Reprinted from SPIE Vol. 848—Intelligent Robots and Computer Vision
Sixth in a Series

TRAFFIC SPATIAL MEASUREMENTS USING VIDEO IMAGE PROCESSING

Application of mathematical morphology to vehicles detection

S. BEUCHER*, J.M. BLOSSEVILLE**, F. LENOIR**

* Centre de Morphologie Mathématique, Ecole des Mines de Paris
(CMM/EMP), Fontainebleau - France
** Institut National de Recherche sur les Transports et leur Sécurité
(INRETS), Arcueil - France

ABSTRACT

Algorithms of video traffic image processing, presented below, have been developed by CMM and INRETS using 256x256x6 bits images collected from various scenes covering a 150 m area of a freeway under various weather conditions.

First, road detection is automatically performed. The road and traffic lanes images are used to derive relationships between real and image distances and to build image transformations independant from the perspective view.

Then, markers of the vehicles are extracted using geometrical adapted filters. These markers are joined to define a single marker for each vehicle.

Finally, vehicles trajectories are built and traffic variables are specified.

1. INTRODUCTION

In certain conditions, images of a traffic scene give to the traffic engineer a lot of valuable informations. Due to its efficiency and flexibility video image analysis can be used in the three main domains of traffic management, that is measuring, monitoring and control. Moreover, this technique permits spatial measurements not available by classical means. Finally, portability, and ease of installation are substantial advantages.

On the other hand the use of video imaging raises a lot of problems, in particular perspectives distortion, variations of illumination, great variability in the shape of the vehicles. Furthermore it is necessary to collect and to analyse successive images with high frequency in order to obtain accurate measurements.

Despite the fact that many researchers have been involved in that problem, there exists currently no system working correctly under all the various traffic conditions. We shall present below the main results of a two years study performed in collaboration between the French National Research Institute of Transports and their Security (INRETS) and the Mathematical Morphology Center of Paris School of Mines (CMM/EMP).

2. GOALS

The presented algorithms have been developed using motorway traffic images which have been collected with a video recorder and then digitized (256x256x6 bits). We considered various traffic scenes under daylight conditions for various weather conditions, each scene covering a 100-200 m area of a three lanes freeway. Pictures can be front or rear views. In a first approach, moderately dense traffic has been analysed. In order to use already installed cameras, we prohibited the use of more sophisticated sensors than CCD or infra-red cameras. On this connection, we wanted to develop a system able to work with remote controlled cameras. This means that the image of the same scene may change if an operator has moved the camera in the meantime for incident detection purpose for instance.

The image analysis algorithms use widely the tools of the mathematical morphology. This image processing methodology provides quick, efficient, and flexible means of treatment. Furthermore, this treatment can be easily implemented in hardware processors [1].

Three main steps can be defined in the process. In the first one (lanes detection), we try to automatically select the region of interest of the scene, composed of the different lanes of the road. Then, the vehicles are detected in a second step (vehicles detection). In that step, images are individually considered, our foremost aim being to detect each vehicle and to associate to it a single marker. In the last step (tracking), trajectories of all the detected vehicles are built lane by lane and traffic measurements are derived from them.

3. LANES DETECTION AND REAL DISTANCES COMPUTATION

The interest of the lanes detection is twofold. First, we focus the process on the region of interest of the scene. This enables a faster processing, prevent from false vehicle detections and makes the trajectories computation quick and easy. The second interest is to take into account the image distortion due to the perspective view. As a matter of fact, on the image provided by a camera covering few hundred meters of road, the apparent size of the vehicles varies greatly between the foreground and the background. So, it is compulsory to compensate this distortion by computing for every pixel of the image a size of transformation related to the distance between the camera and the corresponding point in the real scene.

3.1. Lanes detection

All this procedure uses a generated image called difference image (D-image). This image is computed by averaging the differences of successive images :

$$\text{D-image} = 1/n \sum_{i=1}^{n} |I_i - I_{i-1}|$$

where I_i is the current image at time i and n the number of images chosen in the range [80-100].

This provides a picture of the moving parts of the scene (Fig. 1). D-image is filtered by morphological opening transformations (Fig. 2) and thresholded.

