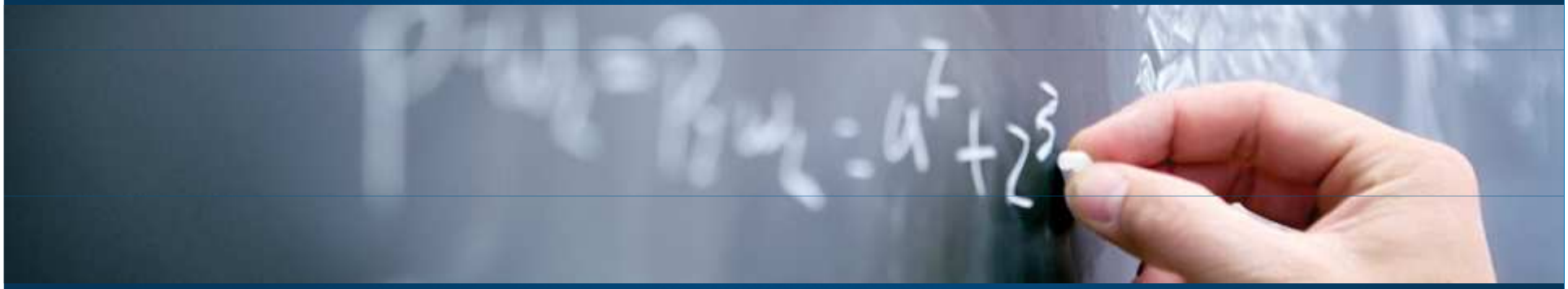
Content Based Image Retrieval

Natalia Vassilieva
nvassilieva@hp.com
HP Labs Russia

Tutorial outline

- Lecture 1
  - Introduction
  - Applications
- Lecture 2
  - Performance measurement
  - Visual perception
  - Color features
- Lecture 3
  - Texture features
  - Shape features
  - Fusion methods
- Lecture 4
  - Segmentation
  - Local descriptors
- Lecture 5
  - Multidimensional indexing
  - Survey of existing systems

Lecture 1
Introduction to Image Retrieval Applications

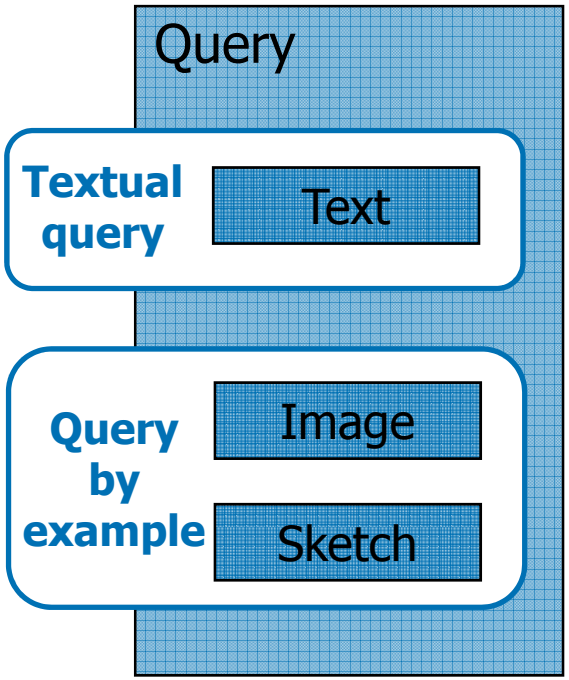


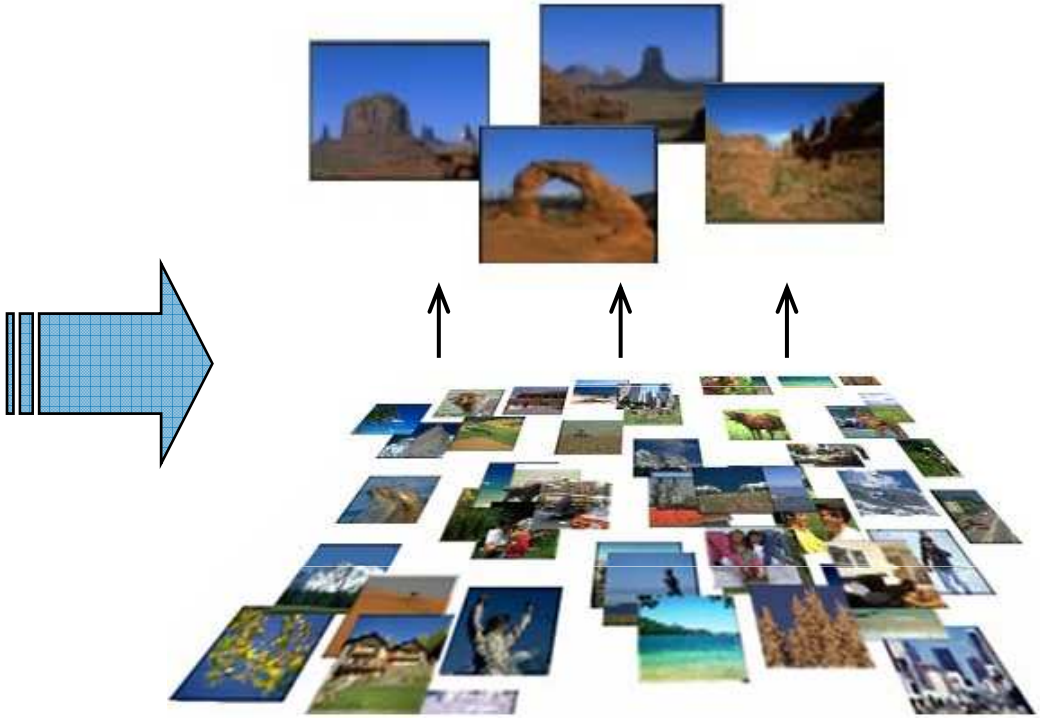
Lecture 1: Outline

- What is and Why image retrieval?
- How to compare and retrieve images?
  - Digital image representation
  - Common components of the CBIR systems
  - Main problems and research directions
- What are applications?

What is image retrieval?

- Description Based Image Retrieval (DBIR)
- Content Based Image Retrieval (CBIR)





DBIR v. s. CBIR

	DBIR	CBIR
+	▪ Fulltext search algorithms are applicable ▪ Search results corresponds to image semantics	▪ Automatic index construction ▪ Index is objective
–	▪ Manual annotating is hardly feasible ▪ Manual annotations are subjective	▪ Semantic gap ▪ Querying by example is not convenient for a user

Levels of image retrieval

- Level 1: Based on color, texture, shape features
  - Images are compared based on low-level features, no semantics involved
  - A lot of research done, is a feasible task
- Level 2: Bring semantic meanings into the search
  - E. g. identifying human beings, horses, trees, beaches
  - Requires retrieval techniques of level 1
  - Very active and challengeable research area
- Level 3: Retrieval with abstract and subjective attributes
  - Find pictures of a particular birthday celebration
  - Find a picture of a happy beautiful woman
  - Requires retrieval techniques of level 2 and very complex logic
  - Is far from being developed with modern technology available now

Why image retrieval?

- Huge amounts of images are everywhere: how to manage this data?
- “A Picture is worth thousand words”
- Not everything can be described in text
- Not everything is described in text

Why content based image retrieval?

- Automatic generation of textual annotations for a wide spectrum of images is not feasible.
- Annotating images manually is a cumbersome and expensive task for large image databases.
- Manual annotations are often subjective, context-sensitive and incomplete.
- Google, Yandex and others use text-based search.
  Results are not perfect.
  However, now it is much better, than a couple of years ago!

