2. Summary

- Last week:
  - What are multimedia databases?
    - Multimedia, Medium
  - Multimedia database applications
    - Image, Audio, Video, Hybrid queries
  - Evaluation of retrieval techniques
    - Relevance = Precision + Recall

- Both measures only make sense, if considered at the same time
  - E.g., get perfect precision by returning just one obvious correct document, but the recall is then low (assuming more than one document are relevant)
  - Or, get perfect recall by simply returning all documents, but then the precision is extremely low…

2. Summary

- Prepare a test set: all documents are tagged by experts with regard to a certain query
  - Precision: correctly returned documents relative to all returned documents
    - \( P = \frac{ca}{ca + fa} \)
  - Recall: correctly returned documents relative to all relevant documents
    - \( R = \frac{ca}{ca + fd} \)

- Both measures only make sense, if considered at the same time
  - E.g., get perfect precision by returning just one obvious correct document, but the recall is then low (assuming more than one document are relevant)
  - Or, get perfect recall by simply returning all documents, but then the precision is extremely low…

2. Summary

- Precision-Recall-Curves
  - Average precision of the system 3 at a recall level of 0.2
  - Which system is the best?
  - What is more important: recall or precision?

2 Color-based Retrieval

- 2.1 Basics of image retrieval
- 2.2 Introduction to color spaces
- 2.3 Extracting color features
- 2.4 Matching
2.1 Multimedia Data Retrieval

• Information retrieval (text)
  – Words carry semantic information
  – Texts with similar words are of similar content
• Now: complex multimedia objects
  – What carries semantic information?
  – How do we define (dis-) similarity?
  – Extraction of content-based characteristics!
    • E.g., identify/recognize persons and objects in images or videos

2.1 Textual Metadata

• Metadata describing content is difficult
  – Is really everything described so you can respond correctly to any request?
  – Accuracy and chosen words for the description?
    • Granularity of the description e.g., rodent vs. mouse, …!
  – Can images be described so easily?

2.1 Retrieval on Image, Video and Audio

• Textual metadata
  – Relational (author, size, …)
  – Content descriptive (picture of a white mouse)
• Advantages
  – Good quality
  – Uses existing procedures
• Disadvantages
  – Manual annotation is costly
  – Can everything be found? (e.g., scooter)

2.1 Textual Metadata

• Example: Description of a wallpaper pattern on the phone!

2.1 Multimedia Data Retrieval

• Essential Components
  – Text (→ full-text search, IR methods)
  – Image, video
  – Audio
• Retrieval of image, video, audio
  – Textual (descriptive) metadata
  – Content-based features
2.1 Close your eyes!

• Describe the wallpaper pattern!

\[ \begin{array}{c}
\text{Image of wallpaper pattern}
\end{array} \]

2.1 Used categories

• Color
  – Pink and white
  – Foreground pink, white background

• Shapes
  – Little flowers in different sizes
  – Petals on stems with 2 leaves

• Texture
  – High contrast
  – Spread evenly over the surface

• Typical examples of (low-level) features!

2.1 Similarity Search

• Features
  – Evaluate different (and not all) characteristics
  – Are often not comprehensible
  – Return more or less relevant results with respect to the query
  – But allow for some queries that would otherwise be very complicated

2.1 Remember?

\[ \begin{array}{c}
\text{Diagram of query flow}
\end{array} \]

2.1 Retrieval of Images

• Images are two-dimensional arrays
• Each tuple is a pixel characterized by ...
  – Coordinates
  – Color

\[ \begin{array}{c}
\text{Diagram of pixel array}
\end{array} \]

2.1 Description of Images

• Low-level Features:
  – Color
  – Texture
  – Shapes

• High-level features:
  – The whole image as the input signal
  – Fourier transformation
  – Wavelets
  – …
2.1 Example: Low-level Color Features

- **Assumption:** If two images share similar colors then also their content may be similar
- **Loss of information through low-level features**
- **Example:** red sunset (orange, yellow)

2.1 Significance

- Results are often quite good ...
  - A frog is not a sunset
- ... but not always
  - Also orange frogs are no sunsets

2.1 Differentiation

- Combination of several low-level features usually provides better differentiation
- **Semantics is not always obvious**
  - Sunrise ↔ sunset
  - Red ball on the beach

2.1 Example: High-level feature

- **Fourier transformation**
  - **Image as signal**
  - Transform from **position space** (normal visible image) to **frequency domain** (description of the image by overlapping 'intensity oscillations')
  - **No loss** of information
  - Difficult to interpret

2.1 Position Space and Frequency Domain

More about it, later...

