High accuracy traffic monitoring using road-side line-scan cameras

Damien Douxchamps, Benoît Macq and Kunihiro Chihara

***Abstract*— This paper presents a new technique for simultaneous traffic monitoring and law enforcement using a road-side rig head based on two line-scan cameras. As the vehicle travels across the vertical scanning lines, two images are formed and can be correlated with each other to determine the vehicle speed. Our first tests showed that the resulting videos are easily interpreted by computers in real-time and that the measures are both accurate (less than 1% error) and stable. In addition to speed, the system is able to extract a broad range of traffic information, among others the vehicle size, acceleration and inter-vehicle distances.**

I. INTRODUCTION

As the traffic density increases over the years, our society looks for ways to optimize the throughput of roads not only by using appropriate traffic management but also by making the roads safer, all of which requires a certain level of law enforcement. In that perspective, this paper presents a new sensor based on line-scan cameras that is able to measure several parameters of the traffic, among others the vehicle speed, acceleration and length. Moreover, the target accuracy of the proposed approach is less than 1% which makes it more than suitable for law enforcement applications.

Classic video monitoring does not have the resolution and framerate necessary to reach this kind of accuracy and lack in stability due to complex segmentation, among others [1] [2]. On the other hand, radars do have an acceptable accuracy for law enforcement but lack the rich output of a video system which makes them unsuitable in most traffic monitoring applications. Our technique elegantly combines the advantages of the radars with the convenience of a video system which can be easily interpreted by humans or computers; a far cry from classic video-based road monitoring.

Our idea for this new multi-purpose system is to use two line-scan cameras to scan the road perpendicularly to the axis of motion of the vehicles (Fig. 1a). The latter are scanned as they pass in front of the cameras and we can calculate the time $T_2 - T_1$ needed by a vehicle to go from the first to the second camera. Because the line rate of these cameras is very high (up to 50000 lines/sec) we can expect an accuracy that is better than current solutions.

We will start by laying out the principles behind our new line-scan approach and proving its feasibility. After a description of the experiments, a section reviews the image processing tools used to extract all information from the line sequences. The results of this process are then presented and discussed.

D. Douxchamps and K. Chihara are with the Image Processing Laboratory, Nara Institute of Science and Technology, Japan d.douxchamps@ieee.org

B. Macq is with the Laboratoire de Télécommunications et Télédétection, Université catholique de Louvain, Belgium

II. PRINCIPLE AND FEASABILITY

The setup presented in Fig. 1a consists in two line-scan cameras separated by a distance b that we call the baseline. The field of view of each camera is a vertical plane, orthogonal to both the road plane and the direction of the traffic. The beginning and end of the scan of a vehicle by each camera correspond to four key instants $T1...T4$, represented in Fig. 1b. The resulting line-scan images are shown in Fig. 1c. Note that the vehicles will appear 'stretched' or 'compressed' depending on their speed and acceleration. Given this simple geometry, the camera baseline b, the speed s of the vehicle and the time $\Delta T = T_2 - T_1$ needed by the vehicle to go from one camera to the other are related with the expression:

$$s = \frac{b}{\Delta T} = \frac{b}{n/f}, \tag{1}$$

where n is the number of lines of delay between the appearance of the vehicle on each line-scan sequence and f is the camera line rate (in Hertz). Using this simple expression we can already verify that the parameters involved are within the capabilities of current hardware. Let us suppose as a first approximation that the error on the detection of points in the image is one pixel. Since two indexes are used in the calculation of ΔT the error on its evaluation is roughly $\sqrt{2}/f$ and at least $n\sqrt{2}$ lines will be needed for a relative error of $1/n$. The relative error ϵ_s on the vehicle speed s can be estimated as a function of the baseline, the sampling rate and the vehicle speed:

$$\epsilon_s = \frac{\Delta s}{s} = \frac{\frac{bf}{n} - \frac{bf}{n+2}}{\frac{bf}{n}} = \frac{2s}{bf + 2s}, \tag{2}$$

where Δs is the absolute speed error. This leads to a required line rate of 6600Hz for a typical baseline of one meter and an accuracy of 1% at 120km/h, which is well within the capability of line-scan cameras.

