Application of a Cumulative Method for Car Borders Specification in Image

Martin DOBROVOLNÝ, Pavel BEZOUŠEK, Martin HÁJEK

Faculty of Electrical Engineering and Informatics, Univ. of Pardubice, Studentská 95, 532 10 Pardubice, Czech Republic

martin.dobrovolny@upce.cz, pavel.bezousek@upce.cz, martin.hajek@upce.cz

Abstract. *The paper deals with low-level car detection methods in images. The car detection is an integral part of all intelligent car cruise systems. This article describes the method for suppression of the edges produced by classical edge's operators, based on application of the cumulative method. The designed method uses the non-stationary property of the picture background in time-realizations of the image signal.*

Keywords

Image processing, cumulative method, Intelligent Driver Systems (IDS), car detection, edge detection.

1. Intelligent Driver Systems

The Intelligent driver systems (IDS**)** will certainly contribute to the enhanced safety on roads in future as well as to the further reduction the negative consequences of accidents. One part of the *IDS* systems is a subsystem based on an on-board camera observing the car vicinity. This system with a build-in camera represents an intelligent fellow-rider, monitoring the driver, the car and the situation around the vehicle. It also gives navigation information to the driver and informs him about critical situations.

A lot of partial problems should be solved in this area. One of the most important problems is a great over-complexity of the captured traffic scene. The usually applied edge detection methods [1], [2], lead to over-segmentation. The whole problem is much more complicated by the non-stationary character of the background [3]. The standard methods applied in image analysis with the non-stationary background use the optical flow analysis for instance [5], [6], but these methods often fail in the case of a highly complicated image. The optical flow linearly decreases towards the axonometric center of the image where it completely vanishes [4]. Another approach to the segmentation is often based on a suitable color conversion. The input image is converted into a convenient color space (mostly *HSV*), which suppresses the brightness sensitivity [7]. These methods are unusable by the infra-vision systems [8], [1].

In the next part of the article, the original method for suppression of the edges produced by the standard edge detectors most often based on application of the cumulative method is described in detail. At the input of the present method the vector of approximate cars coordinates is used and at the output we get an image with only car's outlines depicted while the other edges are suppressed.

2. The Cumulative Method for Car's Outlines Search

Traffic scene objects can be split into several groups according to the speed of the intensity function variations in equidistantly captured pictures. The main presumption is a difference between a small optical flow in the area of the car and a massive optical flow in background pixels. Consequently only a minimal optical flow inside the outlines of the moving object and its constant size could be assumed. Due to the movement of the camera carried by a moving car the background is also non-stationary. For this reason the background including the road surface and other objects embodies a massive optical flow and the fast sizes variations due to an axonometric distortion in time. But these assumptions are not quite satisfied. For instance the intensity function progression inside the car outlines is affected by variable lighting conditions during the car movement and the size of the car picture variations depend on the relative speed of the object and the car with camera. The background optical flow is variable, too. Due to the axonometric distortion the flow is zero at the middle of the horizon, whereas at the edges of the scene it reaches its maxima.

If suitable transforms suppressing some of these effects are applied it is possible to separate the objects in successive images and to emphasize the car borders and at the same time to suppress the background by application of the cumulative method.

2.1 The Method Description

The method starts with an initial frame I_n and with coordinates $S=\{x_1, y_1..x_p, y_p\}$ of p potential objects obtained by one of the Car-detection methods described in [4]. For each object O^i, $i = 1…p$ it is then necessary to find the equivalent image segments in the next *n* frames with relevant shifted objects. For this reasons the template of size $T^i=(x_i \pm \lambda_i, y_i \pm \lambda_i)$ is stripped from the initial frame, where λ_i is the smallest estimated object *i* neighborhood radius and x_i, y_i are the inaccurate coordinates of the object center. This template is searched in the next maximum *k* frames to I_{n+k}. If the object O^i maximum shift from the initial point in the image plane could be restricted to ΔS_{max}, then the selective scanning window can be defined as:

$$W^i = \left(S^i \pm n \cdot \Delta S_{\max}\right). \qquad (1)$$

The searching is then realized only in the window W^i, which increases the computation speed.

