

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number:

0 403 193
A2

(12)

EUROPEAN PATENT APPLICATION

(21) Application number: 90306317.0

(51) Int. Cl.5: G08G 1/04

(22) Date of filing: 11.06.90

(30) Priority: 16.06.89 GB 8913946

(43) Date of publication of application:
19.12.90 Bulletin 90/51(84) Designated Contracting States:
AT BE CH DE ES FR GB GR IT LI NL(71) Applicant: UNIVERSITY COLLEGE LONDON
Gower Street
London WC1E 6BT(GB)(72) Inventor: Hoose, Neil
2, Heron Court, Bishops Stortford
Hertfordshire CM23 2AY(GB)(74) Representative: Boon, Graham Anthony et al
Elkington and Fife Beacon House 113
Kingsway
London, WC2B 6PP(GB)

(54) Method and apparatus for traffic monitoring.

(57) A method of traffic monitoring is described which comprises forming at least first and second scene images of a scene in which traffic may be present, the images being formed at instants of time separated by a time interval. At least one of the first and second scene images is processed to form an edge image representing the occurrence of edges in the scene, and a determination is made on the basis of the edge image of the presence or absence, and spatial location, of traffic in the scene. A difference image is formed in which each pixel represents the difference between the intensity of the pixels of the first and second scene images at the corresponding point in the image, and a determination is made from the distribution of pixels of different intensities in the difference image of the presence or absence of movement in the scene.

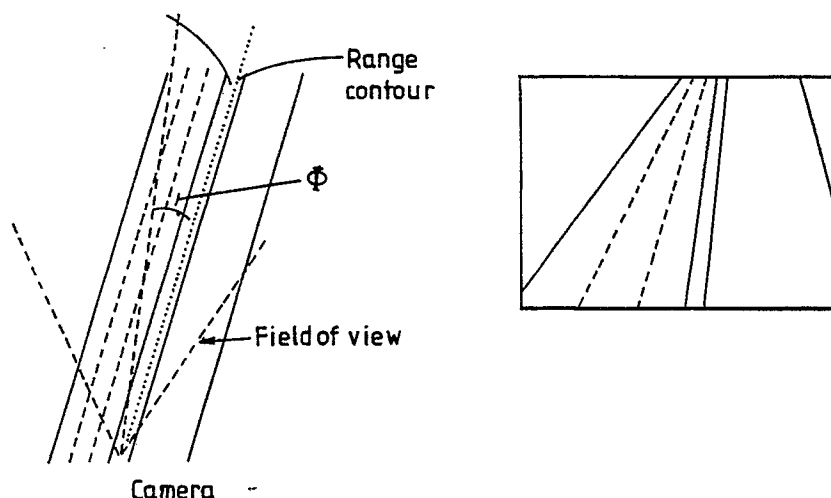


Fig.1.

EP 0 403 193 A2

Method and apparatus for traffic monitoring

This invention relates to a method of traffic monitoring.

Closed circuit television systems are increasingly being installed to allow the monitoring of critical sections of a highway. In some cases these act in tandem with a loop based incident detection system as in the HIOCC system described by J.F. Collins in "Automatic incident detection - experience with TRRL
5 algorithm HIOCC", TRRL, U.K. 1983, Supplementary report 775.

The monitoring of the pictures from the cameras by a human operator is both costly and, for the person employed, tedious. There is also a limit to the number of cameras any one person can sensibly be expected to deal with. Therefore, the use of a computer to monitor camera images and draw an operator's attention to a particular scene when an abnormality in traffic behaviour is detected would have considerable
10 utility and it is an object of the present invention to provide a method and apparatus which uses a computer for this purpose.

When human observers are faced with traffic scenes it is apparent that they can describe the state of the traffic without actually measuring any traffic parameters, e.g. they can say that there is congestion without counting the number of vehicles. An object of this invention is to mimic that ability in a computer,
15 that is, to provide an procedure whereby some description, albeit qualitative, of the current traffic state over the image can be obtained directly from image data.

A method which has that ability to some extent is described in the following paper: Hoose N. (1989), "Queue detection using computer image processing", 2nd International Conference on Road Traffic Monitoring, Institution of Electrical Engineers Conference Publication 229, London, UK.

The present inventor there described a method of traffic monitoring which comprises forming at least
20 first and second scene images of a scene in which traffic may be present, the images being formed at instants of time separated by a time interval, each scene image being an array of pixels, processing at least one of the first and second scene images to form an edge image representing the occurrence of edges in the scene, determining on the basis of the said edge image the presence or absence, and spatial location,
25 of traffic in the scene, forming a difference image in which each pixel represents the difference between the intensity of the pixels of the first and second images at the corresponding point in the image, and determining from the distribution of pixels of different intensities in the difference image the presence or absence of movement in the scene, wherein the edge image and difference image are each subdivided into an array of cells, each edge cell and its related difference image cell corresponding to a given sub-area of
30 the image constituting a scene image cell, the presence or absence of traffic and the presence or absence of movement being separately determined for each scene image cell.

However, the output from the aforesaid method is difficult to interpret directly, since it does not correspond with descriptions of traffic behaviour which would be given by an ordinary observer. According to the present invention, there is provided a method as aforesaid, characterised in that each of a plurality of
35 scene image cells lying along the image of the line of a traffic lane is analysed as aforesaid, and the presence of predetermined traffic objects detected on the basis thereof.

The above time interval between successive scene images should be short, preferably not more than 0.2 seconds.

It is to be understood that the difference and edge images mentioned above need not be, and generally
40 will not be, visual images, but rather representations of such images in electronic form.

Preferably images of the scene are repeatedly formed, for example by a video camera, at short intervals of, say, 80 ms, each image being examined for edges and each image being compared with its predecessor to form a difference image.

As already indicated, each of the edge and difference images is divided into a plurality of cells, and
45 analysis is carried out separately for each cell. For example, a video camera image 512 x 512 pixels in size can be subdivided into square cells each 64 x 64 pixels in size. Although a rectangular grid of cells can be used, it is preferred to use a non-rectangular grid of cells, with the individual cells being of varying size depending on the position of cell in the image. In this way the cells can be arranged along lines which reflect the direction within the image of a traffic lane along which the traffic being monitored is travelling,
50 and the variation in the size of the cells can take account of the fact that the area of image required to represent a given area of the traffic lane decreases with increasing distance of the part of the lane concerned from the video camera or other image forming device. Thus, the variation in the size of the cells can be such that each cell represents, at least approximately, the same area of lane.

Preferably, cells are only analysed if more than a given percentage, say 50%, of their area represents road, as opposed to non-road parts of the image.

The cells themselves need not be square as suggested above, but may be rectangular or even some shape such as trapezoidal, a trapezoidal shape being advantageous in terms of subdividing the image of a lane, which image itself is trapezoidal except in the case of a lane perpendicular to the axis of vision of the camera.

5 The above mentioned traffic objects preferably include at least gaps (a gap is a cell or a succession of cells along the line of a traffic lane in each of which no traffic is detected), and/or platoons (a platoon is a cell or a succession of cells along the line of traffic lane in which in each of which moving traffic is detected), and/or blocks (a block is a cell or a succession of cells along the line of a traffic lane in each of which stationary traffic is detected). Optionally, the images may also be examined for the existence of other
10 traffic objects, as is explained in more detail below.

The invention will now be explained in more detail with reference to the accompanying drawings, in which:

Figure 1 shows a typical camera position, and its associated view of the road, for the method described herein;

15 Figure 2 shows the geometry for the calculation of occlusion values;

Figure 3 is a graph of the inter-vehicle gap at which occlusion will start to occur for a vehicle of high h, at a distance Y from a camera viewing the road from a height H;

Figure 4 shows two histograms derived from an edge image;

Figure 5 shows graphically the difference image intensity produced by movement; and

20 Figure 6 shows two histograms derived from a difference image.

Figure 7 is a graphical representation of the movement parameters and edge parameters of a single cell over a period of time;

Figure 8 shows a number of traffic "objects";

25 Figure 8a is a graph of the weight (or importance) of each of several traffic objects, plotted against the length of the object;

Figure 9 shows diagrammatically the procedure of the invention first with the image divided into a rectangular grid of cells and then with the image divided into cells whose size represents an equal area of traffic lane;

30 Figure 9a shows cell maps which may be used to allow for vehicles driving on either side of the road and in either direction;

Figures 10a and 10b are a more detailed flow diagram of the procedure of the invention, using cells; and

Figure 11 shows diagrammatically a hardware configuration which may be used.

35 The key factor in a vision based system is that a sensor (e.g. a video camera) collects data along a length of road. The actual length over which a digitised video image can be analyzed depends upon several factors such as the relative position of the camera and the road, the camera's field of view and the size of the vehicles within the scene.

When a foreground object in an image partially or totally obscures another object in the image then occlusion is said to be occurring. Occlusion can be considered to be the shadow area of a vehicle that
40 position. The amount of occlusion, or the size of the shadow area (see Figure 2), varies with the size of the vehicle and its relative position to the camera. A large, tall vehicle will produce a large amount of occlusion when compared to a 1.5 Ton van. However, for a camera looking along a length of road the amount of occlusion caused by the latter will increase as it moves further away from the camera. The values for Y_{oc} , the distance at which partial occlusion occurs, will vary with the camera height, the traffic conditions and vehicle mix. Higher camera positions will increase this value. Although camera height has little effect on
45 range the amount of occlusion is clearly dependent on the relative height of the camera compared to the height of the vehicles. This is shown in Figure 3. The value of g_v , the inter vehicle gap at which occlusion starts to occur, increases with increasing vehicle height, h, and lower camera heights H. For example, at a distance of $Y = 50$ metres from the camera a 3.5 metres high vehicle (e.g. a truck) will start to occlude
50 another vehicle if the gap between them, g_v , is less than 38 metres, when viewed from a camera mounted 8 metres above the road. For a single camera position, as traffic volume increases the inter-vehicle gaps will tend to reduce and hence Y_{oc} will reduce. A high proportion of large trucks within a traffic stream will also shorten Y_{oc} and lead to the occurrence of total occlusion of smaller vehicles. Because of this effect it will then be difficult to track vehicles individually as merging will occur. This is one reason why, as explained in
55 more detail below, it is advantageous in the analysis of the traffic to identify groups of vehicles rather than individual vehicles.

