A Vehicle Image Description for On-line Fast Classification

Jan Piecha, Marcin Staniek
Informatics Systems Department of Transport
Faculty of Transport, Silesian University of Technology
ul. Krasinskiego 8, 40-019 Katowice, Poland
{jan.piecha, marcin.staniek}@polsl.pl

Abstract. *The paper discusses an image analysis algorithm implemented in on-line measurement of vehicles video-detectors, where computing processes have to be organised in very short, strictly defined time limit. The data processing simplification is a key problem of image filtering and pre-processing methods finding. The video-recorder interface was constructed using a Programmable Array Logic (PAL) technology [1], processing binary sequences at the control execution unit. The well-known techniques of images processing were taken under consideration, with their preferences for fulfilling the assumed restrictions. The vehicle successful extraction depends on several conditions that were analysed in details; to be used for video-camera detectors development.*

Keywords. Video detection, image processing, on-line traffic control, PLAs applications

1 Introduction

The investigations presented in this contribution concern well known techniques of images processing analysis, with their specific preferences, for vehicles stream recognition and classification. This control systems developer has to consider several restrictions and preferences; for images description, extraction and classification.

One of these analysis criterions concerns time limit for the data stream processing by Programmable Arrays Logic (PAL) computing machine. The incoming data stream, determines sampling and computing time intervals for data processing.

Bellow, several steps of video data analysis were introduced. They were taken under consideration, for adequqte image processing algorithms selection. The investigated video data streams, is possesed by several digital cameras, located on specified points at laines intersections for a traffic proper representation.

2 The image extraction main steps

2.1. The segmentation procedures

The cameras resolution was equal to 768 x 512 pixels, with 8-bits colours coding only. The depth of these colours was also limited into minimal dimension, available at the camera settings. It remarkable reduces the computing time and procedures complexity [2], [3]. First step of the processing algorithm concerns segmentation of the image, where several characteristic features were under analysis. In these segmentation processes separate units of images are extracted, having the same or similar features; pixels denoted as the extracted objects.

The vehicle extraction process was carried out by methods of objects edges finding; overdrawing these objects. The object's edges and their content are described by set of numbers, assigning the object's pixels. These simple and fast algorithms indicate many faults for not continuous edges; implying additional analysis at these points.

For this analysis, the Smith algorithm for segmentation processes was taken under careful consideration [4]. It makes the objects' area filling-up process satisfactory simple and effective. The operation is called „sowing", with multiple checking of the same pixel at one image. In the Smith algorithm, subsequent rows of pixels are analysed. The segment n of the image is denoted by pixels $P(x, y)$ in line y, where neighbourhood of the edge is checked; in accordance with the conditions, given by formula (1). In a processing stack the addresses of pixels $P(x, y+d)$ are recorded.

$$P(x, y) = 0 \cap [P(x, y+d) = 0 \cap \\ (P(x-1, y+d) = 1) \cup (P(x+1, y+d) = 1)] \qquad (1)$$

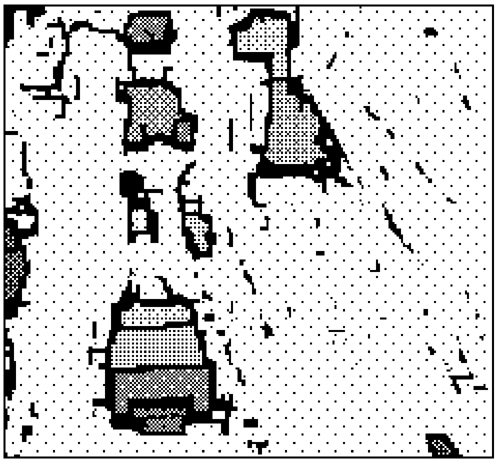
where:
$d = -1$ – for checking pixels in line $y-1$,
$d = 1$ – for checking pixels In line $y+1$.

