

# Video Technologies for Roadway Surveillance and Automated Detection

C. Arthur MacCarley and Leonard Ponce  
California Polytechnic State Univ.

## ABSTRACT

Closed circuit television (CCTV) systems deployed on roadway networks are believed to be among the best mechanisms for providing useful, reliable information for effective traffic management. Video images may be displayed and viewed directly by traffic management personnel for surveillance purposes, or may be processed electronically for detection of traffic metrics such as average vehicle speed, flow volume, and traffic density.

This paper summarizes the current state-of-the art in video imaging and video signal processing technology for traffic surveillance and electronic detection. Technical considerations relevant to the selection of video cameras and computer vision hardware and software for this application are reviewed. Applicable standards are identified, and evaluation criteria and test procedures are described. Commercially available monochrome and color video cameras are examined with respect to sited criteria. The operational characteristics and performance of commercially available and experimental real-time video traffic detection (VTD) systems are summarized. General considerations and evaluation results are reported. Limitations and areas for further development are identified.

## INTRODUCTION

Video technology has become an important component of many traffic management operations. Direct video surveillance of roadways can serve as a valuable aid to traffic control personnel, extending their effectiveness considerably and making possible improved traffic management actions.

At the present time, field deployments of video-based surveillance and automated detection systems have been lim-

ited to high-need applications or experimental evaluations. Both technical limitations and institutional factors may be cited, foremost being the capital and continuing costs of large-scale deployment of video cameras, video transmission networks, monitoring and automated detection facilities.

Video camera technology has improved substantially within the past few years, with the introduction of monolithic semiconductor photosensor arrays for both visible spectrum and infrared (IR) imaging.

Concurrently, advances in microprocessor technology have made possible cost-effective real-time computer processing of video images of highway traffic for automated traffic detection. Video Traffic Detection (VTD) systems are now being considered key components of advanced traffic management systems (ATMS). The processed video image signal can yield traffic flow data that have traditionally been collected using conventional sensors such as in-ground inductive loop detectors. The video image is rich with information, including average vehicle velocity, flow volume (vehicles per unit time), traffic density (vehicles per lane per unit length of roadway), queue length (number of vehicles), license plate identification, and traffic stoppages. Several commercial or near-commercial systems are now available, based upon a number of hardware platforms, each capable of measuring some subset of these metrics. Work is in progress to extend these capabilities to include incident detection, vehicle classification, and detection of aberrant vehicle behavior.

Typical components of a video surveillance and detection system are illustrated in Figure 1.

In the work reported herein, we studied the two most crit-

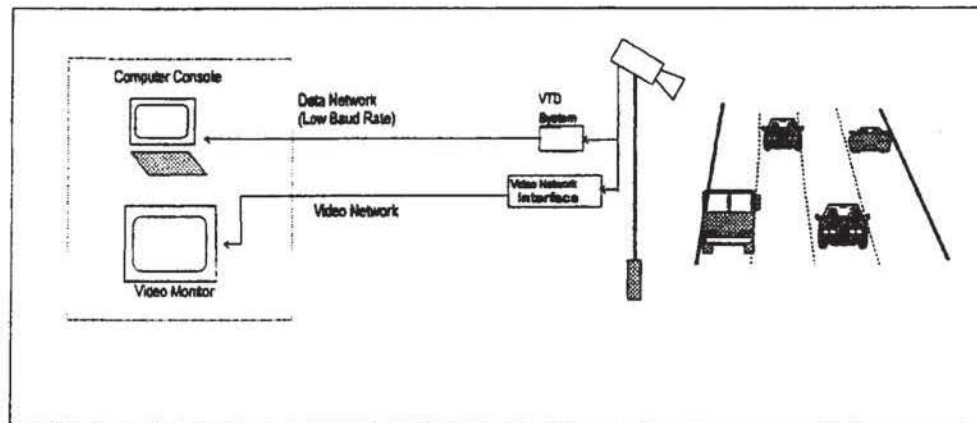


Figure 1. Components of a Video Traffic Surveillance and Detection Network.

ical elements of current and future video technologies: the imaging device (video camera), and the video post-processing hardware and software components comprising a VTD system. Evaluation and selection criteria were established that emphasized factors of greatest relevance to roadway surveillance and automated detection. Tests procedures were designed to address these criteria, including laboratory and field tests involving static and dynamic electronic video test procedures, and real and simulated traffic scenes in the field. Not covered here are other important components of the video system including the optics, electromechanical lens controls, video signal transmission network, video amplifiers, multiplexors or switchers, video signal compression equipment, and monitors.

### SURVEILLANCE AND DETECTION OBJECTIVES

For traffic monitoring applications, information requirements may include:

1. Traffic flow metrics: vehicle speed, traffic flow volume and density.
2. Vehicle classification: auto, truck, tandem, etc.
3. Roadway surface conditions: ice, snow, rain, flood, glare and surface flaws.
4. Visibility: roadway visibility as perceived by drivers.

5. Incident detection: collision or stalled vehicle.
6. Hazardous or impaired drivers: non-conforming vehicle behavior suggestive of driver impairment.
7. Specific vehicle identification: license plate, vehicle make, model, color, etc.

### VIDEO CAMERA TECHNICAL FEATURES

Video cameras designed for surveillance applications differ from general usage or broadcast cameras. Surveillance cameras are designed for optimum imaging of a stationary field of view, containing a very wide range of light intensities. This requires higher-than-normal resolution, and a wide dynamic range (light to dark range). Good sensitivity for best night vision may also be important. Surveillance cameras are often calibrated for a nearly linear response (a proportional relationship between incident light and the corresponding video signal voltage). It is known that this produces images that may be less aesthetically pleasing, and somewhat "flat" in appearance. Some cameras utilize contrast enhancement circuits, which accentuate light-to-dark or dark-to-light transitions in the image. This feature has advantages and disadvantages in traffic surveillance applications: Vehicle outlines are more crisply defined in low light or fog conditions, but signs and license plates become washed out due to the overshoot.



Since the mid-1980's, the majority of surveillance video cameras utilize solid state Charge Coupled Device (CCD) or "chip" technology, replacing older electron tube imaging systems exemplified by the Vidicon camera.

Traditionally, surveillance-type cameras are monochrome rather than color. Monochrome cameras generally provide greater resolution and sensitivity than color cameras. Recently, several high resolution color video cameras specifically designed for surveillance have become available, and are very popular for new traffic monitoring installations.

The spectral response of silicon-detector monochrome (black and white) solid-state cameras extends into the non-visible Infrared (IR) range. Some cameras are provided with removable IR-cut filters to reduce IR sensitivity, that tends to incorrectly image hot surfaces (such as vehicle tires and black roadway surfaces) as bright objects.

Most cameras contain both the focal plane array and associated electronics in a common package. However, for discrete surveillance, some cameras incorporate two separate modules. Some cameras have enclosures that are sufficiently durable and weather-tight, while others require separate environmental enclosures.

