4th WSEAS Int. Conf. on COMPUTATIONAL INTELLIGENCE, MAN-MACHINE SYSTEMS and CYBERNETICS Miami, Florida, USA, November 17-19, 2005 (pp254-259)

Biologically-Inspired Multi-Object Tracking Algorithm Applied to Traffic Monitoring

JOHN A. MOYA and DAVID Z. SÁENZ
Department of Electrical and Computer Engr.
University of Texas at El Paso
El Paso, TX 79968
USA

Abstract: - Various sensing modalities are presently used or conceptualized for use in roadway management projects. These include, among others, bind-type (utilizing inductive loops, reflected light, etc.), acoustical, and radar sensors. However, due to lower cost; flexibility in installation, maintenance and use; high information content; lack of emissions; and wide sensing area, some researchers have argued that the use of video cameras and computer vision techniques offer the best choice. Nonetheless, due to various issues – including lighting condition sensitivity and computational complexity – even the latter approach can be less than optimal. One possible reason that a computer vision scheme can become overly complex and sensitive to lighting conditions is that too much local information needs to be captured/analyzed by the algorithm. Utilizing a patent-pending approach, this paper discusses a just-enough-smart, largely-global technique based on a simplified version of the fly's eye that can address these latter issues. This approach can easily generate vehicular flow data – such as entrance, exit and throughput rates – or serve as a front-end process for yielding pointers to vehicles to assist in the analysis activities of other algorithms. Results from the application of this approach to the tracking of multiple moving vehicles or people in real scenes with fixed and moving cameras are also presented.

Key-Words: - traffic monitoring, security, computer vision, visual technique, global approach, fly's eye

1 Introduction

In hopes of addressing ever-increasing vehicular traffic flow concerns in the larger cities of the USA, various smart roadway projects have been conducted [1, 2, 3]. The general features desired in such smart roads and the various sensor modalities presently conceptualized for use in them are reviewed by Stevens in [4] and Bahler et al. in [5], respectively. Further, Masoud and Papanikolopoulos [6] review different algorithms that can be utilized for a related problem, the protection of pedestrians.

Among the most basic sensors, and also those most traditionally used in traffic monitoring, are inductive loops [7]. When vehicles move over these sensors, vehicle presence is signaled by a change in current. Similar monitoring systems can also be created using fiber-optic loops [8] or via sensors that detect changes from the background electromagnetic field [9, 10] or reflected light [11, 12, 13, 14]. For example, when an infrared laser source is used to illuminate a road and a correspondingly chosen detector is used to distinguish changes in the reflected light, the presence or absence of a vehicle can be signaled at that fixed point on the roadway.

With their limited sensing capabilities, the above bind-type systems are generally designed to only detect vehicles when they move through the sensor's narrow receptive area. For instance, the detection of speed is typically beyond a single bind sensor's function. Such data can only be measured by using multiple detectors. Further, along with the increased costs associated with the multiple sensors in the latter case, when a bind-type sensor is buried in the road surface, non-flexible placement occurs and disruptive installation and maintenance concerns exist.

Wider-sensing area approaches also exist, including systems utilizing acoustical- [15, 16, 17] and radar- [14] based sensors. However, these systems can similarly be less than optimal. Although the above bind-approach problems are in general addressed, these latter techniques are possibly even more susceptible than bind sensors to road clutter and occlusions (parked or multiple moving vehicles) on the roadway and thus these systems can have limitations on their specificity.

With the limitations in the above sensors, some researchers have argued for the use of video camera based systems instead [2, 18, 19]. The use of cameras has a number of advantages starting with their installation/maintenance, which does not necessarily need to disrupt vehicular circulation. Video cameras can also be inexpensive to purchase, provide high levels of information, and can have a wide sensing area. Further, unlike radar-, sonar-, or laser- based systems, they do not emit signals. Finally, given appropriate hardware and computer vision algorithms, clutter and occlusions along with specificity can be a lesser problem than for radar- and acoustical- based sensors.

