Detection of Free Obstacle Region Using Distance Transform and Image Subtraction Method for Monocular Camera Sensor

Muhammad Syazani Nabil Sobri[1], Muhammad Faiz Ramli[2]*, Muhamad Wafi Abdul Aziz[2]

  [1] *Faculty of Mechanical and Manufacturing Engineering (FKMP),*
     Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Johor, MALAYSIA

  [2] *Research Center for Unmanned Vehicle (RECUV), Faculty of Mechanical and Manufacturing Engineering,*
     (FKMP), Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Johor, MALAYSIA

*Corresponding Author: faizr@uthm.edu.my
DOI: https://doi.org/10.30880/paat.2025.05.01.001

Abstract

An area that is free of any barriers is known as a free region. Prior studies have demonstrated that when textureless impediments are used, the pixel volume expansion approach has limitations in identifying free spaces. To address the shortcoming, an approach that combines distance transform and image subtraction techniques is suggested. This algorithm incorporates three primary image processing processes: distance transform, image subtraction, and k-means segmentation. An object that is 170 cm and another that is 200 cm from the camera serve as the algorithm's inputs. Following image processing, the output is separated into 12 sections, one of which is designated as a free region. Two distinct scenarios—a congested atmosphere and an uncluttered environment—were tested by the program. Four different types of obstacles are tested for each scenario. In this case, there are four types of obstacles: texture, textureless, textureless with one obstacle on the left, and textureless with one obstacle on the right. The outcome demonstrates that the technique can accurately identify free zones. Most importantly, it has a 100% success rate in detecting obstructions that lack texture. This demonstrates that the approach can get around the prior method's drawback. Additionally, each obstacle's success rate is compiled and shown in a table. Furthermore, a comparison is made between the proposed method and the pixel volume expansion method. Combining the image subtraction and distance transform method with the volume pixel expansion method yields the same success rate (100%) in texturing barriers for crowded situations.

1. Introduction

Identifying and locating areas in an environment where there are no objects or obstacles that could potentially obstruct or impede the passage of a vehicle is the process of detecting free obstacle regions. In the present day, numerous advancements have been made in the field of obstacle detection, including the investigation of free region detection methods that employ pixel volume expansion [1]. The pixel volume expansion method has been shown to have limitations in the detection of textureless objects in previous research [2][3]. A textureless object

is devoid of any insignia or patterns. The pixel volume expansion method was applied to textureless objects, but pixel expansion only occurred at the borders of the object, resulting in no pixel expansion at the center. This resulted in the free region detection algorithm incorrectly classifying the object's center as a free region.

Two methods were suggested to address this challenge: the Image Subtraction method and the Distance Transform method. The combination of these two methodologies was anticipated to be effective due to their complementary nature. The performance of the proposed combination was assessed in this study. Two research queries were developed in response to the aforementioned issue: Is it possible for a monocular camera sensor to detect free obstacle regions in real-world environments, including textureless obstacles, through the combination of Distance Transform and Image Subtraction? Additionally, is it possible for the combination of Distance Transform and Image Subtraction to detect free obstacle regions more accurately than the pixel volume expansion method? The objective of this research was to create an algorithm that detects free obstacle regions in a single image by utilizing the Distance Transform and Image Subtraction methods with a monocular camera sensor. Lastly, the proposed method was evaluated in terms of its ability to precisely identify free obstacle regions in comparison to the pixel volume expansion method.

