Biuletyn WAT
Vol. LVII, Nr 4, 2008

Recognition of unknown scale objects

WITOLD ŻORSKI

Military University of Technology, Cybernetics Faculty,
Institute of Teleinformatics and Automation,
2 Kaliskiego Str. , 00-908 Warsaw

Abstract. This paper presents an application of the Hough Transform to the task of identifying objects of unknown scale, e.g. within a scene in a robot vision system. The presented method is based on the Hough Transform for irregular objects, with a parameter space defined by translation, rotation and scaling operations. The high efficiency of the technique allows for poor quality or highly complex images to be analysed. The technique may be used in robot vision systems, identification systems or for image analysis, directly on grey-level images.
Keywords: Hough Transforms, robot vision systems, machine vision, pattern matching
Universal Decimal Classification: 681.3.01

1. Introduction to the Hough Transform

The Hough Transform was patented in 1962 as a method for detecting complex patterns of points in a binary image [5]. In 1981, Deans noticed [3] that the Hough Transform for straight lines was a specific case of the more general Radon Transform [9] known since 1917, which is defined as (for the function $I(x, y)$ in two-dimensional Euclidean space):

$$H(\rho,\alpha) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{I}(x,y)\,\delta(\rho - x\cos(\alpha) - y\sin(\alpha))\,dxdy, \tag{1}$$

where δ is the delta function. This result shows that the function $I(x, y)$ is integrated along the straight line determined by the parametric equation $\rho = x\cos(\alpha) + y\sin(\alpha)$. The Radon Transform is equivalent to the Hough Transform when considering binary images (i.e. when the function $I(x, y)$ takes values 0 or 1). The Radon Transform for shapes other than straight lines can be obtained by replacing the delta function argument by a function, which forces integration of the image along contours appropriate to the shape.

Using the Radon Transform to calculate the Hough Transform is simple (almost intuitive) and is often applied in computer implementations. We call this operation **pixel counting** in the binary image.

An (alternative) interpretation of the Hough Transform is the so-called **backprojection** method. The detection of analytical curves defined in a parametrical way, other than straight lines is quite obvious. A set of points lying on a curve determined by n parameters $a_1,\ldots,a_n$ may be presented in the form:

$$\lambda_O = \{(x, y) \in R^2 :\ g((\hat{a}_1,\ldots,\hat{a}_n),(x, y)) = 0\}, \tag{2}$$

where $g((\hat{a}_1,\ldots,\hat{a}_n),(x, y)) = 0$ is the formula that describes the considered curve [6].

By exchanging the meaning of parameters and variables in the above equation we obtain the backprojection relation (mapping image points into parameter space), which may be written down in the following way:

$$\lambda_T = \{(a_1,\ldots,a_n) \in R^n :\ g((\hat{x}, \hat{y}),(a_1,\ldots,a_n)) = 0\}. \tag{3}$$

Based on Eq. (3), the Hough Transform $H(a_1,\ldots,a_n)$ for the image $I(x, y)$ is defined as follows:

$$H(a_1,\ldots a_n) = \sum_{(x_i, y_i)\in I} h(\hat{x}_i, \hat{y}_i, a_1,\ldots,a_n), \tag{4}$$

where

$$h(\hat{x}_i, \hat{y}_i, a_1,\ldots,a_n) = \begin{cases} 1 & \text{if } g((\hat{x}_i, \hat{y}_i),(a_1,\ldots,a_n)) = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

In order to calculate the Hough Transform digitally, an appropriate representation of the parameter space $H(a_1,\ldots,a_n)$ is required. In a standard implementation, any dimension in the parameter space is a subject to quantisation and narrowing to an appropriate range. As a result, an array is obtained where any element is identified by the parameters $(a_1,\ldots,a_n)$. An element in the array is increased by 1 when the analytical curve, determined by co-ordinates $(a_1,\ldots,a_n)$, passes through the point $(\hat{x}, \hat{y})$ of the object in the image I. This process is called **accumulation** and the array used is called an **accumulator** (usually marked with the symbol A).

