Received: 14 December 2025 | Revised: 5 March 2026 | Accepted: 17 March 2026 | Published online: 7 April 2026

RESEARCH ARTICLE

Smart Wearable Technology
2026, Vol. 00(00) 1–16
DOI: 10.47852/bonviewSWT62028788

BON VIEW PUBLISHING

A Method for Determining Obstacle Contours Using Wearable Video Cameras for Orienting Blind and Visually Impaired People in Urban Environment

Oleksandr Poliarus[1,*] and **Yevhen Poliakov**[1]

[1] *Department of Metrology and Life Safety, Kharkiv National Automobile and Highway University, Ukraine*

Abstract: The essence of the method is based on the division of the video camera matrix into submatrices, in which the Euclidean distance is determined between adjacent rows and columns. If there is a gradient of pixel intensities exceeding a certain threshold, this leads to a jump in the Euclidean distance, which indicates the presence of an element of the contour of the entire object in a separate submatrix. The quality of such a contour determination depends on the sizes of the submatrices and the selected threshold. To expand the capabilities of the method in estimating the complex shape of objects, a combined submatrix method is proposed, in which contours are determined at different submatrix sizes and threshold values, and then superimposed on each other. The sizes of two different submatrices are determined based on previous research. The possibilities of optimizing such sizes using a genetic algorithm are considered. The method provides the highest similarity of contours to the original image of steps compared to other methods, and the accuracy of determining the contours of a rectangle, triangle, and circle is comparable to the accuracy of such methods.

Keywords: combined submatrix method, navigation of blind and visually impaired people, object contour recognition, video camera technology

1. Introduction

Scientific and technical achievements of civilized countries can be aimed at improving the lives of people with disabilities. Orientation of blind and visually impaired people (BVIPs) in the environment is often carried out using a white cane, guide dogs, and tactile paths together with ground indicators. The use of a white cane does not provide information about distant dangers. Guide dogs require long-term training and maintenance (feeding, treatment, etc.). Tactile paths in the form of relief patterns on sidewalks or floors guide a person's movement, but they are not available everywhere. Modern electronic aids for people use ultrasonic, laser, or infrared sensors to detect obstacles outside the range of the cane. These include, for example, Sonic Pathfinder and smart canes for visually impaired blind walking. They allow detecting obstacles at head level in a wide range but have a high cost of acquisition and maintenance for a person without a reasonable income. Devices with wearable cameras and artificial intelligence for describing the environment and the ability to identify objects are even more expensive. GPS-based navigation systems are gaining great importance, providing step-by-step instructions to a person using various types of signals. Such systems do not provide high efficiency in all rooms, and therefore, Bluetooth beacons or technologies for automatic identification of objects using radio signals can be used here. This requires appropriate infrastructure, which cannot always be built in the house.

Most technologies for providing orientation of BVIPs in an arbitrary terrain are aimed at generating information about the presence of obstacles for a person and the distance to them. Less attention is paid to the issue of determining the shape of these obstacles using wearable video cameras, which can be pre-installed in a person's clothing. From a systems perspective, a person moving in the environment is part of a human–environment system with continuous information exchange. Equipping visually impaired people with wearable video cameras and an information processing unit creates an artificial sensory subsystem that replaces or supplements vision. As a result, a person can receive better spatial information about the size, position, shape, texture, and dynamics of surrounding objects. To achieve safe and effective interaction with the environment, an artificial subsystem must not only detect the presence of an obstacle but also report what form it takes. Ideally, this creates the conditions for recognizing, for example, a wall, a pillar, stairs, doors, people, animals, or vehicles. This information is useful for contextual decision-making regarding path planning: how to move, what to avoid, and where there may be a passage. For example, a narrow pillar can be bypassed, but the presence of a wall requires choosing a different route. The close-to-rectangular shape of an object can indicate

***Corresponding author:** Oleksandr Poliarus, Department of Metrology and Life Safety, Kharkiv National Automobile and Highway University, Ukraine. Email: poliarus_ol @khadi.kharkov.ua

both the presence of a door or wall on the path and the presence of stairs. They can be distinguished by their orientation in space (vertical or horizontal) and by the ratio of sizes in two planes. Changing the shape of an object during video camera observation can be a sign of its movement. Therefore, the shape helps to determine the type of object, which improves spatial perception and confidence of BVIPs in the process of their navigation on the terrain.

Decision-making by an artificial sensory system about the shape of surrounding objects is provided by audio or vibrotactile signals. Cognitive research results show that users of sensory devices can form accurate mental maps and instructions if the signals are corresponded to spatial geometry. The detected shapes are encoded in the form of patterns that reflect the boundaries of typical objects. Information should be minimal to prevent cognitive overload and loss of user convenience. In conditions of shadows, noise, and local disturbances of the typical shape, the system must provide the required level of probability of correct shape detection at a given probability of false alarms. From a social perspective, the interaction of several visually impaired people on the ground is possible, which can optimize the performance of the overall "human–technology–environment" system. Therefore, determining the shape of obstacles using portable video cameras is justified, since the shape conveys important spatial and semantic information. Integrating video cameras with other sensors built on different physical principles, as well as with wearable and stationary intelligent systems, will allow improving the quality of decision-making regarding the spatial structure of the environment.

2. Overview of Smart Technologies for Determining the Shape of Objects Using Wearable Video Cameras on Blind and Visually Impaired People

Intelligent technologies using video cameras significantly improve the daily lives of BVIPs. These systems use advances in computer vision, mobile technologies, and teleassistance systems to provide navigational aids and improve access to visual information. Numerous studies highlight the effectiveness and innovative design of such assistive systems. One well-known model is the teleassistance tool that supports blind people in their daily activities using smartphones. This model supports remote assistance by allowing caregivers to provide advice through real-time video captured by the user's smartphone camera. Such systems adapt well to existing technologies, demonstrating flexibility and user-centered design, which is important for user decision-making [1, 2]. Furthermore, the integration of Internet of Things (IoT) technologies offers real-time obstacle detection and avoidance, further improving mobility for BVIPs [3]. The combination of smartphone and IoT technology creates a comprehensive foundation for modern assisted navigation.

Another development is the BlindSense interface, which uses the capabilities of a smartphone to create an inclusive experience for blind people. This interface improves the interaction between users and their environment, allowing them to perform everyday tasks with greater independence through the use of built-in accessibility features [4]. Such systems are able to provide feedback and interact with each other seamlessly. Additionally, innovative navigation assistance systems provide haptic feedback and auditory cues to help visually impaired users navigate effectively both indoors and outdoors. For example, a specialized Telenavigation System broadcasts real-time video to caregivers, who then transmit navigation instructions [2]. This approach promotes mobility and creates a sense of security through immediate support, meeting users' psychological needs during navigation.

A variety of assistive technologies aimed at improving vision also holds great promise. Devices that augment existing visual inputs or provide auditory descriptions of the environment allow BVIPs to better understand their surroundings [5, 6]. In addition, systems using augmented reality can convey spatial information through audio cues, which is particularly useful in crowded or complex spaces [7]. This increases the usability of public spaces for BVIPs. Despite the wide range of technological advances, researchers highlight challenges in implementation and usability. While there are many assistive devices available, their effectiveness can vary significantly depending on user preferences, functionality, and specific visual impairment [8, 9]. This justifies the need to involve blind users in the design process so that technologies effectively meet their diverse needs.

Therefore, intelligent camera-based technologies for BVIPs are making significant progress. Combining teleassistance, the IoT, augmented reality, and user-centered design, these innovations not only facilitate navigation and independence but also actively engage people in using technology to improve their quality of life. Future research should focus on increasing the adoption and user satisfaction of these systems while addressing the unique challenges faced by visually impaired users. One such challenge is object shape recognition.

Solving this problem has led to the emergence of various technological innovations aimed at improving object recognition and navigation. These innovations often combine visual input with the use of video cameras and other sensors. Such developments use algorithms based on deep machine learning to detect objects. For example, intelligent visual assistance systems for face recognition use the You Only Look Once (YOLO) algorithm and Haar-Cascade classifiers. This allows visually impaired users to identify their surroundings using auditory feedback [10]. A comprehensive review of the progress in image segmentation using YOLO and its integration with convolutional neural networks is provided in the work of Irfan et al. [11]. New modifications of YOLOv4, YOLOv5, and YOLOv7 optimize the performance of the model for different tasks. For example, in the work of Wu et al. [12], YOLOv5s is proposed for multi-object detection, which is important for municipal waste recycling. This version adapts the classic YOLO architecture to improve real-time processing capabilities while maintaining high detection accuracy, especially in cluttered environments common during recycling operations. Similarly, another significant contribution to the YOLO framework is the YOLOv7 model, which is used to improve the accuracy of road damage detection in images acquired using Google Street View [13]. This adaptation demonstrates the versatility of YOLO in solving specific problems in the subject area, while simultaneously developing its application in practical settings. The improved version of YOLOv8 for road object detection provides a trade-off between accuracy, speed, and computational complexity and also has improved feature extraction, small object detection, attention mechanisms, expanded datasets, and testing [14]. Due to recent advances in computer vision, visual object recognition technologies are being actively implemented by BVIPs, facilitating their interaction with the environment [15].

A complement to the visual approach is the use of tactile interactions, which are often necessary for object recognition [16, 17]. For this purpose, algorithms for interaction between tactile and visual processing systems have been developed [18]. The specific characteristics of tactile technology contribute to the understanding of object geometry for BVIPs [19]. For this, certain areas of the visual cortex are activated, and common blocks for processing visual and tactile information are used [17].

Currently, there are advances in 3D object recognition using robotic systems that help the blind navigate indoors. Such systems, based on effectively segmented collected data, qualitatively recognize and classify objects in real-world situations [20]. The convergence of different technologies should be consistent with the capabilities of BVIPs, which significantly improves their quality of life. The main focus of the paper is on remote methods for determining the shape of objects using wearable video cameras.