Some of the electronic features that distinguish different video cameras include the following:

#### **Video Signal Format**

Several video display and signal formats are in use internationally. The basic frame rate and vertical resolution (number of scan lines) for video signals conforms to one of two international standards:

**EIA** Electronic Industries Association (EIA) standard RS-170 specifies 30 frames per second, each frame displayed as two interlaced fields (half resolution frames) at a rate of 60 fields per second. 525 vertical lines of resolution are specified, each field consisting of 262.5 scan lines [EIA57].

Cameras provide analog signals with video information content in the range of 0 to 0.7 volts, which equilibrates to 0-100 IRE (Institute of Radio Engineers) units.

The color encoding format based upon the EIA display format is established by the National Television Standards Committee (NTSC) of the United States. EIA and NTSC standards are adhered to in the United States, Canada, Mexico, most of South America, and Japan [Benson86].

**CCIR** Video cameras for use in Europe generally conformed to CCIR (International Radio Consultative Committee) display formats, and PAL (Phase Alternation Line-rate) or one of three SECAM (Sequential Color with Memory) color standards. The basic display format is 25 frames per second displayed as fifty interlaced fields per second, and 625 vertical lines (312.5 per field).

#### **Adjustable Gamma**

Most cameras provide either a continuous adjustment or switch-selectable setting for gamma. This parameter affects the camera linearity in translating light levels to voltage levels.

#### **Color Adjustments**

Color cameras differ in their method of adjustment for the color white, which is an equal mix of the red, blue and green primary colors. Some cameras have automatic white balance capability, while some have none or only manual static adjustments. High-end cameras also permit individual color component adjustments for consistent color fidelity.

#### **Automatic Gain Control and Auto-Iris Control**

Automatic Gain Control (AGC) electronically adjusts the overall camera sensitivity in response to the average light level. This has the effect of maintaining a reasonably constant brightness level in the picture. On some cameras, the AGC may be switched off for testing purposes or special applications.

Sensitivity may also be controlled by an electromechanical auto-iris, which controls the aperture (iris) in response to the average light level. Auto-iris control produces a superior image normalization than AGC, but reacts comparatively slower due to the response time of the mechanical components. This can result in a slow recovery time following a brief bright exposure, such as headlight glare, in which the image is saturated (washed out).

#### **Imager Size**

CCD cameras typically utilize imaging ICs with diagonally measured imaging surface dimensions of between 1/3 and 2/3 of an inch, 1/2 inch being typical. Generally, the larger the chip, the better the image resolution capability, although this also depends on the size of each CCD imaging cell or *pixel*. Resolution in CCD cameras is directly proportional to the number of pixels on the



chip, typically between 200,000 and 400,000. Reducing the pixel size reduces cost, which is directly related to the silicon surface area of the chip. Defects in the wafer surface can result in "dead" pixels in the image.

The separation and alignment of the lens and imaging chip is critical for correct "back focus". This is adjustable on some cameras.

### Shutter Speed

Unless specifically designed for high speed (slow motion) photography, mechanical shutters are not used in video cameras. Shuttering is accomplished electronically. EIA/NTSC cameras have an effective shutter speed of less than 1/30th of a second, the rate at which complete video frames are produced (even though they are transmitted as 2 raster fields at 1/60 second each). Some cameras permit selection of faster shutter speeds; however, faster speeds reduce camera sensitivity, due to reduced photon integration time.

### Synchronization

When multiple cameras are integrated into a network, synchronization becomes an issue. If the cameras are not synchronized when switched successively onto the same monitor, picture roll occurs while the monitor is attempting to re-synchronize with the frame rate of the new camera. Surveillance cameras are manufactured with one of three frame timing control options:

#### *Internal clock.*

Camera frame rate is unsynchronized, timed independently from an internal clock.

#### *Phase lock.*

Cameras use the AC line frequency from the power supply for frame synchronization. An initial phase adjustment is usually provided to compensate for phase shift over a large network.

#### *Line-lock or external sync.*

An external sync generator provides a common frame synchronization signal to all cameras in the network.

Cameras using phase lock or external synchronization will switch smoothly without picture roll. Phase synchronization is useful only when all cameras are powered from a common AC source. However, a surveillance network with cameras spread out over miles of freeway would probably not meet this requirement. Line-lock external sync is advised for large network deployments.

### Signal Bandwidth

Commercial broadcast NTSC, PAL and SECAM signals are usually allocated approximately a 6 Mhz signal bandwidth, compatible with the channel separation of broadcast television. The resolution of standard TV receivers is usually consistent with this bandwidth limitation. For closed circuit (CCTV) systems, bandwidth is limited by the video distribution network. Signal bandwidth equilibrates directly to horizontal display resolution expressed in lines, to be discussed later. Commercial broadcast color video signals are usually limited to 200-300 lines of horizontal resolution. By comparison, a high quality monochrome CCTV surveillance camera may provide 600 lines of horizontal resolution.

### CAMERA PERFORMANCE REQUIREMENTS

The performance requirements for surveillance video cameras include consideration of the following:

1. Serviceability. A major component of ongoing system cost.
2. Rain Survival and Removal. Rain removal methods include windshield wipers, spinning windows, forced air deflection, and rain-avoiding enclosures.
3. Snow and Ice Survival and Removal. Snow flakes adhere to the foremost optical element. Ice could present problems with the mechanical components such as pan and tilt mechanism or zoom lens. Snow and ice removal methods include those for rain removal and the use of heated front window.
4. High Temperature Survival. Some mechanism for dissipation of external as well as internally generated heat may be necessary.
5. Dust and Grime Removal and Survival. Dust and grime reduce light transmission, and may cause scoring of the window or damage to the mechanical components. Automatic roll-fed transparent window covering is one available alternative to field service.
6. Ozone and Acidic Pollution Survival. The camera housing must be impervious to the effects of corrosive atmospheric conditions present in some areas.
7. Spectral Filtering. Filters may assist in the elimination of image artifacts. A polarizing filter may reduce glare, an IR filter may correct false imaging caused by IR sensitivity, and an ultraviolet (UV) filter may improve contrast during overcast



conditions.

8. **Projectile Survival.** Outdoor CCTV cameras are often targets of vandalism.
9. **Electromagnetic Noise Immunity.** Immunity to the effects of electromagnetic radiation from automotive ignition systems, high pressure vapor lamps, police radar, and mobile CB or cellular phone transmitters.
10. **Power Supply Noise Immunity.** Tolerance of poor power quality, such as low voltage, noise, spikes, and brief interruptions.
11. **Lightning Survival.** Suitable lightning protection is required to protect both the camera and other electronic devices in the signal path.

