

# Mobile system for road sign detection and recognition with template matching

Michał Maćkowski<sup>1,\*</sup>, Michał Sawiski<sup>1</sup>, and Wojciech Walczyszyn

<sup>1</sup>Silesian University of Technology, Institute of Informatics, Akademicka 16, 44-100 Gliwice, Poland

**Abstract.** This paper explores the effective approach to road sign detection and recognition based on mobile devices. Detecting and recognising road signs is a challenging matter because of different shapes, complex background and irregular sign illumination. The main goal of the system is to assist drivers by warning them about the existence of road signs to increase safety during driving. In this paper, the system for detection and recognition of road signs was implemented and tested with the use of Open Source Computer Vision Library (OpenCV). The system consists of two parts. The first part is the detection stage, which is used to detect the signs from the whole image frame and includes the modules: data-image acquisition, image pre-processing and sign detection. During this stage, the impact of Canny edge detector and Hough transform parameters on the quality-level of sign detection was tested. The second part is the recognition stage, whose role is to match the detected object with *a priori* models of signs in the dataset. In the research, the authors also compared the influence of various image processing algorithms parameters to the time of road sign recognition. The discussion part answers also the question whether the mobile system (smartphone) is robust enough to detect and recognise road signs in real time.

## 1 Introduction

The development of computer science, particularly computer processing of digital images, has led to designing vision systems that support the driver by providing information on the traffic lane or road signs. Many modern cars are equipped with these systems. The main drawback of such solutions is a limited collection of recognised road signs. Moreover, the drivers of older generation cars cannot use such systems without purchasing expensive external devices. However, most motorists have a mobile phone equipped with a video camera and a fast processor. The authors, therefore, decided to use the computing power of modern mobile devices to develop a vision system that allows the user to recognise the road signs in real time.

The issue of recognising road signs became popular in the 2000s. However, it is only in the current decade that a significant progress has been made in this area. There are both commercial and scientific systems that recognise road signs with high efficiency, such as Mobileye or systems dedicated for BMW, Volvo or Opel cars.

There are certain trends in the research works in the field regarding the approach to the problem of road sign detection and recognition. Most solutions are based on three modules – a module for detecting, defining features and sign recognition [1]. Sign detection usually takes place similarly, using image segmentation in the HSV (Hue Saturation Value) model. There are also works describing systems that detect signs based on the RGB (Red Green Blue) model. The sign features are usually

defined as the shape and colour of a given sign. Both the Hough transform [2] and other methods, *e.g.* operating on the contour path, are used to describe the shape [3]. Sign recognition is mostly performed by means of neural networks [4], where the vector of the previously extracted features is given at the input of the network. There are also solutions that use generic algorithms to generate a population of patterns [5]. In some papers, recognition of signs in monochrome images is also possible using the feature matching method [6-7].

The effectiveness of such systems is exceptionally high, as it reaches over 90%, however, typically they focus only on a narrow group of recognised road signs [5, 8]. Most projects, both commercial and scientific, limit the group of recognised symbols to prohibition signs – speed limit and “no overtaking.” Such a narrow set of signs allows the authors to use specific features of these signs to optimise the algorithm.

Most of the existing systems for road signs detection and recognition are built for desktop or dedicated computers. Even if there are solutions for mobile devices, the image acquired *via* a smartphone is usually sent to a computer (server) or another device *via* a radio link and is subsequently analysed there [9-10]. The recognition result is sent back to the device.

The objective of this paper is to explore the possibility of developing a mobile vision system that would be able to perform automated recognition of road signs. The system will be implemented as an application designed for mobile devices with Android system (with the use of Open Source Computer Vision Library –

\* Corresponding author: [michal.mackowski@polsl.pl](mailto:michal.mackowski@polsl.pl)

OpenCV), which analyses the image from the camera in real time and provides information on the recognised road signs. To achieve the goal it was necessary to examine the impact of the camera parameters capturing the image on the quality of the obtained results. The next step was to compare different methods of detecting and recognising road signs. The work discusses the most important issues, which are crucial for the development of the algorithm and the selection of its parameters. The work also contains the analysis of the research results on the impact of parameters on time and the correctness of sign recognition.