3. Analysis of Existing Methods for Determining the Shape of Objects Using Video Cameras

The determination of the shape of objects is carried out within the framework of computer vision methodology, which is described in many textbooks, in particular, in the works of Szeliski [21] and Bishop [22]. To obtain the shape of an object, it is necessary to highlight its boundaries. It is logical to assume that the physical basis for determining the boundary is a sharp change in brightness or its derivatives in neighboring rows or columns of the video camera matrix. The main approaches to solving this problem are the use of gradient methods. Improving the quality of segmentation is achieved by using preliminary image filtering operations, increasing contrast, removing blur, etc. In general, an image may contain a group of objects. Its fragmentation into subsets of pixels (superpixels) corresponding to the shape of predefined objects and preliminary classification are desirable operations. Contour segmentation methods are used to analyze complex images.

The most popular methods for object boundary extraction are based on the use of the Sobel operator and its more complex variants, such as Enhanced Mask R-CNN, which works on complex backgrounds [23, 24]. In fact, the Sobel operator is a discrete differential operator that determines the approximate value of the image gradient brightness. The gradient approximation used in it is quite coarse, especially on high-frequency image fluctuations. The Mini Kirsch edge detection method, in particular, is an example of an effective edge detection technique that not only identifies boundaries but also increases the overall image clarity [23]. The most specific methods for contour segmentation are the chain code method and the specific window method [25]. In the window method, at the initial stage, areas containing low or high contrast are found, and then a window is built based on the approximate size and shape of the searched object. The object is considered selected when the window is within a given range. The contour segmentation method is able to determine the location of the object with good accuracy. Applying a boundary selection filter to an image can significantly reduce the amount of data to be processed, because the filtered part of the image is considered less significant, and the most important structural properties of the image are preserved. Unfortunately, it is not always possible to select boundaries in an image in real time that do not have large or small brightness gradients.

Modern contour segmentation methods play a crucial role in the analysis of complex images in various fields. These methods are important for extracting relevant information from images by identifying object boundaries. They can be divided into three main groups. In the first group, methods are based on active contour models (ACMs). These models provide curve formation based on image data. For this, information about both edges and image regions is used to guide the contour to the shape of the target object. ACMs are actively used in medical imaging, where they help in identifying anatomical structures for further analysis [26, 27]. ACM methods have been successfully applied to ultrasound image processing, where traditional methods suffer from low contrast and high noise [28]. Moreover, the performance of ACMs can be enhanced by saliency-driven methods that incorporate local image attributes to improve segmentation results [29]. Region- and edge-based ACMs achieve advantages when they use intensity, color, and gradient information [30, 31]. However, they are sensitive to noise, which can hinder accurate segmentation [29, 32]. The use of genetic algorithms and fuzzy logic in ACMs to overcome these limitations increases their robustness to noise and inaccurate initialization of contours [33].

The second group includes level set methods, which rely on implicit contour representations and are governed by the structure of a partial differential equation. This approach provides continuous contour deformation [34]. The introduction of hybrid velocity functions that combine boundary and region segmentation has shown promise in segmenting low-contrast and complex images. All this demonstrates the versatility of level set methods [35]. The methods have proven themselves well in medical image processing [36].

The third group includes deep learning methods that use convolutional neural networks to automate the segmentation process [37, 38]. Architectures such as U-Net provide high accuracy in identifying complex anatomical structures, effectively solving the problems of identification and consistency of segmentation results [36, 37]. Additional advantages have been established due to the combination of manual, semi-automatic, and fully automatic methods [39].

Gradient-based methods are sensitive to noise, and as a result, small variations in brightness can be mistaken for edges. Gaussian methods based on the Canny operator include smoothing using a Gaussian filter and thresholding to reduce errors in edge detection. The Canny operator is usually the most accurate edge detector, while the Sobel operator is known for its speed, which is suitable for real-time processing. In general, there are many methods for detecting object boundaries on the background of an arbitrary image from a video camera matrix. In [40], the possibilities of attributing objects to a specific shape based on the calculation of the Hausdorff distance are evaluated. The method allows to estimate the probability of detection and correct classification of landmarks for autonomous mobile robots. At present, each method of estimating the edges of objects is effective in a given range of conditions. The problem of determining objects with a complex shape using video cameras remains relevant for the tasks of BVIPs orientation. The purpose of the paper is to develop a method for determining the contours of objects on a video camera matrix in order to use it for navigation by BVIPs.

4. Method for Determining the Contours of Objects in the Environment Using Video Camera Submatrices

Boundary construction is possible when the difference in pixel intensity between the object and background (gradient) exceeds a certain threshold value. The contours of an object and its shape can be determined by comparing the intensity of pixels in adjacent rows and columns of the matrix. It is advisable to replace such an operation with an estimate of the Euclidean distance between the specified intensities. This method is convenient for detecting the boundaries of objects of regular shape with horizontal and vertical boundaries. If the boundaries of the objects are arbitrary, then the Euclidean distance must be determined not between adjacent rows and columns of the video camera matrix but between adjacent rows and columns of submatrices. To do this, the entire video camera matrix is divided into submatrices, the number of which can be predetermined or selected adaptively depending on the size of the characteristic parts of the objects, the background, the presence of obstacles, etc. The problem of determining fixed or adaptive threshold values of the pixel intensity gradient in adjacent elements of rows and columns of submatrices remains relevant.

For BVIPs, the most important objects are stairs in buildings, vertical elements such as columns, as well as rectangular doors, windows, etc. Next, to demonstrate the method, we will use examples of images with stairs oriented horizontally and vertical elements that are similar in shape to a rectangle. It is important to determine the contours of real images and decide whether the detected image should be classified as a rectangle based on the shape criterion. In this case, it is advisable to use a software method based on the following considerations.

After artificial division, the entire matrix of the video camera is transformed into a cellular one, the rows and columns of which will not be pixels but collections of pixels in the form of submatrices. In each submatrix, the Euclidean distance between all adjacent rows and, separately, between adjacent columns is determined. Sharp changes in the Euclidean distance should be expected at the border of the background and object colors. If the object element is absent in the submatrix, then the Euclidean distance is zero, since there is no gradient in the pixel intensity. If there is a jump in the Euclidean distance that exceeds a predetermined threshold, a horizontal or vertical line is created, depending on whether there is a jump in the Euclidean distance between the rows or columns of the submatrix. The location of the lines is determined by the row or column numbers for which a jump of the specified distance is observed. Let's assume that the rows of the matrix from a video camera mounted on the clothing of a person with poor vision are arranged horizontally, and the columns are arranged vertically. Otherwise, they can be aligned with the equipment for stabilizing the position of the video camera, and then there will be no influence of the person's position on determining the shape of the observed object.

The results of modeling the contours of real objects showed that the described method is able to determine some specific contour elements, for example, horizontal or vertical. To do this, it is necessary to select the dimension of the submatrices and the Euclidean distance threshold. Qualitative determination of the entire contour of an object of complex shape is not always possible. To eliminate this drawback, the paper proposes a combined submatrix method that connects the described technology for obtaining object contours for different sizes of submatrices and thresholds. For example, two options for the sizes of the submatrices $m \times n$ and $n \times m$ are selected. In each variant, the contours of the object are calculated, which are then superimposed on each other. One variant better determines horizontal contours, and the other—vertical. The Euclidean distance thresholds can also be different. Figure 1 shows an illustration representing the technology of the combined submatrix method, which describes the sequence of converting the original image of the stairs into contours.

Figure 2 presents a more detailed version of the stair contour detection. It follows that in simple situations (good lighting, no interfering objects, etc.), the combined submatrix method reliably detects horizontal and vertical elements of stairs.

Figure 1
Example of converting the original image of stairs into contours using the combined submatrix method

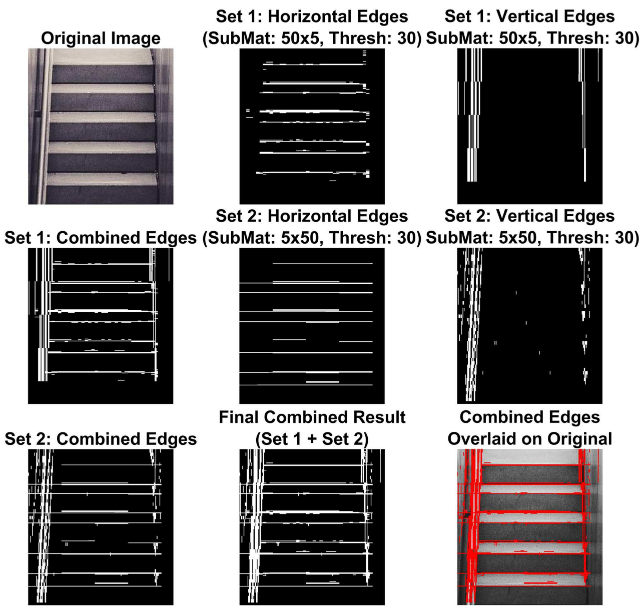
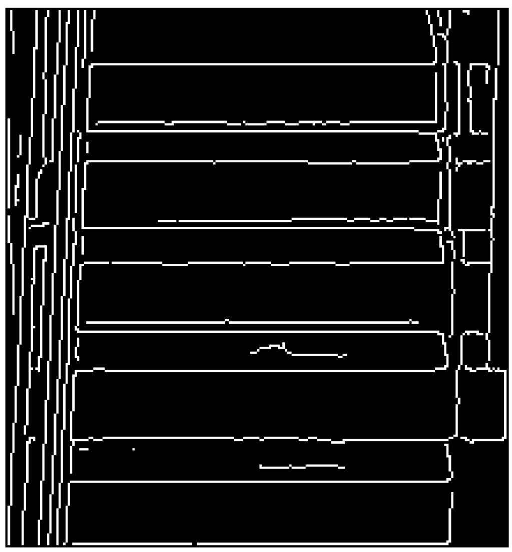


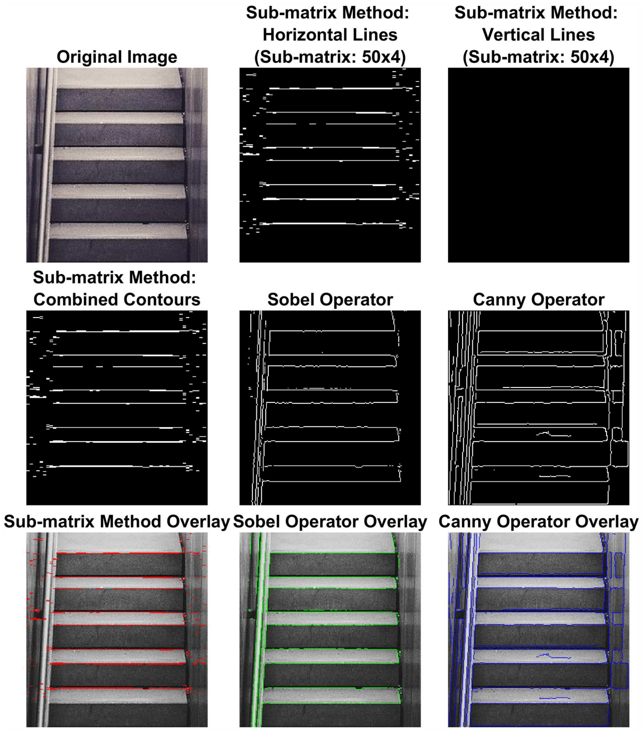
Figure 2
Enlarged image of the staircase contours shown in Figure 1



The authors did not aim to achieve the best quality of stair contour determination, but this can be done by choosing the sizes of the submatrices and thresholds. The presence of clear horizontal contour lines is the basis for identifying stairs, and vertical contours detect their boundaries, which is important for BVIPs.

