SUCCESSIVE IMAGE RECOGNITION WITH CANNY EDGE DETECTION AND ORB ON RGB AND DEPTH IMAGES

Anchit Mulye[1], Jayesh Gangrade[2]

[1] *Student, Computer Science and Engineering, Institute of Engineering & Science IPS Academy, M.P., India*

[2] *Associate Professor, Computer Science and Engineering, Institute of Engineering & Science IPS Academy, M.P., India*

ABSTRACT

Image recognition with multi-level of information extraction is termed as Successive image recognition. The proposed method detects the edges using Canny edge detection algorithm then uses ORB to extract the vital features from the image. The RGB and Depth images are used to recognize images using the proposed new method. The results are also compared with SIFT, the obtained result depicts that the proposed method outruns SIFT in accuracy and time domain. There is an increase in accuracy by 15.46% and decrease in time by 4.08%, of recognizing Depth images by the proposed method. The same method when processed on RGB images produces an increase 6.87% in accuracy.

Keywords: *Successive Image recognition, Canny Edge Detection, ORB, SIFT, Depth Images*

1. INTRODUCTION

Images are computerized rendition of various entities in form of pixels. In only one year, 2017 more than 1.2 trillion images were captured only using smartphones and cameras [1]. This keen exponential growth in digital media has created the data in the vast amount which will still pursue for the coming years. To extract information from these prodigious dataset of images we need to move towards computer vision or machine vision.

It is facile for a human eye to absorb information from an image. The problem becomes arduous when we move from a human to computer. Over the years the evolution in deep learning has resulted in good results in case of object detection. Detecting sundry objects in a single image is a successfully deciphered problem. The next step is to recognize the detected object, things get exceptionally complex in this stage. The recognized object can look like many entities making it very strenuous for computer to identify it exactly.

This paper is centralized on the ways to improve the contemporaneous methods available in image recognition by blending the existing methods with some different methods to gain the optimized results. The main pivot is to improve the overall accuracy in image recognition. The method discussed also reduces the time required and can also be used with low power machines unlike deep learning methods which requires high end configured machines to perform. The method is implemented on RGB and depth images respectively and the improvements in each case are discussed accordingly. The method is also compared with SIFT (Scale Invariant Feature Transform) [2], which is far most used algorithm in image recognition. The problem with SIFT is that, this algorithm is patented by their creators making this quite inconvenient for commercial use. The community is moving for other substitutes than the present SIFT method. In this paper, the method proposed uses ORB (Oriented FAST and rotated BRIEF) with Canny Edge detection for recognition of various objects in RGB and Depth images. Multiclass dataset is used for both RGB and Depth images with 10 classes in the respective category.

Depth images produces better results in image recognition as they are invariant to lighting conditions. It is quite easy to detect the object in focus in depth images which increases the efficiency of object recognition. The present major problem in image recognition is the relative size of the object present in the image. It is difficult for machine to judge a car in the image is the real one or just a toy. The problem of relative size is also addressable with depth images [3]. The depth sensor cameras are also becoming more and more convenient to use. In the field of image recognition the color of image hardly plays any role in recognizing images. The only parameter which is of utmost importance is the depth. Depth can successfully predict the distance of the object, the size of the object and the motion of the object [4]. The two algorithms used for feature extractions from the image are SIFT and ORB. The extracted features are then classified using k-NN algorithm.

The paper is organized in the following sections. In section 2, discussion about already present methods. In Section 3, visual description of the Dataset used and how the images are classified. In section 4, there is in-depth description of the proposed method. In section 5, the results are compared and analyzed. The paper is concluded in section 5 with possible future suggestions where this method can be used.

2. LITERATURE REVIEW

The multiple researches conducted in the area of computer vision using various algorithms are quite beneficial. Handling images is such a difficult task that no method is solely sufficient for every image. There are few algorithms which perform a bit better than others under specified conditions. This section covers the comparison of algorithms in past researches.