In Fig. 1 an example of the input frame I_n is presented including the first object axis S^1 (green). In the figure also the template $T^1=(x_1 \pm \lambda_i, y_1 \pm \lambda_i)$ is depicted in red, which will be searched in the search window W^1 (yellow) in next max. *k* frames.

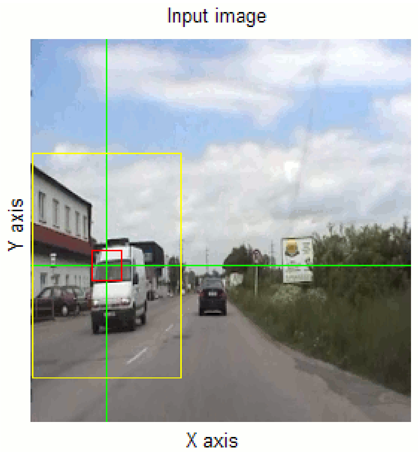

Fig. 1. The example input image with depicted center and the search window W^i

The object T^i position searching in the window W^i with proportions *(u, v)* is usually realized by two-dimensional correlation algorithm [9]. It is possible to partially improve and optimize this calculation regarding to the computational speed. When the mean value of the images I_{n+k} is removed the normalization could be let out (only the relative maximum r^i value of the correlation coefficient is of interest), then it is possible to simplify computations:

$$r^i{}_{corr2}(u,v) = \sum_x \sum_y \left(I^i(x+u,\, y+v) T^i(x,y)\right). \qquad (2)$$

The result of the relation (2) is the matrix $\boldsymbol{R}^i(u, v)$ with calculated correlation coefficients between the search window W^i segments and the template T^i from the initial image (Fig. 2-left). The best template T^i position localization is then realized by a search for the *global maximum* in the matrix $\boldsymbol{R}^i(u, v)$. It is convenient to apply the *Gaussian 2D low pass filter* before the searching process starts up to reduce the possibility of sticking in some local extreme. During the search for the positions of the template T^i in *k* frames by (2) as much as $k \cdot (u-x) \cdot (v-y) \cdot x \cdot y$ multiplications should be realized which slows down the computation speed. The speed improvement may be achieved by the search window W^i size reduction, which is however limited by the condition (1) or by the template T^i size reduction, which restrains the position accuracy.

The significant speed improvement without the undesirable effects connected with the size reduction of matrices W^i or T^i may be achieved by exchanging the standard two-dimensional correlation by one of the popular *SAD*[1] algorithm [10]:

$$r^i{}_{SAD}(u,v) = \sum_x \sum_y \left|\left(I^i(x+u, y+v) - T^i(x,y)\right)\right|. \qquad (3)$$

Due to the *SAD* algorithm simplicity it is extremely fast and effective. It can be fairly implemented on most of today's processors. The main finesse of this algorithm is a substitution of a number of multiplications by the same number of additions requiring much shorter computation time. The further advantage of the SAD algorithm over the two-dimensional correlation is its insensitivity to the mean value of the W^i and T^i matrices, which eliminates the necessity of the mean value computation.

The result of (3) is again the matrix $\boldsymbol{R}^i(u, v)$ similarly to the relation (2). The best template T^i position is realized by the search for the *global minimum* in the matrix $\boldsymbol{R}^i(u, v)$ and again before the searching process start the *Gaussian 2D low pass filter* is applied.

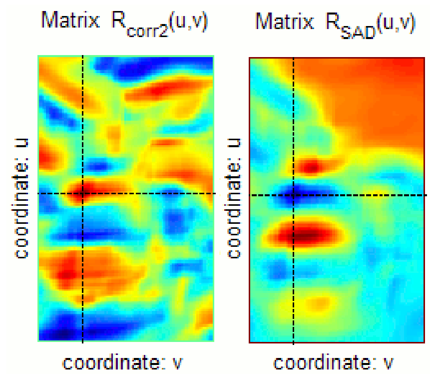

Fig. 2. Examples of the matrix $\boldsymbol{R}^i(u,v)$ obtained by two-dimensional correlation and by the *SAD* algorithm for I_{n+3} (blue marked the minimum, red marked the maximum).