The stairs shown in Figure 1, at first glance, can be attributed to an image with a simple shape, but it should be noted that the pixel intensity gradients here are not large. Figure 3 shows the results of determining the stairs' contours with different submatrix sizes than in Figure 1, as well as examples of determining the contours using the Sobel and Canny operators.

Figure 3
Results of determining the contours of stairs using the submatrix, Sobel, and Canny methods



It should be noted that the improvement of contours can be achieved by adjusting the thresholds for all three methods or by using the combined submatrix method. Therefore, the comparison should be made with the best-chosen parameters of the models underlying these methods. That is why Figure 3 is illustrative. For other stairs (Figure 4), the submatrix method (with a size of 200 × 5) and the Sobel and Canny methods give similar results in the quality of contour detection. It should be emphasized that the submatrix method was used to construct Figure 4. The combined submatrix method, as a rule, significantly improves the result, as can be seen from Figure 1.

The submatrix method (unlike the combined submatrix method) has the disadvantage of not being able to determine the contours of objects of complex shape. This disadvantage turns into an advantage if the contours are simple, such as stairs. It follows from the figure that other methods determine the contours not only of stairs but also of other image elements, which are not always needed by visually impaired people: for them, it is important in many cases to clearly determine only the boundaries of the stairs. Thus, the submatrix method for images has filtering properties in spatial domains and simplifies human decision-making based on the determined contours of objects.

For a complex image, the results of determining the contours of tree trunks in the forest using the Sobel and Canny methods (with different settings) and combined submatrices are shown in Figure 5.

Figure 5 shows that all three methods are able to detect tree trunk contours in a forest, but the combined submatrix method performs this operation visually better than the other methods. In this case, the final image with the detected tree contours is clearer than the original. This is due to the fact that the submatrix method detects vertically and horizontally oriented objects well. If the original image does not have clear object contours, then none of the three methods is effective (Figure 6).

Figure 6(c) is obtained with fixed submatrix sizes, although for other sizes, it is difficult to achieve clear contours of the stairs under the specified conditions. However, for many cases, it is important to optimize the submatrix sizes and the threshold values. The next section presents a method that allows to obtain optimal submatrix sizes and thresholds.

5. Optimization of Parameters of the Combined Submatrix Method

The simplest approach to optimizing the sizes of submatrices and thresholds is to use a brute force of these sizes and visually determine the best quality of the contours. Those sizes that provide the best quality of the contours are considered rational. This approach requires a lot of time, and the results are, in many cases, unreliable. For navigation tasks of BVIPs on the ground, the method is unacceptable. The quality of the contour determination results cannot be assessed by people's preferences. Therefore,

Figure 4
Results of determining stairs contours using submatrix, Sobel, and Canny methods

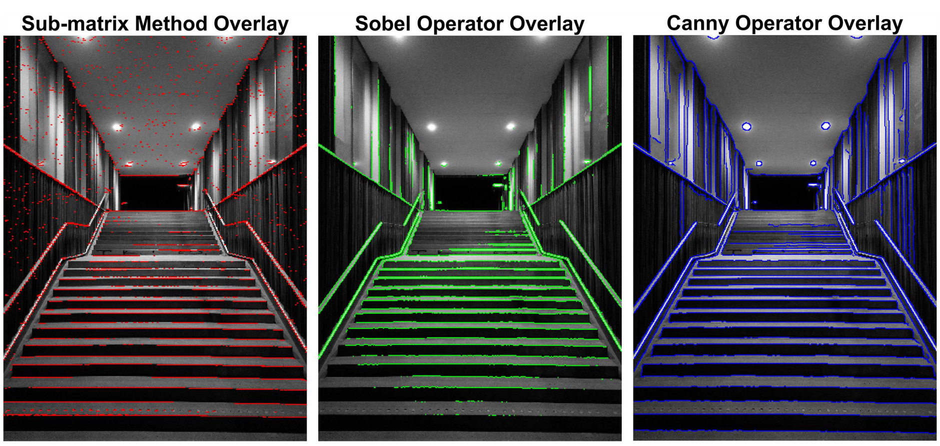


Figure 5
Results of determining tree trunk contours in the forest using the (a) Sobel and Canny methods, (b) Canny method at different thresholds, and (c) combined submatrix method

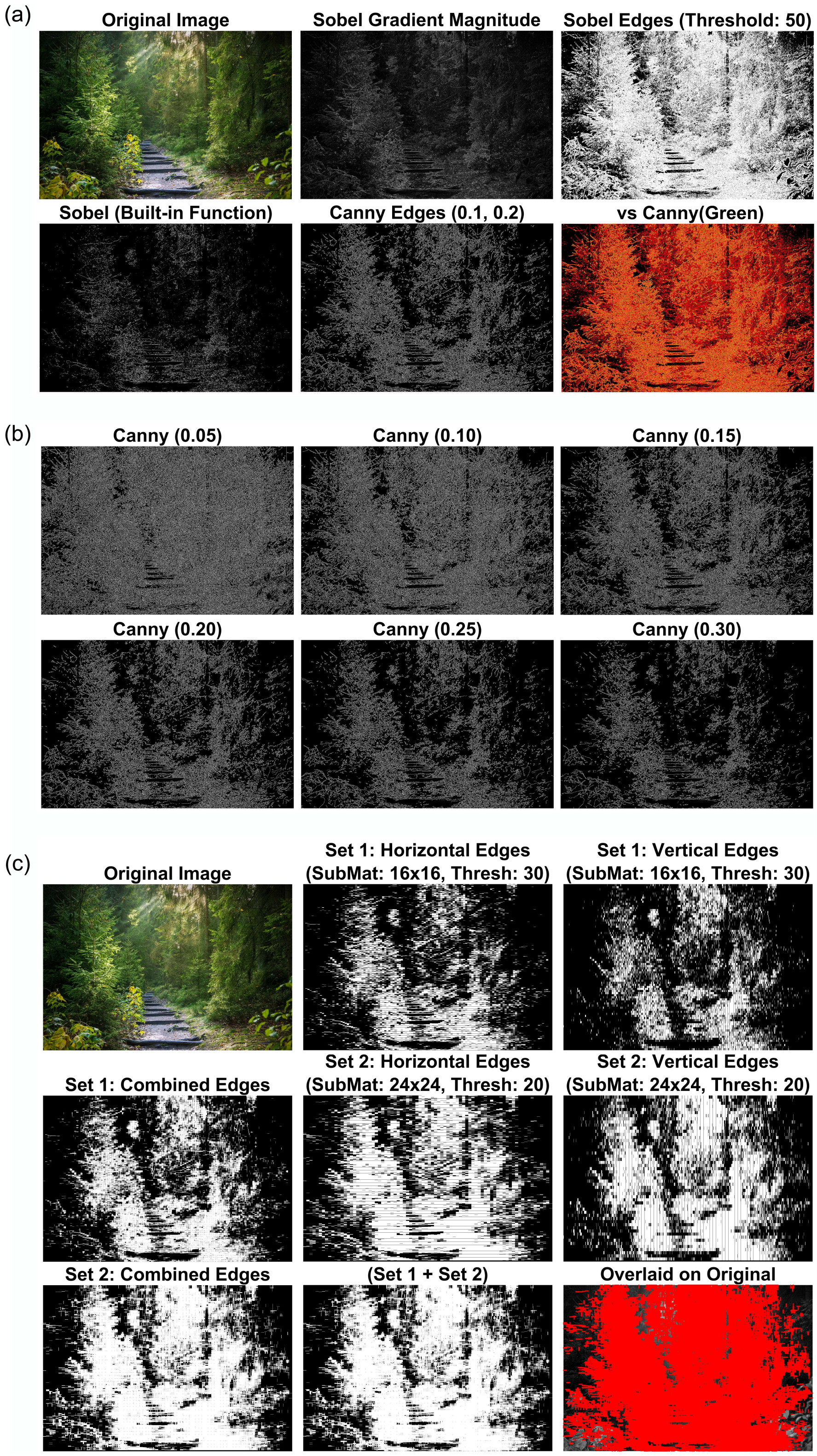
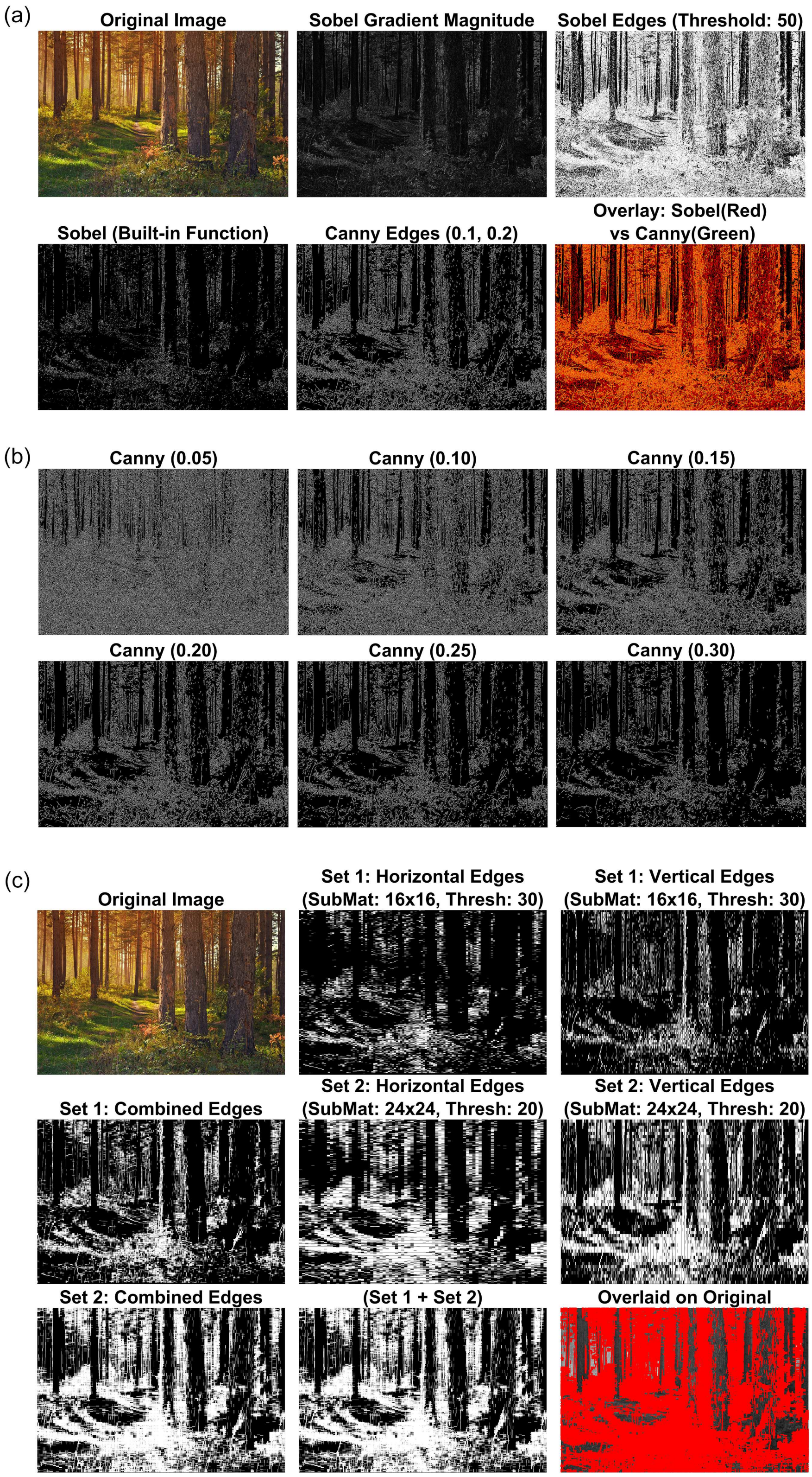