## 2 Methodology

The development of the system for the recognition of signs is based on two main modules:

- a) **Detection:** It is used to detect the signs from a whole image frame. It includes the modules: data-image acquisition, image pre-processing and sign detection. This phase greatly reduces the amount of information to be processed later.
- b) **Recognition:** It identifies road signs by comparing the information provided by the previous phase with the sign pattern stored in a database.

### 2.1 Detection module

#### 2.1.1 Image acquisition

The tests were performed with a standard mobile phone with Android system (main camera parameters: 16 Mpix, f/1.9, 28mm, 1/2.6", 1.12µm, OIS, AF). The tests examined the influence of the input image size on the accuracy of road sign recognition and the algorithm operation time. The optimal size of the frame obtained from the camera and subjected to recognising road signs was 840x630 pixels. This value is approximate because the current frame size on a given device may be different.

There are many sizes of frames available on a given device; hence, it is necessary to make a search so that the picture frame is not smaller than the given size specified as the optimum. A series of tests were also carried out to select the best focus mode. The best choice was the "infinity" mode, which extends the maximum focal length of the lens making distant objects the sharpest and blurring the elements in the foreground of the photo.

#### 2.1.2 Image pre-processing

The image subjected to the sign detection process may be noisy, distorted or too dark, which are the issues addressed at the pre-processing stage. The main element of the developed algorithm in the pre-processing phase is the equalisation of the image histogram.

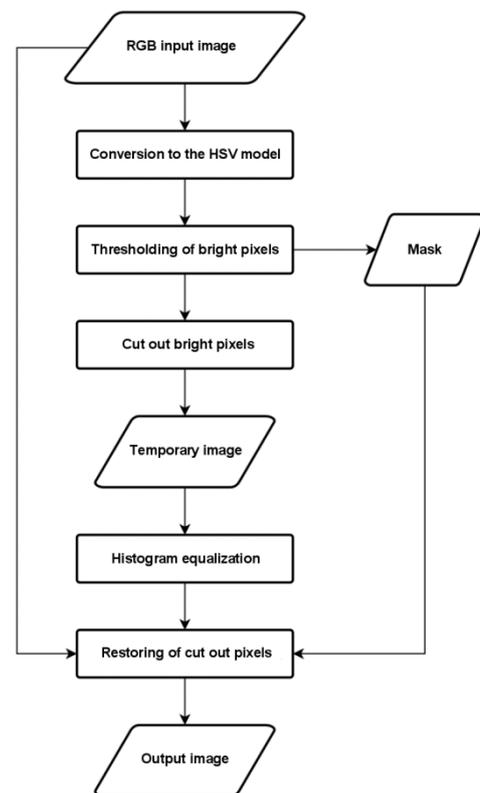
The solution operates on a colour image that was originally stored in the RGB model. Attempts were made to equalise the histogram of such an image using the following methods:

- 1) equalising the histogram of each component of the RGB model;
- 2) equalising the histogram of the brightness and saturation component of the HSV model;

- 3) equalising the histogram of the brightness and saturation component of the HSV model excluding the brightest pixels.

Equalising the histogram using methods 1 and 2 may not always give the desired results. Images captured under the sun are generally well-lit, but objects such as road signs are very dark. Despite the fact that signs are covered with reflective foil, the lack of direct sunlight on the sign means that it does not reflect enough radiation. It happens very often that the standard histogram equalising methods fail because, in general terms, the image dynamics is quite good. To prevent this, the authors developed an algorithm that cuts off the part of the histogram disturbing its balance by removing the lightest pixels from an image subjected to the histogram adjustment. The result of this process is a brighter image. Figure 1 shows a flowchart of the operation of this algorithm. To get a mask containing the location of the brightest pixels the image is subjected to a threshold operation with the threshold value (1).

