A Survey on Video and Vision in Intelligent Transportation Systems

3D Vision and Video Computing
City College
Instructor: Professor Z. Zhu
Spring 2003

Ilknur Icke

May 29, 2003

Introduction

The existing surface transportation systems of many countries are facing the challenge of increasing demands.For instance, United States has approximately 3.7 million miles of roads and 503 public transit systems, accommodating 4 trillion passenger miles and 3 trillion ton miles of freight per year. Demands on the system are growing rapidly, as high as 30% over the next 10 years. In order to prevent congestion from getting worse, the United States would have to increase the capacity of its transportation systems by the same amount. This would require adding new lanes. At the present this is not possible.[2]

Intelligent Transportion Systems are alternative methods to increasing the capacity by adding new lanes. They aim to increase the efficiency of the existing transportation system. In 1991, the Federal Intelligent Transportation Systems(ITS) program has been founded. The goals were as follows[2]:

- Enhance public safety
- Reduce congestion
- Improved access to travel and transit information
- Generate cost savings to motor carriers,transit operators,toll authorities, and government agencies
- Reduce environmental impacts

Vehicle Detection and surveilance technologies are useful in various areas, such as speed monitoring, vehicle presence detection and counting, vehicle classfication. The sensors used in such technologies can be broken into two groups: Intrusive Sensors and Non-intrusive Sensors.

Intrusive sensors are installed under the road surface, or under the pavement surface. Most common sensors of this type are inductive loops, magnetometers, microloop probes, pneumatic road tubes, piezoelectric cables. Their working principles are well studied and these techniques are well established since they have been in use for a long time.Installation of this type of sensors require disruption of the traffic.

Non-intrusive sensors provide minimal disruption of the traffic while being able to work as accurate as the inductive loop detectors which are the most generally used intrusive sensors. Non-intrusive sensors are installed above the ground. Most common sensors of this type are video image processors,microwave radar, laser radar, passive infrared, ultrasonic, passive acoustic array.

The following figure summarizes the comparison of the sensor technologies used in Intelligent Transportation Systems [2]:

Technology	Strengths	Weaknesses
Inductive Loop	• Flexible design to satisfy large variety of applications. • Mature, well understood technology. • Provides basic traffic parameters • (e.g., volume, presence, occupancy, • speed, headway, and gap). • High frequency excitation models provide classification data.	• Installation requires pavement cut. • Decreases pavement life. • Installation and maintenance require lane closure. • Wire loops subject to stresses of traffic and temperature. • Multiple detectors usually required to instrument a location.
Magnetometer (Two-axis fluxgate magnetometer)	• Less susceptible than loops to stresses of traffic. • Some models transmit data over wireless RF link.	• Installation requires pavement cut. • Decreases pavement life. • Installation and maintenance require lane closure. • Some models have small detection zones.
Magnetic (Induction or search coil magnetometer)	• Can be used where loops are not feasible (e.g., bridge decks). • Some models installed under roadway without need for pavement cuts. • Less susceptible than loops to stresses of traffic.	• Installation requires pavement cut or tunneling under roadway. • Cannot detect stopped vehicles.
Microwave Radar	• Generally insensitive to inclement weather. • Direct measurement of speed. • Multiple lane operation available.	• Antenna beamwidth and transmitted waveform must be suitable for the application. • Doppler sensors cannot detect stopped vehicles.
Infrared	• Active sensor transmits multiple beams for accurate measurement of vehicle position, speed, and class. • Multizone passive sensors measure speed. • Multiple lane operation available.	• Operation of active sensor may be affected by fog when visibility is less than »20 ft or blowing snow is present. • Passive sensor may have reduced sensitivity to vehicles in its field of view in rain and fog.
Ultrasonic	• Multiple lane operation available.	• Some environmental conditions such as temperature change and extreme air turbulence can affect performance. Temperature compensation is built into some models. • Large pulse repetition periods may degrade occupancy measurement on freeways with vehicles traveling at moderate to high speeds.