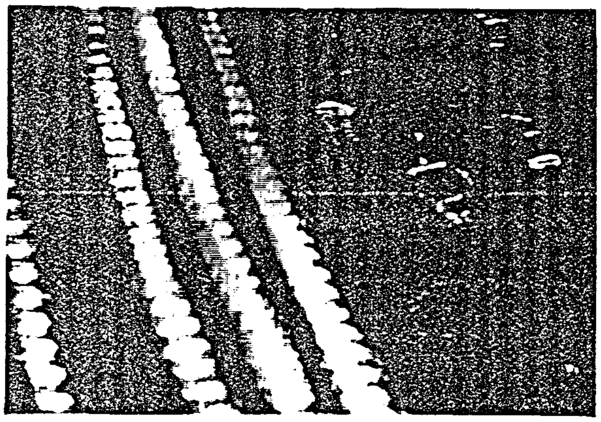

Figure 1. Difference image (D-image)

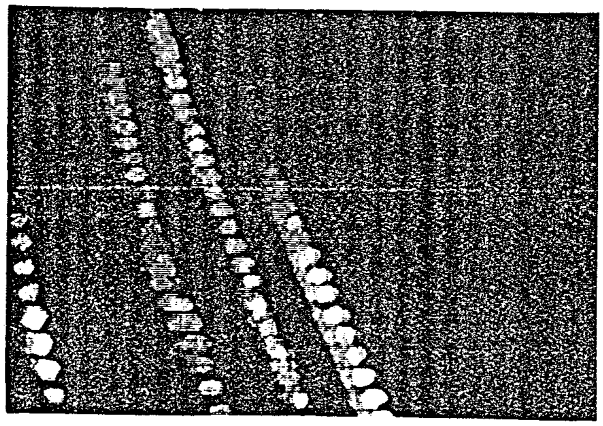

Figure 2. Filtered D-image

The resulting binary image is then processed : erosion and picture reconstruction for small features elimination, holes closing, the final result being a binary picture where every lane is marked by a simply connected component : the traffic areas (Fig. 3).

Then, the image of the outside of the road is built by dilating the previous one and addition of the two pictures gives an image where every important part of the scene is simply marked : one marker per lane and one marker per sideway area (Fig. 4). This image and the morphological transform called skeleton by zones of influence (SKIZ) are finally used to entirely partition the scene (Fig. 5).

Sometimes, more refined algorithms may be used especially when the ground layout of the road is available. In that case, we perform a SKIZ conditioned by the image of the ground layout derived from the following average image (Fig. 6) :

$$\text{A-image} = 1/n \sum_{i=1}^{n} I_i$$

A binary image of the ground layout is obtained from the A-image by using the morphological Top-hat transformation followed by a thresholding (Fig. 7). This more sophisticated procedure leads to a better positionning of the traffic lanes (Fig. 8).

