ICT ILLINOIS CENTER FOR TRANSPORTATION

LEVERAGING TRAFFIC AND SURVEILLANCE VIDEO CAMERAS FOR URBAN TRAFFIC

Prepared By
Jakob Eriksson
University of Illinois at Chicago

Research Report No. FHWA-ICT-14-022

A report of the findings of
ICT-R27-131
Leveraging Traffic and Surveillance Video Cameras for Urban Traffic

Illinois Center for Transportation

December 2014

Technical Report Documentation Page

1. Report No. FHWA-ICT-14-022	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Leveraging Traffic and Surveillance Video Cameras for Urban Traffic		5. Report Date December 2014
		6. Performing Organization Code
7. Author(s) Jakob Eriksson		8. Performing Organization Report No. ICT-14-024 UILU-ENG-2014-2026
9. Performing Organization Name and Address Department of Computer Science College of Engineering University of Illinois at Chicago Room 1120 SEO (M/C 152) Chicago, IL 60607		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. R27-131
12. Sponsoring Agency Name and Address Bureau of Materials and Physical Research Illinois Department of Transportation 126 East Ash St. Springfield, IL 62704		13. Type of Report and Period Covered
		14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The objective of this project was to investigate the use of existing video resources, such as traffic cameras, police cameras, red light cameras, and security cameras for the long-term, real-time collection of traffic statistics. An additional objective was to gather similar statistics for pedestrians and bicyclists. Throughout the course of the project, we investigated several methods for tracking vehicles under challenging conditions. The initial plan called for tracking based on optical flow. However, it was found that current optical flow–estimating algorithms are not well suited to low-quality video—hence, developing optical flow methods for low-quality video has been one aspect of this project. The method eventually used combines basic optical flow tracking with a learning detector for each tracked object— that is, the object is tracked both by its apparent movement and by its appearance should it temporarily disappear from or be obscured in the frame. We have produced a prototype software that allows the user to specify the vehicle trajectories of interest by drawing their shapes superimposed on a video frame. The software then tracks each vehicle as it travels through the frame, matches the vehicle's movements to the most closely matching trajectory, and increases the vehicle count for that trajectory. In terms of pedestrian and bicycle counting, the system is capable of tracking these "objects" as well, though at present it is not capable of distinguishing between the three classes automatically. Continuing research by the principal investigator under a different grant will establish this capability as well.

17. Key Words bicycles, optical flow tracking, pedestrians, traffic cameras, traffic statistics, video surveillance, urban traffic	18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 12	22. Price

Form DOT F 1700.7 (8-72)   Reproduction of completed page authorized

ACKNOWLEDGMENT AND DISCLAIMER

EXECUTIVE SUMMARY

The objective of this project was to investigate the use of existing video resources, such as traffic cameras, police cameras, red light cameras, and security cameras for the long-term, real-time collection of traffic statistics. An additional objective was to gather similar statistics for pedestrians and bicyclists.

Use of existing video resources includes a number of challenges, such as poor video quality, challenging perspective, occlusion between vehicles or stationary objects and vehicles, and lighting conditions. Due to the long-term nature of permanent installations, traffic counting should require minimal user input but facilitate easy validation of counts when so desired.

Throughout the course of the project, we investigated several methods for tracking vehicles under challenging conditions. The initial plan called for tracking based on optical flow. However, it was found that current optical flow–estimating algorithms are not well suited to low-quality video—hence, developing optical flow methods for low-quality video has been one aspect of this project.

The method eventually used combines basic optical flow tracking with a learning detector for each tracked object—that is, the object is tracked both by its apparent movement and by its appearance should it temporarily disappear from or be obscured in the frame.

We have produced a prototype software that allows the user to specify the vehicle trajectories of interest by drawing their shapes superimposed on a video frame. The software then tracks each vehicle as it travels through the frame, matches the vehicle's movements to the most closely matching trajectory, and increases the vehicle count for that trajectory.

In terms of pedestrian and bicycle counting, the system is capable of tracking these "objects" as well, though at present it is not capable of distinguishing between the three classes automatically. Continuing research by the principal investigator under a different grant will establish this capability as well.

