RGB-D Images for Object Segmentation, Localization and Recognition in Indoor Scenes using Feature Descriptor and Hough Voting

Abrar Ahmed
Dept. of Computer Science
Air University
Islamabad, Pakistan
180159@mail.au.edu.pk

Ahmad Jalal
Dept. of Computer Science
Air University
Islamabad, Pakistan
ahmadjalal@mail.au.edu.pk

Kibum Kim
Dept. of Human-Computer Interaction
Hanyang University
Seoul, South Korea
kibum@hanyang.ac.kr

***Abstract*— With the development in technology, object localization and recognition systems in RGB and depth have become the essential part of vision systems. There are different ways of objects localization and recognition which segment the whole scene in parts. RGB-D and depth information is best way for the extraction of objects of interest. Researchers and scientists from all our world, continuously trying to improve the vision and detection property of computer systems for improving the people lives. Till now, the computers can sense using sensors and get involved in interesting conversations but lack the capability of understanding a real scene like humans. In our system, we allow the machines to detect and recognize the indoor RGB-d scenes using novel methodology. Inspired by the significance of human-computer interaction, we have presented a technique to recognize and localize the multiple objects present in RGB-D indoor scenes taken from RGB-D object dataset. We have proposed Saliency map based RGB-D object segmentation along with the multiple features and Hough voting to evaluate the performance of our proposed system with other proposed systems. The proposed system should be used in autonomous driving systems, security systems, violence detection, traffic monitoring, games, defense and sports scenes.**

Keywords*— RGB-D object segmentation, saliency map, RGB-D object recognition, Hough Voting, point cloud, feature extraction and matching.

I. INTRODUCTION

Object recognition and their uses in real life has become an interesting task in the field of image processing and computer vision as it is basic task in many commercial and industrial applications. In recent years, it has gained a lot of interests. A variety of systems have been developed in this literature but still there is lot of challenges in this field because of large variation in colors and illumination [1], different view-points [2], resolution, different textures and occlusion [3] in the real-world application. Most existing object detection and recognition [4] systems use visual features for objects matching and localization, but objects are captured in different conditions and view-points. Many objects have similar texture and color. Many real-world applications are affected, and their performance is not satisfactory with these challenges.

In recent years, RGB_D object recognition and localization [5] systems have been proposed to handle these challenges. These systems used the depth information [6] provided by high quality depth cameras such as Microsoft Kinect [7]. Depth information of RGB-D images is robust to resolution, view-points and illumination variation. A lot of Scientists are working for the generation of RGB-D datasets [8] and experimental results on RGB-D datasets show better experimental results, when depth information is combined with visual features of RGB images. Existing RGB-D object recognition and localization system focused on learning features or feature descriptors for recognition of objects. Feature extraction and matching [9] system achieved high performance over object recognition.

In this literature, different systems have been proposed for the RGB-D object recognition and localization. Some system performs object localization using depth map, such as [10] use depth map for building the CRF model and developed a system for scene understand using indoor scenes. [11] improved performance of detection using 3D features. High accuracy is achieved by [12] using 3D features. Image classification is performed by [13], using kernel features. Object detection and localization in 3D space for RGB-D images was performed by [14]. Foreground object segmentation of RGB-D images is performed by [15]. In [16], 3D structure features are extracted by using the point feature histograms. Bo et al. [17] proposed a system that extract combination of feature such as depth edges, 3D shapes and size feature and achieved remarkable performance over RGB-D datasets. Socher et al. [18] developed a system that learn depth and color channel separately then merge them for achieving final classifier. His model combines recursive neural network (RNN) and convolution filters. Unique shape context [19] descriptor was proposed by Tombari et al. that use local reference frame (RF), which greatly expands the accuracy and robustness of 3DSC descriptor. But the dimensions size of USC descriptor is very high such as 1960 and very high compotation cost.

As mentioned above, there are different method related to salience segmentation, feature extraction and matching and object recognition and localization but these tasks used by some of them efficiently. In this paper, we have proposed an well-organized system for localization and recognition of RGB-D objects. Our proposed system is divided in four phases. Initially, Segmentation of RGB-D images is achieved by the fusion of saliency map and centered darker channel. Secondly, segmented object is converted in point cloud and features are extracted by using Histogram of Oriented Gradients (HOG). Thirdly, Nearest Neighbour Search (NNS) algorithm is applied for feature matching and object detection. Finally, Hough Voting (HV) algorithm is applied for the recognition of objects. Accuracy of proposed object localization and recognition system in complex scenes is validated by experimental results.