Figure 3. Lane markers

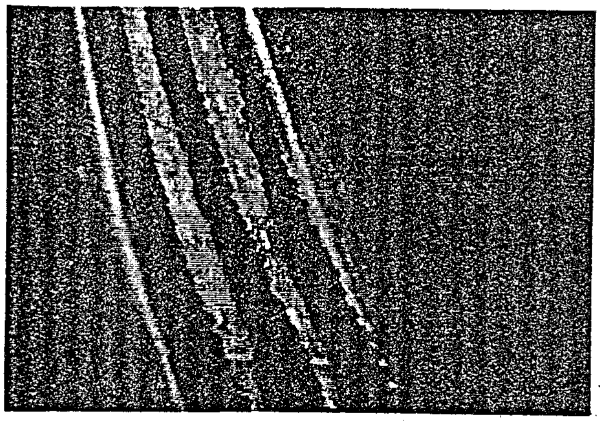

Figure 4. Lane and outside markers

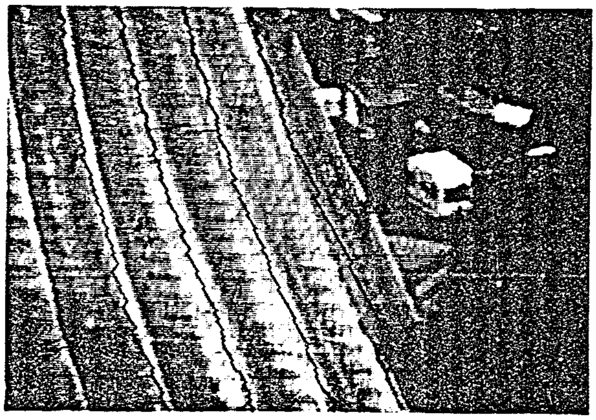

Figure 5. Scene segmentation

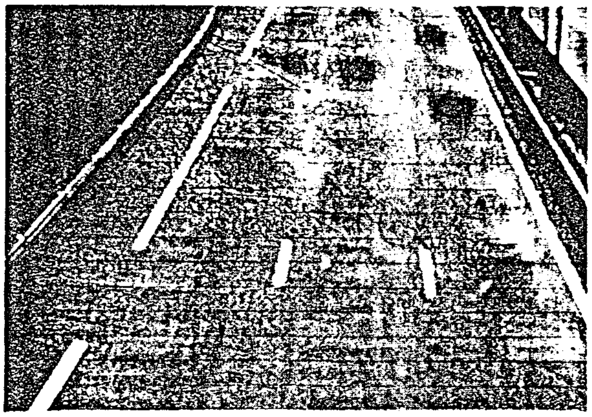

Figure 6. Average image (A-image)

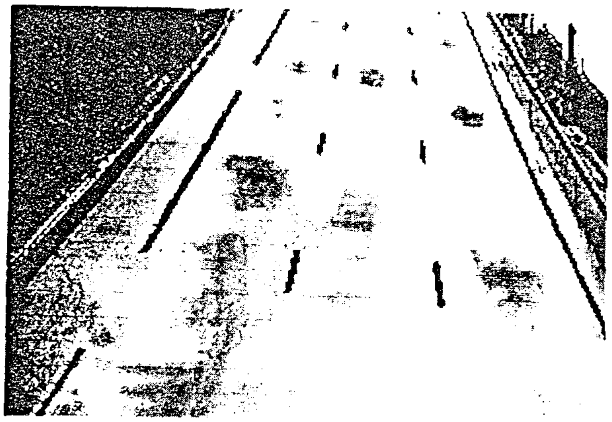

Figure 7. Ground layout detection

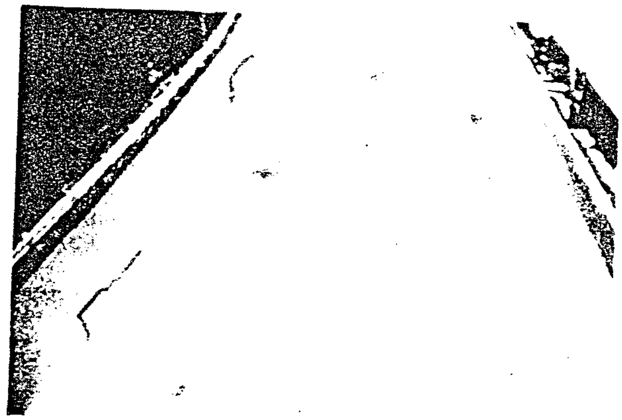

Figure 8. Scene segmentation final result

3.2. Real distances computation

The real distances computation uses a scale factor. This scale factor associates to every pixel of the image and to its neighbourhood the real distances of the corresponding points in the scene :

$$d_p(x_i,x_j) = e_{ij} \quad d_r(x_i,x_j)$$

d_p and d_r are respectively the distances between the two adjacent pixels x_i and x_j and the two corresponding points of the road x_i, x_j...