CONTENTS

CHAPTER 1   COUNTING VEHICLES IN OPPORTUNISTIC VIDEO

Traffic sensing and monitoring has a wide variety of applications, with users ranging from the professional planners in the department of transportation, to soccer moms planning a picnic. On many controlled-access highways, the majority of sensing is done with magnetic loops embedded in the road surface. However, for arterial roads, that is typically not available. Chicago, having one of the world's largest networks of traffic and surveillance cameras, is uniquely positioned to take advantage of recent advances in computer vision.

Performing traffic counts through video analysis is not a new idea, and commercial solutions exist for performing such traffic counts today. However, these solutions rely on custom cameras, carefully mounted in ideal observing locations for traffic counts. We propose to extend the scope of these methods to existing cameras, typically mounted for the purpose of manual traffic monitoring or law enforcement. Challenges include (1) variations in camera angle, focal length, and occlusion by obstructions; (2) intermittent changes in observation direction and focal length due to actuated cameras; and (3) variations in video quality: frame rate, resolution, and compression artifacts as well as transmission-induced stutters, glitches, and other abnormalities.

This project's aim was to overcome these challenges by developing novel techniques for object tracking (vehicular, pedestrian, and bicyclist), optical flow, road and lane detection, occluding object and foreground/background detection in opportunistic video, or video from existing camera resources.

1.1 PROBLEM SETTING

Figure 1 shows a frame from a typical road-facing camera operated by the Illinois Department of Transportation (IDOT). This frame also highlights some of the challenges in producing vehicle counts from existing video sources. The perspective is such that half of the frame is dedicated the ground close to the camera, and a third of the remainder shows the sky. The remaining, and central, 100 × 300 pixels hold the primary information of the image. Here, only one vehicle is clearly discernible as a car. The remainder can be construed to be vehicles based on their size, overall shape and location on a road surface. Figure 2 shows a zoomed in version of the white vehicle facing the camera in the upper left corner of Figure 1 in isolation. Here, even a person would have great difficulty in identifying the object as a vehicle. Looking a little closer at Figure 1, we see several more vehicles, possibly three white vehicles and several dark vehicles facing away from the camera, and perhaps one more vehicle approaching the camera in the distance. The fact that the author of this report cannot reliably judge the number of vehicles present in the scene is witness to the challenge at hand.

The goal of the project is to produce both traffic and turn counts. Here, traffic counts lend themselves to somewhat naïve solutions, focusing on the part of the image where the vehicles are large and clearly visible. This has been the primary focus of prior published work in this area. However, for accurate turn counts, it is necessary to follow vehicles beyond this zone of relative comfort and into the blurry regions farther into the distance, producing as long and accurate a vehicle trajectory as possible, in order to accurately classify the movement as belonging to one turn type or another.

Figure 1. A frame from a typical road-facing camera.

Figure 2. Zoomed in picture of white car in left of Figure 1.

1.2 INITIAL OPTICAL FLOW–BASED APPROACH EVALUATED AND EVENTUALLY USED FOR SECONDARY OBJECTIVES

The project went through several iterations before eventually settling on a vehicle tracking technique for this type of video. Initially, the plan was to develop an object tracker based primarily on optical flow, a computer vision technique that estimates the movement of each pixel in a frame. The thought was that the additional information provided by optical flow would lend robustness to the object tracker.

