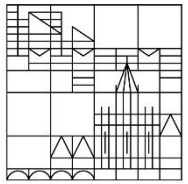
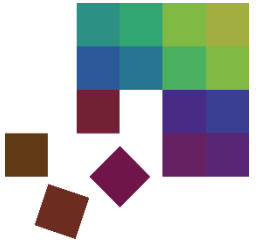


University of Konstanz
Data Analysis and Visualization Group



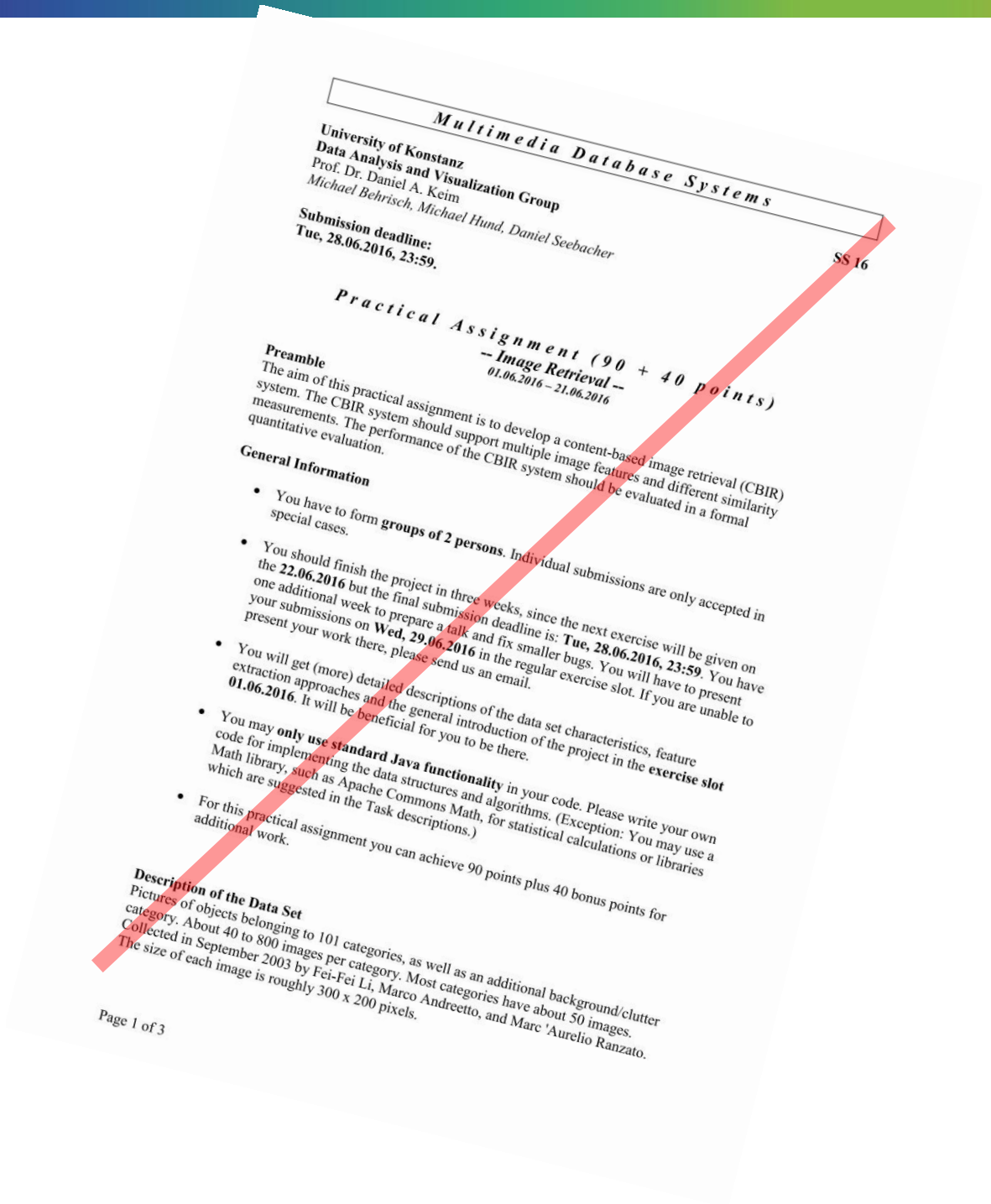
MMDBS
Practical Assignment

Michael Behrisch, Michael Hund, Daniel Seebacher

Prof. Daniel A. Keim

Disclaimer

Assignment sheet will be updated on Ilias



Multimedia Database Systems

University of Konstanz
Data Analysis and Visualization Group
Prof. Dr. Daniel A. Keim
Michael Behrisch, Michael Hund, Daniel Seebacher

Submission deadline:
Tue, 28.06.2016, 23:59.

SS 16

Practical Assignment (90 + 40 Points)
– Image Retrieval –
01.06.2016 – 21.06.2016

Preamble
The aim of this practical assignment is to develop a content-based image retrieval (CBIR) system. The CBIR system should support multiple image features and different similarity measurements. The performance of the CBIR system should be evaluated in a formal quantitative evaluation.

General Information

- You have to form **groups of 2 persons**. Individual submissions are only accepted in special cases.
- You should finish the project in three weeks, since the next exercise will be given on the **22.06.2016** but the final submission deadline is: **Tue, 28.06.2016, 23:59**. You have one additional week to prepare a talk and fix smaller bugs. You will have to present your submissions on **Wed, 29.06.2016** in the regular exercise slot. If you are unable to present your work there, please send us an email.
- You will get (more) detailed descriptions of the data set characteristics, feature extraction approaches and the general introduction of the project in the **exercise slot 01.06.2016**. It will be beneficial for you to be there.