Location of moving vehicles in a video sequence has been tackled by using a variety of non-statistical and/or statistical computer vision approaches. Common non-statistical techniques include the use of optical flow discrimination [19], image differencing for target/ground separation [20, 21], and edge detection and joining to outline vehicles [21]. Statistical techniques have included the use of mixture of Gaussian models [22] or other a-priori statistical information on image background [3] that allows the separation of likely background and vehicle pixels. Once vehicles are located, their tracking has then been accomplished using various techniques, including learning approaches [2, 20], reconfigurable templates [6, 18, 22] and Kalman filtering [6, 22].

However, although application of these computer vision techniques to traffic problems can improve the situation, some disadvantages do exist for these approaches [2]. For instance, the above computer-based approaches can be computational intensive and further can be sensitive to lighting conditions.

The authors, in previous efforts [23, 24], have suggested that one possible reason that a computer vision algorithm can become overly complex and sensitive to lighting conditions is that too much local information needs to be captured/analyzed by the algorithm. Further, an approach inspired by the fly's eye was presented that may be able to address these issues in complex images. The developed approach utilizes a global, just-enough-smart, quickly-updating technique. To improve its computational speed, it was suggested that one could also consider creating a new type of camera containing a sensor that could electro-optically implement at least a portion of the algorithm. The latter results in an algorithm runtime that may be of linear or lesser order.

In Section 2, this paper extends the latter technique. Section 3 and 4 then report results from the use of this extension as a front-end process for the detection of vehicles and people with a fixed and moving video camera. The presented approach yields a low level of scenery understanding, but could directly be used to measure aspects such as vehicular in/out/through flow rates, speed, queue length and closeness. Further, this technique may be capable of reducing the area of analysis for follow-up smarter software stages which could for instance assess overloading of streets, redirect traffic via the closing of entry/exit pathways, track the trajectory of a specific car, recognize vehicle class characteristics, or assist in warning drivers of accidents and alternate routes. Security monitoring applications are also possible.

2 Tracking Algorithm Extension

The problem of vehicular tracking is in some sense similar to the processing problem faced by the fly when it is trying to acquire a mate via visual tracking [25]. Moreover, since the tracking maneuvers of flies must be carried out in nature at high speed and with possibly changing backgrounds (effectively different lighting conditions), the use of a tracking algorithm suggested by this insect could likely be fast and may be amenable to the various lighting conditions on a roadway. Thus, potential exist to apply aspects of the fly's structure and processing to the vehicular tracking problem.

By analogy to the neuro-ommatidium, the basis structure in the fly's eye [26, 27, 28, 29], the tracking approach presented in [23] and [24] gathers a set of image powers $\{P_m(t)\}$ via four (i.e., m= 1 to 4) Gaussian-shaped position sensitivity function. Further, given these power values over time, it is shown in [23] and [24] that the radial position of a single moving object (of sufficient contrast to the background) in a sequence of square images can be found relative the center of each of the Gaussian filters using

$$R_m(t_n)=N\left(\sqrt{2}+u_m\right)\left(1-u_m\right)\left(1-\frac{1}{\pi}\cos^{-1}\left(1-\frac{2y_m(t_n)}{A(t_n)}\right)\right), \quad (1)$$

where N is the length of an image side; u_m is the shortest, normalized diagonal distance from a corner in the image to the center of the Gaussian filter $h_m(i,j)$; $y_m(t_n)$ is the absolute change in the power received via $h_m(i,j)$ and given by

$$y_m(t_n) = \left|P_m(t_n) - P_m(t_0)\right|; \quad (2)$$

and $A(t_n)$ estimates the maximum power loss that could occur from movement of the object across $h_m(i,j)$. Further discussion concerning $A(t_n)$ is reserved for [23] and [24]. With knowledge of the latter radii, a combined estimate of the position of the moving object can then be determined using triangulation.

Extension of the above for the purpose of multiple object tracking may be carried out by replicating the Gaussian sensing quad filter sets (producing a tiling pattern of overlapping Gaussian filter quads) and executing Eqn. 1 for each of the tile filter groupings. With overlapping, it is expected that multiple tiles will detect the same moving object. Thus, some form of merging must be implemented. This could be carried out using averaging, various clustering algorithms, a neural network, etc. [24].

