Deeper and Wider Siamese Networks for Real-Time Visual Tracking

Zhipeng Zhang and **Houwen Peng**

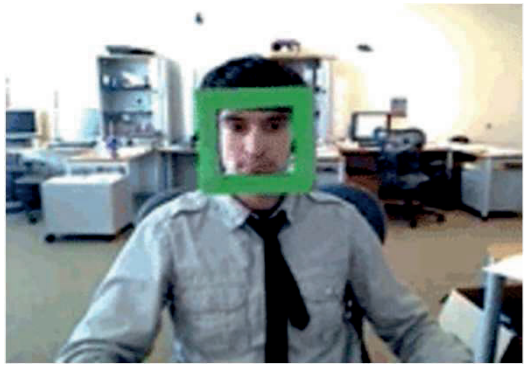
Microsoft Research Asia (MSRA)

CVPR 2019 Oral

June 16th - June 20th, Long Beach, CA

Visual Object Tracking

- Definition
  - It aims to estimate the position of arbitrary targets in a video sequence, given only the location in initial frame.
- Category
  - **Single object tracking**
  - Multiple object tracking





https://www.learnopencv.com/object-tracking-using-opencv-cpp-python/

Visual Object Tracking

- Challenges

Illumination Variation



Occlusion



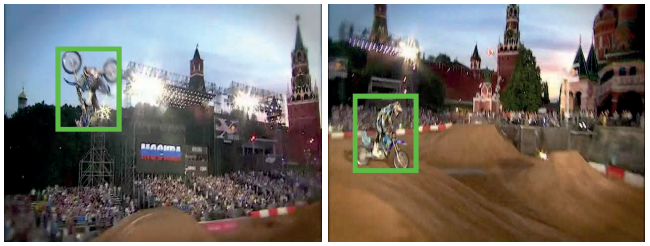
Background Clutters



Scale Variation



Rotation



Motion Blur

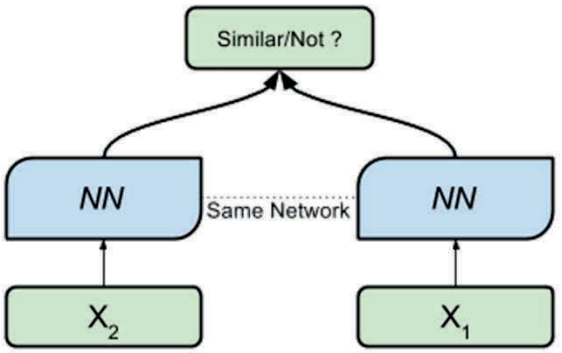


Outline

- Background on Siamese Trackers
- Motivation
- Analysis and Guidelines
- Method
- Experiments

Background on Siamese Trackers

- Siamese network architecture
  - Network and weight sharing
  - Metric learning, loss
  - Increase training samples naturally
- Applications
  - Face verification
  - Person re-ID



The **Distance Function** decides if the output vectors are close enough to be similar

The **Neural Network** transforms the input into a properties vector

Input Data (image, text, features…)

Parkhi, O.M., Vedaldi, A., Zisserman, A.: Deep face recognition. BMVC 2015 1(3) (2015)
https://aws.amazon.com/cn/blogs/machine-learning/combining-deep-learning-networks-gan-and-siamese-to-generate-high-quality-life-like-images

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：https://d.book118.com/887153104060006143