 DOI: 10.21307/tp-2015-006

Keywords: binocular technical vision; wheeled robot; controlling using camera

Anatoliy KULIK*, Kostiantyn DERGACHOV, Oleksandr RADOMSKYI
National Aerospace University "Kharkiv Aviation Institute"
Chkalova str. 17, Kharkiv, 61070, Ukraine
**Corresponding author.* E-mail: anatoly.kulik@gmail.com

BINOCULAR TECHNICAL VISION FOR WHEELED ROBOT CONTROLLING

Summary. Proposed model of technical vision system use for supplementation and clarifying information about surround objects classes and distances, received by driver.

БИНОКУЛЯРНОЕ ТЕХНИЧЕСКОЕ ЗРЕНИЕ ДЛЯ УПРАВЛЕНИЯ КОЛЕСНЫМ РОБОТОМ

Аннотация. Предложена модель применения системы технического зрения для дополнения и уточнения информации, получаемой водителем транспортного средства, о характере и расстояниях до окружающих транспортное средство объектов.

1. INTRODUCTION

Over the past few decades there has been considerable progress in the field of digital image processing, the use of high-speed small photomatrixes as well as new methods of information processing opens up the possibility of new technologies that can be applied in various fields of science and technology.

The main task of the movement of vehicles is a safe move to different conditions, for example, in town.

Studies of human thought shows that the average person cannot simultaneously control more than 5-8 parameters, with the proviso that person is rested, if a person is tired, its reduced to 2 - 3 parameters. Thus, when driving in town, this limit is only limited to vehicle parameters, and surrounding objects without the proper amount of attention, resulting in traffic accidents of varying severity. This is confirmed by statistical studies of the causes of accidents on the roads.

The paper proposes the use of a vision system to supplement and clarify the information, obtained by driver of the vehicle, about surrounding objects. As a result of a vision system functioning, can be constructed view of the object-surrounding world, it is the basis for a vehicle mobility map calculating.

This wills both expand the range of application of ultrasonic parking system, and will provide an opportunity not to use sonar or radar to map objects, due to the significant deficiencies last when used in heavy city traffic.

2. STATEMENT OF DESIGNING

To address the problems of road safety management system it is necessary to develop a mobile unit using binocular vision, the provision of information about the objects that could potentially cause an accident - vision system, which includes several pairs of cameras.

Sufficient quality of the solution can be provided by using multiple video cameras located, depending on the measurement range. With sufficient resolution, speed and precision of the relative fair chambers, the mathematical model and the principles described in Piecha [1] as well as a number of studies [2, 3, 23], allows to build a three-dimensional picture of the observed scene with high accuracy. There are two ways to obtain a series of images of the observed scene for further transformation with the model:

– The use of a mobile camera.

– The use of a few calibrated cameras.

Both approaches have both advantages and disadvantages, but in the case of using such a system for a vehicle, it makes sense to use multiple cameras, because while vehicle driving there is a lot of periodic and aperiodic oscillations with different frequencies and amplitudes. A series of images obtained by a single camera will be noisier, by aperiodic movements of the vehicle that can be adequately exclude by using of multiple cameras. This software has the following tasks:

– Calculation of the objects coordinates that surrounds vehicle in vehicle coordinate system.

– Making a map where will be indicated vehicle and surrounding objects, distances to the nearest objects and their speed.

Known work Piecha [1] in which we solve the problem of monitoring moving objects at a fixed motionless video source. A significant advantage of the approach to solving the problems of identification of moving objects in the observed scene is the simplicity of the mathematical apparatus and the low demand for computing resources that allows real-time tracking of up to ten objects of different types. Some details of the method discussed in [3]. Direct usage of the methods, described in [1, 3] for objects in a mobile video source controlling is not possible due to a number of basic principles on which the decision was based Piecha [1]. Such of the principles underlying the method Piecha [1], as immobility video source, a priori known dependence of the direction of movement of the object in the camera view when it straight or circular motion, decision to deny the possibility of applying to the mobile video source. Circumvent the difficulties encountered allows principles, described in the works [5, 6]. Thus, as the material carrying information about moving objects, it is proposed not to use the raw data from a video source, are processed and based on the known parameters of the main types of objects as well as the properties of the objects displayed on the camera plane, but using of the pre-prepared data, obtained from multiple cameras.

