International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-4, April 2019

Object Identification to Assist Visually Challenged

Sreenu Ponnada, Srinivas Yarramalle

Abstract*: Our recent survey on problems faced by the visually challenged suggested self-reliant movement in urban spaces as a major challenge. In this paper, the authors propose a novel way of assisting the visually challenged to identify various public transport means and also help get onboard with little or no assistance from others. We use an integrated system of a mobile phone connected wirelessly via Bluetooth to Arduino controlled array of uniquely placed ultrasonic sensors complemented with vibro motor for haptic feedback. The system detects obstacles in all four directions and helps navigate through crowded spaces. We employ image-based recognition based on the visual information obtained from a mobile phone camera to detect vehicles like bus, car, truck, two-wheelers, auto-rickshaw (three-wheeler) as well as objects. The results are converted to audio feedback via the mobile device. We experiment with detectors like Multivariate Generalized Gaussian Mixture Model (MGGMM) based on features from Histogram of Oriented Gradients (HOG). Results indicate 96.89% accuracy.

***Index Terms*: Computer Vision, GMM, HOG, MGGMM, Object Detection.**

I. INTRODUCTION

Navigation in unfamiliar and dynamic spaces is a challenge for people with visual impairments. In regions with increased traffic flow, visually challenged people have trouble overcoming barriers. This necessitates the visually challenged to be dependent on the help of others, resulting in a reduced quality of life. We conducted a survey on the problems faced by the visually challenged in which over 150 people participated. One of the major challenge reported is to utilize public transport in a self-reliant manner.

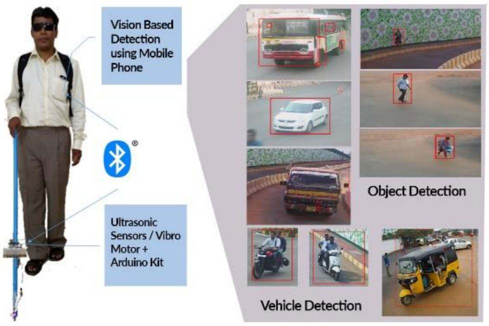


Fig. 1. ***Integrated System Overview.***

Unfortunately, there are about 441 million people who are either fully or partially visually impaired people across all countries[1]. Therefore, it is an important topic to help the visually challenged live a quality life as a normal person. In this paper, we propose an integrated framework for helping the visually challenged utilize the road public transport effectively. Fig 1 gives a quick overview of the integrated system. We build an integrated framework [2] which consists of a group of three sonars (right, left and front) attached on a white cane. This framework is backed up by the Arduino processor. There is one vibro-feedback attached to the cane to assist the user about the hurdle as well as a Bluetooth module to connect to a mobile phone for acoustic feedback and more advanced image-based object detection.

The visual information captured from the mobile device drives a software-based identification system. The captured frames are pre-processed and HOG features are extracted to drive a Multivariate Generalized Gaussian Mixture Model (MGGMM) to identify a variety of vehicles (bus, three-wheelers, two-wheelers, truck etc.).

II. RELATED WORK

Various engineering aids are there to increase the quality of life of the Visually Challenged. These systems use GPS, cameras, infrared, laser, and ultrasonic sensors for recognition of various items in near by field and communictes the data to the user through the tactile interfaces or haptics.

The Smart Cane [3] and UltraCane/Batcane [4,5] are systems that entirely rely on a white cane and use sonar sensors attached to the cane for detecting impediments in front of the cane by even giving a direction cue. However, these systems cannot recognize the type of obstacle efficiently. [2] present a hybrid approach using the cues from the ultrasonic sensor to trigger a vision based detection of manholes and staircases with better accuracy.

Manuscript published on 30 April 2019.
***** Correspondence Author (s)
Sreenu Ponnada*, Department of Computer Science and Engineering, GITAM Institute of Technology, Visakhapatnam, India.
Srinivas Yarramalle, Department of Computer Science and Engineering, GITAM Institute of Technology, Visakhapatnam, India.

