2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

A Tracker for Multiple Dynamic Targets Using Multiple Sensors

Adam Feldman, Summer Adams, Maria Hybinette and Tucker Balch

Abstract—**We describe a clustering-based algorithm for tracking a dynamically varying number of targets observed by multiple sensors. The algorithm relies on discrete target detections (e.g., laser ``hits'') and a simple model of the targets to be tracked (e.g. a human is modeled in 2-D as a circle). The algorithm is evaluated in the context of a 4 versus 4 basketball game (8 targets) using 4 SICK LMS291 laser scanners as input. Our evaluations show that the sensor system correctly reports the number of targets roughly 99% of the time. We also demonstrate use of the tracker with two video datasets of multiple changing numbers of ants and fish, respectively.**

I. Introduction

Tracking humans, robots and animals is becoming increasingly important in order to analyze and understand behavior in domains ranging from biology to computer vision and robotics research. Our focus is to automatically track the number and locations of multiple animals, objects or people (hereafter, "targets") in a dynamic environment either indoors or outdoors in uncertain lighting conditions as they move rapidly through the environment over time.

To address this task we focus on multiple laser range finders. Laser range finders have been used in other research, for example in robot soccer to track a single ball [1]. Laser range finders are exceptionally reliable because they are not particularly susceptible to 'false positives' and 'false negatives'. In other words, a detected object (laser hit) almost certainly corresponds to an actual object in the world, and the lack of a hit reliably indicates that there is no corresponding object in the world. Further, laser range finders have very high spatial accuracy; the laser hit corresponds to the object's actual location, within 1.5 cm (according to the manufacturer).

We use multiple sensors (lasers) placed at different viewpoints to address occlusion and to cover a large area at a greater density. For example, a single laser placed at a location provides only one view. Hence, it cannot detect hidden objects - objects that are behind another object when viewed from the viewpoint of the laser. Multiple lasers overcome this problem by providing additional viewpoints that may discover these 'hidden' objects. Another advantage is that multiple lasers provide more laser hits, consequently covering an area with a greater density. Finally, supporting a number of lasers makes tracking more robust with respect to equipment failure.

In this paper we introduce an approach that accurately computes the tracks of a varying number of moving objects as they move through time. Our approach uses multiple laser range finders that record objects' positions. It removes 'uninteresting objects' and accounts for individual targets in close proximity. The ultimate result is a series of snapshots of positions of objects as time unfolds. Individual objects in these snapshots (or "detections") are strung together, creating tracks of an individual object's location over time.

To illustrate our approach, we track players on a basketball court, and evaluate it's performance across different metrics. Although the focus is on laser-based data, we also test the approach with two video datasets.

II. Related Work

Traditionally, tracking of humans has focused on using vision. Machine vision is a well-studied problem, with a wide variety of approaches, such as particle filter based tracking 0 and color segmentation [2]. Research has also been done specific to tracking indoor team sports [3]. Some of these techniques have trouble dealing with situations in which the number of tracked targets is unknown or changing over time (for example, when a target temporarily leaves the field of view of the sensors). Vision trackers are also typically very computationally intensive. This limits their ability to function in real-time. Finally, challenges for vision-based solutions also include the difficulty of dealing with changing lighting conditions and distinguishing foreground from background.

Alternatively, lasers provide accurate information about the environment, requiring less computation than computer vision. Laser data is more accurate than other range sensors, such as ultrasound and infrared. One example of earlier research on laser-based tracking [4] uses occupancy grids and linear extrapolation of occupancy maps to estimate trajectories. Another system [5] uses multiple lasers in the environment and advanced trajectory estimation algorithms to perform real-time tracking.

The objective of many of these tracking systems is to be able to achieve the ultimate goal of identifying people or

Manuscript received September 15, 2006. This work was supported in part by the Center for Behavioral Neuroscience venture grant and NSF CAREER Award IIS-0219850.

Adam Feldman is with the Georgia Institute of Technology, Atlanta, GA 30332 USA (phone: 404-385-6602; e-mail: storm@cc.gatech.edu).
Summer Adams is with the Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: summer@cc.gatech.edu).
Maria Hybinette is with the University of Georgia, Athens, GA 30602 USA (e-mail: maria@cs.uga.edu).
Tucker Balch is with the Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: tucker.balch@cc.gatech.edu).

1-4244-0602-1/07/$20.00 ©2007 IEEE.

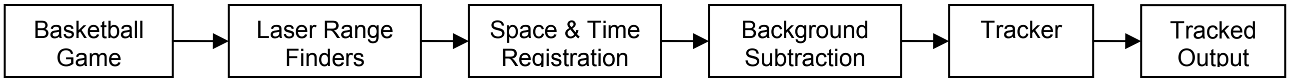

Fig. 1. Overview of our system.

other targets and their activities. One system [6] built upon their earlier work [5] to detect anomalous interactions between people. Likewise, our system will play a part in the ultimate goal of uniquely identifying tracks of individuals.