978-1-7281-4676-8/20/$31.00©2020 IEEE

The remainder of this paper is settled as follows: Section II show an detailed overview of the system architecture of object recognition system over RGB-D dataset, which includes foreground object segmentation, feature extraction, feature matching and object recognition. Experimental results and their comparison with other state-of-the-art object recognition systems are shown in Section III. Finally, contributions of this paper concluded in section IV.

II. System Methodology

The detail and flow of our RGB-D object recognition system is detailed in this section. Fig.1 shows an overview of flow diagram of our proposed system. Initially, segmentation of RGB-D images is achieved by the fusion of centered darker channel and saliency map of RGB-D images. After segmentation, foreground objects are detected and separated from background objects. Secondly, the segmented images are passed through feature extraction step, then object of interest (OOI) [20] is detected in the feature matching step. Finally, recognition of OOI is achieved by using the Hough voting algorithm.

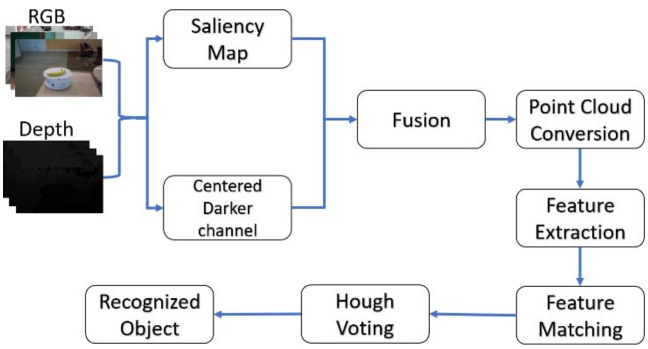


Fig. 1. Flow diagram of our proposed system.

A. Saliency Map Generation

To generate the saliency map [21], we extract the color feature of RGB image and depth feature of depth images then fused them for the better results. Fig. 2 shows the saliency map of RGB-D images. The feature of all components such as R, G and B are color components [22] and D is depth in RGB-D images. For the features in RGB-D, we use quaternion [23] to presents feature of a pixel at (x, y) as

$$F(x, y) = \omega_1 F_1 + \omega_2 F_2 i + \omega_3 F_3 j + \omega_4 F_4 k \quad (1)$$

where ω_1, ω_2, ω_3 and ω_4 are the weights of each four feature i, j and k are imaginary numbers and

$$F_1 = D, F_2 = \frac{R+G+B}{3}, F_3 = R_1 - G_1, F_4 = B_1 - Y_1 \quad (2)$$

where r, g and are the componens of RGB image and D is depth.

$$R_1 = R - \frac{B+G}{2}, G_1 = G - \frac{R+B}{2}, B_1 = B - \frac{R+G}{2}$$
(3)

$$Y_1 = \frac{R+G}{2} - \frac{|R-G|}{2} - B$$
(4)

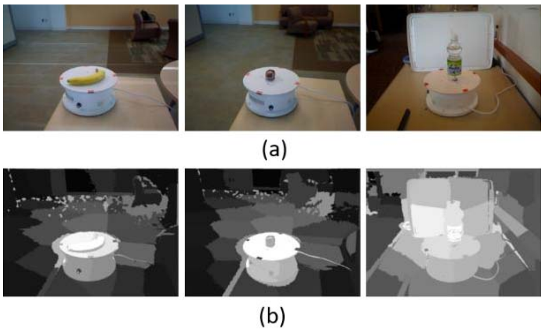


Fig. 2. Saliency map of RGB-D images (a) RGB-D images and (b) Saliency maps

B. Centered Darker Channel

We implement the central darker channel [24] by using two tasks such as dark channel detection [25] and center saliency map [26]. We achieved more accurate results by the fusion of these two tasks.

1) Dark channel detection

Dark channel detection is task that is used in haze removal field [27]. The dark channel identifies the most haze [28] containing region and improve it. As the dark channel work on low intensities [29], shadow containing images and colorful images, it ignores high intensities in images. We distinguish the foreground and background [30] of image by estimation of transmission such identifies the salient objects by detection of low intensities and dark colors in images. We extract low intensities foreground objects [31]. Dark channels are shown in Fig. 3.

