CONTENT-BASED IMAGE RETRIEVAL

Multimedia Databases SS 23 (Exercises)

Prof. Dr. Harald Kosch/ Prof. Dr. Mario Döller
Tutor: Kanishka Ghosh Dastidar, Dr. Alaa Alhamzeh

Task 1: CBR

What is CBR?

- A content-based retrieval system processes the information contained in data and creates an abstraction of its content in terms of attributes.
- **E.g.:** Online clothes shopping might allow users to search by traditional categories (brand, price range) and also in terms of visual attributes (color, texture)
- Steps:

  Feature Extraction: The first step in the process is extracting image features to a distinguishable extent.

  Matching: The second step involves matching these features to yield a result that is visually similar.

CBR

- **Compared to each other, what are the benefits and limitations of ABIR and CBIR?**

Limitations of ABIR compared to CBIR

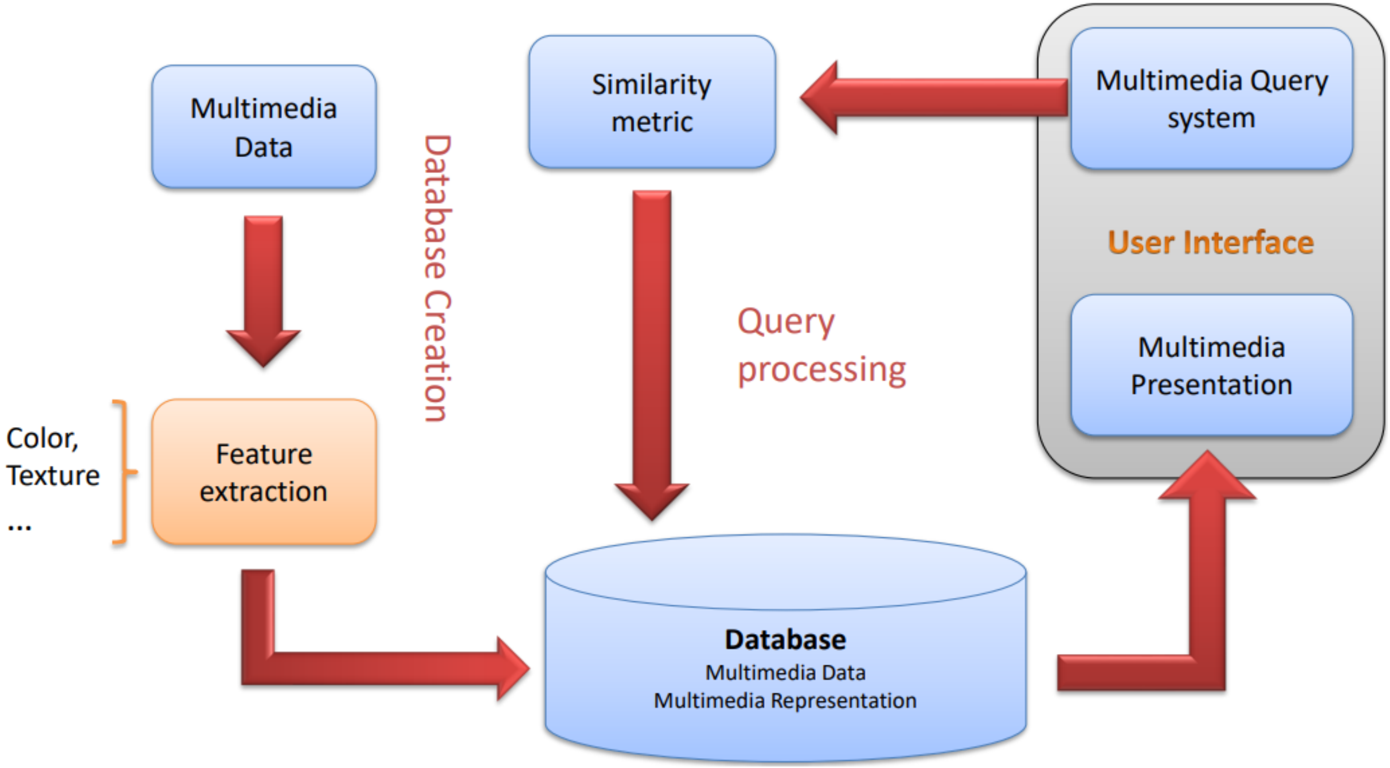
- Problem of Image annotations [Brahmi et al. 2004]
  - Large volumes of DB: manual image annotation is time-consuming and therefore costly.
  - The annotation is language dependent.
- Problem with human perception [Brahmi et al. 2004]
  - Human annotation is **subjective**
  - The accuracy and quality of the annotations are of the responsibility of the annotator and end-user.
- Problem in annotating images with words [Sclaroff et al. 1999]
  - Some images could not be annotated as it is difficult to describe their content with words.