Thus, we may assume that the Hough Transform is based on a representation of the image I into the accumulator array A, which is defined as follows:

$$A : P \to N, \quad \text{where} \quad P = P_1 \times P_2 \times \cdots \times P_p. \tag{6}$$

The symbol $P_i \subset N$ determines a range of i-parameter of the p-dimensional space P. Determining the array A is conducted through the calculation of partial values for points of an object in the image I and adding them to the previous ones [see Eq. (4)] which constitutes a process of accumulation. Initially, all elements of the array A are set to zero.

This paper presents an application of the Hough Transform to the tasks of identifying manipulated objects in a robot vision system with unknown scale of the scene. It is based on the Hough Transform with a parameter space defined by translation, rotation and scaling operations. A fundamental element of this method is a generalisation of the Hough Transform for grey-level images including a solution to the scaling problem. The author tried to test every introduced theoretical element in a realistic situation. With this end in view, an application of the elaborated method dedicated to the robot vision system (see Fig. 1) has been built. It allows users to carry out a variety of experiments in the area of computer vision. The results obtained confirm the efficiency of the method even in the case of manipulated objects which are joined or covered by each other.

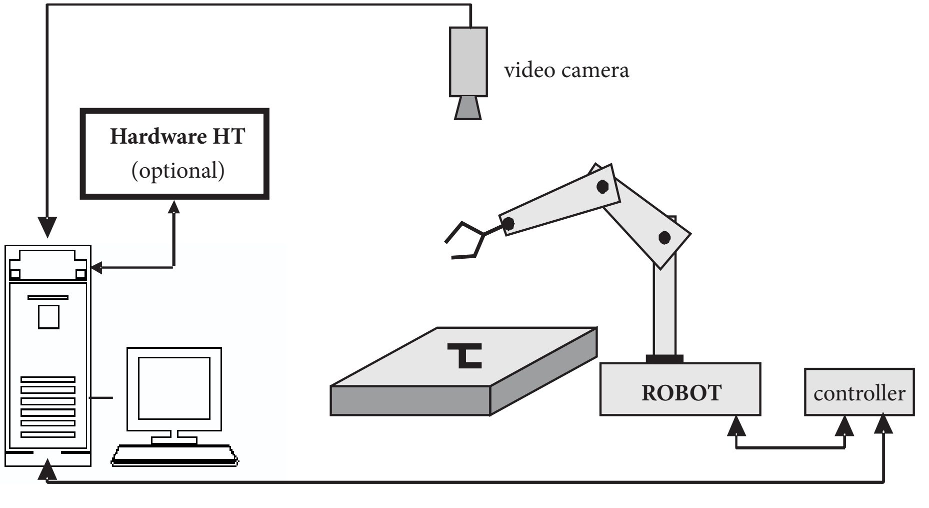

Fig. 1. The considered robot vision system (the scene's image scale is unknown)

2. The Hough Transform for irregular objects

The Hough Transform may be successfully applied to detect irregular objects [1, 11]. In the generalised Hough Transform, an object is represented by a pattern which is a list of the boundary points $\{(x_i, y_i) : i = 1, \ldots, n\}$ (without a concrete analytical description), and the parameter space is defined for the translation $[x_T, y_T]$, the rotation α and (alternatively) the scale s of the pattern in the image.

Each point of the image generates an appropriate hypersurface as a result of backprojection into a parameter space. A number of hypersurfaces that criss-cross a given point (x_T, y_T, α, s) of the parameter space is equivalent to a number of points common for a given object in the image and the fitting pattern.

The Hough Transform for binary images has been already described by the author in details [10].

3. Generalisation of the Hough Transform for grey-level images

Let us first define the concept of a grey-level image, an object appearing in such an image and the concept of a grey-level pattern in a computer vision system.

Definitions

An **image with 256 grey levels** means a set of points, which have a value or "shade" from the set {0,…,255}. Such an image may be presented as follows

$$I_G : D \to \{0, \ldots, 255\}, \quad \text{where} \quad D = [1, \ldots, W] \times [1, \ldots, K] \subset N^2. \tag{7}$$

The **object $b(I_G)$** in the image I_G is any fragment of that image which may be recorded in terms of

$$Q_G : D_Q \to \{0, \ldots, 255\}, \quad \text{where} \quad D_Q \subset D = [1, \ldots, W] \times [1, \ldots, K] \subset N^2. \tag{8}$$

<u>Remark</u>: Identifying an object with an image fragment is a consequence of a set of values taken by the function I_G.