The procedure described in [6] model to determine the distance to each of the discrete, into which the image obtained from camera is split, but does not allow to highlight objects on the image. Integration of methods [1, 3] and [6] will provide an opportunity to apply real-time system Piecha [1] on the vehicle.

3. GOALS, OBJECTIVES, STRUCTURE AND THE WAY FOR PROBLEM SOLVING

Known variants of the installation of such vision systems for mobile objects, but they have some drawbacks. The most significant drawback is the difficulty to obtain a three-dimensional picture of the observed scene and the deposition of images. In such systems, independent video sources are used, and the computer solves task for each of them. Alternatively, this is a system with a pair of cameras, operating range of the system is determined by the manufacturer and the system can be used only for a certain class of vehicles and for certain speed. The main parameter that affects the operating range of vision systems, which implements the principles [6], is the distance between the cameras. Manufacturers offer a range of systems with different numbers of camera modules, but using of a single camera provide the result of a motion control that does not satisfy the wide requirements. Using of a multiple cameras do not solve the problem of the system operating range expanding, distance between pair of cameras, combined in separated functional blocks, is limited by manufacturer.

The aim of this study is to investigate the possibility of using a vision system on the vehicle and investigate ways to increase the operating range of the system.

Known works that has invited the camera in any way, depending on the nature of the work performed by the vehicle. The main ones are, first of all, offers send to the axis of the system in the direction most likely to appear moving objects of interest. On some rail vehicles, in general, such systems are directed toward the passengers, because at the opposite side, in most cases, can be only transport of the other direction, which trajectory and velocity are known. For vehicles are invited to submit axis of functional elements with pairs of cameras, along the axis of the vehicle or, in case when system consists of one element with pair of cameras, to direct its axis with the axis of the long range beam.

This paper proposes the use of four cameras located on the same horizontal line on the bumper of the vehicle, as follows as shown in Fig. 1.

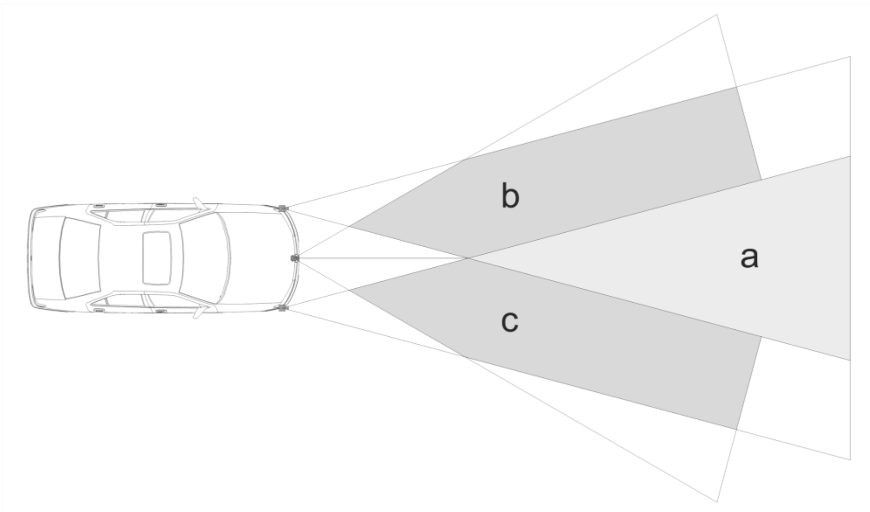

Fig. 1. Operating range of the system
Рис. 1. Рабочий диапазон системы

The two side cameras are directed on the vehicle motion line, combined into a single functional unit, that used to determine the distance to objects, that far than vehicle length (Fig.1 a). The cameras located on the length axis of the vehicle are arranged at angles to the axis of motion and each one, in conjunction with the side camera, allows to determine the distance to objects, that are near to vehicle on left and right side accordingly (Fig.1 b,c).

The developed system will improve the task of a road safety improvement. Having considered the various ways of creating vision systems for vehicles, form kind of a vision system.