However, after several months of investigating and comparing various optical flow methods, it was determined that existing methods were inadequate for the task at hand, producing highly inaccurate flow estimates under our unusually challenging conditions. An example is given in Figures 3 and 4. Here, while the optical flow algorithm in use does detect movement at a crude level, the accuracy is poor enough to be useless for vehicle tracking purposes. The desired result in Figure 4 is a crisp outline of each moving object, in purple for vehicles moving away from the camera and turquoise for those moving toward the camera.
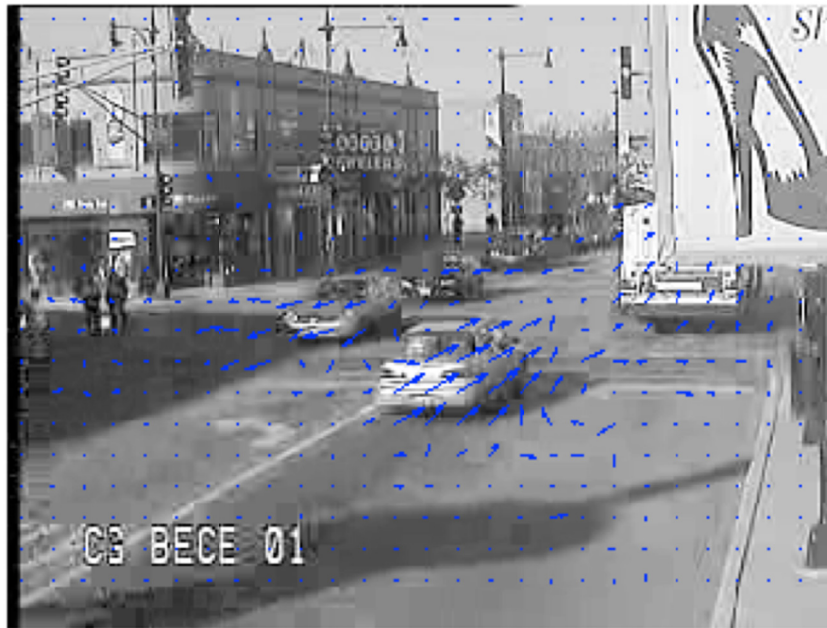

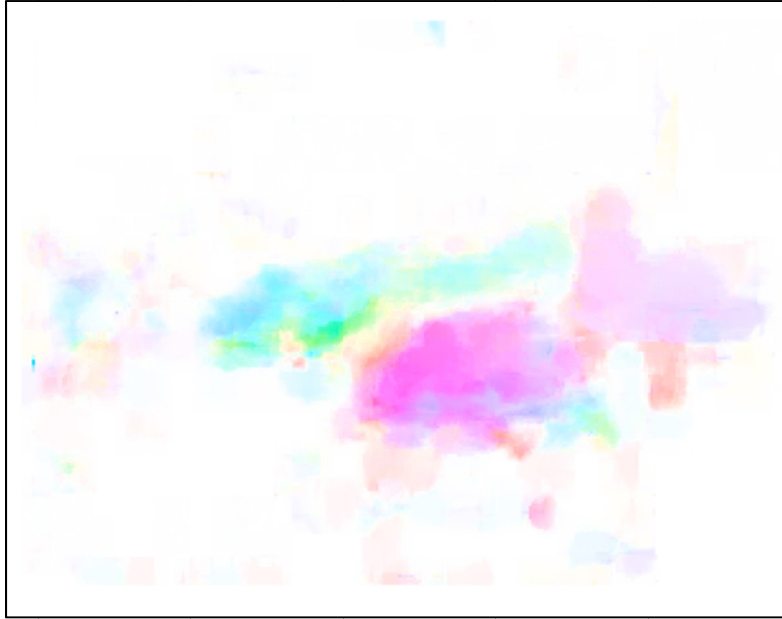Figure 3. Example scene with optical flow annotations (arrows).

Figure 4. Example scene illustrating dense optical flow.
Color indicates direction, and intensity indicates magnitude.

However, for statistical purposes, the relatively low quality of the optical flow produced suffices to generate an accurate picture of the traffic scene. Figures 5 and 6 illustrate a scene from Lakeshore Drive, where a long-term (several minutes) average of the optical flow of the scene is shown in Figure 6. Here, the directionality and extent of the road lanes are accurately determined. Figure 7 shows the mean magnitude of flow, which indicates average driving speeds and could be used to detect speed outliers (speeders).

