

# The Method of Guaranteeing the Separation Between the Recognised Object and Background

Zygmunt Kuś and Aleksander Nawrat

**Abstract** The aim of the following study was to develop a procedure which guarantees the separation between the part of an image where we have the recognised object and the part of the image which corresponds to the terrain where the object moves. This research is conducted for grey scale images. The authors have presented the method which uses moment invariants for creating feature vectors which define the features of the recognised object and the features of the background. The presented method is based on calculating the distance between the values of invariant functions calculated for an object and the background. The distances were calculated for all moment invariants. These moment invariants were elements of the feature vector. In the next step the elements of the feature vector were ordered according to the values of these distances—from lowest to highest. Finally, the moment invariants, for which the distances were highest, were chosen as elements of a new—shorter feature vector. Furthermore, the algorithm of creating features vector was presented in the following paper. The developed algorithm allows to assess if a given invariant function is useful for the classification of the elements of a given set of classes. Owing to this approach, it was possible to choose properly the invariant functions which constitute the features vector. On the one hand, we can decrease the size of the features vector by choosing the invariant functions which separate particular classes in the best way. On the other hand, we know which function is the most proper to be added to the features vector when the size of the features vector is too small. On top of that, this study presents the example of recognising the object moving in some kind of a terrain.

**Keywords** Object recognition • Moment invariants • Pattern vector

---

Z. Kuś (✉) · A. Nawrat  
Institute of Automatic Control, Silesian University of Technology,  
Akademicka 16 Street, Gliwice, Poland  
e-mail: zygmunt.kus@polsl.pl

A. Nawrat  
e-mail: aleksander.nawrat@polsl.pl

## 1 Introduction

Human visual perception provides a wealth of information allowing to function in a real environment. Therefore a computer's role is to provide the visual perception (electronic acquisition and understanding) so that devices are able to function in the same environment [1, 2].

However, computer visual perception is limited by the limited amount of information which can be processed by a computer. Thus, image analysis algorithms narrow down to taking into consideration only a local neighbourhood of the central point of the camera field of view (image of the small part of the environment). By the central point of the camera field of view we mean the point of the environment which is visible in the image centre. It is very difficult to interpret an image when we do not consider environment from outside of the image; worse, we only consider a part of the image which is in the aperture. In the literature of the subject [3–9] one can find a great number of solutions of a pattern recognition problem which are based on image processing methods. One of these solutions are methods based on region moment representations [10, 11].

## 2 Formulating the Problem of Distinguishing an Object and Background

The basic problem of this study is to recognise an object which is moving in a terrain in the case when we assume the involvement of an operator in the first stage of tracking. Therefore in the preliminary stage of tracking after the tracked object appears in the camera field of view [12–14], the operator marks on a screen the area of the image which constitutes the tracked object. Such an approach allows to conduct the examination of the tracked object characteristics (setting the vector of features); moreover, it allows to treat an unmarked area as the background.

Undoubtedly, in a general case, the background—terrain where the object is moving might be inhomogeneous, therefore the analysis of the image could be hindered if we did not have any rationale—indication how to choose the aperture size. Thanks to the operator showing in a starting point the size and location of the object, it is possible to select the aperture size so that it covers the size of the object [15–19]. For such an aperture we will also find the characteristics of the background—terrain.

During object tracking and the appearance of a different background in the camera field of view, we can correct the background features vector on the basis of the marked object image and the new image of the terrain. One of the fundamental assumptions is the support of the image processing system by a rangefinder which allows to measure the distance between the camera and the tracked object in the starting point. Thanks to the knowledge of the focal length of the camera we can compute the size of an object [20–22].

In this way, during the processing of the images which were acquired at different distances between the camera and the observed part of the terrain, we can

accordingly resize the aperture. It must be conducted in such a way that the area in the image covered by the aperture should be approximate to the area covered by an object (if the object was in the observed part of the terrain).

The examples of the objects, which we are going to recognise, are presented in Fig. 1.

The examples of the terrain images where the tracked objects can move are presented in Fig. 2.