The pixel $P(x, y)$ value is equal to 0 – for an internal area of the object, 1 – for the edge. The carried out segmentation, provides us with several image parts that were taken once more under analysis; by another algorithms (for the analysis quality improving).

Fig 1 shows the image map with visible edges as a result of lower level extraction processes; so called morphological operations.

Φιγυρε 1. Τηε εδγεσ μαπ, αφτερ μορπηολογιχαλο περατιονσ

In Fig 2 we observe some indexing results, appointed by different textures, assigning the Smith algorithm products.



Φιγυρε 2. Τηε ιμαγε αφτερ σεγμεντατιον

2.2. The image parameterisation

The analysis is based on so called main image parameters with their shape coefficients, available directly or indirectly. The main parameters express geometrical dimensions (as: surface, edges lengths, projections, Feret diameters, etc.).

The object's shape defines the Feret coefficient and coefficients of object's content [7,16,18]. The list of objects {Oj} is numbered and assigned by their characteristic features (3), as:

$$O_j = (x_s, y_s, L, S, I_z, I_p, R_z, R_F) \qquad (3)$$

where:
j – expresses current index of an object, *(xs, ys)* – are coordinates of the object load centre, *L* – expresses the edge length, *S* – the object surface, I_z – medium value of projections length (according to traffic direction), I_p – medium value of projections length (opposite to traffic direction), R_z – the content coefficient:

$$R_z = \frac{L^2}{4\pi S}, \qquad (4)$$

R_F – the Feret coefficient:

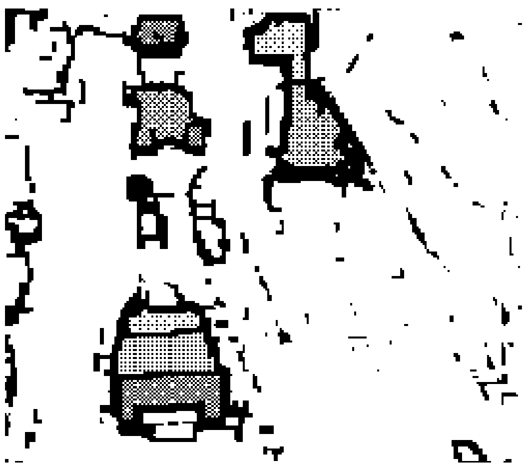
$$R_F = \frac{L_h}{L_v}, \qquad (5)$$

where:
L_h – maximal horizontal dimension, L_v – maximal vertical dimension.

2.3. Some objects selection procedures

Due to reduce objects-number analysis, the pre-processing algorithm runs at the beginning [14,20]. Objects of a size being above or bellow the assumed value (presumable not vehicles) are erased from the data set. The surface S bellow the threshold S_{prog} = 100 pixels, to high density; where the content coefficient value is above the threshold R_{z_prog} = 15.

The example image, after the selection and objects filtering, has been show in Fig 3, where number of remaining data objects was reduced.



Φιγυρε 3. Ιμαγε αφτερ οβφεχτσ φιλτερινγ ανδ σελεχτιον

2.3. Further analysis of the image

Next step of the image processing concerns objects aggregation, with objects found in neighbourhood. The surface of all objects, with their total sum is calculated [21]; due to avoid mistakes in classification, where two neighbour small cars are distinguished instead of one big vehicle. Finally the centre of objects' loads and medium values of vehicles projections lengths was found:

$$OA_i = (x_s, y_s, S, I_z, I_p) \tag{6}$$

On checking the range of OA_i value, the vehicle class is defined. Dimensions and localisations of object are defined on a base of objects' sequence with passing vehicles stream observation.