Fig. 3. Examples of centered darker channels of RGB-D

2) center saliency map

The salient objects are always in the center of the images taken by different cameras. we use the BSCA algorithm [32] for achieving the center saliency map of images. Fig. 4 depicts the center saliency map of images. It works on color dissimilarities in images and distance matrix based on boundaries. Saliency map accurately reserves the salient objects and regions present in the images by removing all the surroundings.

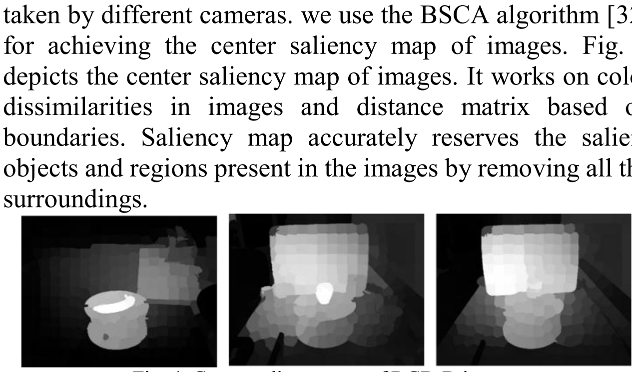


Fig. 4. Center saliency map of RGB-D images

C. Segmentation by fusion

This section figure-out the segmented object by the fusion of saliency map and centered darker object. Visual results of RGB-D object segmentation are shown in Fig. 5. We combined the depth cues [33] from saliency map and centered darker channels for the improved RGB-D objects segmentation [34].

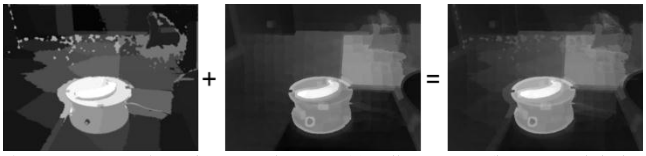
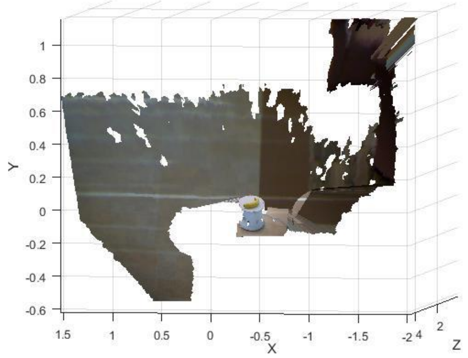


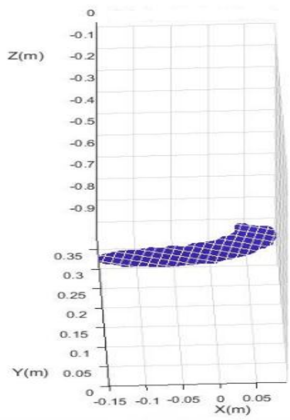
Fig. 5. Segmentation of RGB-D images, (a) saliency map (b) centered darker channel and (c) fused segmented image.

D. Point cloud conversion

After improved segmentation of RGB-D images, 3D point clouds [35] are produced of images for the further recognition process of the system such as feature extraction, feature matching and object recognition. It can greatly reduce the decrease computational cost. Fig. 6 depicts the point cloud representation of RGB-D image.



(a)



(b)

Fig. 6. Point cloud of images (a) point cloud of RGB-G image and (b) point cloud of segmented image.

Then, Nearest Neighbor (NN) [36] resampling is used for down sampling [37] of point cloud for increasing the computational efficiency [38] of the system, because point cloud is represented by pattern of a point and its neighboring points. It preserves the original values of images and did not alter them. Down sampling is achieved as

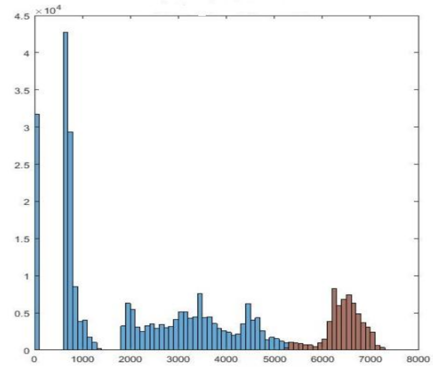
$$k(x) = \begin{cases} 1; \; if \; |x| < 0.5 \\ 0; \; otherwise \end{cases} \tag{5}$$