Image retrieval by Google

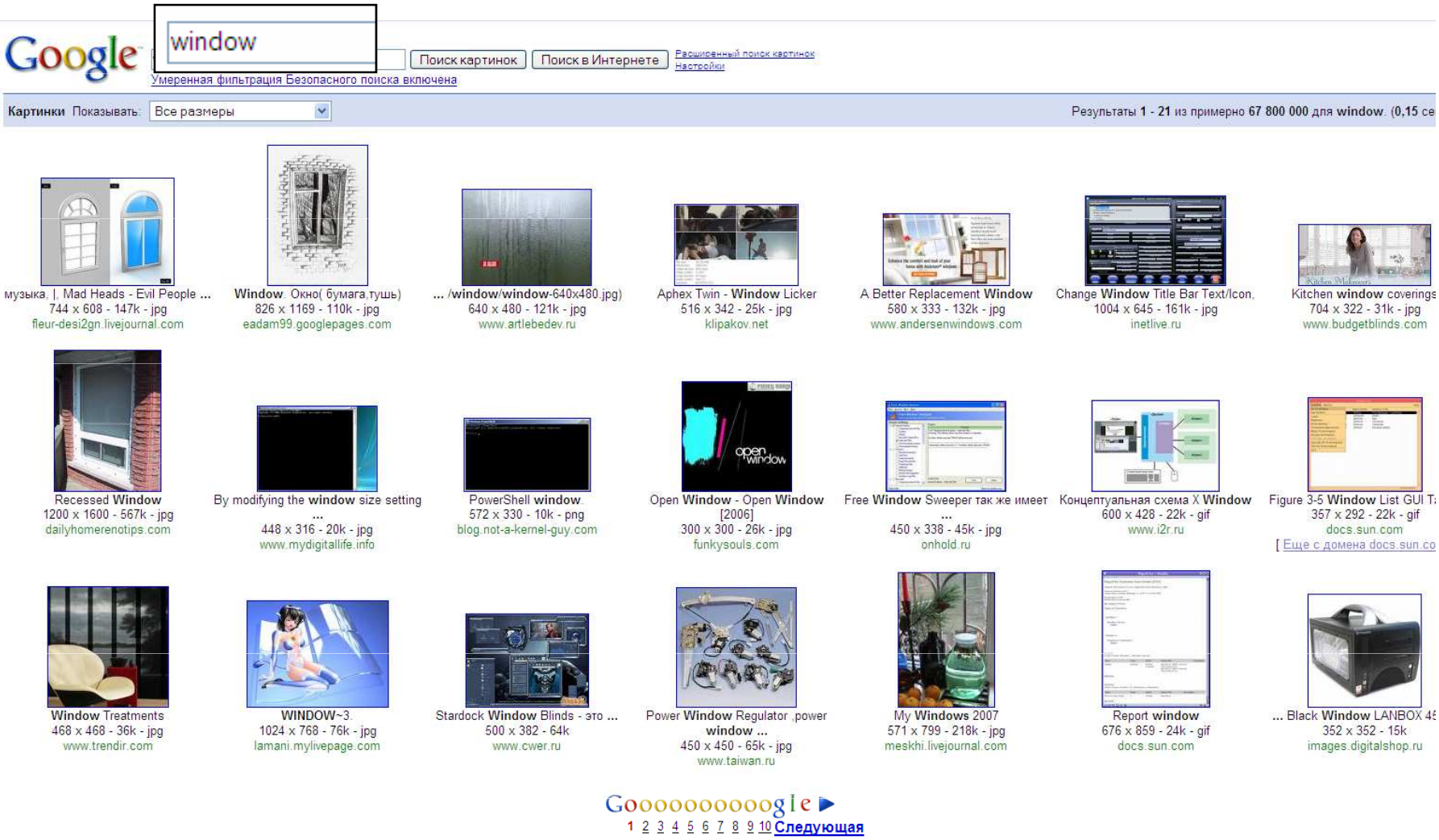
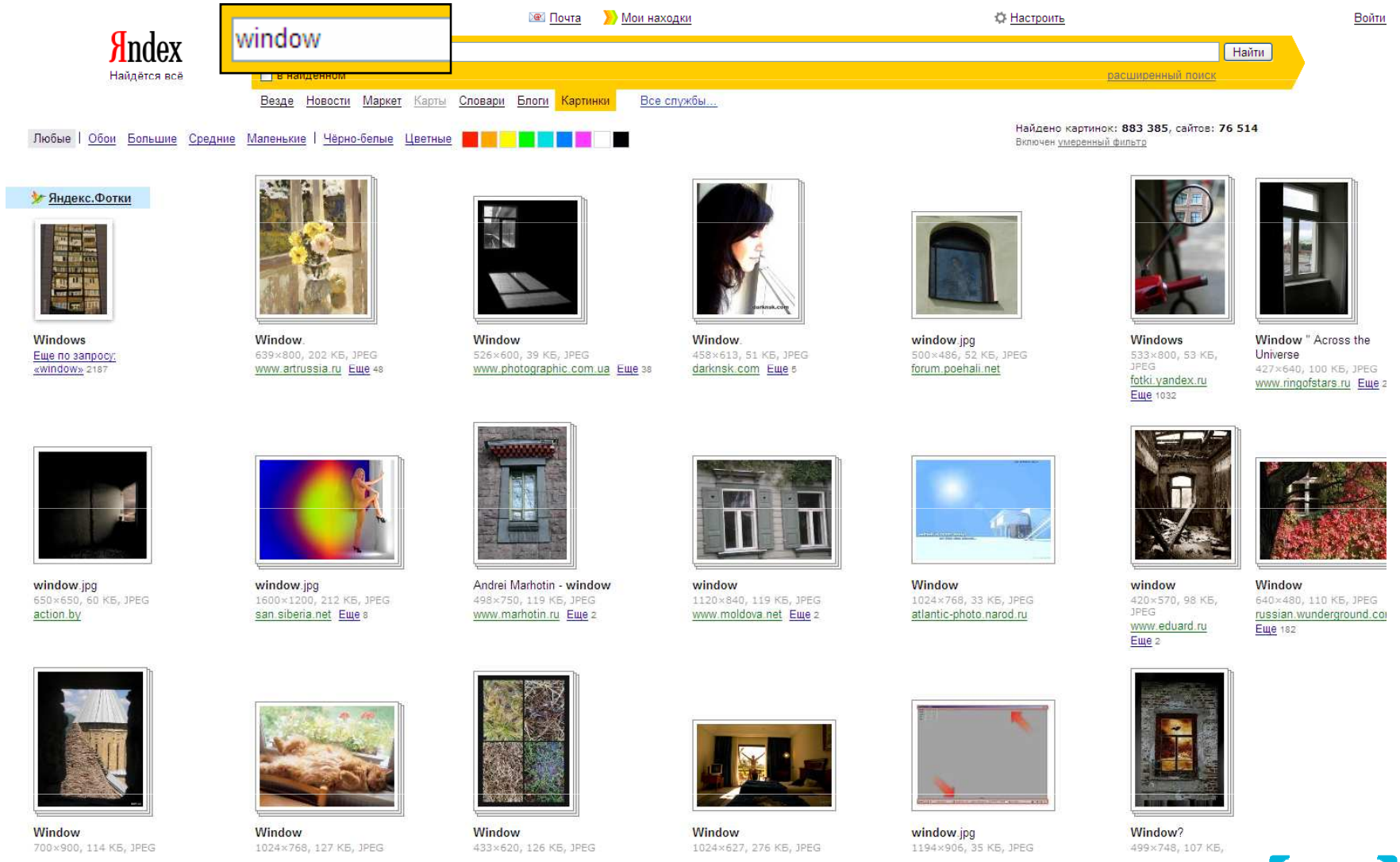


Image retrieval by Yandex

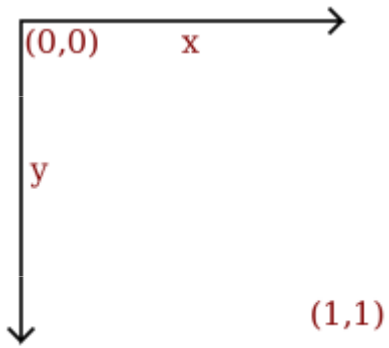


Lecture 1: Outline

- What is and Why image retrieval?
- How to compare and retrieve images?
  - Digital image representation
  - Common components of the CBIR systems
  - Main problems and research directions
- What are applications?