- You may **only use standard Java functionality** in your code. Please write your own code for implementing the data structures and algorithms. (Exception: You may use a Math library, such as Apache Commons Math, for statistical calculations or libraries which are suggested in the Task descriptions.)
- For this practical assignment you can achieve 90 points plus 40 bonus points for additional work.

Description of the Data Set
Pictures of objects belonging to 101 categories, as well as an additional background/clutter category. About 40 to 800 images per category. Most categories have about 50 images. Collected in September 2003 by Fei-Fei Li, Marco Andreetto, and Marc 'Aurelio Ranzato. The size of each image is roughly 300 x 200 pixels.

Page 1 of 3

Project Overview

- Development of a Content-based Image Retrieval (CBIR) System

Query

Analysis Result

Project Overview

- Development of a Content-based Image Retrieval (CBIR) System supporting the following features
  - Feature Extraction [5 + 10 + 15 Points]
  - Similarity Functions [5 + 15 + 10 Points]
  - A GUI [10 Points]
  - An Evaluation [20 Points]
  - **Optional:** Nice Features (weighting, index structures etc.) [20 Points]

Organizational Requirements

- Form groups of **2 persons**

- Working period: 01.06.2016 – 22.06.2016 (3 weeks)
- Submission: 28.06.2016
- **Presentation:** **29.06.2016** (medical certificate if you are not able to present)

Achievable Points

- 30 weeks * 20 Points * 1.5 = **90 Points** (count into Bonus of all exercies)
- 20 Bonus Points

Organizational Requirements

- Programming language: Java

- You may **only use standard Java functionality**
  - No external libraries, if not specified
  - Exception: Apache Commons Math for statistical calculations

