EASR

Engineering and Applied Science Research

https://www.tci-thaijo.org/index.php/easr/index

Published by the Faculty of Engineering, Khon Kaen University, Thailand

Fast obstacle detection system for the blind using depth image and machine learning

Surapol Vorapatratorn, Atiwong Suchato and Proadpran Punyabukkana*

Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok 10330, Thailand

Received 17 October 2020
Revised 19 December 2020
Accepted 8 March 2021

Abstract

Our research proposes a novel obstacle detection and navigation system for the blind using stereo cameras with machine learning techniques. The obstacle classification result will navigate users through a difference directional sound patterns via bone conductive stereo headphones. In the first stage, the Semi-Global block-matching technique was used to transform stereo images to depth image which can be used to identify the depth level of each image pixel. Next, fast 2D-based ground plane estimation which was separate obstacle image from the depth image with our Horizontal Depth Accumulative Information (H-DAI). The obstacle image will be then converted to our Vertical Depth Accumulative Information (V-DAI) which was extracted by a feature vector to train the obstacle model. Our dataset consists of 34,325 stereo-gray images in 7 different obstacle class. Our experiment compared various machine learning algorithms (ANN, SVM, Naïve Bayes, Decision Tree, k-NN and Deep Learning) performance between classification accuracy and prediction speed. The results show that using ANN with our H-DAI and V-DAI reaches 96.45% in obstacle classification accuracy and 23.76 images per second for processing time which is 6.75 times faster than the recently ground plane estimate technique.

Keywords: Computer vision, Scene understanding, Machine learning, Assistive technology, Visually impaired

1. Introduction

There are 285 million visually impaired humans worldwide, 39 million of that are totally blind and 246 million have low vision [1]. Long canes and guide dogs are famous mobility tools utilized by the visually impaired. These tools are unable to guard users' heads and bodies from collisions, such as street signs, bus poles, open windows, tree branches, and fences. However, the blind pedestrians walking speed of 1.61 meters per second [2] which is close to that of sighted pedestrians walking speed.

Various detecting technologies including GPS, lasers, ultrasonic waves, small radar and stereo camera were used in research and commercial products to solve above problems. The GPS-based Navigation Systems [3]. This technique is not always appropriate for indoor environments as high signal loss rates and low. Laser Cane [4-6] have used the laser distance measurement techniques. This detection has a limited coverage, which is hazardous by the point of blind spots. An ultrasonic waves [7-9], a user needs to hold the device to scan the surrounding object all the time. The RGB-D infrared-camera based sensor technique is more coverage and effective in obstacle detection [10, 11]. Nevertheless, this technique has a limitations in outdoor environments, especially in direct sunlight. There are novel approach to avoiding the obstacle with small radar (Light Detection and Ranging: LiDAR) [12-14]. Its good detection performance in a wide range of incidence angles and weather conditions, provide an ideal solution for state-of-the-art of obstacle avoidance sensor. Unfortunately, those systems were designed for smart vehicle which are not appropriate for mounting with the visually impaired due to its size and weight. Stereo camera based [15-19] to detect obstacles indoor and outdoor scenarios with RANSAC, a 3D-based ground plane estimation algorithm. The stereo camera is lightweight and can be mounted on a person. Despite to the reported performance, those system is needed to be improved such that the processing time and detection accuracy to real implement with real-time.

In this paper, we propose a fast and accurate navigation and obstacle detection system for the totally blind using a stereo camera and machine learning techniques. We introduced a novel technique, Depth Accumulative Information (DAI), which enables fast obstacle extraction from depth image with Horizontal-Depth Accumulative Information (H-DAI), and accurate navigation with Vertical-Depth Accumulative Information (V-DAI). Differently from other existing navigation systems, our system informs the user not only obstacle but also surrounding environment. This is achieved by setting seven types of 3D-sound outputs using stereo wireless bone conductive headphone. This voice consists of four voices set of surround environment describing and three voices set of avoiding obstacle. This notification allow the user to get a sense of the surroundings and avoid incoming obstacle. Our method hardware installation and detection area as shown in Figure 1 (left) and Figure 1 (right), respectively.

*Corresponding author. Tel.: +66 2218 6962
Email address: proadpran.p@chula.ac.th
doi: 10.14456/easr.2021.61

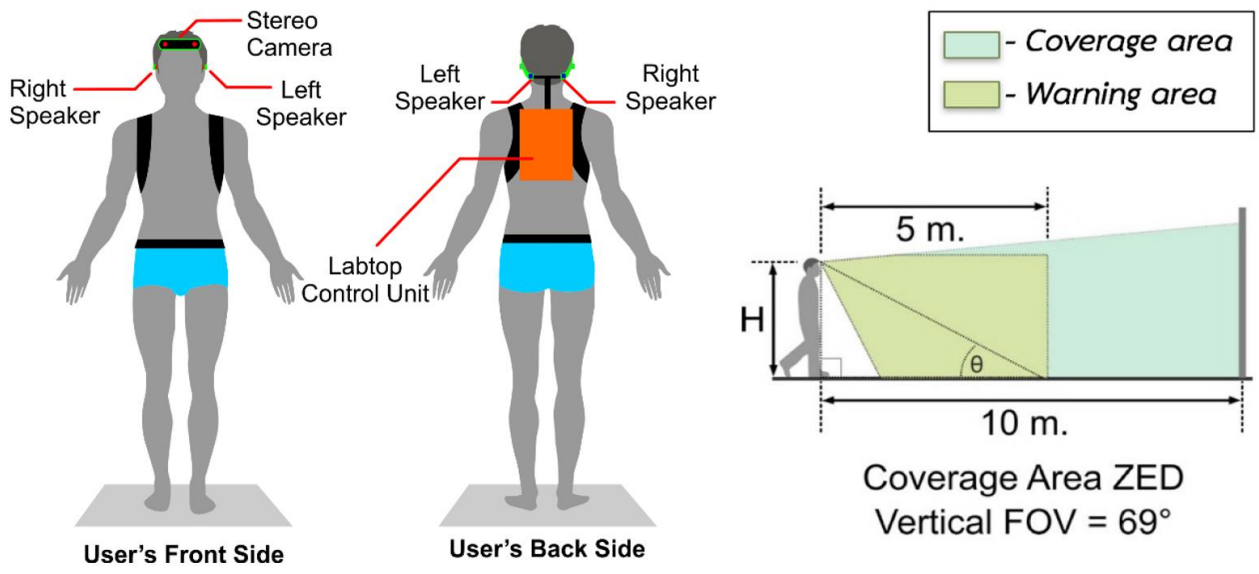

Figure 1 Our proposed method hardware installation (left), Coverage area and warning area (right).