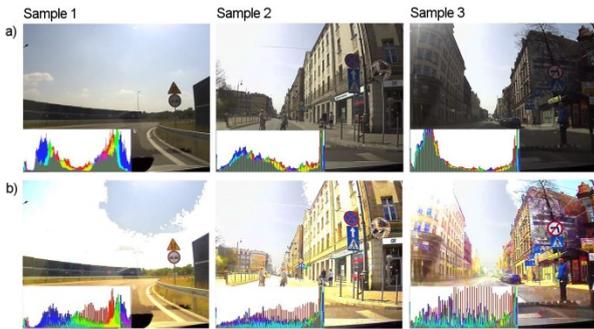
$$T = 90\% * 256 = 230 \quad (1)$$



**Fig. 1.** Diagram of the operation of the modified histogram equalisation method.

Figure 2 presents:

- a) Sample input images with an unbalanced histogram.
- b) Results of histogram equalisation with a modified method based on the HSV model.



**Fig. 2.** Sample input images before (a) and after (b) histogram equalisation with a modified method based on the HSV model.

The study considered two situations – with and without histogram equalisation, and analysed the impact of their presence/absence on the effectiveness of road sign detection and recognition.

### 2.1.3 Detection of Regions Of Interest

To detect road signs and determine their shape, it is essential to know where potential signs are in the image; therefore, the Regions Of Interest (ROI) are located. To do this, special features of road signs that distinguish these objects from the surroundings can be used. The basic colours of most road signs in Poland are yellow, red and blue. Assuming the 40% of brightness threshold value and 65% of saturation threshold value, the image areas of individual colour sign (in the HSV model) are given as (2-4):

$$H \in (22^\circ; 76^\circ) \wedge S \in (65\%; 100\%) \wedge V \in (40\%; 100\%) \quad (2)$$

$$H \in (188^\circ; 250^\circ) \wedge S \in (65\%; 100\%) \wedge V \in (40\%; 100\%) \quad (3)$$

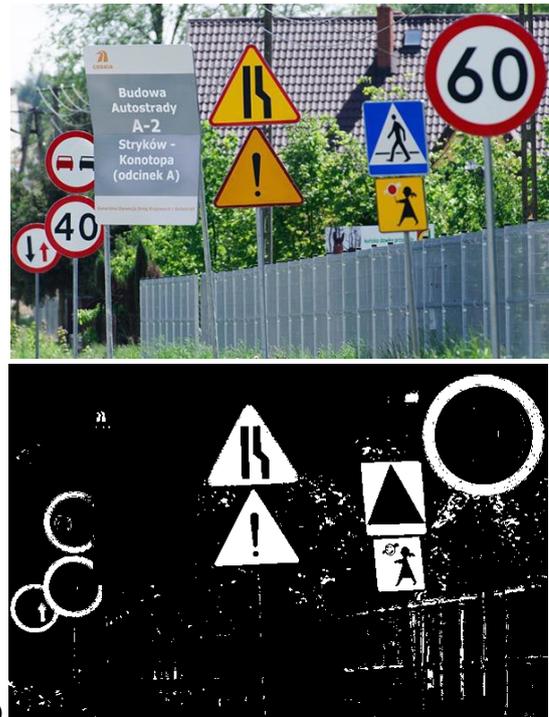
$$H \in (340^\circ; 22^\circ) \wedge S \in (65\%; 100\%) \wedge V \in (40\%; 100\%) \quad (4)$$

where the ranges of hue component in HSV represent the colours of the road signs, respectively: yellow, blue, red.

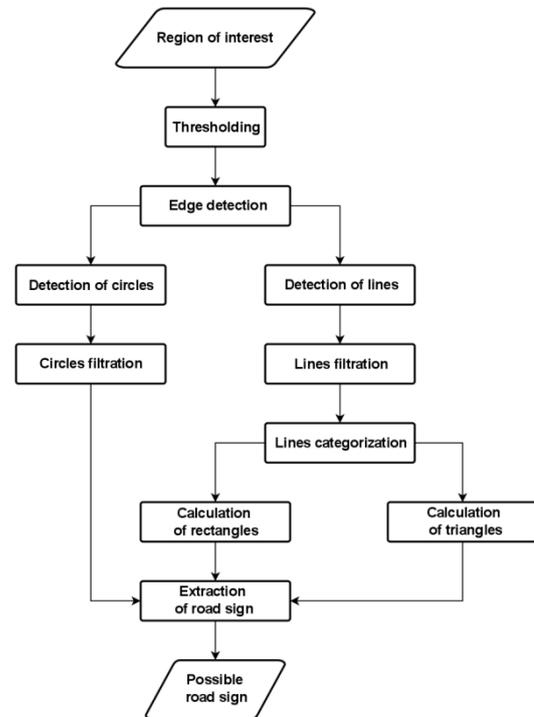
After preliminary tests on the characteristics of road signs, the next step is the image thresholding with the previously determined thresholds values. Separate binary masks are created for each colour, which are then merged by a logical sum operation. Figure 3 shows the result of thresholding of the example image.