Figure 6

Results of determining the contours of the stairs in the forest using the (a) Sobel and Canny methods, (b) Canny method at different thresholds, and (c) combined submatrix method

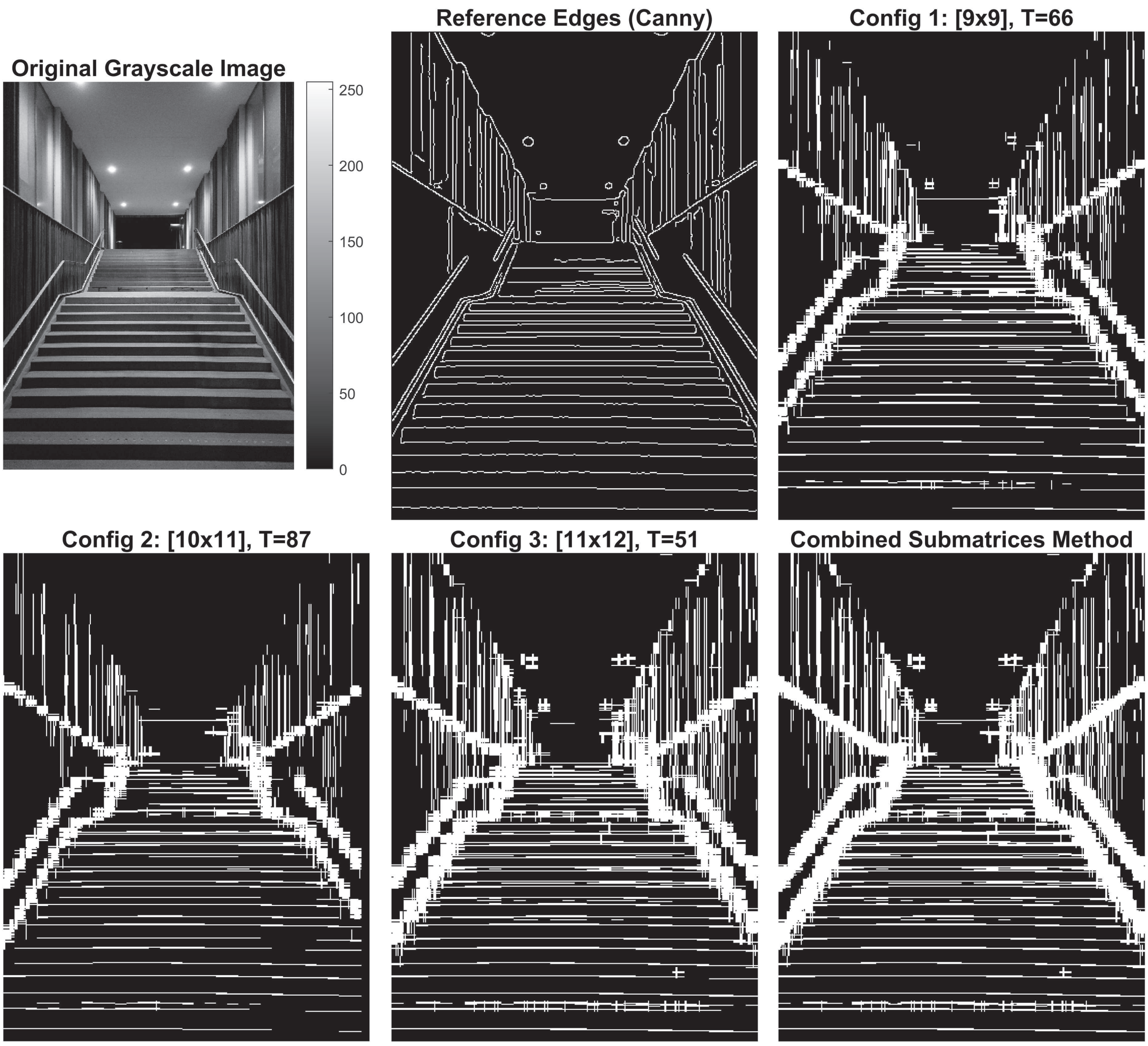


it is necessary to find a criterion for comparing the obtained results. The pixel intensity matrix with determined contours must be compared with another matrix that has the same dimensions. The matrix chosen in the paper is the pixel intensity matrix of the original image. The criterion for the similarity of matrices is the cosine of the angle between them. If the cosine is equal to one, then the matrices are identical, and if it is equal to zero, then the matrices are orthogonal. From the figures given above, for example, Figure 5, it follows that there is actually a similarity between the matrix with the given contours and the original matrix, but it should be expected that the cosine of the angle will be much less than unity. The cosine value is an integral indicator that requires the similarity of the entire matrix. In some cases, for example, for the conditions of Figure 6, it is advisable to evaluate the similarity not of matrices but of submatrices that cover only the stairs in the forest.

Therefore, optimization of the sizes of submatrices and thresholds should be carried out according to the criterion of the maximum cosine of the angle between the pixel intensity matrix with determined contours and the same matrix of the original image. After choosing the optimization criterion, it is necessary to determine the ranges in which the sizes of submatrices and thresholds should change. At the first stage, these ranges can be large, and then, as experience is gained, they can narrow. Since there are many variables in such a problem, it is advisable to use stochastic optimization methods to maximize the cosine of the angle. The paper uses a genetic algorithm. The operation of setting the parameters of the genetic algorithm is important. These parameters affect the speed of its operation. The faster the optimization is carried out, the greater the probability of determining the local maximum of the cosine of the angle between the matrices and not the global one. Memorizing the values of local maxima in the process of searching for the global maximum reduces the probability of incorrect optimization. For this, the level of mutations in the genetic algorithm increases, but the search time also increases.

In the MATLAB program code, the population size, number of generations, mutation and crossover probability, number of elite individuals, number of different submatrix configurations to optimize, minimum and maximum for submatrix dimensions, and bounds for threshold values are specified. Next, the test image

Figure 7
The original image of the stairs in the house, the reference image of the contours, and the image of the stairs' contours obtained by the genetic algorithm with the optimal sizes of the submatrices



is loaded, converted to grayscale, and normalized in the range [0, 255]. For comparison, the edge-detected version of the test image using operator Canny is created. The program simultaneously uses 3–4 different submatrix sizes. The performance of the model for determining the optimal sizes of submatrices and thresholds is evaluated using indicators widely used in machine learning: precision, recall, F1-score, and accuracy. The genetic algorithm maximizes the Fitness Function, that is, the sum of the cosine of the angle between the matrices described earlier and the F1-score. The cosine of the angle is 70% of the sum, and the F1-score is 30%. The maximum value of the Fitness Function is called the Best Fitness Value. The main operators of the algorithm are tournament selection, two-point crossover, and adaptive mutation. The algorithm usually requires 30–50 generations to converge.

Figure 7 (top left) shows the original image of the stairs in the house (same as in Figure 4). The contours of the stairs obtained by the Canny method are taken as a reference image (to the right of the original image). The other images show the contours of the stairs obtained by the genetic algorithm with optimal submatrix sizes: three configurations with different submatrix sizes and threshold values. The bottom right figure describes the best result provided by the combined submatrix method.