2. Depth image

A wide variety of a stereo camera were used for simple to complex applications [20-24]. However, a large variety of the stereo camera cause a various choices. Three features that must be taken into account before using the stereo camera are the framerate, field of view and resolution. In our research, ZED [20] used to generate death image due to its best performance as shown in Table 1. The ZED is sensor to introduce indoor and outdoor long range depth perception along with 3D motion tracking capabilities, enabling new applications in many industries: AR/VR, drones, robotics, retail, visual effects and more as showed in Figure 2 (left). The ZED captures two synchronized left and right videos of a scene and outputs a depth image of the scene, track the camera position and build a 3D map of the area.

Table 1 Comparison of popular stereo camera specification

Stereo camera	Resolution (px.)	Baseline (mm.)	Field of view (degree)	Max frame rate (fps)
ZED [20]	1242p	120.0	110.0	100
iPhone X [21]	1080p	11.0	30.9	60
PCI nDepth [22]	480p	60.0	31.0	60
Bumblebee2 [23]	488p	120.0	66.0	48
MEGA-DCS [24]	960p	9.0	65.2	30

Depth image generation is the process of constructing the depth image from the ZED followed by camera model, rectification, and correlation steps [25]. The ZED was calibrated by the camera modelling and calibration method to obtain an accurate depth image. The depth image results in our paper show as a color jet-map format [26] where hot colors represent near objects and cold colors represent farther away objects, the depth image result from the ZED as shown in Figure 2 (right).

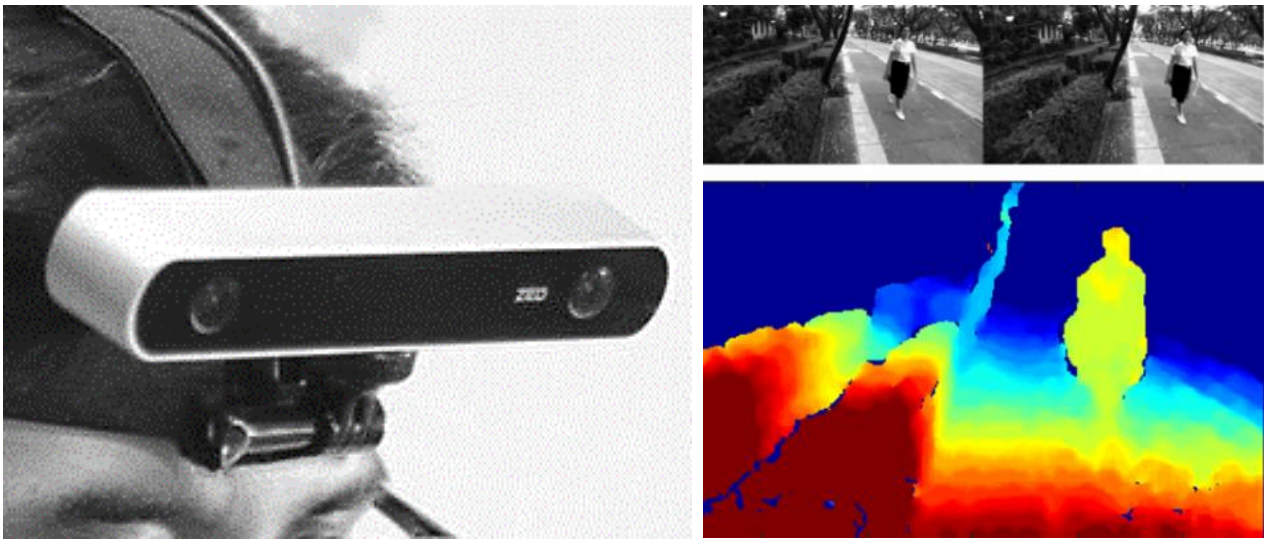

Figure 2 ZED Stereo camera used in this study (left), depth images from ZED (right).

3. Propose method

Our system starts with getting the depth image from the ZED stereo camera, which is mounted on the user's head. The depth image will be transformed into Horizontal-Depth Accumulative Information (H-DAI), which represents a side-view depth image projection. H-DAI is used in our 2D-based fast ground removal process to crop out the background region in the depth image until only the obstacle image remains. Next, the obstacle image is converted to Vertical-Depth Accumulative Information (V-DAI), which is a virtually projected obstacle image in the bird's-eye view. The result of applying V-DAI is used as the input data (feature vector) for classification step. Finally, the resulting interface of the classification is the predicted obstacle class. This information is converted to 3D-tone and voice navigation via user's wireless bone conductive headphone. The system diagram of the proposed methods is shown in Figure 3.

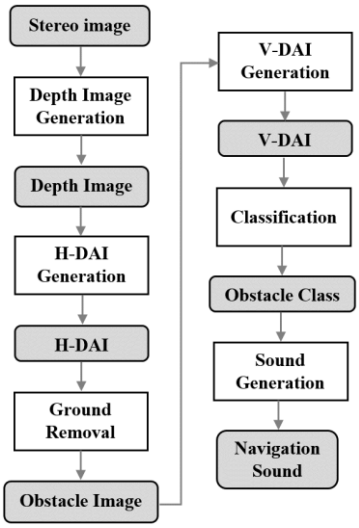

Figure 3 System diagram of the proposed methods.

3.1 Ground removal using Horizontal-DAI

In the first step, the depth image frequency along the Y-axis was observed with the bin count technique to obtain the H-DAI image and side view depth image projection. The maximum value of each H-DAI image column showed the ground curve that represented the pathway. The ground curve information was used to remove the pathway from the depth image into the obstacle image. High depth values represent close objects while low depth values represent long distance objects as shown in Figure 4 (left).