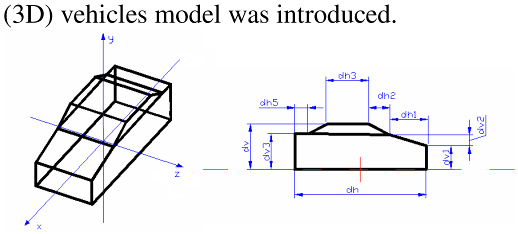
The algorithm characteristics:

– the chosen methods are not using any complex calculations that is why, for vehicles shape extraction, none scaling techniques are needed [9,10,11],
– although the simplest methods were used (for optical flow simplification) they still allow observing the vehicles movement, in any direction by n gates in video sequence analysis,
– some troubles were noticed for distinguishing similar objects, where additional sampling methods for their content analysis, were used,
– the combination of these classification methods satisfactory results, for freely movement of vehicles, were noticed [5],[6],
– some additional recognition procedures are needed for traffic rush states, when vehicles' drawing is aggregated (of close driving vehicles), indicating a higher class of vehicle,
– it is obvious that the method effectiveness also highly depends on an image quality.

3 The vehicle - mask finding

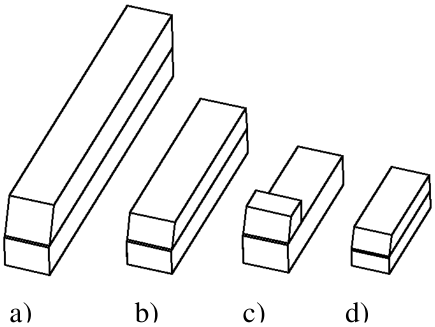
3.1. The vehicle dimensions

The vehicle classification by objects comparison needs their well defined pattern objects (models), of every vehicle class. These models were defined on medium values of vehicles dimensions that are identified as the same group of objects. The vehicle-mask extraction algorithm was also based on finding its medium dimensions [2,8]. In Fig. 4a three-dimension (3D) vehicles model was introduced.



Φιγυρε 4. Τηε τηρεε διμενσιονσ εξαμπλε ϖεηιχλε μοδελ

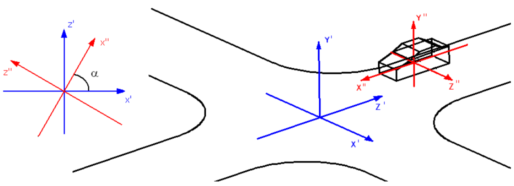
By the carried out modelling method, five classes of vehicles are assigned. The long vehicle is modelled as it is presented in Fig 5a, bus in Fig. 5b, lorry in Fig. 5c and small carrier in Fig. 5d. For the vehicle recognition algorithm the 3D model of every vehicle class was elaborated; placed within road scenery.



Φιγυρε 5. Φουρ εξαμπλε χλασσεσ οφ ϖεηιχλεσ

For these vehicles sufficient description 16 points (P_i) of the model were used. They describe satisfactory the classes of vehicles; used for traffic objects modelling.

The vehicle is symmetrical in surface XY, where vertical length d_{vi} and horizontal d_{hi} define vertexes $P_i \in R^2$ of the model. Converting two-dimension (2D) measures into the 3D space, where $P_i \in R^3$, as in Fig. 6.



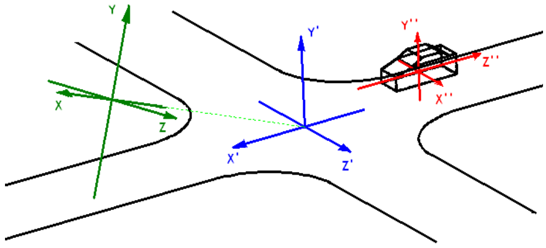
Φιγυρε 6. 3Δ σχενε ωιτη αν εξαμπλε ϖεηιχλε μοδελ

Finally, a mask of the model is obtained, on a base of the vehicle definitions in an observation field of video-camera. The scene of the model is related into coordinates n' =[0, 1, 0] of a road surface.

It was also assumed that the scene and vehicle's axes OY' and OY'' in both models are parallel situated. This fact considerably simplifies the conversion and rotation processes in coordinates XZ:

$$x' = x'' \cos(\alpha) - z'' \sin(\alpha), \qquad (7)$$
$$z' = x'' \sin(\alpha) + z'' \cos(\alpha).$$

These operations are converting the object's coordinates of the scene into coordinates of the video camera (Fig. 7).