The analysis used to estimate the camera range for resolution of vehicles can also be used to estimate the range at which movement would be detectable.

A digital image represents the pattern of light levels across the camera field of view. If the image data is represented as a three-dimensional graph where the vertical axis represents the brightness of a pixel at coordinate x, y then bright regions will show as peaks and plateaux and dark regions will be seen as valleys and troughs. The steepness of a slope in this graph represents how rapidly the light intensity changes and, in general, where there is a significant gradient corresponds to an edge in the image. An image can be transformed so that the pixel intensity represents the size of the gradient at the location by performing a spatial convolution. The image thus produced is referred to herein as an edge image. The spatial convolution referred to can be performed using, for example, a Laplacian operator as disclosed in "Digital Image Processing", by Gonzales & Wintz, published 1987 by Addison-Wesley Publishing Company (see pp 338-340).

In traffic scenes the edge image generally highlights vehicles as complex groups of edges. An individual vehicle will be made up of several regions of differing intensity which in turn are different from the background scene. In most cases the road area in the image has a relatively low edge content, chiefly road markings and kerb lines. The presence of vehicles can thus be detected by the increase in edge content within the road area.

Figure 4 shows two histograms in each of which the number of pixels in a given image cell whose intensity falls in a given range, is plotted against that intensity range. The first histogram shows the case where there are few edges and the second shows the case where there are many edges. The histograms are unsigned (plotted without reference to sign), i.e. no distinction is made between the cases where, in going from one pixel to the next, the intensity increases and the intensity decreases. However, a signed histogram could be used instead, and this would have the convenience of keeping the mean stationary. As can be seen, in the second histogram there is a much greater distribution of intensities, and this is the shape of histogram to be expected where vehicles are present.

Besides detecting the presence of vehicles, the present invention also needs to detect movement. If we subtract an image taken at time t from an image taken at time $t + \delta t$ the differences will be due to four possible causes: movement of the camera, movement of objects within the scene, changes in lighting and electrical noise. For a fixed camera position subject to minimum vibration the first cause can be eliminated. If we choose δt to be sufficiently small then changes in light levels in a real world scene will be negligible. This leaves differences due to moving objects to be differentiated from those due to noise.

The differences caused by moving objects are a result of regions of differing brightness covering or uncovering each other, e.g. a bright region moving over a dark one. If δt is small then these differences will appear at the edges of the region. In real world images most edges are not simple steps but are sloped, and this is shown in Figure 5. The first diagram in Figure 5 is plot of intensity over a region of an image containing an edge, at two instants in time separated by an interval δt . The second diagram shows the result of subtracting one plot from the other. It can be seen that, as the distance moved increases, the size of the difference increases both in magnitude and in area.

An image, referred to herein as a difference image, can be created by subtracting, for each pixel, the intensity of the pixel at a time $t + \delta t$ from the intensity at a time t . A histogram can then be constructed from the difference image, in which the number of pixels in a given image cell whose intensity falls in a given range, is plotted against that intensity range. Increasing movement causes the distribution of intensities to spread. Figure 6 shows two signed histograms, the top histogram showing the distribution of signed differences due to noise, and the lower showing the histogram of differences when movement is taking place.

Using the methods described thus far, we have, for each image cell, a histogram representing the edge content of the cell and a histogram representing the movement content of the cell. These histograms then need to be analyzed. This may be done in various ways. One way is to calculate the value of the variance for each histogram. The greater the edge content or movement content respectively, the greater will be the variance of the respective histogram. A variance above predetermined threshold values is taken to represent the presence of vehicles and the presence of movement respectively. An alternative approach is to sum the area of the histogram above a predetermined positive intensity value and below the corresponding negative intensity value (this is for a signed histogram, for an unsigned histogram only the area above a predetermined positive value is required). A summed area above a predetermined threshold value is taken to represent the presence of vehicles or the presence of movement, depending on which histogram is being considered. The value of the variance or summed area, depending on which approach is being used, is referred to below as the edge parameter or movement parameter, as the case may be. The threshold values, for either method of analysis, are designed below as T_E for edges and T_M for movement.

Two further quantities for each cell can be found from:

$$V1 = \frac{\text{movement parameter}}{\text{area of the cell in pixels}}$$

and

$$V2 = \frac{\text{movement parameter}}{\text{edge parameter}}$$

V1 represents the proportion of movement per unit area and V2 the proportion of movement normalised for the edge content in a cell. These values allow comparison between adjacent cells. V2 gives some measure of the speed of the traffic.

Comparing the values for edges and movement in each cell with the threshold values enables a "state" for each cell can be determined using a simple logical table.

		Movement parameter	
		$<T_m$	$\geq T_m$
Edge	$<T_e$	None	Analyse further
Parameter	$\geq T_e$	Stop	Moving

In the case where further analysis is required, the movement and edge parameters are again compared with the threshold values but this time their closeness to the threshold value is assessed. If the movement parameter is significantly more than T_m the cell state is adjudged to be "Moving". If this is not so the edge parameter is examined to see if it is within a set limit of T_e and if it is the cell state is judged to be "None". If neither of these conditions is satisfied the cell state is undefined.

Any residual undefined cells are examined for their value of V1. If this exceeds a threshold value the cell is designated as moving, otherwise it is set to none. "None" is a representation that no traffic is present in the cell, "Stop" is a representation that traffic is present and stopped, and "Moving" is a representation that traffic is present and moving.

Any cell classified as "Stop" is checked again to confirm the status by closer analysis of the relative values of the parameters and the thresholds. The basis for this is that a "Stop" state is the one which generally triggers an alarm or action, and it is clearly desirable ensure the correctness of the analysis before this occurs.

The result of determining the movement and edge parameters for a single cell for a number of iterations of the above process is shown graphically in Figure 7, each cross representing the result of one cell.

The threshold values are preferably generated by carrying out a "training" run. In this, a number of pairs of images, say 50 or 100 pairs, are analysed in the manner described above and values generated for the movement and edge parameter of each cell. No state definitions are performed and the training run is carried out when the traffic is very light and is moving freely.

To determine the value of T_m , a histogram is constructed of the frequency with which the movement parameter lies in a given range. The mode value of the histogram is determined and T_m is set at a value slightly greater than the mode value.

Whereas it transpires that the mode value for this movement histogram clearly defined, the same is found not to be true if a similar histogram is constructed of edge parameters. Accordingly, T_e is determined differently. What is done there is to consider for each cell only those images in which the value of the movement parameter is less than or equal to the mode value, i.e. for which it assumed that no movement is taking place. Since the traffic is freely moving it follows that one is then only considering, for each cell, images where no vehicles are present. These cell images are examined and the maximum value of the edge parameter determined for these cells. The value of T_e is then set at this value.

A means can be provided whereby the values of T_m and T_e can be displayed to an operator and the

values altered if desired.

The description given thus far has concerned itself largely with analyzing the state of individual cells. However much more information can be obtained by considering groups of cells. By considering cells in groups along each traffic lane a number of "objects" can be defined according to the states of the cells in the group. There are three basic traffic "objects", namely:

GAP	A cell or a succession of cells of state "None".
PLATOON	A cell or a succession of cells of state "Moving".
BLOCK	A cell or a succession of cells of state "Stop".

A number of more complex "objects" can then be defined on the basis of the three basic "objects". Two important ones are:

15 QUEUE: BLOCK followed by a PLATOON with no GAP in between them.

WAVE or HUMP: PLATOON followed by a BLOCK followed by either a GAP or another PLATOON.

Within each object the values of V1 and V2 can be compared between adjacent cells. If the difference in these values between adjacent cells exceeds a threshold value then an error in cell state classification is deemed to have occurred. The data of the cell thus identified is then re-examined and the cell is re-classified according to its own data and the type of the object currently being considered. The object classification for that lane is then repeated. This process continues recursively until a stable object classification is achieved.

Figure 8 shows the identification of traffic "objects" in four exemplary lengths of lane.

The system of the invention may be provided with means for displaying to an operator the presence of the traffic "objects".

Figure 9 shows in summary the steps involved in the procedure of the invention. On the left, starting from the bottom, are the steps when the image is divided into a rectangular grid of cells, and on the right are the steps when the image is divided into cells whose size represents an equal area of traffic lane. The division of image into cells is referred to below as a cell map. The letters P, G and Q stand for "platoon", "gap" and "queue" respectively.

If desired, the analysis may be further refined to give a measure of the speed and acceleration of the travel or formation of the "object". Also, an indication can be given of the location where each "object" starts and finishes.

The state of the traffic in each lane can be found by an analysis of the objects.

35 A lane is described by:

Spatial Occupancy =

$$\left(1 - \frac{\text{Total GAP length}}{\text{Total lane length}}\right) \times 100$$

Distribution factor =

$$\left(1 - \frac{\text{Number of GAPS}}{\text{Number of objects}}\right) \times 100$$

The spatial occupancy expresses the percentage of the length of a lane which is occupied by vehicles. The distribution factor is a measure of the extent to which traffic is bunched together in a lane.

Each object type has an associated weight factor which is used in calculating a weight for each object. Object weight (W) = (factor x length) x (1 + number of "stops").