From the comparison of the best image of the staircase contours (Figure 7) and the same contours with suboptimal submatrices (Figure 4), it follows that the optimization brings the staircase contours closer to the original image. This is due to the choice of the optimization criterion. On the other hand, small details of the contours are not always needed for BVIPs moving inside the house. The lower central figure demonstrates that in some cases, even optimal submatrix sizes can lead to unsatisfactory contour images. In this case, the reason for such a result is a high threshold. The presence of local extrema of the Fitness Function can lead to similar contours in the images.

Figure 8 shows the dynamics of convergence of the objective function, when the difference between the Best Fitness Value and Average Fitness decreases with the time of operation of the genetic algorithm. On the right, Figure 8 shows the result of comparing the optimal staircase contours with the reference one, which indicates the qualitative determination of contours by the proposed method.

The result of superimposing the detected step contours on the original image is shown in Figure 9 (top left), and the edge intensity map is shown on the top right. It can be seen that the quality of stairs detection is high.

Figure 9 (bottom left) illustrates examples of optimal submatrices and threshold sizes for the case under consideration. If the characteristics of the first two submatrices and thresholds are close, then in the third configuration, the differences are significant, and the result of determining the contours of the stairs, as shown earlier, is unsatisfactory. The results of evaluating the system performance (precision, recall, F1-score, and accuracy) are shown in Figure 9 (bottom right). Only recall reaches high

Figure 8

The dynamics of convergence of the genetic algorithm (left) and the result of comparing the contours of the stairs in the images (right): the reference one (red) and the one obtained by the combined submatrix method (green)

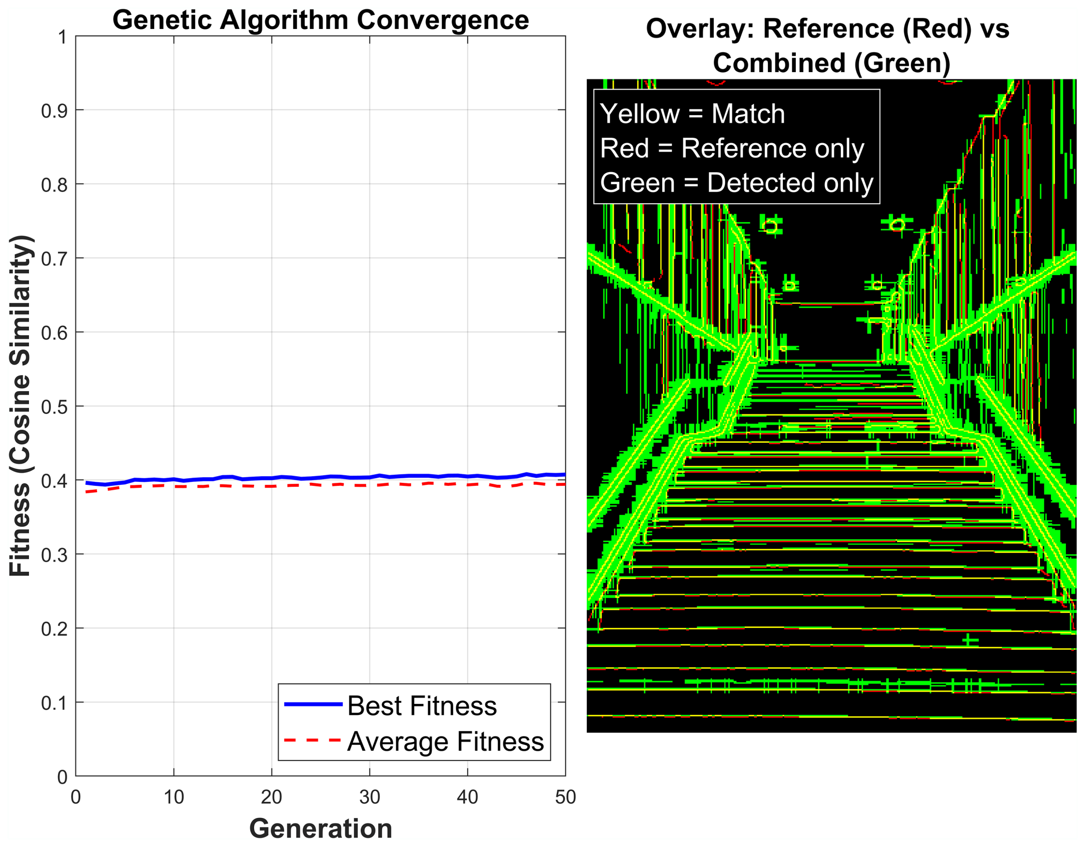
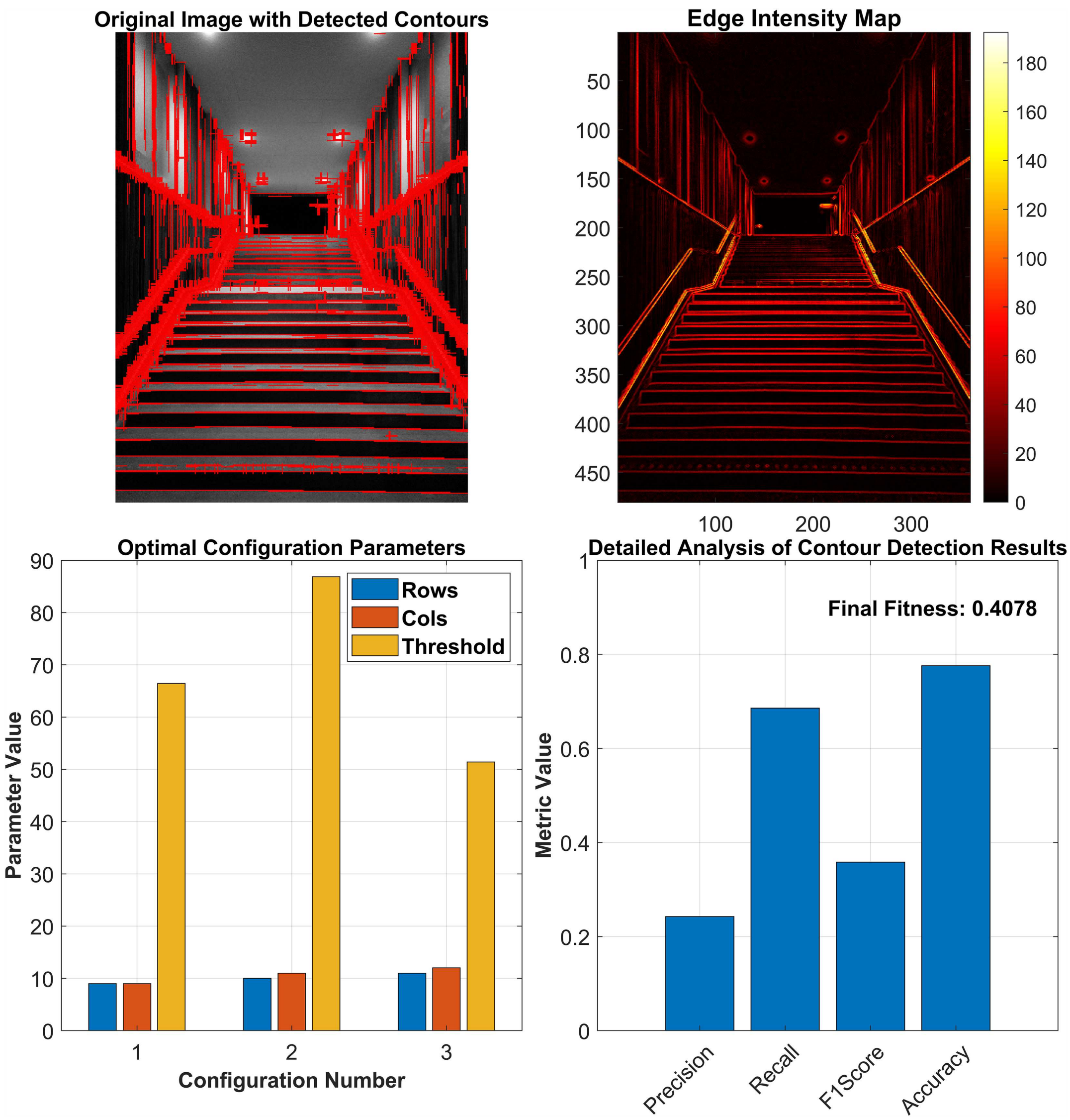


Figure 9
Original image with detected contours, edge intensity map (top), optimal parameters of the three submatrices, metric value of precision, recall, F1-score, and accuracy (bottom)



values, and the most universal indicator, the F1-score, is small for this case. One of the reasons for this result is the fulfillment of the criterion of similarity of the original image with the contours. A significant disadvantage of the genetic algorithm is the long time required to determine the optimal submatrix sizes and thresholds. In this study, it ranges from several minutes to 1.5 h, depending on the complexity of the image and the parameters of the genetic algorithm. There is also a possibility that the objective function will fall into a local rather than a global maximum. Thus, it cannot be used in real time and is not included in the main model of the combined submatrix method. Another problem of the genetic algorithm is the choice of an objective function that is related to the original image. Other approaches to determining this function are needed.

Thus, the obtained results show that optimization of submatrix sizes does not always provide high-quality contour images. Such images, obtained using a genetic algorithm, can be used as a database for training a neural network, which will subsequently evaluate the possibility of making a decision regarding the navigation of BVIPs in the environment. However, based on this study, a simpler direction of work follows. Instead of optimal submatrix sizes and threshold values, rational values can be used, which are obtained on the basis of a preliminary analysis of images that are characteristic of the navigation conditions of a blind person. They are shown, for example, in Figures 1–6. It is advisable to carry out image analysis not only by the combined submatrix method but also by other methods described in sections 2 and 3 of the paper.

