DAM Homework (4)

2019-10-22

Image similarity computing

- Given: 10 images, $I_1, I_2, ..., I_{10}$
- Goal:
  - compute similarities between I_1 and $I_i (i = 2, ..., 10)$
  - 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: $\dfrac{\mathbf{F}_i \cdot \mathbf{F}_j}{\|\mathbf{F}_i\|\|\mathbf{F}_j\|}$

  - 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 or **opencv/pngjs**

More considerations

- How about on 1,000 images?

- How about on 1,000,000 images?

- Other media:

  - Audio ???

  - Video ???

  - HTML pages ???

- Why not try deep learning?