This is used to find the most significant object (e.g. QUEUE is more significant than PLATOON), and to provide some method for judging the warning level to be associated with QUEUE, WAVE and BLOCK.

A message giving lane number, spatial occupancy, distribution factor and most significant object for each lane is then displayed to the operator.

The following is a table setting out some possible values for the "factors" to be associated with various

traffic objects, and the resulting minimum value for W

Object	Factor	Minimum W
GAP	0	0
PLATOON	1	1
BLOCK	6	12
QUEUE	3	12
WAVE or HUMP	3	12

Figure 8a is a graph showing the way in which the value of W increases for various traffic objects with increasing numbers of cells in the object. The graph includes three composite objects made up of a QUEUE plus 1, 2 or 4 stop cells.

The system of the present invention may be connected to means for taking action to control the traffic on the basis of the detection of particular traffic objects. Two particular applications of the invention are in the detection of queues and in the detection of incidents. The term "incidents" is a recognised term in the study of traffic and denotes an abnormal event which has consequences for traffic flow. Such an abnormal event might be, for example, an accident, a breakdown or the presence of debris on the road. Data from the monitoring system of the present invention can be fed, for example, to roadside warning systems to provide drivers with an appropriate message concerning the state of the traffic ahead, to signals (for example traffic lights) controlling access to a road so as, for example, to prevent vehicles entering a road which was over-congested, or to a route guidance system on board a vehicle.

One development of what is described above is to use a number of cameras to observe successive sections of road. The traffic "objects" detected by each camera may then be combined with one another to provide composite "objects" representing the state of the whole road or at least of larger parts of it than the individually observed sections. Also, the "objects" detected by each of the individual cameras can be compared with one another and the operator can then be provided with information relating only to the most significant object, or indeed, with the visual image from just the cameras which has detected the most significant objects. In this way the operator is freed from having to try to observe simultaneously the images provided by a possibly large number of cameras.

Figure 11 shows hardware which can be used for the present invention.

By way, at least in part, of review, attention is now directed to Figures 10a and 10b, which constitute a flow diagram of the steps involved in carrying out the present invention.

Turning first to Figure 10a, the software presents the operator with an option menu at which he is asked to specify whether what is to follow is a training run. If it is the first task is to draw a cell map defining the cells into which the image of the scene is to be divided. If it is not a training run then the system is loaded with cell map data previously provided.

In either case the system then proceeds to grab, digitise and store two successive video images, denoted as image A and image B, the images showing the scene at times separated from one another by approximately 0.2 seconds. A difference image is then formed by subtracting image A from image B, and an edge convolution is performed on image A to form an edge image.

Then, for every cell in the cell map, a histogram is calculated of the difference image, from which the movement parameter for the cell is determined, and a histogram of the edge image is calculated, from which the edge parameter for the cell is calculated. When this has been done for all cells the values of the movement and edge parameters are sent to a file.

In the case where the operator indicated at the outset that this was a training run, the process then passes to a "Break loop?" option which, if exercised, then causes the system to set the threshold values T_e and T_m for the edge and movement parameters. If the Break loop option is not exercised the process of analysing a pair of images is repeated. Typically, a mentioned above, 50 or 100 pairs of images are analysed before the data thus generated is used to set the above-mentioned thresholds.

Where what is being carried out is not a training run then the steps shown in Figure 10b are performed. For each cell in the cell map the cell state is determined, depending on the value of the movement and edge parameters in relation to the movement and edge threshold values. Once this has been done for each cell in the cell map then for each traffic lane in the cell map the line of cells covering that lane is analysed to ascertain the existence of the traffic "objects" described above. For each object the system determines the furthest downstream cell (which gives the start point of the object concerned), the length of the object in cells, the number of "Stop" cells which the object contains, the object weight, and the values of V1 and V2

as defined above. Once this has been done for each object the process then calculates the values for spatial occupancy and distribution factor (see above) for the lane. The overall process just described is repeated for each traffic lane in the cell map.

The data regarding the traffic objects is then output to a file, from where information as to the traffic objects detected in each lane are displayed to the operator. In the case where the system is being used to control traffic, the relevant data is sent from the file of object data to whatever equipment is carrying out the control.

The invention can be conveniently implemented on a commercially available image processing subsystem connected to a 80286 - based microcomputer. The subsystem digitises and processes pixel data at high speed while the host microcomputer controls the programming of the subsystem and processes the cell data. The procedure has been implemented using as the image processing subsystem a Series 151 Image Processor, a subsystem manufactured by Imaging Technology Inc. of Woburn, Massachusetts, U.S.A..

The subsystem is modular and comprises a set of cards connected by a VME bus and a proprietary video bus. Pixel data is transferred between the different function and memory cards via the video bus. An example of the subsystem comprises five cards, a digitiser and controller card, two framestore cards each with 1MB of video storage, an arithmetic/logical processing card and a histogramming card.

The microcomputer host controls the subsystem via a card that connects the host bus to the subsystem VME bus. Instructions can be sent to, and data received from, the subsystem by this route. The subsystem processes the image data up to the histogram stage with the host issuing the appropriate instructions for this lower level stage of the procedure. Analysis of the cell histograms and subsequent stages of the procedure are processes within the host microcomputer.

One application for which the present invention may be regarded as particularly suited is to detect queues of stationary or slow moving traffic on high speed roads such as freeways. The presence of slow moving or stationary traffic on this type of road represents a serious hazard due to the danger of fast moving vehicles running into the rear of the queue. Furthermore, the build up of traffic can be very rapid with the tail end of the queue moving back along the road extremely quickly.

In terms of implementation, high speed roads generally have some constraints which simplify the analysis of special images. In most cases the underlying background scene is relatively simple. Therefore, the presence of a vehicle can be inferred from an increase in scene complexity, i.e. an increase in the number of edges found using an edge detection operator. The pattern of movement within the scene is also fairly straightforward, particularly if attention is restricted to a single carriageway. Movement is then in one general direction and useful information can be derived from the magnitude of such movement. If the road has a high speed limit then the detection of slow traffic is even easier.

A key feature of the approach described above is that its purpose is to provide a qualitative description of the spatial distribution of moving and stationary traffic within a scene. The technique does not attempt to identify individual vehicles nor does it seek to follow the vehicle "clusters" that are identified as they move across the image. Instead, the strategy is to mimic the way in which a human observer might describe the pattern of traffic when viewing a CCTV monitor.

This is a major departure from approaches used to date. Other systems, even when they aggregate individual vehicles into groups or clusters (Abramczuk T (1984), A microcomputer based TV-detector for road traffic. The Symposium on Road Research Program, OECD, Tokyo, Japan, and Kudo, Y. Yamahira, T. Tsurutani T. & Naniwade, M (1983), Traffic flow measurement system using image processing. 33rd IEEE Vehicular Technology Conference, Toronto, Canada. pp 28-34.

45

Video =>Image processing => Vehicle parameters=>Traffic
 image procedure or vehicle group patterns
 parameters

50

The approach proposed here reduces this to three steps:

55

video => Image processing =>Traffic objects
 image procedure

Although this is a very simplified model it does outline the differences. The first approach involves a more "microscopic" analysis of the image followed by a reconstruction of the traffic situation. Here, a more direct approach with the aim of providing a more qualitative description of the traffic has been adopted.

The rationale behind this more direct approach is that by reducing the number of stages and the requirement for detailed data the processing power required is also reduced. However, this reduction has not been effected by reducing the spatial data available for analysis and so the main advantage of using video images has been retained.

Another important feature of the approach is that the image interpretation is carried out on the spatial difference (i.e. gradient) and the temporal difference transforms of the scene and not from grey level values. This reduces the influence of changes in the light distribution across the scene caused either by changes in the ambient light or the action of the auto-iris. The effect of this is to increase the robustness of the technique over time and to widen the range of conditions under which it will perform satisfactory.

The basis for most procedures designed to detect transient effect, e.g. the passage of a vehicle through a camera's field of view, is a comparison of the current image with some predetermined background image. In the procedure described here this comparison between current and background conditions is carried out on parameters measured over small regions from the two transformed images. The use of transform parameters instead of a simple grey level comparison improves the reliability of the procedure because these abstracted values relate to the structure of the scene rather than just the distribution of light intensity.

Describing local regions of the image, i.e. "cells" in terms of two parameter values is both the first data reduction step and the first abstraction from pixel data. These values are still characteristics of the image and the next stage is to interpret these values in a way that relates to vehicles and traffic. This is done by a comparison with predetermined background values for the parameters which specify a "state" for each "cell". The assumption that these states relate to vehicular movement is based on considering only those cells which are over the traffic lanes, which will be true at least for roads where pedestrian movement is prohibited.

Combining adjacent cells based upon their states provides method for further refining the description of the scene and gives a qualitative result that is similar to how a human operator might describe that scene. Each description is entirely independent of any previous descriptions. This means that the processing speed can be governed by the application and the nature of the computing equipment chosen.

Various modifications can be made to the method described thus far. One of these is for the camera to use a plurality of different cell maps. There are two particular circumstances where this may be useful. The first is to enable it to cope either with traffic driving on the left hand side of the road (as in the United Kingdom) or with traffic driving on the right hand side of the road (as in most other countries), and to enable it to cope with traffic flowing in either direction. This is shown in Figure 9a. It will be seen that the four cell maps shown are geometrically identical, but the cells are numbered differently. In each case the cells are numbered starting with the cell furthest downstream in the nearside lane. The second particular circumstance is where the camera may be intended to operate from a number of different positions. In this case, a different cell map may be associated with each such position.

