Real-Time Long-Term Visual Object Tracking

Martin Simon, xsimon14@stud.fit.vutbr.cz, Faculty of Information Technology, Brno University of Technology, April 2015

**Object Tracking**
- **Strong keypoints** like SURF or SIFT are too complex
- **Poor quality** Good Features To Track used
- **Detect up to 16** of these points on object area
- **Mask computed as a difference** between in-box and out-box regions
- **Increase a number** of successfully tracked points during the tracking itself
  - Tracking point forward multiple times
  - Tracking all the images backward
  - Choose the one closest to the original point

**Detector-Tracker Fusion**
- **If we have both** the detector’s and the tracker’s results, it seems reasonable to fuse them
- The fusion with a correction by the detector helps to decrease a tracker’s drift
- It is also a sign that the system could learn safely
- The adaptive learning is performed in these situations:
  1. The fusion is done, which means both the tracker and the detector agreed
  2. Only the tracker has a valid result, but its confidence is high enough

**Object Detection**
- **Variance Filter**
  - First detection stage
  - Reject majority of false positives
- **Random Fern Classifier**
  - Second detection stage
  - Trade-off between speed and quality
- **Normalized Cross Correlation**
  - Final detection stage
  - Similarity with positive and negative samples

**Evaluation**
- Overall success rate is a mean of success of all sequence frames from several runs
- The particular success rate is equal to complement to 1 of normalized Euclidean distance of the tracker’s result and the groundtruth

Better performance in 5 out of 6 tested datasets!