2. Methodology

This paper presents several methods for detecting objects and free regions. The algorithm's flowchart is depicted in Fig. 1. The flowchart commenced with distance detection; the LIDAR sensor identifies an obstacle when the distance reaches 170 cm and 200 cm. Upon reaching 170 cm and 200 cm, the camera automatically captures an image. This indicates that two images are utilized as input prior to proceeding with image processing. One input image is captured at a distance of 170 cm from the item, while the other is at a distance of 200 cm from the object. Both images undergo a k-means segmentation procedure that categorizes the color into three distinct clusters. The image at a distance of 200 cm from the obstacle is combined with the image at 170 cm from the item to form a single image. Following that, the image's color is altered to black and white. Subsequently, the distance transform is executed on the image. To identify the free region, the image is divided into 12 sections. The quantity of black pixels for each area is determined. Among the 12 locations, the one exhibiting the greatest quantity of black pixels is designated as a free region. Finally, the unoccupied area is presented.
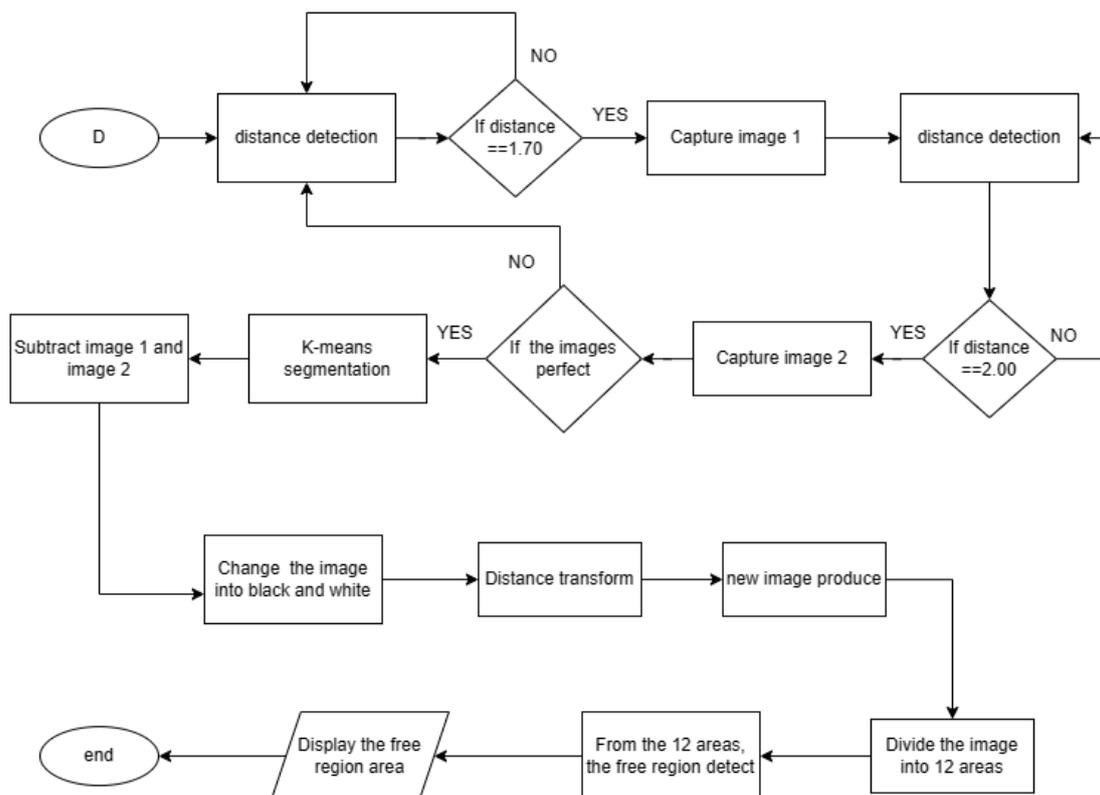

Fig. 1 *Figure shows a flowchart of the algorithm*

2.1 K-mean Segmentation

K-means segmentation is a technique that is classified as image segmentation. The technique of image segmentation is frequently employed in decision-oriented applications to accurately classify the pixels of an image [4][5]. It is a critical instrument that is employed in a variety of sectors, including traffic analysis, image processing, health care, and pattern identification [6][7]. Clustering is accomplished through the utilization of K-means. Clustering algorithms are a type of unsupervised algorithm that denotes the absence of labeled data. It is employed to classify distinct classes or clusters within the supplied data, contingent upon the degree of similarity between the data points.

It commenced with the initialization of the number of clusters (K), which is 3. The k centroid is determined by the K-means algorithm, which subsequently allocates each data point to the cluster with the closest centroid. The Euclidean distance is employed to ascertain the distance to the nearest centroid. Upon the completion of the classification process, the algorithm recalculates the new centroid for each cluster and, using this centroid, calculates a new Euclidean distance between each centroid and each data point. The cluster points with the shortest Euclidean distance are awarded. The centroid and member objects of each partition define its cluster. The centroid of a cluster is the point at which the total distance from all objects is the minimum.