40

Claims

1. A method of traffic monitoring which comprises forming at least first and second scene images of a scene in which traffic may be present, the images being formed at instants of time separated by a time interval, each scene image being an array of pixels, processing at least one of the first and second scene images to form an edge image representing the occurrence of edges in the scene, determining on the basis of the said edge image the presence or absence, and spatial location, of traffic in the scene, forming a difference image in which each pixel represents the difference between the intensity of the pixels of the first and second scene images at the corresponding point in the image, and determining from the distribution of pixels of different intensities in the difference image the presence or absence of movement in the scene, wherein the edge image and difference image are each subdivided into an array of cells, each edge cell and its related difference image cell corresponding to a given sub-area of the image constituting a scene image cell, the presence or absence of traffic and the presence or absence of movement being separately determined for each scene image cell, characterised in that each of a plurality of scene image cells lying along the image of the line of a traffic lane is analysed as aforesaid, and the presence of predetermined traffic objects detected on the basis thereof.

2. A method according to claim 1, wherein the said traffic objects are which the said plurality of scene image cells are analysed include at least gaps (a gap being a cell or a succession of cells along the line of

a traffic lane in each of which no traffic is detected), and/or platoons (a platoon being a cell or a succession of cells along the line of a traffic lane in each of which moving traffic is detected), and/or blocks (a block being a cell or a succession of cells along the line of a traffic lane in each of which stationary traffic is detected).

3. A method according to claim 2, wherein the said traffic objects also include queues (a queue being a block followed by a platoon with no gap in between them) and/or waves (a wave being a platoon followed by a block followed by one of a gap and a platoon).

4. A method according to any preceding claim, wherein the presence or absence of traffic and the presence or absence of movement is determined for each scene image cell in accordance with the following table:

		Movement parameter	
		$<T_m$	$\geq T_m$
Edge	$<T_e$	None	Analyse further
Parameter	$\geq T_e$	Stop	Moving

where the movement parameter is a measure of the amount of movement detected in the cell, the edge parameter is a measure of the amount of edges detected in the cell, T_m is a movement threshold value, T_e is an edge threshold parameter, "None" is a representation that no traffic is present in the cell, "Stop" is a representation that traffic is present and is stopped, and "Moving" is a representation that traffic is present and is moving.

5. A method according to claim 4, wherein a training run is carried out to determine the values of T_e and T_m by monitoring traffic known to be light and moving, T_e being substantially the maximum edge parameter present in scene image cells where no movement is occurring, and T_m being at or slightly above the mode value of a histogram constructed of the frequency with which the movement parameter lies in a given range.

6. A method according to any preceding claim, wherein the presence or absence of the said traffic objects is visually displayed and/or fed to means for controlling the traffic.

7. A method according to any preceding claim, wherein a value is determined for the proportion of road space in a given lane which is occupied by traffic.

8. A method according to any preceding claim, wherein for each traffic object at least one of the following is determined, namely location, the length in cells, a weight expressive of the relative importance of the object, the proportion of movement per unit area and the proportion of movement normalised for the edge parameter.

9. A method according to claim 8, wherein said scene images are formed at a plurality of locations, and a scene image is displayed of that scene in which the traffic object having the greatest weight is to be found.

10. An apparatus for traffic monitoring which comprises means for forming at least first and second scene images of a scene in which traffic may be present, the images being formed at instants of time separated by a time interval, each scene image being an array of pixels; and means for processing at least one of the first and second scene images to form an edge image representing the occurrence of edges in the scene, determining on the basis of the said edge image the presence or absence, and spatial location, of traffic in the scene, forming a difference image in which each pixel represents the difference between the intensity of the pixels of the first and second scene images at the corresponding point in the image, and determining from the distribution of pixels of different intensities in the difference image the presence or absence of movement in the scene, wherein means are provided for subdividing the edge image and the difference image each into an array of cells, each edge image cell and its related difference image cell corresponding to a given sub-area of the image constituting a scene image cell the presence or absence of movement being separately determined for each scene image cell, characterised in that each of a plurality of scene image cells lying along the image of the line of a traffic lane is analysed as aforesaid, and the presence of predetermined traffic objects detected on the basis thereof.

11. An apparatus according to claim 10, wherein the said traffic objects for which the said plurality of scene image cells are analysed include at least gaps (a gap being a cell or a succession of cells along the line of a traffic lane in each of which no traffic is detected), and/or platoons (a platoon being a cell or a succession of cells along the line of a traffic lane in each of which moving traffic is detected), and/or blocks

(a block being a cell or a succession of cells along the line of a traffic lane in each of which stationary traffic is detected).

12. An apparatus according to claim 11, wherein the said traffic objects also include queues (a queue being a block followed by a platoon with no gap in between them) and/or waves (a wave being a platoon followed by a block followed by one of a gap and a platoon).

13. An apparatus according to claim 12, wherein the presence or absence of traffic and the presence or absence of movement is determined for each scene image cell in accordance with the following table:

		Movement parameter	
		$<T_m$	$\geq T_m$
Edge	$<T_e$	None	Analyse further
Parameter	$\geq T_e$	Stop	Moving

where the movement parameter is a measure of the amount of movement detected in the cell, the edge parameter is a measure of the amount of edges detected in the cell, T_m is a movement threshold value, T_e is an edge threshold parameter, "None" is a representation that no traffic is present in the cell, "Stop" is a representation that traffic is present and is stopped, and "Moving" is a representation that traffic is present and is moving.

14. An apparatus according to claim 13, wherein the values of T_e and T_m and values determined in a training run in which traffic known to be light and moving is monitored, T_e being substantially the maximum edge parameter present in scene image cells where no movement is occurring, and T_m being at or slightly above the mode value of a histogram constructed of the frequency with which the movement parameter lies in a given range.

15. An apparatus according to any one of claims 10 to 14, comprising visual display means for receiving data from the said processing means, whereby to display the presence or absence of the said traffic objects, and/or traffic control means connected to receive data from said processing means and operate in accordance therewith.

16. An apparatus according to any one of claims 10 to 15, wherein means are provided for forming said scene images at a plurality of locations, and a scene image is displayed of that scene in which the traffic object having the greatest importance is to be found.

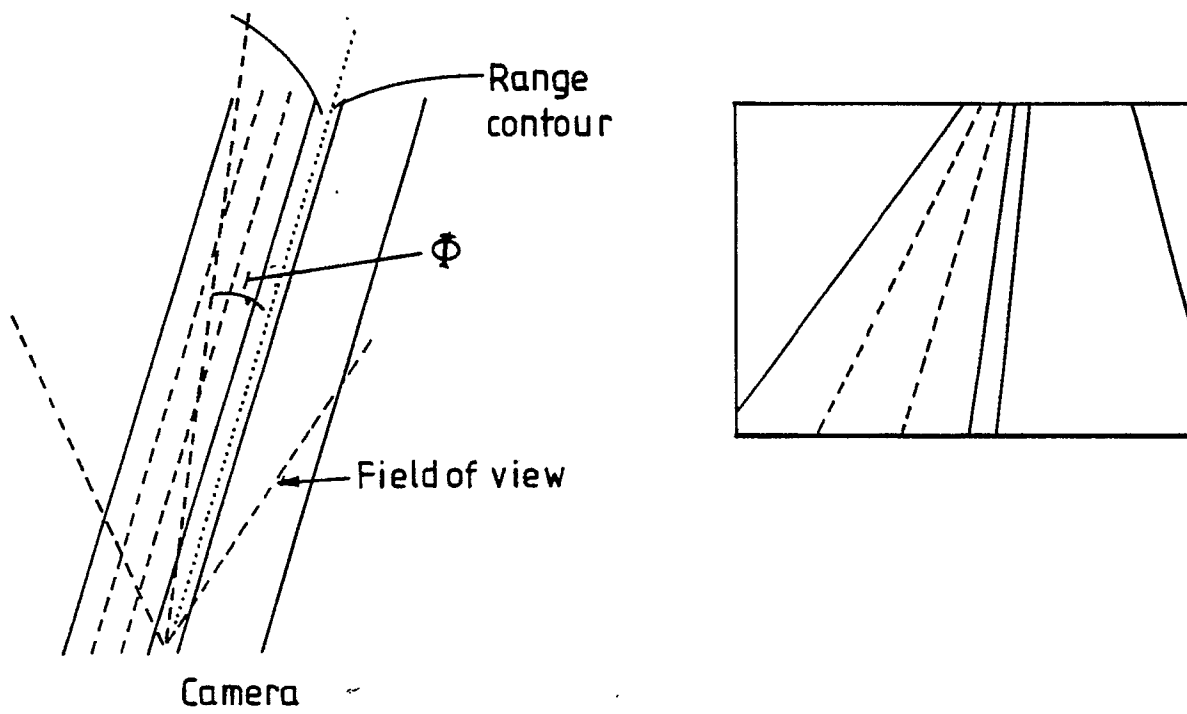


Fig. 1.

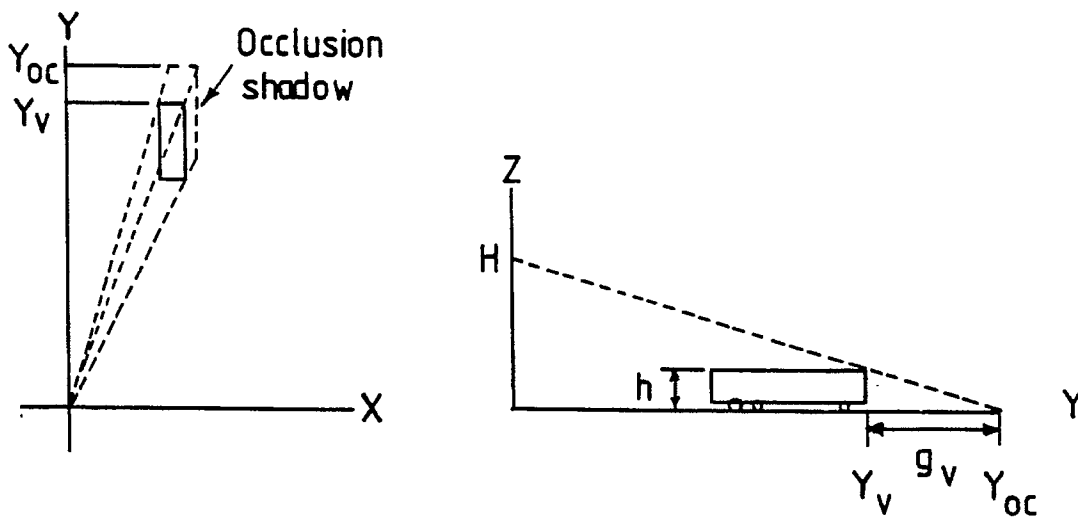


Fig. 2.