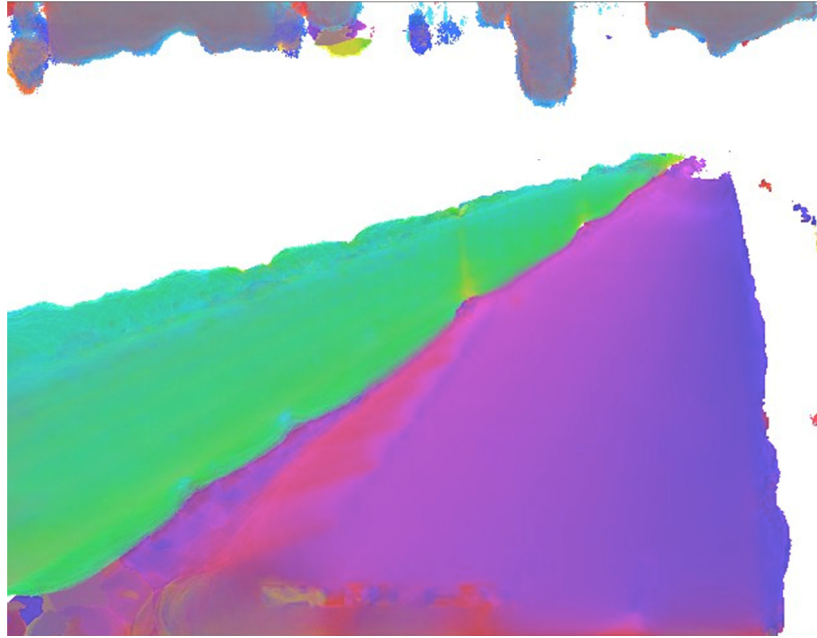Figure 5. Scene from Lakeshore Drive.

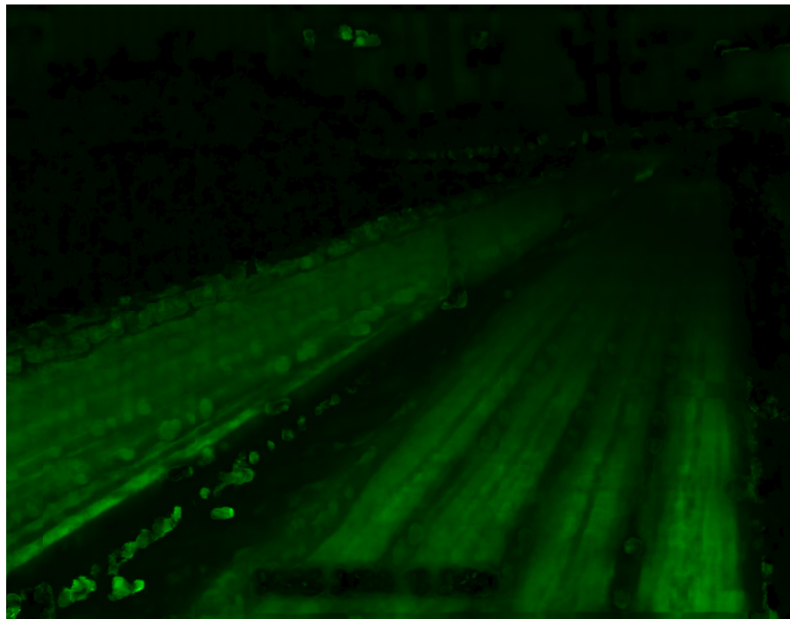Figure 6. Long-term average flow for Lakeshore Drive scene.

Figure 7. Mean flow magnitude for Lakeshore Drive scene. This is an indication of the typical speed of travel in different parts of the image.

Similar results from a more challenging scene are shown in Figure 8, indicating the generality of this approach. Using this method, we demonstrated the ability to automatically detect lanes of travel, which could potentially be used in further automating the final vehicle counting system.
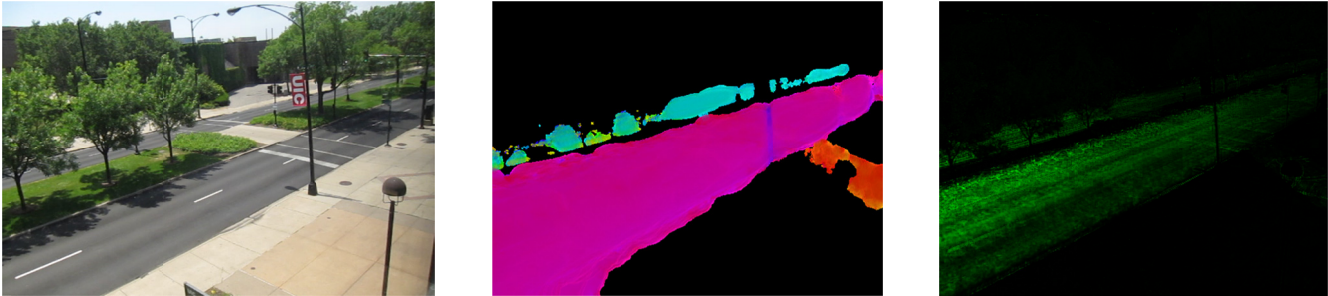
Figure 8. Taylor street scene with challenging perspective, showing mean flow and magnitude plots.