In an image prepared in such a way, the edges and ROI were detected using the Canny edge detector (multi-stage algorithm); next, lines (Hough transform) and circles (a modified version of Hough transform) will be searched in the ROI areas among the edges found in these areas. The lines found are then used to calculate triangles and rectangles – *i.e.* possible road signs. Figure 4 shows the scheme of the applied algorithm.

Detecting rectangular and triangular road sign shapes can be simplified to detecting lines whose intersections would be used to determine the desired shapes. Line detection is performed by the OpenCV library function implementing the original Hough transform and its variation to detect circles. In both cases, the parameters of Hough transform have a significant effect on the number and precision of detected lines.



**Fig. 3.** Image thresholding: a) a sample image with road signs b) binary image after the colour threshold operation [11].

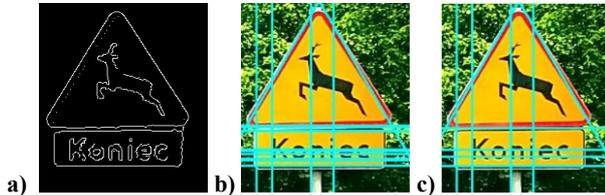


**Fig. 4.** Diagram of the algorithm for detecting shapes in an image.

A series of tests were carried out to determine the optimal value of the following parameters:

- a)  $\rho_{\text{resolution}}$  – distance resolution of the accumulator in pixels,
- b)  $\theta_{\text{resolution}}$  – angular resolution of the accumulator in radians.

*a priori* that the angular resolution of the accumulator  $\theta$  should be  $30^\circ$  because this value is the greatest common divisor of the values of the angles in the rectangle and the equilateral triangle (most warning road signs). Table 1 presents the results of research on the impact of the  $\rho_{\text{resolution}}$  parameter on the number of detected lines at a selected value of  $T = 27$  and  $\theta_{\text{resolution}} = 30^\circ$  for an example image (Figure 5a)



**Fig. 5.** Line detection: a) sample image used for line detection b) detected lines for given parameters of Hough transform c) Reduced set of lines detected in the image.

Analysing the results from Table 1, as well as Figure 5b, it can be noticed that the detected lines with the parameter  $\rho_{\text{resolution}} = 1$  match the most important edges of the object, and the ratio of the number of detected lines to the number of correctly detected lines is the best (this parameter value was chosen for further research). In the example image, 15 lines were detected (Figure 5b), only 5 of which were correct. Therefore, it is necessary to reject as many additional lines as possible at the stage of filtration by merging nearby lines with the same coordinate  $\theta$  (Figure 5c).

**Table 1.** The influence of  $\rho$  Hough transform parameter on the number of detected lines.

$\rho_{\text{resolution}}$	Number of detected lines	Number of correctly detected lines
1	15	5
2	69	7
3	62	7
4	52	6
5	40	5
6	33	5
7	27	5
8	28	4
9	23	4
10	19	3