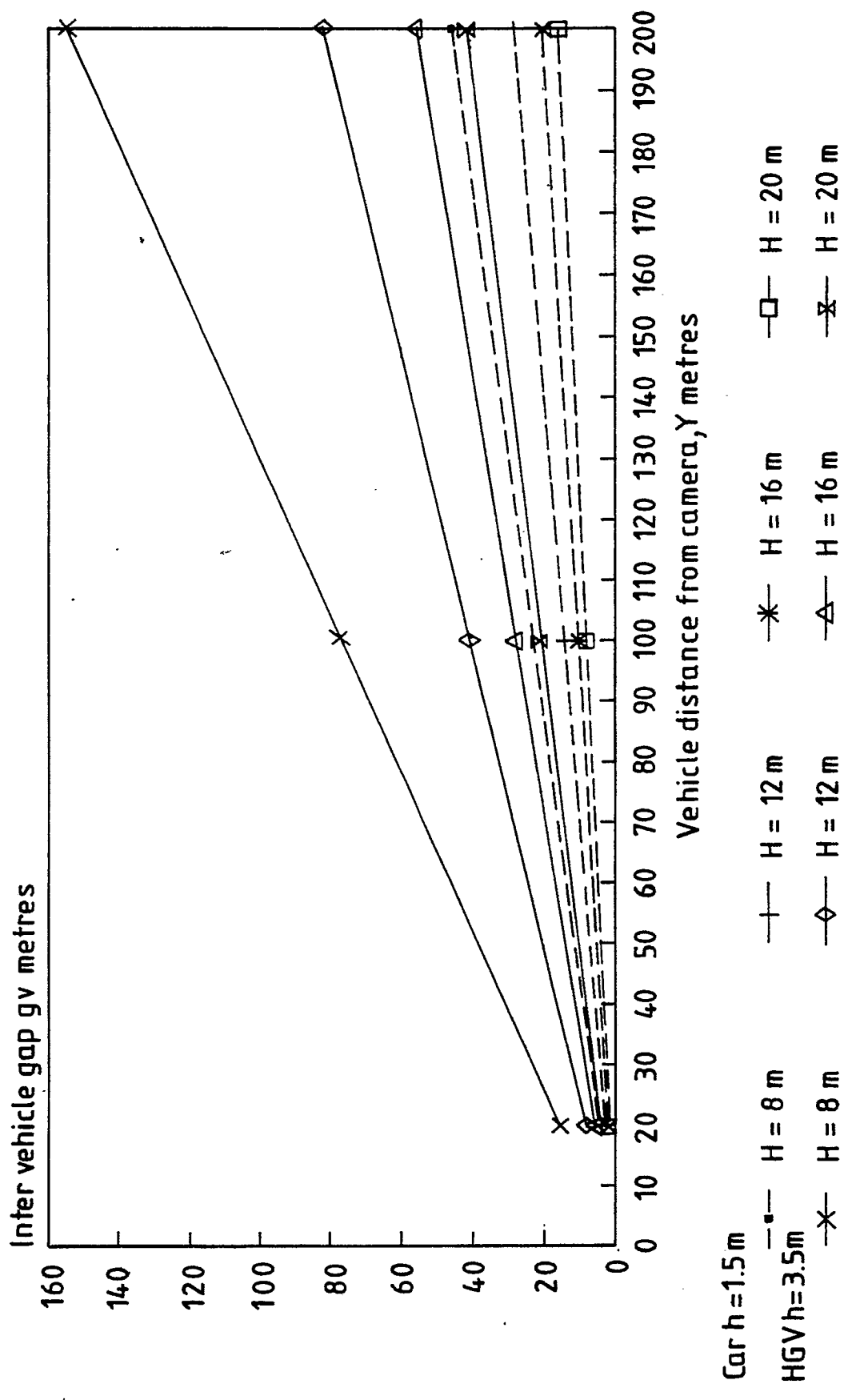


Fig.3.

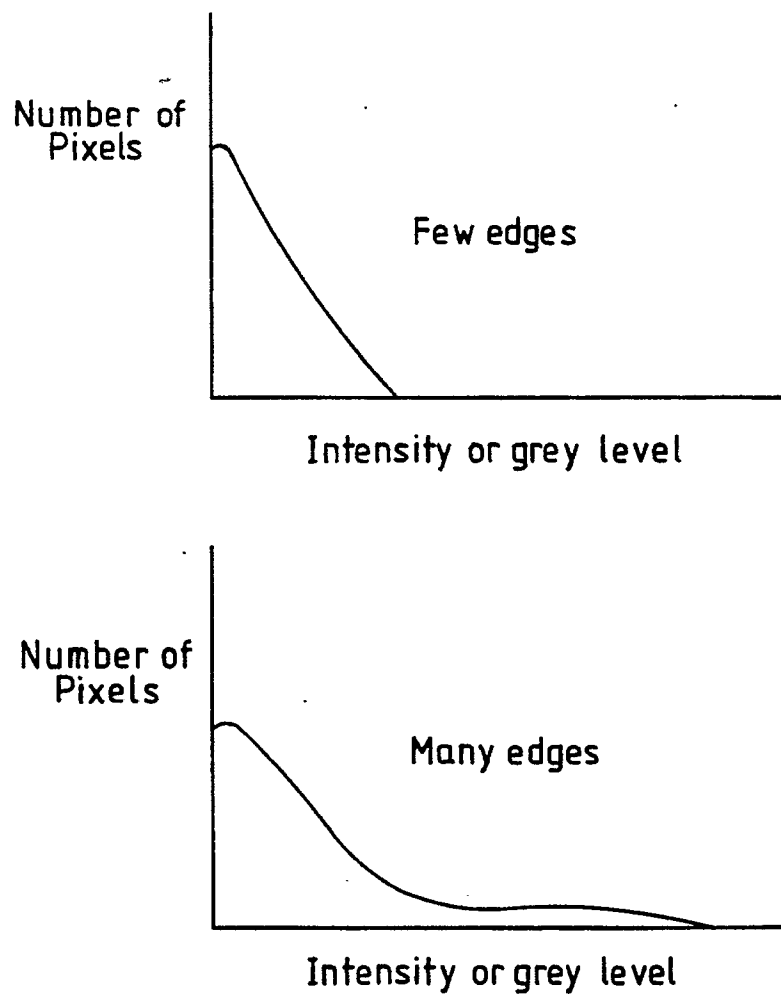


Fig.4.

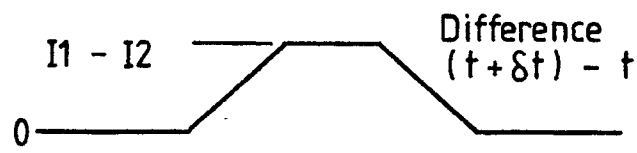
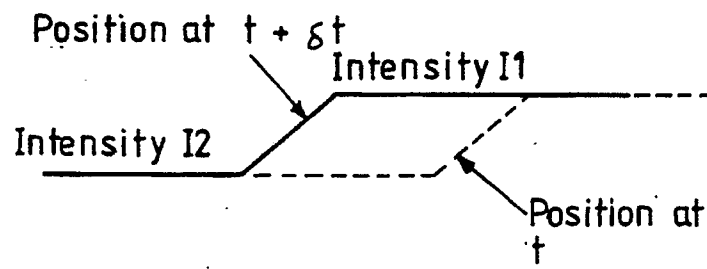


Fig.5.