Φιγυρε 7. Τηε χοορδινατεσ σχηεμε φορ ϖεηιχλε μοδελ χονϖερσιον

3.2. The coordinates conversion

The coordinates of the scene are converted into camera-coordinates by cosine angle multiplication [8,20], as:

$$a_{11}=\cos(\varphi)*\cos(\psi),$$
$$a_{21}=\sin(\varphi),$$
$$a_{31}=\cos(\varphi)\ *\sin(\varphi),$$
$$a_{12}=\cos\left(\pi/2-\varphi\right)*\cos(\pi+\psi),$$
$$a_{22}=\sin\left(\pi/2-\varphi\right),$$
$$a_{32}=\cos\left(\pi/2-\varphi\right)*\sin(\pi+\psi),$$
$$a_{13}=a_{21}*a_{32}-a_{22}*a_{31},$$
$$a_{23}=a_{12}*a_{31}-a_{11}*a_{32},$$
$$a_{33}=a_{11}*a_{22}-a_{12}*a_{21},$$

where: variables φ and ψ define angles values for the camera focus in relation to the observation middle point of the traffic lane. The relation of the cosine angle between respective tensors introduces the table 1.

Ταβλε 1. Τηε ϖαλυεσ οφ ανγλε χοσινε φορ χοορδινατεσ ξψζ ανδ ξ϶ψ϶ζ϶.

	x'	y'	z'
x	a_{11}	a_{12}	a_{13}
y	a_{21}	a_{22}	a_{23}
z	a_{31}	a_{32}	a_{33}

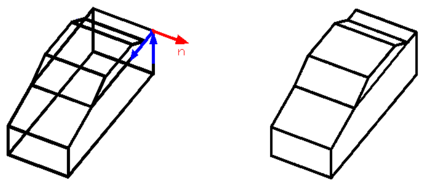
Values of angles φ and ψ are used for video-detector settings calibration:

$$x = a + a_{11}x' + a_{12}y' + a_{13}z'$$
$$y = b + a_{21}x' + a_{22}y' + a_{23}z'$$
$$z = c + a_{31}x' + a_{32}y' + a_{33}z'$$

The selected vehicle-class is put into the sixteen nodes $P_i \in R^3$, defined above (onto the camera's coordinates).

3.3. Removing invisible walls

The set of vehicle descriptors (in 3D space) are put into the image surface. Then, the conversion product contains the vehicle visible walls only (Fig. 8).



Φιγυρε 8. Τηε ινϖισιβλε ωαλλσ ερασινγ

For these invisible walls erasing, the vector notation of the model is needed. The edges of the model define the normal vector n, denoted by relations:

$$n = W_i + W_{i+1}$$
$$n = [\, y_i z_{i+1} - y_{i+1} z_i, x_{i+1} z_i - x_i z_{i+1}, x_i y_{i+1} - x_{i+1} y_i \,]$$

Then, the vector of optical axis k of the camera and normal vector n of the wall are extracted, by the equation (8):

$$n \circ k = |n||k| \angle \cos(n,k) \quad (8)$$
$$n \circ k > 0 \Rightarrow \angle(n,k) < 90 \quad (9)$$

The wall described by the condition (9) is invisible in camera observation field. The walls erasing procedure expresses the equation (10):

$$x_n x_k + y_n y_k + z_n z_k > 0 \quad (10)$$

where: n – is a norm vector of the wall, k – assigns a camera optical axe the perspective projection.

The 3D space vehicle model is not suitable for objects masks comparison algorithm. This operation has been done by more adequate 2D model; therefore the perspective projection of the 3D vehicle model, into the image surface was carried on. Every pixel of the recorded image has its angle product in the camera observation field.

The camera angle observation is obtained by copying each point, from the 3D space into 2D surface.