Figure 1: Comparison of the Sensors

Acoustic	• Passive detection. • Insensitive to precipitation. • Multiple lane operation available.	• Cold temperatures have been reported as affecting data accuracy. • Specific models are not recommended with slow moving vehicles in stop and go traffic.
Video Image Processor	• Monitors multiple lanes and multiple zones/lane. • Easy to add and modify detection zones. • Rich array of data available. • Provides wide-area detection when information gathered at one camera location can be linked to another.	• Inclement weather, shadows, vehicle projection into adjacent lanes, occlusion, day-to-night transition, vehicle/road contrast, and water, salt grime, icicles, and cobwebs on camera lens can affect performance. • Requires 50- to 60-ft camera mounting height (in a side-mounting configura-tion) for optimum presence detection and speed measurement. • Some models susceptible to camera motion caused by strong winds. • Generally cost-effective only if many detection zones are required within the field of view of the camera.

Figure 2: Comparison of the Sensors(continued)

Video Processing Techniques for Traffic Applications

Video sensors are becoming increasingly important in ITS, though they have face some challenges like weather conditions, day-to-night transition, occlusion, shadows, unwanted camera motion caused by strong winds. Despite these challenges, video sensors give fast response thus they can be used in real-time applications, installation and maintenance are non-intrusive and easy and they can monitor wide-areas (for example, multiple lanes)

Next two sections, Traffic Monitoring and Automatic Vehicle Guidance are based on the paper by V. Kastrinaki [4]. This is a very recent paper and it refers to numerious amount of systems and gives a comparison of the systems as well as the open issues and future research directions. There are other uses of video processing in ITS, like License Plate Recognition which can be used at toll-booths to catch the toll violators and Red-light Stop Detection to catch the drivers who ran the red light and send them summons. Following two sections talk about these two application areas of video processing in ITS.

Traffic Monitoring

Video Systems for Traffic Monitoring and Automatic Vehicle Guidance as well, have two main tasks :

- Estimate the road geometry
- Detect the vehicles and obstacles

Estimation of the road geometry generally means estimation of the lane or the central line on the road which is also called Automatic Lane Finding(ALF). The fundamental aspects of the ALF approaches are given as follows:

- Lane-Region detection This method assumes stationary camera, and a map of significant even changes are built. This map distinguishes between the active areas in the scene where motion is occuring and non-active areas where there is little or no motion. Once formed this map helps find the lanes on the road.
- Feature-driven approaches The road images are edge detected and the edges are organized into meaningful structures(lanes or lane markings). In general, these techniques suffer from noise effects and irrelevant feature structures also they depend on the feature extraction technique used.
- Model-driven approaches The aim is to matcha deformable template to the image. The pavement edges and lane markings are approximated by circular arcs on a flat-ground plane. More flexable approaches use snakes and splines to model road segments. Model based techniques provide powerful means for the analysis of road edges and markings.However, the difficulty is to choose a proper model that fits the road structure and the high computational complexity.

The second task in hand is the Object Detection. There are a number of approaches to this problem:

- Thresholding
- Multigrid identification of regions of interest
- Edge-based detection(spatial differentiation)
- Space Signature
- Background frame differencing
- Inter-frame differencing
- Time signature
- Feature aggregation and object tracking
- Optical-flow field

- Motion Parallax
- Stereo Vision
- Inverse perspective mapping
- 3D Modeling and forward mapping

Future trends in this field, as it's described in [4] could be :

- Improvements in Image Processing stages, for example, more usage of Morphological operators in both segmentation and detection of the edges.
- More usage of multiresolution techniques
- The ability to fuse different sources of information, for example, multisensory data.
- Systems that can tolerate uncertainty.For example, under rainy conditions illusion patterns might cause misinterpretation of the scenes. Fuzzy-set theory could be used here.

Automatic Vehicle Guidance

Video Processing is important in the field of Automatic Vehicle Guidance. It is used in finding the vehicle's position in the lane and locating the obstacles on the road. Vehicle guidance is a more complex task because the aim is to reconstruct a 3D representation of the environment from the observed 2D images and the environment dynamically changes making it more challenging for the system to adapt. Similar principles described above for Traffic Monitoring are also applicable to this field.Bertozzi gives an overview of such systems in [1]

License Plate Recognition(LPR)

There are a number of commercial products offered for LPR and in other resources these systems may be referred as Automatic Vehicle Identification (AVI),Automatic Number Plate Recognition(ANPR),Car Plate Recognition (CPR),Car Plate Reader (CPR) and Optical Character Recognition (OCR) for Cars. These systems have wide application areas, some of which are the following[6] :