Limitations of CBIR compared to ABIR

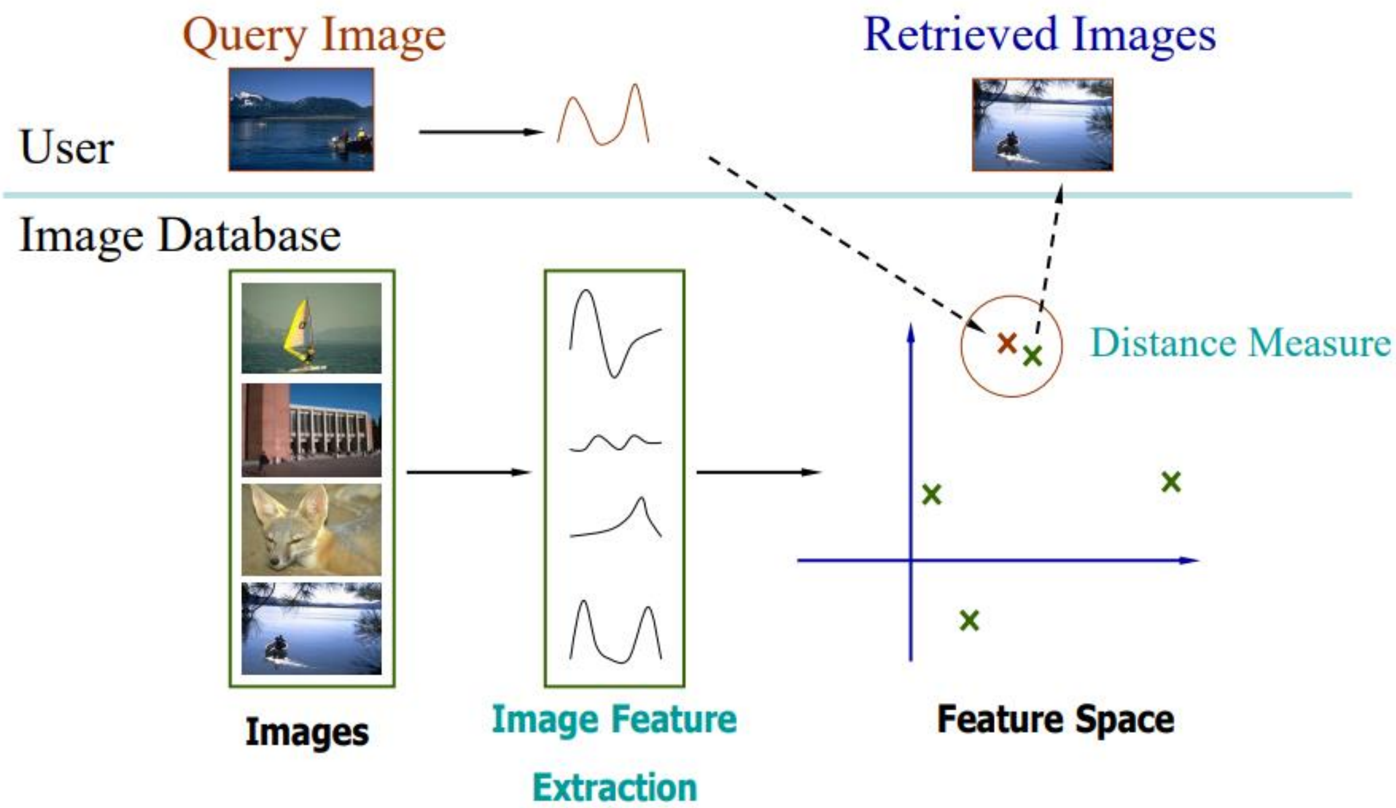
- Problem of Semantic Gap [Inoue et al. 2004, Brahmi et al. 2004]
  - Visual features cannot fully represent concepts.
- Problem of the availability of the image query [Inoue et al. 2004]
  - Users must have an example image in their hands

Task 1: CBR

- What are the components of a CBIR architecture? Explain the basic principle of each component using an example?



Task 1: CBR


Query Image
Retrieved Images
User
Image Database
Distance Measure
Images
Image Feature Extraction
Feature Space

Task 1: CBR

- **What is a feature vector?**
  - Numerical representation of object **features** in an n-dimensional **vector**
  - E.g. feature = avg color, fv = (255,0,0)
- **Which problem arise when indexing features vector?**
  - 'The query performance of access methods decreases , if the dimensionality of the underlying data set becomes high.'

    by Berchtold et al.
  - A large number of features diminishes the distinguishing power of each feature and the indexing structures.
- **How to avoid the 'Curse of Dimensionality'?**
  - Development of specific index structures that are specialized for high-dimensional data
  - Reduction of the vector dimension:
    - Feature selection
    - Feature extraction

Task 2: CBR Definitions

1. Dominant Color

- The Dominant Color Descriptor allows specification of a small number of dominant color values + statistical properties like distribution and variance.
- In contrast to histogram color, only the representative colors are selected from each region or image.
- Its purpose is to provide an effective, compact and intuitive representation of colors present in a region or image.