1.3 THE HYBRID TRACKING/DETECTION APPROACH SETTLED UPON

The approach that we finally settled on is a hybrid optical flow and detection method, with several enhancements for the problem at hand. Here, a very basic optical flow method (the Lukas–Kanade method) is used to provide short-term object tracking. Let us call this "flow-based" tracking. Meanwhile, a learning detector is used to "rediscover" the object once this weak tracker loses it. The detector learns the appearance of the object while the flow-based tracker has a firm grasp on its location, allowing the detector to adapt to changes in scale, rotation, and lighting. More details on this are provided in Chapter 2.

1.4 OTHER METHODS INVESTIGATED

Beyond object tracking and counting, which constitute the primary parts of the proposed system, we investigated several other related components and have made some worthwhile progress in these areas. We briefly describe them below.

Because we found existing optical flow methods inadequate, we have spent, and continue spending, time and effort on developing a more robust optical flow–estimation method that explicitly accounts for occlusion and is able to accommodate imagery of low resolution and poor quality. The basic approach we have been pursuing is simultaneous motion and image segmentation.

The video is divided into segments based on appearance, and movement is computed for entire segments together rather than each pixel separately. This approach allows us to produce more accurate estimates due to the diminishing effect of error when averaging over more pixels. However, as a result of projection effects, parts of a segment may appear to move, grow, or shrink at different rates. We are investigating the use of projective geometry to address this issue. We have made good headway but do not have publishable results at this point.

CHAPTER 2   HYBRID TRACKING ALGORITHM DETAILS

Our current tracking system consists of several key components, which we will outline below, together with a brief discussion of the research that led to this design.

2.1 INITIAL DETECTION

As a first step of any tracking system, we must locate and select moving objects to track. Under ideal conditions, this would be done using a detector trained specifically on vehicles, so that only vehicles would be tracked. However in practice, as shown in Figure 2, because of poor image quality, it is often not possible to discern a vehicle from something else, at least not when it first appears, with any reasonable accuracy.

Instead, our design is based on foreground separation. For any given frame, the pixels that deviate substantially from the normal (background) color values are marked as foreground. We investigated several different foreground separation methods before eventually settling on the one we use today—an algorithm called ViBe.

Figure 9 shows the results of one evaluation experiment. Here, the ViBe method clearly identifies the red vehicle as foreground. Moreover, both the pedestrians and the vehicle entering the scene at the top right, behind the trees, are identified. The strong showing in this scene comes at a cost of additional false positives.
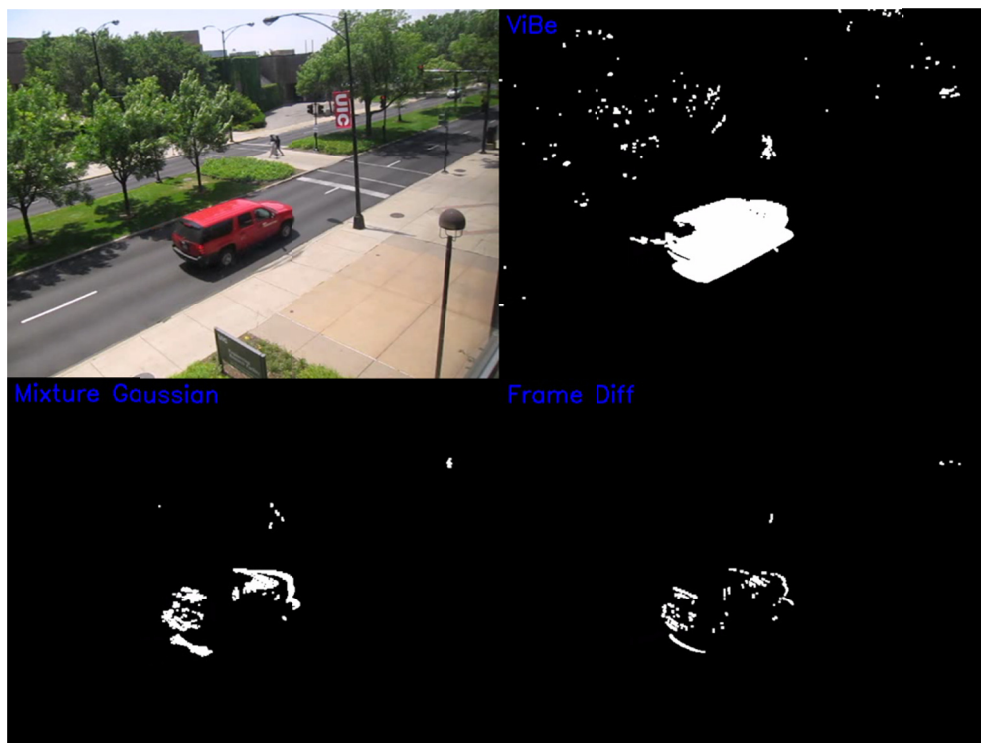

