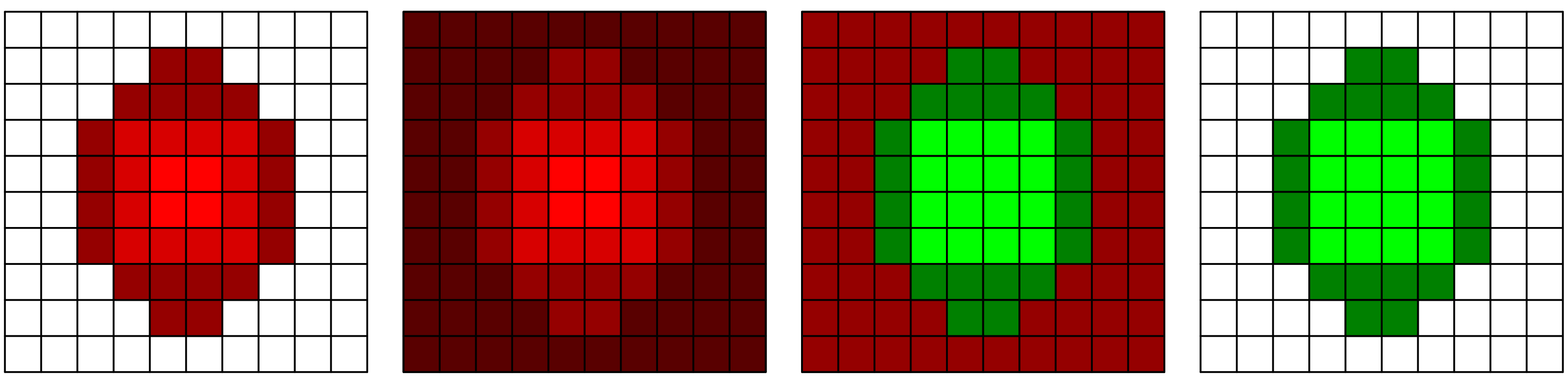
Color Histogram

Bilder:

- 1: [0, 4, 12, 20, 64, 0, 0]
- 2: [66, 4, 12, 20, 0, 0, 0]
- 3: [0, 0, 0, 64, 0, 20, 16]
- 4: [0, 0, 0, 0, 64, 20, 16]

Distanzfunktion d: L1

- d(1,2) = 130
- d(1,3) = 160
- d(1,4) = 52



Dominant Color

- Reduce histogram to dominant colors
  - e.g. for 64 colors c0-c63:
    - image 1: c12 -> 23%, c33 -> 6%, c2 -> 2%
    - image 2: c11 -> 43%, c2 -> 12%, c54 -> 10%
- Distance function in 2 aspects:
  - Difference in amount (percentage)
  - Difference between colors (c11 vs. c12)
- Further aspects:
  - Diversity and distribution

Dominant Color

- Benefits:
  - Small feature vectors
  - Easily understandable & intuitive
  - Similarity of color pairs (light vs. dark red, etc.)
  - Invariant to rotation & reflection
- Disadvantages
  - Similarity of color pairs no trivial problem
  - Dominant colors might not represent semantics

Color Distribution

- Index dominant color in image segment
  - e.g. 8*8 = 64 image segments
  - feature vector has 64 dimensions
    - One for each segment
  - color index is the entry on segment dimension
    - e.g. 16 colors [2, 0, 3, 3, 8, 4, ...]

Color Distribution

- Similarity
  - L_1 or L_2 are commonly used
- Benefits
  - Works fine for many scenarios
    - clouds in the sky, portrait photos, etc.
  - Mostly invariant to scaling
- Disadvantages
  - Colors might not represent semantics
  - Find quantization fitting to domain / perception
  - Rotation & reflection are a problem

Color Correlogram

- Histogram on
  - how often **specific colors** occur
  - in the **neighbourhood** of each other
- Histogram size is (# of colors)^2
  - For each color an array of neighboring colors

Color Correlogram

- Extraction algorithm
  - Allocate array h[#colors][#colors] all zero
  - Visit next pixel *p*
  - For each pixel *q* in neighborhood of *p*:
    - increment h[color(p)][color(q)]
  - IF pixels left THEN goto line 2
- Algorithm is rather slow
  - Depends on size of neighborhood
  - Typically determined by city block (L1) distance