Task 2: CBR Definitions

- **Dominant Color**
  - The dominant color descriptor is defined by :

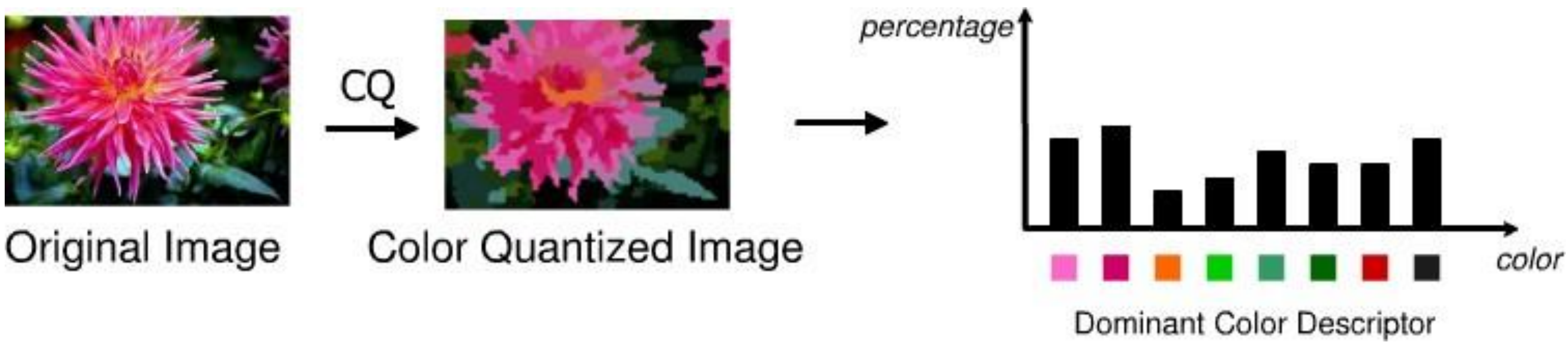
$$F = \{(c_i, p_i, v_i), s\}, (i = 1 \ldots N)$$

    - N is the total number of color clusters (= bins in the histogram) in the image region, such that $1 \leq N \leq 8$
    - c_i is the dominant color vector (3d RGB vector),
    - p_i is the percentage for each dominant color, such that
      - $P_i \in [0,1]$ and their sum is equal to 1,
    - v_i is its color variance,
    - s is a scalar that represents the overall spatial coherency of the dominant colors in the image.

Task 2: CBR Definitions

- **Dominant Color**
  - Often, the color variance v and spatial coherency s are not considered, which simplifies the definition of the dominant color descriptor to:

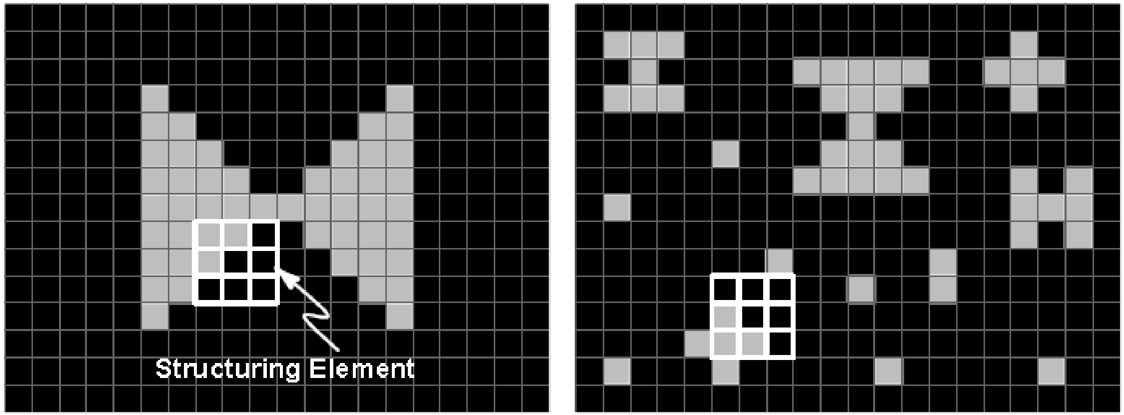
$$F = \{(c_i, p_i)\}, (i = 1 \ldots N)$$



Task 2: CBR Definitions

2. Spatial Coherency

- The spatial coherency of a given dominant color is measured with **the normalized average number of connected pixels of this color.** It is computed using a 3x3 mask
- The overall spatial coherency is a linear combination of the individual spatial coherencies weighted with the corresponding percentages `pi`.