A detailed review of various assistive technology solutions is presented in [6]. However, most of the solutions are relatively expensive than the simple system presented in [2]. Stereo vision systems [7] and depth sensor based systems [8,9] can determine more accurately the location of the obstacles, however, fall short of accurately identify the type of object. Computer vision based approaches combined with machine learning pose a better alternative for accurate object identification. For example, [10] proposes a stereo camera mounted helmet paired with an Android phone that connects to a cloud computing platform that applied a vision-based approach for OCR, object detection and recognition etc. We, however, extend the hybrid system [2] to further perform object detection specific to use of public road transport.

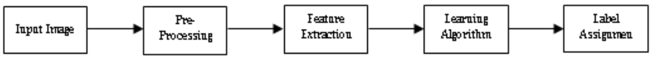


Fig. 2. *Pipeline for vision-based object classification*

Public road transport aid should involve bus detection. Fig 2 describes the pipeline for vision-based object detection. The accuracy of detection varies depending upon the features extracted and learning algorithm used. Commonly used features are Haar [11–13], SIFT [14], SURF [14,15], and HOG [15, 16]. In any image, edges and corners provide a lot more information about object shape than flat regions. Recently, saliency maps along with Gaussian Mixture models are shown to work well with image segmentation [17]. Often the edges and their directions contribute to the saliency map and these are clearly visible in gradients (x and y derivatives) of the captured image. For example, [18] uses HOG features & Cascaded SVM (Support Vector Machine) model for bus detection on road. We use HOG feature descriptions in our work for vehicle and object identification.

The final stage of the pipeline usually a classifier. On a busy road, there will be all kinds of vehicles and people moving around, thereby rendering the input dataset to be multi-dimensional. Discerning the needed patterns and information in multidimensional data requires the selection of a suitable statistical model and knowledge of its parameters. Mixture models are extensively used statistical methods in several applications and they permit a recognized approach for unsupervised learning [19]. Gaussian distribution is isotropic and has the ability to denote information efficiently using a covariance matrix and mean vector. This makes the Gaussian mixture (GM) decomposition a widespread method. Nevertheless, the rigidity of the form and symmetry around the mean prevents it to be ineffective. Generalized Gaussian distribution (GGD) is a better model for non-Gaussian data [20, 21]. GGD has one extra dimension called λ used to control the tail of distribution. In this article, we propose Multivariate Generalized Gaussian Mixture Model (MGGMM) for segmenting the input image and further classify the regions into known categories like a vehicle, person, etc.

III. PROPOSED METHODOLOGY

A. Multivariate Generalized Gaussian Mixture Model

Our target set of object images viz. object, vehicle and various categories of vehicles, may exhibit prototypes which are non-symmetric and partly symmetric i.e., the shape of pixels inside the object image regions may be exhibiting a Gaussian, mixture of Gaussian or non-Gaussian. Therefore to understand these sorts of object image variations, one needs to think about the statistical models which can house mesokurtic, platykurtic and leptokurtic, i.e., the statistical models ought to be developed to classify the pattern of the pixels more aptly. This requires to consider the generalizations of the Gaussian mixture models.

This paper presents a statistical model based on MGGMM for efficient segmentation of object images. At any point in time, as the numbers of vehicles or people moving around are high in number, the proportionality of the number of feature vectors becomes a monotonous task. Therefore we approximate each of the individual object images to follow an M-component mixture distribution with the supposition that the feature

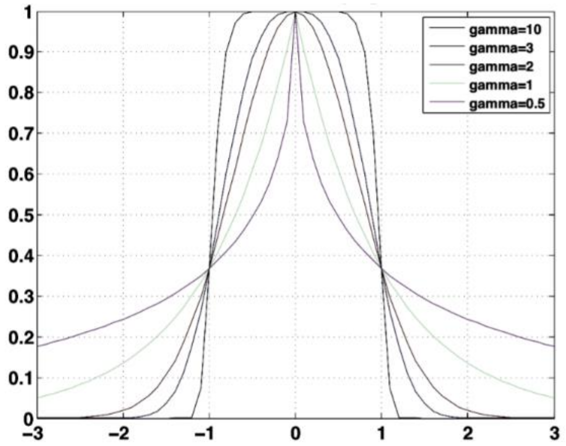


Fig. 3. ***The Plot of GGM PDF with shape parameters [22]***

vectors in each one of these object image regions exhibit the proposed mixture model. Object image may contain several subimages and each subimage assumed to have similar features following M-component multivariate generalized Gaussian mixture model. The cumulative probability density function of each one of the object image regions under consideration will be of the form as shown in Fig 3. In order to acquire the dynamic features of image, the model based on MGGMM will be more suitable. This model by default sets the lower and upper ranges of the image being acquired, and any information beyond this range will be automatically discarded. In dynamic scenarios, the video running for a period of time will be considered and the average number of pixels will be identified. These average values are assumed to be the background values and the static values are mapped against foreground values.