Composition of the system is determined by its functions and includes:

- Video cameras delivers images of the observed scene behind the object;
- A computer that provides image processing, implementation of the scene recovery model, making a digital map, display the output of a digital map to the user;
- The car.

The concept of construction of the circuit shown in Figure 1, is characterized by the following fundamental features. Where: α - Angle of camera view, β - Angle between view directions of a two center cameras, A - Distance between side cameras, B - Distance between side and center cameras.

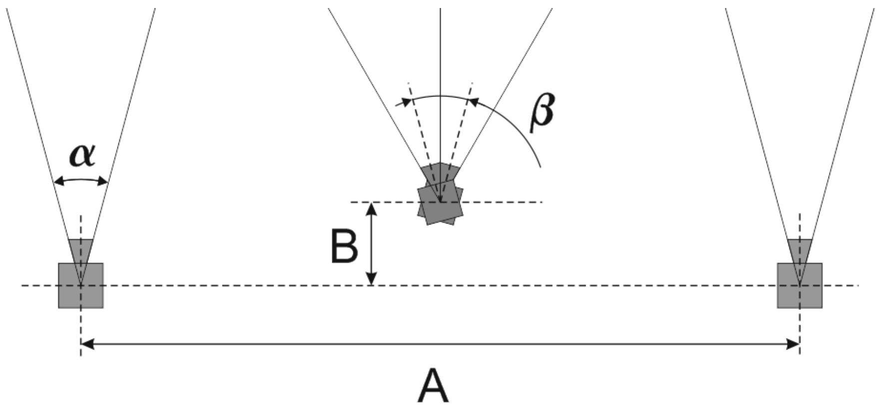

Fig. 2. Main geometric parameters of a technical vision system
Рис. 2. Основные геометрические параметры системы технического зрения

To increase the range of the system can be brought additional pairs of cameras. One of the methods to expand the range to the right and to the left of the base due to the mutual arrangement of the side cameras, is the method of using additional technical resources - additional cameras. Obviously fashion add a pair of cameras on each side respectively. Such a system will consist of six cameras.

The peculiarity of the proposed system operation, the two cameras are located on the axis of the vehicle at an angle to this axis. This arrangement makes it possible to detect objects that are outside the range of a pair of side cameras without additional technical resources. This reduces the amount of resources, required to implement the system, reduces the load on the computer as well as the power consumption of the system. In contrast to the intensive techniques integral accumulation of information use was requested [11, 13]. This approach allows us to obtain not only the number of moving objects in the measurement, but also their characteristics such as speed, direction of travel and the class of the object.

To solve the problem of the use of video as a source of information about the dynamic characteristics of the observed objects asked a series of algorithms for data preparation.

The first step is to determine the size of the data that is sufficient to ensure the quality of the results of calculations [7]. There are various approaches to the preliminary size reduction coming from the video source data [14].

To reduce the amount of data to be processed in paper proposes the use of the packet data samples at regular intervals. Such an approach to reduce the required computing power is used in a variety of theories of computation. The method allows to share data packets to the class, the clusters subset of computational intervals and [15].

Instead of sequential processing of information about all the number of moving objects visible in view of the system, the method allows to process information about objects of the same type as a whole.

Numerous studies aimed at identifying the knowledge, embedded in the data [16-19], the approach is considered in computing systems, data mining. Thus, it is possible to directly determine the temporal characteristics of the test to a group of objects.

Experiments show that the sampling frequency detector of moving objects may be significantly lower with decreasing vehicle's own speed. This is an indirect sign of addiction to the sampling rate, which is not related to the number of observed objects [16].

The detection system of moving objects and algorithms for decision-making to provide not only an ordered list of mobile objects, their speed and direction of movement, but also possible solutions to prevent critical situations, such as collisions [13].

In order to get a satisfactory solution using this method uses some of the principles of fuzzy logic in the processing of information. This combination provides the findings in a fuzzy range of values [12, 20]. In the object detection system using the video camera field of view is divided into multiple zones. Thus, the detection results are used directly in the model automatically detecting objects using the partition for observation zone [22]. In many cases, it can automatically identify the class of the selected objects to the discrete model of the observed situation where the union of objects in classes with traffic management plays a key role. Parameters of road traffic should be calculated for the specified region of the observed scene during the sampling interval.