Note: these two images have the same colors but the distribution of the colors is not the same.

Task 2: CBR Definitions

3. Distance metrics

- Degree of similarity between two points is measured by distance in data space. Several metrics have been established, i.e. euclidian distance.

4. Curse of Dimensionality

- CBR performs worse when dimensionality of the FVs increases.

5. Types of content-based queries

- **Point Query**: Retrieve all points with identical feature vector
- **Range Query**: Retrieve all points with a maximum distance from query point.
- **K-Nearest Neighbor Query**: Returns the k most similar results

Task 3: CBIR System

- What are the necessary conditions that must be fulfilled in order to be able to issue the following query to a CBIR system:
  - *Give me all images which contain a red car!*

- What are the problems that can occur?

Task 3: CBIR System

- **What are the necessary conditions that must be fulfilled in order to be able to issue the following query to a CBIR system:**
  - **Give me all images which contain a red car!**

To resolve this question using a ”query by feature”, we could:

1. Define the set of colors that correspond to the concept of ”red”
2. Define a set of contours (shape descriptors), which correspond to the shape of a car.
3. Perform a query by features based on these two features.

Task 3: CBIR System

- **What are the problems that can occur?**

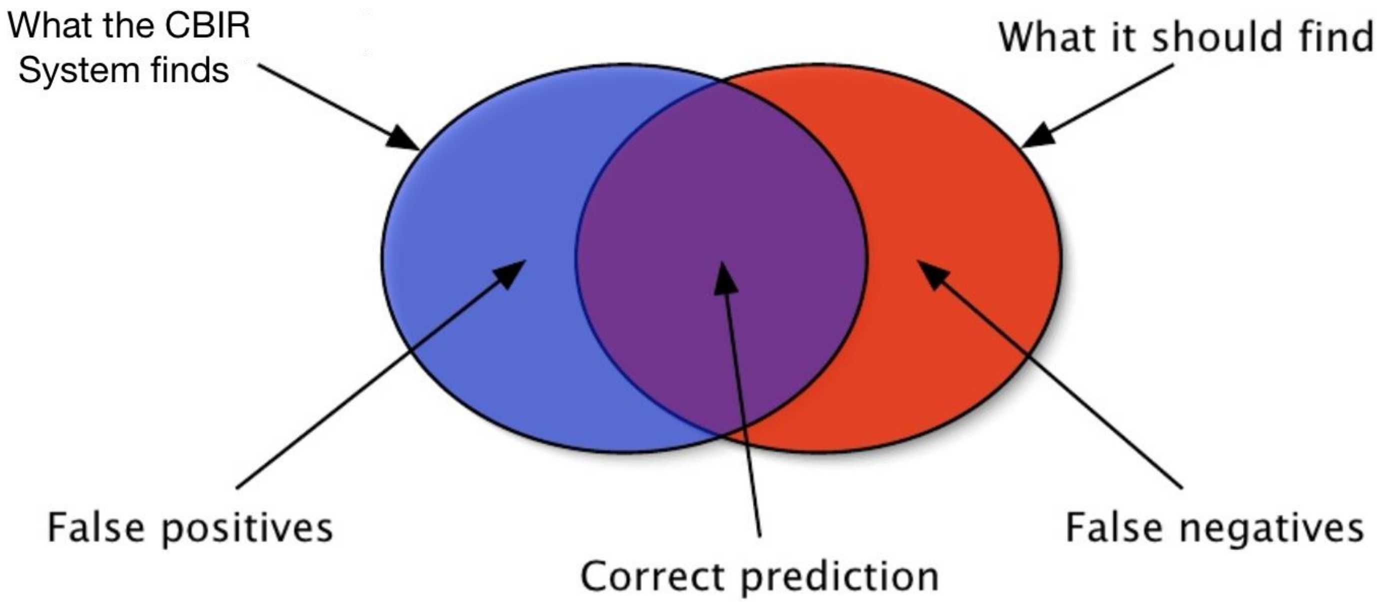
- In the QBF-solution, the difficulty lies in the definition of the shape, as cars can have very different contours (e.g. formula 1 cars, BMW cars, etc.).

- To which extent the definition of the shape should be precise?
  - precise definition will lead to a good precision but a bad recall.
  - less precise definition will provide bad precision (false positives) but a better recall.

- Thus, the appropriate strategy depends on the exact goals of the query.
  - But, it is very difficult to get this information from the user.

Task 3: CBIR System

- **How good is the result?**



Precision How many of the returned entities are relevant?