Figure 9. Foreground masks for Taylor scene, using three different separation methods.

After denoising the foreground mask of each video frame, foreground pixels are grouped together into contiguous objects. Any object that is not already being tracked is then forwarded to the tracking algorithm.

2.2 TRACKING

Once initialized with the boundary box and foreground mask of an object to track, our tracking algorithm, based on Tracking-Learning-Detection (TLD), starts to follow the object through the scene. Two different algorithms compete against each other during regular operation—one is an optical flow–based tracker, which simply updates the expected location of the object based on the mean optical flow observed within the bounding box of the object. This works well in the short term, but it incurs significant drift over time, leading to the eventual loss of the object. Flow-based tracking is also incapable of dealing with occlusion: once an object, or even part of an object, is lost from view, optical flow–based tracking cannot, by design, recover the object.

The competing method is a detector. The detector is initially trained only with the initialization image provided by the foreground separation. However, as long as the flow-based tracker maintains a confident estimate of the vehicle's location, new images are continuously fed to the detector for training. The detector constantly scans the image for objects resembling the one it was trained on. Thus, when the flow-based tracker loses an object, the detector will recover it after it reappears in the scene. To accommodate multiple moving objects in the scene simultaneously, the system supports multiple independent trackers but shares resources such as computed flows between trackers to maintain efficiency. Figure 10 illustrates simultaneous tracking of two vehicles in a typical scene. Here, the red boxes indicate the estimated location of the vehicles, and the red lines describe their estimated trajectory through the scene up to the present frame. The small images in the upper right show the respective appearances of two vehicles that the detector is looking for throughout the scene, and the green dots on top of the vehicles illustrate features currently used to track the vehicle from frame to frame.
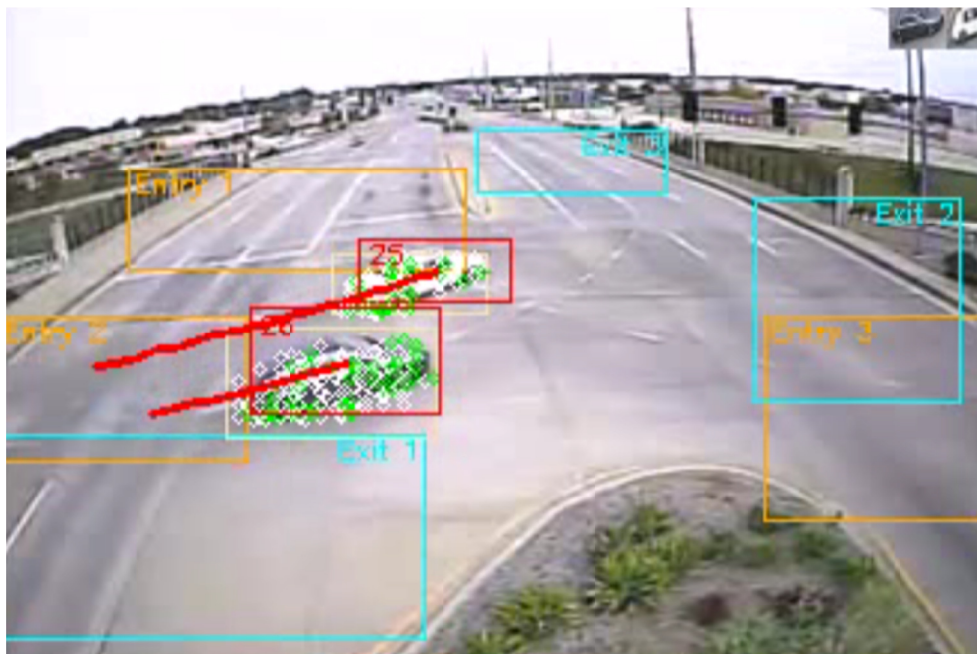

