THE SYSTEM OF VIDEO-DATA PROCESSING FOR THE AUTONOMOUS CONTROL OF MOBILE ROBOT

Denis Vershok [1), Rauf Sadykhov [1), Andrei Selikhanovich [1), Klaus Schilling [2), Hubert Roth [2)

1) United Institute of Informatics Problems National Academy of Sciences of Belarus, System Identification Laboratory, Surganova str. 6, 220012, Minsk, Belarus,
E-mail: vershok@newman.bas-net.by, sadyhov@newman.bas-net.by, sel@newman.bas-net.by
2) Steinbeis Transferzentrum ARS, FH Ravensburg-Weingarten, Postfach 1261, D-88241 Weingarten Germany,
Tel: 0049-751-48542, Fax 0049-751-48523,
E-mail:schi@ars-sun1.ars.fh-weingarten.de, roth@fbe.fh-weingarten.de

Abstract: *This paper describes the system of video-data processing based on monocular vision for autonomous control of mobile robot. The system allows detecting obstacles in a robot environment modeled as a set of straight-line segment. The given system consists of three basic stages and uses original algorithms, ensuring the required precision and realization of the system in the real time. The first stage uses a fast edge detection algorithm on the basis of two-dimensional Walsh transform. The algorithm of modified Hough transform is used for detection of straight-line segments. The third stage «segment tracking» uses Kalman filtration for tracking segments in a monocular sequence of images.*

Keywords: *- Robotics, Hough transform, Walsh transform, edge detection, monocular vision*

1. INTRODUCTION

Traditional sensors for mobile robots can be divided into two broad categories: position sensors and visual sensors. Position sensors, which include both on-board range sensors (sonar, infrared, tactile) as well as on-board and off-board localization sensors (laser, GPS), operate in real-time but provide very low dimensional information, such as the distance to a target or the location of the robot within a room. In contrast, visual sensors can provide much richer and high-dimensional information about the environment, but only few vision processing schemes exist that can operate in real-time on a mobile platform in an unstructured environment [1-2].

The fact that visual sensors contain an enormous amount of information does not mean that all algorithms based on vision must use all visual information. It is often possible to extract useful informative features to perform tasks that cannot be done easy with other sensors, or to use vision in a way that duplicate functions of other sensors, but with different characteristics.

The system of video-data processing which enters into the composition of the intellectual system for the mobile robot control is proposed in the paper. The whole control system [7] consists of several neural networks and uses two independent informational streams to receive an information about environment: 1) video camera data, 2) data from ultrasonic sensors and infrared scanner. It is known that each of these types of sensors has both advantages and limitations that specified above. Besides, additional restrictions of sensing systems exist. They are depended from a type and model of a mobile robot. Fig. 1 shows the Walter mobile robot used for our experiments.

Fig. 1 – The Walter mobile robot.

Walter′s ultrasonic sensors and infrared scanner are located in a single plane along the perimeter of the robot that allows receiving an information about distances to obstacles only in this plane.

Also this robot is equipped by a single video camera covering approximately 55 deg. of visual field that forbids to use a stereo vision technique for an evaluation of distances to objects within the robot environment (fig. 2.)

2. VIDEO-DATA PROCESSING SYSTEM

Clearly, that in the case of using only one type of sensors the creation of a satisfactory control system is impossible. Therefore, it is natural to use two independent information flows which complement each other. The data from these streams are entered to an input of a system of path searching for the mobile robot.

In our case the output of the first information flow (ultrasonic sensors and infrared scanner) represents a map of distances from the robot center to obstacles in the single plane. The map is divided into 32 sectors and covers a range of ±90° relatively to the current direction of robot motion.

For processing information from a video camera (second information flow) the system of detection of "dangerous" features from a robot environment is offered. For this purpose a monocular vision is used. Rectilinear segments approximating contours of objects from the robot environment are used as "dangerous" features. The given system consists of three basic stages and uses original algorithms permitting to achieve required accuracy and to realize system in real time.
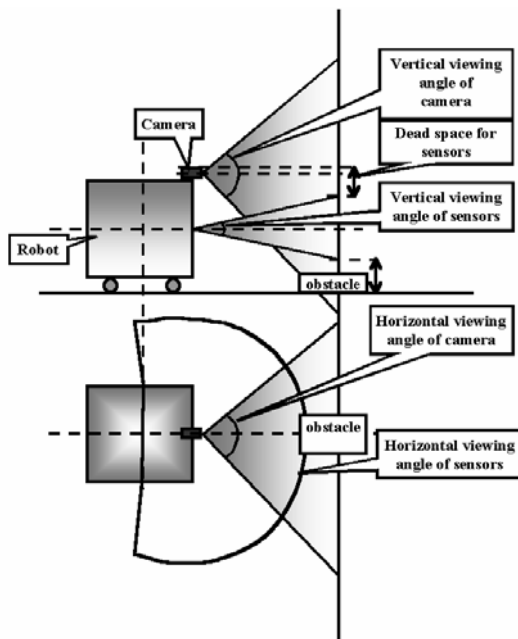

Fig. 2 – Fields of viewing for Walter`s sensors.