- Parking. A Parking facility can use this technology to let the prepaid customers enter the lot for example airport parking.
- Access Control to secure areas.
- Tolling. At the toll-booths, to reduce the waiting time to pay the toll.
- Border Control.
- Stolen Cars.
- Enforcement. Extracting the license plate number of the speeding cars or red-light runners.
- Traffic Control. Cars can be directed to different lanes according to their license plate numbers.
- Marketing Tool. A list of frequent visitors to a place can be made based on the plate numbers of the cars.
- Travel. LPR systems can be installed at different points on a road, and two images of the same car can be taken and then analyzed to determine the speed of the car for speed limit enforcement.

A typical LPR system consists of the following components[6]:

- Cameras– to take the images of the car(front and rear)
- Illumination–A controlled light source to light up the license plate, in most cases Infra-Red(IR) light source is used.
- Frame Grabber–Hardware Interface between the PC and the camera
- Computer–Normally a PC running windows or Linux.
- Software–The application and the Recognition components
- Hardware–Such as networking components
- Database–Events are recorded onto a database.For example, the recognition results and also the the driver's face image.

According to Setchell[5], License Plate Recognition Systems consist of two sequential processes: Locate the license plate in the image of the car and character recognition(OCR). As he states, even though there are a number of commercial systems, the technical details of these systems are not disclosed by the companies. The few systems that can be found in the literature are based on two approaches in order to locate the license plate[5]:

- Thresholding the image such that the license plate characters are black and background is white. Then the image is searched for black blobs which are of similar dimensions.
- Neural Networks

The Character Recognition stage has been fairly well studied in Computer Vision. There are various techniques:

- Template Matching. This method uses a separate template for each possible input character and the set of valid input characters are known. Current character is compared with all templates in order to find the one that fits the best. There are several matching criteria used in literature: Euclidean Distance,Cross Correlation,Normalized Correlation, City Block distance.
- Feature Based Character Recognition. Significant features from the input character are extracted and then compared to a database of feature descriptors for all valid input characters. The best matching description determines the character. The feature extraction stage can be done in a number of ways :
  - Global Transformations to reduce dimensionality of feature vector. For instance, Fourier, Walsh, Haar, Hadamard, Hough transforms and etc.
  - Statistical Distribution of points
  - Geometrical or Topological Features
- Neural Networks.

Setchell[5] describes a Number-Plate Recognition System in his PhD dissertation.The system runs on a Transputer based vision hardware.

The figure below shows the components of his system.The Finder component locates the plate by first thresholding the image to produce a binary image with black number plate characters and white background. Then the image is scanned from bottom to top. For each black region,its dimensions and aspect ratio are determined and if they match to the predetermined values for valid license plate characters they are considered as candidate license plate characters. When a number of such candidate characters are found in a horizontal row then the system assumes the location of the license plate has been found and this location is passed to the reader.