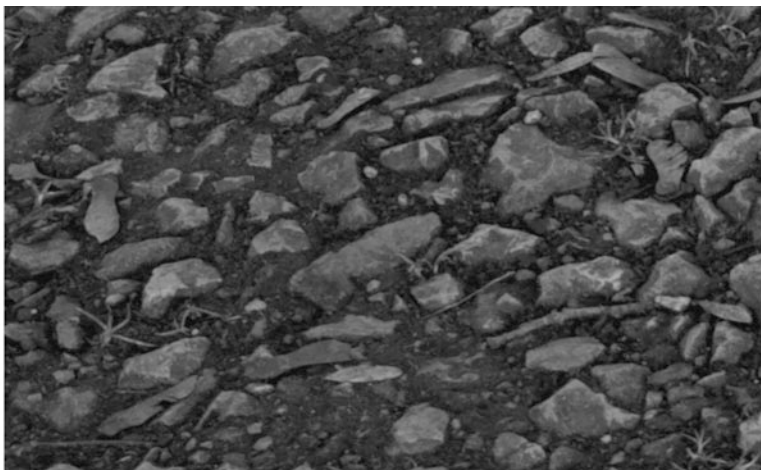


Object 1



Object 2

**Fig. 1** The examples of the objects

**T I****T II****T III**

**Fig. 2** The examples of the terrain images

### 3 Invariant Functions Used for a Description of an Object and Background Features

Examining the usefulness for recognising the objects from a given test set required that we chose invariant functions (1) (moment invariants) as it was presented in literature:

$$\begin{aligned} J_1 &= \frac{I_1}{m_{00}^2}; \quad J_2 = \frac{I_2}{m_{00}^4}; \quad J_3 = \frac{I_3}{m_{00}^5}; \quad J_4 = \frac{I_4}{m_{00}^5}; \\ J_5 &= \frac{I_5}{m_{00}^{10}}; \quad J_6 = \frac{I_6}{m_{00}^7}; \quad J_7 = \frac{I_7}{m_{00}^4}; \quad J_8 = \frac{I_8}{m_{00}^5}; \end{aligned} \quad (1)$$

where

$$\begin{aligned} I_1 &= M_{20} + M_{02} \\ I_2 &= (M_{20} - M_{02})^2 + 4M_{11}^2 \\ I_3 &= (M_{30} - 3M_{12})^2 + (3M_{21} - M_{03})^2 \\ I_4 &= (M_{30} + M_{12})^2 + (M_{21} + M_{03})^2 \\ I_5 &= (M_{30} - 3M_{12})(M_{30} + M_{12})((M_{30} + M_{12})^2 - 3(M_{21} + M_{03})^2) \\ &\quad + (3M_{21} - M_{03})(M_{21} + M_{03})(3(M_{30} + M_{12})^2 - (M_{21} + M_{03})^2) \\ I_6 &= (M_{20} - M_{02})((M_{30} + M_{12})^2 - (M_{21} + M_{03})^2) \\ &\quad + 4M_{11}(M_{30} + M_{12})(M_{21} + M_{03}) \\ I_7 &= M_{20}M_{02} - M_{11}^2 \\ I_8 &= M_{30}M_{12} + M_{21}M_{03} - M_{12}^2 - M_{21}^2 \end{aligned} \quad (2)$$

Central moments of order  $(p + q)$ :

$$M_{pq} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (i - x_s)^p (j - y_s)^q f(i, j)$$

The co-ordinates of the region's centre of gravity (centroid):

$$x_s = \frac{m_{10}}{m_{00}}; \quad y_s = \frac{m_{01}}{m_{00}} \quad (3)$$

Moments of order  $(p + q)$ :

$$m_{pq} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} i^p j^q f(i, j)$$

where  $f(i, j)$ —image pixel;  $M$ —the height of the image,  $N$ —the width of the image.

The abovementioned equations, usually formulated for black and white images, were herein modified in such a way that their values can be calculated for the grey scale images.

A great number of the abovementioned definitions of the IF were presented in [10]. In [11] the use of these IF in image recognition was discussed.

The abovementioned IFs should be invariant regarding rotation, resizing and the shift. It concerns transformations which are made in the image plane on the parts of the image which are the very object. In the case of 3D objects, which change their location in relation to the camera, we obtain different images for the same object [23–25]. In this case, it is necessary to treat each view of the object as a separate subclass and to create the class of the object consisting of the subclasses corresponding to different views of a given object. This paper, for the sake of clarity, will narrow down to one view for each object in the presented example.