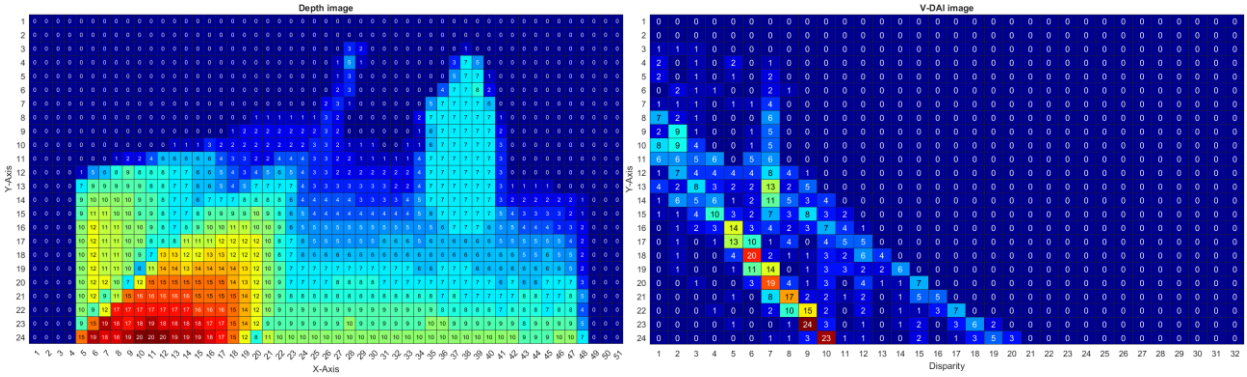

Figure 4 Depth image for convert to H-DAI (left), the H-DAI after depth level counting process (right)

Creating a horizontal–DAI (H-DAI) image is carried out by sequentially counting the frequency of depth values from the bottom rows of depth image to the top rows, sequentially. The warm color represents the high frequency of depth value and the cold color represents the low frequency of depth value. The H-DAI image is a virtually projected image from the side used to find the pathway area as shown in Figure 4 (right). The ground removing process is creates an obstacle image that is cropped out of the pathway until only the harmful objects remain or ground plane estimation process. The pathway region can be identified from the ground curve of H-DAI. Firstly, we estimate the ground curve from the maximum value position of each H-DAI image column with the polynomial function as shown in Equation (1):

$$G = \sum_{i=1}^{n+1} p_i x^{n+1-i} \tag{1}$$

where G is ground curve position depend on disparity value x, n is the degree of the polynomial, p_i is the polynomial constant. The order gives the number of fit coefficients, and the degree gives the highest power of the predictor variable. However, an example of identify ground curve is shown in Figure 5.

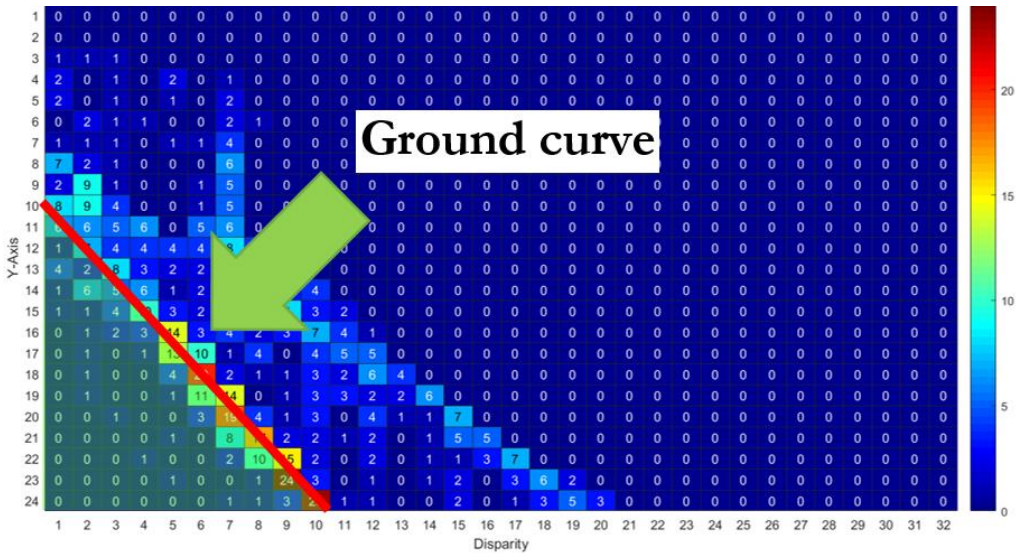

Figure 5 H-DAI with ground curve in red line

Next, we select the H-DAI data that is less than the ground curve to create the ground mark. This data is stored in binary format. The pathway region is logic 1 and the rest is logic 0 as shown in Figure 6 (left). Then we delete the data at the location of the ground mark of the depth image as shown in Figure 6 (right). Based on the results of the depth image ground removal step, it appears that there is still an obstacle region which indicates the object 5 meters far away. Finally, we have to use the un-region of interest mark, shown in Figure 7 (left), to remove the obstacle that is beyond the considered range. The result is an obstacle image as shown in Figure 7 (Right).

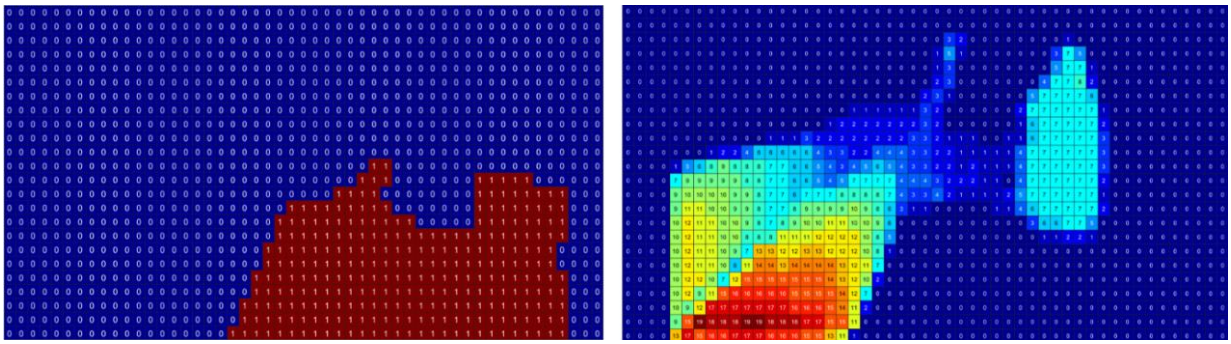

Figure 6 Ground mark (left), removed ground depth image (right)