2.1. EDGE DETECTION

At the first stage the detection of objects contours in the pictures is carried out. For this purpose the method based on orthogonal Walsh transform is offered. The discrete two-dimensional Walsh functions are used as an edge operator in this method. Four such functions [3] are illustrated in fig. 3.
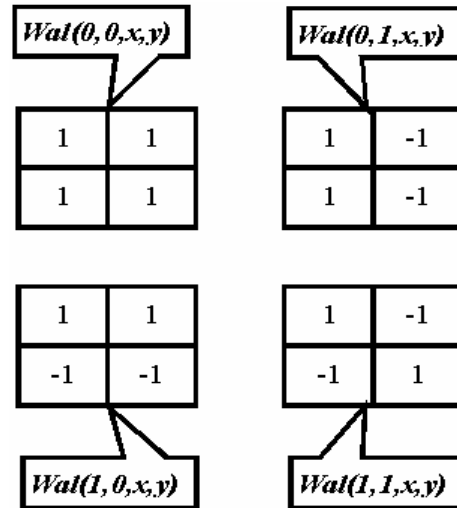

Fig. 3 – Two-dimensional discrete Walsh functions.

This basis can be easily generated if two-dimensional Walsh functions are represented as follows:

$$wal(m,n,x,y) = wal(m,x) \otimes wal(n,y),$$
$$m,n = \overline{0, N-1} \qquad (1)$$

where $wal(m,x)$ and $wal(n,y)$ are row-vector and column-vector of the Walsh functions, $\otimes$ is the Kronecker multiplication.

At the edge detection stage the initial image is scanned by a 2×2 window with a step equal to 1. The direct and inverse Walsh transforms in the two-dimensional basis are calculated in each of these windows:

$$f_{mn} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} s_{ij} \cdot wal_{ij}^{mn}, \quad m,n = \overline{0 \div N-1},$$

$$s_{mn} = \frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} f_{ij} \cdot wal_{ij}^{mn} \qquad (2)$$

where s_{ij} and wal_{ij}^{mn} – elements of matrices of a signal and two-dimensional Walsh function $wal(m,n,x,y)$ accordingly; (m,n) – number of the two-dimensional Walsh function.

The set of Walsh coefficients C_{00}, C_{01}, C_{10}, C_{11} corresponding two-dimensional Walsh

functions $wal(0,0,x,y)$, $wal(0,1,x,y)$, $wal(1,0,x,y)$, $wal(1,1,x,y)$ is obtained as a result of direct transform realization. The first coefficient C_{00} represents a constant constituent of the transform and corresponds to an average intensity of the image fragment. Another coefficients C_{01}, C_{10}, C_{11} correspond to directional derivatives. The first coefficient is eliminated after realization of the direct transform (set to zero) and then inverse Walsh transform is applied to the obtained matrix.

The unique feature of the method is that all pixels from a window can be modified after transform, instead of single pixel. The new values of pixels take part in the next iteration step. This approach allows to get effect of more expressed peaks in the resulting function. The peaks of resulting function correspond to intensity change in original image.

The fast algorithm [3] can be used for computation of Walsh transform. The scheme of algorithm for the basis of two-dimensional Walsh functions (fig. 3) and input matrix $Z = \begin{bmatrix} z_1 & z_2 \\ z_3 & z_4 \end{bmatrix}$ is shown in fig. 4. This algorithm determines the four coefficients of two-dimensional Walsh transform of Z.
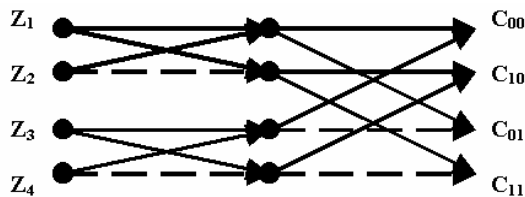

Fig. 4 – Fast computation algorithm in basis of two-dimensional Walsh functions.

2.2 RECTILINEAR SEGMENTS DETECTION

The second stage - detection of rectilinear segments is based on the modified Hough transform (MHT). As a parametric equation the normalized representation of a line is used:

$$R = x \cdot \cos Q + y \cdot \sin Q \qquad (3)$$

where R is a distance from an origin to a line, Q is an angle between normal and OX - axis.

The following modifications have been done in the original Hough transform:

1) A weighted accumulative value is used instead of unit when accumulative array is formed. For selection of the weight of accumulative value the analysis of neighborhoods in a window of 3×3 points is performed. The rules of selection are based on the technique [4]

2) For reducing a number of operations and removing noise the rules for selection of most probable directions of lines passing through a point have been introduced. For this purpose, as well as in the first case, the image analysis in the window of 3×3 is produced [4]. This modification allows for every point of an image to reduce the set: $\mathbf{Q}=\{Q_i\}$, $i=0..2\pi$, ΔQ of possible directions of straight-line to a new set: $\mathbf{Q'}=\{Q_i| P(Q_i)>C\}$, $\mathbf{Q'}\subseteq\mathbf{Q}$ where $P(Q_i)$ is a likelyhood of location of line in Q_i direction, C is a constant

The consequent procedure of detection of peaks allows to mark a straight-lines in the accumulative matrix. Then for marked straight-lines an inverse Hough transform is performed and rectilinear segments located on these lines are detected. For representation of segment we use the following parameters: *(x,y,Q,L)* where *(x,y)* are the coordinates of the midpoint, *Q* is the orientation, and *L* is the length of the segment.

