Machine-Man Collaboration for Counting Pedestrians, Bicycles, and Cars in Video

Ben Schilling, Gagan Mirchandani, School of Engineering, University of Vermont

Introduction

Most existing automated systems are not well suited to the task of counting pedestrians in outdoor environments, and little is known about their effectiveness and accuracy [1]. Currently, the most accurate method for counting pedestrians in vision-based systems is manual; however this is a slow and tedious process. Our program automates the counting process of over 122 hours of video in over 112,000 files. With efficient error checking methods and manual post-processing, estimates to within 79 and 80 % of the ground truth are obtained for pedestrians and bicyclists respectively. The results improve further to 100 and 88 % with the excellent camera placement used in TRC23.

Background and Problem

Pedestrians are a vitalizing force in the activity of traffic-generating centers of a city [4]. Interest in pedestrian traffic flows and activities has been growing since the 1960s. In the past the only method available to attain this information was through the pencil and paper approach, standing at intersections and manually counting and recording. New methods have emerged including buried pressure pads, inductive loops, pneumatic tubes, infrared sensors, and computer vision [2]. While each method has its own strengths and weaknesses, they are preferable to manual counts. At the TRC, video records for 24 hour periods at desired locations have been obtained and scanned off-line to generate a manual count. We automate this task using a combination of machine-man techniques. While computer vision algorithms can be quite powerful, real-life errors due to occlusion, drift, illumination variation, and background changes make the resulting count less than robust. Even the most sophisticated video algorithms for pedestrian detection miss 20-30 % in the best of conditions with performance degrading rapidly with small scales and occlusion [3]. In the spirit of emerging machine-man collaboration algorithms, we apply very efficient manual post-processing to our algorithm to considerably increase count accuracy.

Video Procedure

Although a motion activated camera was used to collect data, there were still many hours of video containing no desired objects. There were also many scenes where undesirable motion occurred in the background or foreground such as branches blowing and shadows moving. To mitigate these noise factors several simple filters were used:

- **Outlier pixel filter**: removes excess noise pixels greater than 2 standard deviations away from the object's mean (x,y) coordinates
- **Small object filter**: removes detected objects that are below the threshold size variable
- **Immobile object filter**: removes detected objects that are below the threshold movement distance variable
- **Short segment filter**: removes detected objects that are below the threshold number of frames

Program Implementation

The Matlab program can be broken down into 4 main sections:

1. **Pre-Processing**
2. **Video Analysis**
3. **Training & Classification**
4. **Post-Processing**

**Pre-Processing**
- Set simulation variables such as minimum movement distance, minimum object size, and minimum segment length
- Crop the desired video to remove excess pixels not in the region of interest

**Video Analysis**
- Load the desired videos from folder
- Detect motion of an object between frames in each video
- Calculate and save feature values for every object in each video including size, speed, aspect ratio, and density

**Training & Classifying**
- Train the classifier with videos containing prime examples of each different object
- Classify the detected objects from every video based upon the trained feature values for each object type

**Post-Processing**
- Manually check the classified objects by viewing sequentially saved images of pedestrians and bicyclists

Strengths and Weaknesses

**Strengths**
- Very user-friendly program with a simple GUI that allows the user to perform all functions with the click of a few buttons
- Capable of analyzing many hours or even days of video at multiple locations
- Single Matlab file as opposed to several separate files needed to be organized a specific way

**Weaknesses**
- Difficulty with nighttime scenes
- Difficulty with tightly packed groups of objects or occluded objects

Program Results

Conclusions

The program performs quite well in detecting, classifying, and counting pedestrians and bicyclists. In nearly all cases the program over-counts pedestrians and bicyclists. This is primarily due to the presence of cars that constitute 95 % of the objects detected.

In cases where over-counting occurred the usual cause was in nighttime videos where cars were often mistaken for bicyclists and double counting where pedestrians would stop or change directions randomly. Under-counting occurred in the presence of occlusion where a car would block vision of a pedestrian or when groups of bicyclists were counted as only one object.

By performing the very efficient post-processing step of viewing the saved images of each classified bicyclist and pedestrian, the results improve to the point where the counts are very reliable. We hope to further improve the machine-man collaboration by providing feedback to the program after completion of the manual error check.

References