For the sake of simplicity, this scale factor which varies with the direction in the image, is calculated in the horizontal and vertical directions. Moreover, these two scale factors e_h and e_v are supposed to only depend on the vertical coordinate of the point in the image. They are given by the following formulas (Fig. 9) :

$$e_v(x_i) = \frac{|[L^2\ (x_1 - x_2)^2 - (\ell_2 - \ell_1)^2\ h^2]|^{1/2}[\ell_2(x_1 - x_i) - \ell_1\ (x_2 - x_i)]^2}{(x_1 - x_2)^2\ (x_1\ell_2 - x_2\ell_1)\ L^2}$$

$$e_h(x_i) = \frac{[\ell_2(x_1 - x_i) - \ell_1(x_2 - x_i)]}{(x_1 - x_2)L}$$

The parameters h and L are the height of the camera and the width of the road respectively. Other parameters can be easily calculated from the lanes image.

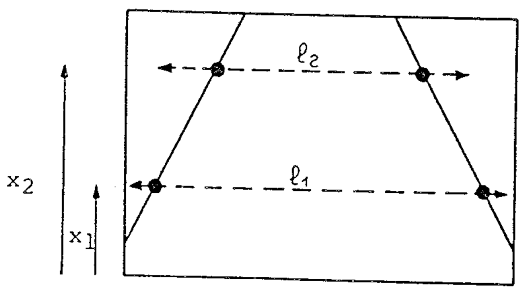

Figure 9. Perspective view parameters

These scale factors are used to segment the region of interest of the image into different strips such that in each one there corresponds a fixed image distance expressed in pixels to a given real distance in meters. The distance variation in the image between adjacent strips is equal to one pixel. Thus, all transformations involving the size of the objects to be detected can be defined in terms of real distances, the "true" transforms being performed in the different zones of the image using the corresponding sizes in pixels. These transformations are defined by means of a generalization of the morphological geodesic operators [2].

4. VEHICLES DETECTION

The vehicles detection is the result of the analysis of the scene one picture at a time. This way of doing, that is not to take into account the movements of the objects, seems to give better results especially for incident or accident detection, situations where vehicles are not moving any more.

Vehicles are detected by use of visual and geometrical features which may exist in the picture. Vehicles are made of a collection of dark and white zones (front shadow and radiator gril, hood, windscreen, roof, etc...). Among them, the front shadow is the most important and permanent feature. For that reason, the vehicles detection process always starts with the extraction of this front zone. This feature being a shadow, a local grey-tone minimum of the image [3]. Thus, these minima are extracted and filtered : those which are wide enough (1 m is the lower limit) and dark enough (darker than the average grey value of the road) are retained and thresholded (Fig. 10).

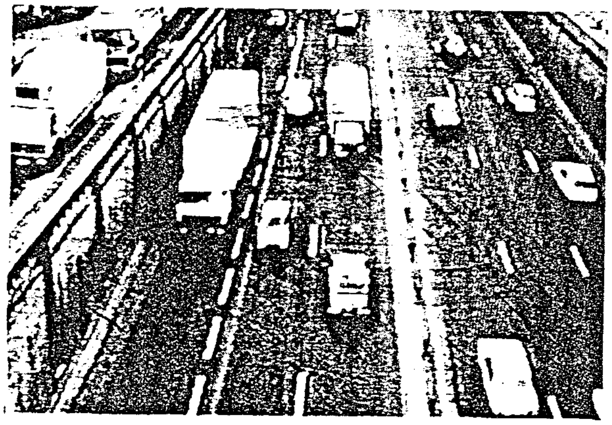

Figure 10. (a) Original image

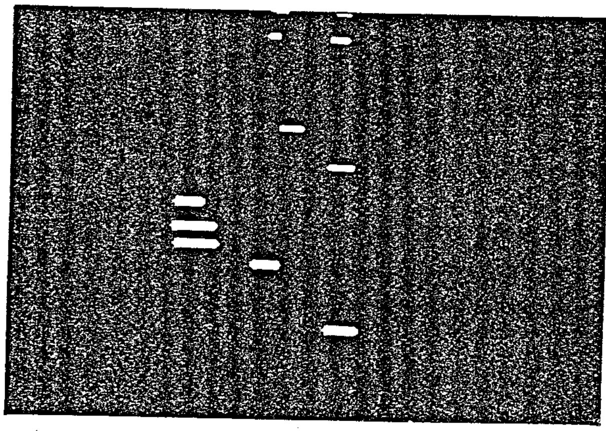

Figure 10. (b) Detected dark markers

The white features are then detected since they correspond always to geometrically well-shaped objects (roofs and hoods), they are revealed by a Top-hat transform. Here again this image is thresholded and a filtering procedure using linear openings eliminates every too small and to narrow components (Fig. 11).