High precision = few false positives

Recall How many relevant entities are returned?

High recall = few false negatives

Task 4: Image Indexing by Colours

- Apply a uniform colour quantization of 8 bins.
- Which quantization area (range) do the colours in the two images belong to?

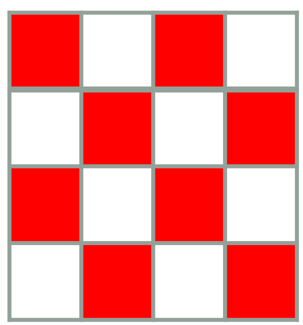


Image 1

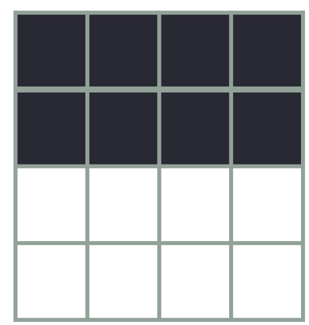
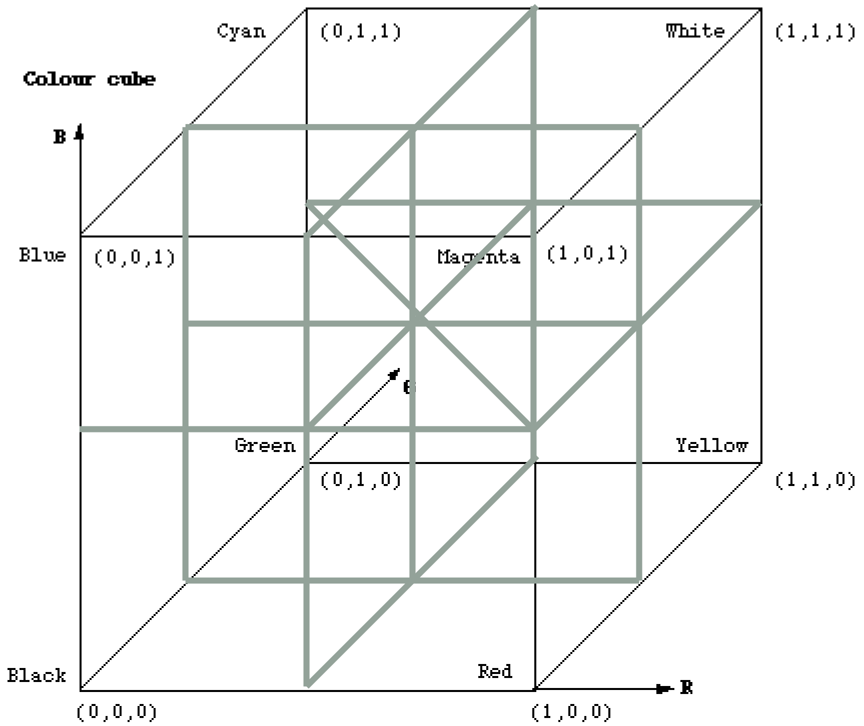
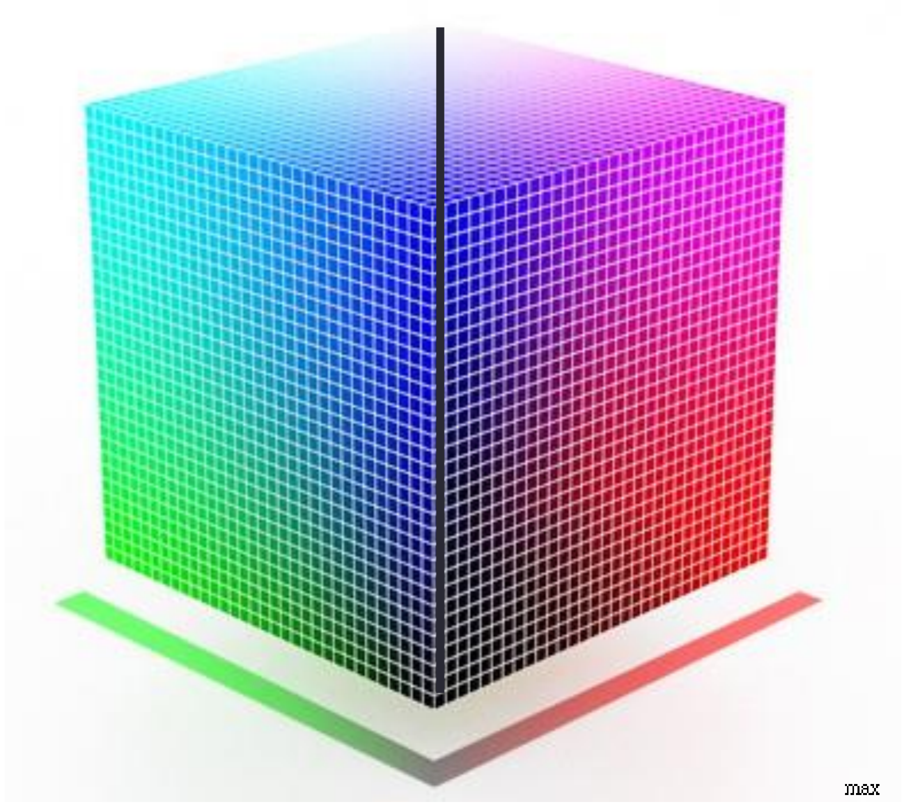