3 Tracking Results

Given an 8x8 array of Gaussian filters, i.e. a 7x7 uniform tiling of overlapping Gaussian quad filter groupings, five single minute videos acquired from the Texas Department of Transportation were processed to locate moving vehicles using the above concepts. In this processing, the video image size was reduced from the original 320x240 pixels to 128x128 pixels and u_m was set to 0.25. Further, $A(t_n)$ was selected via an estimate of object size, as suggested in [23]. Typical processed frames from the video sequences are shown in Fig. 1. Note that in certain cases groups of vehicles are detected as a single blob, and that in others, vehicles were not detected at all.

Taking both grouped and individual detections as successful, Table 1 shows that the overall detection rate for vehicles in these videos was approximately 93%. Where detections did not occur, insufficient contrast was present between the vehicle and the background to exceed a particular noise-limiting threshold (1.0% power change from the background). It is noted that in some of these latter cases partial detections did occur but the detection was intermittent, producing durations of detection that were judged to be similar to those potentially produced by camera noise. Thus, these would reasonably be ignored for instance in a vehicle-counting task. In any case, when the power change threshold was not exceeded, Eqn. 1 was not executed.

4 Related Applications

Fig. 2 shows results from applying the above approach to a security monitoring situation (the detection of people moving about parking lots or fields outside a particular facility). The resolution of Figs. 2B and 2C were intentionally reduced to exhibit the fact that high-resolution cameras are not a required aspect of this approach. The goal is to locate targets of possible interest. After these have been acquired a higher resolution zoomed-in view may be utilized for instance in an identification algorithm.

Although the details are reserved for separate publication, it is worthwhile to also note that the presented algorithm can be used in moving-camera, moving-object, collision-avoidance applications. An example of such processing can be seen in Fig. 3. However, unlike the above fixed background situations, an application such as the latter may require a few periodic updates, such as the periodic re-initialization of the value $P_m(t_0)$ used in Eqn. 2.

Table 1: Vehicle Detections in Traffic Videos

Video	Vehicles Detected Individually	Groups of Vehicles Detected as a Single Blob	Vehicles not Detected
1	17	1	3
2	33	3	0
3	16	5	1
4	20	4	3
5	21	3	2
Total	107	16	9

5 Conclusions

Various sensing modalities are presently used or conceptualized for use in roadway management. Further, based on the opinions of some researchers, the use of video cameras and computer vision techniques offer the best choice. Nonetheless, due to various issues – including lighting condition sensitivity and computational complexity – even the latter approach can be less than optimal.

One possible reason that the latter technique can become overly complex and sensitive to lighting conditions is that too much local information needs to be captured/analyzed by the algorithm. Utilizing a patent-pending approach, this paper discussed a just-enough-smart, largely-global technique based on a simplified version of the fly's eye that may be able to address these latter issues. Results from the successful application of this approach to the tracking of multiple moving vehicles and people in real scenes using fixed and moving cameras were also presented. Moreover, it was shown that this

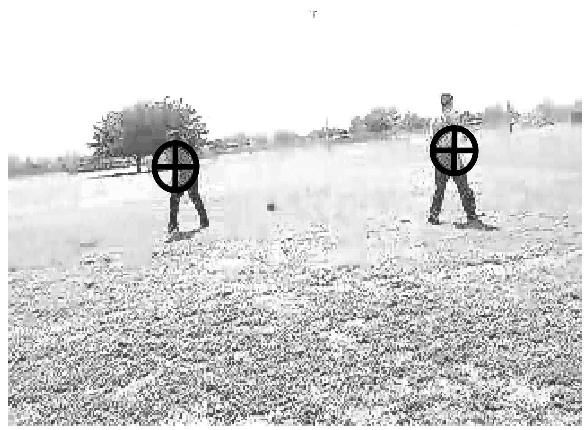
A

B

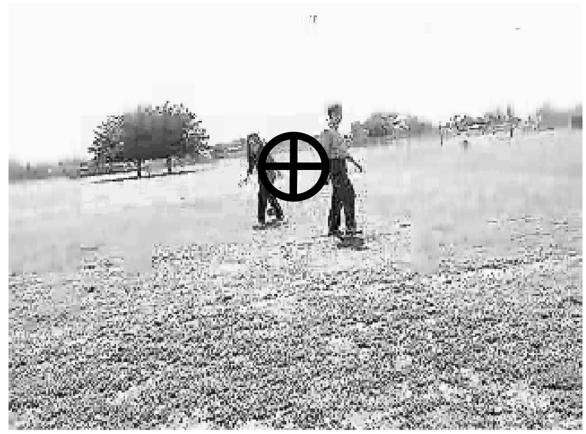
C

Fig. 1 Output of Vehicle Detector. **A.** Vehicle undetected. **B.** Two individual vehicles detected. **C.** A pair of vehicles detected but merged. **Note:** The color of the detection indicator was adjusted to improve viewability.