2.2 Color Features

- **Today:** color features
  - Important element of human perception
  - Important for detection and differentiation of visual information
  - Relatively easy to extract and compare
  - It requires defining a **color space**
2.2 What is Color

• Color features, color spaces…but what is color?
  – What is the color of this apple?
  – So then...
  color is a property of objects!

• Imagine it is night, and there is no light
  – What is the color of this apple?
  – So then...
  color is a property of light!

• Close your eyes!
  – What is the color of the apple I am displaying?
  – So color happens in the observer!
    • If you see a red apple then you should get your eyes checked!

• So color is an event which occurs among three participants:
  – An object
  – A light source
  – And an observer

2.2 Color Spaces

• Multi-dimensional spaces in which, various dimensions describe various color components
• Correspond to the perception of colored light by three independent receptors that are stimulated at different wavelengths

  • Range of visible light: [380, 780] nm
  • Blue: 435.8 nm, green: 546.1 nm, red: 700 nm

• 3-dimensional Euclidean vector space
• Each component corresponds to the degree of stimulation (0-255)
• Additive color mixing with basic colors red, green and blue (primary colors)
  – Perception: imagine each color is radiated by a flashlight on a surface. The more light from each bulb the brighter the mixture (hence additive)
### 2.2 RGB Color Space

- \((0, 0, 0)\) black
- \((255, 255, 255)\) white
- \((255, 0, 0)\) red
- \((0, 255, 0)\) green
- \((0, 0, 255)\) blue
- \((255, 255, 0)\) yellow
- \((0, 255, 255)\) cyan
- \((255, 0, 255)\) magenta

Good representation of the visible light
- But **poor usability of the similarity search**
  - No consistent change in the perception of color (un-) similarity
  - Equal distances in different areas or different dimensions do not lead to the same color similarity

### 2.2 Example

- Magenta: more red then blue?

In RGB-space, the same!

### 2.2 CMYK

- Subtractive color model
- Reflects the mixing of paint or inks

- cyan
- magenta
- yellow
- black (key)

### 2.2 Optimal Color Space

- **Idea:** transform the RGB color information to achieve better partition of the color space with respect to human perception
- **Problem:** there is no single known color space resulting in uniform perception
- We aim for a “best possible” space, for color features

- **Supposition:** distance in spaces represents also distance in perception

### 2.2 Perceptional Color Spaces

- Attempt to sort the colors based on the human perception
  - Stretching of distances between dissimilar colors
  - Contraction of distances between similar colors
- Conservation of distances in space as the distances in the perception
2.2 Munsell Color System

- Albert H. Munsell: American Painter
- Book of Colors (1905)
- Discrete space based on perceived color similarity

Adjacent colors have the same perceptual distance
- Supported by psychological tests
- Variants of the Munsell color system used for color classification
  - Hair color and skin color
  - Colors of liquids (especially beer)
- Disadvantages:
  - Distances between non-adjacent colors do not respect perception
  - No simple transformation from RGB

2.2 CIE Color Spaces

- Commission International de l’Eclairage (Standardization Commission on Illumination) is proposing a better perceptual spaces with non-linear transformation of RGB values:
  - CIE 1976 (L*, a*, b*)
    - L - lightness, L = [0; 100];
    - a* - negative values indicate green, positive values magenta;
    - b* - negative values indicate blue, positive values yellow
- CIE models are surprisingly successful
  - CIE LAB implemented in Photoshop and in most color management systems