Color Correlogram

- Similarity
  - L_1 or L_2 are commonly used
- Benefits
  - Integrates color as well as distribution
  - Works fine for many scenarios
  - Mostly invariant to rotation & reflection
- Disadvantages
  - Find appropriate neighborhood size
  - Find quantization fitting to domain / perception
  - Rather slow indexing / extraction

Color Correlogram

- Auto Color Correlogram
  - Just indexing how often *color(p)* occurs in neighborhood of pixel *p*
  - Simplifies the histogram to size # of colors

Color Correlogram

- Integrating different pixel features to correlate
  - **Gradient Magnitude** (intensity of change in the direction of maximum change)
  - **Rank** (intensity variation within a neighborhood of a pixel)
  - **Texturedness** (number of pixels exceeding a certain level in a neighborhood)

Content Based Image Retrieval

- Motivation & Semantic Gap
- Perception
- Color Based Features
- Texture & Shape Features
- Exercise 5

Texture & Shape Features

- Indexing non color features in image
  - Outlines, edges of regions
  - Overall characteristics like coarseness and regularity

Spatial Filtering

- Methods for *enhancing* the image
- Normally a kernel or filter is used:
  - A matrix which is applied to the image
  - In a linear transformation

Spatial Filtering

194	128	102	197	69
162	68	103	144	115
121	85	57	27	14
24	183	192	239	150
92	93	154	138	170

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

194	128	102	197	69
162	68	103	144	115
121	85	122	27	14
24	183	192	239	150
92	93	154	138	170

Spatial Filtering

- This is a simple smoothing kernel
- Other operations
  - Sharpen
  - Unsharp Mask
  - Gradient

Edge Detection

- Based on the gradient
  - Denotes the amount of change at specific point
  - Can be estimated with Gradient $\nabla f \approx |G_x| + |G_y|$
- Different kernels based on estimation
  - Test e.g. with Gimp

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Edge & Texture Features

- Edge based on gradient map
  - Aims to represent edges in number vector
  - e.g. Edge Histogram (MPEG-7)
  - Problems with rotation & reflection
- Texture features
  - Statistics on the image representing
  - Heavily depends on domain
  - Mostly invariant to rotation & reflection
  - Problems with scaling

Tamura Features

Tamura & Mori (1978)

- Widely used in CBIR
  - E.g. IBM QBIC
- 6 texture features
  - Coarseness, contrast, directionality
  - Line-likeness, regularity, and roughness
- Good overview is provided in:
  - Thomas Deselaers, “Features for Image Retrieval”, Thesis, RWTH Aachen, Dec. 2003

Tamura Features

Tamura & Mori (1978)

- Coarseness
  - Pixel diversity in neighborhoods
- Contrast
  - Using mean and variance of an image
- Directionality
  - Horizontal and vertical derivatives (like Sobel)

Shapes

- Indexing of
  - **Boundaries** and **Regions**
  - Invariant to scaling, rotation and translation
- Features depend on domain, e.g.
  - Length of outline w.r.t. the image
  - Convexity & Concavity
  - Holes & Connectivity

Joint Histograms

- Determine color & texture properties per pixel
  - E.g. color (64 bins) + gradient (4 bins)
- Use a color histogram per texture bin
  - E.g. 4 times a color histogram
- Classify pixel based on texture
- Add value to the corresponding color histogram.

Joint Histograms

- Strategies
  - Find good texture properties
  - Minimize overall number of bins
- Benefits
  - Works better than pure color histograms
- Disadvantages
  - Typically slower extraction
  - Higher number of bins

Content Based Image Retrieval

- Motivation & Semantic Gap
- Perception
- Color Based Features
- Texture Based Features
- Exercise 5

Exercise 05

- Rank all 7 images according to your subjective quality rating.
- Assume image 1 is the query and all other images are results
  - Rank the 6 result images according to their similarity to the query image
  - Use your own subjective similarity rating.






Exercise 05

- You will also find the images on the course homepage
- Send me an Excel / Calc / PDF until next lesson.

Thank you …

… for your attention