#### **4 The Method of Creating Features Vector Which Guarantees the Separation of an Object and Background**

The aim of the following study is to develop a method which will enable to select the invariant function  $F_k$  (calculated for the part of the picture taken from the aperture). This method should guarantee the best distinction between two situations: (a) the object occurs in the aperture and (b) the background occurs in the aperture. It is one of the problems which appear in the pattern recognition task which is realised by the image processing system. We assume that the images are grey scale images. Furthermore, we assume that the size of the aperture is adjusted to the size of the object (pattern) which we are looking for. It also needs to be assumed that the picture of the terrain where the object is located and various pictures of the object are available. We will examine the invariant functions presented in Sect. 3.

This paper will aim at achieving the results which are adjusted to the particular object and particular background; therefore the research will be conducted for an exemplary terrain image and exemplary object pattern which we are looking for.  $F_k(\text{aperture})$  is the value of the  $k$ th invariant function calculated for the part of the picture taken from the aperture. For each  $k$ th invariant function we define the difference of the function calculated for the background and object as:

$$\Phi_k(\text{background, object}) = F_k(\text{background}) - F_k(\text{object}) \quad (4)$$

On the basis of this difference calculated for different exemplary object and background images, we are going to assess the usefulness of the  $k$ th invariant function in order to look for a given object which is located in the given background.

### The algorithm

- A. We take the photos of an object and select the set of images which will define a desirable pattern vector  $\bar{P}_o = \{P_o(i)\}$  where  $P_o(i)$  is the  $i$ th image of the object.
- B. We take the photos of a terrain corresponding to different kinds of the backgrounds in which the recognised object can occur. Next we select the set of images which will define a backgrounds vector  $\bar{P}_T = \{P_T(j)\}$  where  $P_T(j)$  is the  $j$ th image of the background.
- C. We create a set of all possible pairs  $P_{oT} = \{P_o(i), P_T(j)\}$ .
- D. We calculate  $\Phi_{kl}(i, j) = |F_k(P_T(j)) - F_k(P_o(i))|$  (in our case:  $k = 1, \dots, 8$  and  $l = 1, \dots, 6$ ) and create the set  $\Phi_k = \{\Phi_{kl}(i, j)\}$  for each  $i$  and  $j$ .
- E. For each invariant function (a vector of the invariant functions)  $F_k$  we order the elements of the set  $\Phi_k$  from the lowest to the highest. We select half of the elements of the set—elements placed on the left of the median—lower than the median. Next we calculate the mean value  $\Phi_{kmean}$  of the selected elements.
- F. We compare the values  $\Phi_{kmean}$  and select this function  $F_k$ , for which  $\Phi_{kmean}$  is the highest. In order to make the comparison of different invariant functions possible.

## 5 Examples

We examine the images of the objects presented in Fig. 1 as a set of the images of the objects, whereas the set of the terrain images was shown in Fig. 2. In this way we obtain 6 pairs ‘object type’—‘terrain type’ as presented in (5).

$$\begin{array}{llll} 1. & \text{'ob1'—TI,} & 3. & \text{'ob1'—TII,} & 5. & \text{'ob1'—TIII,} \\ 2. & \text{'ob2'—TI,} & 4. & \text{'ob2'—TII} & 6. & \text{'ob2'—TIII} \end{array} \quad (5)$$

For each object and terrain type, we calculate the invariant functions  $J_1$  to  $J_8$ . These values are shown in Table 1.

Table 2 illustrates the values of the differences modules of the IFs for all possible pairs ‘object type’—‘terrain type’ for the subsequent invariant functions  $J_1$  to  $J_8$ . Owing to this fact, we will be able to assess how much a given function  $J$  distinguishes an object from the background.

**Table 1** The values of the invariant functions for the tested objects and terrain types

	$J_1$	$J_2$	$J_3$	$J_4$	$yJ_5$	$J_6$	$J_7$	$J_8$
Object 1	0.3150	0.0380	0.0002	0.0001	0.0000	0.0000	0.0153	0.0000
Object 2	0.2599	0.0128	0.0002	0.0005	0.0000	0.0001	0.0137	0.0000
Terrain I	0.6545	0.0906	0.0001	0.0001	0.0000	0.0000	0.0844	0.0000
Terrain II	0.3336	0.0233	0.0001	0.0001	0.0000	0.0000	0.0220	0.0000
Terrain III	0.6834	0.0933	0.0010	0.0002	0.0000	−0.0001	0.0934	−0.0001