E. Feature extraction from point cloud

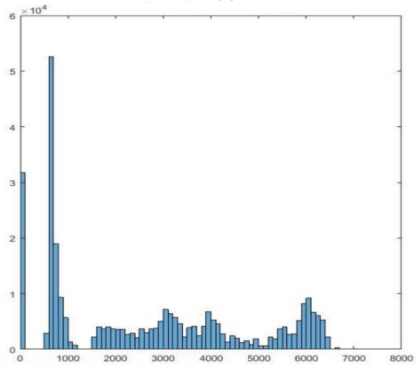
Although point cloud give a fine local representation of object class from dataset, yet some confusions about the similar shap object classes from dataset such as apple class and ball class. To give flexibility to our proposed system, we have extracted local features [39] from all classes of dataset such as point to point and point to plane features. We use Histogram of Oriented Gradients (HOG) [40] for the feature extraction from the point clouds of segmented object classes of dataset. Computational cost [41] greatly reduced, when we use neighboring point for producing these feature. We found points, which minimum distance to their neighboring points. Each point show different features descriper [42] such as histogram. These feature varies point to point in the same image. But these are similar in different images for same points. Fig. 7 depicts the histogram of distance from a point to all neighboring pixels and distance from point to plane. Fig. 8 show the point cloud of distance from a point to all neighboring pixels and distance from point to plane. Connectivity of point clouds C_{pc} of objects can be verified as

$$C_{pc}(pc_1, pc_2) = \begin{cases} Yes; \; if \; |d_{pc}| < t \\ No; \; otherwise \end{cases} \tag{6}$$

Where pc_1, pc_2 are point cloud 1 and 2 respectively, d_{pc} is distance between point clouds and t is threshold.



(a)



(b)

Fig. 7. Distance histogram using HOG (a) point to point distance histogram and (b) point to plane distance histogram.

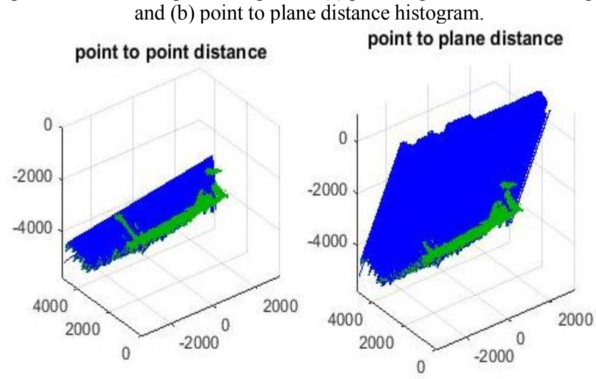


Fig. 8. Point cloud representation of feature, left is point cloud of point to point and right is point to plane distance

F. Feature matching

After proposed features algorithm extracts the points with minimum distance to their neighboring points, descriptor matching [43] is applied for the detection of OOI from the RGB-D images. Using a feature descriptor each point can be described such as histogram. We can match 2 descriptors based on shortest distance between them such as Euclidian distance [44]. To match the feature descriptor extracted for 2 images, we used Nearest Neighbor Search (NNS) [45] algorithm.

G. *Hough voting for object recognition*

After 3D matching between segmented object model and current scene image, Hough voting [46] is used to filtering the outliers and recognition of OOI using a process of voting. Fig. 9 shows the final recognized objects in RGB-D complex scenes. Hough voting identifies the correspondence between current scene and the OOI model. The purpose of proposed Hough voting is to calculate the feature vote [47] to insure the presence of OOI by using corresponding. Then, we made a vote vector of the correspondence between the OOI model and the current scene image. The feature point which shows maximum vote of correspondence is center of OOI model and those feature points show minimum vote in the vector of correspondence are the outliers [48]. If a feature is matched in OOI and current scene image, then it cost a vote in vote vector. Initially, we started the voting process from reference point of OOI model. We have selected centroid [49] of model as reference point R. Vector $Vec_{k,J}$between reference point and all feature point F_k calculated as

$$Vec_{k,J} = R - F_k \tag{7}$$

We calculated voting in Local Frame Reference, which is rotation [50] and translation invariant [51]. We identified the correspondence of all feature point of OOI model with the scene image and detect the presence of OOI by thresholding the peaks of Hough voting space.

Fig. 9. Object recognition using Hough voting over RGB-D objects scenes.

III. Experimental Results

This section contains dataset description, experimental results, recognition accuracy and a comparison of our approach with other state-of-the-art [52] object recognition systems over RGB-D.