## 2.2 Recognition module

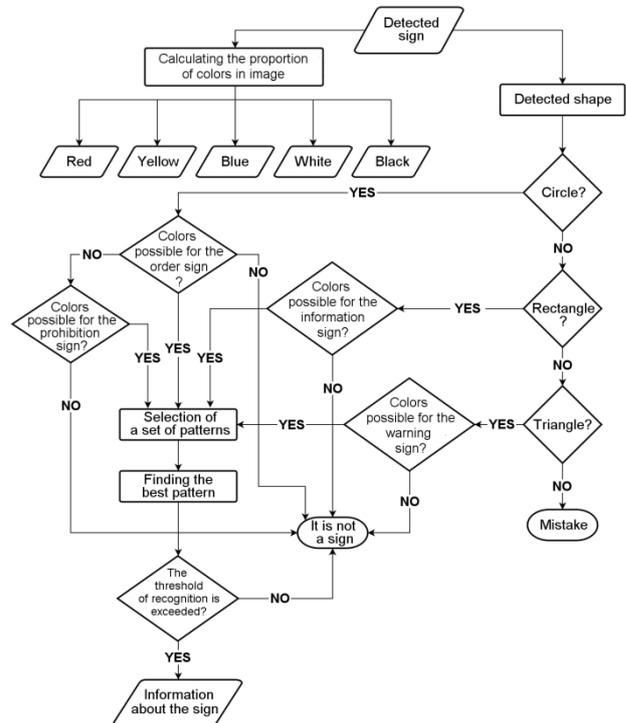
The sign recognition takes place in two steps. The first is colour analysis, which allows determining which colours appear in the image. By combining information about the shape of a sign and its colours, the sign can be qualified for further recognition or rejected at this stage (algorithm optimisation). Figure 6 presents the general flowchart of the algorithm for road sign recognition.

### 2.2.1 Road signs template matching

Road signs are strictly defined; their content is known and predictable. In addition, the symbols appearing on the signs are easily distinguishable and their collection is different for each category of road signs. Therefore, to recognise symbols on detected signs, it was decided to use the pattern matching method. 152 patterns – binary

images were prepared, depicting symbols appearing on signs. Figure 7 presents examples of binary images. The database of pattern images has been divided into 4 sets, one for each shape of the road signs:

- rectangle: 21 patterns,
- circle: 88 patterns,
- triangle: 42 patterns,
- octagon (stop sign): 1 pattern.



**Fig. 6.** Flowchart of the road sign recognition algorithm.



**Fig. 7.** Sample road signs symbols used during the recognition phase (pattern matching).

The results of detecting road signs in the image provide information about the shape of the sign being recognised. This allows reducing the pattern images to those that may appear in the signs of a given shape.

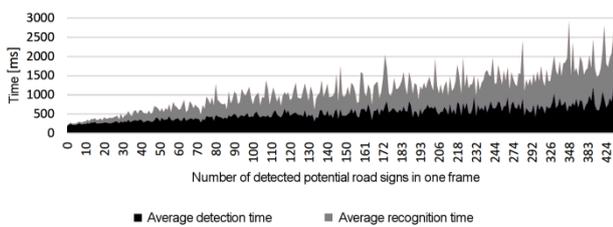
The OpenCV library implements a search algorithm on a given image of a given pattern. To put it simply, the result of such an operation is information about the position of the matching pattern in the input image and the accuracy of this match. Using this mechanism, an algorithm that finds the best fit for a given image and a vector of features was developed. The result of the algorithm operation is the index of the best-fitted template and the accuracy it was matched to. Based on the test results, the threshold for signs recognition was chosen specified at  $T_{\text{matching}} = 0.65$ . The signs whose matching pattern did not exceed this threshold are rejected.

### 3 Results and discussion

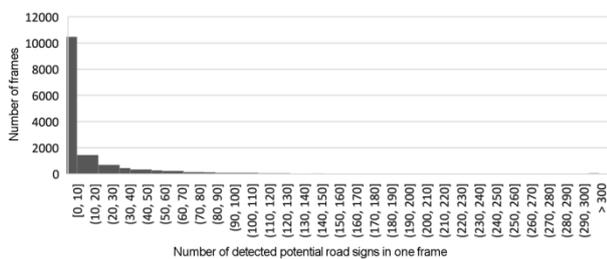
During the tests of the proposed solution for detecting and recognising road signs in the real time, data from nearly 15,000 picture frames were collected, where over 8000 potential signs were detected and classified. These data were analysed and grouped. One group contains the results obtained for images whose HSV histogram was equalised (excluding the lightest pixels), the second group – images whose histogram was not equalised. The histogram equalisation mechanism was turned on and off during the experiment to get differentiated data.