2.2 Distance Transform

Distance transformation is a technique in image processing that has a wide range of applications [8]. It is primarily employed in binary images to determine the distance between the adjacent object and the background points (or vice versa). The distance transform is a technique that has typically been employed to determine the distance between the highlighted pixel and the non-highlighted pixel. It is a method that converts a binary image that includes both feature and non-feature elements into a picture in which each feature element is assigned a value that approximates its distance from the next closest feature element.

The initial step of this approach was to input the images. The initial step in employing distance transform is to convert the image to grayscale. This is due to the fact that the pixel is limited to 1 and 0 in order to convert the image to a binary format. Subsequently, the image is binarized by employing a threshold. Then, the image undergoes a distance transform, and the distance is subsequently normalized within the [0, 255] range.

2.3 Image Subtraction

The technique of image subtraction involves the use of two images as input. The first image subtracts the second image, resulting in the third image, which is a residual pixel after the pixel on the first image subtracts the pixel on the second image. There are numerous applications for this procedure. One of them is that this method can detect discrepancies between two images and rectify errors in an image, such as when one-half of it is obscured by darkness. The process entails subtracting a specific value from each pixel and utilizing a unique image as input [1]. The purpose of image subtraction in this investigation is to quantify the pixel expansion by contrasting the increase in pixels surrounding an obstacle between the distances of the first and second frames.

$$IF_{\text{subtract}}(x, y) = IF_1(x, y) - IF_2(x, y) \tag{1}$$

An image frame is defined as IF, and (x, y) represents the coordinates of the pixels where x represents the horizontal coordinate of the pixel while y is the vertical coordinate of the pixel from which the output pixel values are derived.

2.4 Applying Grid

After the image is subtracted, it is divided into 12 areas, as illustrated in Fig. 2. The image that has been subtracted has dimensions of 480x640 in both width and height. The objective is to partition the image into open regions and obstacle regions. The obstacle is situated in the center of the image. Obstacle regions are defined as regions 6, 7, 10, and 11, while unencumbered regions are defined as regions 1, 2, 3, 4, 5, 8, 9, and 12.

1	2	3	4
5	6	7	8
9	10	11	12

Fig. 2 *Figure shows the grid image*

2.5 Pixel Volume Expansion

Two primary stages comprise the pixel volume expansion method: image subtraction and k-means segmentation. This method is intended to identify obstacles or changes in the environment by examining the differences in pixel information between two images that were captured at varying distances. The procedure commences with the acquisition of two sequential images of the same scene, one at a distance of 170 cm and the other at 200 cm from the camera sensor. Subsequently, these images are employed as inputs for additional processing.

K-means segmentation, a clustering algorithm that organizes pixels into distinct classes based on their color similarity or intensity, is applied to both images in the initial stage. By reducing the variability in pixel values, this segmentation simplifies the scene and emphasizes critical regions, thereby improving the visibility of object boundaries. The second stage of the procedure, image subtraction, is performed on the images after the segmentation is complete. In this stage, the image captured at 200 cm is subtracted pixel-by-pixel from the image captured at 170 cm. This operation underscores the distinctions or modifications that exist between the two scenarios. It specifically emphasizes the regions where there are substantial pixel intensity differences, which are typically associated with objects or obstacles that appear more defined or larger in the closer image (170 cm).

The pixel expansion, a phenomenon in which objects that are closer to the viewer occupy more pixels than those that are farther away, is visually represented in a new image as a result of this subtraction. This pixel expansion is a critical indicator of the presence or movement of an object. The algorithm can identify and isolate potential obstacles in the environment by analyzing these expanded pixel regions. In general, the pixel volume expansion method combines techniques of differential image analysis and unsupervised segmentation to identify obstacles by utilizing the relationship between the scale of objects in the image and the distance of the camera.