The very high line rate also means that we will have a very short exposure time. To compensate for this we use a camera with very large rectangular pixels. Their surface is 40 times larger than pixels from a regular camera which increase the camera sensitivity by the same factor. Binning was also used to increase the camera response to levels compatible with outdoor lighting conditions and a line rate around 5000Hz.

III. FIELD EXPERIMENTS

Before starting the road-side trials the camera rig was calibrated offline once for all. Due to the special linear

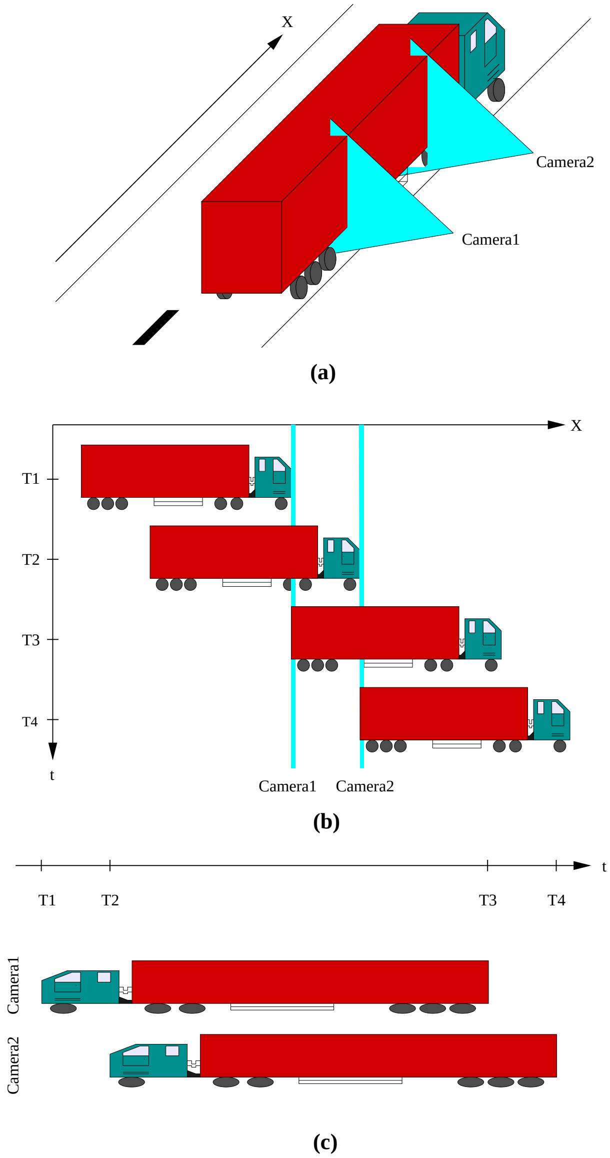

Fig. 1. Principles of the line-scan speed measurement system: (a) three dimensional view of the setup, with the scanning planes in light gray; (b) side view of the setup at entering/leaving times $T_1 ... T_4$; (c) the two image scans of the vehicle

Fig. 2. An excerpt of a typical sequence: camera 1 (top) and camera 2 (bottom).

sensors involved we have designed a specific procedure for their calibration. This aspect is not detailed here but the interested reader will find more information about it in [3]. The principal result of the calibration is the determination of the baseline and the obtention of parallel imaging planes. For our tests the baseline was measured as 593mm and the parallelism of the imaging planes is such that 10 meters away from the camera, the baseline will not have changed more than 2mm.

The calibrated line-scan monitoring rig can then be used in two different basic configurations: with the camera on the roadside or with the camera under a bridge, looking downward. (Other configurations are possible but one then looses the ability to measure the height or the width of the vehicles.) For our tests we used the lateral view for several reasons: i) it is more interesting if a classification of the vehicles is required, ii) it is easier to setup and iii) it contains more features than the top view, providing more matching opportunities because logos and window frames are mostly located on the vehicle side. The tests were performed on a small road on a sunny day with a typical distance of 3 meters between the camera and the vehicles. A resulting sequence is presented in Fig. 2. It is important to note that in line-scan images the horizontal axis represents the time (Fig. 1c). Since the line rate is high the horizontal size of the image will be large too: it stretches on 15000 pixels for a little less than three seconds of recording at a line rate of 4882.8Hz. Given the baseline calculated above and a typical speed of 70km/h at the location chosen for our tests we can use (2) to estimate the error ϵ_s at 1.3% for speed measurements and 1.8% for length estimations.