B. Histogram of Oriented Gradients

HOG can be employed for detecting objects from the given image. The steps for HOG descriptor model is as follows:

1. Image is divided into small cells. In each cell, calculate a histogram of gradient directions.
2. Discretize every cell into angular bins as per the direction of the gradient.
3. Every cell contributes a weighted gradient.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

4. Each adjacent cell is called as block, cells are to be grouped into blocks for normalization and grouping of histograms.
5. Histogram block is denoted by the cluster of histograms and each histogram is normalized. The descriptor is denoted by group of histogram blocks.

For the given input image I, 2-D gradient is calculated for each pixel of the image [15]. The detection window of *64x64* pixels is separated into non-overlapping cells. Each cell considered to be of *8x8* pixels size. The block is made by 4 adjacent cells. One cell is overlapped in a block both horizontally and vertically, and leading to 7*15 = 105 blocks for a specific detection window. Block normalization is performed by using the L2-norm as given below.

$$\text{L2-norm: } f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}} \tag{1}$$

C. Obstacle Detection Process

This paper proposes a methodology which employs HOG and GGMM to identify the vehicles and track the vehicle information based on the acquired information. Considering this information to the model there by the output of the model generates the signal through the Android mobile phone. GGMM is used because of its ability to track the background more reliably and its capability in detecting the objects during changes in illumination in the outside world. In this methodology, every pixel of the video is considered and is given as an input to the GGMM. In this proposed methodology the background and foreground information is considered based on standard deviation (SD) and weight factors. The SD of the pixels are considered and the pixels having a higher SD than threshold values are considered as background pixels and low SD are considered as foreground pixels. Each of the pixels is then considered and is grouped into either a background or foreground image. Every pixel categorized as the background is given a color black and foreground pixels are assigned a color white or given the value 3. The unmatched pixels are considered as foreground and are given as label 3. During the process, if any holes or missing information is obtained, these holes are filled with morphological techniques such are erosion and dilution.

In order to identify the features, we have considered the HOG, histogram of the oriented gradient. The HOG is captured and generally, HOG's are used to characterize the shape and appearance of the object and hence it serves as a descriptor for the object identification. In order to identify the HOG, the gradient is considered which points the highest scalar value among the edge pixels. Proposed methodology and workflow are shown in Figures 4 and 5 respectively.

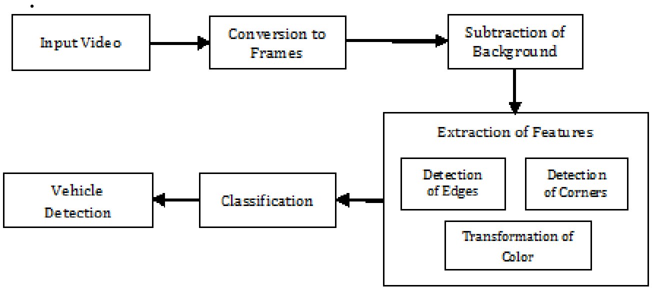


Fig. 4. Methodology

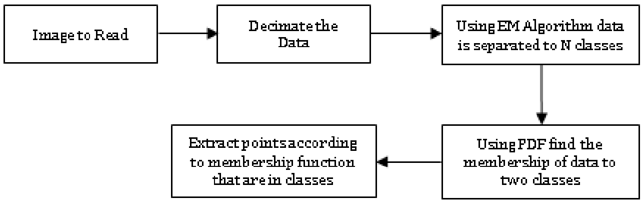


Fig. 5. ***The flow of the work process***

Fig. 6. ***Inputs of Object***

Fig. 7. Segmented Outputs of using Object GGMM

IV. RESULTS

The experimentations are carried on the Windows 10 OS with 3.00 GHz 4-core processor having higher than 2 GB RAM. To investigate the efficiency and accuracy of the suggested method, experimentations are performed under diverse conditions. The results are compared with GMM algorithm. GMM algorithm barely identifies the moving things, and it gives incorrect detection of objects in the result. GMM algorithm cannot segment the whole object region, and it does not have the sensitivity for the low-speed object. Different inputs of objects and vehicles that are considered in the experimentation are shown in Fig 6 and Fig 9 respectively. Segmented outputs of objects using GGMM are shown in Fig 7. Object detection using the sensors stick & mobile phone is shown in Fig 8. Segmented outputs of the vehicle using GGMM and the corresponding vehicle detection are shown in Fig. 10 and Fig. 11 respectively.