6. Accuracy of the Method

Analysis of the capabilities of the combined submatrix method should include the calculation of accuracy indicators. For this, it is appropriate to compare the obtained contours with the corresponding contours of objects located in the image; that is, it is necessary to have a map of real, accurate contours. At the first stage, the content of such a map can be used as a real image

Figure 10
Comparison results of cosine similarity with original image and Jaccard similarity with original edges for four contour detection methods

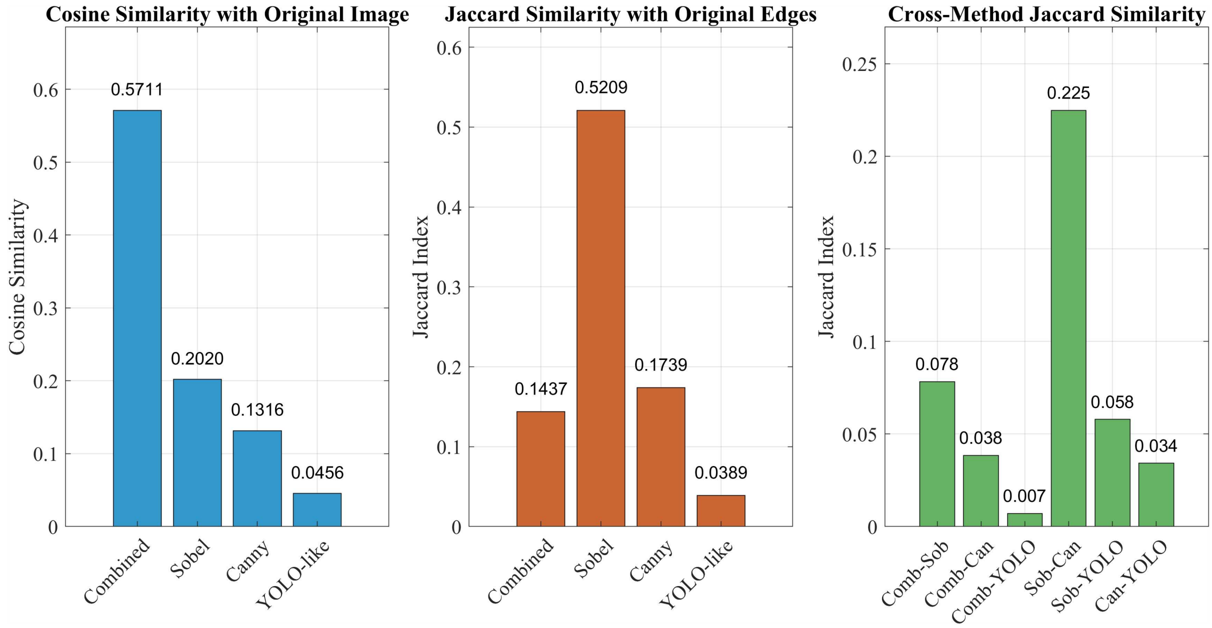
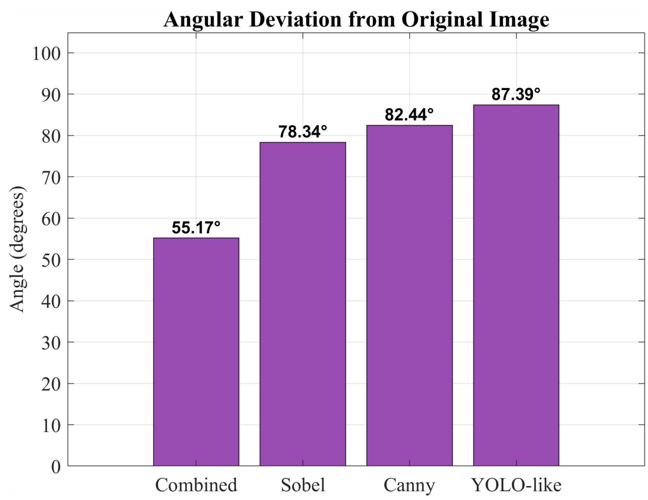


Figure 11
Results of angle comparison between the original image matrix and contour matrices obtained by four methods



(Figure 7) and compare the contours determined by the combined submatrix method, the Sobel, Canny, and YOLO-like methods. The similarity of the matrices of this image with the contour matrices determined by the four methods will be estimated using the cosine of the angle between the matrices and Jaccard similarity (Figure 10). The contours determined by the combined method are the closest to the contours of the original image by the cosine criterion between matrices compared to the other three methods. The Jaccard similarity of this method is lower than that of the Sobel method, and the cross Jaccard similarity of the combined method with the Sobel method is the highest compared to the other methods. Therefore, the combined method may be better than the other methods in some indicators. The numerical values of the angles between the matrices are shown in Figure 11.

From Figures 10 and 11, it follows that the most similar matrices of contours to the original image are given by the combined method of submatrices (the smaller the angle, the greater the similarity). Complete similarity is ensured when the angle is zero. For this example, the angle is significantly different from zero, since in the original image, in addition to the contours of the steps, there are many elements that do not belong to them. This simple method does not fully appreciate the capabilities of the methods. First, cosine treats images as vectors and measures the angle between them. Edge maps are fuzzy, and therefore, cosine is insensitive to small spatial displacements. Second, it ignores the localization tolerance: a contour displaced by a few pixels is still correct. Third, cosine requires careful normalization for comparison, which is associated with scale and pixel intensity issues. This is why it is necessary to use other comparison criteria without involving the original image.

These include the generally accepted indicators: precision, recall, F1-score, and IoU (Intersection over Union). Instead of the original image, in the absence of an accurate contour map, it is advisable to take the contours of simple geometric shapes, in particular, a rectangle, a triangle, and a circle, the coordinates of each point of which are known in advance. The results of the analysis of the accuracy of determining the contours of such objects are shown in Figures 12, 13, and 14.

The performance metrics summary is given in Table 1.

Thus, the combined submatrix method has approximately the same accuracy as other methods used to detect the contours of objects such as steps. Its advantage is the ability to adjust the size of the submatrix and the threshold for better contour detection (Figure 12). If a blind or visually impaired person uses the same route, then high contour detection accuracy can be achieved by pre-adjusting the size of the submatrices and the threshold. In other cases, these parameters can be selected based on experience. The study showed that it is not possible to unambiguously choose any one method for contour detection in all environmental conditions, which is typical for navigation by people with poor vision. The widespread use of neural networks and a combination of

Figure 12
Results of comparison of object contour detection methods

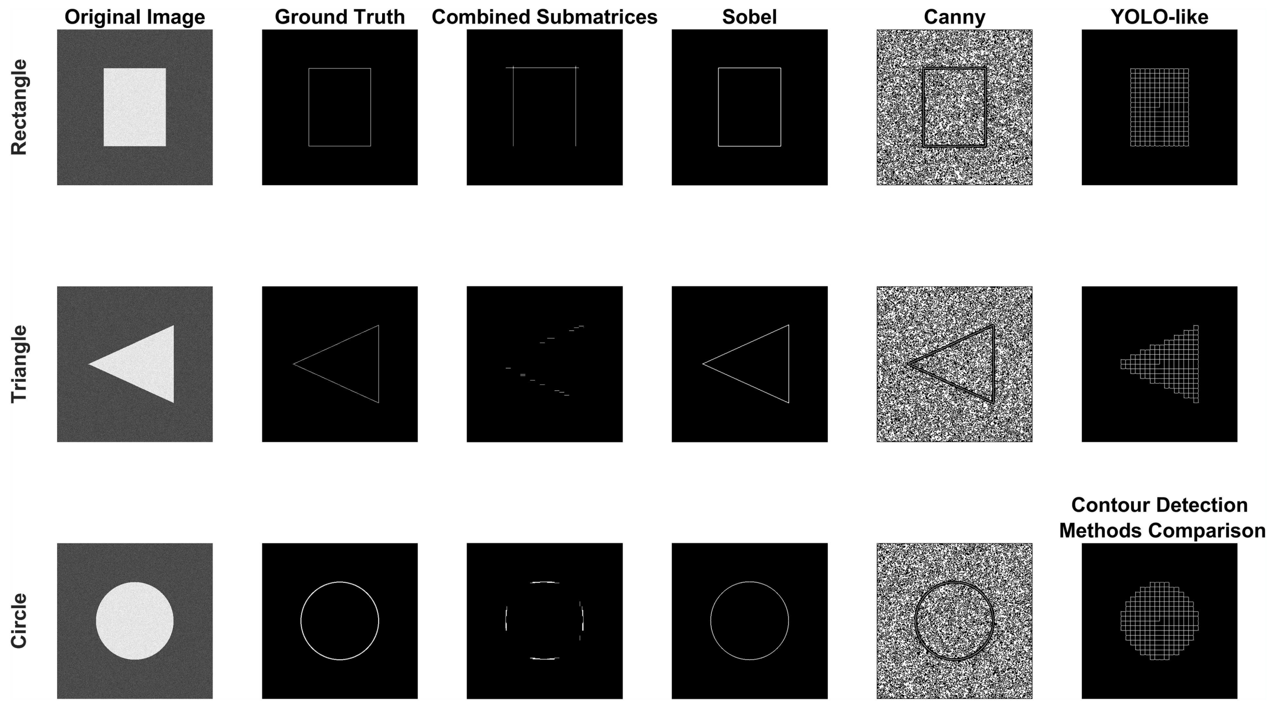


Figure 13
Precision, recall, F1-score, and performance metrics comparison, characterizing the effectiveness of the four methods in detecting the contours of a rectangle, triangle, and circle