III. Approach

Our system accurately tracks a variable number of targets using multiple sensors. To track targets efficiently we have developed a clustering algorithm that exploits previously estimated positions while clustering detection-based data. We process the data in several phases, namely – data collection, registering, background subtraction and tracking.

Fig. 1 provides an overview, illustrating the flow of data from one phase to the next. First in the data collection phase, we use four SICK LMS-291 laser range finders to record the targets in the area of interest. In the registering phase the data is passed to several modules, which register the data in space and time to create a single global ‘picture’ of all laser data. The data is then run through the background subtraction module to remove extraneous data points not related to the targets. Finally, in the tracking phase the processed and formatted data is passed to the tracker, which computes tracks representing the location of each target using our clustering algorithm.

As a detection-based tracker, video data can also be used. For such data, each pixel is determined to be either a detection or not, using an algorithm such as adaptive background subtraction or color segmentation. This data can then be tracked, as described above.

The goal of the tracker is twofold. First, it must determine which clusters of data points in a given frame correspond to one of the target to be tracked. Here, a cluster refers to a grouping of data points in one frame, which together represent a target to be tracked; the target is considered to be located at the cluster’s geometric center. Second, it must recognize these clusters from frame to frame in order to build tracks representing the same agent over time. Our tracker accomplishes these goals in parallel, using the information about the clusters found in one frame to help find the corresponding cluster in the next.

The tracker has three main elements. The first uses the location of the clusters in previous frame(s) and iterates on a given frame to find all of the clusters in the current frame. The second builds up tracks across multiple frames, also eliminating clusters that are no longer present in the data. Finally, the third outputs the track information, filtering tracks which do not meet minimum requires of being a track.

To test this tracking system’s accuracy, we began with a data set consisting of one basketball game about 20 minutes in length with 8 people (4 players per team). Accuracy was measured in two ways; the number of tracks found in each frame and the correct handling of multi-track collisions. Also, the system’s ability to perform on two video datasets, multiple fish in an aquarium and ants in a potential nest, is examined. Quantitative results are presented for the ant set.

IV. Results

With the laser-based basketball data, the best accuracy the tracker achieves when detecting the correct number of tracks is 99.17%.

We observed favorable results regarding our collision experiment as well. The human observer detected 37 instances of collisions in which the tracker should have split the tracks. With the appropriate parameter setting, the tracker achieved 100% accuracy. It should be noted that higher accuracy results for collision detection does reduce the average length of tracks.

On the ant data, the track detection accuracy was 96.75%. This result compares favorably to other vision-based tracking. For example, [7] achieved an 89% accuracy examining the same metric with a different vision-based tracker applied to a similar ant dataset. Again, a split accuracy of 100% could be achieved, at the cost of shorter average track length.

Acknowledgment

We would like to thank Professor John Bartholdi of the Georgia Institute of Technology and Professor Kim Wallen of Emory University.

References

[1] S. Sclaroff and R. Rosales, “Improved tracking of multiple humans with trajectory prediction and occlusion modeling,” in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, Workshop on the interpretation of Visual Motion*, Santa Barbara, USA, 1998.

[2] J. Bruce, T. Balch, and M. Veloso, “Fast and inexpensive color image segmentation for interactive robots,” in *Proc. of 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-2000)*, Takamatsu , Japan, October 2000.

[3] J. Perš, S. Kovačič, “Computer vision system for tracking players in sports games,” in *Proc. of the First Int'l Workshop on Image and Signal Processing and Analysis (IWISPA 2000)*, Pula, Croatia, June 2000, pp. 81-86.

[4] E. Prassler, J. Scholz, A. Elfes, “Tracking people in a railway station during rush-hour,” in *Proc. of the Computer Vision Systems, First International Conference (ICVS-1999)*, Las Palmas, Spain, January 1999, pp. 162-179.

[5] A. Fod, A. Howard, and M. J. Mataric, “A laser-based people tracker,” in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA-2002)*, Washington, DC, USA, May 2002.

[6] A. Panangadan, M. J. Mataric, and G. S. Sukhatme, “Detecting anomalous human interactions using laser range-finders,” in *Proc.of 2004 IEEE/RSJInternational Conference on Intelligent Robots and Systems (IROS-2004)*, September 2004, pp. 2136-2141.

[7] T. Balch, Z. Khan, M. Veloso, “Automatically tracking and analyzing the behavior of live insect colonies,” in *Proc. of Autonomous Agents*, 2001.