The **pattern** M_p means an image (square matrix) of the size $N_p \times N_p$ which is as

$$M_P : D_P \to \{0, \ldots, 255\}, \quad \text{where} \quad D_P = [1, \ldots, N_P] \times [1, \ldots, N_P] \subset N^2. \tag{9}$$

An example of a grey-level pattern is shown in Fig. 2.

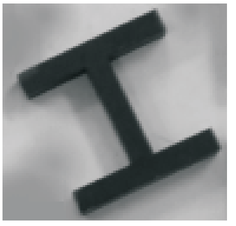

Fig. 2. An example of grey-level pattern M_p

In a computerised robot monitoring system, the identification process of manipulated objects is carried out with the use of previously learned patterns. The task is aimed at identifying (i.e. determining the location and rotation angle) a given object in the image. We assume that the given object is represented by the pattern M_p. The task to identify the pattern M_p in the image I_G may be regarded as determining parameters which uniquely describe its location and orientation in the given image. Some obtained results are shown in Figs. 3, 4 and 5.

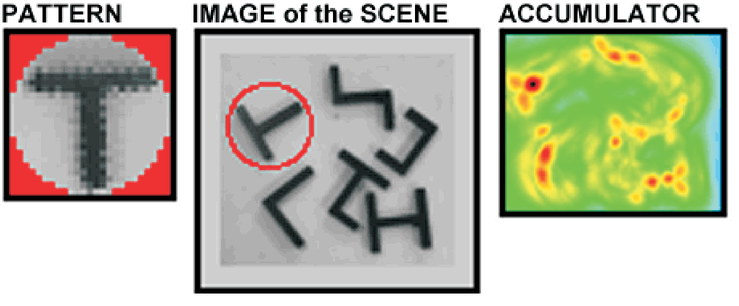

Fig. 3. Pattern recognition for a scene with separated objects

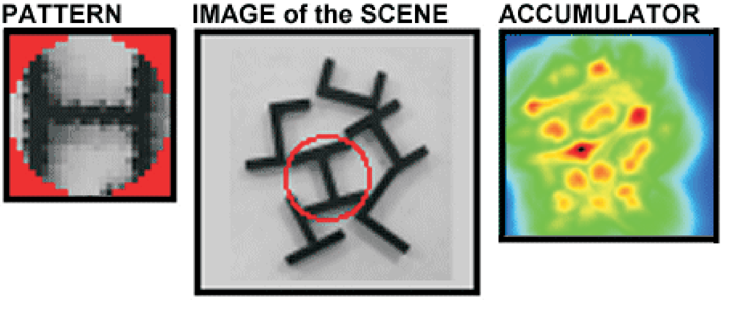

Fig. 4. Pattern recognition for a scene with joined objects

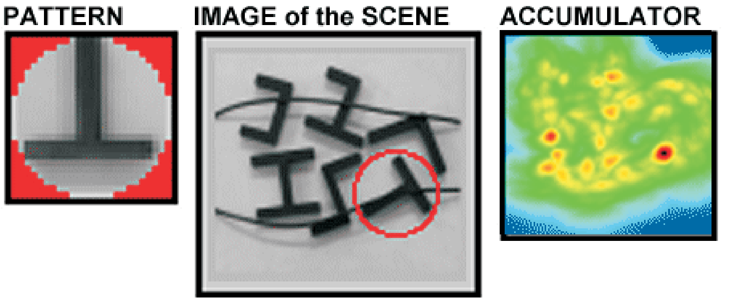

Fig. 5. Pattern recognition for a scene with partly covered objects

The Hough Transform $H(x_T, y_T, \alpha, s)$, which takes into account translation, rotation and scaling, for the image $I_G(x, y)$ [see Eq. (7)] in the process of identifying the pattern M_p determined by Eq. (9) may be defined as [compare Eqs. (4) and (5)]