The conversion procedures are available by math equations presented below:

$$-0{,}5\frac{\alpha_V}{r_V} + \left(i \cdot \frac{\alpha_V}{r_V}\right) \le \frac{x}{z} < 0{,}5\frac{\alpha_V}{r_V} + \left(i \cdot \frac{\alpha_V}{r_V}\right),$$
$$-0{,}5\frac{\alpha_H}{r_H} + \left(j \cdot \frac{\alpha_H}{r_H}\right) \le \frac{y}{z} < 0{,}5\frac{\alpha_H}{r_H} + \left(j \cdot \frac{\alpha_H}{r_H}\right).$$

where: α – is the angle of the camera viewing line, r – the camera resolution, i, j – pixels indexing, K – point in the space.

4 The image comparsion

4.1. The model mask finding

After the invisible walls erasing the objects perspective projection is carried out, this way the 2D image is obtained. In the observation field of the digital camera the discussed above parameters are extracted from the image, similarly to segmentation procedures.

Thanks to parameterisation of the objects vertexes description, belonging to relevant walls. The segmentation operations are not needed in the discussed procedures.

Simple math operations provide us with pattern data that are responsible for parameterisation of the models' walls. Parameters p describe walls w_m in the model (11); similarly as in object OA_j in the image segmentation.

$$w_m = \{x_s', y_s', S', I_z', I_p'\} \quad (11)$$

The set of parameters, of every wall w_m, in the model of visible and covered walls, define the class M_c of the vehicle (12). For the defined five classes of vehicles five sets of pattern parameters are defined, in accordance with:

$$M_c = \{w_m\} \quad (12)$$

where: c – the vehicle class, m – index of the wall in vehicle model c.

Using the same description items of the objects and vehicles patterns in observation field of the camera, the simple comparison can be done; finally the vehicle class is extracted.

4.2. The pattern measures

For the above vehicle description items comparison, the Euclidean measures were used.

The distance between parameters sets that describe the vehicles patterns is used for the comparison procedures. The minimal distance appoints the most adequate (expected) class of the vehicle.

The Euclidean measure for the vehicles comparison is defined as follows:

$$d(M_c, OS_c) = \sqrt{\sum_{p=1}^{5} (M_c - OS_c)^2} \qquad (13)$$

where: OS_c – is a set of objects OA_j in number of visible walls of the vehicle's c pattern.

After the distance for every object OS_c is defined, the minimal distance $d(M_c, OS_c)$ defines the class of the current vehicle.

Due to use this method in an on-line measuring system the simplified measures are offered:

$$d(M_c, OS_c) = \sum_{p=1}^{5} |M_c - OS_c| \qquad (14)$$

The disadvantage of the Euclidean measure concerns domination of parameters describing the compared objects. Thanks to normalisation (objects scaling) of the objects parameters, into one space, the Euclidean domination was avoided.

5 Conclusions

Although the selected processing algorithm looks very simple, the method was implemented with satisfactory results. For all objects, similar to the defined patterns, the recognition level was almost 100%; the investigated simple and fast operators are recommended for image-processing unit development. They effectively support a very fast processing machines elaborated in Programmable Logic Arrays (PLA) technology; used for the on-line video controlling interface construction.

The elaborated and evaluated methods, with coefficient M_c of measure comparison, work more effectively for longer object edges. The method will also work better for a good contrast level of the image.

The critical conditions concern time limitations for all computing procedures defined for the on-line traffic control; the fastest but enough effective methods finding were subjects of these investigations.

The electronic equipment (Programmable Logic Arrays) implemented for this camera interface construction put the developers with these very restrictive parameters demands.

The contribution discusses our investigations final stage of new technology transfer into small industry, supported by EC funds.

Bibliography

[1] Battle J., Marti K., Ridao P., Amar J.: New FPGA/DSP - Based Parallel Architecture for Real-Time Image Processing, Real Time Imaging, Elsevier, 2002, Vol.8, pp.345-356.