In Fig. 2 we may compare results of the $\boldsymbol{R}^i{}_{corr2}(u, v)$ matrix with marked maximum and $\boldsymbol{R}^i{}_{SAD}(u, v)$ matrix with the minimum for the I_{n+3} frame after the initial frame in Fig. 1. We may see that *SAD* algorithm achieves the same accuracy with more than six time shorter[2] computation time.

When the maximum realistic object shift speed is assumed in the image plane between the frames, the search window of the objects near the image edges can fall par-

---

1 SAD – Sum of Absolute Differences

2 By the window size W^i = 151 x 101 pixels and the template size T^i = 21 x 21 pixels

tially out of the search window due to the real 3D-scene motion. This situation is typical particularly in the case of arriving or turning cars. In such a case it is useful to restrict the maximum change of coordinates by the following condition:

$$S^i_{new} = \left(S^i \pm n \cdot \Delta S_{\max}\right), S^i_{new} \in \left|\left(size\left(I_{n+k}\right) - \xi\right)\right| \qquad (4)$$

where ξ represents the protective zone on the image I_n edge. The protective zone allows a partial shift of the search window W^i outside the I_{n+k} image coordinate. In this case only a part of the searched object in the search window is found and the image extension by pixels with null brightness should be realized. The further search in the next I_{n+k} images is then stopped (Fig. 3- the last snap I_{n+4}).

The result of the procedure described above is the matrix of images representing objects shifted against their starting positions over the non-stationary background. In Fig. 3 the result of the search for the first object from Fig. 1 is presented. The object centre (marked by red color crossing) stay always in the center of the new founded window W^1. The shifting border ξ, which limits the calculation process at image edges, is indicated by the green line. The situation with the object extruded partially out of the image is presented in the last frame (I_{n+4}). In this case the calculation process is stopped (the condition (4) is not satisfied).

Fig. 3. The resulted series of a tracked object.

2.2 The Scale of the Founded Objects Normalization

Fig. 3 represents the object moving in the image plane. The realistic movement of objects proceeds in a 3D space comprising objects coming near as well as receding objects. Due to this reason the objects picture sizes are changing.

The object size should then be normalized to the size from the initial frame. The normalization of the I_{n+k} frame is realized by creating the series of images with variable zoom in the range from z_{min} to z_{max} (typically 0.5 – 2). These are compared with the initial image.

$$Z^i_{SAD}(z) = \sum_x \sum_y abs\left(I^i_n(x,y) - zoom\left(I^i_{n+k}(x,y)\right)\right). \qquad (5)$$

In Fig. 4 the function $Z(z_{min} - z_{max})$ response is presented showing a minimum for images with equivalent scale.

Application of this procedure for all the images from Fig. 3 leads to the series $J^i(1..k)$ of identical objects with the same shift and scale in the image (Fig. 5).

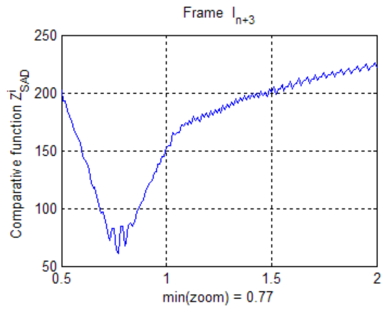

Fig. 4. The example of function Z^i_{SAD} for the frame I_{n+3}.

Fig. 5. The resulting images with the normalized size of objects.