IV. Image Processing for Line-Scan Sequences

Image processing is usually a computer-intensive operation that is difficult to perform in real-time without expensive hardware or the use of simpler, often limiting algorithms. Our approach solves this quandary thanks to the degenerated view of the scene provided by the line-scan cameas which is easily interpreted without major limitations by simple, stable and real-time algorithms. Moreover, the scope of metric information inferred from the image measurements is broad and accurate.

The first step is a rough detection of the front and back boundaries of the vehicles which is performed with a background comparison in the sequence of lines. Since we use a linear sensor the reference background is actually a single line which is obtained regularly as the average of a bundle of typically a hundred lines. As can be seen in Fig. 2 the background is very stable across the whole vehicle length. Indeed, on the contrary to real images where such long averaging of data is not possible, line-scan cameras operate at a much higher frequency so that the 'large' average on 100 lines has a temporal length of only a few hundredth of a second. The detection will therefore not be subject to problems like illumination changes or other slow background changes because such events happen on a much slower time scale [4]. The only significant issue is the shadow that is detected as part of the vehicle. To eliminate it in the future we plan to limit the background detection to a upper part of

Fig. 3. The detected boundaries of the vehicle. Overestimation of the length due to shadow is clearly visible in the front of the vehicle.

the image as well as lower the camera closer to the ground. An example of detected boundaries is shown in Fig. 3.

From the vehicle boundaries in the two images we can obtain a first estimation of the vehicle speed S_v, length L_v and acceleration A_v using (1):

$$S_v = bf/(B_f^2 - B_f^1), \quad L_v = S_v(B_r^1 - B_f^1)/f \tag{3}$$

$$A_v = \frac{1}{B_r^1 - B_f^1}\left(\frac{b}{B_r^2 - B_r^1} - \frac{b}{B_f^2 - B_f^1}\right) \tag{4}$$

where B_f^i and B_r^i are respectively the front and rear line indexes of a vehicle in camera i.

The first speed estimation that we just obtained above is not accurate and may show instabilities because it is based on a single boundary detection. To obtain denser measurements we compute the whole speed profile of the vehicle along the horizontal axis (which represents the time). This time-evolution will allow us to obtain more stable measures.

As (3) shows, the speed depends on matching features in both images. To this effect we use a subpixel Block Matching Algorithm (BMA) to compute the optical flow between the two scans of the vehicle [5] [6]. Given the very similar images obtained from the two cameras we can use the simple Sum of Absolute Difference (SAD) criteria to find the vehicle speed for a specific position along its profile. Moreover the displacement is purely horizontal which limits the complexity to a fast unidimentional search using large vertical image blocks. We can see that the degenerated view of the line-scan cameras considerably simplifies the processing of the sequences. We also use a modified Moravec operator [7] to compute the BMA only when a block contains enough horizontal variance to further enhance the speed of the algorithm and avoid spurious estimations. From the speed profile we obtain a more stable measure of the initial vehicle speed and its acceleration through the use of a linear regression [8]. The standard deviation from this fit is a measure of the error and is shown in Table I. This speed estimation is then used to infer a better measure of the length which will then include a compensation for the vehicle acceleration.

Besides the speed, acceleration and length, a line-scan system can provide other extra information like the lateral position of the vehicle on the road, its height and the inter-vehicle distance. The two first ones are straightforward but require to calibrate the camera on site which was not performed for these first tests. The enforcement of the minimum inter-vehicle distance is quite difficult today especially in the extreme cases where vehicles follow each other so closely that they will appear as one vehicle for most monitoring systems. This typically happens with area-scan cameras but line-scan cameras are able to differentiate the vehicles and measure their distance thanks to their high speed and proper orientation with respect to the road.

Suppose that two vehicles have been detected, one after the other, and that their respective speeds $S_{v,1}$ and $S_{v,2}$ have been measured. The delay between their appearance in the first camera is

$$\Delta T = \left(B_{f,2}^1 - B_{r,1}^1\right)/f \tag{5}$$

where the second subscript indicates the vehicle index. It leads to the following estimation of their distance D:

$$D = \Delta T\left(\frac{S_{v,1} + S_{v,2}}{2}\right) \tag{6}$$

V. RESULTS

The results for a few selected vehicles are shown in Fig. 4 to Fig. 6. Table I summarises some important measurements performed on the vehicles. A known test vehicle passed a few times in front of our setup and in that case an indication of its speed or acceleration is shown in the table.