**Table 2** The values of the differences modules of the IFs for an object and terrain type

	$ J(TI) - J(O1) $	$ J(TI) - J(O2) $	$ J(TII) - J(O1) $	$ J(TII) - J(O2) $	$ J(TIII) - J(O1) $	$ J(TIII) - J(O1) $
$J_1$	0.3394	0.3945	0.0186	0.0737	0.3684	0.4235
$J_2$	0.0526	0.0778	0.0148	0.0105	0.0553	0.0805
$J_3$	0.0001	0.0000	0.0000	0.0000	0.0008	0.0008
$J_4$	0.0000	0.0004	0.0001	0.0004	0.0001	0.0003
$J_5$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$J_6$	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001
$J_7$	0.0691	0.0707	0.0067	0.0083	0.0781	0.0797
$J_8$	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001

In the next step, as shown in Table 3, for each invariant function  $J$  we order the values of the function  $\Phi(k, l)$  from the lowest to the highest.

Next, we will calculate the mean value for three lowest values of  $\Phi(k, l)$ . They are shown in Table 4.

According to Table 4, we are able to assess which invariant function  $J$  suits the best for distinguishing an object from the background. The best one is the one for which the mean of  $\Phi(k, l)$  is the highest. In Table 4 the best functions  $J$  are in the

**Table 3** The ordered values of the differences  $\Phi(k, l)$  of the invariant functions

	$ J(Tj) - J(Oi) $					
$J_1$	0.0186	0.0737	0.3394	0.3684	0.3945	0.4235
$J_2$	0.0105	0.0148	0.0526	0.0553	0.0778	0.0805
$J_3$	0.0000	0.0000	0.0000	0.0001	0.0008	0.0008
$J_4$	0.0000	0.0001	0.0001	0.0003	0.0004	0.0004
$J_5$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$J_6$	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001
$J_7$	0.0067	0.0083	0.0691	0.0707	0.0781	0.0797
$J_8$	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001

**Table 4** The mean value for three lowest values of  $\Phi(k, l)$ 

	$ J(Tj) - J(Oi) $			Mean value
$J_4$	0.0186	0.0737	0.3394	0.1439
$J_7$	0.0067	0.0083	0.0691	0.028033
$J_5$	0.0105	0.0148	0.0526	0.025966
$J_1$	0.0000	0.0001	0.0001	0.00007
$J_6$	0.0000	0.0000	0.0000	0.0000
$J_3$	0.0000	0.0000	0.0000	0.0000
$J_2$	0.0000	0.0000	0.0000	0.0000
$J_8$	0.0000	0.0000	0.0000	0.0000



top half of the table. If we create the features vector using a few functions  $J$ , it will guarantee a better separation of an object from the background. After choosing four functions  $J_4$ ,  $J_5$ ,  $J_7$  and  $J_1$ , we will conduct the recognition of a given object on a given background.

Figures 3, 4 and 5 presents the selected objects on the backgrounds TI, TII and TIII.



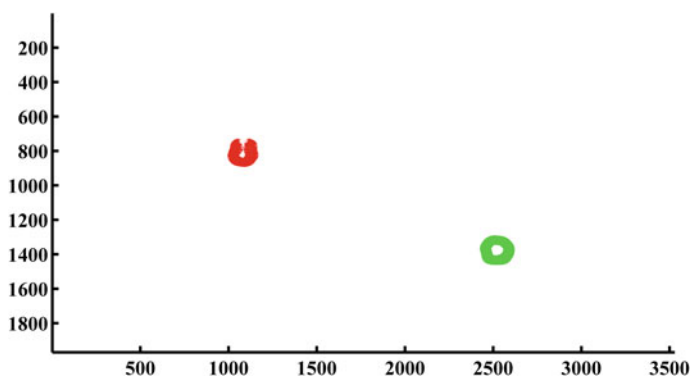
**Fig. 3** The example of the objects '1' and '2' on the background I



**Fig. 4** The example of the objects '1' and '2' on the background II



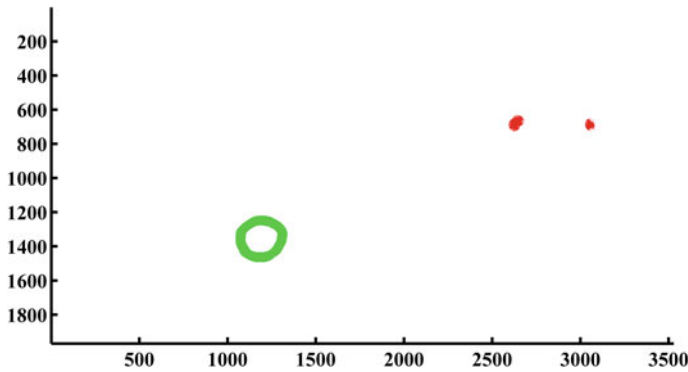
**Fig. 5** The example of the objects '1' and '2' on the background III