$$H(x_T, y_T, \alpha, s) = \sum_{(x_i, y_i) \in M_P} h(x_i, y_i, x_T, y_T, \alpha, s), \tag{10}$$

where

$$h(x_i, y_i, x_T, y_T, \alpha, s) = 255 - \left| I_G(x_i'', y_i'') - M_P(x_i, y_i) \right| \tag{11}$$

and the values x_i'', y_i'' are calculated from

$$\begin{cases} x_i'' = x_r + s(x_i - x_r)\cos(\alpha) - s(y_i - y_r)\sin(\alpha) + x_T \\ y_i'' = y_r + s(x_i - x_r)\sin(\alpha) + s(y_i - y_r)\cos(\alpha) + y_T \end{cases}. \tag{12}$$

The above equations relate to the situation given in Fig. 6.

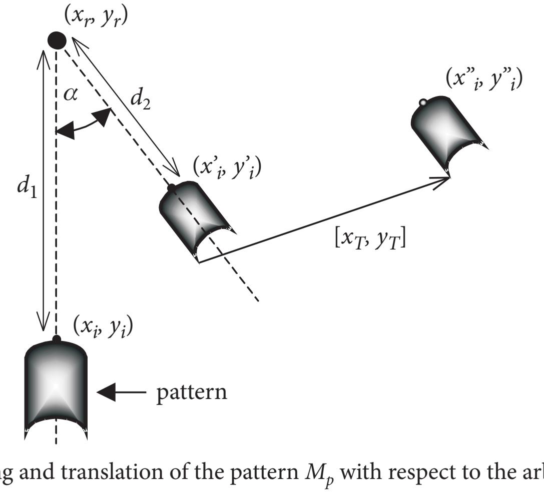

Fig. 6. Rotation, scaling and translation of the pattern M_p with respect to the arbitrary point (x_r, y_r)

4. The Hough Transform and the scaling issue

Taking into consideration pattern scaling adds an extra dimension to the parameter space. However, because the scale range is commonly known and it is often not too large, only a few values of the scale factor *s* are enough to achieve the process of identification.

If we assume that the following *n* values of the scale *s* factor must be taken into consideration

$$\xi_1,\ldots,\xi_n, \qquad n \in N, \tag{13}$$

then, the parameter space may be determined in the following way:

$$P = P_1 \times P_2 \times P_3 \times P_4 = [1,\ldots,W] \times [1,\ldots,K] \times [0,\ldots,L-1] \times [\xi_1,\ldots,\xi_n], \\ (\Delta\alpha = \frac{2\pi}{L}). \tag{14}$$

In order to accelerate calculations (applying the histogram study), the set of patterns

$$\{M_P^1,\ldots,M_P^n\}, \tag{15}$$

must be generated first by scaling the given pattern M_p within a range determined by the values $\xi_1,\ldots,\xi_n$. The pattern localisation process for any pattern formed from the set of Eq. (15) can then be applied. Such an approach can drastically reduce the number of calculations required.

However, this method has one disadvantage that results from having to calculate histograms for the initial image *n* times (as the size of each pattern is different). A solution is to create a new set of the patterns $\{\bar{M}_P^1,\ldots,\bar{M}_P^n\}$ of the same size but without losing information connected with the scale of the patterns $M_P^1,\ldots,M_P^n$. Note that the size of the pattern M_P^1 is $N_P^1 \times N_P^1$. The appropriate patterns $\bar{M}_P^1,\ldots,\bar{M}_P^n$ from the patterns $M_P^1,\ldots,M_P^n$ can be obtained by separating their central part of the size $N_P^1 \times N_P^1$. As a result, we have the patterns that "remember" the scale they were created at, but are of the same size. A graphical illustration of this process is shown in Fig. 7.

Since the received patterns are of the same size, it is sufficient to calculate their histograms once and compare them with the calculated histogram of the initial image. As the size of the patterns decreases, the time to create the histograms is reduced. Decreasing the patterns size also results in shortening the CPU time for the accumulator calculation. Unfortunately, the patterns $\bar{M}_P^1,\ldots,\bar{M}_P^n$ carry less information than the patterns $M_P^1,\ldots,M_P^n$.