## CAMERA EVALUATION CRITERIA AND PROCEDURES

A sample of thirty-two surveillance video cameras were evaluated. Laboratory tests involved measurements of electronic parameters that underlie many of the surveillance requirements. These parameters included resolution, sensitivity, noise, dynamic range, grayscale linearity, geometric linearity, flair, bloom, lag, comet tail, vertical or horizontal smear, and backfocus accuracy. Color cameras were also tested for color fidelity as indicated by color vector magnitude and phase accuracy, and white balance. EIA/NTSC standard laboratory test procedures, apparatus, and illumination sources were used for all tests [Benson86]. A detailed description appears in [MacCarley92.1,93]. We address here selected test metrics and their relevance to the information needs of a TMC operator or the input requirements of a VTD system.

The tests may be divided into two categories: static tests that involve images containing no motion, and dynamic tests which utilize images with moving objects or light sources.

### Resolution

The *horizontal resolution* of the camera generally correlates with the amount of information present in the video signal generated by the camera. Greater resolution means that either 1) for a given angular resolution requirement, a larger field of view may be imaged, or 2) for a given field of view, a finer grain in the image may be discerned.

Resolution is quantified by the number of "television lines" that can be distinguished electronically in image. This is measured as the maximum density of black and white bars of equal width that can be distinguished along

the entire width (horizontal) or height (vertical) dimension of the television picture.

It is a factor of primary importance affecting the ability of a TMC operator to interpret information in the image. While the camera optics may be used to trade surveillance area for the minimum resolvable feature size in the image, the electronic resolution of the camera is a constant representing a product of these two factors.

Perceived resolution can also be limited by the monitor or the bandwidth of the communications path from the camera to the monitor. Therefore, resolution is important, but only up to the resolution-related limits of the other components of the distribution and display system.

*Vertical resolution* is fixed by the EIA/NTSC vertical line specification (525 lines interlaced). Since solid state cameras separate line scans with separate rows of pixels, the vertical resolution is some number slightly less than 525 (depending on the number of scan lines displayed), divided by an integer (usually one or two).

For solid state cameras, horizontal resolution is fundamentally limited by the horizontal pixel density of the imaging chip. However, bandwidth limitations in the signal path may also limit horizontal resolution.

### Sensitivity and Dynamic Range

*Sensitivity* is an indication of the ability of the camera to form an image in low light conditions. Daytime illumination levels greatly exceed the lower sensitivity limits. At night, the brightness of vehicle headlights is much greater than the reflected light from the vehicles or roadway features. The ability to detect features in the image other than just the headlight spots depends primarily upon the dynamic range of the camera, and secondarily on the actual low-light limit, assuming at least some minimum level of reflected light from the features.

Most manufacturers specify sensitivity as the minimum illumination level necessary for either full or usable video. However, the definition of full or usable video is often manufacturer-specific or nonrigorously defined. Measurement of sensitivity is further complicated by automatic gain control (AGC), IR-cut filters, and the spectral characteristics of the illumination. Ambiguities can be avoided by measuring camera sensitivity relative to the intrinsic camera noise level, an approach that cancels the effect of any gain in the signal path that acts upon both the image information and the noise.

The *dynamic range* of is measured as the response range



from the sensitivity limit to the saturation limit.

The *signal to noise (S/N) ratio* of a camera system is defined as the ratio between the camera peak signal output and the root mean square (RMS) noise output. S/N is evaluated by measuring the RMS noise output of the system when no light is permitted to enter the lens, and comparing this with the rated camera output. This measurement cannot be reliably made unless the AGC and black clip circuits of the camera can be disabled, which was not possible for all cameras.

*Bloom* is the spread of the image around the original image due to charge leakage in the pickup device. Bloom can also be observed as a result of faulty optics, sometimes due to poor or non-existent lens coatings. Although bloom can be a significant problem for tube cameras, solid state cameras are usually unsusceptible.

*Flare* is manifested as fluctuations in the black level of an image related to varying white levels. Flare is not known to be a common problem with solid state cameras.

#### **Gamma/Grayscale Linearity**

*Gamma* is a metric of the linearity of the relationship between the incident light intensity and the signal voltage produced by the camera, with  $\gamma = 1.0$  corresponding to a truly linear relationship. However, a unity setting is not always desirable, since the human eye, and often the monitor also, have nonlinear responses.

Linearity does not appear to be a factor of primary concern in traffic surveillance. From a TMC operator's point of view, the shade of gray representing a particular object in the scene is probably of secondary relevance (monochrome assumed). The relative intensity differences between features in the image convey the greatest information, provided that the image is not overly flattened out or binary due to excessive contrast.

#### **Geometric Linearity**

The *geometric linearity* of a camera is a measure of its tendency to introduce dimensional distortion in the image. This could be an important factor in the inference of distances or shapes in a traffic scene. Monitors in the TMC also introduce geometric distortion in the displayed image. Geometric linearity may be more critical for computer vision applications, since distances in the image may be inaccurately rendered.

Geometric nonlinearity is not a typical problem for solid-state (CCD) cameras, due to precise photolithography that

locates the pixels in the focal plane array. Distorted optics are more often responsible for any dimensional distortion.

#### **Vertical and Horizontal Smear**

*Vertical or horizontal smear* are problems common to MOS/CCD cameras. Smear is manifested as a white vertical or horizontal bar extending from a bright point light source in the image, across the entire image. This usually occurs only at sufficiently wide aperture settings such that the light source is saturated while the background is dark.

A camera exhibiting smear could be seriously limited for traffic surveillance at night, since the field of view contains numerous bright point light sources (headlights). Smear artifacts make the output of the camera unusable for most (possibly all) computer vision-based detection systems.

#### **Color Fidelity**

For color cameras, the TMC operator would expect a reasonably faithful reproduction of the colors and their relative intensities in the image. While *color fidelity* is only an aesthetic issue in entertainment, it could become a critical issue in traffic surveillance. For example, a TMC operator might need to identify a vehicle by color. Poor color reproduction might cause the vehicle color to be incorrectly reported.

*White balance* is an indication of a color camera's ability to faithfully produce the neutral color white.

For surveillance, the other half of the color reproduction system is the monitor. Color monitors provide adjustments for both color hue and intensity. The monitor adjustments can be used to some degree to compensate for the poor color fidelity of a camera. However, in a TMC, the capability must exist for any monitor to switch to any camera. Inconsistent color fidelity between cameras could yield distorted color reproduction on all but the original setup camera.

#### **Dynamic Problems**

Some metrics of camera performance are related to motion in the image. *Comet tail*, describes a problem when a bright object moves across a dark field, leaving a decaying after-image. Similarly, *lag* refers to the after-image visible when a nonsaturated (gray) object moves across a dark background. These problems are not common in solid state cameras, but are sometimes observed.



## Field Tests

Field tests were conducted at two sites. A 0.5 mile straight roadway section instrumented for vehicle position and velocity measurements was used for the *daytime field tests*. The cameras under test were mounted on a 25 meter tower at one end of the track. Vehicles and test symbols were placed or driven along the track at various speeds and distances from the camera.