[2] Earnshaw R. A.: Fundamental algorithms for computer graphics, Springer, Berlin, Germany 1985.

[3] Porikli F., Li X., Traffic congestion estimation using HMM models without vehicle tracking, IEEE Intelligent Vehicles Symposium, Parma, Italy, 2004.

[4] Fathi M., Atiquzzaman M.: A parallel pipeline based multiprocessor system for real-time measurement of road traffic parameters, Real-Time Imaging, Academic Press, 2000, Vol.3, pp.241-249

[5] Tan T. N., Baker K. D.: Efficient image gradient based vehicle localisation, IEEE Transactions on Image Processing, 2000 Vol.9, No. 8, pp. 1343-1356.

[6] Moon H.: Shape-encoded particle filtering for object detection and tracking, University of Maryland, USA, 2001.

[7] Obinata G., Dutta A.: Vision-Systems Segmentation and Pattern Recognition, I-Tech Education and Publishing, Vienna, Austria, 2007, ISBN 978-3-902613-05-9

[8] Laszlo S.: Theory of Three Dimensional Computer Graphics, Publishing House of the Hungarian, Academy of Science, Budapest, Hungary, 1995.

[9] Koller D., Daniilidis K., Nagel H. H.: Model-based object tracking in monocular image sequences of road traffic scenes, International Journal of Computer Vision Vol.10, No.3, 1993, pp.257-281.

[10] Böckert A.: Vehicle detection and classification in video sequences LiTH-ISY-EX-3270, 200.2

[11] Ferryman J. M., Worrall A. D., Sullivan G. D., Baker K. D.: A generic deformable model for vehicle recognition, Proceedings of British Machine Vision Conference, Birmingham, England, 1995, Vol.1, pp. 127-136.

[12] Malik J., Belonge S., Leung T., Shi J.: Contour and texture analysis for image segmentation, International Journal of Computer Vision, 2001, Vol.43, No 1, pp.7-27.

[13] Korn, A.F.: Towards a symbolic representation of intensity changes in images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1988, PAMI-10:610-625.

[14] Płaczek B., Staniek M., Piąstka K.: The hardware interface solution for video camera vehicles tracking algorithm implementation. Proc. of Int. Conf. Telematics and Transport Safety Vol.5, Katowice, Poland, 2005, pp. 30-38.

[15] Płaczek B. Staniek M. Research reports: "Moduły wideo-detektorów pojazdów ZIR-WD do sterowania i nadzoru ruchu drogowego, 512/11/475/06/FS-11. 2007 (in polish).

[16] Tadeusiewicz R., Korohoda P.: Komputerowa analiza i przetwarzanie obrazów, Wydawnictwo Fundacji Postępu Technicznego, Kraków, 1997 (in Polish).

[17] Nieniewski M.: Segmentacja obrazów cyfrowych. Metody segmentacji wododziałowej, Wydawnictwo ELIT, 2005, ISBN: 83-87674-82-6, (in Polish).

[18] Choraś R.: Komputerowa wizja. Metody interpretacji i identyfikacji obiektów, Wydawnictwo Elit, 2005, ISBN: 83-87674-89-3, (in Polish).

[19] Piecha J. Staniek M. The vehicle description method by means of its excreted image recorder by video camera. Proc. of Int. Conf. Transportation and Logistic Integrated System, Cracow, Poland, 2007, pp. 204-215.

[20] Płaczek B., Staniek M.: The vehicles tracking algorithm for hardware video-detection implementation. Transactions on Transport Systems Telematics. Modeling, Management and Image Processing, Ed. J. Piecha Monography, Gliwice, Poland, 2007, ISBN 978-83-7335-422-7.

[21] Płaczek B., Staniek M.: Model based vehicle extraction and tracking for road traffic control. International Conference on Computer Recognition Systems 2. Wroclaw University of Technology. Systems and Computer Networks. Wroclaw, Poland, 2007, pp. 844-852.