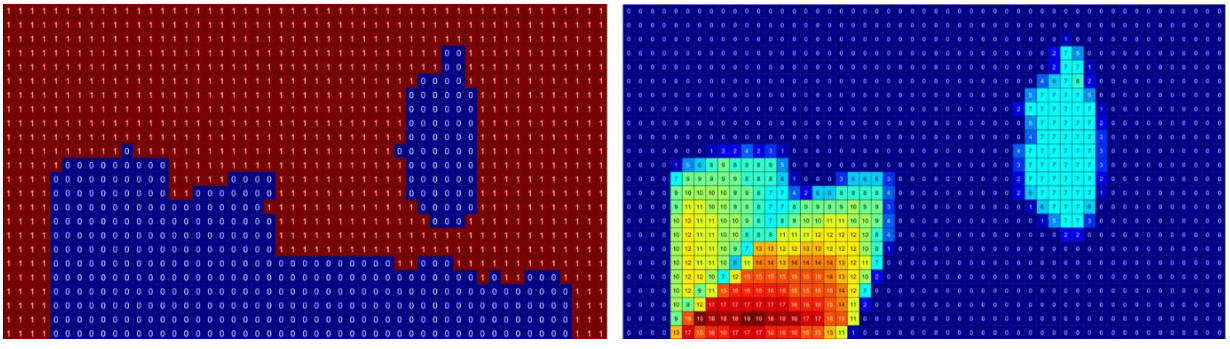

Figure 7 Un-region of interest mark (left), obstacle image (right)

3.2 Classification using vertical-DAI

In order to use machine learning to classify obstacles, we created an obstacle model to guide users to avoid various obstacles. This is our second proposed contribution, fast and accurate obstacle classification using Vertical-Depth Accumulative Information (V-DAI) feature. In order to use machine learning to classify obstacles, we created an obstacle model to guide users to avoid various obstacles, the diagram as shown in Figure 8. In the first step, data acquisition stores a lot of stereo images from the stereo camera in-person view, as the individual walked along the path in various routes. The second step is data preparation in which the obstacle image is converted by the V-DAI feature and prepared for experts to label the training obstacle models. In the final step, our machine learning is trained in obstacle classification steps in a process called obstacle model training.

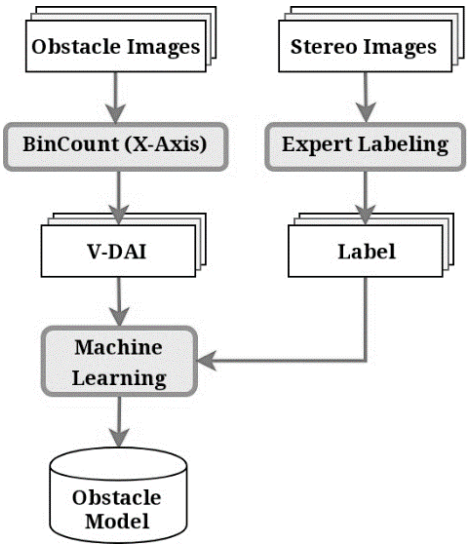

Figure 8 Obstacle model training diagram

Creating a Vertical-Depth Accumulative Information (V-DAI) feature for the training set is done by finding the frequency of depth value that occurs in each column or width of the obstacle image (depth image ground removed) by counting from the left column along the width of the obstacle image until the right column as shown in Figure 9 (left). The completed V-DAI is shown in Figure 9 (right). This V-DAI is a virtually projected image from the top or birds-eye view used to train obstacle models.

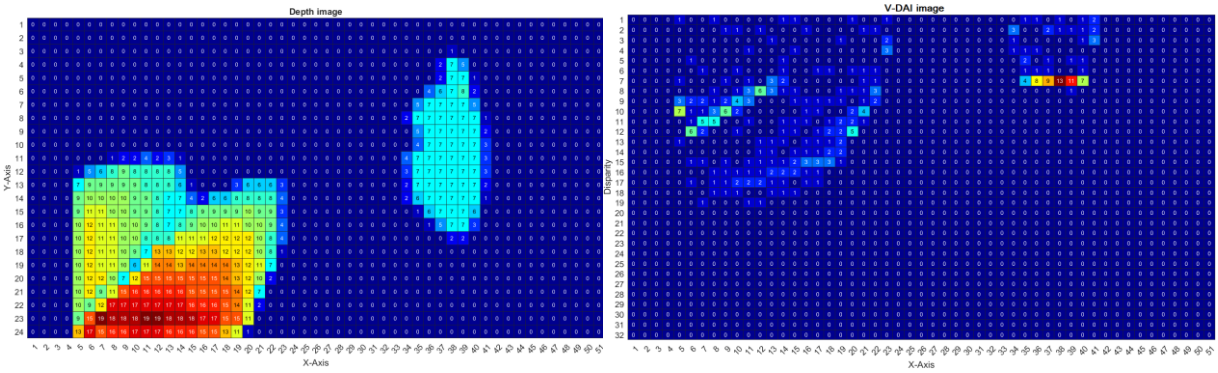

Figure 9 Obstacle image for convert to V-DAI (left), the V-DAI after depth level counting process (right)

In the model training step, we use classification learner tool [27], starting with importing the V-DAI feature vectors obtained from the feature extraction step and the obstacle class label which is from the expert labeling step. In the next step, we randomly divide all dataset into 3 parts including training set, validation set and test set. For the training set is a set of data used to adjust the data in the neural structure to minimize errors. In the validation set is a set of data that is used to test the ability to generalize characteristics and must stop training models as well. Finally, the Test set is used to measure the performance of the model, due to it being a set of data that has not been trained before. The division of the dataset for model training for 70% training set, 15% validation set, and 15% testing set. Then, we trained an obstacle model with various learning algorithms, including Artificial Neural Networks (ANN) [28], Support Vector Machines (SVM) [29], Naïve Bayes [30], Decision Tree [31], k-Nearest Neighbor (k-NN) [32] and Deep Learning (CNNs) [33]. The final step for model training is to measure the efficiency of each machine learning algorithms. We measure three classification performance metrics. The first is accuracy in the ability to classify obstacles. The second is prediction speed, or the time that machine learning uses to process the obstacle class in the classification step. The final performance metric is training time, or time spent on training models. All the three metrics were evaluated in order to identify the most suitable machine learning algorithms for the obstacle classification.