A



B



C

Fig. 2 Security Detector. **A. and B.** Individuals detected. **C.** A pair of individuals detected but merged.

approach could easily generate vehicular tracking data and possibly serve as a front-end process for yielding pointers to moving objects to assist in the analysis activities of other algorithms.

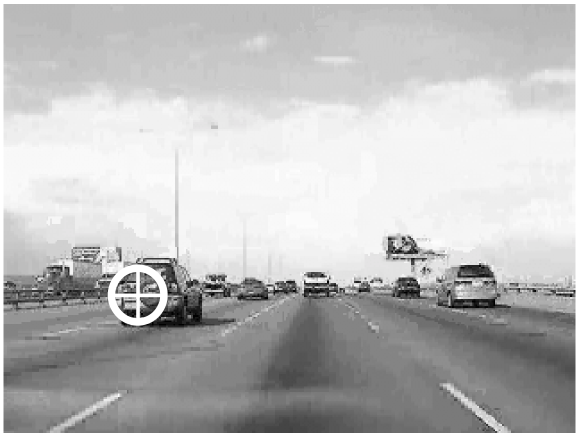


Fig. 3 Collision Avoidance Detector. Cars near a moving vehicle are located.

References:

[1] R. Horowitz and P. Varaiya, Control Design of an Automated Highway System**,** *Proc. IEEE*, Vol. 88, 2000, pp. 913-925.

[2] D. Bullock, J. Garrett Jr., and C. Hendrickson, A Neural Network for Image-Based Vehicle Detection, *Transport. Res. C: Emerging Techn.*, Vol.1C, 1993, pp. 235-247.

[3] P. G. Michalopoulos, Vehicle Detection Video Through Image Processing: The Autoscope System, *IEEE Trans. Vehicular Techn.***,** Vol. 40, 1991, pp. 21-29.

[4] W. B. Stevens, Evolution to an Automated Highway System, *Automated Highway Systems*, P. A. Ioannou, Ed., Plenum, NY, 1997.

[5] S. J. Bahler, J. M. Kranig, and E. D. Minge, Field Test of Nonintrusive Traffic Detection Technologies, *Transport. Res. Record*, no. 1643, 1998, pp. 161-170.

[6] O. Masoud and N. P. Papanikolopoulos, A Novel Method for Tracking and Counting Pedestrians in Real-Time using a Single Camera, *IEEE Trans. Vehicular Techn*., Vol. 50, 2001, pp. 1267-1278.

[7] J. Baras, A. Dorsey, and W. Levine, Estimation of Traffic Platoon Structure from Headway Statistics, *IEEE Trans. Auto. Control*, Vol. 24, 1979, pp. 553-559.

[8] D. Donlagic and M. Hanc, A Simple Fiber-Optic Vehicle Axle Detector for Roadways, *IEEE Trans. Vehicular Techn.*, Vol. 52, 2003, pp. 401-405.

[9] T. Uchiyama, K. Mohri, H. Itho, K. Nakashima, J. Ohuchi, and Y. Sudo, Car Traffic Monitoring System using MI Sensor Built-In Disk Set on the Road**,** *IEEE Trans. Magnetics*, Vol.36, 2000, pp. 3670-3672.

[10] J. Scarzello, D. Lenko, R. Brown, and A. Krall, SPVD: A Magnetic Vehicle Detection System using a Low Power Magnetometer, *IEEE Trans. Magnetics*, Vol. 14, 1978, pp. 574-576.

[11] T. M. Hussain, T. N. Saadawi, and S. A. Ahmed, Overhead Infrared Sensor for Monitoring Vehicular Traffic, *IEEE Trans. Vehicular Techn.*, Vol. 42, 1993, pp. 477-483.