These two procedures provide, most of the time, more than one marker associated to each vehicle. Hence, the third part of the vehicles detection algorithm consists in joining these various markers. To achieve this, vertical upward dilations of the dark marker are performed. The dilation size is approximately equal to the length of a vehicle. White markers not reached by this process are then joined in a second step in vertical downward dilations of greater size (Fig. 12, 13). The final result is a single marker associated to each vehicle. This unique marker is placed in front of the vehicle (Fig. 14).

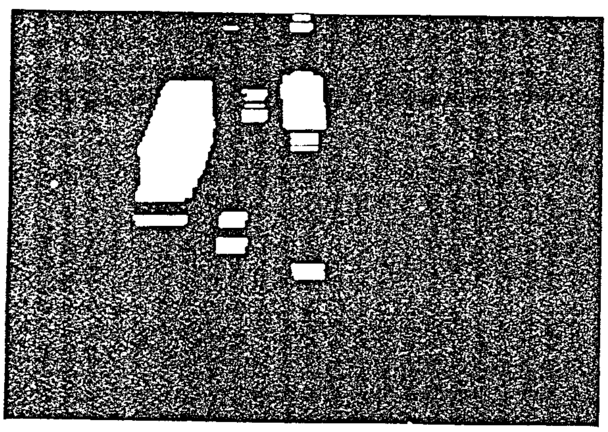

Figure 11. White filtered markers

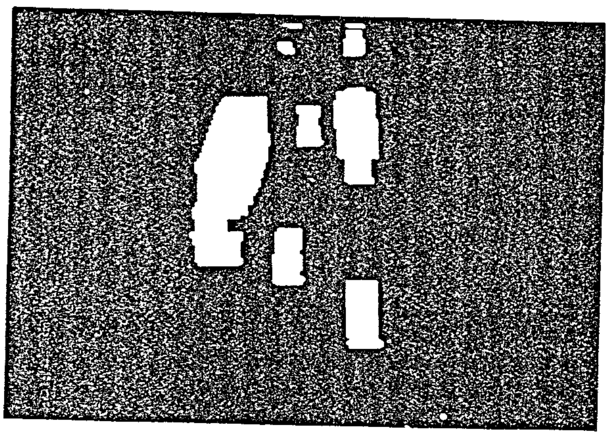

Figure 12. Markers concatenation

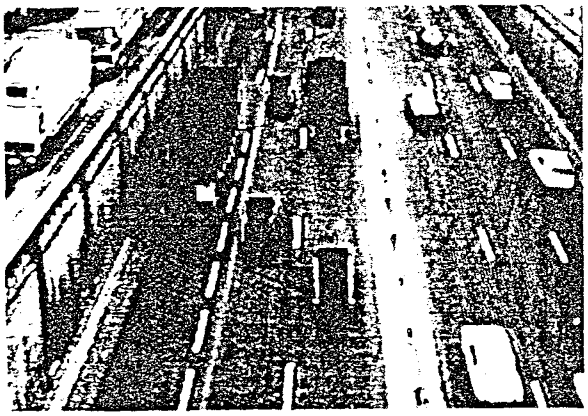

Figure 13. Markers concatenation (black-zones) and the original image

Figure 14. Detected vehicles (marked with a white token

5. VEHICLES TRACKING

The above markers are used, lane by lane, to build vehicles trajectories. The position of each vehicle is plotted, time versus road axis in a 2D-map (Fig. 15).