A. Precision-Recall (PR) curves

The precision and recall are computed using standard formulas which in turn uses true positive, false positive and false negative. Comparative analysis in terms of Precision and Recall are demonstrated and shown in Fig. 12 and Table 1 respectively. The evaluation is further carried out using statistical metrics like kappa metric, Entropy, and Figure of Merit. The above-said metrics are presented below [23].

B. Cohen's Kappa

Kappa value is calculated using the following formula:

$$Kappa = \frac{(a-b)}{(1-b)} \tag{2}$$

Where 'a' is the detected pixels of the processed image and 'b' is the likelihood of getting pixel by chance. Kappa is normalized from 0 to 1. Closer to 1 indicates better segmentation of the image.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

C. Entropy

Entropy is estimated using,

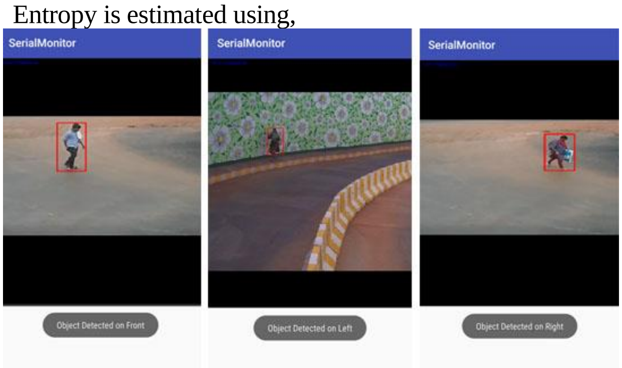


Fig. 8. *Object detection using sensors stick & mobile phone*

Fig. 9. *Inputs of Vehicle*

Fig. 10. *Segmented Outputs of using Vehicle GGMM*

$$E = -\sum_{i=1}^{a} e_i \log e_i \quad (3)$$

Here 'e' signifies number of pixels and 'i' signifies the pixel intensity. A lower value of entropy indicates good segmentation as the degree of randomness would be less and would thus result in better segmentation.

D. Figure of Merit

The figure of Merit is computed using,

$$R = \frac{1}{I_R}\sum_{i=1}^{I_E}\frac{1}{1+\beta d^2} \quad (4)$$

Where

$I_R = \max(I_A, I_E)$

I_A = total no. of absolute of edge points.

I_E = total no. of true edge points.

d = dispersion of true edge points from absolute edge points.

β = fixed scaling value.

Fig. 11. *Vehicle detection using sensors stick & mobile phone*

Higher values of the figure of merit indicate optimal segmentation and in general, R-value lies between 0 and 1.

In order to acquire the dynamic features of the image, the model based on MGGMM will be more suitable. This model by default sets the lower and upper ranges of the image being acquired, and any information beyond this range will be automatically discarded. In dynamic scenarios, the video running for a period of time will be considered and the average number of pixels will be identified. These average values are assumed to be the background values and the static values are mapped against foreground values. Compared the proposed methodology with other existing SVM and GMM based models based on the parameters such as entropy metric and figure of merit, which are illustrated in Table 2 and Fig. 13, 14 & 15 respectively.