Digital image representation

Vector image

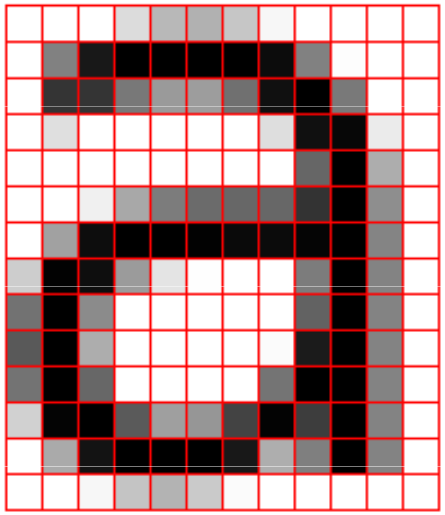



```
draw circle
        center 0.5, 0.5
        radius 0.4
        fill-color yellow
        stroke-color black
        stroke-width 0.05
draw circle
        center 0.35, 0.4
        radius 0.05
        fill-color black
draw circle
        center 0.65, 0.4
        radius 0.05
        fill-color black
draw line
        start 0.3, 0.6
        end 0.7, 0.6
        stroke-color black
        stroke-width 0.1
```

Digital image representation

Bitmap (raster) image

a



1.0	1.0	1.0	0.9	0.6	0.6	0.6	1.0	1.0	1.0	1.0	1.0
1.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.0	1.0	1.0
1.0	0.2	0.2	0.5	0.6	0.6	0.5	0.0	0.0	0.5	1.0	1.0
1.0	0.9	1.0	1.0	1.0	1.0	1.0	0.9	0.0	0.0	0.9	1.0
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.5	0.0	0.5	1.0
1.0	1.0	1.0	0.5	0.5	0.5	0.5	0.5	0.4	0.0	0.5	1.0
1.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.0
0.9	0.0	0.0	0.6	1.0	1.0	1.0	1.0	0.5	0.0	0.5	1.0
0.5	0.0	0.6	1.0	1.0	1.0	1.0	1.0	0.5	0.0	0.5	1.0
0.5	0.0	0.7	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.5	1.0
0.6	0.0	0.6	1.0	1.0	1.0	1.0	0.5	0.0	0.0	0.5	1.0
0.9	0.1	0.0	0.6	0.7	0.7	0.5	0.0	0.5	0.0	0.5	1.0
1.0	0.7	0.1	0.0	0.0	0.0	0.1	0.9	0.8	0.0	0.5	1.0
1.0	1.0	1.0	0.8	0.8	0.9	1.0	1.0	1.0	1.0	1.0	1.0

$$f(x,y)=\begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}.$$

$0 \le f(x,y) \le L$, and typically $L = 255$

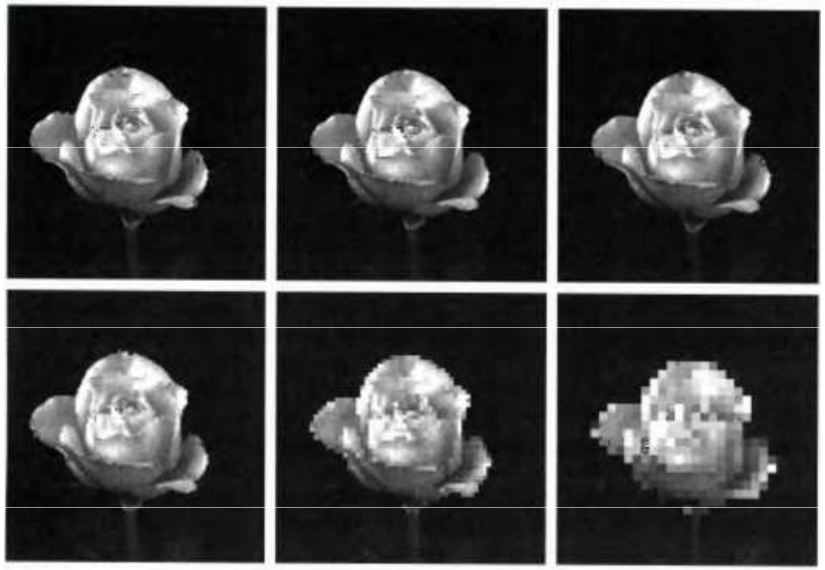
- Bitmap image is an array of pixels
- The value of each array element corresponds to the color of the appropriate pixel

Digital image representation

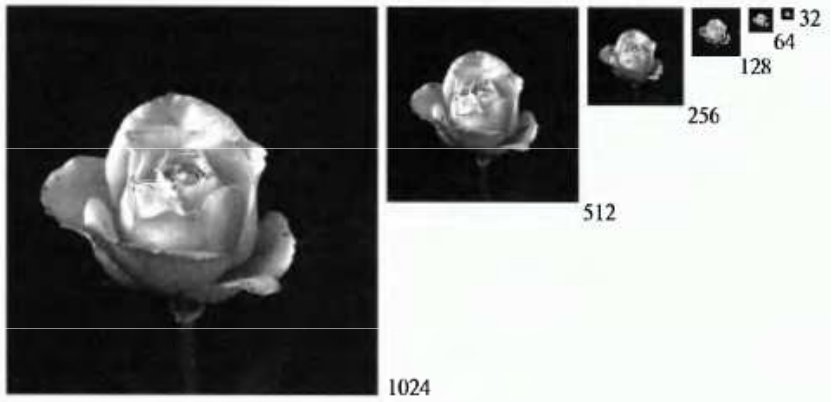
Bitmap (raster) image

Important parameters of raster image:

- Raster dimensions
- Resolution (ppi)
- Sample depth (usually 2^k)



Fixed dimensions, varying resolution

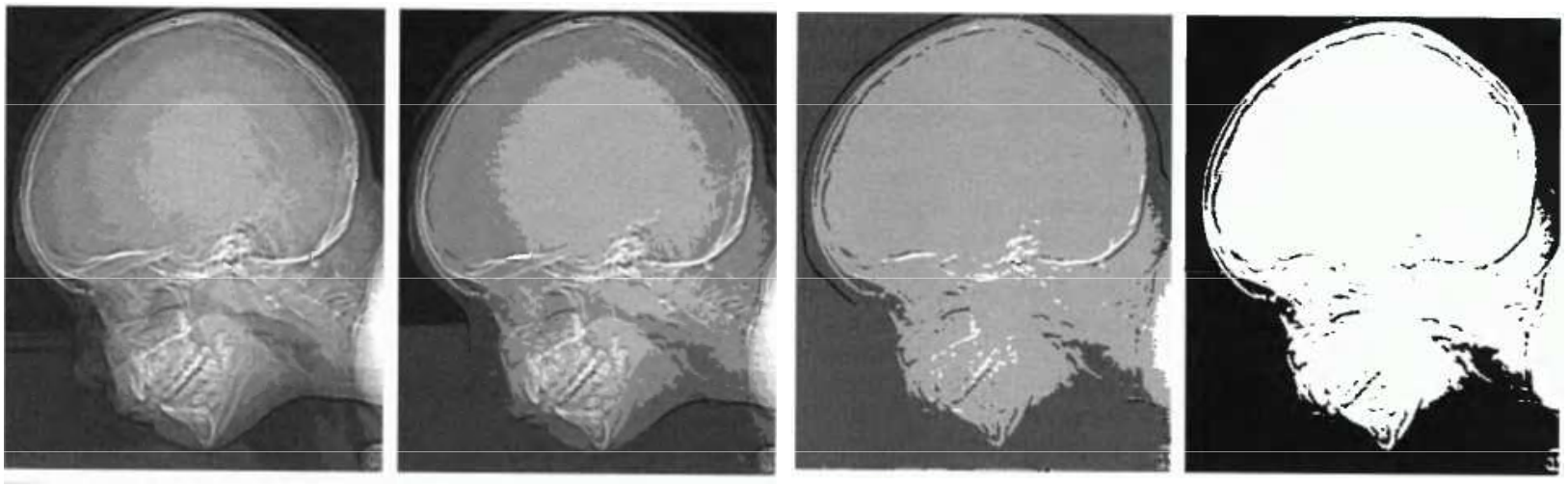


Fixed resolution, varying dimension

Digital image representation

Bitmap (raster) image

The same image with varying sample depths:



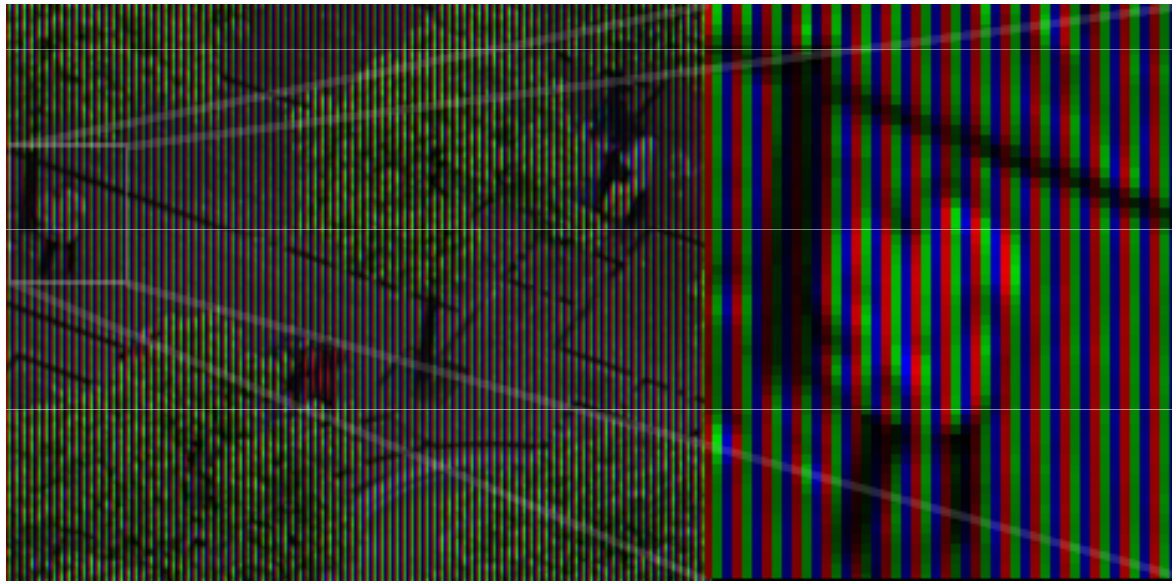
16 levels | 8 levels | 4 levels | 2 levels

Typical levels: 8 bit (256 levels), 16 bit – png, tiff

Digital image representation

Bitmap (raster) image: color

- RGB – the most common color model (CRT monitors, LCD screens/projectors)
- Each pixel represented by 3 values: red, green, blue



RGB bands:
color image built up of bands of red, green and blue color

Digital image representation

Bitmap (raster) image: color

- Pixel-interleaved format (chunky) – is a common one

$$f(x,\ y) = \begin{pmatrix} r_{0,0},\ g_{0,0},\ b_{0,0} & r_{0,1},\ g_{0,1},\ b_{0,1} & \cdot & r_{0,N-1},\ g_{0,N-1},\ b_{0,N-1} \\ r_{1,0},\ g_{1,0},\ b_{1,0} & r_{1,1},\ g_{1,1},\ b_{1,1} & \cdot & r_{1,N-1},\ g_{1,N-1},\ b_{1,N-1} \\ \cdot & \cdot & \cdot & \cdot \\ r_{M-1,0},\ g_{M-1,0},\ b_{M-1,0} & r_{M-1,1},\ g_{M-1,1},\ b_{M-1,1} & \cdot & r_{M-1,N-1},\ g_{M-1,N-1},\ b_{M-1,N-1} \end{pmatrix}$$

- Color-interleaved format (planar)

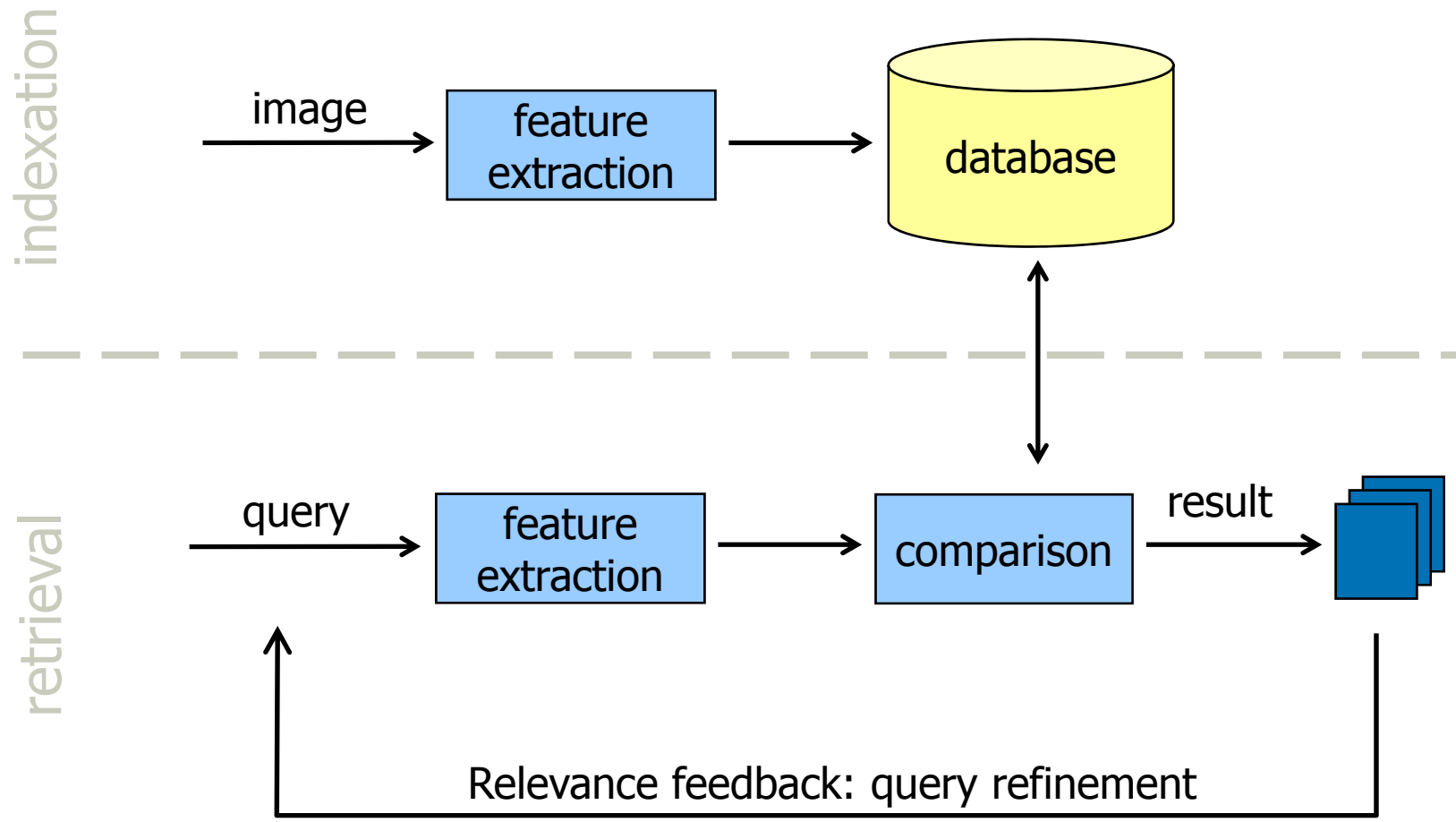
$$f(x,\ y) = \left(\begin{pmatrix} r_{0,0} & r_{0,1} & \cdot & r_{0,N-1} \\ r_{0,1} & r_{1,1} & \cdot & r_{1,N-1} \\ \cdot & \cdot & \cdot & \cdot \\ r_{M-1,0} & r_{M-1,1} & \cdot & r_{M-1,N-1} \end{pmatrix}, \begin{pmatrix} g_{0,0} & g_{0,1} & \cdot & g_{0,N-1} \\ g_{0,1} & g_{1,1} & \cdot & g_{1,N-1} \\ \cdot & \cdot & \cdot & \cdot \\ g_{M-1,0} & g_{M-1,1} & \cdot & g_{M-1,N-1} \end{pmatrix}, \begin{pmatrix} b_{0,0} & b_{0,1} & \cdot & b_{0,N-1} \\ b_{0,1} & b_{1,1} & \cdot & b_{1,N-1} \\ \cdot & \cdot & \cdot & \cdot \\ b_{M-1,0} & b_{M-1,1} & \cdot & b_{M-1,N-1} \end{pmatrix} \right)$$