For the *night field tests*, a camera platform was set up on a four-lane highway overpass. Both approaching and departing traffic scenes were viewed.

Human evaluators compared video images displayed on reference monitors, and completed written questionnaires intended to determine the information they could extract from the image, and qualitative issues such as sharpness, clarity, and color accuracy (when applicable). The ability of the human observers to identify specific features in a scene is duplicative of the more precise laboratory resolution and sensitivity tests. However, the relative values of color or grayscale linearity to a TMC operator are addressed in these tests - assessments that could not be done in a laboratory.

## CAMERA TEST RESULTS

Table 1 summarizes the test results. Individual cameras are identified by descriptor codes of the format vv:cb where vv is the vendor code number, b is nonzero for monochrome cameras and c is nonzero for color cameras. The following notes apply to Table 1:

- 1) Horizontal line resolution is compared at the -15dB point and is reported as an equivalent number of lines resolvable in the image along a single horizontal scan.
- 2) Low light sensitivity is the illumination at 0dB S/N ratio, reported in Lux.
- 3) Grayscale (gamma) linearity is stated as average absolute deviation from the ideal, reported in percent.
- 4) Geometric linearity is measured as the magnitude of the spatial misregistration over five points on the test chart. It is reported as a percentage.
- 5) Vertical smear (VS), lag and comet tail (L/C) are stated as yes or no, indicative of whether these problems were observed or not.

- 6) Field test scores are reported as ratios of the total points received to the maximum number of points possible.
- 7) Color fidelity measurements are reported as the absolute phase error in degrees, and magnitude error in percent, over six standard color vectors.
- 8) Cameras are numerically rated on a scale of 1 (worst) to 3 (best) according to overall performance in the laboratory tests, field tests, and finally a composite of all tests, indicative of the overall suitability of the camera for traffic surveillance applications.

The majority of the video cameras that we evaluated would be suitable for traffic surveillance applications. Cameras that received high ratings provided resolution above 400 horizontal lines, adequate sensitivity, and were free of excessive operational limitations. (All monochrome cameras exhibited IR sensitivity, and some smear at high sensitivity).

Operational problems of critical concern are those related to the basic usefulness of the camera in its intended application: synchronization problems, serious image distortion, extreme grayscale nonlinearity, very poor color true-ness (phase error), uncorrectable backfocus problems, excessive dead pixels, unusually poor resolution, or unusually low saturation limits.

Operational issues of less concern include moderate grayscale nonlinearity, a few dead pixels, minor image artifacts (such as dot-grid pattern noise), color intensity fidelity (magnitude error), poor sensitivity, marginal resolution (at least 250 lines), and lag/comet tail problems, if not excessive.

Vertical smear could potentially be a serious impediment to nighttime traffic surveillance, since it prevents the use of wide apertures at night. A wide aperture is necessary to image complete vehicles rather than just headlight pairs. The excellent low-light sensitivity of most cameras is of no value if bright headlight spots in the image cause vertical or horizontal smear. With as many as 100 cars in the field of view, 200 bright vertical smear lines renders the image useless, especially for automated detection by a VTD system.

All monochrome cameras tested that were not equipped with IR block filters were sensitive to IR radiation, at least in the 0.9 to 1.2 micron near-IR range. IR sensitivity causes false intensity levels in the image: black tires and hot asphalt surfaces appear white. A red car

TABLE 1 SUMMARY OF CAMERA TEST RESULTS

Camera Code	Perf. Class	Cost Class *	Horizontal Resolution (Lines at -15 dB of dynamic range)	Low Light Sensitivity (Lux at 0 dB S/N ratio)	Gamma Linearity (scaled avg. dev.)	Geometric Linearity	Color Fidelity (color cameras only) (mean nls. error)		Dynamic					Field Test 1 (comp. score)	Field Test 2 (comp. score)	Rating by Class			Comments
							Mag.(%)	base(Deg)	Saturated		Unsat.	Lab Tests	Field Tests			Over-all			
									Horiz. Scan	Vertical Scan									
																	VS	L/C	
101	Med.	Low	280	0.001	1.70%	<0.5%			yes	no	yes	no	no	2102	1202	1	2	1	Low res., back focus problem
102	Med.	Mod.	500 (-6.78 dB)	0.001	1.80%	<0.75%			yes	no	yes	no	no	2802	1002	3	2	2	
103	Med.	Mod.	400 (-10.35 dB)	0.003	1.80%	<0.75%			yes	no	yes	no	no	2702	902	3	1	2	
210	Color	High	290	0.81	0.80%	<0.5%	28.39	7.33	yes	no	no	no	no	1505	1505	2	2	2	
202	High	High	600 (-6.16 dB)	0.001	3.00%	<0.5%			yes	no	yes	no	no	2702	1802	3	3	3	Highest rated monochrome camera
201	Med.	Mod.	362	0.006	3.30%	<0.5%			yes	no	yes	no	no	1902	1702	2	2	2	
310	Color	High	278	5.13	1.20%	<0.5%	45.94	4.17	no	no	no	no	no	1605	1405	2	1	2	Immune to vertical smear
301	High	High	203	0.001	2.20%	<0.5%			yes	no	yes	no	no	1902	1302	1	2	1	Poor res., very high IR sensitivity
302	Low	Mod.	250 (-10.95 dB)	0.856	1.60%	<0.5%			no	no	no	no	no	802	1302	2	1	2	Immune to vertical smear
1401	Med.	Mod.	400	0.003/0.284	1.90%/2.10%	<0.5%			yes	no	yes	no	no	2502	1502	2	3	2	
1402	High	High	376	0.004/0.155	1.90%/1.00%	<0.75%			yes	yes	yes	yes	no	2102	902	2	2	2	Lag and comet tail
1410	Color	High	450 (-4.27 dB)	0.585	1.20%	<0.5%	14.85	5.5	yes	no	no	no	no	1905	1805	3	2	3	Excellent res. for color camera
901	High	Mod.	485	0.002/0.001	1.40%	<0.5%			yes	yes	yes	yes	no	2702	1602	2	3	2	Lag and comet tail
1101	Med.	Low	450 (-8.08 dB)	0.003	3.30%	<0.5%			yes	no	yes	no	no	2302	1402	3	2	2	
1102	Color	Mod.	489	0.003/0.074	4.10%	<0.5%			yes	no	yes	no	no	2302	1402	3	2	2	Very nonlinear grayscale
1110	Color	Mod.	396	1.976	3.80%	<0.5%	54.1	10.83	yes	no	yes	no	no	1805	1705	2	2	2	High res. but poor color fidelity
1001	Med.	Mod.	467	0.005	1.40%	<0.75%			yes	no	yes	no	no	2202	1502	2	2	2	
1002	High	High	512	0.005	1.80%	<0.5%			yes	no	yes	no	no	2102	1102	3	1	2	Dead pixels
610	Color	High	283	0.49	1.60%	<0.5%	21.6	11.83	yes	no	yes	no	no	2105	1405	2	2	2	Good sensitivity for color, low res.
601	Med.	Mod.	400	0.005	1.40%	<0.5%			yes	no	yes	no	no	2402	1102	2	2	2	
1301	Med.	Low	450 (-11.90 dB)	0.01	3.40%	<0.75%			yes	no	yes	no	no	2802	1302	2	3	2	
401	High	Mod.	470	0.346	2.10%	<0.5%			yes	no	yes	no	no	2302	2002	2	3	2	IR-immune, noise problem
410	Color	Mod.	450 (-8.09 dB)	1.113	0.90%	<1.0%	35.37	9	yes	no	yes	no	no	2105	2205	2	3	2	
501	High	Low	350 (-12.74 dB)	0.009/0.014	1.60%	<0.5%			yes	no	yes	no	no	2102	1502	2	3	2	
502	Med.	Mod.	450 (-10.42 dB)	0.004/0.004	2.20%	<0.75%			yes	no	yes	no	no	2102	1402	2	2	2	
510	Color	High	294	1.722/1.722	1.80%	<1.0%	18.38	13.17	yes	no	yes	no	no	1905	1805	2	2	2	Poor color "tintness" (phase error)
810	Color	High	263	0.319/0.314	1.10%	<0.5%	28.98	10	yes	no	yes	no	no	1605	1905	1	2	1	unstable color balance
801	Med.	Mod.	400 (-11.83 dB)	0.002/0.002	3.20%	<0.75%			yes	no	yes	no	no	2002	1502	3	2	2	Electronic auto-iris
1201	Med.	Mod.	439	0.005	1.60%	<0.5%			yes	yes	yes	yes	yes	2402	1302	2	2	2	Lag and comet tail
1210	Color	High	328	1.58	1.20%	<0.75%	29.96	10	yes	no	no	no	no	2105	2105	2	3	2	
701	Med.	Low	254	0.004	1.30%	<0.5%			yes	no	yes	no	no	2302	1102	1	2	1	Low resolution, image distortion
710	Color	Low	260	2.47	1.60%	<0.5%	n/a see test		yes	no	no	no	no	1305	1005	1	1	1	Poor resolution, sensitivity, sync