V. CONCLUSION AND FUTURE WORK

The authors present an enhanced moving object recognition algorithm grounded on Generalized Gaussian mixture model and using the HOG Features in this paper. The suggested algorithm can spontaneously choose the total components for every pixel. This amendment intensely increases the precision of background subtraction and convergence while retaining the sequential adaptability. Apart from this, it preserves the comprehensive edge information of the moving object by means of edge detection technology. A complete study of experimental outcomes illustrates that this algorithm can identify moving objects in the complex scenes effectively and has good robustness. A comparison of the derived outputs using SVM and using the metrics like Precision, Recall, Kappa, Entropy and figure of merit has also been presented

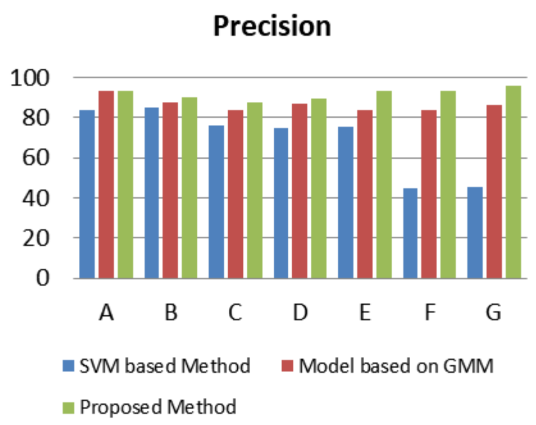
Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

Table 1: Analysis with respect to recall and precision

Vehicles	SVM based method		Model based on GMM		Proposed Method	
	Precision	Recall	Precision	Recall	Precision	Recall
A	83.55	93.55	93.45	93.55	93.44	95.90
B	84.75	93.56	87.67	93.35	89.83	94.85
C	75.74	86.58	83.45	89.47	87.53	91.77
D	7458	87.49	86.56	93.85	89.73	94.98
E	75.55	88.47	83.55	93.55	93.35	96.89
F	44.55	84.35	83.47	95.85	93.54	94.98
G	45.54	88.75	86.48	93.38	95.49	93.98

Table 2: Comparision of SVM based model, GMM based model and the proposed methodology, in terms of the kappa, entropy metric and figure of merit

	SVM BASED			Model based on GMM			Proposed Methodology		
Vehicle	Kappa	Entropy	Figure of Merit	Kappa	Entropy	Figure of Merit	Kappa	Entropy	Figure of Merit
A	0.74	0.91	0.65	0.78	0.86	0.74	0.89	0.76	0.87
B	0.71	0.8	0.7	0.74	0.81	0.73	0.83	0.66	0.84
C	0.64	0.79	0.69	0.69	0.84	0.78	0.82	0.71	0.8
D	0.69	0.87	0.71	0.76	0.9	0.78	0.78	0.69	0.82
E	0.76	0.85	0.41	0.81	0.89	0.7	0.87	0.78	0.89
F	0.79	0.88	0.66	0.84	0.88	0.71	0.85	0.79	0.86
G	0.68	0.79	0.63	0.74	0.82	0.7	0.89	0.72	0.84