Image 2

Task 4: Image Indexing by Colours

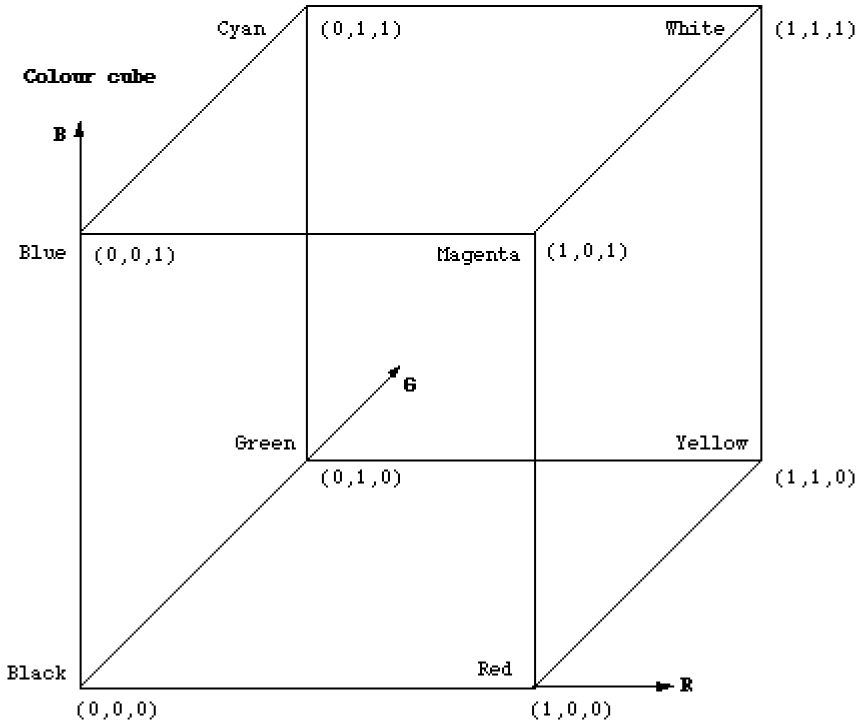
- Apply a uniform colour quantization of 8 bins.





Task 4: Image Indexing by Colours

- Apply a uniform colour quantization of 8 bins.



- (0,0,0)
- (0,0,1)
- (0,1,0)
- (0,1,1)
- (1,0,0)
- (1,0,1)
- (1,1,0)
- (1,1,1)

Task 4: Image Indexing by Colours

- Which quantization area (range) do the colours in the two images belong to?

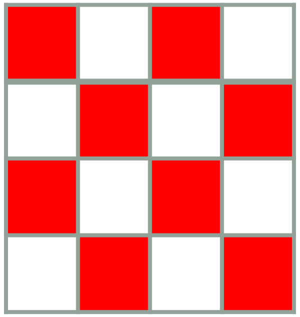


Image 1

- Red pixels are part of the (1,0,0) bin
- White pixels are part of the (1,1,1) bin

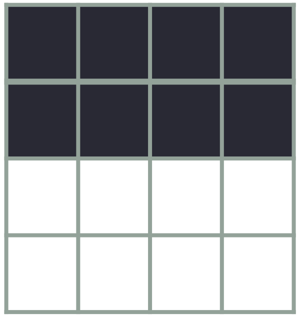


Image 2

- Black pixels are part of the (0,0,0) bin
- White pixels are part of the (1,1,1) bin

Task 4: Image Indexing by Colours

- Create a color histogram for both images

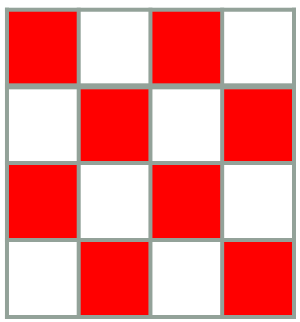


Image 1

- $H_1 = (0,0,0,0,8,0,0,8)$

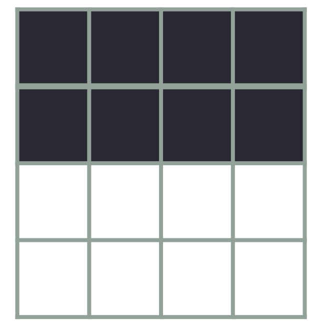


Image 2

- $H_2 = (8,0,0,0,0,0,0,8)$

Task 4: Image Indexing by Colours

- Apply a uniform bin quantization for 2 bits.