\* Price Class: &lt;\$500 Low, \$500-&lt;\$1000 Moderate, &gt;\$1000 High



appears whiter than a green car of equal visible reflectivity. It is difficult to say whether this is a real problem or not in traffic surveillance, since enough other visual queues exist in the image to correctly identify surfaces regardless of temperature.

Three of the cameras tested exhibited lag and comet tail, but none to a degree significant enough to be of concern in traffic surveillance. Usability as inputs to a VTD system might be compromised.

Human subjects seemed to accept color information in exchange for decreased resolution. While color information will never substitute for the resolution required to read a sign or identify a vehicle model, it could aid considerably in identifying particular vehicles, or distinguishing a vehicle from its own shadow, an important requirement for VTD systems.

Camera costs generally correlated well with performance, although a few exceptions were encountered. High cost is often associated with special features such as a ruggedized housing or accessible controls. In view of the overall system cost, installation and maintenance expense, and the projected service lifetime, the camera cost is probably a secondary consideration.

The ideal video camera for roadway surveillance would probably be a solid-state color camera with at least 450 lines of horizontal resolution, 0.5 Lux sensitivity, and complete immunity to bloom, lag, comet tail and smear. At the time of the evaluation, such a camera was not commercially available. The smear problem is a noteworthy deficiency of most monochrome cameras tested, and further development is suggested. Improved image array designs and process methods now entering the market may meet all above stated requirements.

## VIDEO TRAFFIC DETECTION SYSTEMS

VTD systems are designed to detect various traffic objects and measure traffic behaviors. The objects are usually vehicles, but could include pedestrians and vehicle queues [Rourke91]. The types of behavior include congestion, incidents, and average flow volume [Inigo89,D'Agostina92]. To perform these tasks, the VTD system must mechanize either individual vehicle (micro) analysis, global vehicle (macro) analysis, or both. These analyses require measurements of various complexity. A simple measurement might be simple vehicle detection. A complex measurement might be traffic pattern characterization.

The detail of the traffic data can be classified as being fine or coarse. A fine measurement is one that continuously involves all positions in the field of view. A coarse measurement is one that measures only at a few local areas and at periodic intervals [Bielik94]. Ideally it would be advantageous to carry out micro measurements of traffic to provide full information about the vehicles in the field of view. This is usually considered impractical due to its computational intensity. In many situations full traffic information is not required. Macro descriptions of traffic are cumulative measurements that take coarse micro measurements and average them over time.

## Detection Algorithms

A common attribute of all algorithmic approaches is the detection of motion in an image. Motion is typically detected by comparing successive frames of an image and observing pixels that have changed. To reduce the amount of calculations necessary, many VTD systems process only specific detection zones or features in the field of view.

All systems are software based. Some require specialized hardware platforms or components; others run on IBM PC compatible platforms requiring only video digitizing cards for the camera interface. Algorithmic sophistication is usually limited by the processing capability of the computational platform and video digitizing interface.

Two fundamental algorithmic approaches are generally used. We designate these as Type 1 and Type 2, and segregate systems into two classes based upon the algorithm type.

### Type 1

This approach is computationally simple, and uses only a small subset of the image information. Two or more virtual gates a known distance apart and perpendicular to the roadway are designated in the scene by the operator during the setup process. A vehicle moving down each lane causes an intensity change at the first gate, then the second gate. This pair of events is interpreted as the passage of a single vehicle. The vehicle's velocity is determined by measuring the elapsed time between the two gate-crossing events.

The accuracy of Type 1 speed measurements are related to the separation between the gates. This technique restricts a maximum of one vehicle passing between the gates at any time. Therefore the gates must be placed

Table 2. Video Traffic Detection (VTD) Systems.