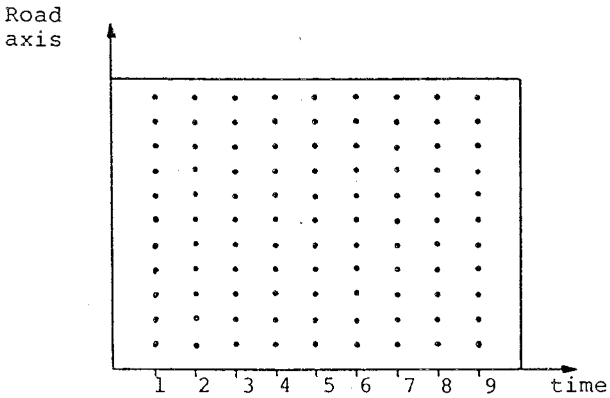

Figure 15. 2D-mapped markers

The various points are joined to form the time trajectory of each vehicle. The concatenation algorithm is predictive : using the average speed of the vehicles in the scene calculated from their previous positions, every point at time t_i is joined to the best fitting point at time t_{i+1} (Fig. 16).

This time versus position representation of the markers and the drawing of the trajectories increase the accuracy of vehicles detection, by eliminating false markers or, on the other hand by joining markers using a larger interval of time, despite the fact that a marker may have been lost.

Doing so, some particular events (vehicle moving from its traffic lane, standstill vehicle, etc...) may be recognized.

Finally, many traffic variables may be derived from that mapping : volume of traffic, vehicles concentration, speed (instant or mean) of each vehicle, average speed of the traffic in the scene, percentage of the road occupied by the vehicles, and so on.

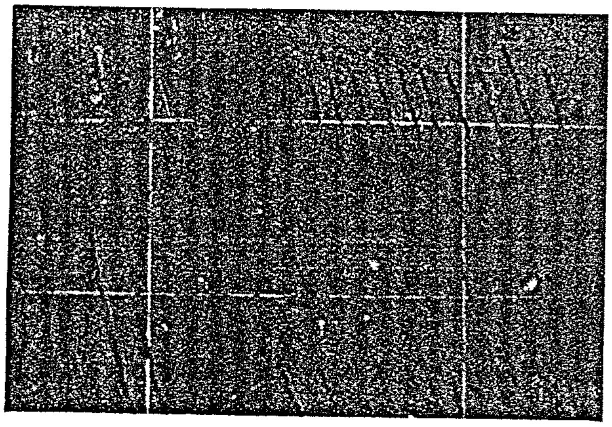

Figure 16. Built trajectories (forty frames)

6. RESULTS

These image processing algorithms have been tested on various traffic situations. More than ten sites were analysed with more than a hundred images in each, images being collected at a rate of four per second. Figures 17 to 22 illustrate some of these scenes and the result of the detection.

The average vehicle false detection percentage of the method (calculated on more than twenty thousand vehicles) is less than 3 % after the second phase (vehicles detections) and falls down close to 0 % after the third phase (vehicles tracking), this last step allowing the complete elimination of false detections.

7. CONCLUSION

Presented method gives good results in medium traffic density. But it is necessary to adapt it for high density conditions. Since now results are good enough to lead INRETS and CMM to build up a real time traffic sensor. These two subjects : detection in high density conditions and development of a real time processor are both our new objectives.

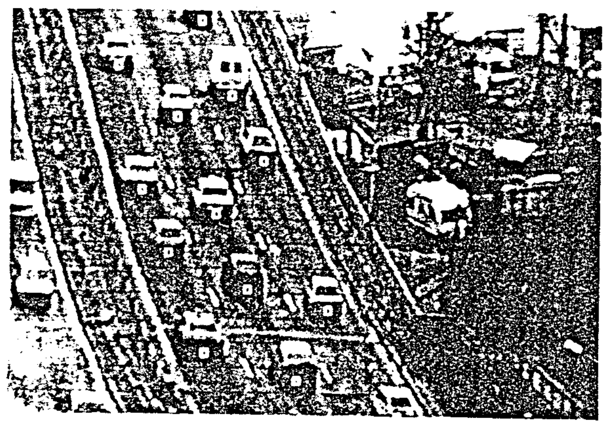

Figure 17. View of a 110 meters stretch of a 3 lanes freeway from upstream the camera is 30 meters high and the traffic density is a medium one