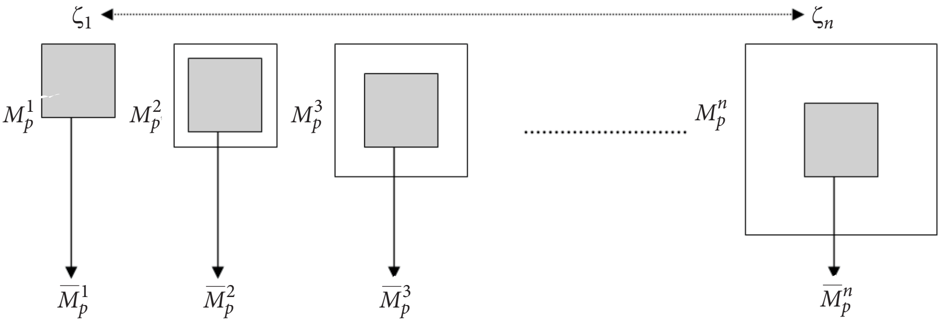

Fig. 7. Graphical illustration of the process of the patterns $\bar{M}_p^1,\ldots,\bar{M}_p^n$ creation

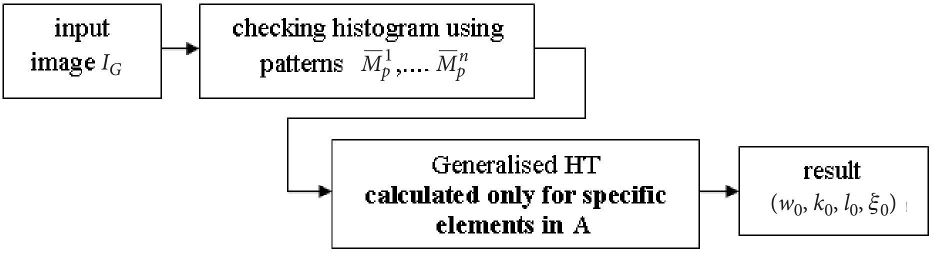

Fig. 8. The way of conduct while identifying objects in images of unknown scale

Test results based on this method are presented in Figs. 9 and 10. Due to the difficulty of illustrating a four-dimensional accumulator array it has been split (for every scale s) and reduced to the flat image by taking cross-section for the maximal

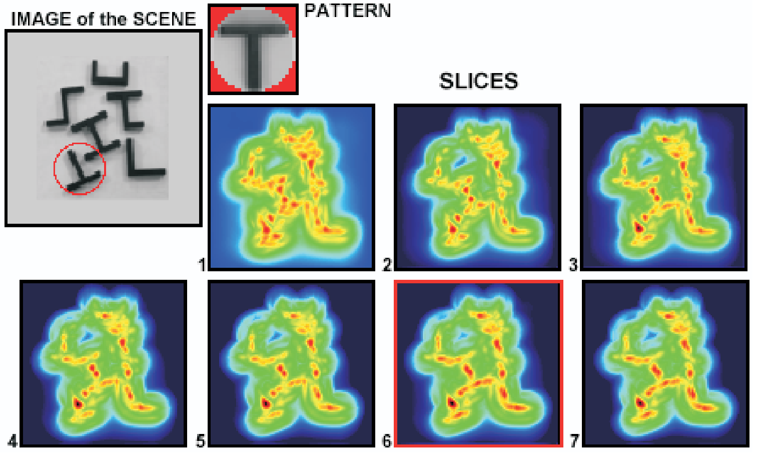

Fig. 9. Result for a scene increased about 13% compared to the pattern's resolution