System	Algorithm Type	Development Level	Institution / Company
VTDS	1b	Research Proto	Odetics
CCATS	1a	Commercial Product	Devlonics Control NV/Trafficon
IDSC	1a	Field Proto	Univ. of Tokyo
TRIP I	1a	Research Proto	Sheffield Univ./ Univ. of Manchester Institute of Science and Technology
TULIP	1a	Commercial Product	Univ. of Newcastle upon Tyne, Transport Operations Research Group/ Wootton Jeffreys Consultants Ltd
IPVD	2a	Field Proto	Japanese National Police Agency
TRIP II	2a	Demonstration	Napier Polytechnic of Edinburgh, Transportation Engineering Research Unit
Autoscope	1b	Commercial Product	Econolite Control Products Inc.
"Shigeta"	1b	Production Model	Mechanical Engineering Laboratory
VADAS/VDDAS	1b	Research Proto	Australian Road Research Board
SYD-TRAF 3	1b	Research Proto	Royal Inst. of Technology, Instrumentation Laboratory
"Becker"	2b	Research Proto	Univ. of Karlsruhe
CLIP	2b	Demonstration	Univ. College London, Transport Studies Group
"Deparis"	2b	Demonstration	INRETS-CRESTA (Centre de Recherche et d'Evaluation des Systemes de Transport Automatisés)
"Kobatake"	2b	Field Proto	Tokyo Univ. of Agriculture & Technology, Dept. of Electrical Engineering
"Koller"	2b	Demonstration	University of Karlsruhe
"Meccucci"	2b	Research Proto	Data Processing System
RAPAC	2b	Research Proto	Sheffield University, Dept of ELEE
"Takato"	2b	Field Proto	Hitachi, Ltd., Research Laboratory
TITAN	2b	Demonstration	INRETS: Inst. National de Recherche sur les Transports et leur Sécurité
IDET 100	2a	Commercial Product	Sumitomo Electric Ltd.
EVA	2a	Commercial Product	Eliop Traffic
CCIDS	1c	Research Proto	Trafficon
VDS	n/a	Commercial Product	Hughes Aircraft Company
ATAS	1a	Research Prototype	ASPEX
TAS	2	Commercial Product	CRS
Traffic Tracker	2	Prototype	Sense and Vision Electronics Systems



## Type 2

Type 2 algorithms might be referred to as "vehicle tracking" algorithms since they first detect the presence of cohesive objects (blobs) moving in the image, and then measure the velocity along its trajectory. Once a blob (vehicle) is acquired, its speed is determined along the plane of the roadway by reference to at least two scene features with known separation. Type 2 algorithms are more sophisticated and require significantly greater computer processing power. They are generally more robust at positive vehicle detection. The vehicle detection and tracking features of Type 2 algorithms are conceptually identical to well-established methods for military target acquisition and tracking.

Type 2 algorithms can also be divided into two categories based on complexity. Type 2a algorithms utilize only a subset of the image area and information, with reduced computational complexity. One example is "linear tracking" [Bielik94]. One line of pixels parallel to the flow of traffic is monitored. Successive frames are differenced to identify vehicles as "line-blobs" traveling along a virtual line in the center of each lane.

Type 2b algorithms analyze the entire image and track vehicles in any direction. This type of algorithm uses predictive filters, and is capable of tracking vehicles even when partially occluded by other vehicles [Chen92]. More accurate speed measurements are possible since a vehicle's position is known over a continuous interval, which might not be exactly parallel with the axis of the roadway [Takato89].

Type 2b algorithms are computationally demanding but usually amenable to code vectorization, making them suitable to parallel processing hardware for implementation in real-time.

One Type 2b tracking method involves the use of *optical flow* to separate moving objects from a stationary background. Position gradients for groups of pixels are calculated together to find a common flow vector. Once a flow vector is found for an image, objects can be segmented. Velocities and next position estimates are predictable from the flow vector [Weber93,Rao92].

The use of artificial neural networks has been suggested for detection of macro-vehicle behavior such as congestion, and vehicle identification. The neural network implements a learning machine that can be trained to make decisions based on inputs that resemble trained inputs. To be completely effective, every type of vehicle

must be used to train the net [Dougherty93,Wan92].

Model based tracking is also being studied, primarily as an improved means for dealing with shadows and vehicle occlusion. This technique uses line-models of several types of vehicles. The image is processed to detect edges, and the vehicle model is scanned until a close match is found. Difficulties arise when the vehicle orientation changes, or multiple vehicles proximate [Kilger92,Koller93].

## EVALUATION OF VTD SYSTEMS

Table 2 lists 21 commercial and prototype VTD systems, available at the time of our study (1993). Eight of these systems were subjected to comprehensive performance evaluation: ASPEX ATAS, CRS TAS, Devlonics (now Traficon) CCATS, Eliop EVA, ISS (now Econolite) Autoscope, INRETS TITAN, Sense & Vision Systems Traffic Tracker, and the University of Newcastle TULIP system. In addition, field tests using live video feeds from existing roadway surveillance cameras were conducted on the Sumitomo IDET 100, and upgraded versions of the Traficon CCATS, Econolite Autoscope, and Eliop EVA systems.

We focused on the ability of the systems to accurately detect, count and determine velocities of vehicles as the primary metrics of performance.

All systems tested were designed to handle oncoming traffic, although most could also handle departing traffic. Detection of departing traffic is now usually considered more reliable, especially for Type 2 algorithms, since the vehicle occupies the largest number of image pixels at the start of its track, rather than the end.

All systems utilized monochrome video images and were designed to operate with standard EIA or CCIR monochrome video cameras.

All systems required full-bandwidth video inputs, and were incompatible with lossy compressed video signals, since temporal and spatial relationships are upset in the compression process. This restriction implies that all systems must be installed at camera location in the field, unless full-bandwidth video is available at the TMC.

Video images of roadway traffic were acquired, time coded, and recorded in both EIA and CCIR formats, using S-VHS video recorders. Monochrome high-resolution MOS/CCD cameras were used, with variable focal length lenses and mechanical aperture adjustments.

Table 3. Summary of Video Test Suite for VTD System Evaluation.

Test#	Parameter Tested
1.	Large Number of Lanes
2.	Small Number of Lanes
3.	Day to Night Transitions
4.	Shallow Camera Angle
5.	Steep Camera Angle, Departing Traffic
6.	Shallow Camera Angle, Departing Traffic
7.	Night, Steep Camera Angle, Approaching
8.	Night Shallow Camera Angle, Approaching
9.	Night Steep Camera Angle, Departing Traffic
10.	Night Shallow Camera Angle, Departing Traffic
11-18.	Same as 3-10 (above), Side Camera Mounting
19.	Weather - Fog
20.	Weather-Rain, Daytime
21.	Weather-Rain, Night-time
22.	Unstable Camera mount - Sway
23.	Heavy Traffic - Capacity Operations
24.	Congested Traffic
25.	Heavy Shadows from Vehicles
26.	Heavy Shadows from Environment
27-28	Ignition and Electromagnetic Noise

A suite of 28 test conditions was defined for evaluation of the systems, described in Table 3. Each test segment is twenty minutes in length, including a ten minute initial period to permit the system under test to cancel the background and adapt to the ambient light level.