2.1 Canny Vs Prewitt

Canny edge detection is used to detect the edges in the image. The major advantage of using this algorithm is that significantly reduces the amount of data to be processed without compromising with important information [5]. Prewitt edge detection is another algorithm which calculate the gradient of the intensity in an image. The sudden change in the intensity is used to detect the edges in the image [6]. The previous research conducted clearly depicts that Canny surpasses Prewitt in accuracy and time [7]. This provides a solid ground on our approach to take image recognition to a new level.

2.2 Canny Vs Sobel

Sobel edge detection algorithm is another algorithm which is very similar to Prewitt. It takes approximation of the gradient vector to detect the edges [8]. Canny edge detection surpasses the Sobel by huge margin in accuracy whereas both are nearly same in time domain [9]. The complete analysis on past work clearly states that the Canny Edge detection algorithm is the most efficient in various scenarios.

2.3 ORB Vs SIFT

The comparison between two different feature extraction methods is tedious task. Each algorithm is designed differently keeping various scenarios in mind. It nearly impossible to design a single algorithm which focuses on all the elements in any image. Studies have shown that SIFT produces much more features than compared to ORB [10]. These two algorithms calculate quite varying features in the same image. ORB is not yet patented like SIFT, this gives a greater uphold on the reliability of this algorithm.

2.4 RGB Vs Depth

RGB images carries information regarding the color of any particular entity. The major disadvantage is that color is not very good parameter when it comes to image recognition. Moreover most of the algorithms works on Grey scale images. The information supplied in the form of RGB images can’t be utilized to its maximum by feature extraction algorithms. On the other hand the Depth images contains vital information about how distant a particular object is which tells the object in focus. This information results out to be very helpful for feature extraction algorithms. The evident results from many researches tell about the effectiveness of depth images over RGB images [11]. With new technology depth cameras are getting more convenient and powerful. This open an entire new era in computer vision for future work.

3. DATASET

3.1 RGB

In RGB images each pixel is decoded by combination of three colors namely Red, Green, Blue. With the combination of only these color all other colors can be made. In this paper RGB images are used from Caltech-256 Dataset. The dataset used in this is made of different classes with each class containing 200 images in total. Then each class is later divided into training and testing images of different sizes and then studied accordingly. There are total 10 different classes of different images. The number of classes increases the level of difficulty for training of images. Fig -1 shows the sample images from each class.

Fig -1 Classes in RBG Dataset

3.2 Depth

The depth images in used in this paper are all captured using Microsoft Kinect Depth sensor. The images are captured in various different environments. The images are different hand gesture which are captured in multiple angles. The entire dataset is of 10 classes of 200 images each. The depth images are then converted into edged images to build a copy of the entire dataset. Then the two datasets are trained and compared accordingly. The depth images from each class is shown in Fig-2.