LABS hp

Lecture 1: Outline

- What is and Why image retrieval?
- How to compare and retrieve images?
  - Digital image representation
  - Common components of the CBIR systems
  - Main problems and research directions
- What are applications?

Common components of CBIR system



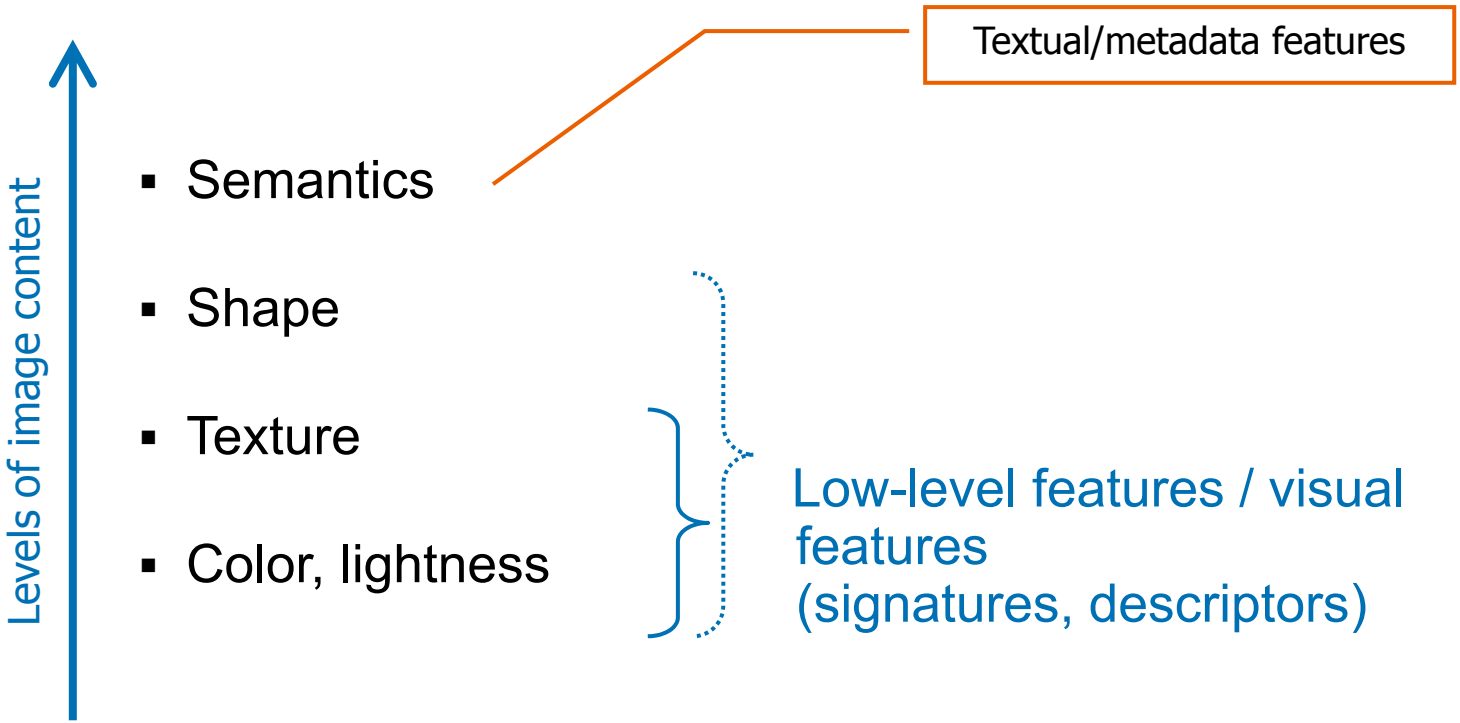
Lecture 1: Outline

- What is and Why image retrieval?
- How to compare and retrieve images?
  - Digital image representation
  - Common components of the CBIR systems
  - Main problems and research directions
- What are applications?

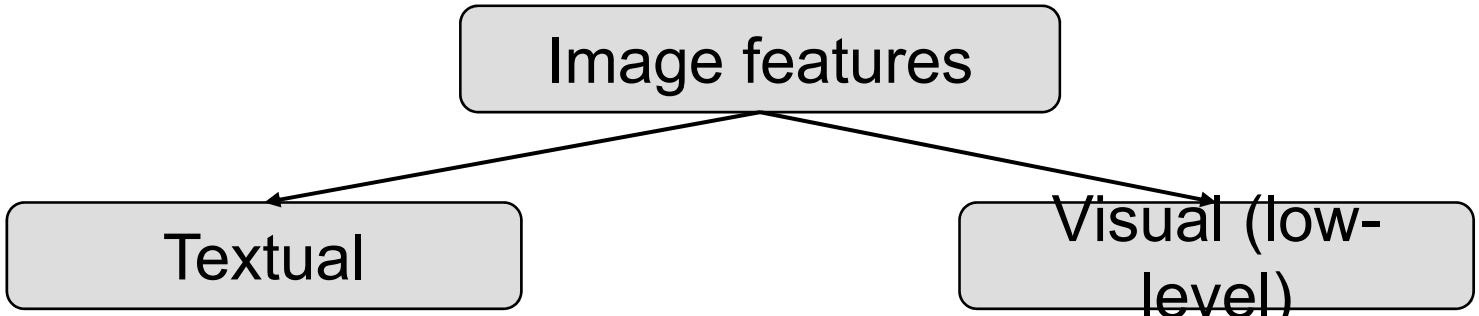
Problems and directions

- Low-level feature extraction
  - How to represent an image in a compact and descriptive way?
  - How to compare features, and, thus, images?
- High dimensional indexing
  - How to index huge amounts of high dimensional data?
- Visual interface for image browsing
  - How to visualize the results?