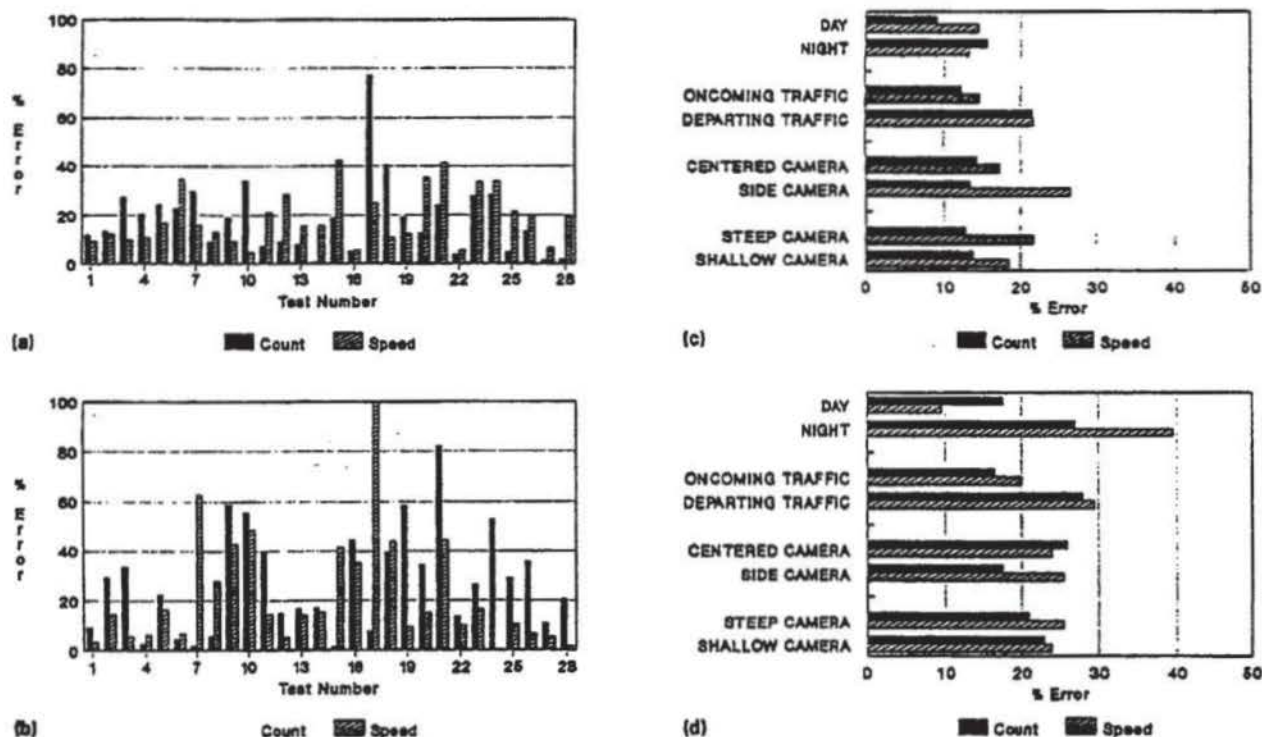
Parameters included day and night illumination levels, variable numbers of lanes (2 to 6), various camera elevations and angles to the roadway, rain and fog conditions, camera vibration and sway, traffic conditions ranging from free flow through heavy congestion, long shadows from vehicles or stationary objects, and the effects of simulated ignition noise and 60 Hz electromagnetic noise combined with the video signal. Tests were performed on both approaching and departing traffic. As a practical matter, only those combinations of variables most representative of standard deployment scenarios were included in the test suite. Table 3 indicates the parameter or combination of parameters emphasized in each of the 28 standard tests.

Actual traffic counts and vehicle velocities on a per-lane basis were determined from the videotaped images over the duration of each segment. This was done manually, on a frame-by-frame basis.

Most systems were designed for camera placement directly above the roadway centerline, at a height of between 10 and 15 meters. An exception to this was the TITAN (INRETS) system, designed for very high mounting well off the side of the roadway, with a large field of view [Blosseville89]. A high camera position minimizes vehicle occlusion, but is more prone to sway and vibration. A centered camera minimizes perspective distortion, while a roadside placement is easier to install and maintain, and provides a greater field of view.

All test suite images were acquired from freeway overpasses, with cameras placed above the roadway centerline, and also aligned with the roadside edge. Camera heights varied from 8.3 to 14.2 meters above the roadway surface, measured using an ultrasonic range finder.





Summary of test results: average absolute error by test for (a) Type 1 systems and (b) Type 2 systems; average by grouping for (c) Type 1 systems and (d) Type 2 systems.

Figure 2. VTD System Average Performance, Grouped by Algorithm Type.

cameras were solid state. The cameras were mounted on various structures: a 17 meter pole, a 9 meter overpass, an 8 meter changeable message sign, and a 17 meter cross member of a suspension bridge.

The video feeds were time-lapse recorded to provide a visual record of the actual traffic conditions and verification of reported VTD data. Tests consisted of continuous data collection: 1-2 hour periods during day-to-night transitions, and 6-12 hour periods in other cases. At most of the test sites inductive loop detectors were available and used for comparison to the VTD system count data. A doppler Radar speed measurement gun was used to randomly check speeds reported by the systems.

A qualitative evaluation of system human factors was also performed, considering issues of ease-of-setup and use, quality of graphical interface of data display.

## VTD SYSTEM TEST RESULTS

Figure 2 summarizes the average performance of the systems, classified by algorithm type. Average detection accuracies are reported by algorithm class for each test condition. Further details are available in [MacCarley92.2].

For all systems, we observed error rates usually less than 20% for vehicle count and speed measurements over a mix of low, moderate, and high traffic densities, with optimum camera placement and clear/daylight/nonshadow conditions. No system was clearly superior to the others under optimum conditions.

Systems designed for high camera placement were usually intolerant of partial occlusion of vehicles, yielding high error rates for tests with lower camera heights.

Tests with slow-moving, high traffic densities usually yielded reduced accuracy and occasionally, complete detection failure, probably attributed to the background subtraction method employed. Another problem for Type 2a algorithms is the assumption of at least some road surface separating vehicles. If the vehicles are traveling too close together (dense traffic), the road surface may not be visible between vehicles. These situations were emphasized in tests 23 and 24 (Table 3).

Lighting changes at sunrise and sunset caused reduced accuracy. During these periods, the systems must make a transition from daytime algorithms, which detect entire vehicles, to nighttime algorithms, which detect headlight groups. This is a notable deficiency, since peak traffic periods usually coincide with sunrise and sunset. Systems that used adaptive techniques to automatically switch based on the average scene intensity fared better during the transitions. If a system switched to its night algorithm early, it would miss vehicles without headlights. If the algorithm switched too early, it would frequently over-count lighted vehicles. We observed that the overcounting followed by undercounting usually compensate for each other over longer intervals, leading to incorrect assumptions of accuracy during transition periods.