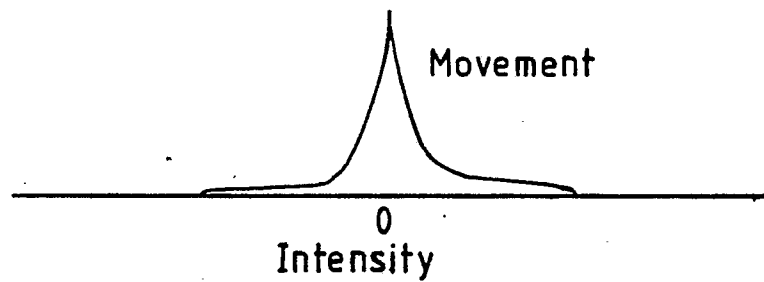
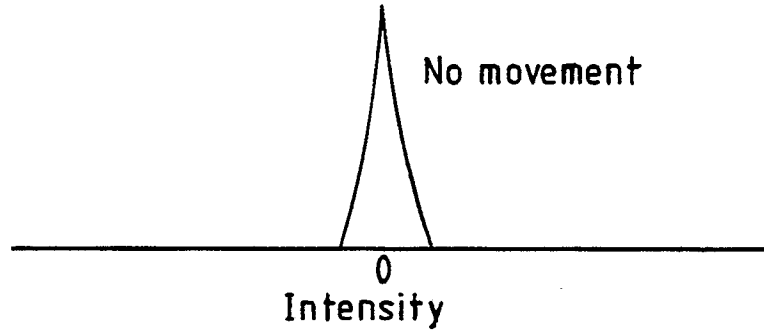
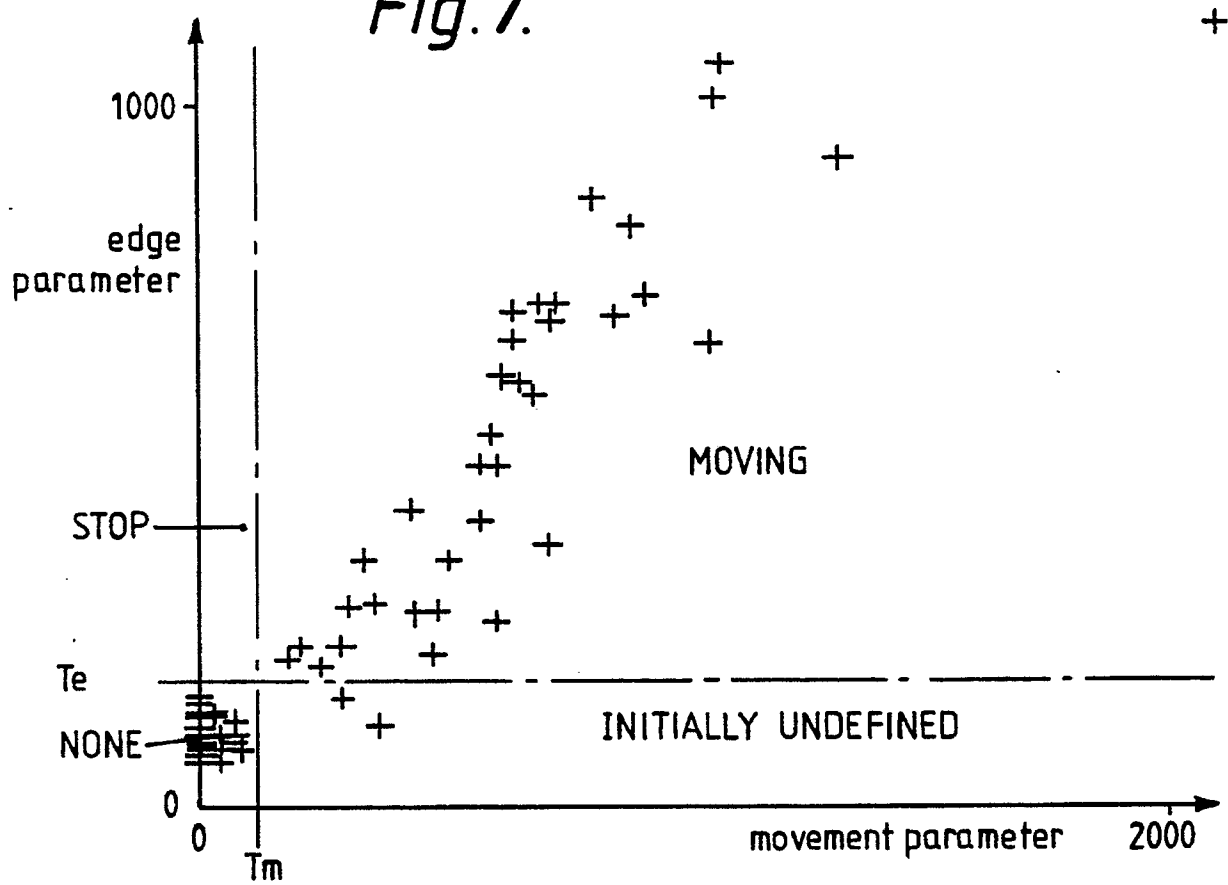


Fig.6.

Fig. 7.



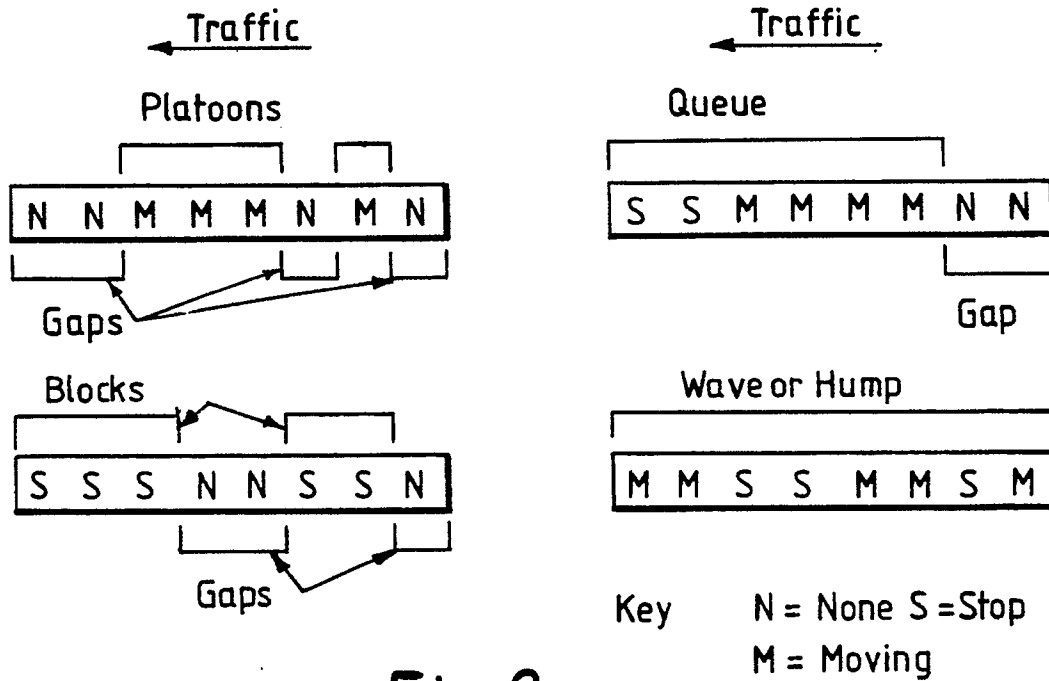


Fig.8.

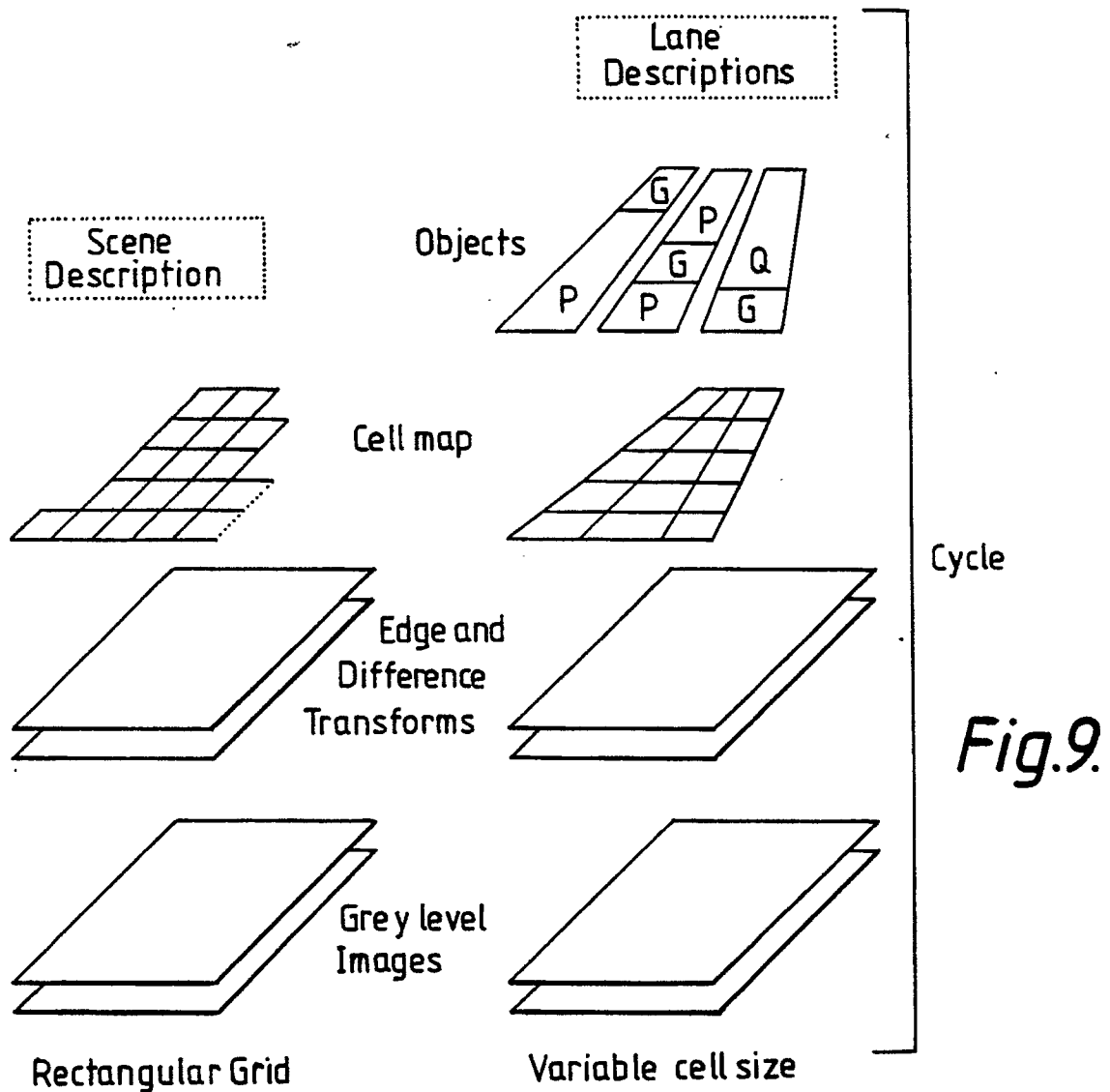


Fig.9.