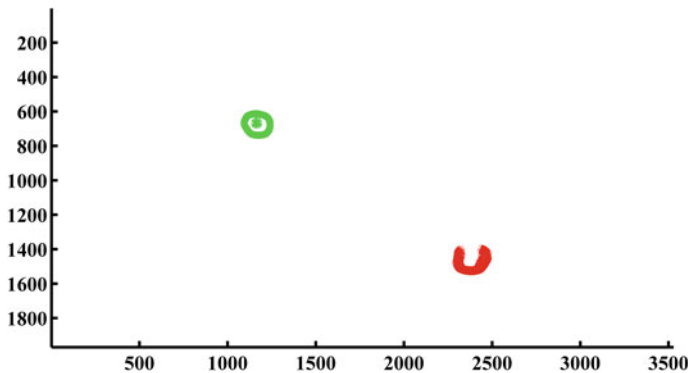
**Fig. 6** The location of the object '2' (*green*) and object '1' (*red*) in the Fig. 3 (Color figure online)

Having the values of the pattern vector of the invariant functions  $[J_4, J_5, J_7, J_1]$  both for an object and the background, we can carry out the recognition of the object on the background by assigning each point of the image to the background or the object. In order to assess the distance between the pattern vectors and current vectors for the object and background, we will use Euclidian metric. Figures 6, 7 and 8 presents the calculated location of the recognised object.

The areas, which were marked by appropriate colours, correspond to the minimal values of the distances between the model pattern vector for a given object and the pattern vector for the current location of the aperture. When we compare Figs. 3, 4 and 5, we can observe that the location of the object was correctly determined.



**Fig. 7** The location of the object '2' (green) and object '1' (red) in the Fig. 4 (Color figure online)



**Fig. 8** The location of the object '1' (green) and object '2' (red) in the Fig. 5 (Color figure online)

## 6 Conclusions

This paper has presented an algorithm of creating the invariant functions vector which guarantees the best separation of the given object classes. In this very case, these classes consisted of different objects and different terrain types where these objects were moving. This algorithm allows to assess the usefulness different invariant functions used for classifying the given set of classes.

At the same time, owing to the ordering of the invariant functions from 'the best' to 'the worst', we can when necessary increase the size of the features vector if this size is too small to classify properly the elements belonging to the particular classes. The presented example illustrates that the proposed method works properly.