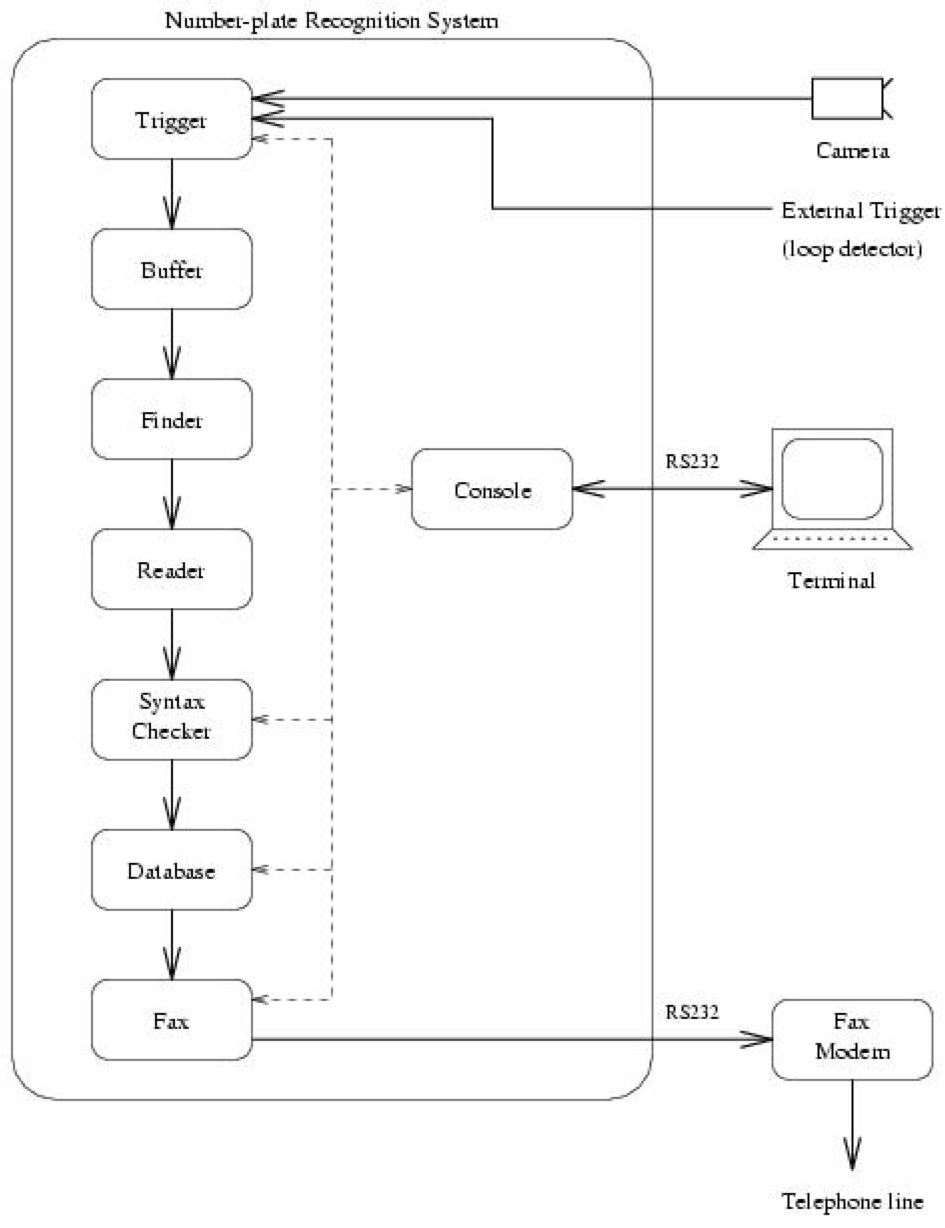

Figure 3: Components of the Number-plate recognition system

If one pass through the entire image fails then the search is repeated with a new threshold value.If after several threshold values are tried and there is no plate found then it's assumed that the image does not contain a license plate. This system requires controlled lighting conditions as it depends on a threshold, also the license plate should be clean and the letters should not appear to be touching each other.

The reader component of the system does the actual recognition. Setchell specifies that tried three approaches:

- Template Matching
- A large neural network having n outputs, one for each of the possible n characters
- n small neural networks, each having one output. One network is trained to recognize each individual character.

As he describes, the single large neural network gives him the best results– around 85% accuracy.

	Template Matching	n small neural networks	Single, large neural network
Plate completely correct	59.5%	69.4%	85.4%
1 character wrong	27.2%	8.6%	5.8%
2 characters wrong	7.1%	3.8%	1.8%
3 characters wrong	1.7%	2.6%	1.0%
4 characters wrong	0.9%	4.4%	1.2%
5 characters wrong	1.0%	1.4%	0.7%
6 characters wrong	0.3%	1.6%	0.5%
7 characters wrong	0.1%	1.5%	0.1%
Plate not read	2.3%	6.8%	3.5%

Figure 4: Reader performance

Red-light Stop Detection

It is estimated that 22% of the traffic accidents in the U.S. are caused by drivers who ran red lights. To fight with this, local authorities have chosen to install red-light cameras to catch the red-light runners. These systems consist of three components: one or more cameras,one or more triggers,a computer.[3]

Generally there are four cameras, one at each corner of the intersection. Some systems use film cameras which require people to change the film periodically, but most recent systems use digital cameras.

Figure 5: Red-light cameras

The purpose of the trigger component is to detect when a car passed a particular point on the road. Generally, induction-loop triggers are used for this purpose.