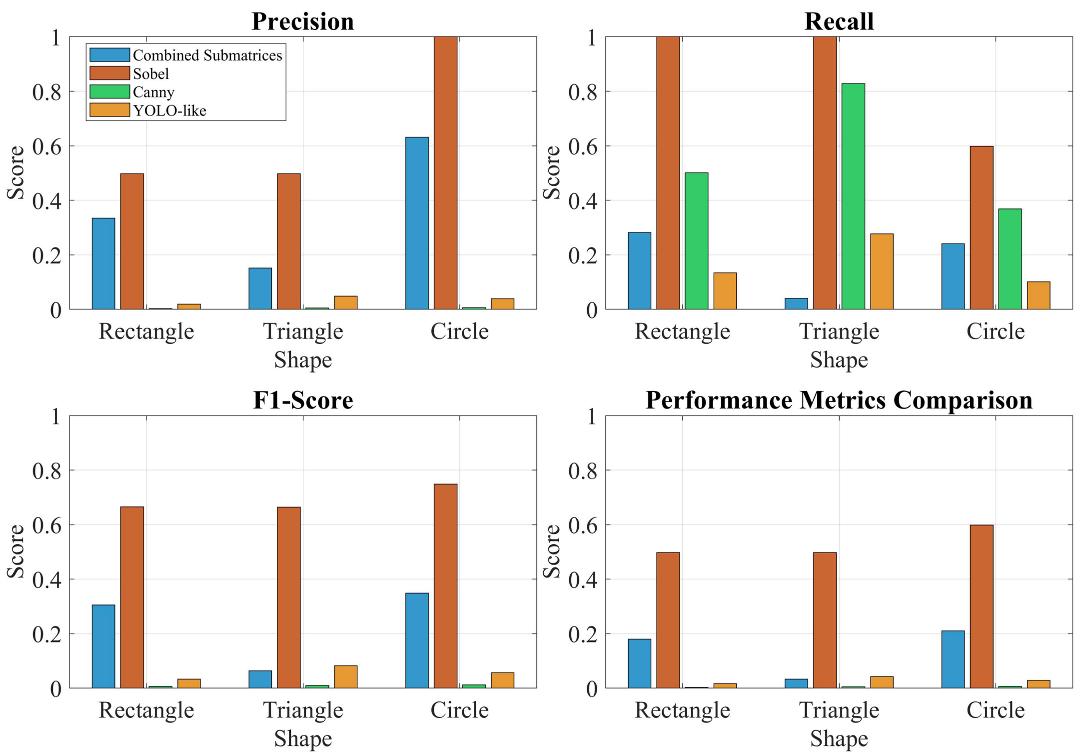


Figure 14
Results of robustness assessment of four methods (sensitivity to noise) using F1-score and IoU

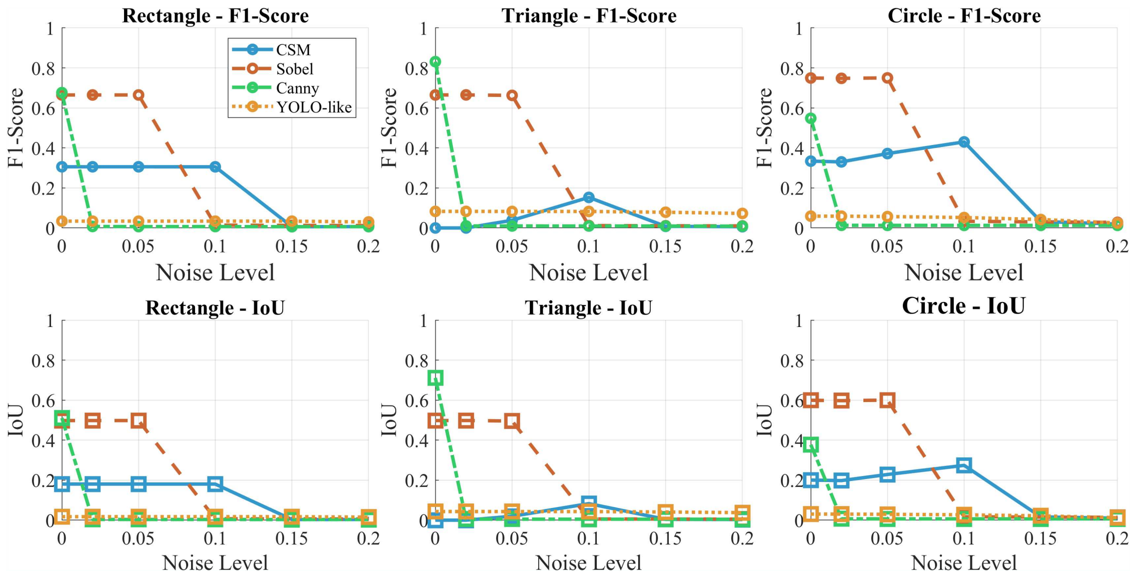


Table 1
The performance metrics summary

Method	Precision	Recall	F1-score	IoU	Time (s)
Rectangle					
CSM	0.3346	0.2815	0.3058	0.1805	0.1416
Sobel	0.4978	1.0000	0.6647	0.4978	0.1278
Canny	0.0034	0.5011	0.0067	0.0034	0.2257
YOLO	0.0195	0.1337	0.0340	0.0173	0.0878
Triangle					
CSM	0.1521	0.0403	0.0637	0.0329	0.0814
Sobel	0.4976	1.0000	0.6645	0.4976	0.0223
Canny	0.0049	0.8278	0.0098	0.0049	0.0765
YOLO	0.0487	0.2772	0.0828	0.0432	0.0568
Circle					
CSM	0.6308	0.2406	0.3483	0.2109	0.0927
Sobel	1.0000	0.5985	0.7488	0.5985	0.0299
Canny	0.0065	0.3686	0.0128	0.0065	0.1368
YOLO	0.0395	0.1018	0.0569	0.0293	0.0308

different methods can improve the results, but this requires further research.

7. Conclusions and Prospects for Future Research

Bringing the living conditions of BVIPs closer to the level of ordinary people is an important problem of modern society. One approach to solving this problem is to use technical means, in particular video cameras, to guide people in areas with many obstacles. There are many methods of finding obstacles based on determining their contours. The paper proposes another one, which is called the method of combined submatrices. It is based on dividing the pixel intensity matrix of the source image into submatrices, each of which determines the Euclidean distance between adjacent rows and columns. For those row and column coordinates where a pre-selected threshold of such distance is exceeded, a horizontal or vertical line is drawn, depicting part of the contour of some object on the video camera matrix. This is part of the combined submatrix method, which is simply called the submatrix method. It well detects the contours of the horizontal and vertical elements of the entire object. The combined submatrix method performs the same operations, but in this case, the entire matrix is divided into submatrices of two different sizes. The contours of obstacles are determined for these submatrices and then superimposed on each other. A visual comparison of the effectiveness of the proposed method with known methods based on the use of YOLO, Sobel and Canny operators is carried out. All three methods are effective when there are large gradients of pixel intensity in adjacent rows and columns of the matrix.

For blurred images, the problem of qualitatively determining contours is not always solved. It is shown that in some situations, the method of combined submatrices is more effective than others. By optimizing the sizes of submatrices using a genetic algorithm, better results can be achieved. The optimization criterion is the cosine of the angle between the matrix with the determined contours and the matrix of the original image. Since the operating time of the genetic algorithm varies widely from minutes to hours (depending on the parameters of the genetic algorithm, image, etc.), the optimization cannot be carried out in real time. Predetermined contours of many objects on the matrices of video cameras using submatrices of optimized sizes and thresholds can be included in the database with training samples for neural networks. It is promising to estimate the probability of correctly assigning contours to a given shape of an object in the interests of obtaining information for BVIPs.

The main directions of future research on the topic of the article may be, first, the widespread use of deep learning and artificial intelligence methods, in particular, the edge detector based on MobileNet, ConvNext-SE-Attn models, and models based on DarkNet [41–43]. MobileNet is excellent for detecting object contours in real time. ConvNeXt-SE-Attention provides improved localization and suppresses noise. DarkNet is integrated with real-time YOLO-type detectors. The second is the modernization of the method by optimizing the sizes of submatrices depending on the specific image. The main problem in this case is the choice of a reliable optimization criterion.

Recommendations

The conclusions showed that there are now many methods for determining the contours of objects on video camera matrices. The combined submatrix method proposed in the paper is also based on the use of pixel intensity gradients. It has advantages when the shape of objects consists mainly of horizontally and vertically oriented elements. This is important for navigation of BVIPs in an urban environment. For other practical needs, it is close in quality of results to other methods.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Author Contribution Statement

Oleksandr Poliarus: Conceptualization, Methodology, Validation, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Yevhen Poliakov:** Software, Validation, Formal analysis, Resources, Data curation, Writing – original draft, Visualization.

References

[1] Yu, R., Lee, S., Xie, J., Billah, S. M., & Carroll, J. M. (2024). Human–AI collaboration for remote sighted assistance: Perspectives from the LLM era. *Future Internet*, *16*(7), 254. https://doi.org/10.3390/fi16070254

[2] Chaudary, B., Pohjolainen, S., Aziz, S., Arhippainen, L., & Pulli, P. (2021). Teleguidance-based remote navigation assistance for visually impaired and blind people—Usability and user experience. *Virtual Reality*, *27*(1), 141–158. https://doi.org/10.1007/s10055-021-00536-z

[3] Abdel-Jaber, H., Albazar, H., Abdel-Wahab, A., Amir, M. E., Alqahtani, A., & Alobaid, M. (2021). Mobile-based IoT solution for helping visual impairment users. *Advances in Internet of Things*, *11*(4), 141–152. https://doi.org/10.4236/ait.2021.114010

[4] Khan, A., Khusro, S., & Alam, I. (2018). BlindSense: An accessibility-inclusive universal user interface for blind people. *Engineering, Technology & Applied Science Research*, *8*(2), 2775–2784. https://doi.org/10.48084/etasr.1895

[5] Okolo, G. I., Althobaiti, T., & Ramzan, N. (2025). Smart assistive navigation system for visually impaired people. *Journal of Disability Research*, *4*(1), 1–10. https://doi.org/10.57197/JDR-2024-0086

[6] Gao, Y., Wu, D., Song, J., Zhang, X., Hou, B., Liu, H., & Zhou, L. (2025). A wearable obstacle avoidance device for visually impaired individuals with cross-modal learning. *Nature Communications*, *16*(1), 2857. https://doi.org/10.1038/s41467-025-58085-x

[7] Ricci, F. S., Liguori, L., Palermo, E., Rizzo, J.-R., & Porfiri, M. (2024). Navigation training for persons with visual disability through multisensory assistive technology: Mixed methods experimental study. *JMIR Rehabilitation and Assistive Technologies*, *11*, e55776. https://doi.org/10.2196/55776

[8] Muhsin, Z. J., Qahwaji, R., Ghanchi, F., & Al-Taee, M. (2024). Review of substitutive assistive tools and technologies for people with visual impairments: Recent advancements and prospects. *Journal on Multimodal User Interfaces*, *18*(1), 135–156. https://doi.org/10.1007/s12193-023-00427-4

[9] Nikanfar, S., Hebri, A., Ram Nambiappan, H., Nale, G., Siddiqua, M., Farhanipad, F., & Makedon, F. (2025). A survey on assistive technologies for visually impaired individuals: Recent innovations, limitations, and future directions. In *Proceedings of the 18th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, 429–434. https://doi.org/10.1145/3733155.3734895

[10] Syed Ameer Abbas, S., Hareharan, M., & Sivakumar, I. (2024). *Integrated intelligent assistant for visually impaired*. Research Square. https://doi.org/10.21203/rs.3.rs-4283447/v1

[11] Irfan, S. M., Islam, M. M., Mia, M. S., Islam, M., Islam, S., & Islam, T. (2023). Advancements in object detection and tracking algorithms: An overview of recent progress. *EPRA International Journal of Research & Development*, *8*(9), 134–140. https://doi.org/10.36713/epra14418

[12] Wu, Q., Liang, T., Fang, H., Wei, Y., Wang, M., & He, D. (2023). A lightweight deep learning algorithm for multi-objective detection of recyclable domestic waste. *Environmental Engineering Science*, *40*(12), 667–677. https://doi.org/10.1089/ees.2023.0138

[13] Pham, V., Nguyen, D., & Donan, C. (2022). Road damage detection and classification with YOLOv7. In *2022 IEEE*

international conference on big data (Big Data), 6416–6423. https://doi.org/10.1109/BigData55660.2022.10020856.

