

Title: Multi-target cross-camera tracking

Industrial Applications □Intelligent Manufacturing □Smart Driving □Intellectual Life □Smart Medical ■Smart City

[Overall background]

Face recognition has been widely applied in smart city. However, restricted by the camera's field of view and angle, in many cases, the face cannot be obtained clearly. Cross camera tracking technology can supplement the limitations of face recognition and make the recognition of pedestrians more scalable.

[Business background]

Cross camera tracking, which is associated with the same target in different cameras, depends on a variety of information and technologies, such as spatiotemporal relationships, such as pedestrians recognition. At the same time, because of the uncertainty of all kinds of information, we need to deal with them comprehensively.

Project description:

[Problem description]

The venue consists of 3 buildings with 18 cameras, 2 roads and 1 square. The cameras installed at the 2 entrances of each building can take a clear face to enter the building target.

200 goals enter to at least one building after entering the arena.

The organizer provides similar ID photos for all targets and video captured by all cameras in the stadium. The player gives the time and location of the target appearing in all videos.

18 cameras are all 2 million pixel network gun camera, 1/2.8 "Progressive Scan CMOS; effective pixel 2 million; star level, minimum illuminance 0.001Lux@F1.2; C lens interface; DC12V + 10% working voltage, PoE; power 4W MAX; mm): 62 x 62 x 88.

[User expectations]

It can recognize face correctly and get the target at the time of entering a building. It can detect pedestrians in video correctly. According to the time when the target enters the building and the human shape information at that time, it can associate the target to the pedestrians appearing in other videos.

[Expected economic effect]

Face recognition combined with cross camera tracking has great application value in smart city.

[Technical path]

Face recognition, pedestrian tracking, pedestrian ReID

[Technical indicators]

A ID is assigned to each target, and the ID is assigned to each pedestrian in the video by face recognition and cross camera tracking.

If the pedestrian detection results (bounding box) is highly coincided with one owning its ID

(IoU > 0.5), the result is correct (TP); otherwise, the result is wrong (FP).

If a correct pedestrian with its ID does not have a test result that highly coincided with (IoU > 0.5), the result is considered to be missing (FN).

Through two indicators to measure the achieved effect of the task:

$$IDP = TP / (TP + FP)$$

$$IDR = TP / (TP + FN)$$

The final ranking is determined by IDF.

$$IDF = 2 * IDR * IDP / (IDR + IDP)$$

[Standard Submission]

For each face photo, the following structure is provided JSON:

```
[{video_id: 0,  
  [{frame_id: 5, x: 100, y: 200, w: 32, h: 64},  
    ... ],  
},... ]
```

Where video_id is the number of video, frame_id is the number of frames in the video, x and y are the coordinates of the upper left corner of the pedestrian, w and h are the width and height of pedestrian results.

[Task list]

The organizers provide 220 face photos and 18 videos to the contestants marked with ID. The corresponding pedestrians in 200 of the 220 photos appear in some 18 videos. The contestant submits the JSON file and gives the time and location of each ID in each video, in which the time is represented by the frame number, and the position is represented by the coordinates and height of the upper left corner and the width. The specific format is as shown by the "submission standard".

Reference information

[Reference tool]

None

[Reference data]

None

[Data interface]

None