

# DAM Homework (4)

2012-10-30

# Image similarity computing

- Given: 10 images,  $I_1, I_2, \dots, I_{10}$
- Goal:
  - compute similarities between  $I_1$  and  $I_i (i = 2, \dots, 10)$
  - find the most similar image to  $I_1$

# The simplest solution

- Image Feature vector:  $I_i \rightarrow \mathbf{F}_i$ 
  - RGB based moments, 9 float numbers
- Similarity:
  - dot product:  $\frac{\mathbf{F}_i \cdot \mathbf{F}_j}{\|\mathbf{F}_i\| \|\mathbf{F}_j\|}$
  - distance based:  $\exp(-\|\mathbf{F}_i - \mathbf{F}_j\|^2)$

# Better ways ...

- Position and structure
- Better color spaces, Lab/HSV/Yuv ...
- Texture features, Gabor filter bank
- Better similarity computing
  - advanced machine learning methods

# Constraints

- Use **python** and **PIL**

# More considerations

- How about work on 1,000 images?
- How about work on 1,000,000 images?
- Other media:
  - Audio ???
  - Video ???
  - HTML pages ???

**ftp://10.214.0.111/2012/homework-04/**

**user: stu**

**pass: 2012**

**deadline: 2012-11-10**