For the test shown in Fig 2 the corresponding metric speed S_v is 63.3km/h, which is below the 70km/h displayed by the vehicle meter but still plausible given the tolerance of the latter. As we do not have a ground truth measure of the speed we will use the length measure as an indicator of the quality of our results. The true length of the vehicle is available from the manufacturer, itself determined by visual inspection of the line-scan video. The error that we will calculate from this true length will be an upper boundary of the speed error because the length estimation depends on the speed (see (3)). The full vehicle length cannot be used, however, because it is subject to shadow errors: the estimated length is 4.66m while the true value is 4.39m. To get rid of the shadow effect we use the more recognisable wheelbase as test feature. It was measured as 2.64m which is very close to the true value of 2.61m and within the expected accuracy found in section III.

A first observation is that the speed profile sometimes shows a decreasing value on the windshield of the car, as in Fig. 6 between lines 600 and 750. The reason behind this problem lies in the large difference between the horizontal and vertical image resolution. Indeed, a vertical misalignement of the cameras of one pixel will yield a much greater error in displacement estimation when the features are slanted. To limit this effect we allow the block matching to perform a limited vertical scan but this does not cancel the error entirely. A better calibration and a higher vertical image resolution would be needed to tackle this issue.

Even with this problem in mind, the results of table I are consistant with our expected accuracy. They also show a very good repeatability: four cars have a true length close to 3.83m and all measured lengths are very close to 4.04m. The test car, for which the in-car counter speed or its tendency is indicated, is 4.39m long and was also measured with