2.2 HSV Color Spaces

- Simpler than CIE spaces
- Based on hue, saturation, value
- Non-linear transformation from RGB, but easy to compute

HSV color space is
  - Intuitive and easy to understand
  - Nearly perceptual
- Good color spectrum for similarity search
- Used in MPEG-7 standard as one of the color spaces for image descriptors
2.3 Comparison of Image Material

- Compare images based on the color? Extract **color features** first
  - Each pixel of an image contains color information
- Images consist of many pixels
  - Pixel by pixel?
- Aggregation for comparisons?
  - Average color
  - Color histograms
  - Color layout (regions)

2.3 Average Color

- Calculate the average RGB values of all pixels and normalize by the number of pixels
  
  \[
  R_{avg} = \frac{1}{N} \sum_{p} R(p) \]
  
  \[
  G_{avg} = \frac{1}{N} \sum_{p} G(p) \]
  
  \[
  B_{avg} = \frac{1}{N} \sum_{p} B(p) \]

2.3 Average Color

- Comparison of 2 images \(x\) and \(y\) by using the Euclidean distance for the average color
  
  \[
  d_{avg}(x,y) = \sqrt{(R_{avg,x} - R_{avg,y})^2 + (G_{avg,x} - G_{avg,y})^2 + (B_{avg,x} - B_{avg,y})^2} \]

- **Very bad** similarity measure
- E.g., magenta image and red-blue image are the same according to average color

2.3 Average Color

- Perceptionally somewhat **questionable**…
- But…
  - Quick and easy to calculate and compare
- Best to use as a filter: exclude images
  - Dominant color influences the average color, the opposite is not valid
  - E.g., search for mostly blue images: exclude all images with red, yellow or green color averages

2.3 Average Color

- Specification, either directly through color values or by color wheel, sliders, etc.

2.3 Average Color

- Example query (QBIC tool from IBM)
2.3 Color Histograms

- A key measure for the occurring colors in the image material are **color histograms**
  - Partitioning of the color space
  - Usually 256 values per axis in 24-bit color images (i.e. $2^{24}$ colors, RGB) 16 Mio colors
  - A histogram column for each color
  - Height of the column corresponds to the normalized number of pixels with the specified color in the image
  - Normalization: scaling, so that the sum of the heights of histogram columns is 1

2.3 Color Quantization

- Better than average color
  - All colors in histogram columns really appear in the image
  - Average color (127,0,127)
  - (0,0,254)
  - (254,0,0)

2.3 Color Quantization

- Reduce histogram size through **quantization**
  - **Basic step**: disjoint partitioning of the color space by vector quantization
    - Mapping from a color to a color partition
    - A color is given through a $k$-dimensional real-valued vector - $K$ is usually 3 (RGB)
    - A mapping $Q: \mathbb{R}^k \rightarrow C$, $C := \{y_0, \ldots, y_{m-1}\}$
    - $C$ is called the set of code words (also Codebook)
    - $Q$ is called encoding (also known as code)

2.3 Color Quantization of HSV model

- **Requirements** for the coding/partitioning
  - Group only perceptionally similar colors in each partition
  - Each codeword should represent the best possible description of the colors it is assigned to (The centroid of the partition spaces are very often used as code words)
  - Minimize the number of partitions ($\rightarrow$ search efficiency)

- E.g., IBM’s QBIC-Tool (64 colors)
  - HSV is cylindrical
  - Hue is most important (divided into 18 segments, each of 20°)
  - For saturation and grey value 3 steps are sufficient
  - Four additional values for gray color

© Heron Project, Augsburg university
2.3 Color Quantization of HSV model
• $Q_c^{166}$ partitions the color space in
  $18 \cdot 3 \cdot 3 + 4 = 166$ different color values

2.3 Color Histograms
• E.g., an image and its histograms

2.3 Color Histograms
• Color histograms indicate the number of pixels
  in each color partition
• **Normalization** (e.g., dividing by the total
  number of image pixels) provides comparability to
  other pictures
• Comparison with other histograms is possible with
  different metrics