3. System Configuration

The hardware utilized in this research is the Raspberry Pi 4 Model B, as depicted in Fig. 3(a). The Raspberry Pi is a single-board computer that has achieved significant appeal for educational applications in programming and computer science. The purpose of the Raspberry Pi is to implement the method for free region detection. Figure 3(b) illustrates the SeeedStudio Grove-TF Mini LiDAR Time of Flight (ToF) Distance Sensor employed in the project. The LiDAR measures the distance from the camera to the item at a predetermined range, at which the camera captures an image. Furthermore, Figure 3(c) illustrates the Raspberry Pi Camera Rev. 1.3. The camera captures photos at a resolution of 480 x 640.
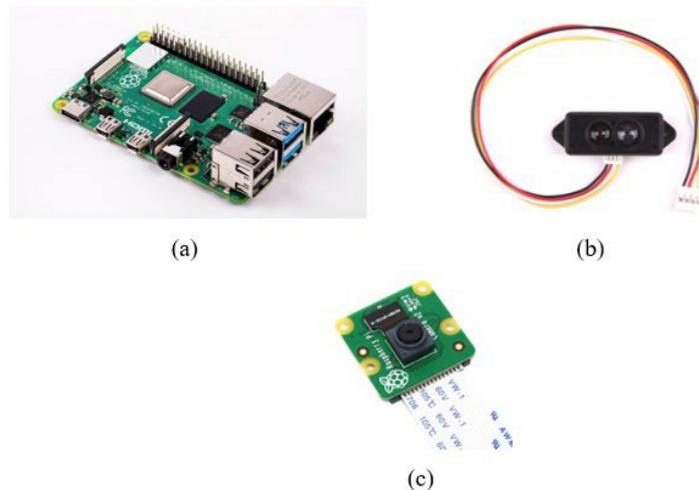

(a)   (b)

(c)

Fig. 3 *The hardware used in the research. (a) Raspberry Pi 4 model B; (b) LiDAR sensor; (c) Raspberry Pi camera*

Various categories of obstacles are employed to assess the proposed method. The textureless obstacle (a), texture obstacle (b), textureless obstacles with a single obstacle on the right (c), and textureless obstacles with a single obstacle on the left (d) are the obstacles. These impediments are evaluated in two distinct environments: a crowded environment and an uncrowded environment. Five repetitions of each experiment were implemented. The objective of performing the task five times is to assess the effectiveness and inadequacy of free region detection. With a cumulative total of 100%, each trial contributes 20% to the success rate. The experiment settings that have been conducted in this investigation are illustrated in Fig. 4.
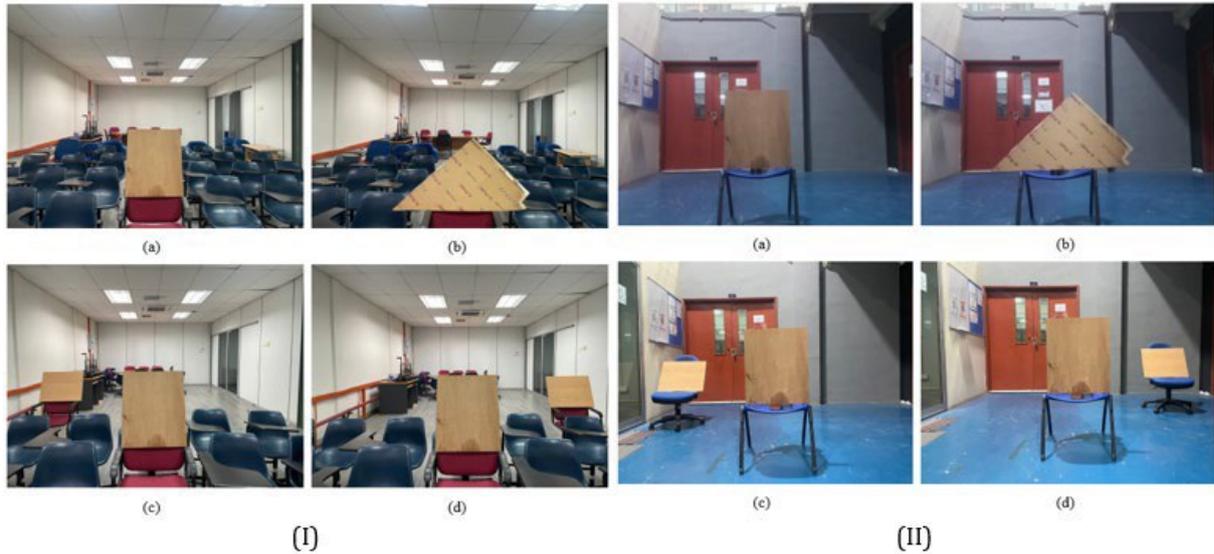
Fig. 4 *The experiments conducted. (I) Crowded environment; (II) Uncrowded environment*