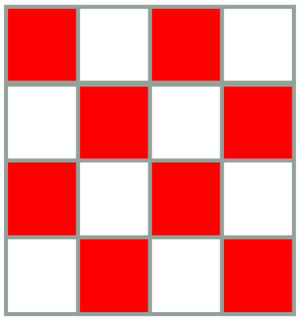


Image 1

- $H_1 = (0,0,0,0,8,0,0,8)$

=> $H_1 = (0,0,0,0,2,0,0,2)$

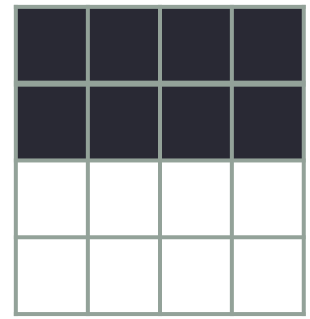


Image 2

- $H_2 = (8,0,0,0,0,0,0,8)$

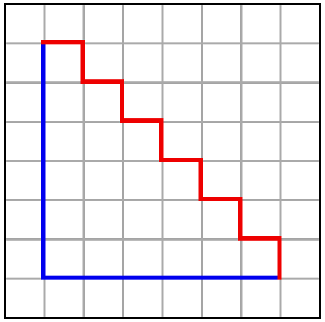
=> $H_2 = (2,0,0,0,0,0,0,2)$

Task 5: Similarity of images because of colour distribution

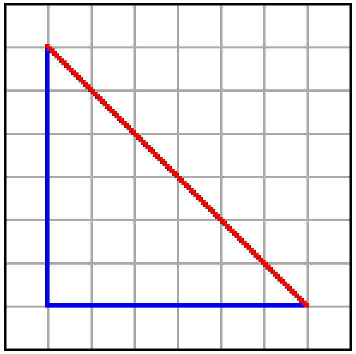
- **Minkowski Distances**

  Use the two images of exercise 4 and the results of exercise 4.b (without bin quantification). Determine the similarity of the images with the help of Minkowski distances:

  - L_1: Manhattan distance
  - L_2: Euclidian distance
  - L_∞: Maximal distance (also called Tschebyschow distance)



Manhattan Distance



Euclidean Distance

Task 5: Similarity of images because of colour distribution

- **Minkowski Distances**

Starting with $P = (x_1, x_2, \dots, x_n)$ *and* $Q = (y_1, y_2, \dots, y_n) \in R^n$

$$L_p(P,Q) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p}$$

- $H_1 = (0,0,0,0,8,0,0,8)$
- $H_2 = (8,0,0,0,0,0,0,8)$

$$L_1(H_1, H, 2) = |0-8| + |0-0| + |0-0| + |0-0| + |8-0| + |0-0| + |0-0| + |8-8| = 8+8 = 16$$
$$L_2(H_1, H_2) = \sqrt{(0-8)^2 + \cdots + (8-0)^2 + \cdots + (8-8)^2} = \sqrt{128} \approx 11,3$$
$$L_\infty(H_1, H, 2) = \max_i |x_i - y_i| = 8$$

Task 5: Similarity of images because of colour distribution

- Which result would you obtain if the red colour in the left image was black? Which conclusions do you draw?

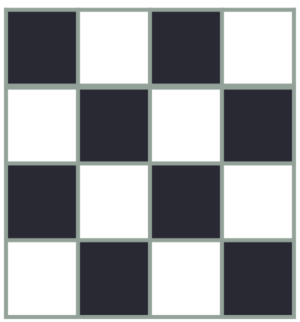


Image 1

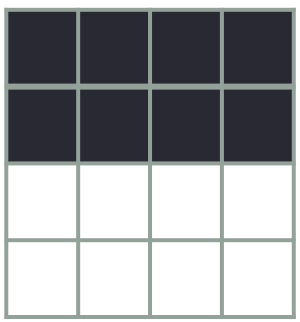


Image 2

Spatial Distribution of colors ignored!

Task 5: Similarity of images because of colour distribution

- **Statistical Distances**

Starting with the following colour distribution: $H_1 = (4, 4, 4, 4)$ and $H_2 = (8, 3, 4, 5)$

Non-parametrical distances:
Calculate the distances between H_1 and H_2 with the help of the following functions: Kolgomorov-Smirnov Distance, Chi-squared Distance.