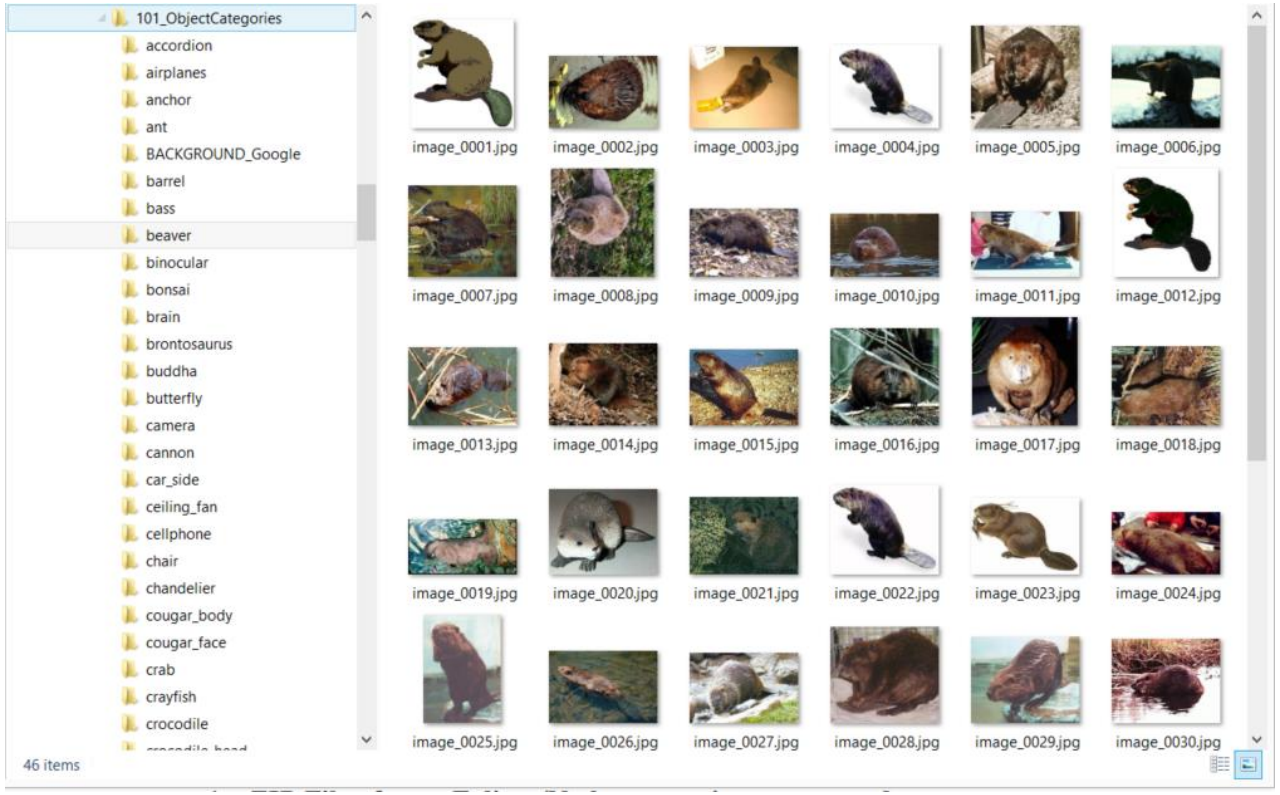
Deliverables @ 28.06.2016 / Presentation 29.06.2016

A single ZIP-File containing the following parts

- ZIP File of your Eclipse/Netbeans project source code
- A PDF describing a short summary what you implemented and what is missing.
- A PDF for the evaluation with descriptions of your findings + screenshots of your tool + precision/recall curves etc.
- A PDF describing how to start and use your tool (short with a few screenshots)
- Your presentation

Dataset

- Image data of 101 different categories / classes + 1 noise class
- 40-800 (typically 50) images per class
- Average size: 300x200 pixels
- Download at http://www.vision.caltech.edu/Image_Datasets/Caltech101/Caltech101.html
- Delete noise class



Task 1: Data Acquisition [0 Points]

- Download data

- Get an overview

Task 2: Implementation of Feature Descriptors [30 Points]

- **Color Histogram** [5 Points] with the following parameters
  - #bins
  - #cells

- **Global Edge Histogram** [10 Points]
  - Convolve image with a Sobel-Operator beforehand
    Java provide methods for the convolution of images (c.f. ConvolveOp)

- **Texture Haralick Features** [15 Points]
  - Implement the gray-level co-occurrence matrix (see exercise 7)
  - Implement the first 12 Haralick features (Descriptions are in the paper [1])

[1] Haralick, Robert M., Karthikeyan Shanmugam, and Its' Hak Dinstein. "Textural features for image classification." Systems, Man and Cybernetics, IEEE Transactions on 6 (1973): 610-621.