4. Results and Discussion

One of the objectives is to compare the pixel volume expansion method with the combination of image subtraction and distance transform. The results of both methodologies are presented in Tables 1 and 2. The results of the combination of pixel volume expansion methods with image subtraction and distance transform methods in a congested environment are presented in Table 1. The success rate for each form of obstacle is consistent across these methods. The achievement rate is 100% for both textureless and textured obstacles. Concurrently, the success rate is 80% for textureless obstacles with a single obstacle on the left and textureless obstacles with a single obstacle on the right. Consequently, both of these methods are capable of accurately identifying the free region, and they are equally effective in detecting the free region in this particular scenario.

Table 1 *Crowded environment*

Type of obstacle	Combination of distance transform and image subtraction		Pixel volume expansion	
	Free region detection	Success rate	Free region detection	Success rate
Texture-less obstacle	5/5	100%	5/5	100%
Texture obstacle	5/5	100%	5/5	100%
The texture-less obstacle with a single obstacle on the right	4/5	80%	4/5	80%
The texture-less obstacle with a single obstacle on the left	4/5	80%	4/5	80%

Table 2 *Uncrowded environment*

Type of obstacle	Combination of distance transform and image subtraction		Pixel volume expansion	
	Free region detection	Success rate	Free region detection	Success rate
Texture-less obstacle	5/5	100%	5/5	100%
Texture obstacle	4/5	80%	4/5	80%
The texture-less obstacle with a single obstacle on the right	5/5	100%	3/5	60%

Additionally, the results of the combination of the distance transform and image subtraction techniques in an uncrowded environment are presented in Table 2. The success rate for both methods is the same for the texture obstacle, the textureless obstacle, and the textureless obstacle with a solitary obstacle on the left, as shown in Table 4-6. For textureless obstacles, the success rate is 100%, while for textured obstacles, it is 80%. However, for textureless obstacles with a solitary obstacle on the left, the success rate is 100%. The textureless obstacle with a singular obstacle on the right is the significant distinction between the two methods. Pixel volume expansion results in a 60% success rate, while the success rate is 100% when the distance transform and image subtraction are combined. The success rate of the distance transform and image subtraction combination is higher than that of the pixel expansion method for this type of obstacle. The pixel expansion method's constraints are the cause of the discrepancy in success rates. It is unable to accurately identify a textureless object due to the absence of pixel expansion at its center.

The results indicate that both methods are effective in a crowded environment, while the combination of the image subtraction and distance transform method is more effective in an uncrowded environment. In an uncrowded environment, the pixel volume expansion method is not effective due to a variety of factors. Initially, the number of pixel expansions is low in an uncrowded environment, as there are numerous vacant regions in the vicinity. Therefore, the algorithm has a propensity to identify an obstruction as an unrestricted region.

5. Conclusion

In summary, the objectives of this investigation have been satisfactorily accomplished. The proposed method's precision in identifying the free zone is illustrated in this research. Numerous experiments have been conducted in both congested and uncrowded environments. Four types of obstacles have been employed in each environment: textureless obstacles, textured obstacles, textureless obstacles with a single obstacle on the left, and textureless obstacles with a single obstacle on the right. The success rate of the free region was determined by conducting five tests for each category of obstacle. The first objective of this investigation is to develop an algorithm that employs image subtraction and Distance Transform techniques to detect obstructions in a single image acquired by a monocular camera sensor. We have effectively accomplished this objective, and the results illustrate the algorithm's capacity to accurately identify the unoccupied region. In this method, the achievement rate ranges from 80% to 100%. Additionally, the initial phase of algorithm development involves the development of the complete algorithm flowchart. The flowchart provides a visual representation of the algorithm's intended functionality. As soon as the flowchart has been completed, the algorithm development procedure commences. The second objective of this research is to evaluate the effectiveness of Distance Transform and Image Subtraction techniques for identifying free obstacle locations in comparison to alternative methods, including pixel volume expansion. Both strategies have been effectively implemented in the algorithm, with the results presented in Table 1 and Table 2. The comparison of these methods showed that employing distance transform and image subtraction yields a more precise identification of the free region than the pixel volume expansion technique.