Figure 10. Tracking screen for an example intersection scene.
Two vehicles, labeled 25 and 26, are currently being tracked.

2.3 COUNTING

The vehicle tracker produces a large number of trajectories in a video recording. To produce a final vehicle count, our algorithm accepts a number of pre-specified trajectory templates (Figure 11). These trajectory templates indicate movements of interest and provide a rough outline of what the movement looks like in the scene. The bright green lines in Figure 11 indicate tracking results, and the remaining lines show template trajectories. After a vehicle leaves the scene, its movement is credited to the template that best matches the actual tracking result.
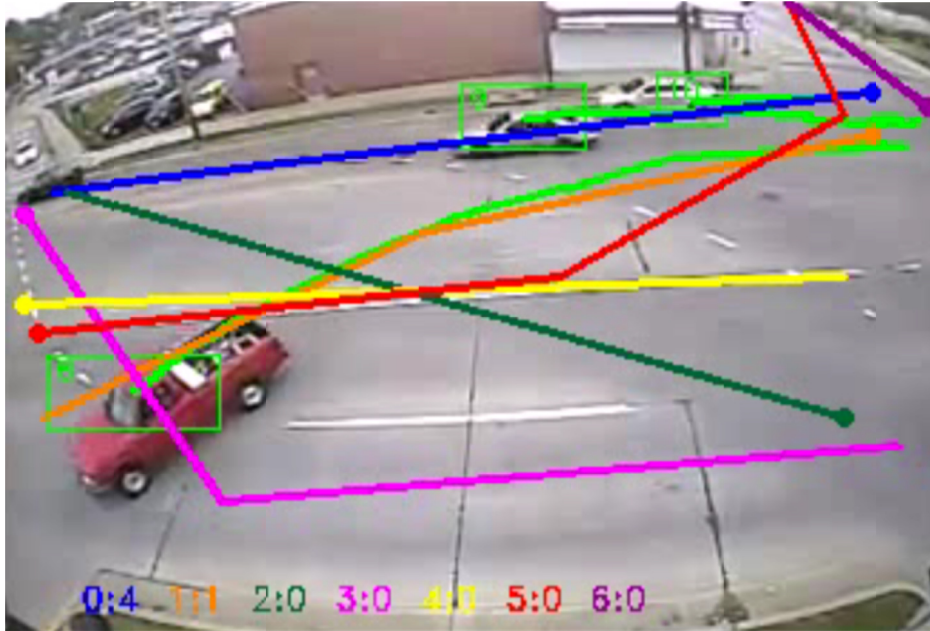

Figure 11. Vehicle counting from provided templates and computed vehicle trajectories.

In Chapter 3, we briefly discuss the software prototype produced as part of this project. The software currently runs under Linux and is capable of producing counts from a wide range of videos. However, limitations include lighting and atmospheric conditions and severe occlusion.

CHAPTER 3   SOFTWARE PROTOTYPE

Our software prototype executes the foreground separation, tracking and counting tasks as described previously. It also provides a simple user interface in which the user may draw template trajectories, load and store such template drawings, and perform vehicle counts on individual video files. A screen shot of this software is shown in Figure 12.