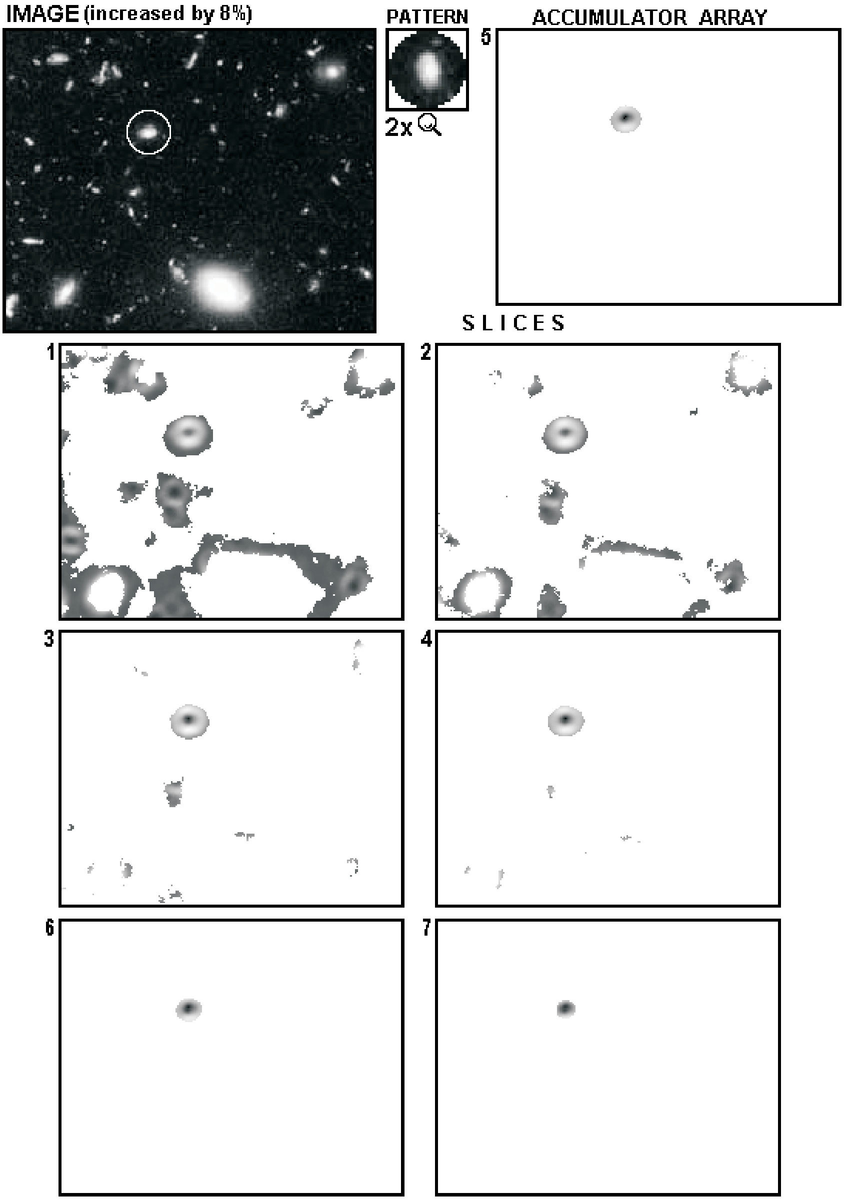

Fig. 10. Location of a stellar object in an image (increased by 8%) obtained from HST

values of the rotation angle α. Presented figures show slices generated by the scale factor s. The main slice, i.e. containing the global maximum, is distinguished. The numbers indicate patterns created with different scale parameters (see Fig. 8). The scale values change in size within the range (–25%, +25%).

Figure 9 shows an identification result for a scene with manipulated objects. The scale of the scene is about 13% increased compared to the pattern. The global maximum has been found in the slice number 6.

Figure 10 shows a result of an object location in an astronomy image (increased by 8%). The global maximum has been found in the fifth slice, i.e. in accordance with the scale change. The behaviour of the accumulator array is very typical in such situations. The larger the scale of the pattern the more unclear contents of the adequate slice appear. The considered figure shows the advantage of using a histogram analysis technique described by the author in [13] and [14].

5. Conclusions

The results obtained confirm the efficiency of the method in the case of detecting the location and identification of manipulated objects. The method may be used in the case of image analysis without skeletonisation and binarisation. This method may be successfully applied to the identification of objects in satellite, aerial or astronomy images too (see Fig. 10). The high efficiency of the technique allows for poor quality or highly complex images to be analysed. The Hough Transform may be supported by its hardware implementation which is suggested in Fig. 1 and described by the author in details [11, 12].