2.3 The Cumulative Method Application

After application of the previous operations it is possible to use a *cumulative method* on the image signal to enhance the pixels representing cars and their borders. The *cumulative* addition method is based on relatively simple principle of correlated signals enhancement in non-correlated background. The maximum effect could be obtained in the case of ideally correlated object pixels in totally uncorrelated background (white noise) [9]. In our case the ideal signal is represented by the high correlated car's pixels encased with the non-correlated moving background's pixels.

But neither of the presumptions is fully satisfied in the real signal. Though the shape and the size of the objects are equalized after previous transforms, the car pixels have some properties, reducing the pixels correlations:

- The capture system produces the noise which negatively affects the image signal.
- The car pixels intensity is affected by variable local and global lighting conditions.
- The previous transforms add complex and hardly to describe degradations to the signal (change of the scale followed by the filtration, image extensions...).

Also the background's pixels have not purely the white noise character. When the car is not moving and the background is stationary then the *cumulative method* cannot be used because there is only a small difference between the car pixels after fixation of the moving objects by the above-described methods and the background pixels. If the vehicle is moving then we may suppose a massive optical flow also at the background pixels. Due to the car fixation by the previously described methods in individual frames the entire optical flow is concentrated mainly to the background pixels. With the growing speed difference between the car and the background the correlation falls down and the method becomes effective.

Before the cumulative method is applied the mean value from each signal realization is removed (Fig. 5). In the next step the individual realization are added to the cumulative signal. Regarding to the fact that the signal is not ideal the signal to noise ratio (*SNR*) enhancement decreases with an increasing number of cumulated frames. That is why the *SNR* parameter is evaluated in each step and if the *SNR* enhancement is less than a defined limit α, the addition process is stopped. In the same way also special situations are handled for instant when the light conditions change causes that the further addition does not substantially increase the output signal quality. The *SNR* parameter is represented by edge number acquired by edge operator in the region of interest (in the vicinity of the middle choice window W^i) against the number of edges in a cumulative image S created from J^i_n images:

$$SNR_i = \frac{\left|E\left[edge(W_i)\right]\right|}{\left|E\left[edge(S)\right]\right|}. \tag{6}$$

The result of this procedure is the cumulative image *S*, with greatly highlighted car pixels. The resulting accumulated image of the first object from Fig. 1[3] is shown in Fig. 6. In the figure the car body segments acquired by previous process can be simply recognized. The other parts of the car like the chassis, the front window, etc. are only minimally emphasized due to the variations in the lightening. The spacing between the highlighted parts and the background is minimally about one order. Also the sky segment is considerably correlated and the optical flow here is minimal. Using this fact the sky can be very well suppressed by using the *mask of sky* obtained by the clustering methods as described in [4]. Since the sky does not contain abrupt variations of the intensity function it does not cause over-segmentation of the image.

2.4 The Outlines Extraction from the Summary Image

In Fig. 6 the procedure of the car outlines extraction is presented. The horizontal and vertical car outlines are extracted separately (Fig. 6 - 1st column). In the first step the Sobel's edge operator with forced orientation extracts the edges. There are artifacts in the picture, made by the summary of the extended images (Fig. 5). To suppress over-segmentation the correction of the boundary artifacts is first applied (Fig. 6 - 2nd column). Then the morphological opening using a structural element in a shape of a vector of *1x3* a *3x1* pixels is applied on the images of the vertical and horizontal edges (Fig. 6 - 3rd column) removing the edges of an undesirable shape. In the image produced this way a segment with the most number of edges is chosen. In this segment the gravity centre is calculated (Fig. 6 - 4th column) representing the newly specified vertical and horizontal center of the car. If vertical and horizontal object symmetry is assumed, then it is possible to tilt the object image over the centre and to add it up. After this the car boundaries with maximum accuracy could be determined (Fig. 6 - 5th and 6th column).