Kolgomorov-Smirnov Distance:

$$KS(P, Q) = \max_i |F^r(i; P) - F^r(i; Q)|$$

$F^r(i; P)$ is equivalent to *cumulative histogram* of P in place i.

- Cumulative H_1 = (4, 8, 12, 16) and Cumulative H_2 = (8, 11, 15, 20)
  $\Rightarrow$ KS (H_1, H_2) =4

Task 5: Similarity of images because of colour distribution

- **Chi-squared Distance**

$H_1 = (4,4,4,4)$ and $H_2 = (8,3,4,5)$

$$D_\chi(P,Q) = \sum_i \frac{(x_i - f'(i))^2}{f'(i)}$$

$$f'(i) = \frac{x_i + y_i}{2}$$

$$f'(i_1) = 6, f'(i_2) = \frac{7}{2}, f'(i_3) = 4, f'(i_4) = \frac{9}{2}$$

$$D_\chi(H_1,H_2) = \frac{(4-6)^2}{6} + \frac{(4-\frac{7}{2})^2}{\frac{7}{2}} + 0 + \frac{(4-\frac{9}{2})^2}{\frac{9}{2}} = \frac{4}{6} + \frac{\frac{1}{4}}{\frac{7}{2}} + \frac{\frac{1}{4}}{\frac{9}{2}}$$

$$D_\chi(H_1,H_2) = \frac{2}{3} + \frac{1}{14} + \frac{1}{18} = \frac{50}{63}$$

$$D_\chi(H_1,H_2) \approx 0,794$$

Task 5: Similarity of images because of colour distribution

- **Parametrical Distance Function**
- Calculate the distance between H1 and H2. Use Weighted-mean-variance and the following training data:
  - V_1 (8,8,4,12), V_2 (4,0,0,16), V_3 (2,3,8,7), V_4 (4,4,6,10)

Weighted-mean-variance:

$$WMV(P,Q) = \frac{|\mu(P) - \mu(Q)|}{|\sigma(\mu(Ref))|} + \frac{|\sigma(P) - \sigma(Q)|}{|\sigma(\sigma(Ref))|}$$

μ: *Average*
σ: *Standard deviation*
$\mu(Ref)$: *Average calculated from training data*
$\sigma(Ref)$: *Standard deviation calculated from training data*

Task 5: Similarity of images because of colour distribution

- **Parametrical Distance Function**
  - V_1 (8,8,4,12), V_2 (4,0,0,16), V_3 (2,3,8,7), V_4 (4,4,6,10)

$$WMV(P,Q) = \frac{|\mu(P) - \mu(Q)|}{|\sigma(\mu(Ref))|} + \frac{|\sigma(P) - \sigma(Q)|}{|\sigma(\sigma(Ref))|}$$

- **How to calculate it?**

$$H_1 = (4,4,4,4) \text{ and } H_2 = (8,3,4,5)$$

  - Numerator:
    - Mean values of $\mu(H_1)$, $\mu(H_2)$
    - Standard deviation: $\sigma(H_1)$, $\sigma(H_2)$

$$\sigma(P) = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu(P))^2}{n}}$$

Task 5: Similarity of images because of colour distribution

- **Parametrical Distance Function**
  - V_1 (8,8,4,12), V_2 (4,0,0,16), V_3 (2,3,8,7), V_4 (4,4,6,10)
  - Ref = $\{V_1, V_2, V_3, V_4\}$

$$WMV(P,Q) = \frac{|\mu(P) - \mu(Q)|}{|\sigma(\mu(Ref))|} + \frac{|\sigma(P) - \sigma(Q)|}{|\sigma(\sigma(Ref))|}$$

- **How to calculate it?**
  - Denominator: $\sigma(\mu(Ref)) \approx 1{,}224$
    - $\mu(Ref) = (\mu(V_{d1}), \mu(V_{d2}), \mu(V_{d3}), \mu(V_{d4}))$
    - $\sigma(Ref) = (\sigma(V_{d1}), \sigma(V_{d2}), \sigma(V_{d3}), \sigma(V_{d4}))$

$$\sigma(P) = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu(P))^2}{n}}$$

Task 5: Similarity of images because of colour distribution

- **Parametrical Distance Function**
  - V_1 (8,8,4,12), V_2 (4,0,0,16), V_3 (2,3,8,7), V_4 (4,4,6,10)

$$WMV(P,Q) = \frac{|\mu(P) - \mu(Q)|}{|\sigma(\mu(Ref))|} + \frac{|\sigma(P) - \sigma(Q)|}{|\sigma(\sigma(Ref))|}$$

- **How to calculate it?**
  - Denominator: σ (σ (Ref)) ≈ 1,716
    - Mean value of μ (σ (Ref))

$$\sigma(P) = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu(P))^2}{n}}$$

- **Result:** WMV(H_1, H_2) ≈ 1,907