A. *RGB-D Object Dataset*

RGB-D dataset [53] have multiple and different object classes form indoor scenes. Dataset has different and complex scene images. Working over RGB-D was really challenging. We have performed experiments over 10 classes of RGB-D object such as *apple*, *ball*, *banana*, stapler, water-*bottle*, *cell phone*, *shampoo, lemon, plate* and *tomato*. The images in the dataset have dimension of 640*480. We divided 70% of the dataset for training; while rest 30% is used for testing of our proposed system. The proposed object segmentation and recognition system was implemented using point cloud and we have performed experiments in Matlab on a PC Intel(R) Core-i3 having a Ram of 6 GB along with 2.5 CPU. Fig. 10 presents some examples classes of the RGB-D object dataset. Table I shows the RGB-D object segmentation results and Table II represents the confusion matrix of RGB-D object recognition.

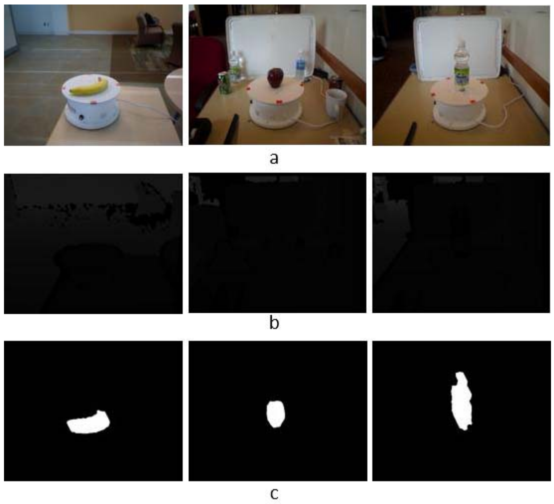


Fig. 10. RGB-D object dataset and recognition results (a) RGB image (b) Depth image (c) recognized objects

B. *Experimental Results*

For the evaluation the performance of the proposed system, we performed experiments over RGB-D benchmarked dataset. In our system, we used simple features for object detection as compare to state-of-the-art [54] method. Our approach decreases the computational cost for the recognition of objects in RGB-D complex scenes. In the experiments, class labels [55] are compared with predicted class [56] identifiers from the dataset to determine recognition results over RGB-D dataset as shown in Table II and.

TABLE I. MEAN ACCURACY OF RGB-D OBJECT SEGMENTATION

Class Name	**AP**	**BL**	**BA**	**ST**	**BO**
Accuracy (%)	84.1	79.6	78.4	82.2	86.7
Class Name	**CE**	**SH**	**TO**	**LE**	**PL**
Accuracy (%)	78.3	86.1	83.3	83.5	80.3
Accuracy of RGB-D Object Segmentation = 82.15%					

AP=apple; BL = ball; BA = banana; ST = stapler; BO = water-bottle; CE =cell-phone; SH =shampoo; TO =tomato; LE=lemon; and PL =plate class of the msrc dataset.

TABLE II. CONFUSION MATRIX OF OBJECT RECOGNITION OVER RGB-D DATASET.

Object Classes	AP	BL	BA	ST	BO	CE	SH	TO	LE	PL
AP	**.90**	0.0	.06	0.0	0.0	0.0	.04	0.0	0.0	0.0
BL	0.0	**.88**	0.0	.03	.03	0.0	0.0	0.0	.06	0.0
BA	.05	0.0	**.87**	0.0	0.0	.02	.04	0.0	.02	0.0
ST	0.0	.03	0.0	**.88**	.02	0.0	0.0	.03	.04	0.0
BO	0.3	.04	0.0	.03	**.86**	0.0	0.0	0.0	.04	0.0

CE	.03	0.0	.03	.01	0.0	**.89**	.03	0.0	0.0	0.0
SH	.04	0.0	.05	0.0	.03	0.0	**.86**	0.0	.02	0.0
TO	.02	0.0	.02	0.0	.03	0.0	.02	**.84**	0.0	.06
LE	0.2	0.4	0.0	.02	.04	0.0	0.4	0.0	**.84**	0.4
PL	.02	0.0	.01	0.0	0.0	0.0	.03	.07	0.0	**.87**

Finally, Table IV shows the performance comparison of proposed sysytem for ODR with other state of the art methods over RGB-D object dataset.