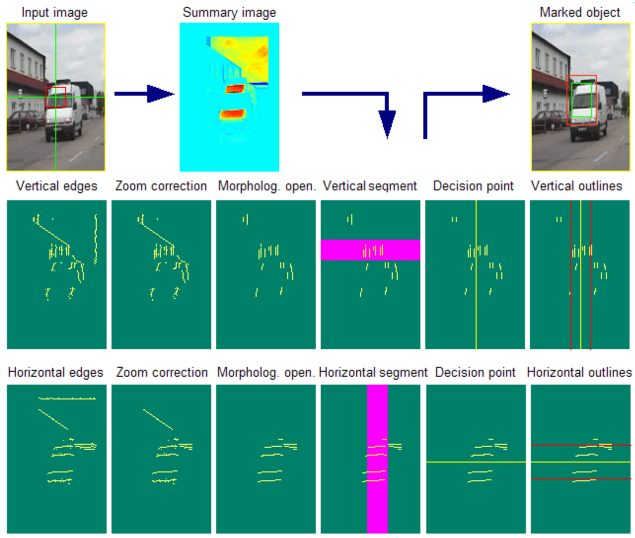

Fig. 6. The boundaries extraction from the summary image.

By this way the *new corrected center and shape* of the car is obtained. The obtained outlines of the car represent only those areas highlighted by the cumulative method. The boundary parts of the cars have in the real *3D* scene more variable lighting during the car motion due to the carriage curvature and that is why they are not highlighted by the previous method as much as the inner parts. This fact is corrected by multiplicative factor η (typically $\eta = 1.2$), which magnifies the predicted boundary range (Fig. 6- red box) that are the better representation of the real car outlines.

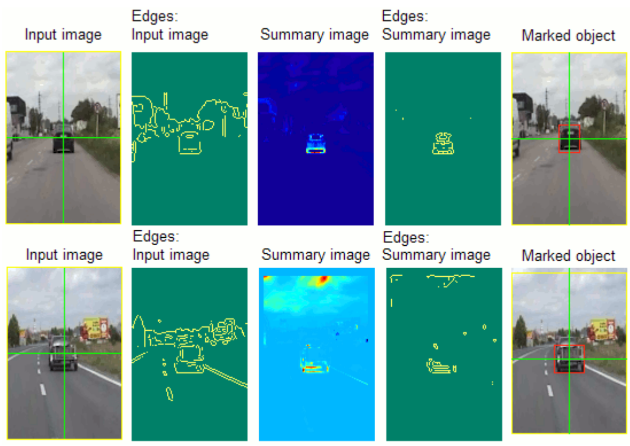

Fig. 7. The car outlines detection by the cumulative method.

3 For a better contrast the color map is chosen emphasizing of the maxima (red) and the minima (blue) of the signal applied.

3. Conclusion

Fig. 7 presents examples of detection of the other objects (the red outlines) together with the summary images (the 3rd column) acquired by the above described techniques. In the top line of the images in Fig. 7 the second objects from Fig. 1 is presented. In the bottom line the object moving in a slightly right–hand curved road is presented. For the demonstration of the procedure contribution the images created with the *Sobel dual direction edge operator* on the original and on the cumulative picture (the 2nd column) are displayed. The picture shows the great difference between the number of edges in the car area and number of edges in the suppressed neighborhood of the car (the 4th column).

In both cases the leading cross (green) is also depicted indicating the car centre. The position of the cross is successively corrected and the car contours are calculated using the above described procedures. By these procedures the car borders are correctly specified even in the case of highly inaccurate input centre coordinate.

The further examples of the objects detections are presented in Fig. 8. For clearness the whole input picture I_1 including the search window W^1 is presented. The new centers of the objects are indicated by green lines and their outlines by the red ones.

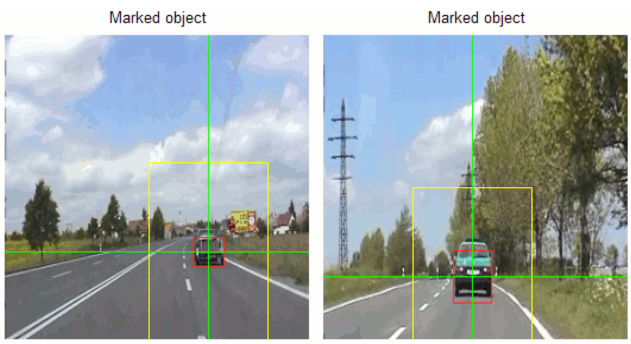

Fig. 8. The other objects detection examples.