#### 3.1 Analysis of detection time and sign recognition

Figure 8 shows the dependence of the number of signs detected on the analysis time of one image frame. The chart is based on all data collected during the detection of potential signs, without division into groups. Based on the results in Figure 8, the detection time is less dependent on the number of potential signs detected, and even if no sign is detected, the time is comparatively high. This results from the fact that most operations in the detection phase are always performed before the signs are detected. The graph comprises all data, including extreme cases, as shown in Figure 9. It illustrates that most samples contain from 0 to 10 detected signs. Based on all the data, without division into groups, the total average road sign recognition time was calculated, and amounted to  $t = 225.76 \text{ ms}$



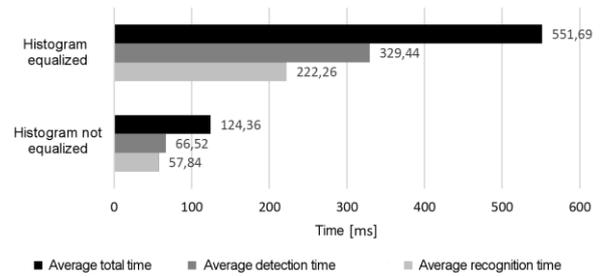
**Fig. 8.** Influence of the number of potential detected road signs on the analysis time.



**Fig. 9.** Histogram of the number of detected road signs on one frame of the image based on all frames acquired during the experiment.

It should be noted how the algorithm execution times vary depending on the way the image is processed. An analysis of the execution times of the following stages of the algorithm for two groups of samples was made, and its results are shown in Figure 10. The execution time of the algorithm for images that require histogram equalisation is almost 5 times higher than the

corresponding results obtained for images without histogram equalisation, and 2 times longer than the average time. This difference results from the various number of signs detected in the images in both groups.

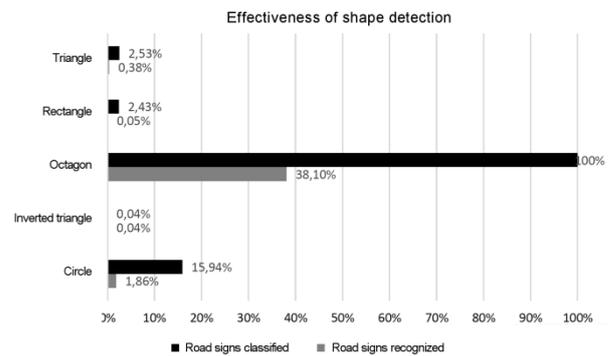


**Fig. 10.** The execution time of subsequent stages of the algorithm in the proposed system.

#### 3.2 Analysis of road sign detection

Following the analysis of the obtained data, it is possible to determine the average number of detected signs in the frame of the image. Histogram equalisation results in the detection of a much larger number of road signs (average: 50.83) compared to the situation without histogram equalisation (6.39). Considering the average number of potential signs detected in the image and the effectiveness of such a process, it becomes apparent that the use of histogram equalisation of the input image increases the efficiency of sign recognition more than twice.

Based on the analysis of the collected data, the algorithm shows the lowest efficiency for shapes whose position is calculated from the detected lines (Figure 11). The largest is for octagons, which are detected in a different way from the rest of the shapes.



**Fig. 11.** The number of classified and recognised samples among all detected road signs of a given shape.

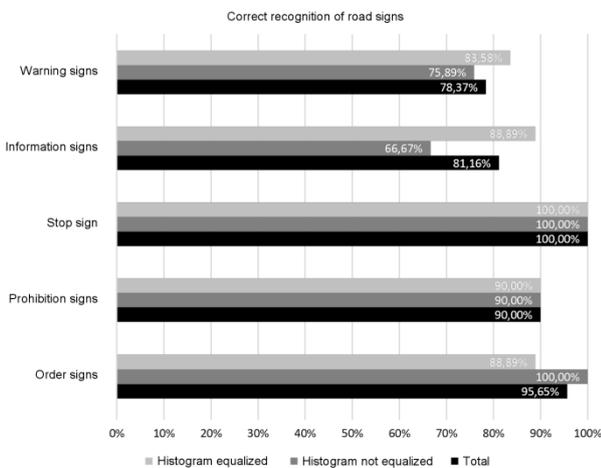
#### 3.3 Analysis of the correctness of sign recognition

Overall, average correctness of sign recognition is presented in Table 2. Equalisation of the input image histogram, although it extends the analysis time and causes the detection of additional areas requiring recognition, provides a better quality of results. In images with equalised histogram, objects are brighter

and their contrast is greater, which in turn leads to more accurate image segmentation and a more robust algorithm in terms of making fewer mistakes in recognising signs. Figure 12 presents the correctness assessment for particular sign categories.