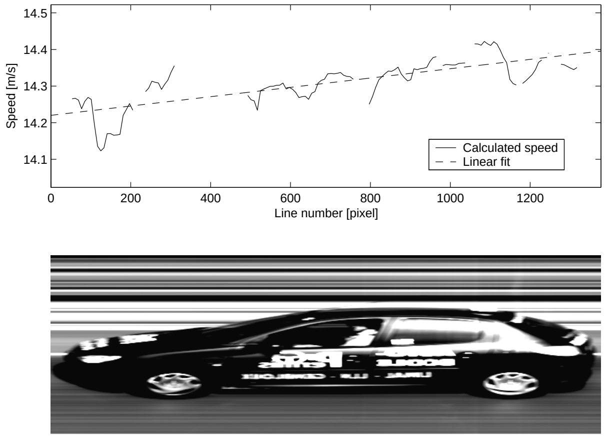

Fig. 4. Results for car number 6676

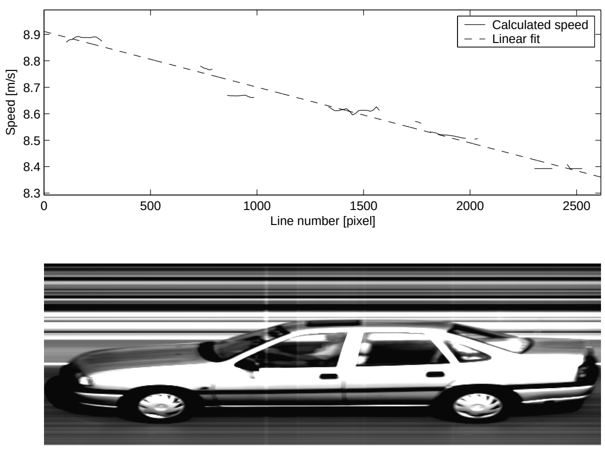

Fig. 5. Results for car number 64991

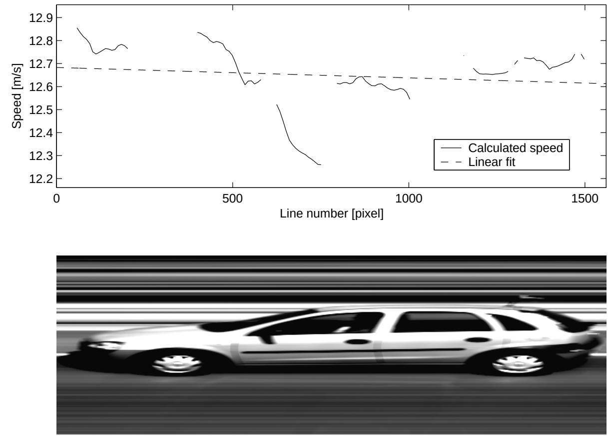

Fig. 6. Results for car number 68407

Car ID	Speed $[km/h]$	Accel. $[m/s^2]$	Size $[m]$	Counter speed $[km/h]$	True length $[m]$	Rel.fit error [%]
2123	62.1	0.44	4.04		3.84	0.63
6676	51.5	0.62	4.04		3.84	0.22
11874	54.7	0.58	3.65		3.54	0.54
25774	39.3	-1.42	4.56	breaking	4.39	0.35
27896	63.1	-0.09	4.63	70	4.39	0.41
37050	59.4	0.85	4.36		4.15	0.49
38416	62.8	-0.39	4.63	70	4.39	0.46
50580	59.5	0.24	5.05		4.73	0.82
61176	31.8	-1.38	4.59	breaking	4.39	0.60
64991	31.1	-1.03	4.64	breaking	4.39	0.23
68407	45.5	-0.22	4.04		3.82	0.81
74948	65.1	0.29	4.03		3.83	0.79
89378	58.5	-0.02	4.85		4.51	0.76
94881	41.9	-0.09	5.89		5.40	0.63
97647	62.6	-0.83	4.63	70	4.39	0.42

TABLE I

METRIC RESULTS FOR A FEW CARS

consistent results: all tests involving a constant speed yield the same length of 4.63. Due to the higher error in the estimation of the acceleration the results from braking tests are not as accurate but still within 1%.

VI. CONCLUSIONS

We have presented a new system for traffic monitoring and law enforcement. The approach we took of using a geometrically constrained camera setup pays off because the degenerated view of the scene makes it much easier to interpret the vehicles recordings. The system has several advantages over currently existing systems. First it is able to provide much more than speed and a visual output is available. The measures are also more accurate than existing techniques and can be verified by hand in the case of a legal challenge. At last, the proposed solution can run in real-time on a laptop, is cost effective and can even monitor multiple lanes simultaneously if the traffic is not too dense. A sensor could be mounted over each lane to scan it vertically in the case of dense traffic on wide highways.

VII. ACKNOWLEDGMENTS

The authors wish to thank Hervé Capart for his insightful comments on this manuscript, the students of UCL who performed early research on this topic and Toshiyuki Umeda for assistance during the field tests.

REFERENCES

[1] P. Piscaglia, A. Cavallaro, M. Bonnet, and D. Douxchamps, “High level description of video surveillance sequences,” in *Proc. European Conf. on Multimedia Applications, Services and Techniques (ECMAST’99)*, Madrid, Spain, May 1999, pp. 316–331.

[2] B. Abreu, L. Botelho, A. Cavallaro, D. Douxchamps, P. Figueiredo, B. Macq, *et al.*, “Video-based multi-agent traffic surveillance system,” in *Proc. Intelligent Vehicles Conf. (IV’00)*, Dearborn, MI, Oct. 2000.

[3] D. Douxchamps, “Multidimensional photogrammetry of short-lived events,” Thèse de Doctorat, Université catholique de Louvain, Louvain-la-Neuve, Belgium, Oct. 2004.

[4] A. Cavallaro, D. Douxchamps, T. Ebrahimi, and B.Macq, “Segmenting moving objects: the MODEST video object kernel,” in *Proc. Workshop on Image Analysis For Multimedia Interactive Services (WIAMIS’01)*, Tampere, Finland, May 2001.

[5] F. Devernay, “Vision stéréoscopique et propriétés différentielles des surfaces,” PhD Thesis, INRIA, Sophia Antipolis, 1997.

[6] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, Cambridge, Massachusetts, 1993.

[7] H. P. Moravec, “Towards automatic visual obstacle avoidance,” in *Proc. Intl. Joint Conf. on Artificial Intelligence*, Cambridge, MA, Jan. 1977, p. 584.

[8] W. H.Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipies*. Cambridge University Press, 1989.