The method described in this article effectively filters the undesirable edges produced by edge operators in the background pixels area. The detection of the cars is robust even in the case of a very complex image. The efficiency of the described method increases with the increasing difference of the objects speed against the non-stationary background. The positive advantage of the method is the possibility to change the accuracy/speed ratio by scaling of the searching window W^i and the searching template T^i. The computation time was improved using the *SAD* algorithm whereas the accuracy does not change substantially.

Acknowledgements

The research described in the paper was financially supported by the Czech traffic department under grant No. CG743-037-520.

References

[1] BETKE, M., HARITAGLU, E., DAVIS, L. Multiple vehicle detection and tracking in hard real time. In *IEEE Intelligent Vehicles Symposium*, pp. 351–356, 1996.

[2] ZHAO, T., NEVATIA, R. Car detection in low resolution aerial images. In *IEEE Int. Conf. Computer Vision*, p. 710-717, 2001.

[3] YING, R., CHIN-SENG, C., YEONG, K. Motion detection with nonstationary background. *Machine Vision and Applications*, Springer-Verlag, 2003

[4] DOBROVOLNY, M. On-road vehicle detection based on image processing. *Doctoral dissertation*, University of Pardubice, 2008

[5] CHRISTOPHE, P. B. Discrete wavelet analysis for fast optic flow computation. *Applied and Computational Harmonic Analysis*, July 2001, vol. 11, no. 1, pp. 32-63.

[6] BRUHN, A., WEICKERT, J., SCHNORR, C. Lucas/Kanade Meets Horn/Schunck: Combining local and global optic flow methods. *International Journal of Computer Vision*, 2005, vol. 61, no. 3, p. 211–231.

[7] CHATUVERDI, P. Real-time identification of driveable areas in semi-structured terrain. *Proc. SPIE*, vol. 4364, p. 302-312, *Unmanned Ground Vehicle Technology III*, 2001, 2001SPIE.4364..302C.

[8] ANDREONE, L., ANTONELLO, P. C., BERTOZZI, M., BROGGI, A., FASCIOLI, A., RANZATO, D. Vehicle detection and localization in infra-red images. In *The IEEE 5th International Conference on Intelligent Transportation Systems*, 3-6 September 2002, Singapore. pp. 141-146.

[9] CASTELMAN, K. *Digital image processing.* Prentice Hall, September 2, 1995, ISBN: 0-132114-674

[10] FORSYTH, D., PONCE, J. *Computer Vision: A Modern Approach*. Prentice Hall, August 24, 2002, ISBN: 0-130851-98-1.

About Authors...

Martin DOBROVOLNÝ was born in the Czech Republic in December 1976. He received his M.Sc. (2003) and Ph.D (2008) degrees from the Jan Perner Transport Faculty, University of Pardubice, Czech Republic. He is interested in image and signal processing and the computer networks.

Pavel BEZOUŠEK was born in 1943. He received his M.Sc. degree from the CTU, Prague in 1966 and his Ph.D. degree from the same university in 1980. He was with the Radio Research Inst. of Tesla Pardubice from 1966 till 1994, where he was engaged in microwave circuits and systems design. Since then he has been with the University of Pardubice, now at the Inst. of Electrical Engineering and Informatics. Presently he is engaged in radar systems design.

Martin HÁJEK was born in the Czech Republic in January 1978. He received his M.Sc. degree from Jan Perner Transport Faculty, University of Pardubice in 2003. Since 2003, he has been a Ph.D student at Jan Perner Transport Faculty, University of Pardubice. His research interests include digital signal processing, applications of FMCW radars and implementing various methods of signal processing into real hardware.