Thus obtained sample interval, after which it is used for further analysis. The identification of data packets allows for batch computing methods.

Analysis of the sample rate of the data consists of several phases:

a) possibility of packing the data used for control.

b) the statistical parameter prediction of the observed scene, based on the history records data changes.

c) the speed of image sampling necessary for a satisfactory analysis in this traffic.

First, a description of moving objects has to be done by microscopic pattern, wherein the motion parameters assigned as fuzzy numbers [12, 21]. The random nature of the behavior of objects of interest is taken into account in terms of probability of the speed of their movement at every step of the simulation.

In the discrete model of mobile objects each weighting coefficient is assigned as a set of fuzzy coefficients.

The status of each of the observed scene determined by the number of moving objects. The current state of the observed scene in general can be written as a set of cells, each of which is in his state:

$$C_L = \{C_{1L}, C_{2L}, ... , C_{n(L),L}\} \tag{1}$$

where: L - denotes the granularity level, C_{1L} - is a state of cell number i, at granularity level L, $n(L)$ - is a number of cells at granularity level L

At the lowest granularity level (L = 1) states of the cells have binary values: $C_{i1} = 0$, if the cell i is empty and $C_{i1} = 1$, if object is in the cell i.

At higher granularity levels ($L > 1$) the cells state values denote number of vehicles belonging into the cell:

$$C_{iL} = \{0, 1, ... , L\} \tag{2}$$

3.1. Mathematical model

In Piecha [1] describes a method of selecting objects in the image obtained with the help of a typical small-sized CMOS cameras, which are used for the registration of road traffic. It is also known to work [4] proposed a model in which the classification of such objects in the four groups.

The proposed model involves the use of a digital image Sobel Orthogolal Mask Operator

$$Mi,j = |Mhi,j| + |Mvi,j| \tag{3}$$

where: Mhi,j- lightness of the pixel attribute value for horizontal edges, Mvi,j- lightness of the pixel attribute value for vertical edges $i = 0 .. n-1$, $j = 0 .. m - 1$

Image binarization can significantly reduce the amount of memory required to store and process the image.

Well-known mathematical model [6] reconstruct the scene for a number of images used as the input color images, scenes shot for greater accuracy, with a maximum resolution of video cameras.

This approach is fully justified if sufficient computational power and the absence of restrictions on the speed of your camcorder. In the case of this work, to video cameras apply strict requirements for speed and dimensions.

Proposed to improve the algorithms that implement the recovery model of the observed scene for a number of images, the introduction of an additional module that implements the model binarization image. The work is to investigate the change of accuracy determine the distance to objects of interest in the dependence of the image as well as the time required for the simulation result.

3.2. Camera model

The general model of a perspective camera that will be used here is that represented by an arbitrary 3×4 matrix, P, known as the *camera matrix*. The camera matrix transforms points in 3-dimensional projective space to points in 2-dimensional projective space according to the equation $\mathbf{u} = P\mathbf{x}$.

The camera matrix P is defined up to a scale factor only, and hence has 11 independent entries. As shown by Strat [22], this model allows for the modeling of several parameters, in particular: the location and orientation of the camera; the principal point offsets in the image space; and unequal scale factors in two directions parallel to the axes in image space.

This accounts for 10 of the total 11 entries in the camera matrix. It may be seen that if unequal stretches in two directions not aligned with the image axes are allowed, then a further 11-th camera parameter may be defined. In practical cases, the focal length (magnification) of the camera may not be known, and neither may the principal point offsets. Strat [22] gives an example of an image where the camera parameters take on surprising values.

Our purpose in treating general camera transformation is to avoid the necessity for arbitrary assumptions about the camera. If the camera is not placed at infinity, then the left-hand 3×3 sub matrix of P is non-singular. Then P can be written as $P = (M \mid -MT)$ where T is a vector representing the location of the camera. By the method of QR-factorization [23], M may be written as a product $M = KR$ where K is upper triangular and R is a rotation matrix.

