

Multimedia oral examination - Video Retrieval and Instance Search

Multimedia oral examination - Video Retrieval and Instance Search

1.Video Retrieval

- (1)What is Video Retrieval and its challenges?
- (2)Current approaches.
- (3)How to do Video Retrieval based by Frame based approach

2.Instance Search

- (1)What is instance search?
- (2)Major Challenges in Instance Search?
- (3)NII INS system
 - (3-1)What is NII INS system?
 - (3-2)Pros and cons?
- (4)Deep Features

1.Video Retrieval

(1)What is Video Retrieval and its challenges?

Challenges are about scalability.The increasing scale of data brings challenges to speed and memory and search quality.

(2)Current approaches.

There are video signature based, trajectory based and frame based approach

Video signature based compute global feature for each frame and aggregate them as one feature vector to video level It is intuitively but only good for nearly identical videos.

Trajectory based approach extract feature from tracking trajectories of objects in video sequence.It has high complexity in memory and computation but bad performance.

Frame based approach regard video as a collection of images,decompose video query to frames and convert frame rankings to to video ranking,which is simple, work but low scalability.

(3)How to do Video Retrieval based by Frame based approach

Frame based approach samples frames from videos and extracts BoW or VLAD features from them,then caculate similarities between frames. To convert frame similarities to video similarity,we aggregate frame similarities within same time window and take the highest similarity score as video similarity.

2.Instance Search

(1)What is instance search?

Instance Search search for given instances of a specific object or persons, and localize them by bounding box.

(2)Major Challenges in Instance Search?

First,it faces scalability challenges because the computation cost is higher than image search if there are too many instance.

It also faces representation challenges. Global representation does not work.

Although keypoint features cover most of the local regions and very distinctive.But sometimes too distinctive and do not cover an object exactly,and they are in big number and sensitive to deformations.

And Bounding boxes produce too many meaningless candidates.So an object level representation is required for instance search.(Deep feature)

(3)NII INS system

(3-1)What is NII INS system?

NII INS system is a representative method and baseline framework for instance Search Task. This approach use detect point features by Hessian-Affine and describe them by RootSIFT.Then quantize features by BoVW,using tf-idf weighting and inverted file to organize them.Fused by average pooling.

(3-2)Pros and cons?

The advantages of this approach is it combines the promising side of BoVW with inverted file by using very-large vocabulary The disadvantages are ,first Keypoint features performs bad on nonrigid instances. And very-large vocabulary training is time-consuming.

(4)Deep Features

Deep Feature combined with convolutional neural network are used for instance search,which works pretty well but requires a well pixel level annotation training set.