2.3. SEGMENT TRACKING

The third stage allows determining a position of "dangerous" features in 3D space. It is based on transference of stereo vision principles from a spatial domain in the time domain. The last stage of video processing consists of the following steps: at first in the neighborhood images obtained with an interval Δt the appropriate "dangerous" features are detected, i.e. those segments, which describe the similar portions of object from environment. For this purpose we have used method suggested by B.Giai-Checa, P.Bouthemy and T.Vieville [6]. Their algorithm of segment tracking is realized by recursively estimating four segments parameters by means of Kalman filters [5]. The prediction stage allows us to define a search area in which lie the segments candidates to be matched. We choose from among them the segment that minimizes the Mahalanobis distance for the four parameters of the segments. Then the displacements of these matched features on an image at a moment $t + \Delta t$ relative to its position in an image at a moment t are calculated. Then on the basis of the known values (displacements of "dangerous" features and distance which the robot has passed in a moment Δt) the distances from the camera to these features in 3D-space are calculated.

These distances and data from the first information flow are used by the control system on the basis of neural nets for searching the free way [7].

3. CONCLUSION

Proposed video-processing scheme allows to increase the robustness of the system for mobile robot control. System based on principles of monocular vision uses original algorithms on stages of edge detection and straight-line extraction.

The selection of Walsh based method on the stage of contour detection is stipulated by the following factors:

1) low computational complexity (note that the algorithm can be effectively implemented using hardware);

2) in comparison with the Roberts operator the offered operator is less sensitive to noise and takes into account the slanting component of a gradient;

3) as compared with the Sobel operator, suggested method allows to obtain more thin contour lines of objects and operates better with textured pictures that is necessary for successful execution of the following stage.

On the stage of straight-line detection the algorithm based on modified Hough transform is used. The selection of above method is validated by such advantages:

1) robust stability of the method relative to noise;

2) the method of MHT has a high characteristics of performance;

3) the result of transform has a suitable form for the next stage – searching correspondence between "dangerous" features in neighborhood pictures which snapped with interval Δt.

4. REFERENCES

[1] J. Borenstain, L. Feng. *Navigating Mobile Robots. A.K. Peters, Wellesley, MA, 1996.*

[2] I. Horswill. *Collision Avoidance by Segmentation. In Proceedings of the 1994 Int. Conf. On Intelligent Robots and Systems (IROS-94), IEEE-press.*

[3] A. Machnev, A. Selikhanovich. *Walsh operators of edge detection for gray-scale images. In Proceedings of Int. Conf. On Neural Networks and Artificial Intelligence, 12-15 October 1999, Brest, Belarus, pp. 178-182.*

[4] R.H. Sadyhov, D.A. Vershok. *Algorithm of invariant features extraction based on modified Hough transform. In Proceedings of Fourth International Conference Pattern Recognition and Information Processing (PRIP'97), Minsk-Szczecin, Belarus-Poland, 20-22 May, 1997.*

[5] P. Ruymgaart, T. Soong. *Mathematics of Kalman-Bucy filtering. Spring Verlag, Berlin, 1985.*

[6] B. Giai-Checa, P. Bouthemy, T. Vieville. *Detecting of moving objects. Research report, INRIA Sophia Antipolis, 1996.*

[7] V. Golovko, A. Klimovich, K. Schilling, J. Pico. *Neural Networks for Reactive Control of Mobile Robot. In Proceedings of Int. Workshop on Intelligent Signal Processing, 4-7 September 1999, Budapest, Hungary, pp. 310-311.*

Rauf Kh. Sadykhov, in 1967 graduated from Azerbaijan Polytechnic Institute (Baku) on the specialty "Mathematical and Computing the Instruments and Devices". After graduation from the Institute he attended the postgraduate course at the Institute of Engineering Cybernetics in Minsk. In 1991 he defended his thesis for a scientific degree of a doctor of engineering science in the field of computing science and in 1992 has obtained a professor,s scientific rank.

Since 1995 R.Kh. Sadykhov is a head of Computer System Department in Belarusian State University of Informatics and Radioelectronics and simultaneously he is a head of System Identification laboratory, Institute of Engineering cybernetics of the Belarusian Academy of Sciences. R.Kh. Sadykhov has published more than 350 scientific works, including books, patents, papers and reports at the International Conferences, Symposiums and Workshop. The area of the scientific investigations includes: digital signal and image processing, recognition of handwritten symbol and signature identification, remote-sensing object recognition, computer vision system for the control and recognition, intellectual neural systems, multi-agent systems, parallel architectures for digital signal and image processing.

Prof. R.Sadykhov is the vice-chairman of Belarusian Association of Pattern Recognition (IAPR) and Belarus SIG of International Neural Network Society (INNS), the member of IEE (United Kingdom).