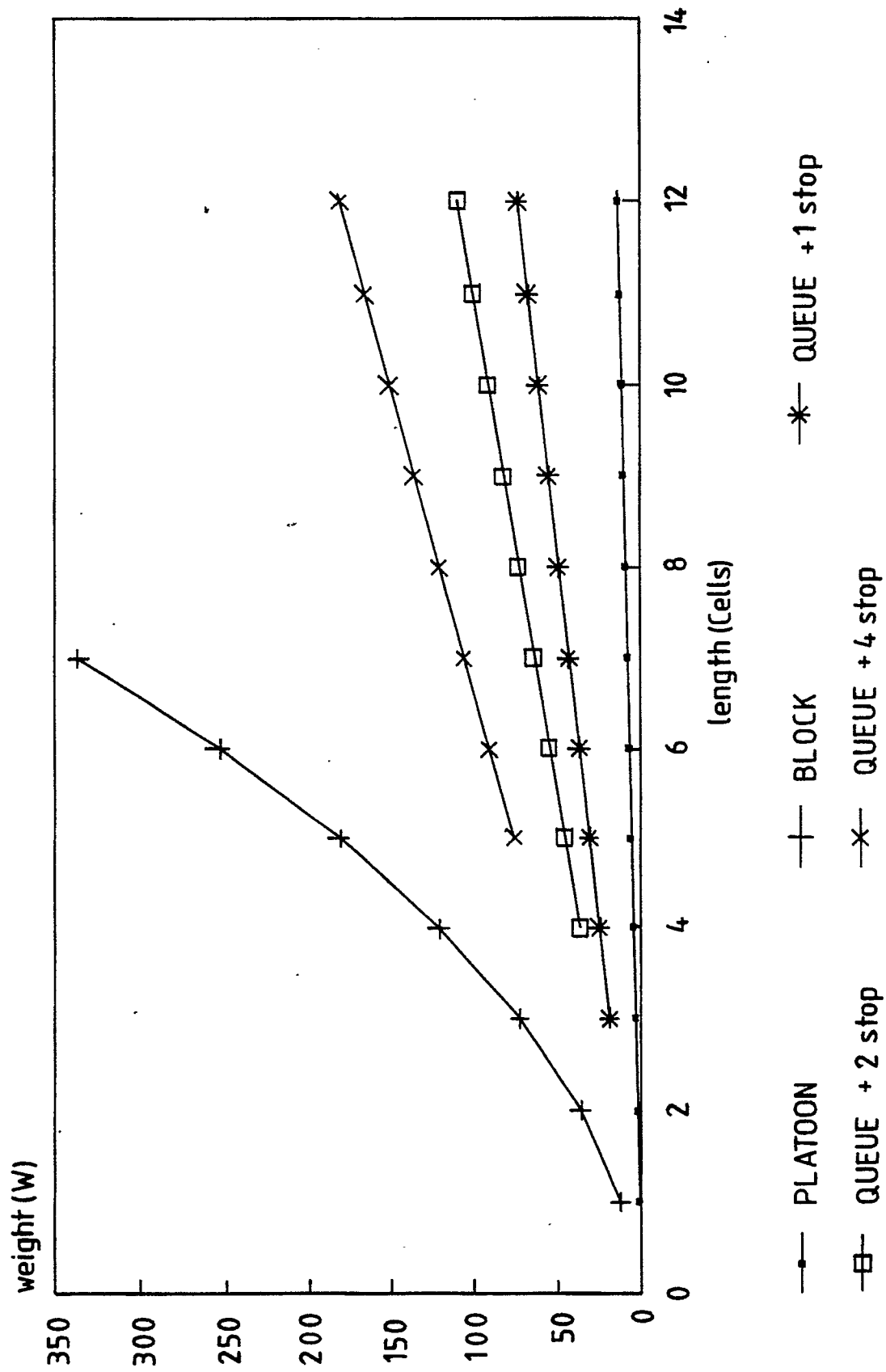
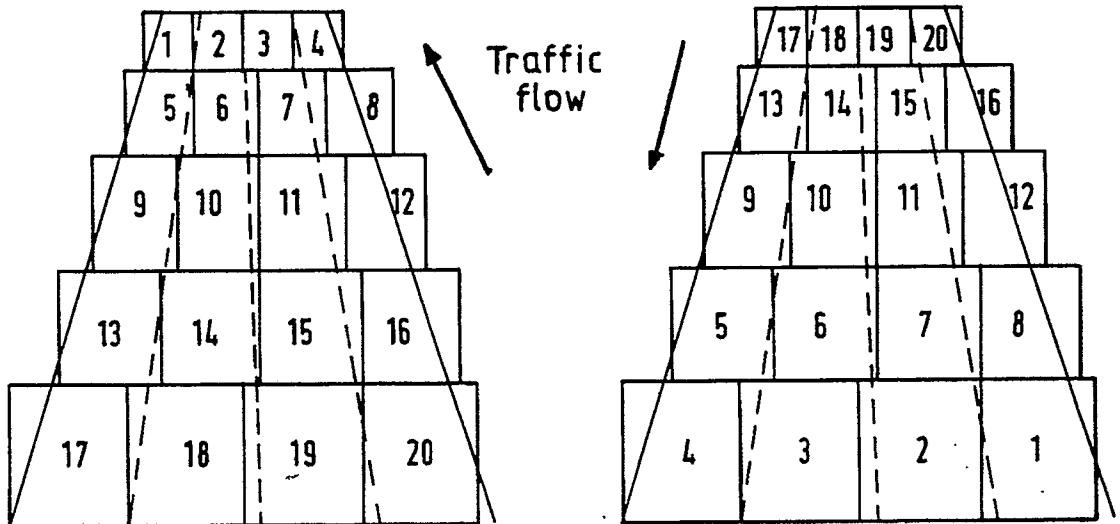
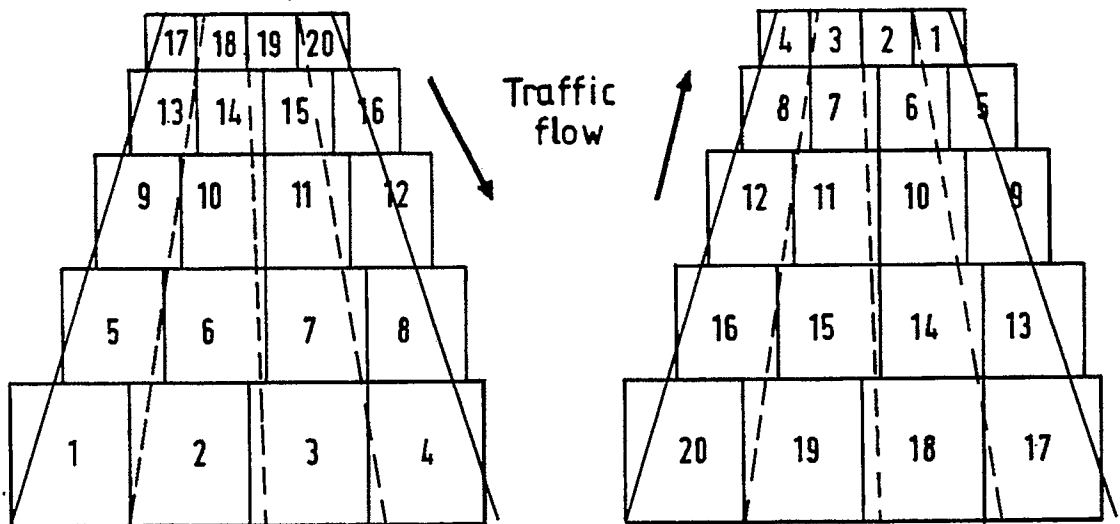


Fig. 8a. Weight v length for different object types.



Driving on left



Driving on right

Fig.9a.