3.3 Navigation

In navigation step starts by obtaining the obstacle class from classification step. The example resulting interface of our system is shown in Figure 10. The user's navigation starts by using the 7 obstacle classes converted to an English-speaking female navigation voice. This voice consists of 4 voice set of surround environment describing and 3 voice set of avoiding obstacle. "All clear" (free space, no obstacles), "Left object" (objects on left side, no need to avoid), "Right object" (objects on right side, no need to avoid) and "Parallel object" (objects on left and right side, no need to avoid), "Keep left" (dodge to the left), "Keep right" (dodge to the right), and "Slow down" (stop walking). When the obstacle detected, the navigation voice will be generate to wav sound format and inform the user via a bone conductive stereo headphone. Three-dimensional tone pattern including sound spectrum and left-right altitude as shown in Figure 11.

Figure 10 The example resulting interface of our obstacle classification and navigation system

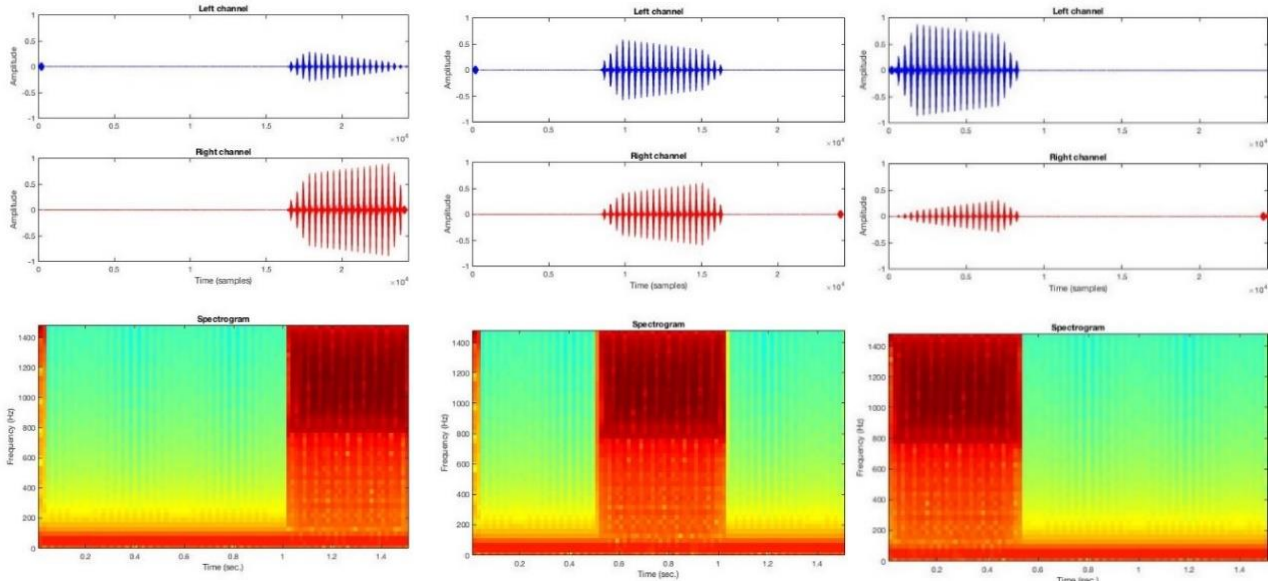

Figure 11 Three-dimensional tone pattern for keep left (Left), slow down (Center) and keep right (Right)

When the Obstacle class status changes, the voice navigation will be played via the bone conductive stereo headphone. When the obstacle class remains the same, the three-dimensional tone pattern is played repeatedly, which consists of 7 sound patterns with different frequencies and different left-right loudness except for "All clear" (free space is not obstructed) will be a silent sound.

4. Experiment and evaluation

4.1 Data acquisition

Data collection for the training set is done by mounting a stereo camera on the head position of the data collector to get a person's perspective view according to Figure 12 (left) and then recording a stereo image in 1,344 × 376 pixels with a sampling rate of 15 frames per second in mp4 gray video format. Furthermore, all stereo image that have been recorded from Chulalongkorn University's pedestrian pathways which contain obstructions along the path, such as electricity poles, bushes, traffic signs, walls, etc., as shown in Figure 12 (right). The total walking distance is 2.3 kilometers with a duration of 39.61 minutes and containing 34,325 stereo images, the dataset information shown in Table 2. Those stereo images will be converted into obstacle images for the feature extraction process.

Table 2 Stereo image dataset information

Dataset	Distance (m.)	Duration (sec.)	Sample (image)	Percentage (%)
Chula-1	240	208	3,048	8.88
Chula-2	400	404	5,909	17.21
Chula-3	160	193	2,756	8.03
Chula-4	500	440	6,565	19.13
Chula-5	350	452	6,774	19.73
Chula-6	650	680	9,273	27.02
Total	2,300	2,377	34,325	100.00

Figure 12 Stereo camera mounted on data collector head (left), the example of stereo image for data acquisition process (right)

4.2 Expert labeling

Expert labeling is processed to generate labels to be used in model training steps for a supervised learning technique. Labeling in this research starts with preparing videos for experts who specify the video frame numbers of each stereo image. Moreover, we are attaching obstacle images for more information and the dataset images for the expert is shown in Figure 13.

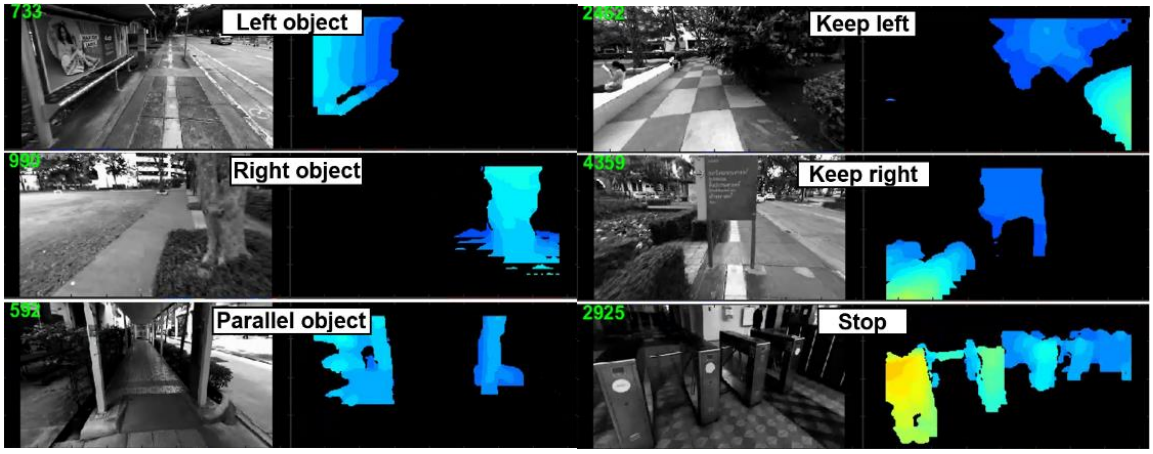

Figure 13 The example of image and depth image for expert labelling process

The expert will specify the type of obstacle class in every image frame. They will identify the obstacle into 7 different classes consisting of a left object, right object, parallel object, keep left, keep right, all clear, and stop class. The labeling results will be stored in the feature table in CSV format. The labeling results will be stored in the feature table in CSV format. The columns 1 to 400 of the Table will be the V-DAI feature vector obtained from the feature extraction step. The last column is the obstacle class where the expert will complete this all. The results of the expert labeling are shown in Table 3.