TABLE III. PERFORMANCE COMPARISON OF OBJECT LOCALIZATION RECOGNITION ACCURACY ON RGB-D DATASET WITH DIFFEENT METHODS

Methods	**Accuracy of Object Recognition (%)**
Rand. Forest [57]	79.6
Kernel SVM [57]	83.9
DKM [58]	80.88
HKDES [59]	86.03
Our approach	**86.90**

IV. CONCLUSION

This paper presents a simple and novel approach for localization and recognition of complex objects from RGB-D object dataset for providing necessary intelligence to the computers. Point clouds and feature matching are present high computational efficiency. The targets the domains of proposed object recognition are object tracking event recognition, security, scene recognition and medical images. The performance of our proposed object detection approach along with features and Hough voting is evaluated on RGB-D object dataset. As comparison with other state of the art systems, experimental results of our system show good results over benchmarked RGB-D dataset.

In future, we will target scene understanding in depth images using other classifiers with some more complex datasets of scenes.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2018R1D1A1A02085645).

REFERENCES

[1] B. Martinkauppi, S Maricor and P Matti, “Detection of skin color under changing illumination: a comparative study,” *in Proc. of IAP*, pp. 652-657, 2003.

[2] A. Jalal and S. Kim, S, “Advanced performance achievement using multi-algorithmic approach of video transcoder for low bit rate wireless communication,” *ICGST International Journal on graphics, vision and image processing*, Vol. 5(9), pp. 27-32, 2005.

[3] A. Jalal and A. Shahzad, “Multiple facial feature detection using vertex-modeling structure,” *in Proceedings of the IEEE conference on Interactive computer aided learning*, pp. 1-7, 2007.

[4] M. Shehnaz, and N. Naveen, “An object recognition algorithm with structure-guided saliency detection and SVM classifier,” *in Proc. of PICC*, pp. 1-4, 2015.

[5] A. Jalal, and S. Kim, "The Mechanism of Edge Detection using the Block Matching Criteria for the Motion Estimation," *Proc. Human Computer Interaction*, pp.484-489, Jan. 2005.

[6] A. Jalal, and S. Kim, “A complexity removal in the floating point and rate control phenomenon,” *Proc. Korea multimedia society*, pp.48-51, 2005.

[7] A. Jalal, S. Kim and B. J. Yun, “Assembled algorithm in the real-time H.263 codec for advanced performance,” in *Proceedings of the IEEE workshop on enterprise networking and computing in healthcare industry*, pp.295-298, 2005.

[8] A. A. Rafique, A. Jalal, and A. Ahmed, “Scene Understanding and Recognition: Statistical Segmented Model using Geometrical Features and Gaussian Naïve Bayes”, *IEEE conference on International Conference on Applied and Engineering Mathematics*, 2019.

[9] A. Jalal, M. Mahmood and A. S. Hasan, “Multi-features descriptors for human activity tracking and recognition in Indoor-outdoor environments,” *IEEE International Conference on Applied Sciences and Technology*, 2019.

[10] A. Jalal and M. Mahmood, “Students’ Behavior Mining in E-learning Environment Using Cognitive Processes with Information Technologies,” *Education and Information Technologies*, *Springer*, 2019.

[11] A. Jalal and A. Nadeem and S. Bobasu, “Human body parts estimation and detection for physical sports movements,” *IEEE International Conference on Communication, Computing and Digital Systems*, 2019.

[12] S. Gupta, R. Girshick, P. Arbeláez, and J. Malik. “Learning rich features from RGB-D images for object detection and segmentation.” *in Proc. of ECCV*, pp. 345-360, 2014.

[13] A. Zeng, S. Song, M. Nießner, M. Fisher, J. Xiao, and T. Funkhouser, “3dmatch: Learning local geometric descriptors from rgb-d reconstructions,” *in Proc. of CVPR*, pp. 1802-1811. 2017.

[14] R. Palmer, G. West, and T. Tan, “Using depth to extend randomised hough forests for object detection and localisation,” *in Proc. of DICTA*, pp. 1-8, 2013.

[15] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, “Indoor segmentation and support inference from rgbd images,” *in Proc. of ECCV*, pp. 746-760, 2012.

[16] A. Zeng, S. Song, M. Nießner, M. Fisher, J. Xiao, and T. Funkhouser. “3dmatch: Learning local geometric descriptors from rgb-d reconstructions,” *in Proc. of CVPR*, pp. 1802-1811. 2017.

[17] C. Y. Tsai, and S. H. Tsai, “Simultaneous 3D object recognition and pose estimation based on RGB-D images,” *IEEE Access*, pp. 28859-28869, 2018.