How to: Image features



How to: Image features



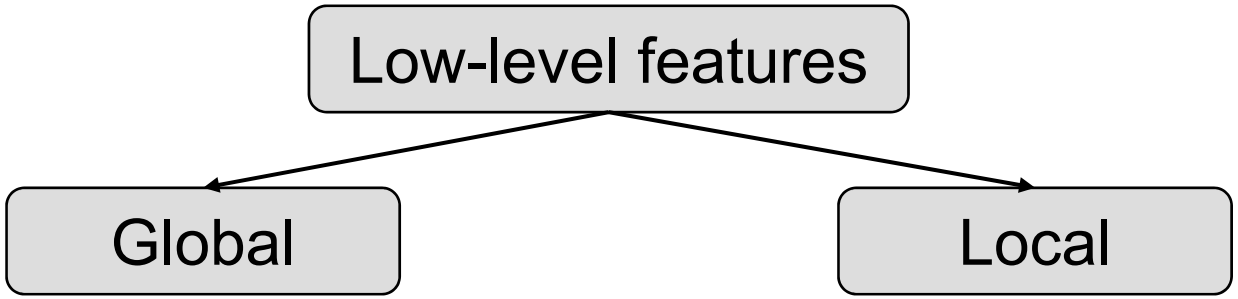
Annotations and metadata:

- tags/keywords;
- creation date;
- geo tags;
- name of the file;
- photography conditions (exposition, aperture, flash…).

Features extracted from pixel values:

- color descriptors;
- texture descriptors;
- shape descriptors;
- spatial layout descriptors.

How to: Image features



Describes the whole image:

- average intensity;
- average amount of red;
- …

All pixels of the image are processed.

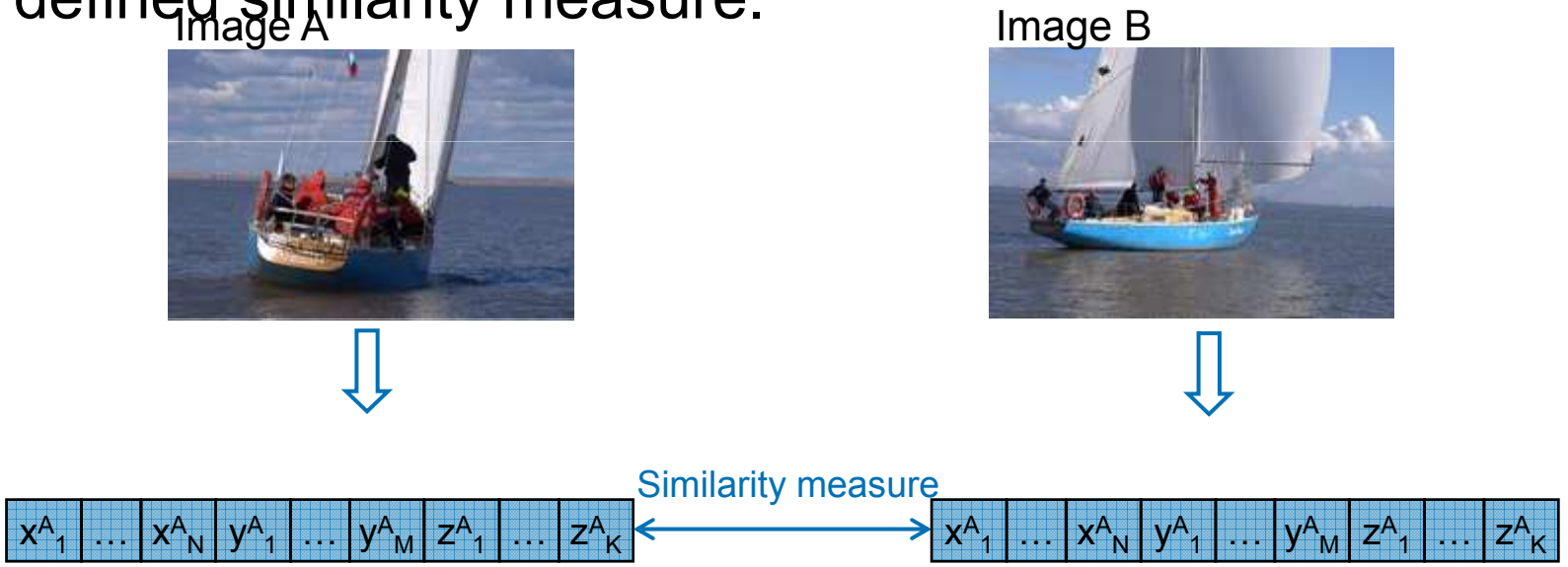
Describes one part of the image:

- average intensity for the left upper part;
- average amount of red in the center of the image;
- …

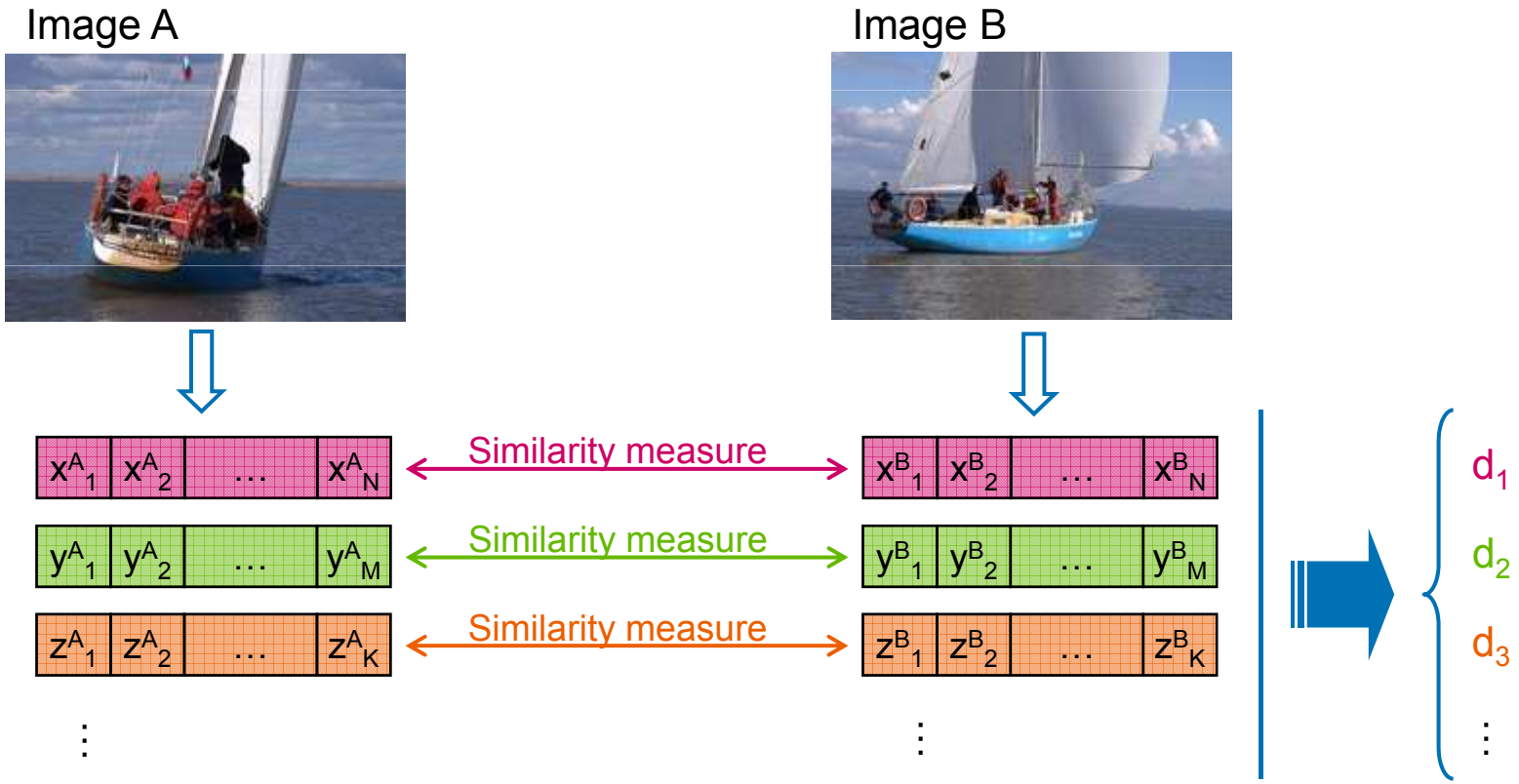
Segmentation of the image is performed, pixels of a particular segment are processed to extract features.

