Short course
An introduction to action recognition and tracking in videos

Tracking in videos

Massimo Piccardi
University of Technology, Sydney, Australia

These notes are still very preliminary
Agenda

- Definition of tracking
- Example
- Model-based tracking
- Models and features
- State-based models

Tracking

- Tracking aims to extract the trajectory of a moving object ("target") from sequential data (radar, sonar, video etc)
- The word "tracking" seems to have taken on different meanings in different communities:
  - in the signal processing community, tracking means inferring the exact trajectory of a moving object out of noisy data; if multiple targets are present, associating a measurement with an object is known as "data association" (or data correspondence)
  - the computer vision community seems to equate tracking to data association
Model-based tracking

• It is generally very useful to create a **model** of the tracked target

• The model is not directly observed; rather, inferred from the measurements of the tracked target. It may contain a "denoised" version of the measurements and should support precise tracking and effective data association

• The main steps of tracking include repeated:
  – *Prediction* of where the target is expected to be
  – *Data association* between measurements and targets
  – *Update* of all the targets' models
Models and features for tracking in videos

• Models and features should be chosen based on the nature of the target (e.g. a car vs a human) and the complexity of the scene (frequent occlusions, changes in viewpoint etc)

• Both the model and features should enable precise tracking and effective data association from frame to frame

  • human shape model: 3D ellipsoid
  • human appearance model: texture template
  • features: pixels’ values in the predicted area
  • prediction: Kalman filter with constant velocity model

State-based models

• In the following, we will assume that the target’s model is represented as a continuous, multivariate random variable (called the target’s state)

• Each measurement of the target is a continuous, multivariate random variable, too, yet often of different nature and dimensionality

• We will adopt the framework of recursive Bayesian estimation to infer the states from the measurements
Main references

A survey paper:


A recent book:


On data association:


Main references

Some notable and/or recent approaches (my selection):


• Porikli, F.; Tuzel, O.; Meer, P., “Covariance Tracking using Model Update Based on Lie Algebra”, CVPR 2006, 728 – 735

• Williams, O.; Blake, A.; Cipolla, R., “Sparse Bayesian learning for efficient visual tracking”, T-PAMI, vol. 27, no. 8, pp. 1292 - 1304, Aug. 2005