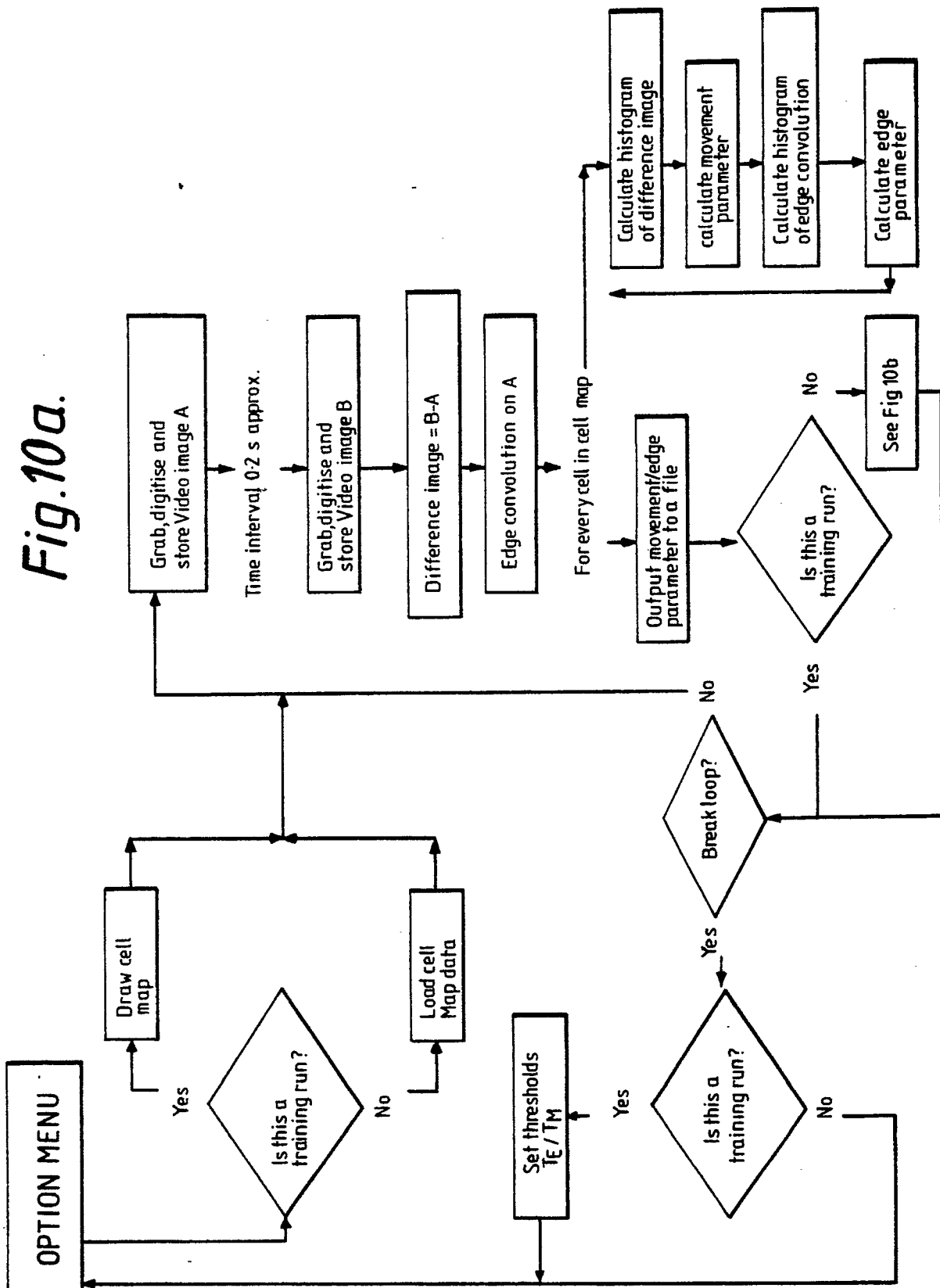
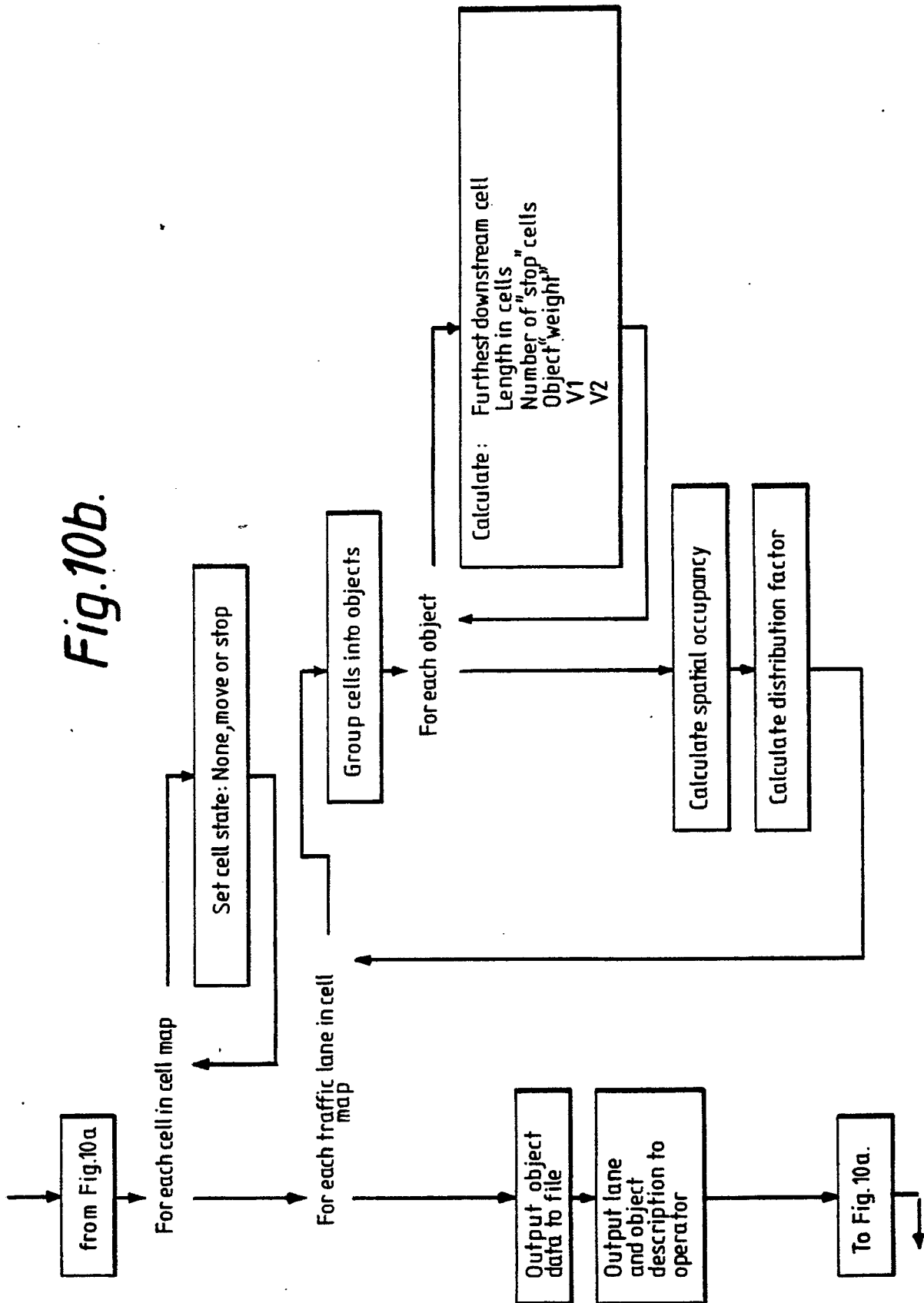


Fig. 10b.



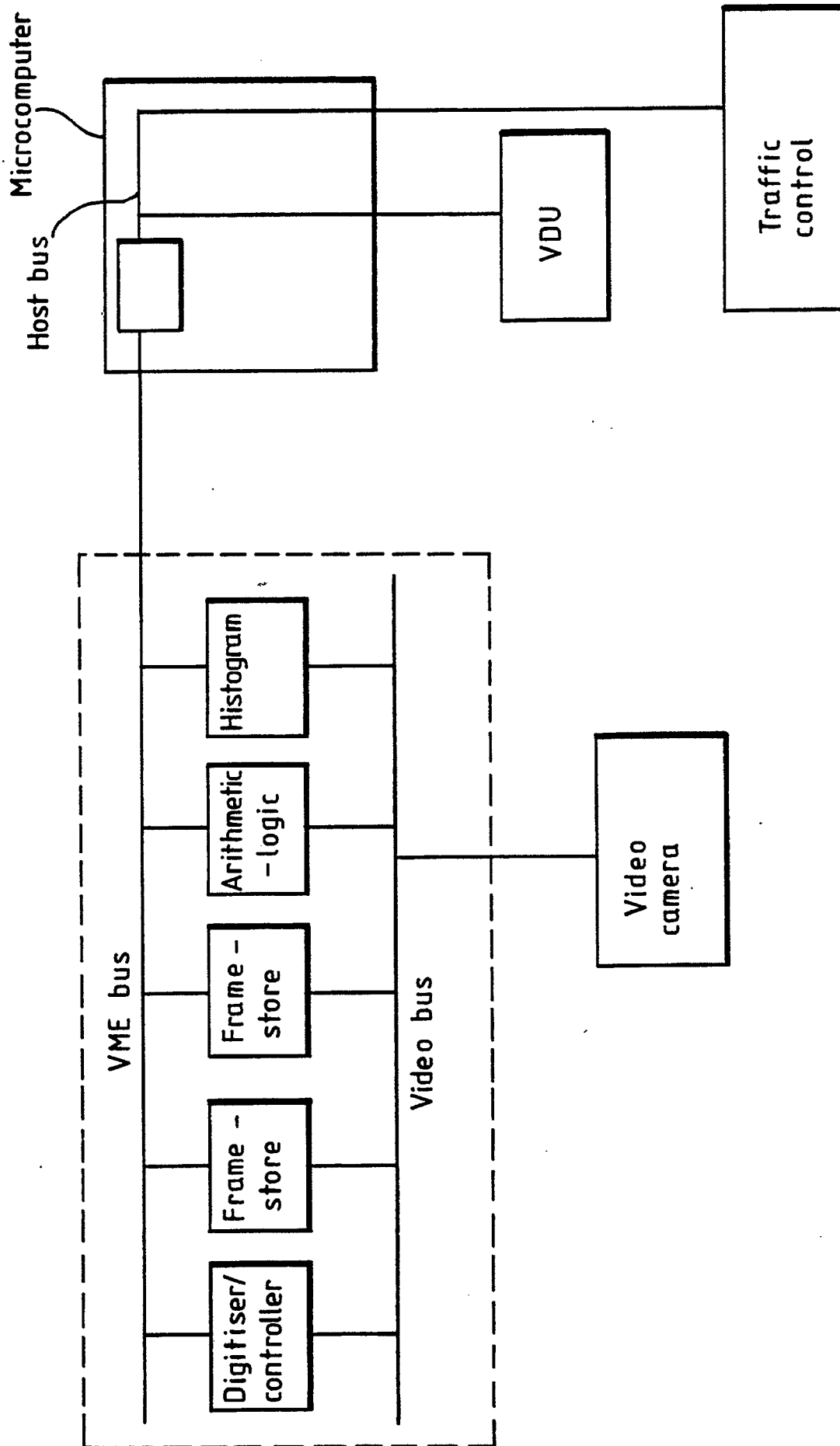


Fig.11.