Table 3 Number of sample in each class and coverage

Class no.	Obstacle class	Number of image	Percentage (%)
1	All clear	6,985	20.35
2	Left object	6,502	18.94
3	Right object	8,733	25.44
4	Parallel object	7,519	21.91
5	Keep left	1,389	4.05
6	Keep right	2,412	7.03
7	Stop	785	2.29
	Total	34,325	100

4.3 Ground removal with H-DAI

The depth sampling of the H-DAI depth levels affects processing time and. Thus, the performance of the depth levels from 8 to 128 levels was compared in Table 4 to find the optimal depth level. The results showed that the low depth sampling provided a low accuracy result. However, the highest depth sampling result showed high processing time. Thus, the optimal depth sampling will be used at 64 levels due to the highest level of classification accuracy.

Table 4 Comparison performance of depth level

Depth sampling (level)	Runtime (ms./image)	ANN-Classification	
		F-measure (%)	Accuracy (%)
8	156.2940	85.3928	74.5091
16	157.4902	89.6734	82.8199
32	161.7783	97.0075	95.4469
64	165.9861	97.8666	96.7671
128	178.7246	97.4006	96.0393

We conducted an experiment to determine the optimal degree of the polynomial, which found that a high degree of polynomial increases the processing time due to the complexity of polynomial equations and also causes errors in removing passages. Therefore, we selected the degree of the polynomial of 2 degrees as indicated by the results shown in Table 5.

Table 5 Comparison performance of degree of polynomial

Degree of polynomial (degree)	Runtime (ms./image)	ANN-classification	
		F-measure (%)	Accuracy (%)
1-degree	168.8248	97.7952	96.6655
2-degree	178.0120	97.8666	96.7671
3-degree	188.8431	97.5316	96.2593
4-degree	197.8748	96.9926	95.4299
5-degree	210.9794	96.1156	94.0420

4.4 Classification using V-DAI converting

To find the most suitable V-DAI size, we tested sizes 40 × 40, 20 × 20, and 10 × 10 pixels as shown in Table 6. The results showed that the smaller V-DAI image size can effectively processing time well. Unfortunately, using too low of a size, decreased classification accuracy. Therefore, we used the V-DAI of 20 × 20 pixels which is the 400-dimensional vector in binary format.

Table 6 Model performance for each V-DAI size

V-DAI size (pixels)	Training time (sec.)	Predict speed (f/s)	F-measure (%)	Accuracy (%)
10 × 10	94.3750	158,380	54.3254	58.3214
20 × 20	135.7868	101,180	95.2069	96.4457
40 × 40	230.8376	61,322	91.1273	94.3751

To ensure that the V-DAI feature is a suitable feature for our classification, we also tested the depth image and obstacle image ability to extract the feature vector by reducing the image to 20 × 20 pixels in binary format. Subsequently, these features were tested to compare the performance between V-DAI and depth image and obstacle image. It is not surprising that the V-DAI feature has the highest processing speed and precision as shown in Table 7.

Table 7 Performance for each feature comparison

Feature vector (20 × 20 px.)	Training time (sec.)	Predict Speed (f/s)	F-measure (%)	Accuracy (%)
Depth image	314.9313	60,455	56.2997	60.8827
Depth obstacle	639.4254	47,667	76.478	80.5594
V-DAI	135.7868	101,180	95.2069	96.4457

To provide the most obstacle model performance both in processing speed and accuracy, we have conducted experiments to find the most suitable machine learning as well. In this experiment, we trained the model with various machine learning algorithms, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), Naïve Bayes, Decision Tree, k-Nearest Neighbor (k-NN) and Deep Learning (CNNs) with our V-DAI feature. The result of machine learning performance which is using averaged value from 10 rounds training, the result as shown in Table 8. The obstacle model that reaches the best performance will be used in the obstacle classification step. Based on the experiment result, the ANN algorithm has the highest accuracy in obstacle classification as 96.45 % and also has the fastest prediction speed as 101,180 frames per second.

Table 8 Machine learning algorithms performance comparison

Machine learning algorithm	Training time (sec.)	Predict speed (f/s)	F-measure (%)	Accuracy (%)
ANN [28]	135.78	101,180	95.2069	96.4457
SVM [29]	242.61	1,600	91.0263	94.4268
CNNs [33]	48.6162	9,639	87.6119	92.5617
k-NN [32]	0.00	74	88.6198	91.1610
Decision Tree [31]	14.72	55,000	78.6993	86.5637
Naïve Bayes [30]	1.63	3,200	78.3175	85.0066

4.5 Benchmark testing

To ensure that our propose method is better than recently technique, we compare the processing time and classification efficiency between our propose method and RANdom Consensus algorithm (RANSAC); 3D-based obstacle segmentation technique [16]. The RANSAC is a ground plane estimation method processing on 3-D world point coordinates or point cloud which can be generated from depth image conversion. After that, RANSAC will fit the ground plane and remove the ground data from the point cloud. The last step is to convert the 3-D world point coordinate point cloud to a 2-D image or obstacle image which consume the processing time. The processing time between our propose method and RANSAC based. The results of the performances comparison showed that in our propose method, take a 42.0837 milliseconds for one image processing. In other words, it reached a frame rate as 23.76 images per second which 6.75 times faster than the RANSAC based as shown in Table. 9. The various scenarios detection result between RANSAC and our proposed method are summarized in Table 10.