The matrix K represents the so-called internal parameters of the camera. If K is known *a priori*, then we say that the camera is *calibrated*. For calibrated cameras, a common simplification is to assume that the matrix K is the identity, so that M is a rotation matrix.

3.3. Epipolar geometry overview

Corresponding to a pair of cameras, there exists a 3×3 matrix Q known as the essential matrix [23, 25, 26], such that if $\mathbf{u}$ and $\mathbf{u}'$ are a pair of matched points expressed in homogeneous coordinates, then $\mathbf{u}'^T Q\mathbf{u} = 0$. If a sufficient number of matched points are known, the matrix Q may be computed by the solution of a (possibly overdetermined) set of linear equations.

If the internal calibration of the cameras is known, then it is possible to determine from Q the relative placement of the cameras and hence the relative locations of the 3-D points corresponding to the matched points. It is shown in [25] that this is also true even when the focal lengths of the two cameras are unknown. Unfortunately, for uncalibrated cameras, it is not possible to compute the camera parameters or the point locations unambiguously. However, we prove in Theorem 1 that the various solutions (i.e., the 3-D location of points and the camera transformations matrices) that are compatible with the given set of matched points are related with each other via a 3-dimensional projective transformation H.

Then, we show how one can compute (non-uniquely) two camera transformations $P1$ and $P2$ from Q and use them to find a tentative set of 3-D points locations. Since both $P1$ and $P2$, and the set of points may be off by an unknown projective transformation H, ground control points are used to compute the true 3-D location of the points.

Thus, the 3-D point locations are found by considering both matched points and ground control points using linear methods. Other purely non-iterative methods (e.g. those by Sutherland [27] or Longuet-Higgins [26]) are not able to handle ground-control and matched points simultaneously. In our approach, we avoid the explicit computation of internal or external camera parameters though they may easily be obtained if needed.

4. CONCLUSIONS

During the work, proposed a model of machine vision for applications that used to increase road safety. Proposed a method of using the model of object recognition in the implementation of the system with mobile cameras.

Were the methods of image processing and construction of three-dimensional picture of the observed scene.

It is also proposed a system of functional composition and positioning of elements vision systems.

We propose a method for expanding the working range of the system without additional technical resources, thereby reducing the time of the decision.

References

1. Piecha, J. Digital camera as a data source of ITS solutions in traffic control and management. *Transport Problems*. 2012. Vol. 7. No. 3. P. 57-70.
2. Piecha, J. & Staniek, M. The syntactic alphabet analysis for vehicles simple and fast assignment. *Proc of Int. Conf. on Telematics, Logistics & Transport Safety*. Silesian University of Technology. Katowice. 2009. P. 51-57.
3. Gnyla, P. & Piecha, J. The Transportation Network Rough Description for an Adaptive Traffic Control Process by Means of Video Detection Technology. *IV Int. Conf. "Transport Problems"*. Silesian University of Technology. Katowice. 2012. P. 155-163.
4. Pamuła, W.: Object Classification Methods for Application in FPGA Based Vehicle Video Detector. *Transport Problems*. Vol. 4. No. 2. 2009. P. 5-14.
5. Piecha, J. & Staniek, M. Vehicles Trajectories Movement Description by Means of Syntactic Method. *Transport Problems.* 2009. Vol. 4. No. 4. P. 53-60.
6. Chilian, A & Hirshmüller, H. Stereo Camera Based Camera Naviagation of Mobile Robots on Rough Terrain. *IEEE International Conference of Intelligent Robots and Systems (IROS).* October 2009. St. Louis. MO. USA.
7. Beymer, D. & McLauchlan, P. & Coifman B. & Malik, J.*A* Real-time Computer Vision System for Measuring Traffic Parameters. *Proc. of Int. Conf. Computer Vision and Pattern Recognition Workshops*. 1997. P. 495-501.
8. Bos, I. & Ettema, D. & Molin, E. Modeling effect of travel time uncertainty and traffic information on use of park-and-ride facilities. *Transportation Research Board*. 2004. Vol. 1898. P. 37-44.
9. Burzyński, M. & Kosiński, W. & Schulz, T. & Zając, P. A highway traffic modeling by means of cellular automaton. In: *Transactions on transport systems telematics and safety*. Gliwice: Silesian University of Technology. 2009. P. 11-18.
10. Deng, L. & Tang, N. & Lee, D. & Wang, Ch. & Lu, M. Vision Based Adaptive Traffic Signal Control System Development. *Proc. of Int. Conf. Parallel and Distributed Systems*. Vol. 2. 2005. P. 634-638.
11. Chitturi, M.V. & Medina, J.C. & Benekohal, R. F. Effect of shadows and time of day on performance of video detection systems at signalized intersections. *Transportation Research Part C*. 2010. Vol. 18. P. 176-186.
12. Degang, C. & Zhang, L. & Zhao, S. & Hu, Q & Zhu, P. A Novel Algorithm of Finding Reducts with Fuzzy Rough Set. *Transactions on Fuzzy Systems*. 2011. No. 99. P. 385-389.
13. Gnyla, P. & Piecha, J. The Transportation Network Rough Description for an Adaptive Traffic Control Process by Means of Video Detection Technology. *IV Int. Conf. "Transport Problems"*. Katowice: Silesian University of Technology. 2012. P. 155-163.