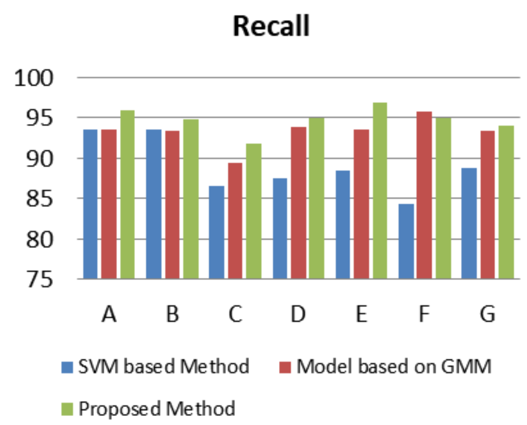


Fig. 12. Analysis with respect to recall and precision in graphical form

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

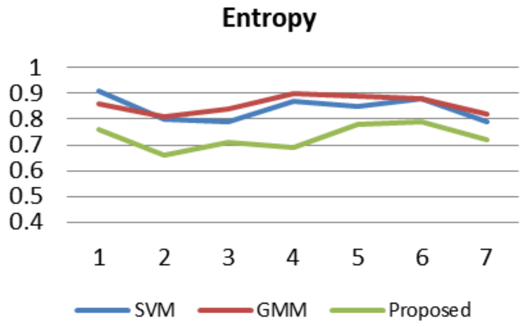


Fig. 13 ***Entropy Metric***

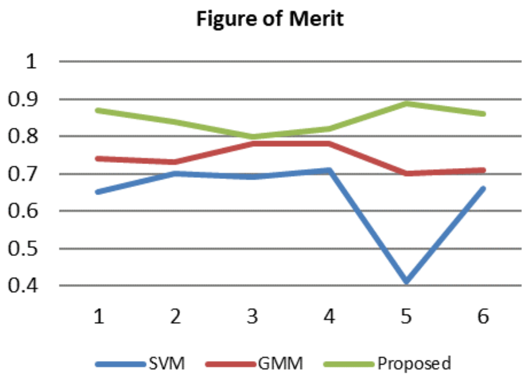


Fig. 14 ***Figure of Merit***

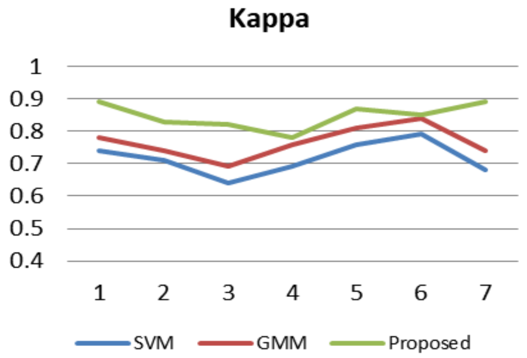


Fig. 15 ***Kappa***

REFERENCES

1. Bourne, R.; Flaxman, S.; Braithwaite, T.; Cicinelli, M.; Das, A.; Jonas, J.; Keeffe, J.; Kempen, J.; Leasher, J.; Limburg, H.; et al.. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. Lancet Global Health 2017, 5, e888–e897.
2. Ponnada, S.; Yarramalle, S.; MadhusudhanaRaoT., V. A Hybrid Approach for Identification of Manhole and Staircase to Assist Visually Challenged. IEEE Access 2018, 6, 41013–41022.
3. Paul.; Garg.; Singh, M.; Mehra.; Manocha. Smart Cane for the Visually Impaired : Technological Solutions for Detecting Knee-above Obstacles and Accessing Public Buses. 2007.
4. Penrod, W.; Corbett, M.D.; Blasch, B. A Master Trainer Class for Professionals in Teaching the Ultracane Electronic Travel Device. Journal of Visual Impairment and Blindness 2005, 99, 711–714.
5. Hoyle, B.; Waters, D., Mobility AT: The Batcane (UltraCane). In Assistive Technology for Visually Impaired and Blind People; Hersh, M.A.; Johnson, M.A., Eds.; Springer London: London, 2008; pp. 209–229. doi:10.1007/978-1-84628-867-8_6.
6. Chanana, P.; Paul, R.; Balakrishnan, M.; Rao, P. Assistive technology solutions for aiding travel of pedestrians with visual impairment. Journal of Rehabilitation and Assistive Technologies Engineering 2017, 4, 1–16. doi:10.1177/2055668317725993.
7. Bourbakis, N.; Makrogiannis, S.K.; Dakopoulos, D. A System-Prototype Representing 3D Space via Alternative-Sensing for Visually Impaired Navigation. IEEE Sensors Journal 2013, 13, 2535–2547. doi:10.1109/JSEN.2013.2253092.
8. Joseph, S.L.; Zhang, X.; Dryanovski, I.; Xiao, J.; Yi, C.; Tian, Y. Semantic Indoor Navigation with a Blind-User Oriented Augmented Reality. 2013 IEEE International Conference on Systems, Man, and Cybernetics, 2013, pp. 3585–3591. doi:10.1109/SMC.2013.611.
9. Hernández-Aceituno, J.; Arnay, R.; Toledo, J.; Acosta, L. Using Kinect on an Autonomous Vehicle for Outdoors Obstacle Detection. IEEE Sensors Journal 2016, 16, 3603–3610. doi:10.1109/JSEN.2016.2531122.
10. Bai, J.; Liu, D.; Su, G.; Fu, Z. A Cloud and Vision-based Navigation System Used for Blind People. Proceedings of the 2017 International Conference on Artificial Intelligence, Automation and Control Technologies; ACM: New York, NY, USA, 2017; AIACT '17, pp. 22:1–22:6. doi:10.1145/3080845.3080867.
11. Choi, J. Realtime On-Road Vehicle Detection with Optical Flows and Haar-Like Feature Detectors. Computer science research and tech reports, University of Illinois at Urbana-Champaign, 2012.
12. Haselhoff, A.; Kummert, A. A vehicle detection system based on Haar and Triangle features. 2009 IEEE Intelligent Vehicles Symposium, 2009, pp. 261–266. doi:10.1109/IVS.2009.5164288.
13. Bai, H.; Wu, J.; Liu, C. Motion and haar-like features based vehicle detection. 2006 12th International Multi-Media Modelling Conference, 2006, pp. 4 pp.–. doi:10.1109/MMMC.2006.1651348.
14. Juan, L.; Gwun, O. A Comparison of SIFT, PCA-SIFT and SURF. International Journal of Image Processing 2009, 3, 143–152.
15. Hsieh, J.; Chen, L.; Chen, D. Symmetrical SURF and Its Applications to Vehicle Detection and Vehicle Make and Model Recognition. IEEE Transactions on Intelligent Transportation Systems 2014, 15, 6–20. doi:10.1109/TITS.2013.2294646.
16. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, Vol. 1, pp. 886–893 vol. 1. doi:10.1109/CVPR.2005.177.
17. Bi, H.; Tang, H.; Yang, G.; Shu, H.; Dillenseger, J.L. Accurate image segmentation using Gaussian mixture model with saliency map. Pattern Analysis and Applications 2018. doi:10.1007/s10044-017-0672-1.
18. Pan, H.; Yi, C.; Tian, Y. A primary travelling assistant system of bus detection and recognition for visually impaired people. 2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), 2013, pp. 1–6. doi:10.1109/ICMEW.2013.6618346.
19. Elguebaly, T.; Bouguila, N. Finite asymmetric generalized Gaussian mixture models learning for infrared object detection. Computer Vision and Image Understanding 2013, 117, 1659 – 1671. doi:https://doi.org/10.1016/j.cviu.2013.07.007.
20. Allili, M.S.; Bouguila, N.; Ziou, D. Finite Generalized Gaussian Mixture Modeling and Applications to Image and Video Foreground Segmentation. Fourth Canadian Conference on Computer and Robot Vision (CRV '07), 2007, pp. 183–190. doi:10.1109/CRV.2007.33.
21. Elguebaly, T.; Bouguila, N. Bayesian Learning of Generalized Gaussian Mixture Models on Biomedical Images. Proceedings of the 4th IAPR TC3 Conference on Artificial Neural Networks in Pattern Recognition; Springer-Verlag: Berlin, Heidelberg, 2010; ANNPR'10, pp. 207–218. doi:10.1007/978-3-642-12159-3_19.
22. M. A. Saad; A. C. Bovik and C. Charrier, "Blind Image Quality Assessment: A Natural Scene Statistics Approach in the DCT Domain," in IEEE Transactions on Image Processing, vol. 21, no. 8, pp. 3339-3352, Aug. 2012. doi: 10.1109/TIP.2012.2191563
23. Arie, B.; A lot of randomness is hiding in accuracy. Engineering Applications of Artifical Intelligence., Vol. 20 (7), pp. 875-887, 2007. Doi:10.1016/j.engappai.2007.01.001