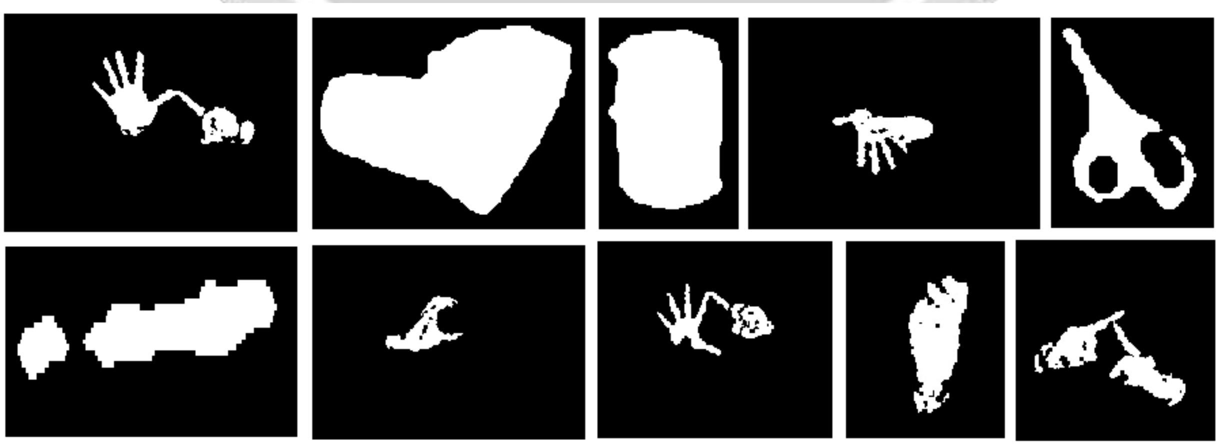


Fig-2 Classes in Depth Dataset

4. METHODOLOGY

This paper presents a new approach to detect the images in computer vision. The images from two different datasets namely RGB and Depth are all processed using Canny Edge detection algorithm. These images are differently classified in separate directory. The images from respective dataset are then processed for extracting features using SIFT and ORB. The features extracted are then used to create histogram of extracted features. Then k-NN classification algorithm is applied to classify the features using supervised machine learning. The detailed process is covered in respective sections.

4.1 Edge Detection

In this paper the extensive focus in on Canny Edge Detection method as this is one the finest method available to detect edges in various images. Canny algorithm works by applying Gaussian filter to remove the noise from images. Then it calculates the intensity gradients of the image. Two step threshold is applied to determine the potential edges in the images. Finally the hysteresis is used to target only the main edges, eliminating the less potential edges [5]. The major advantage of Canny algorithm is that it reduces the data to be processed and this was clearly observed while extracting features from the images. Fig-3 shows the sample of RGB dataset after detecting the edges. Fig-4 shows the sample of each class from Depth dataset after detecting edges.