[18] R. Socher, B. Huval, B. Bath, C. D. Manning, and A. Y. Ng, “Convolutional-recursive deep learning for 3d object classification,” *in Proc. of ANIPS*, pp. 656-664. 2012.

[19] F. Tombari, S. Salti, and L. D. Stefano, “Unique signatures of histograms for local surface description,” *in Proc. of ECCV*, pp. 356-369, 2010.

[20] A. Jalal, J. T. Kim, and T.-S. Kim, “Human activity recognition using the labeled depth body parts information of depth silhouettes,” *in Proc. of ISHB*, pp. 1-8, 2012.

[21] M. Batool, A. Jalal, and K. Kim, “Sensors Technologies for Human Activity Analysis Based on SVM Optimized by PSO Algorithm”, *IEEE ICAEM conference*, 2019.

[22] A. Jalal, and M. A. Zeb, “Security and QoS Optimization for distributed real time environment,” *In: Proc. of the IEEE conference on Computer and Information Technology*, pp. 369-374, 2007.

[23] Y. Kong, Y. Fu, “Close Human Interaction Recognition Using Patch-Aware Models,” *IEEE Transactions on Image Processing*, vol. 25, pp. 167-178, 2016.

[24] N. Cho, S. Park, J. Park, U. Park and S. Lee, “Compositional interaction descriptor for human interaction recognition,” *Neurocomputing*, vol. 267, pp. 169-181, 2017.

[25] K. Houda, Yannick, and Feriel, “Human interaction recognition based on the co-occurrence of visual words,” *in Proceedings of CVPR Conference*, pp. 455–460, 2014.

[26] A Jalal, Md. Zia Uddin and T.-S. Kim, “Depth Video-based Human Activity Recognition System Using Translation and Scaling Invariant Features for Life Logging at Smart Home”, *IEEE Transaction on Consumer Electronics*, vol. 58, pp. 863-871, 2012.

[27] A. Jalal, Y. Kim, and D. Kim, “Ridge body parts features for human pose estimation and recognition from RGB-D video data,” *in Proceedings of the IEEE International ICCCNT*, pp. 1-6, 2014.

[28] A. Jalal, M. Mahmood, and M. A. Sidduqi, “Robust spatio-temporal features for human interaction recognition via artificial neural network, *IEEE Conference on FIT*, 2018.

[29] A. Jalal, M. A. K. Quaid and M. A. Sidduqi , “A Triaxial acceleration-based human motion detection for ambient smart home system,” *IEEE Conference on Applied Sciences and Technology*, 2019.

[30] A. Jalal, “IM-DailyDepthActivity dataset,” imlab.postech.ac.kr/ databases.htm, 2015, [Online; accessed 14 April- 2015].

[31] A. Jalal and S. Kamal, “Real-Time Life Logging via a Depth Silhouette-based Human Activity Recognition System for Smart Home Services,” *in Proc. AVSS*, pp. 74-80, 2014.

[32] A. Jalal, MZ. Uddin, J. T. Kim, and T. Kim, “Daily Human Activity Recognition Using Depth Silhouettes and R Transformation for Smart Home,” *in Proc. ICOST*, pp. 25-32, 2011.

[33] A. Sony, K. Ajith, K. Thomas, T. Thomas, and P. L. Deepa, “Video summarization by clustering using euclidean distance,” *in Proc. Conference on Signal Processing, Communication, Computing and Networking Technologies*, 2011.

[34] A. Jalal, S. Kamal, and D. Kim, “A depth video sensor-based life-logging human activity recognition system for elderly care in smart indoor environments,” *Sensors*, vol. 14(7), pp. 11735-11759, 2014.

[35] A. Jalal, S. Kamal, and D. Kim, “Shape and motion features approach for activity tracking and recognition from Kinect video camera,” *in Proc. IEEE AINA*, pp. 445-450, 2015.

[36] A. Jalal, M. A. K. Quaid, and A. S. Hasan, “Wearable Sensor- Based Human Behavior Understanding and Recognition in Daily Life for Smart Environments, *Proc. IEEE conference on International Conference on Frontiers of information technology*, 2018.

[37] A. Jalal and S. Kim, “Algorithmic implementation and efficiency maintenance of real-time environment using low-bitrate wireless communication,” *in Proc. IEEE STFEUS*, 2006.

[38] A. Jalal, S. Kamal, A. Farooq and D. Kim, “A spatiotemporal motion variation features extraction approach for human tracking and pose based action recognition,” *in Proc. IEEE ICIEV*, 2015.