In addition to turn counts, the software is also capable of producing a validation video, as shown in Figure 13. This feature is useful for increasing user confidence in the output of the software, as well as for identifying remaining weaknesses in the current design. The validation video clearly marks each vehicle by its corresponding template and shows the current count at all times.
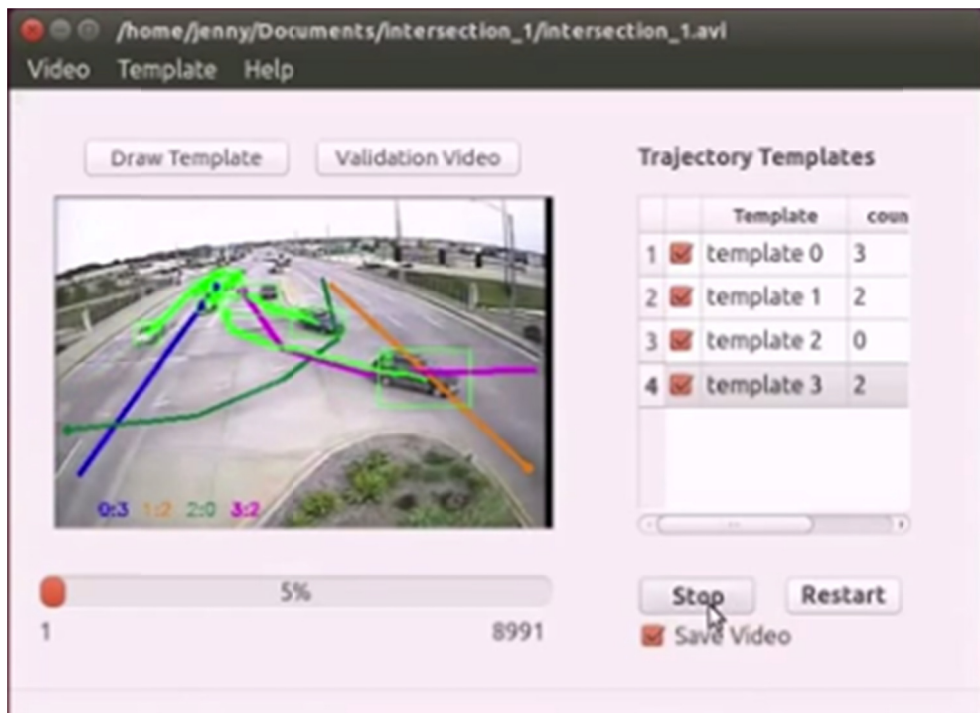

Figure 12. Primary user interface from vehicle counting software prototype.
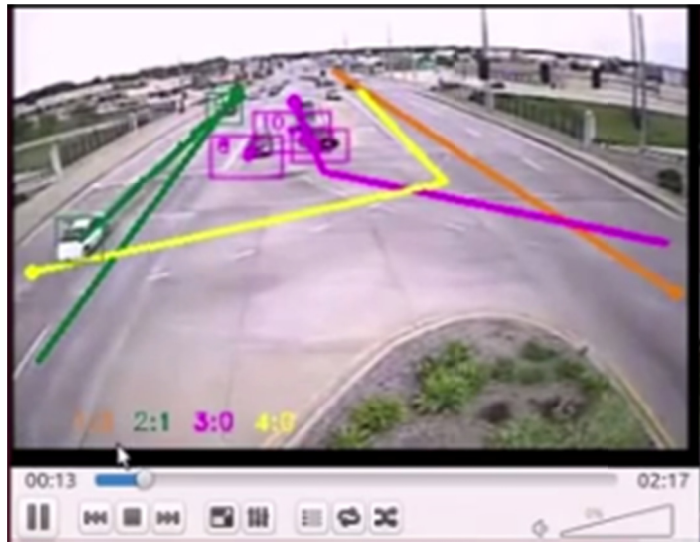
Figure 13. Validation video produced by prototype software.

CHAPTER 4   CONCLUSIONS

While this project originally set out to do more than what was finally accomplished, we have made significant progress in the intended direction. We are able to produce usable turn counts from a wide variety of cameras and perspectives that are robust to both poor resolution and poor overall image quality. However, more research is needed to further strengthen the system in the face of severe occlusion, both in terms of stationary objects obstructing the view and in terms of vehicles obscuring each other. More work (and data) is also needed to extend this system to accurately classifying vehicle types as well as vehicles versus pedestrians and bicyclists. Here again, image quality constraints make contemporary approaches to these problems impractical.

Overall, the project results have been both above and below expectations. We have learned a great deal and hope and plan to capitalize on these lessons in ongoing and upcoming projects in this exciting area of research.

For more up-to-date information, and video recordings of the system in action, please see our project web page at http://www.cs.uic.edu/Bits/VehicleCounting.