Although the results are encouraging, this research has several limitations that should be addressed in future research. One significant constraint is the grid resolution employed during the free region detection procedure. The current algorithm presently divides the image into only 12 grids, which limits the granularity of the analysis. The omission of smaller or narrower free paths that may be navigable, particularly in complex environments, can result from this coarse segmentation. In addition, the investigations were conducted at fixed distances (170 cm and 200 cm) using a static camera setup. This controlled scenario does not accurately represent real-world applications in which the camera, such as one affixed on a UAV, would be in a state of movement that is continuous. Stable lighting conditions and a comparatively uniform background are also presupposed by the method. The image subtraction technique may be subject to degradation in practice due to fluctuations in illumination, shadows, and complex or dynamic backgrounds. The future work will concentrate on improving the resolution of free region detection by increasing the number of grids beyond the current 12. This may result in improved accuracy and a more flexible path selection. Additionally, the algorithm should be evaluated in dynamic environments by employing a moving monocular camera to replicate real-world applications such as UAV navigation. Furthermore, the implementation of adaptive thresholding could potentially enhance performance in response to changing lighting conditions. To further enhance robustness, the integration of this method with other sensors or lightweight machine learning models could improve obstacle detection in complex scenarios.

Acknowledgement

Conflict of Interest

Authors declare that there is no conflict of interest regarding the publication of the paper.

Penerbit
UTHM

Author Contribution

*The authors confirm their contribution to the paper as follows: **study conception and design:** Muhammad Faiz Ramli, Muhamad Wafi Abdul Aziz; **data collection:** Muhammad Syazani Nabil Sobri, Muhamad Wafi Abdul Aziz; **analysis and interpretation of results:** Muhammad Syazani Nabil Sobri, Muhammad Faiz Ramli, Muhamad Wafi Abdul Aziz; **draft manuscript preparation:** Muhammad Syazani Nabil Sobri, Muhamad Wafi Abdul Aziz. All authors reviewed the results and approved the final version of the manuscript.*

References

[1] Badrloo, S., Varshosaz, M., Pirasteh, S., & Li, J. (2022). Image-based obstacle detection methods for the safe navigation of unmanned vehicles: A review. *Remote Sensing*, *14*(15), 3824.

[2] Abdul Aziz, M. W., & Ramli, M. F. (2024). Region Detection Technique Using Image Subtraction and Pixel Expansion Cue for Obstacle Detection System on Small–Sized UAV. *International Journal of Computing and Digital Systems*, *15*(1), 451-463.

[3] Pirasteh, S., Varshosaz, M., Badrloo, S., & Li, J. (2024). Developing an expansion-based obstacle detection using panoptic segmentation. *Journal of Field Robotics*, *41*(5), 1245-1264.

[4] Likas, A., Vlassis, N., & Verbeek, J. J. (2003). The global k-means clustering algorithm. *Pattern recognition*, *36*(2), 451-461.

[5] Naseer, A., Al Mudawi, N., Abdelhaq, M., Alonazi, M., Alazeb, A., Algarni, A., & Jalal, A. (2024). CNN-based object detection via segmentation capabilities in outdoor natural scenes. *IEEE Access*, *12*, 84984-85000.

[6] Altini, N., De Giosa, G., Fragasso, N., Coscia, C., Sibilano, E., Prencipe, B. & Bevilacqua, V. (2021, June). Segmentation and identification of vertebrae in CT scans using CNN, k-means clustering and k-NN. In *Informatics* (Vol. 8, No. 2, p. 40). MDPI.

[7] Al Bashish, D., Braik, M., & Bani-Ahmad, S. (2011). Detection and classification of leaf diseases using K-means-based segmentation and. *Information Technology Journal*, 10(2), 267-275.

[8] Bailey, D. G. (2004, December*). An efficient Euclidean distance transform. In International Workshop on Combinatorial Image Analysis* (pp. 394-408). Berlin, Heidelberg: Springer Berlin Heidelberg.