[14] Nath, A., Patel, J. B., Patel, G. K., Patel, P. J., & Patel, D. S. (2025). Enhanced YOLOv8 for road object detection with advanced AI-based real-time precision algorithm. In M. S. Kaiser, J. Xie, & V. S. Rathore (Eds.), *Intelligent strategies for ICT: Proceedings of ICTCS, 5*, 47–59. Springer. https://doi.org/10.1007/978-981-96-4148-2_5

[15] Yu, X., & Saniie, J. (2025). Visual impairment spatial awareness system for indoor navigation and daily activities. *Journal of Imaging*, *11*(1), 9. https://doi.org/10.3390/jimaging11010009

[16] Miller, R. L., & Sheinberg, D. L. (2022). Evidence for independent processing of shape by vision and touch. *eNeuro*, *9*(3), 1–14. https://doi.org/10.1523/eneuro.0502-21.2022

[17] Quinn, R., Murtough, S., de Winton, H., Ellis-Frew, B., Zane, S., de Sousa, J., . . . , & Spiers, A. J. (2024). A shape-changing haptic navigation interface for vision impairment. *Scientific Reports*, *14*(1), 29223. https://doi.org/10.1038/s41598-024-79845-7

[18] Chow, J. K., Palmeri, T. J., & Gauthier, I. (2021). Haptic object recognition based on shape relates to visual object recognition ability. *Psychological Research*, *86*(4), 1262–1273. https://doi.org/10.1007/s00426-021-01560-z

[19] Gkanidi, M. I., & Drigas, A. S. (2021). Tactile maps and new technologies for blind and people with visual impairments. *International Journal of Management and Humanities*, *5*(8), 1–5. https://doi.org/10.35940/ijmh.e1208.045821

[20] Ye, C., & Qian, X. (2018). 3-D object recognition of a robotic navigation aid for the visually impaired. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(2), 441–450. https://doi.org/10.1109/TNSRE.2017.2748419

[21] Szeliski, R. (2022). *Computer vision: Algorithms and applications* (2nd ed.). Switzerland: Springer.

[22] Bishop, C. M. (2006). *Pattern recognition and machine learning*. USA: Springer.

[23] Sia, J. S. Y., Tan, T. S., Yahya, A. B., Tiong, M. F. T., & Sia, J. Y. X. (2021). Mini Kirsch edge detection and its sharpening effect. *Indonesian Journal of Electrical Engineering and Informatics*, *9*(1), 228–244. https://doi.org/10.52549/ijeei.v9i1.2597

[24] Bello, R., Mohamed, A. S. A., & Talib, A. Z. (2021). Contour extraction of individual cattle from an image using enhanced Mask R-CNN instance segmentation method. *IEEE Access*, *9*, 56984–57000. https://doi.org/10.1109/ACCESS.2021.3072636

[25] Al-Asadi, T. A., & Witefee, D. M. (2020). Geometrical fusion based on chain code representation. *IOP Conference Series: Materials Science and Engineering*, *928*(3), 032033. https://doi.org/10.1088/1757-899X/928/3/032033

[26] Hemalatha, R. J., Thamizhvani, T. R., Dhivya, A. J. A., Joseph, J. E., Babu, B., & Chandrasekaran, R. (2018). Active contour based segmentation techniques for medical image analysis. In R. Koprowski (Ed.), *Medical and biological image analysis* (pp. 17–36). InTech. https://doi.org/10.5772/intechopen.74576

[27] Hemalatha, R. J., Vijaybaskar, V., & Thamizhvani, T. R. (2018). Performance evaluation of contour based segmentation methods for ultrasound images. *Advances in Multimedia*, *2018*(1), 4976372. https://doi.org/10.1155/2018/4976372

[28] Talebi, M., Ayatollahi, A., & Kermani, A. (2011). Medical ultrasound image segmentation using genetic active contour. *Journal of Biomedical Science and Engineering*, *4*(2), 105–109. https://doi.org/10.4236/jbise.2011.42015

[29] Iqbal, E., Niaz, A., Memon, A. A., Asim, U., & Choi, K. N. (2020). Saliency-driven active contour model for image segmentation. *IEEE Access*, *8*, 208978–208991. https://doi.org/10.1109/ACCESS.2020.3038945

[30] Fang, J., Liu, H., Liu, H., Zhang, L., & Liu, J. (2016). Localized patch-based fuzzy active contours for image segmentation. *Computational and Mathematical Methods in Medicine*, *2016*(1), 1064692. https://doi.org/10.1155/2016/1064692

[31] Shyu, K.-K., Pham, V.-T., Tran, T.-T., & Lee, P.-L. (2012). Global and local fuzzy energy-based active contours for image segmentation. *Nonlinear Dynamics*, *67*(2), 1559–1578. https://doi.org/10.1007/s11071-011-0088-1

[32] Niu, S., Chen, Q., Sisternes, L., Ji, Z., Zhou, Z., & Rubin, D. L. (2017). Robust noise region-based active contour model via local similarity factor for image segmentation. *Pattern Recognition*, *61*, 104–119. https://doi.org/10.1016/j.patcog.2016.07.022

[33] Afifi, A., Nakaguchi, T., Tsumura, N., & Miyake, Y. (2010). A model optimization approach to the automatic segmentation of medical images. *IEICE Transactions on Information and Systems*, *E93-D*(4), 882–890. https://doi.org/10.1587/transinf.e93.d.882

[34] Zhao, M. R., & Zhang, X. W. (2012). Based on the topography image segmentation level set method. *Applied Mechanics and Materials*, *236–237*, 404–408. https://doi.org/10.4028/www.scientific.net/amm.236-237.404

[35] Park, J., Park, S., & Cho, W. (2012). Medical image segmentation using level set method with a new hybrid speed function based on boundary and region segmentation. *IEICE Transactions on Information and Systems*, *E95.D*(8), 2133–2141. https://doi.org/10.1587/transinf.e95.d.2133

[36] Zhu, J., Liu, Y., Zhang, J., Wang, Y., & Chen, L. (2019). Preliminary clinical study of the differences between interobserver evaluation and deep convolutional neural network-based segmentation of multiple organs at risk in CT images of lung cancer. *Frontiers in Oncology*, *9*, 627. https://doi.org/10.3389/fonc.2019.00627

[37] Lustberg, T., Soest, J., Gooding, M., Peressutti, D., Aljabar, P., Stoep, J., . . . , & Dekker, A. (2018). Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer. *Radiotherapy and Oncology*, *126*(2), 312–317. https://doi.org/10.1016/j.radonc.2017.11.012

[38] Wooten, Z. T., Yu, C., Court, L. E., & Peterson, C. B. (2022). Predictive modeling using shape statistics for interpretable and robust quality assurance of automated contours in radiation treatment planning. *Pacific Symposium on Biocomputing 2023*, *28*, 395–406. https://doi.org/10.1142/9789811270611_0036

[39] Khan, A., Iskander, D. N. B. A. I., Chai, W. Y. C., Lim, P., Ullah, F., Ullah, J., & Ahmad, T. (2022). Segmentation model approaches using cardiac magnetic resonance images: A review. *Research Square*.https://doi.org/10.21203/rs.3.rs-2368607/v1

[40] Poliarus, O., & Poliakov, Y. (2024). Methodology of assigning terrain object images to the class of landmarks for autonomous mobile robots. In J. C. Rodríguez-Quiñonez, W. Flores-Fuentes, M. J. Castro-Toscano, & O. Sergiyenko (Eds.), *Scanning technologies for autonomous systems* (pp. 3–32). Springer. https://doi.org/10.1007/978-3-031-59531-8_1

[41] El Habchi, A., Khiati, W., Zerrouk, I., Aissi, M., Berrich, J., & Bouchentouf, T. (2025). Fast drone obstacle detection

approach based on MobileNet classification CNN architectures. In *Communication and Information Technologies through the Lens of Innovation: Proceedings of the 5th International Conference on Advanced Technologies and Humanity*, 67–74. https://doi.org/10.1007/978-3-031-74470-9_9

[42] Nath, A., Roy, O., Silveri, P., & Patel, S. (2025). Deep learning approach with ConvNeXt-SE-attn model for in vitro oral squamous cell carcinoma and chemotherapy analysis. *MethodsX*, *15*, 103519. https://doi.org/10.1016/j.mex.2025.103519

[43] Yang, L., Chen, G., & Ci, W. (2023). Multiclass objects detection algorithm using DarkNet-53 and DenseNet for intelligent vehicles. *EURASIP Journal on Advances in Signal Processing*, *2023*(1), 85. https://doi.org/10.1186/s13634-023-01045-8

How to Cite: Poliarus, O., & Poliakov, Y. (2026). A Method for Determining Obstacle Contours Using Wearable Video Cameras for Orienting Blind and Visually Impaired People in Urban Environment. *Smart Wearable Technology*. https://doi.org/10.47852/bonviewSWT62028788