14. Haag, M. & Nagel, H.H. Combination of Edge Element and Optical Flow Estimates for 3D-Model-Based Vehicle Tracking in Traffic Image Sequences. *International Journal of Computer Vision.* 1999. Vol. 35. No. 3. P. 295-319.
15. Yao, J.T. & Yao, Y.Y. Induction of classification rules by granular computing. *Proceedings of the Third International Conference on Rough Sets and Current Trends in Computing (TSCTC'02)* London: Springer-Verlag. 2002.
16. Bargiela, A. & Kosonen, I. & Pursula, M. & Peytchev, E. Granular analysis of traffic data for turning movements estimation. *International Journal of Enterprise Information Systems*. 2006. Vol. 2. No 2. P.13-27.
17. Bargiela, A. & Pedrycz, W. The roots of granular computing. *Proceedings of the 2006 IEEE International Conf. on Granular Computing*. 2006. P. 806-809.
18. Płaczek, B. The method of data entering into cellular traffic model for on-line simulation. In: Piecha, J. (ed.) *Transactions on Transport Systems Telematics*. Gliwice. 2006.
19. Płaczek, B. Selective data collection in vehicular networks for traffic control applications. *Transportation Research Part C*. 2012. Vol. 23. P. 14-28.
20. Piecha, J. & Gnyla, P. & Baca, M. Some traffic control proposals by means of fuzzy sets theory. *Proc. of Central European Conf. on Information and Intelligent Systems – CECIIS*. Zagreb. Sept. 2011.
21. Zhao, Y. & Kockelman, K.M. The propagation of uncertainty through travel demand models: an exploratory analysis. *Annals of Regional Science*. 2002. Vol. 36. P. 145-163.
22. Strat, T.M. Recovering the camera parameters from a transformation matrix. *Proc. of DARPA Image Understanding Workshop*. New Orleans. LA. 1984. P. 264-271.
23. Atkinson, K.E. *An Introduction to Numerical Analysis*. John Wiley and Sons. 1989.
24. Tilneac, M. & Dolga, V. & Grigorescu, S. & Bitea, M.A. 3D stereo vision measurements for weed-crop discrimination. *Elektronika ir Elektrotechnika.* 2012. Vol 123. No. 7. P. 9-12.
25. Hartley, R. Estimation of Relative Camera Positions for Uncalibrated Cameras. *Proc. of ECCV-92*. In: Sandini G. (Ed.). LNCS-Series. 1992. Vol. 588. Berlin: Springer-Verlag.
26. Longuet-Higgins, H.C. A computer algorithm for reconstructing a scene from two projections. *Nature*. 1981. Vol. 293. P. 133-135.
27. Sutherland, I.E. Three dimensional data input by tablet. *Proceedings of IEEE*. 1974. Vol. 62. P. 453-461.

Received 05.10.2013; accepted in revised form 23.02.2015