2.4 Matching
• Comparison of histograms
  - **Metric**: function $d: A \times A \rightarrow \mathbb{R}$
    with the following characteristics:
    • Non-negativity:
      $- d(x, y) \geq 0$
    • Identity:
      $- d(x, y) = 0$ iff $x = y$
    • Symmetry:
      $- d(x, y) = d(y, x)$
    • Triangular inequality:
      $- d(x, z) \leq d(x, y) + d(y, z)$

2.4 Comparison of Histograms
• Given: histograms $h_1$ and $h_2$
• **Minkowski distance** with parameter $r$:
  $$d_r(h_1, h_2) := \sum_{i \in C} |h_1(i) - h_2(i)|^r$$
  • $r = 1$: Histogram-$L_1$-norm
    (also: city block distance, Manhattan distance)
  • $r = 2$: Histogram-$L_2$-norm (Euclidean)

2.4 Minkowski Distance
• $h_1$: $h_2$: $\cdots$
2.4 Minkowski Distance

• It is efficient to compute, but does not take the similarity of colors into account
  – The distance between a red and a bright red image is the same as between a red and a blue one

• Works poorly in the case of color shifts because all columns are individually compared

2.4 Quadratic Distance Measure

2.4 Mahalanobis Distance

• If all colors are not correlated:
  – The covariance matrix is a diagonal matrix and the metric is therefore a weighted $L_2$-norm (Weights: reciprocals of the covariances)

• If some colors are correlated:
  – The coordinate system can always be transformed so that in the resulting system there are no more correlations (principal component analysis)
  – The Mahalanobis distance in the original coordinate system corresponds to a weighted $L_2$-norm in the new system

2.4 Comparison of Histograms

• Quadratic distance measures
  – Evaluates the relationship between different colors in the histogram
  – Cross-talk matrix: $A$ expresses pairwise similarity $a_{ij}$ between color $i$ and color $j$ ($a_{ii} = 1$ and $a_{ij} = a_{ji}$):

$$d_A(h_1, h_2) = (h_1 - h_2)^T \cdot A \cdot (h_1 - h_2) = \sum_{i \in C} \sum_{j \in C} a_{ij} \cdot (h_1(i) - h_2(i)) \cdot (h_1(j) - h_2(j))$$

• Color channel metrics aggregate the values for the R, B and G-Channel
  – Each image is thus represented by a vector with three components (rather than by a histogram)
  – Somewhat vague, however, can be computed efficiently

• Mean color-distance (average color)

• Also possible for higher moments
  – Variance-Color-Distance: $(\sigma_R^2, \sigma_G^2, \sigma_B^2)$
  – Skewness-Color-Distance
  – Comparison is made using Euclidean distance
2.4 Comparison of Histograms

- **Consideration**: complexity of the calculation vs. accuracy of the description
  - E.g., a simple Euclidean distance of three-dimensional vectors of color channel moments vs. the multiplication (166 x 166)-matrices for quadratic distance measures
  - But the cross-talk matrices naturally contain more semantics than the color channel moments

- Experiments for color queries in (Castelli / Bergman, 2002, Chapter 11):
  - Color channel metrics generally provide relatively poor retrieval accuracy
  - Surprisingly, the accuracy of the Minkowski distance and the one of the quadratic distance measures are quite similar for many image collections

2.4 Experimental Results

- Minkowski distances are usually enough for multimedia databases

2.4 Color Layout

- **Global** description vs. description of individual image segments (color layout)
  - Images are compared as the weighted sum of the (dis-) similarity of each region
  - First steps with simple grid distribution (Hsu and others, 1995)
  - Later approaches to common compositions (e.g., foreground motif)

- E.g.: (Stricker and Dimai, 1996)

2.4 Color Layout

- Fuzzy Regions
- Weighted Regions

2. Summary

- Color spaces
  - RGB, CYMK, HSV
- Extracting color features
  - Average color, color histogram, quantization
- Matching
  - Comparison of histograms, Minkowski distance, Quadratic distance, Mahalanobis distance
  - Color Layout

Next lecture

- Using texture for image retrieval
  - Basic texture features
  - Probabilistic models