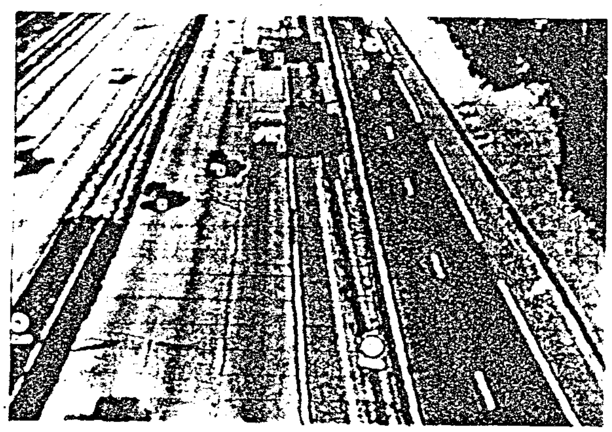

Figure 18. Only the 3 main lanes are considered Cars and trucks are well detected in spite of lateral shadow

Figure 19. Successful detections of different types of trucks and cars on the 3 left lanes

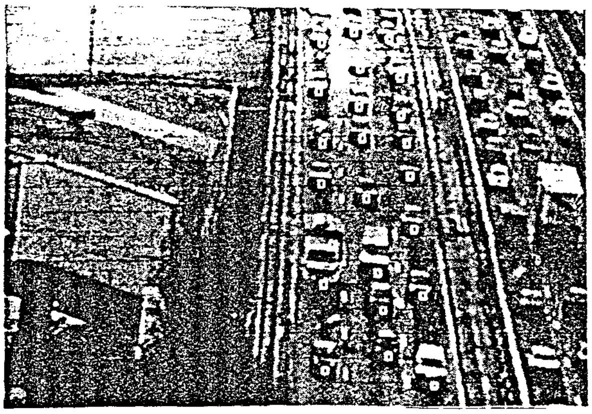

Figure 20. Wide area (220 m) and a rather high density of traffic

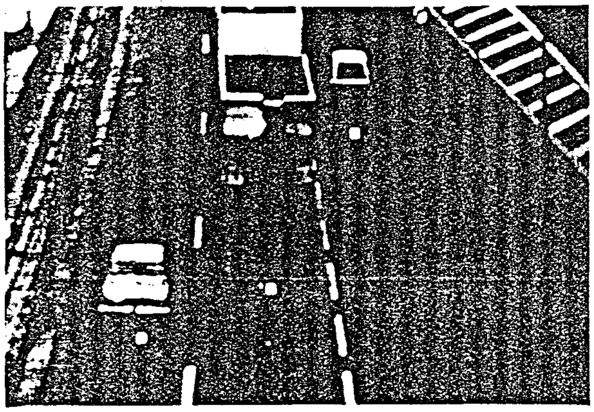

Figure 21. Successful detection in spite of a very bad original image

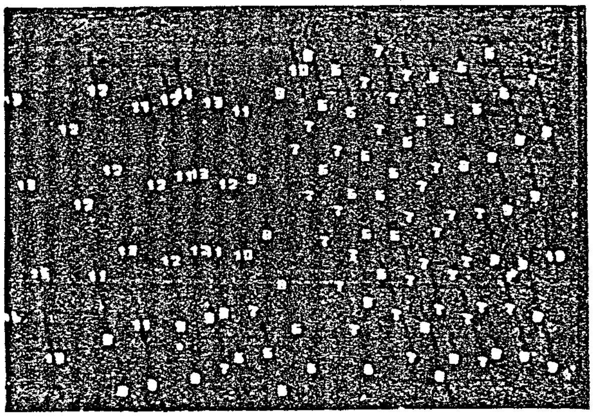

Figure 22. 2D-mapped (time versus axis road) trajectories. The instantaneous speeds (meters per second) are showed on each detection point

8. REFERENCES

1. J. SERRA, Image analysis and Mathematical Morphology, Academic Press, London (1982)
2. S. BEUCHER, C. LANTUEJOUL : On the use of geodesic metric in image analysis, Journal of Microscopy, Vol. 121, Part 1 (Jan. 1981)
3. S. BEUCHER, J. SERRA : Shapes and patterns of microstructures considered as grey-tone functions, Proc. of the 3rd European Symposium on Stereology, Ljubljana, Yougoslavie, 3 suppl. 1, 43 (1981)