Task 3: Implement Similarity Functions

- **Euclidean Distance** [5 Points]

- **Quadratic Form Distance** [15 Points]
  - Similarity matrix must not be hard coded
  - You will get a matrix with a weighting based on human perception in about a week

- Quadratic Form Distance (**Lower-bound property**) [5 + 5 Points]
  - Singular Value Decomposition of Matrix + initial Eigenvector multiplication
  - Eigenvalue limitation / dimension reduction

Query Processing for Quadratic Forms

- i.e. there is a diagonal matrix $W = diag(w_1, \dots, w_n)$ and an orthonormal matrix V with: $A = VWV^T$
- for all vectors (histograms) p and q this means:

$$D_A(p,q) = \sqrt{(p-q) \cdot VWV^T \cdot (p-q)^T} = \sqrt{(pV - qV) \cdot W \cdot (pV - qV)^T} = D_W(pV, qV)$$

- D_A is therefore equivalent to the weighted Euclidean Distance transformation D_W, after transforming all vectors by the basis V
- The weights are the Eigenvalues of matrix A, the columns in V are the Eigenvectors of A

Query Processing for Quadratic Forms

- Dimension reduction
  - As A is positive definite, all Eigenvalues w_i are positive. That is $w_i > 0$; and a truncation of the d-dimensional Vectors pV and qV to $r < d$ dimensions provides a guaranteed lower bound for $D_W = D_A$:

$$D_{W,r}(pV, qV) = \sqrt{\sum_{i=1}^{r} w_i(pV_i - qV_i)^2} \leq \sqrt{\sum_{i=1}^{d} w_i(pV_i - qV_i)^2} = D_W(pV, qV)$$

  - The lower bounding property provides completeness of query processing
  - Observation: The transformation depends on the similarity matrix A

Task 4: GUI [10 Points]

- Your user interface needs to have at least the following features
  - Load a folder with your image dataset
  - Select a query image
  - Choose a feature descriptor (incl. parameters, e.g. #cells in color histograms)
  - Choose a similarity function (incl. parameters, e.g. #eigenvalues in QFD)
  - Show most similar images to query image

Task 5: Evaluation [20 Points]

- Test at least 5 different query objects
  - Create a precision/recall plot

- **Interpret your results** (reason on the performance of the feature descriptors, and all parameter settings)

Optional Task: Surprise us! [20 Points]

- Improve your CBIR System by nice features, e.g.
  - Combine / weight feature descriptors and distance functions
  - Implement a more advanced image feature descriptor (e.g. check OpenCV, OpenIMAJ, ...)
  - Implement a multi-scale approach that takes into account differently scaled version of the input image to compute the descriptors
  - Use an index structure to speed up the calculation.
  - ...

- You may use external libraries for this task.
- Show the usefulness (e.g. speedup + higher accuracy) with a short evaluation.

Plan for the next weeks

- 01.06.2016 Project Introduction
- 08.06.2016 Optional exercise: you can ask questions
- 15.06.2016 Optional exercise: you can ask questions
- 21.06.2016 Exercise 8 online
- 22.06.2016 Optional exercise: you can ask questions
- 28.06.2016 Exercise 9 online
- 29.06.2016 Presentation of Practical Assignment
- 06.07.2016 Discussion Exercise 8 + 9 + Questions before Exam
- 11.07.2016** Oral Exam