AUTHORS PROFILE

SREENU PONNADA is currently working as an Assistant Professor in the Department of Computer Science and Engineering at Vishnu Institute of Technology, Bhimavaram, India. He holds a B.Tech degree in Computer Science Engineering from Maharaj Vijayaram Gajapathi Raj College of Engineering (affl. JNTUK), Vizianagaram, India; and M.Tech in Computer Science Engineering from Avanthi's St.

Theressa Institute of Engineering and Technology (affl.JNTUK), Vizianagaram, India. He is currently pursuing Ph.D. in Computer Science Engineering from GITAM University, Visakhapatnam. With over three years of teaching and research experience, Sreenu has contributed to more than seven international and national publications in the areas of artificial intelligence and mobile networks. He is particularly interested image processing and assistive technologies involving both hardware and software components to assist the visually challenged in their daily activities.

DR. Y. SRINIVAS is currently working as a Professor in the Department of Information Technology at GITAM University, Visakhapatnam, India. He received his PhD in Computer Science with Specialization in Image Processing from Acharya Nagarjuna University, Guntur, India. He has 22 years of teaching and research experience. He has guided 18 students for Ph.D. and currently guiding eight scholars for Ph.D. His current research interests are in the areas of image processing, knowledge discovery and data mining, computer vision and image analysis. He has published more than fifty papers in various reputed international journals and conference.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)