

Literature Review

Willie

August 18, 2011

Overview

The phrase ‘multiple target tracking’ (MTT) has been used to refer to two distinct but related problems: the first is the problem of detecting and tracking the positions of multiple targets of interest in a video, and the second is the problem of associating data points, where each data point is assumed to represent the centroid of some target whose position varies with time. The first is a problem in the field of machine vision and may use data association techniques from the second to accomplish MTT after targets of interest are detected and their positions represented as points.

The goal of this project is to apply the generalized Polya urn Dirichlet process mixture model (GPUDPM) to videos containing multiple targets in a way that accomplishes data association without requiring explicit target detection. We hope to allow for a flexible MTT algorithm that can handle videos with different numbers of targets, with diverse backdrops, whose targets display a wide range of characteristics, and where there are time-varying shifts in targets’ characteristics, perspectives, or orientations.

The following are summaries of papers focusing on MTT as it relates to the association of target centroids, MTT as it relates to machine vision, and the generalized Polya urn Dirichlet process mixture model.

Data Association Related MTT

Papers on this topic often involve algorithms that take a set of data, where each point has a time-value and is assumed to represent the centroid position of a target at the given point in time, and aim to return the most probable set of tracks formed by ‘associating’ these data points into sequences.

In Bar-Shalom’s classic MTT survey paper (Tracking Methods in a Multitarget Environment, 1978), he discusses the common algorithms of the time, and how they attempt to overcome issues caused by false positive and true negative data points. Blackman (Multiple Hypothesis Tracking for Multiple Target Tracking, 2004) provides an overview of multiple hypothesis tracking (MHT), a classic and perhaps most-widely-used MTT algorithm, which is extended by many of the machine-vision algorithms described in the following section. Blom and Bloem (Interacting Multiple Model Joint Probabilistic Data Association avoiding track coalescence, 2002) give a good overview of another classic data association algorithm, the joint probabilistic data association filter (JPDA). Pulford (Taxonomy of multiple target tracking methods, 2005), aiming to provide a modern update of Bar-Shalom’s paper, gives a thorough overview of ‘classic’ and ‘modern’ data association algorithms (and gives the details of 35 different MTT algorithms). Hue et al. (Sequential Monte Carlo Methods for Multiple Target Tracking and Data Fusion, 2002) introduce particle filters for MTT. Oh et al. (Markov Chain Monte Carlo Data Association for General Multiple-Target Tracking Problems, 2004) uses an MCMC approach to perform general data association. Fox et. al (Nonparametric Bayesian Methods

for Large Scale Multi-Target Tracking, 2006) uses a Dirichlet Process prior to estimate the number of targets in an MTT scheme.

Machine Vision Related MTT

Papers on this topic often involve algorithms that have a target detection phase (where the centroids of potential targets are collected, often by looking for specific target features) and a data association phase (where the centroids are sequenced in order to carry out tracking).

Qu et al. (Distributed Bayesian Multiple-Target Tracking in Crowded Environments Using Multiple Collaborative Cameras, 2007) term machine vision related MTT ‘visual multiple-target tracking’, and provide a good overview of the topic (the following mirrors this overview). MacCormick and Blake (A Probabilistic Exclusion Principle for Tracking Multiple Objects, 2000) focus on distinguishing between foreground and background objects and propose a probabilistic exclusion principle to help the problem where objects occlude each other (ie where one object overlaps and blocks another). Isard and MacCormick (BramBLE: A Bayesian Multiple-Blob Tracker, 2001) present a people-tracking algorithm that deduces foreground and background objects probabilistically, without the typical strategy of selecting foreground objects prior to MTT; the authors also talk about ways in which their techniques allow for tracking variable numbers of objects and objects that enter and exit the videos. Zhao and Nevatia (Tracking Multiple Humans in Crowded Environment, 2004) make use of human shape models to better the MTT accuracy when tracking crowds of people. Tao et al. (A Sampling Algorithm for Tracking Multiple Objects) use sampling to perform Bayesian inference and focus on handling occlusion, addition, and deletion of targets. Rasmussen et al. (Probabilistic Data Association Methods for Tracking Complex Visual Objects, 2001) apply a version of the joint probabilistic data association filter to certain shape and texture features they are able to extract from video targets. Khan et al. (An MCMC-based Particle Filter for Tracking Multiple Interacting Targets, 2004), present an MCMC-based particle filter that also introduces a Markov random field to model the interactions of targets when they are within certain ranges of other targets (applied to videos of ants, arguing that an ant’s behavior is highly influenced by nearby ants). Smith et al. (Using Particles to Track Varying Numbers of Interacting People, 2005) present an MCMC particle filter to increase efficiency in the multiple hypothesis tracker, and also incorporate target features such as color into their model. McKenna et al. (Tracking Groups of People, 2000) use color and gradient information to perform background subtraction on videos of people; their model then has a hierarchical structure where groups of occluding people are first tracked, and then later separated to track each person separately, using color information. Yamamoto et al. (Realtime Multiple Object Tracking Based on Optical Flows, 1995) uses optical flows (calculated based on a so called ‘generalized gradient model’ of an image) to carry out multiple target tracking.

Generalized Polya Urn Dirichlet Process Mixture Model

Caron et al. (Generalized Polya Urn for Time-varying Dirichlet Process Mixtures, 2007) introduce this model. Gasthaus et al. (Dependent Dirichlet Process Spike Sorting, 2009) apply this model to the spike-sorting problem.