Received July 9 2008, revised September 2008.

REFERENCES

[1] D. H. Ballard, *Generalizing the Hough Transform to Detect Arbitrary Shapes*, Readings in Computer Vision: Issues, Problems, Principles, and Paradigms. Los Altos, CA., 1987, 714-725.

[2] E. R. Davies, *Machine Vision: Theory, Algorithms, Practicalities*, Academic press Ltd, 24/28 Oval Road, London NW1 7DX, United Kingdom, 1990.

[3] S. R. Deans, *Hough transform from the Radon transform*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 3, no. 2, 1981, 185-188.

[4] K. S. Fu, R. C. Gonzalez, C. S. G. Lee, *ROBOTICS: Control, Sensing, Vision, and Intelligence*, McGraw-Hill, New York, 1987.

[5] P. V. C. Hough, *Method and means for recognizing complex patterns*, U.S. Patent 3,069,654, Dec. 18, 1962.

[6] J. Illingworth, J. Kittler, *A survey of the Hough Transform*, Computer Vision, Graphics and Image Processing 44, 1988, 87-116.

[7] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, New Jersey, 1989.

[8] V. F. Leavers, *Shape Detection in Computer Vision Using the Hough Transform*, Springer, London, 1992.

[9] J. Radon, *Uber die Bestimmung von Funktionen durch ihre Integralwerte langs gewisser Mannigfaltigkeiten*, Berichte Sachsische Akademie der Wissenschaften Leipzig, Math Phys Kl., 69, 1917, 262-267.

[10] W. Żorski, B. Foxon, J. Blackledge, M. Turner, *Irregular Pattern Recognition Using the Hough Transform*, Machine Graphics & Vision, 9, 2000, 609-632.

[11] W. Żorski, W. Rabczuk, *Multi-structural Hardware Implementation of the Hough Technique for Grey-level Pattern Recognition*, Proceedings of the 10th IEEE International Conference MMAR 2004, vol. 2, 947-952.

[12] W. Żorski, A. Żak, M. Turner, *Hardware Implementation of the Hough Technique for Irregular Pattern Recognition*, Proceedings of the 8th IEEE International Conference MMAR 2002, 2-5 September 2002, vol. 1, 561-566.

[13] W. Żorski, *Application of the Hough Technique for Irregular Pattern Recognition to a Robot Monitoring System*, Proceedings of the 11th IEEE International Conference MMAR 2005, 725-730.

[14] W. Żorski, *The Hough Transform Application Including Its Hardware Implementation*, Advanced Concepts for Intelligent Vision Systems: Proceedings of the 7th International Conference, Lecture Notes in Computer Science, Springer-Verlag, vol. 3708, 2005, 460-467, http://www.springerlink.com/content/50yk3q0fw71x1qld

[15] W. Żorski, *Fast Hough Transform Based on 3D Image Space Division*, Advanced Concepts for Intelligent Vision Systems: Proceedings of the 8th International Conference, Lecture Notes in Computer Science, Springer-Verlag, vol. 4179, 2006, 1071-1079, http://www.springerlink.com/content/6216256332x1p166

W. ŻORSKI

Rozpoznawanie obiektów o nieznanej skali

Streszczenie. Artykuł prezentuje zastosowanie transformaty Hougha do zadań identyfikacji obiektów o nieznanej skali, np. na scenie w systemie widzenia robota. Metoda wykorzystuje transformatę Hougha dla wzorców nieregularnych, z przestrzenią parametrów określoną przez operacje: translacji, obrotu i skalowania. Wysoka sprawność techniki pozwala na analizę obrazów o słabej jakości lub dużej złożoności. Technika może być przykładowo zastosowana w systemach widzenia robotów, do bezpośredniej analizy obrazów pozyskanych w poziomach szarości.

Słowa kluczowe: Transformata Hougha, systemy widzenia robotów, widzenie maszynowe, dobieranie wzorców

Symbole UKD: 681.3.01