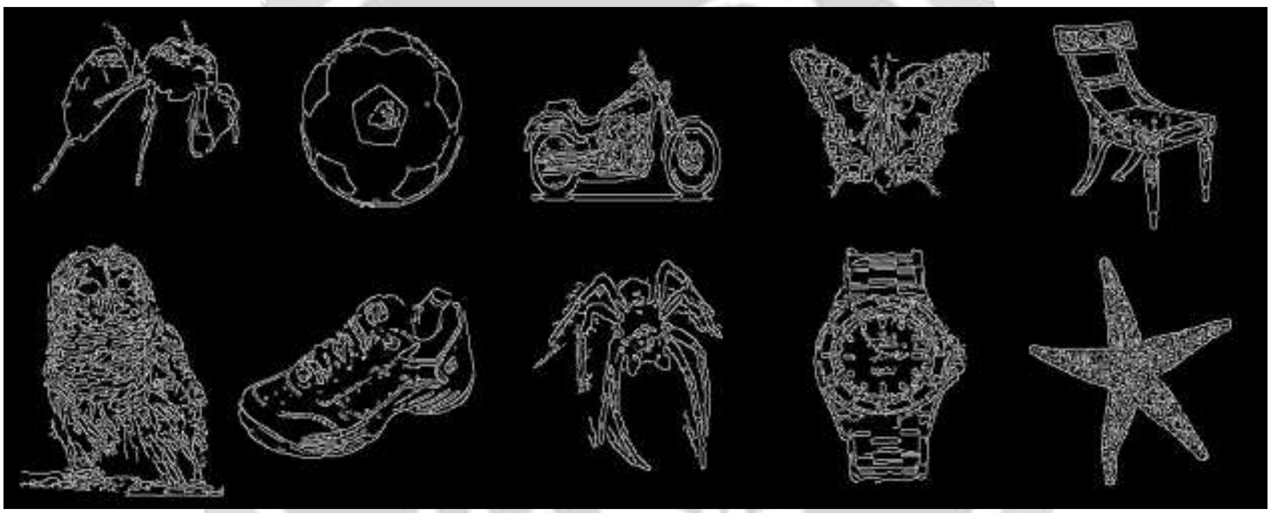


Fig -3 RBG images after edge detection

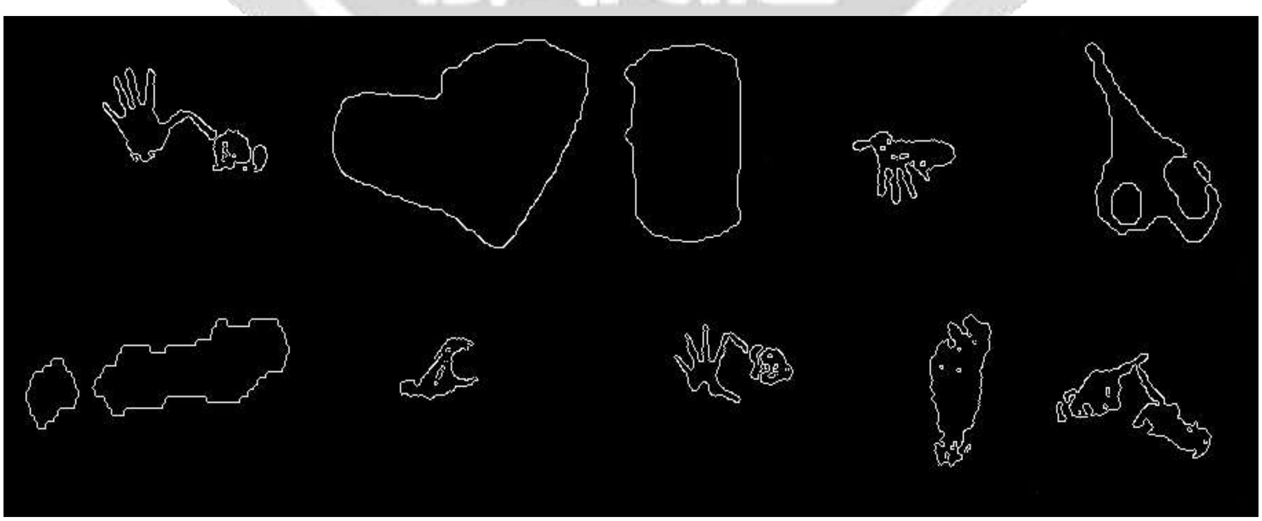


Fig -4 Depth images after edge detection

4.2. Feature Extraction

Feature extraction can be done using either shape descriptor or feature descriptor. In this paper we use two feature descriptor methods to achieve the task. The important keypoints are located in any image by these algorithms. The number of keypoints in any image varies a lot, even if the image belongs to same class. To normalize the collected features, bag of visual words method is used [12]. This the slowest process in entire the method as the collected features are processed by algorithm to extract the vital information.

4.2.1 SIFT

The SIFT algorithm is known for its scale and rotation invariance as it's named, Scale Invariant Feature Transform. The algorithm produces fine results even there is huge change in the angle of images. If one two images is completely upright then too the algorithm works accurately. Once we have detected any object from image the chances of the object in completely upright position are very rare. Therefore this feature is not of quite importance when we have successfully detected the object. The number of features extracted by SIFT are quite more compared to other algorithms which makes it quite slower to produce results. The number of features extracted also reduces when an edged image is provided for the particular same normal image of that type. The method proposed it much better for computer vision as there no significant loss with recognition. The Fig -5 shows the difference in the features extracted on normal image and edged image in both RGB and Depth dataset. The shown sample image can be clearly observed to analyze to difference in the extracted features in two cases. It is clearly noticed that edged images reduces the vital keypoints to be processed.

4.2.2 ORB

This algorithm is made by combining two different algorithms and was presented as an alternative to SIFT [13]. The name Oriented FAST and rotates BRIEF only suggest the areas of dominance of this algorithm. ORB is much faster than SIFT and it also calculates much accurate features in case of Depth images. As the number of features calculated are much less, it also makes it much faster. The method proposed in this paper also reduces the number of keypoints extracted and also minimizes the time required. The Fig -6 shows the comparison of features extracted in edged image and normal RGB image.