[12] T. M. Hussain, A. M. Baig, T. N. Saadawi, and S. A. Ahmed, Infrared Pyroelectric Sensor for Detection of Vehicular Traffic using Digital Signal Processing Techniques, *IEEE Trans. Vehicular Techn.*, Vol. 44, 1995, pp. 683-689.

[13] R. A. Olson, R. L. Gustavson, R. J. Wangler, and R. E. McConnell II, Active-Infrared Overhead Vehicle Sensor, *IEEE Trans. Vehicular Techn*., Vol. 43, 1994, pp. 79-85.

[14] K. Whitehouse, Keeping an Electronic Eye on the Road, *IEEE Comp. Graphics and Appl*., Vol.15, 1995, pp. 16-17.

[15] Hyungjin Kim, Joo-Hyune Lee, Sung-Wook Kim, Jae-In Ko, and Dongil Cho, Ultrasonic Vehicle Detector for Side-Fire Implementation and Extensive Results including Harsh Conditions, *IEEE Trans. Intell. Transport. Sys.*, Vol.2, 2001, pp. 127-134.

[16] N. Furstenau, H. Horack, and W. Schmidt, Extrinsic Fabry-Perot Interferometer Fiber-Optic Microphone**,** *IEEE Trans. Instru. and Meas.*, Vol. 47, 1998, pp. 138-142.

[17] Shiping Chen; Ziping Sun; and B. Bridge, Traffic Monitoring using Digital Sound Field Mapping, *IEEE Trans Vehicular Techn.*, Vol. 50, 2001, pp. 1582-1589.

[18] C. Setchell and E. L. Dagless, Vision-Based Road-Traffic Monitoring Sensor, *IEE Proc.-Vis., Image and Sig. Proc.*, Vol. 148, 2001, pp. 78-84.

[19] C. E. Smith, C. A. Richards, S. A. Brandt, and N. P. Papanikolopoulos, Visual Tracking for Intelligent Vehicle-Highway Systems, *IEEE Trans. Vehicular Techn.*, Vol. 45, 1996, pp. 744-759.

[20] Shu-Ching Chen; Mei-Ling Shyu; S. Peeta, Chengcui Zhang, Learning-Based Spatio-Temporal Vehicle Tracking and Indexing for Transportation Multimedia Database Systems, *IEEE Trans. Intell. Transport. Sys.*, Vol. 4, 2003, pp. 154-167.

[21] D. J. Dailey, F. W. Cathey, and S. Pumrin, An Algorithm to Estimate Mean Traffic Speed using Uncalibrated Cameras**,** *IEEE Trans. Intell. Transport. Sys.*, Vol. 1, 2000, pp. 98-107.

[22] H. Veeraraghavan, O. Masoud, and N. P. Papanikolopoulos, Computer Vision Algorithms for Intersection Monitoring, *IEEE Trans. Intell. Transport. Sys.*, Vol. 4, 2003, pp.78-89.

[23] D. Saenz and J. Moya, Just-Enough-Smart Sensing Applied to Printed Circuit Board Inspection, *IEEE Trans. Automated Sci. and Engr*., (in review).

[24] D. Saenz, Dissertation, University of Texas at El Paso, 2005.

[25] C. Wehrhahn, Sex-Specific Differences in the Chasing Behavior of Houseflies (*Musca*), *Biol. Cyber.*, Vol. 32, 1979, pp. 239-241.

[26] V. Braitenberg, Patterns of Projections in the Visual System of the Fly. I. Retina-Lamina Projections, *Exp. Brain Res.*, vol. 3, 1967, pp. 271-298.

[27] K. Kirschfeld, Die Projektion der optischen Umwelt auf das Raster der Rhabdomere im Komplexauge von *Musca*, *Exp. Brain Res.*, Vol. 3, 1967, pp. 248-270.

[28] J. Scholes, The Electrical Responses of the Retinal Receptors and the Lamina in the Visual System of the Fly *Musca*, *Kybernetik*, Vol. 6, 1969, pp. 149-162.

[29] J. Smakman, J. van Hateren, and D. Stavenga, Angular Sensitivity of Blowfly Photoreceptors: Intracellular Measurements and Wave-Optical Predictions, *J. Comp. Physiol. A*, Vol. 155, 1984, pp. 239-247.