Tests 21, 25 and 26 (Table 3) emphasized two aberrant conditions that caused particularly high error rates for most systems, rain at night, and long vehicular and stationary shadows, respectively. Long shadows are particularly a problem at sunrise and sunset, adding to the transition difficulties just mentioned. Headlight reflections, especially from a wet road surface, cause similar detection errors. Problems with headlight reflections are exacerbated by shallow camera angles due to low camera positions. As a vehicle approaches, its headlight reflection would change in size and position relative to the vehicle, appearing to accelerate or decelerate.

These problems are related in the sense that they are challenges to the ability of the systems to discriminate actual vehicles from other moving areas of high contrast (either light or dark) in the camera image.

Type 1 algorithms attempt to cancel headlight reflections or vehicle shadows by rejecting detection events that occur in too brief a time interval. Type 2 systems attempt to correlate a shadow or reflection with an associated vehicle. However, the source of the shadow or light may be outside the field of view, say a car off the detected area of roadway, aircraft overhead or the shadow

of a tall object or tree. In these situations, both algorithm classes usually fail.

The effects of added electronic (ignition or RF) noise were studied in tests 27 and 28. Generally, low noise levels had little effect on count or speed accuracy, up to a threshold where detection failed completely. Type 1 systems seemed to fail at a lower noise threshold than Type 2. A similar observation was made in test 19 for atmospheric fog. Significant vibration or other movement of the camera, simulated in test 22, caused count and speed errors for Type 1 systems, but often caused loss of track and complete detection failure for Type 2 systems.

Under optimum daytime conditions, the Type 1 systems generated more accurate vehicle counts, while the Type 2 systems generated more accurate speed measurements. Aberrant conditions yielded high error rates for both algorithm classes. Overall, Type 1 systems showed somewhat lower error rates in both vehicle count and speed measurements. It should be noted that the two Type 2 (out of eight) systems subjected to the test suite were prototype versions at the time, and the two Type 2 (out of four) systems subjected to field tests were relatively new commercial systems following approximately two years of development, compared with approximately ten years for the two Type 1 systems.

## SUMMARY

Video camera and computer video processing technologies can play a valuable role in improved traffic management. Monochrome video cameras generally excel in resolution and sensitivity, but remain susceptible to vertical or horizontal smear at high sensitivity levels which could limit their usefulness for highway surveillance at night. Color information is a valuable feature for both surveillance and automated detection.

VTD systems extend the utility of video surveillance components, by providing non-contact measurements of standard traffic metrics. No clear advantage was demonstrated for more sophisticated Type 2 algorithms over Type 1 for basic traffic measurements, but Type 2 methods are required for other detection objectives such as incident detection and vehicle classification.

Conditions that degraded detection performance were:

- 1) Non-optimum camera placement.
- 2) Transition from day to night.
- 3) Headlight reflections on wet pavement.



- 4) Shadows from vehicles or objects outside the detection area.
- 5) Obscured atmospheric conditions (fog or heavy rain).
- 6) Camera motion.

Further development is needed to better handle non-ideal conditions. Incompatibility with video data compression methods is also a limitation, since signals from field deployed cameras are most economically transmitted to a TMC in compressed format.

## REFERENCES

- [Benson86] Benson, K. Blair, ed. "Television Engineering Handbook," Section 2.3, CIE System, McGraw-Hill, New York, 1986.
- [Bielik94] Bielik, A. "Image Measurement Systems for Traffic Analysis: Task and Operation, State-of-the-Art, Accuracy," Licentiate Thesis TRITA-ILA 94.01, The Royal Institute of Technology, Stockholm, April 1994.
- [Blosseville89] Blosseville J. M., et al. "Titan: A Traffic Measurement System Using Image Processing Techniques," IEEE Second International Conference on Road Traffic Monitoring, London, 1989.
- [Chen92] Chen, L., Chang, S. "A Video Tracking System With Adaptive Predictors," Pattern Recognition, v.25, n.10, February 1992.
- [Cypers90] Cypers L., et al. "CCATS: The Image Processing-Based Traffic Sensor," Traffic Engineering and Control, June 1990.
- [D'Agostino92] Agostino92'u>(518u+ln).br D'Agostino, S. "Machine Vision and Intelligent Highways," Photonics Spectra, April 1992.
- [Doughery93] Doughery, M. and Kirby, H. "The Use of Neural Networks to Recognize and Predict Traffic Congestion," Traffic Engineering and Control, June 1993.
- [EIA57] Electronic Industries Association, EIA Standard RS-170, Rev. TR-135., Washington D.C., 1957.
- [Inigo89] Inigo R.M. "Application of Machine Vision to Traffic Monitoring and Control," IEEE Transactions on Vehicular Technology, August 1989.
- [Kilger92] Kilger M. "A Shadow Handler in a Video-based Real-time Traffic Monitoring System," IEEE Workshop on Applications of Computer Vision, Palm Springs, California, November 1992.
- [Koller93] Koller, D., Daniilidis, K., and Nagel, H.H. "Model-Based Object Tracking in Monocular Sequences of Road Traffic Scenes," International Journal of Computer Vision. v.10 n.3, June 1993.
- [MacCarley92.1] MacCarley, C.A. "Evaluation of Closed-Circuit Television Technology for Application in Highway Operations," Final Project Report, Caltrans Contract 51J932, California Polytechnic State University, San Luis Obispo, CA., 1992.
- [MacCarley92.2] MacCarley, C.A., Hockaday, L.M., Need, D. and Taff, S. "Evaluation of Video Image Processing Systems for Traffic Detection," Transportation Research Record, No. 1360, Transportation Research Board, National Research Council, Washington, D.C. 1992.
- [MacCarley93] MacCarley, C.A., Need, D. and Nieman, R.L. "Video Cameras for Roadway Surveillance: Technology Review, Test Methods, and Results," Transportation Research Record, No. 1410, Transportation Research Board, National Research Council, Washington, D.C. 1993.
- [Michalopoulos90] Michalopoulos P. "Automated Extraction of Traffic Parameters through Video Image Processing," ITE 1990 Compendium of Technical Papers, Presented August 1990.
- [Rao92] Rao, B. "Data Association Methods for Tracking Systems," Active Vision, edited by Blake A. and Yuille A., Massachusetts Institute of Technology, 1992.
- [Rourke91] Rourke, A., and Bell, M.G.H. "Queue Detection and Congestion Monitoring Using Image Processing," Traffic Engineering and Control., v.32, n.9, September 1991.
- [Takato089] Takato M., et al. "Traffic Flow Measuring System Using Image Processing," SPIE Vol.1197 Automated Inspection and High-Speed Vision Architectures III, 1989.
- [Wan92] Wan, C.L., and Dickinson, K.W. "Computer Vision and Neural Networks for Traffic Monitoring," IEE Conference Publication, Institution of Electrical Engineers (U.K.), n.355, p.143, 1992.
- [Weber93] Weber J. and Malik J. "Robust Computation of Optical Flow in a Multi-Scale Differential Framework," University of California Institute of Transportation Studies, UCB-ITS-PWP-93-4, July 1993.