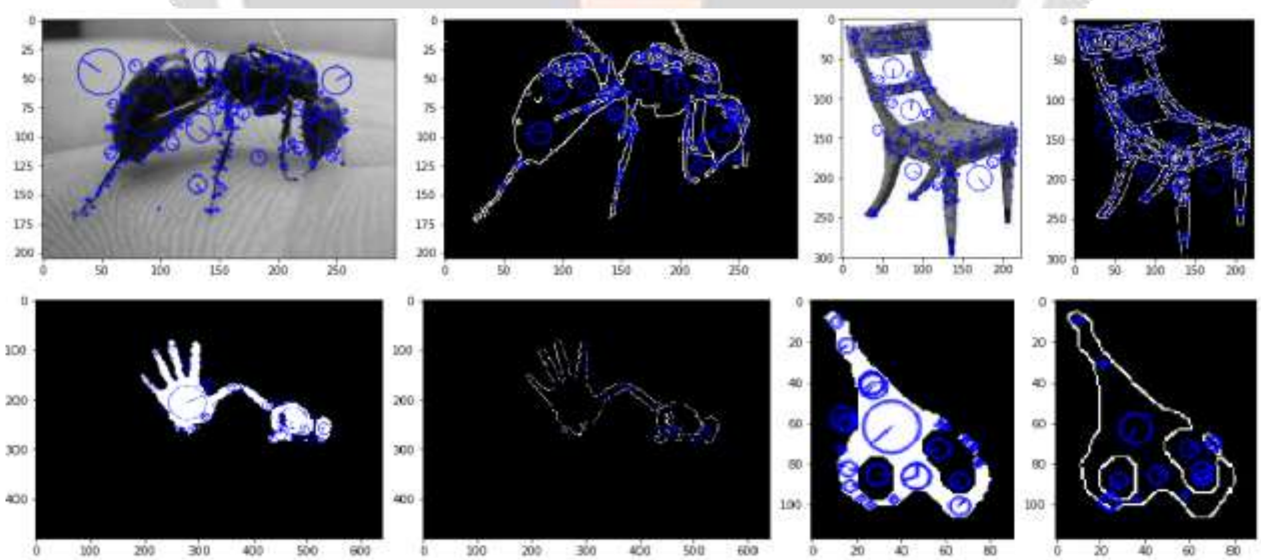


Fig -5 Features extracted using SIFT

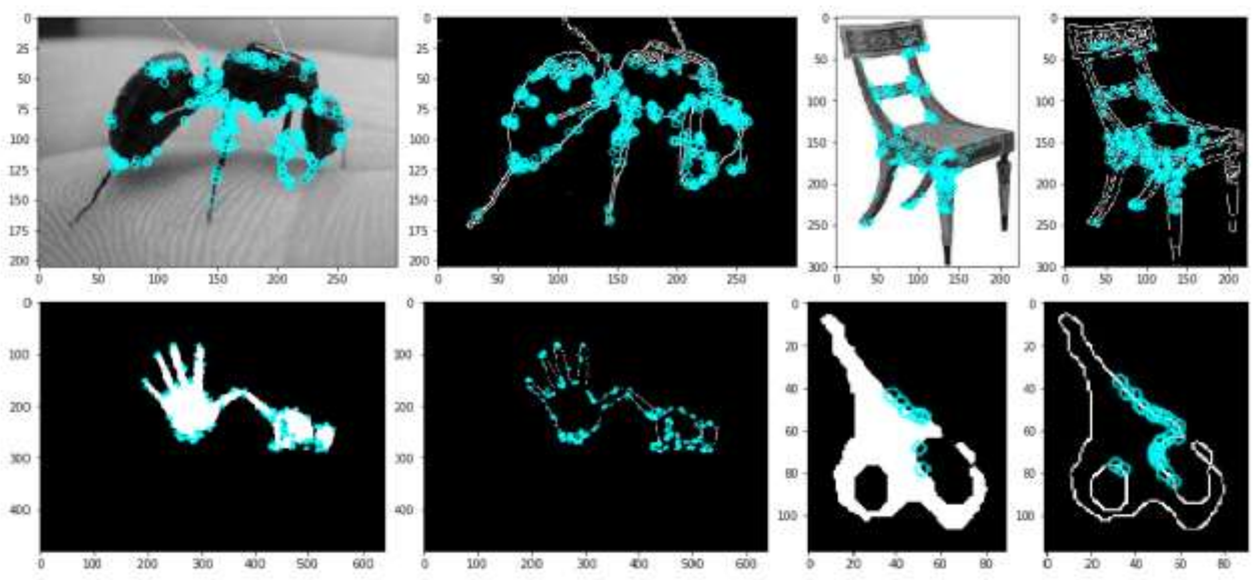


Fig -6 Features extracted using ORB

4.3. Classification

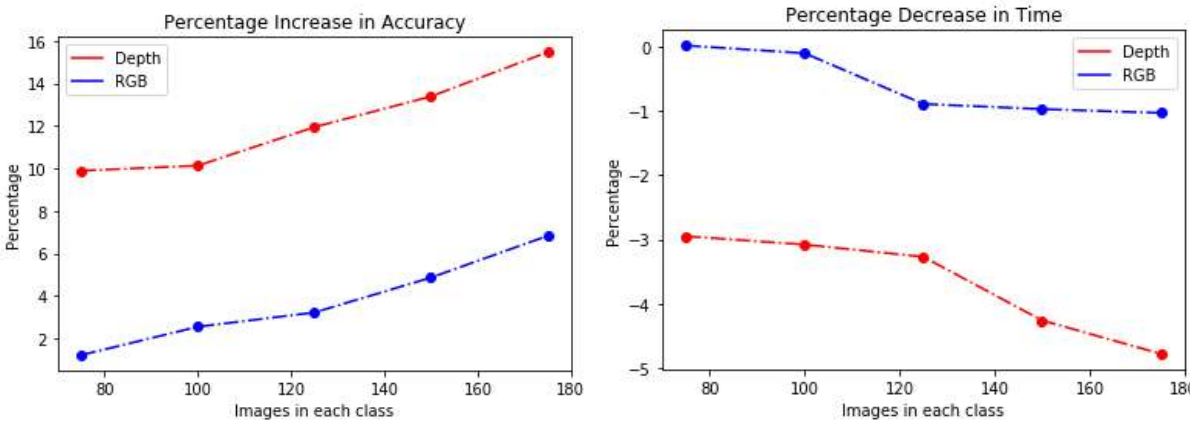
The process of distinguishing the data plays a dominant role in machine learning and computer vision. There are many algorithms which works best in this domain. There is no single algorithm which works in all scenarios. The further thing required is to set the parameters in the classification algorithms to gain the optimum result. The algorithm used to classify is k-NN which non parametric in nature [14]. The external parameter provided to this algorithm is k which is number of nearest neighbors to be referred to classify a data point.

5. RESULT AND ANALYSIS

The method proposed produces fine and constant results on continuous iteration. The results are classified in two sections for RGB and Depth images. The success rate of the proposed method is more favorable in Depth images as there is substantial increase in accuracy. The RGB images also show increase in efficiency but due to more noise the increase is not steep. The processing of image after edge detection also reduces the overall training time. Similarly the decrease in training time is more prompt for Depth images than RGB images. The table-1 shows the numerically calculated results. Fig 7 shows the visual representation of how the accuracy increases and time decreases for RGB and Depth images.