Table 9 The obstacle detention performance comparison between RANSAC and our propose method

Performance	RANSAC [16] (3D-based)	Our propose method (2D-based)	Diff.
Accuracy (%)	86.1069	96.7671	+10.6602
F-measure (%)	85.1804	95.2069	+10.0265
Precision (%)	83.0081	95.8018	+15.1314
Recall (%)	80.6704	94.6458	+13.9754
Runtime (ms./image)	282.3457	42.0837	-240.2620

Table 10 Detection obstacle result in various scenario between benchmark technique (RANSAC) and our proposed method

Scenario	Obstacle detection result	
	RANSAC [16] (3D-based)	Our propose method (2D-based)
Patterned walkway, no obstacle.	Can correctly separate walkway.	Can correctly separate walkway.
No patterned walkway, no obstacle.	Partially separates the walkway.	Partially separates the walkway.
Reflected surface walkway, no obstacle.	Interprets the reflection as hole.	Interprets the reflection as walkway.
Water surface walkway, no obstacle.	Interprets the reflection as hole.	Interprets the reflection as walkway.
Difference level walkway, no obstacle.	Can correctly separate walkway.	Interprets the different level as obstacle.
Spacious walkway, Including wall/big sign.	Interprets the wall/big sign as walkway.	Can correctly separate walkway.
Narrow walkway, Including wall/big sign.	Interprets the wall/big sign as walkway.	Can correctly separate walkway.
General walkway with stairs.	Interprets the stairs as walkway.	Interprets the stairs as obstacle.
General walkway with holes.	Interprets the holes as obstacle.	Interprets the holes as walkway.
General walkway, dense obstacle.	Unable to find walkway.	Partially separates the walkway.

4.6 System testing with unseen area

To ensure that our systems has no the machine learning over-fit problem or still work outside the training area which is used in training the obstacle model, we tested our system with outside the training area. For the testing route, we chose the walkway from the Phaya-thai police station home to the Phaya-thai skytrain BTS station in Pathumwan, Bangkok as shown in Figure 14

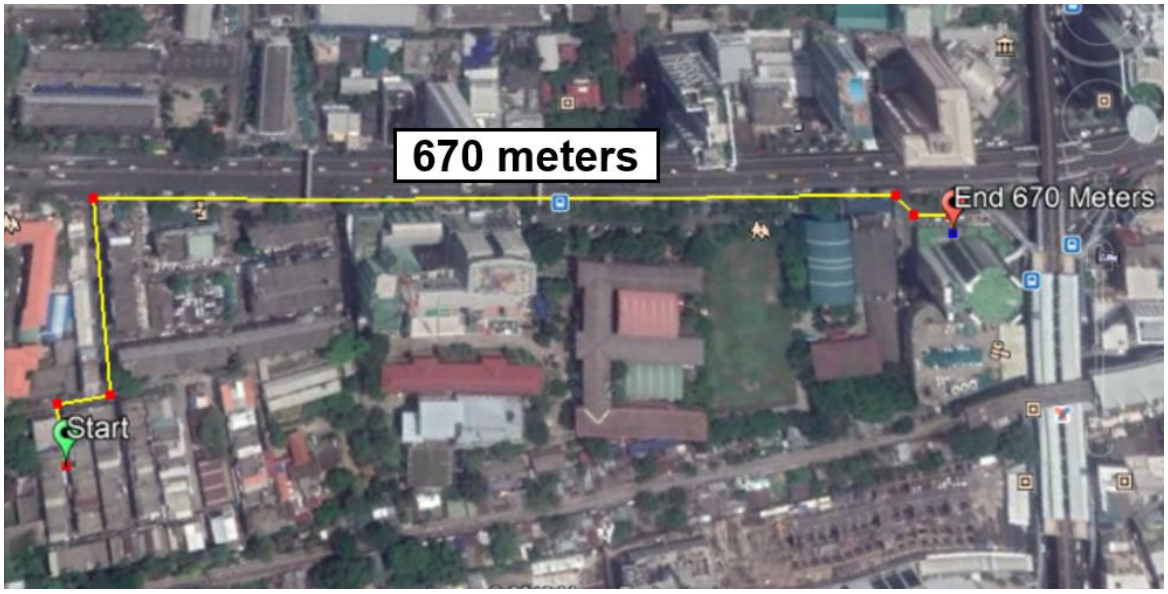

Figure 14 Real-world testing 670 meters walking route

We chose this route for the experimental because the route has many obstacles, such as pedestrians, motorcycles, cars, traffic signs, electric poles, walls, trees, shrubs, and others. The testing took 7 minutes 57 seconds at a distance of 670 meters, consisting of a total of 7,155 images. After testing, we found that the system was able to effectively navigate and warn the user of those obstacles. The accuracy of classification is 87.43% which F-measure score is 79.48%. The example of testing scenarios are shown in Figure 15.

Figure 15 The example of real-world testing result

5. Conclusions and discussion

Our research proposes a real-time navigation and obstacle detection system for the blind using a stereo camera. Our detection is based on a stereo camera with a machine learning technique. Our first contribution is fast obstacle segmentation method using our Horizontal-Depth Accumulative Information (H-DAI) which is 6.75 times faster than the traditional 3D-based ground plane estimate technique (RANSAC). The second contribution is a fast and accurate obstacle classification technique where our Vertical-Depth Accumulative Information (V-DAI) feature reaches 96.45% in accuracy and 23.76 images per second in processing speed. We conducted many tests to ensure the highest level of accuracy and speed of obstacle detection as errors in detection or excessive processing time have the potential to cause enormous harm to the user.

For future direction, limitations of this work such as the ability to use in a dark or foggy environment can be considered. This may be addressed using fusion sensors between the stereo camera and LiDAR that are available in modern mobile devices. Such sensors can improve the detection in low-vision scenarios. Another potential future work may include further enhancement in result accuracy by considering the trajectory direction of incoming objects. This could be done by incorporating object trajectory prediction based on object's previous and current positions. Furthermore, object recognition could be integrated to enhance our navigation system. The user would be informed of the type of an incoming object. This will help them to be prepared for quick avoidance of dangerous or extremely harmful objects.

6. Acknowledgements

This research has been financially supported by "The 100th Anniversary Chulalongkorn University Fund for Doctoral Scholarship" and "The 90th Anniversary Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund)". However, we would like to thank the blind and visually impaired volunteers from the "Department of Ophthalmology, Faculty of Medicine Siriraj Hospital" for their useful information and collaboration.

