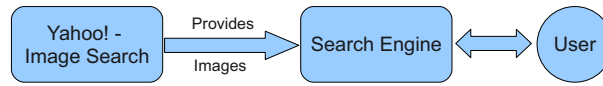


### Overview

Due to the strictly growing amount of multimedia data in the internet, there is a high demand for efficient search engines. Up to now mostly textual attributes of images are used for a search. The task of Content-Based Image Retrieval is to use the pure image content for retrieval. Since there are some drawbacks because of the complex semantics in images, search strategies like Relevance Feedback must be used for good retrieval results. This leads to an increased effort for the user because he has to decide on the relevance of search results. The task of Active Learning is to minimize the human effort in this process by an intelligent choice of the user's incorporation.

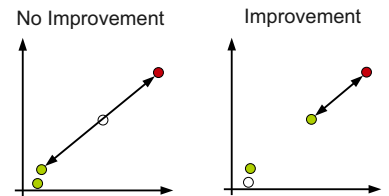
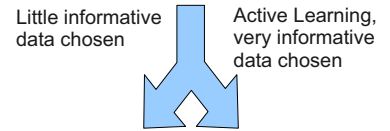
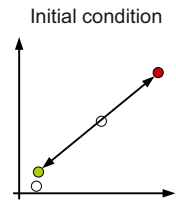
### Development of an Online Image Search Engine

- Created in 2008 at the NI-Group of TU-Berlin
- Capable of query-based search for images
- Improvement of search results by **Relevance Feedback** and **Active Learning**



### Active Learning

- Classification algorithms need preclassified data
- Human classification work is very expensive
- Unclassified data numerously available
- **Choose the most informative data for classification by the user**
- Classification algorithm needs fewer examples for good results
- Necessity of human work is minimized



- Many different approaches to Active Learning existent
- Mostly based on heuristics
- Active field of research in the Neural Information Processing Group of TU-Berlin.

### Content-Based Image Retrieval

- Current Search Engines are mostly based on a text search
- Content-Based Image Retrieval uses the content information of images

#### Use of Computer Vision



- Exploit of content information leads to abstract mathematical description of images

#### MPEG-7 Standard

- Created to find a way for standardized metadata representations of multimedia data

- Among its purposes are:
- Fast and efficient retrieval of multimedia data
  - Comprehensive description of multimedia contents

Drawback:

- Semantic gap
- Not capable of representing the complete semantic information of images

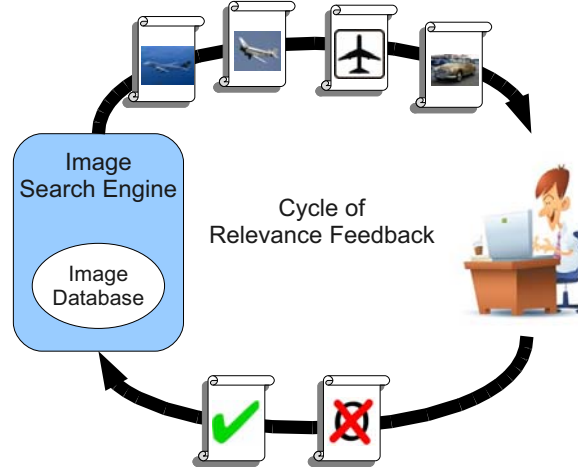
#### Different semantics, similar appearance



### Relevance Feedback

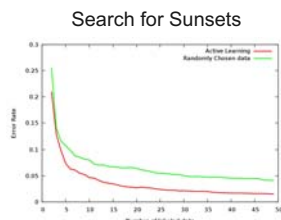
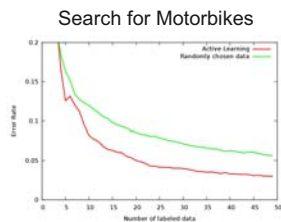
- Iterative process of search
- Provided information about the user's search intention continuously grows.

1<sup>st</sup> step: The user receives the search results

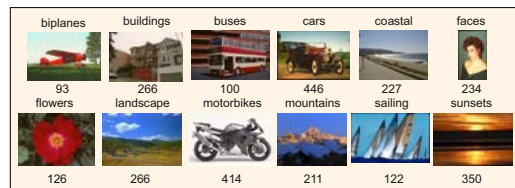


2<sup>nd</sup> step: The user provides relevance information

### Results



- Test of Active Learning in Content-Based Image Retrieval on a set of 12 classes of images



- Test scenario:
- One class relevant, the others irrelevant
  - Simulated Relevance Feedback Process

- Active Learning algorithm based on the use of Support Vector Machines
- Asked for the relevance of the image with the most uncertain label

- Active Learning leads to significant improvement of search results
- Encouragement for further research