Table- 1 Results for RGB and Depth images

Images	**Accuracy**		**Time**	
	Percentage	Remark	Percentage	Remark
RGB	6.87%	Increase	1.03%	Almost Same
Depth	15.46%	Increase	4.08%	Decrease



(a) **(b)**

Fig -7 Percentage change in the results

The results obtained are compared on the ground of results from simple image recognition using SIFT. The results for the comparison are unbiasedly calculated using the same dataset and constraining all the other parameters. The Canny algorithm performs quite well with Depth images as compared to RGB image. The difference is more sudden when Canny algorithm is applied to RGB dataset. There is high decrease in number of features calculated and also high increase in accuracy. The noise is clearly suppressed by Canny algorithm with RGB. Although the Depth images give more accurate results with our method. The increase in accuracy with this method can be very helpful. The smallest of increase can change the game in computer vision. This increase using the proposed noble method is quite beneficial in other domains.

6. CONCLUSION

This paper covers a new approach to increase the efficiency in image recognition. The proposed method suggest to apply edge detection before feature extraction. This proposed method increases the efficiency, reduces the time and also reduces the computational complexity. The edge detection is very famous method which has lost its way in present time with neural networks. The image recognition is moving on another level with unsupervised recognition gaining more attention. The advantage of edge detection can also be utilized with unsupervised learning. The detection of edges is important as there are very few exceptions where two completely different objects produces same edge. The edge is itself a unique feature of an entity in any image.

6. REFERENCES

[1]. Statistical data from infotrends. http://blog.infotrends.com/our-best-photos-deserve-to-be-printed

[2]. Lowe D. G., "Distinctive image features from scale-invariant keypoints" in International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, Nov. 2004

[3]. A. Anwer, A. Baig and R. Nawaz, "Calculating real world object dimensions from Kinect RGB-D image using dynamic resolution," 2015 12th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, 2015, pp. 198-203.doi: 10.1109/IBCAST.2015.7058504

[4]. J. Lee , Leyvand T., Peeper C., "Motion detection using depth images" U.S. application Ser. No. 13/410,546

[5]. J. Canny, "A Computational Approach to Edge Detection," in IEEE Transactions on Pattern Analysis and Machine Intelligence, *vol. PAMI-8, no. 6, pp. 679-698, Nov. 1986. doi: 10.*1109/TPAMI.1986.4767851

[6]. J. M. S. Prewitt, “Object enhancement and extraction,” Picture Processing and Psychopictorics, B. Lipkin and A. Rosenfeld, Eds., New York: Academic Press, 1970, pp. 75-149.

[7]. Nisha, Rajesh Mehra, Lalita Sharma, “Comparative Analysis of Canny and Prewitt Edge Detection Techniques used in Image Processing”, International Journal of Engineering Trends and Technology (IJETT) – Volume 28 Number 1 - October 2015

[8]. Sobel, Irwin. An Isotropic 3x3 Image Gradient Operator. Presentation at Stanford A.I. Project 1968

[9]. Dr.S.Vijayarani, Mrs.M.Vinupriya, “Performance Analysis of Canny and Sobel Edge”, Detection Algorithms in Image Mining”, International Journal of Innovative Research in Computer and Communication Engineering Vol. 1, Issue 8, October 2013

[10]. Andersson, O., & Reyna Marquez, S. (2016). A comparison of object detection algorithms using unmanipulated testing images : Comparing SIFT, KAZE, AKAZE and ORB (Dissertation). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-186503

[11]. C. Jing, J. Potgieter, F. Noble and R. Wang, "A comparison and analysis of RGB-D cameras' depth performance for robotics application," 2017 24th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), Auckland, 2017, pp. 1-6. doi: 10.1109/M2VIP.2017.8211432

[12]. Gabriella C., Christopher R., Lixin Fan D., Williamowski Ju., Bray C. (2004) Visual Categorization with Bags of Keyopints. In workshop on Statistical Learning in Computer Vision, ECCV.

[13]. E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," *2011* International Conference on Computer Vision, Barcelona, 2011, pp. 2564-2571. doi: 10.1109/ICCV.2011.6126544

[14]. Altman, N.S.. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. American Statistician - AMER STATIST. 46. 175-185. 10.1080/00031305.1992.10475879.