7. References

[1] WHO. Draft action plan for the prevention of avoidable blindness and visual impairment 2014-2019: universal eye health: a global action plan 2014-2019. Geneva: World Health Organization; 2013.
[2] Clark-Carter DD, Heyes AD, Howarth CI. The efficiency and walking speed of visually impaired people. Ergon. 1986;29(6): 779-89.
[3] Loomis JM, Golledge RG, Klatzky RL. GPS-Based navigation systems for the visually impaired. In: Barfield W, Caudell T, editors. Fundamentals of wearable computers and augmented reality. New Jersey: Lawrence Erlbaum Associates; 2001. p. 429-46.
[4] Ritz M, Konig L. Laser technique improves safety for the blind. MST NEWS. 2005;5:39.
[5] Inc NRI. The laser cane, model N-2000 [Internet]. 2020 [cited 2020 Sep 20]. Available from: http://www.nurion.net.
[6] Ulrich I, Borenstein J. The guide cane-applying mobile robot technologies to assist the visually impaired. IEEE Trans Syst Man Cybern Syst Hum. 2001;31(2):131-6.
[7] Research G. The miniguide mobility aid [Internet]. 2020 [cited 2020 Sep 20]. Available from: http://www.gdp-research.com.au/minig_1.htm.
[8] Ito K, Okamoto M, Akita J, Ono T, Gyobu I, Takagi T, et al. CyARM: an alternative aid device for blind persons. Extended abstracts proceedings of the 2005 conference on human factors in computing systems; 2005 Apr 2-7; Portland, USA. New York: Association for Computing Machinery; 2005. p. 1483-8.
[9] Vorapatratorn S, Nambunmee K. ISonar: An obstacle warning device for the totally blind. J Assist Rehabil Ther Tech. 2014;2(1):23114.
[10] Damaschini R, Legras R, Leroux R, Farcy R. Electronic travel aid for blind people. In: Pruski A, Knops H, editors. Assistive technology: from virtuality to reality. Amsterdam: IOS Press; 2005. p. 251-5.

[11] Vorapatratorn S, Suchato A, Punyabukkana P. Real-time obstacle detection in outdoor environment for visually impaired using RGB-D and disparity map. Proceedings of the international convention on rehabilitation engineering & assistive technology; 2016. p. 1-4.
[12] Kumar AD, Karthika R, Soman K. Stereo camera and LIDAR sensor fusion-based collision warning system for autonomous vehicles. In: Jain S, Sood M, Paul S, editors. Advances in computational intelligence techniques. Singapore: Springer; 2020. p. 239-52.
[13] Gao H, Cheng B, Wang J, Li K, Zhao J, Li D. Object classification using CNN-based fusion of vision and Lidar in autonomous vehicle environment. IEEE Trans Ind Informat. 2018;14(9):4224-31.
[14] Catapang AN, Ramos M. Obstacle detection using a 2D LIDAR system for an autonomous vehicle. 2016 6th IEEE International conference on control system, computing and engineering (ICCSCE); 2016 Nov 25-27; Penang, Malaysia. New York: IEEE; 2016. p. 441-5.
[15] Li B, Zhang X, Munoz JP, Xiao J, Rong X, Tian Y. Assisting blind people to avoid obstacles: an wearable obstacle stereo feedback system based on 3D detection. 2015 IEEE International conference on robotics and bio mimetics (ROBIO); 2015 Dec 6-9; Zhuhai, China. New York: IEEE; 2015. p. 2307-11.
[16] Rodriguez A, Yebes JJ, Alcantarilla PF, Bergasa LM, Almazan J, Cela A. Assisting the visually impaired: obstacle detection and warning system by acoustic feedback. Sensors. 2012;12(12):17476-96.
[17] Correal Tezanos R. Stereo vision-based perception, path planning and navigation strategies for autonomous robotic exploration. Madrid: Universidad Complutense de Madrid; 2015.
[18] Emani S, Soman K, Variyar VS, Adarsh S. Obstacle detection and distance estimation for autonomous electric vehicle using stereo vision and DNN. In: Wang J, Reddy G, Prasad V, Reddy V, editor. Soft computing and signal processing. Singapore: Springer; 2019. p. 639-48.
[19] Guindel C, Martin D, Armingol JM. Traffic scene awareness for intelligent vehicles using Conv Nets and stereo vision. Robot Autonom Syst. 2019;112:109-22.
[20] Stereolabs. ZED stereo camera [Internet]. 2020 [cited 2020 Sep 20]. Available from: https://www.stereolabs.com/zed/.
[21] Inc A. iOS Device Capture Format Specifications [Internet]. 2020 [cited 2020 Sep 20]. Available from: https://developer.apple.com/library/archive/documentation/DeviceInformation/Reference/iOSDeviceCompatibility/Cameras/Cameras.html.
[22] AZOROBOTICS. PCI nDepth™ Vision System from Focus Robotics [Internet]. 2020 [cited 2020 Sep 20]. Available from: https://www.azorobotics.com/equipment-details.aspx?EquipID=467.
[23] Ltd VSP. The Bumblebee®2 stereo vision camera [Internet]. 2020 [cited 2020 Sep 20]. Available from: https://www.voltrium.com.sg/en/cameras/machine-vision-cameras/stereo-products/bumblebee2.
[24] Design V. MEGA-DCS stereo head [Internet]. 2020 [cited 2020 Sep 20]. Available from: http://users.rcn.com/mclaughl.dnai/sthmdcs.htm.
[25] Cochran SD, Medioni G. 3-D surface description from binocular stereo. IEEE Trans Pattern Anal Mach Intell. 1992;(10):981-94.
[26] The MathWorks I. Jet colormap array [Internet]. 2020 [cited 2020 Sep 20]. Available from: https://www.mathworks.com/help/matlab/ref/jet.html.
[27] The MathWorks I. Classification Learner [Internet]. 2020 [cited 2020 Sep 20]. Available from: https://www.mathworks.com/help/stats/classificationlearner-app.html.
[28] Jain AK, Mao J, Mohiuddin KM. Artificial neural networks: a tutorial. Comput.1996;29(3):31-44.
[29] Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B. Support vector machines. IEEE Intell Syst appl. 1998;13(4):18-28.
[30] Murphy KP. Naive bayes classifiers. Vancouver: University of British Columbia; 2006.
[31] Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. IEEE Trans Syst Man Cybern. 1991;21(3):660-74.
[32] Keller JM, Gray MR, Givens JA. A fuzzy k-nearest neighbor algorithm. IEEE Trans Syst Man Cybern. 1985;(4):580-5.
[33] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Adv neural inform process syst. 2017;60(6):84-90.