Some other types of triggers include radar,laser or air-tube sensors.A recent type of trigger mechanism is called a video loop. In such a system, a video image is analyzed in order to see if a car moved through the intersection.If the light is red and a car is detected in the intersection within the video image, the red-light cameras are activated. The advantage of this system is that it's easy to install and non-instrusive.

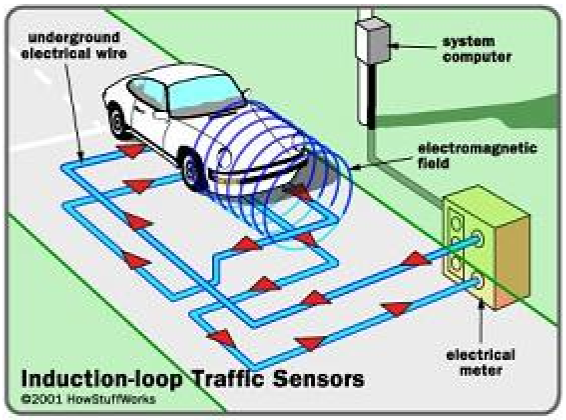

Figure 6: Induction-loop sensor diagram

Figure 7: Induction-loop sensor installation under the road surface

The third and the last component is the computer which is the brain of the system. If the light is red and the trigger is set off, the computer signals the cameras to take two pictures of the car. The first picture is taken when the car is about to enter the intersection and the second one is taken when the car is in the intersection.

When the light is yellow or green, the computer does not activate the red-light cameras even though the passing cars would set the triggers off. When the right turns red, the system turns on. If a car is already in the middle of the intersection when the light turns red, the red-light camera would not be activated.

In most systems, there are two rows of inductive loops, this way the computer can decide if the car is moving fast towards the edge of the intersection. If the car has passed the first row of loop but not the second row, the computer would infer that the car has stopped at the edge. If the light is red and a car activated both rows of triggers, the computer signals the red-light camera to take the first picture showing the car entering the intersection, then after a slight delay, second picture is taken which shows the car in the middle of the intersection.

Figure 8: The computer and the camera

The computer also includes the date,time, location of the violation, speed of the car and how many seconds after the light turned red, has the car entered the intersection. Some systems just take the picture of the license plate, whereas others try to take the picture of the driver as well.

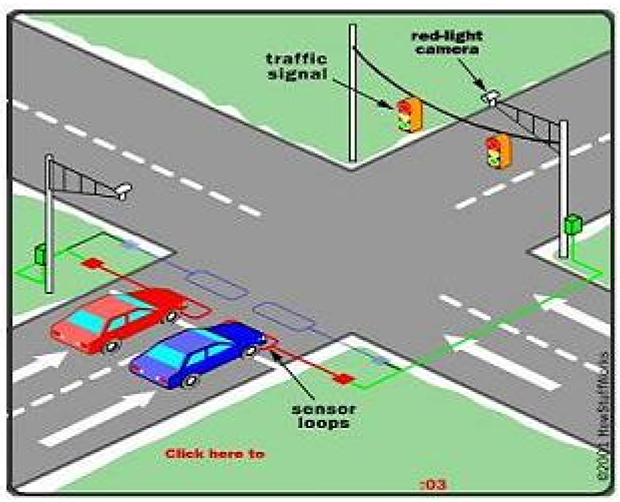

Figure 9: Red-light Camera System Diagram

[1]

[1] All figures in this section are from [3]

References

[1] Bertozzi, M., Broggi, A., and Fascioli, A. Vision-based intelligent vehicles :state of the art and perspectives. *Robotics and Autonomous Systems* (2000).

[2] Clearinghouse, V. D. Traffic report: A summary of vehicle detection and surveilance technologies used in intelligent transportation systems. *http://www.nmsu.edu/ traffic* (November 2000).

[3] Harris, T. How red-light cameras work. *http://www.howstuffworks.com/red-light-camera.htm* (2003).

[4] Kastrinaki, V., Zervakis, M., and Kalaitzakis, K. A survey of video processing techniques for traffic applications. *Image and Vision Computing 21* (2003), 359–381.

[5] Setchell, C. J. Applications of computer vision to road-traffic monitoring. *PhD Dissertation* (September 1997).

[6] Yoram Hofman, R. M. License plate recognition–a tutorial. *http://www.licenseplaterecognition.com* (2003).