How to: Feature spaces

- Feature vector – a vector of features, representing one image.
- Feature space – the set of all possible feature vectors with defined similarity measure.



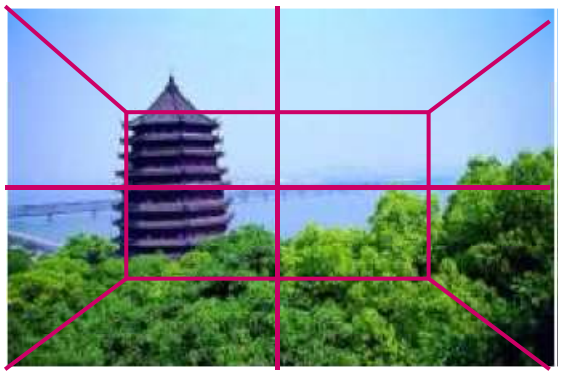
How to: Combine results

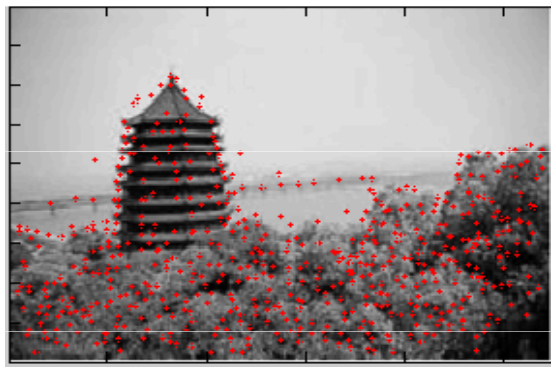


$$D = \sum_i c_i d_i$$

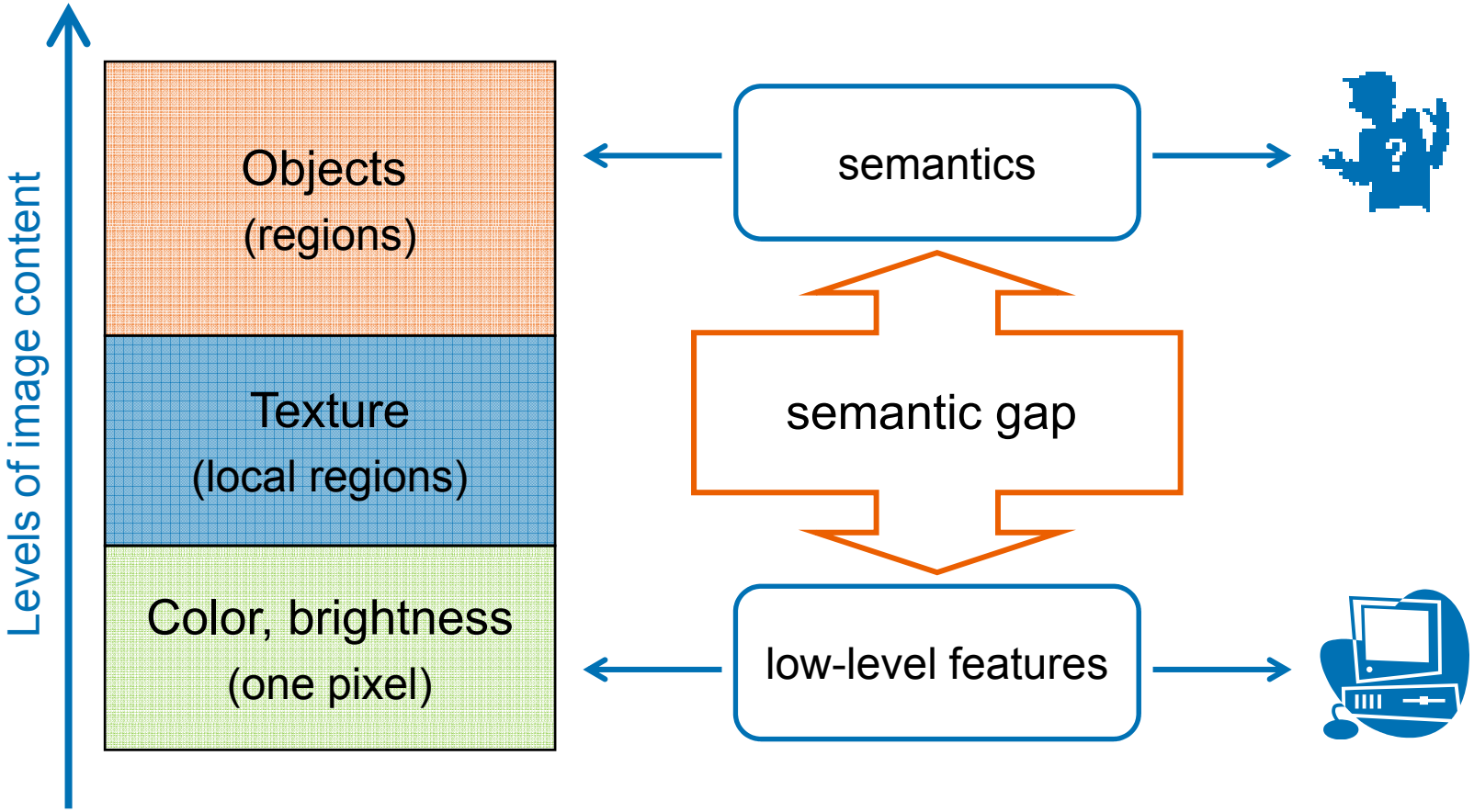
How to: Image segmentation

- Fixed regions
  - The same region boundaries for all images.
- Segmentation
  - Boundaries depends on image content.
- Key points (point of interest) detection
  - Points of particular interest in the image, feature extraction for areas around key points.





Problems: semantic gap



How to understand what's on the images?

Problems: what's on the images?

- Sometimes it is not easy to understand the image even for humans!
- What do we want from machines?

Problems: what's on the images?


















- How do we now that all these objects are lamps?

Problems: subjectivity of perception

Let’s compare our perception!

- Copy test application and test images from CD or from common share \\lampai.tsure.ru\RUSSIR\CBIR
- Evaluate the results of CBIR systems
- Give me your results on Thursday, Sep 4
- I’ll share the statistics calculated based on your results on Friday, Sep 5

Problems: high dimensional data

- More information in feature vectors – better search results.
- Local features are usually more precise than global -> more feature vectors.
- The dimensionality of the feature vectors is normally of the order 10^2.
- ~200-500 keypoints per image
- Non-Euclidean similarity measure

How to: high dimensional indexing

- Perform dimension reduction
  - The dimension of the feature vectors is normally very high, the embedded dimension is much lower.
- Use appropriate multi-dimensional indexing techniques, which are capable of supporting Non-Euclidean similarity measures
  - Trees (k-d tree, VP-tree and others)
  - Hashing