## References

1. Bibik, P., Narkiewicz, J.: Helicopter optimal control after power failure using comprehensive dynamic model. *J. Guid. Control Dyn.* **35**, 1354–1362 (2012)
2. Bibik, P., Narkiewicz, J.: Helicopter modeling and optimal control in autorotation. In: *Annual Proceedings—American Helicopter Society*, vol. 64, no. 2, pp. 986 (2008)
3. Davies, D., Palmer, P.L., Mirmehdi, M.: Detection and tracking of very small low contrast objects. In: *Proceedings of the 9th British Machine Vision Conference*, Sept 1998
4. Zhang, S., Karim, M.A.: Automatic target tracking for video annotation. *Op. Eng.* **43**, 1867–1873 (2004)
5. Irani, M., Peleg, S.: Improving resolution by image registration. *CVGIP Graph Models Image Process.* **53**, 231–239 (1991)
6. Chesnaud, C., Refegier, P., Boulet, V.: Statistical region snake-based segmentation adapted to different physical noise models. *IEEE Trans. Patt. Anal. Mach. Intell.* **21**, 1145–1157 (1999)
7. Gordon, N., Ristic, B., Arulampalam, S.: *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, Boston (2004)
8. Sharp, C., Shakernia, O., Sastry, S.: A vision system for landing an unmanned aerial vehicle. In: *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*, vol. 2, pp. 1720–1727. IEEE, Los Alamitos (2001)
9. Casbeer, D., Li, S., Beard, R., Mehra, R., McLain, T.: *Forest fire monitoring with multiple small UAVs*, Portland, OR, Apr 2005
10. Papoulis, A.: *Probability, Random Variables, and Stochastic Processes*, 3rd edn. McGraw-Hill, New York (1991)
11. Sonka, M., Hlavac, V., Boyle, R.: *Image Processing, Analysis and Machine Vision*. Thompson, Stamford (2008)
12. Babiarz, A., Bieda, R., Jaskot, K.: Vision system for group of mobile robots. In: *Vision Based Systems for UAV Applications. Studies in Computational Intelligence*, vol. 481, pp. 139–156 (2013). ISBN: 978-3-319-00368-9
13. Ryt, A., Sobel, D., Kwiatkowski, J., Domzal, M., Jedrasiak, K., Nawrat, A.: Real-time laser point tracking. In: *Computer Vision and Graphics. Lecture Notes in Computer Science*, vol. 8671, pp. 542–551 (2014)
14. Nawrat, A., Jedrasiak, K.: Fast colour recognition algorithm for robotics. In: *Problemy Eksploatacji*, pp. 69–76 (2008)
15. Jedrasiak, K., Nawrat, A., Daniec, K., Koterias, R., Mikulski, M., Grzejszczak, T.: A prototype device for concealed weapon detection using IR and CMOS cameras fast image fusion. In: *Computer Vision and Graphics. Lecture Notes in Computer Science*, vol. 7594, pp. 423–432 (2012)
16. Bieda, R., Grygiel, R.: Wyznaczanie Orientacji Obiektu w Przestrzeni z Wykorzystaniem Naiwnego Filtru Kalmana. *Przegląd Elektrotechniczny* **90**, 34–41 (2014)
17. Galuszka, A., Bereska, D., Simek, K., Skrzypczyk, K., Daniec, K.: Wykorzystanie Elementów Teorii Grafów w Systemie Analiz Kryminalnych. *Przegląd Elektrotechniczny* **86**, 278–283 (2010)
18. Daniec, K., Jedrasiak, K., Koterias, R., Nawrat, A.: Embedded micro inertial navigation system. *Appl. Mech. Mater.* **249**, 1234–1246 (2013)
19. Barnat, W., Niezgoda, T., Panowicz, R., Sybilski, K.: The influence of conical composite filling on energy absorption during the progressive fracture process. *WIT Trans. Model. Simul.* **51**, 625–633 (2011)
20. Bereska, D., Daniec, K., Fras, S., Jedrasiak, K., Malinowski, M., Nawrat, A.: System for multi-axial mechanical stabilization of digital camera. *Vision Based Systems for UAV Applications. Studies in Computational Intelligence*, vol. 481, pp. 117–189 (2013). ISBN: 978-3-319-00368-9, 2013

21. Sroka, M., Sciegienka, P., Babiarz, A., Jaskot, K.: Prototyp bezzalogowego pojazdu podwodnego - układ stabilizacji i utrzymania zadanego kursu. *Przegląd Elektrotechniczny* **89**, 205–217 (2013)
22. Jaskot, K., Babiarz, A., Sroka, M., Sciegienka, P.: Prototyp bezzalogowego pojazdu podwodnego - konstrukcja mechaniczna, panel operatora. *Przegląd Elektrotechniczny* **89**, 52–67 (2013)
23. Babiarz, A., Bieda, R., Jedrasiak, K., Nawrat, A.: Machine vision in autonomous systems of detection and location of objects in digital images. In: *Vision Based Systems for UAV Applications. Studies in Computational Intelligence*, vol. 481, pp. 3–25 (2013). ISBN: 978-3-319-00368-9
24. Grzejszczak, T., Mikulski, M., Szkodny, T., Jedrasiak, K.: Gesture based robot control. In: *Computer Vision and Graphics. Lecture Notes in Computer Science*, vol. 7594, pp. 407–413 (2012)
25. Jedrasiak, K., Andrzejczak, M., Nawrat, A.: SETH: the method for long-term object tracking. In: *Computer Vision and Graphics, Lecture Notes in Computer Science*, vol. 8671, pp. 302–315 (2014)

<http://www.springer.com/978-3-319-21117-6>

Innovative Simulation Systems

Nawrat, A.; Jedrasiak, K. (Eds.)

2016, XIII, 444 p. 382 illus., 162 illus. in color.,

Hardcover

ISBN: 978-3-319-21117-6