**Table 2.** Average correctness of recognising road signs.

Histogram not equalised	Histogram equalised	Total
79.64%	88.76%	83.59%



**Fig. 12.** The correctness of recognising road signs.

## 4 Conclusion

Based on the test results, it can be concluded that the proposed method of equalising the image histogram in the HSV model that does not include the brightest pixels of the image, positively influenced the quality of the obtained results. This method enables the detection of signs in unfavourable lighting conditions, in which such an operation would be impossible or difficult for images with no aligned histogram. Unfortunately, the increased accuracy of road sign detection negatively affects the time it takes to perform the operation.

The process of detecting the sign shapes unfortunately involves certain drawbacks. The main disadvantage of such a solution is the excess number of areas that are to be recognised; as it is confirmed by test results, which clearly state that 0.38% of detected triangles and only 0.05% of detected rectangles are finally recognised with an accuracy exceeding the set threshold. This has a negative effect both on the duration of the algorithm realisation and the correctness of the results. On the other hand, detection and recognition of the "STOP" sign is flawless in the examined data set. This sign is very characteristic, and the sets of patterns taken into account when calculating the match are the smallest of all.

The correctness of signs recognition that is over 83% is a satisfying value – the system recognises 152 road signs. A necessary condition – the signs are to be recognised before the vehicle passes them – was also fulfilled. The application analyses at an average speed of 4 frames per second, which allows the user to move at a

maximum speed of 130 km/h in built-up areas and 260 km/h on highways (this is possible due to the difference in road sign sizes).

Among the problems encountered during the development and testing of applications, the most notable is the huge consumption of power by a mobile phone. The application analyses the image in real time, and the amount of data is not limited in any way. This means that the data is processed at maximum speed throughout the application lifetime, and as a consequence, the battery is discharged rapidly. The results of the appropriate calculations were confirmed experimentally, and it was proved that the working time on the battery power supply of the average smartphone with the developed application on amounts to mere 2-3 hours. Such a result is unacceptable for most users. In addition, the continuous load of the processor results in heating up the entire device structure, which in turn leads to overheating and power cut. Therefore, although the results are very promising, further research and optimisation of the developed system are needed.

The research presented in this paper was done with the use of equipment purchased within the Motorola Solutions Foundation grant realised by Institute of Informatics at Silesian University of Technology, no 02/020/ ZZZD17/0122 (ZZD/3/Rau2/2017/507).

## References

1. S. Aigerim, Z. Askar, Z. Meirambek et al., IEEE 8th International Conference on AICT, 1-6 (2014)
2. E. Onat and Ü. Üzdil, 23rd Signal Processing and Communications Applications Conference (SIU) 2161-2165 (2015)
3. A. Salhi, B. Minaoui, M. Fakir, ICMCS, 433-437 (2014)
4. P. Shopa, N. Sumitha, P. S. K. Patra, ICICES2014, 1-6 (2014)
5. M. Kobayashi, M. Baba, K. Ohtani et al., IEEE/SICE International SII, 455-460 (2015)
6. J. Duan, M. Viktor, ICCAIS, 107-112 (2015)
7. F. Ren, J. Huang, R. Jiang, et al., 24th International Conference Image and Vision Computing, 409-414 (2009)
8. K. Fu, I. Y. H. Gu, A. Ödöblom, IEEE Intelligent Vehicles Symposium, 443-448 (2015)
9. C. H. Lai, C. C. Yu, International Conference on Technologies and Applications of Artificial Intelligence, Hsinchu City, 195-202 (2010)
10. B. Yavas, M. M. Özbilen, Y. Yilmaz, IEEE 8th International Conference on AICT, 1-4 (2014)
11. T. Zamojski, (Image source) KONTROLA: Droga przez mękę. NIK skontrolowała znaki drogowe. <http://www.tygodnikzamojski.pl/artykul/47938/kontrola-droga-przez-meke-nik-skontrolowala.html> (2014)