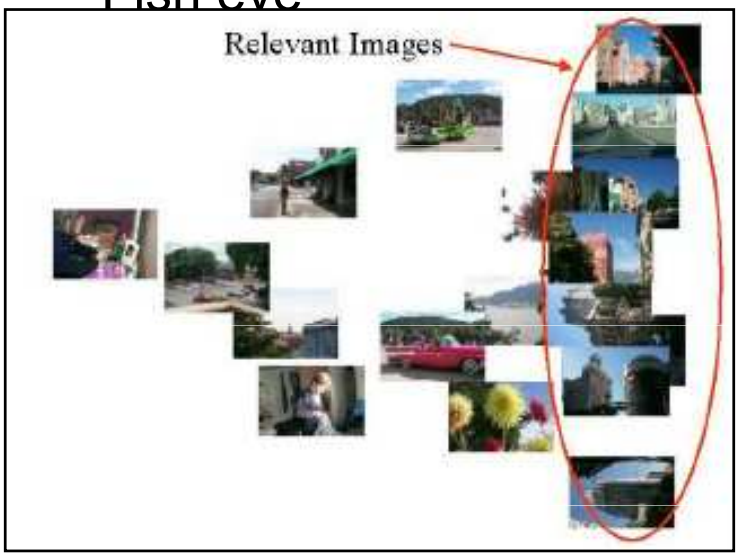
Problems: visualization

- Image content is very rich and its interpretation is very contextual and subjective.
- Many independent similarity measures are commonly used. How about to let user influence the choice of these parameters?
- Which images to show as a result (result diversity)?
- Interactive search and relevance feedback.

How to: visual interfaces

- 1-D visualizations
  - As a list (standard way)
- 2-D visualizations
  - Based on dimension reduction techniques
- “3-D” visualizations
  - Fish eye

1-D



2-D

3-D

Neighbour research areas

- Image processing
  - Features extraction
  - Pattern recognition and machine learning
    - Faces, handwritings, thumbprints, …
    - Classification tools
  - Image enhancement
  - Image classification
    - The same features are used
    - Classification helps to retrieve
- Information retrieval
  - Scalability
  - Performance measurement
  - Fusion of multiple evidences

Lecture 1: Outline

- What is and Why image retrieval?
- How to compare and retrieve images?
  - Digital image representation
  - Common components of the CBIR systems
  - Main problems and research directions
- What are applications?

What are applications? – Image Archives.

- Manage image archives
  - Personal photo collections (many thousands of photos in mine)
  - Professional photograph archives (millions of photos)
  - Art collections (millions of photos)
- Browse images
- Organize image collection: delete duplicates, classify images, select “the best” from the group of similar images
- Posters creation, auto cropping, album creation (www.snapfishlab.hpl.hp.com)
- Better organization of search-by-text results

LABS hp

What are applications? – Image Archives.

- Manage image archives





  - …
  - Search for particular image (by its smaller version, by its fragment)
  - Search for similar images (landscape paintings, sea views, paintings by the same author)
  - Search for a painting with particular colors (“I want a sea view painting to my bedroom with an orange carpet and yellow walls”)
  - Search for group photos of my family
  - Search for an image that will be a good illustration to my article/presentation
  - … a lot of other use cases

What are applications? – Copyrights.

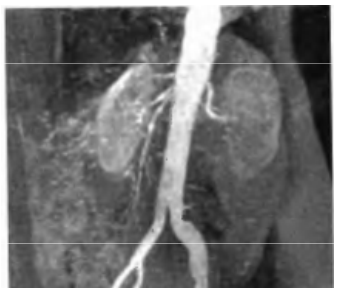
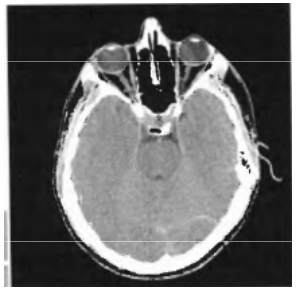
- Trademark and copyright application
  - World Wide Web
  - Enterprise network

- Copyright detection without watermarking and protect intellectual property
- Forged images detection and sub-image retrieval
- Trademark image registration: a new candidate is compared with existing marks to ensure no risk of confusing property ownership
- Search if confidential images are included into public presentations

What are applications? – Medical.

- Medical diagnosis
  - Collection of X-ray images



  - Search for similar past cases
  - Is it similar to the “healthy” case?
  - Classification of X-ray images

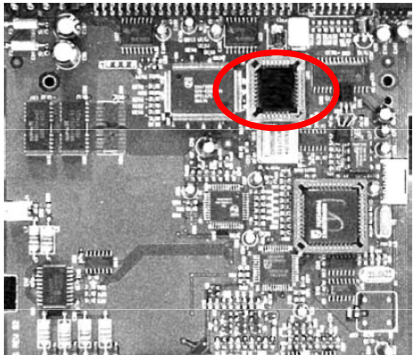
What are applications? – Security.

- Security issues
  - Video surveillance material
  - Faces, fingerprints, retina images

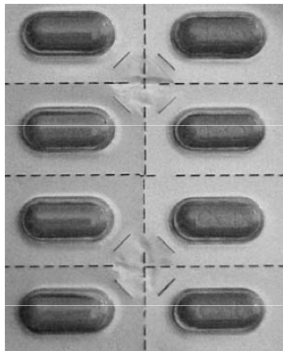
- Detect suspicious objects during the video surveillance
- Detect “wanted” faces during the video surveillance
- Grant or deny access based on fingerprints/retina scanning

What are applications? – In industry.

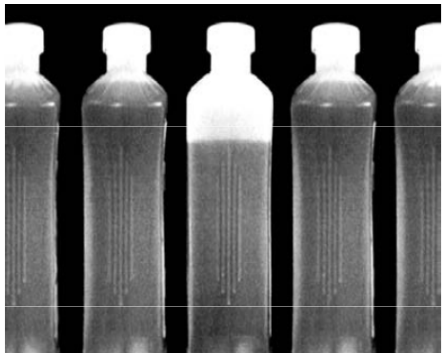
- Quality assurance



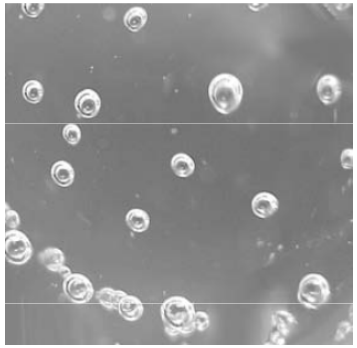
(a) CD-ROM controller



(b) Pack of pills

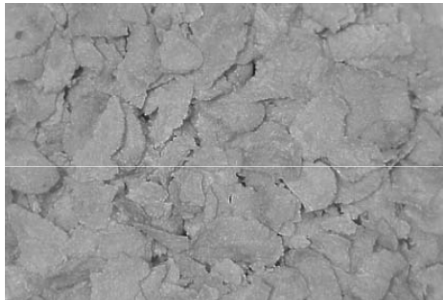


(c) Level of liquid



(d) Air-bladders in plastic

- Control that all parts of the product are on place (a)
- Control if all places in pill pack are filled (b)
- Control the level of liquid in bottles (c)
- Control the quality of plastic details (d)
- And even control the corn flakes! (e)



(e) Corn flakes

What are applications? – Others.

- Military-related issues
  - Auto aiming, tracking systems
- Image-based modeling and 3-D reconstruction
  - Medical imaging
  - Indoor scene reconstruction from multiple images
  - Outdoor scene reconstruction from aerial photography
- Geographical information and remote sensing
  - Process satellite data: climate variability, sea surface temperatures, storms watch.

Lecture 1: Resume

- CBIR is an actual problem and an active research area
- Main research directions are:
  - Feature extraction
  - Multidimensional indexing
  - Visualization
- CBIR combines research results of image processing, information retrieval, database communities
- CBIR has many applications in various areas

Lecture 1: Bibliography

- Gonzalez R, Woods R. Digital Image Processing, published by Pearson Education, Inc, 2002.
- Rui Y., Huang T.S., Chang S.-F. Image Retrieval: Past, Present and Future. In Proc. of Int. Symposium on Multimedia Information Processing, Dec. 1997.