[39] A. Farooq, A. Jalal and S. Kamal, “Dense RGB-D Map-Based Human Tracking and Activity Recognition using Skin Joints Features and Self-Organizing Map ,” *KSII*, vol. 9(5), pp. 1856-1869, 2015.

[40] A. Jalal, S. Kamal, and D. Kim, “Depth Map-based Human Activity Tracking and Recognition Using Body Joints Features and Self-Organized Map,” *in Proc. IEEE ICCCNT*, pp. 1-6, 2014.

[41] A. Jalal, S. Kamal and D. Kim, “Depth Silhouettes Context: A new robust feature for human tracking and activity recognition based on embedded HMMs,” *in Proc. IEEE URAI*, pp. 294-299, 2015.

[42] A. Jalal and Y. Kim, “Dense Depth Maps-based Human Pose Tracking and Recognition in Dynamic Scenes Using Ridge Data,” *in Proc. of the IEEE AVSS*, pp. 119-124, 2014.

[43] A. Jalal, S. Kamal and D. Kim, “Individual Detection-Tracking-Recognition using depth activity images,” *in Proc. IEEE URAI*, 2015.

[44] S. Kamal and A. Jalal, “A hybrid feature extraction approach for human detection, tracking and activity recognition using depth sensors, *AJSE*, vol. 41(3), pp. 1043-1051, 2016..

[45] A. Jalal, Y.-H. Kim, Y.-J. Kim, S. Kamal and D. Kim, “Robust human activity recognition from depth video using spatiotemporal multi-fused features, *Pattern recognition*, vol. 61, pp. 295-308, 2017.

[46] S. Kamal, A. Jalal and D. Kim, “Depth Images-based Human Detection, Tracking and Activity Recognition Using Spatiotemporal Features and Modified HMM, *JEET*, pp. 1921-1926, 2016.

[47] A. Jalal, S. Kamal and D. Kim, “Human depth sensors-based activity recognition using spatiotemporal features and hidden markov model for smart environments, *Journal of computer networks and communications*, vol. 2016, pp. 1-11, 2016.

[48] A. Jalal and S. Kamal, “Improved behavior monitoring and classification using cues parameters extraction from camera array images,” *International Journal of Interactive multimedia and Artificial Intelligence*, vol. 5(2), 2018.

[49] A. Jalal, S. Kamal and D. Kim, “Facial Expression recognition using 1D transform features and Hidden Markov Model, *Journal of Electrical Engineering & Technology*, vol. 12(4), 2017.

[50] A. Jalal, S. Kamal, and D. Kim, “A depth video-based human detection and activity recognition using multi-features and embedded hidden Markov models for health care monitoring systems,” *International Journal of Interactive multimedia and Artificial Intelligence*, vol. 4(4), pp. 54-62, 2017.

[51] F. Farooq, A. Jalal, and L. Zheng, “Facial expression recognition using hybrid features and self-organizing maps,” *in Proc. IEEE International Conference on multimedia and expo*, Jul. 2017.

[52] A. Jalal, S. Kamal and D. Kim, “Detecting Complex 3D Human Motions with Body Model Low-Rank Representation for Real-Time Smart Activity Monitoring System,” *KSII Transactions on Internet and Information Systems*, vol. 12, no. 3, pp. 1189-1204, 2018.

[53] M. Mahmood, A. Jalal, and H. Evans, “Facial Expression Recognition in Image Sequences Using 1D Transform and Gabor Wavelet Transform, *Proc. IEEE ICAEM*, 2018.

[54] A. Jalal, S. Kamal and C. Cecer, “Depth maps-based human segmentation and action recognition using full-body plus body color cues via recognizer engine, *JEET*, 2018.

[55] A. Jalal, Y. Kim, S. Kamal, A. Farooq and D. Kim, “Human daily activity recognition with joints plus body features representation using Kinect sensor,” *in Proc. IEEE ICIEV*, pp: 1-6, 2015.

[56] M. H. Khan, F. Li, M. S. Farid, M. Grzegorzek, “Gait recognition using motion trajectory analysis”, *International Conference on Computer Recognition Systems*, pp. 73-82, May 2017.

[57] K. Lai, L. Bo, X. Ren, D. Fox, “A large-scale hierarchical multiview rgb-d object dataset,” *in Proc. of* ICRA, pp. 1817-1824, 2011.

[58] L.F. Bo, X.F. Ren, D. Fox, “Depth kernel descriptors for object recognition,” *in Proc. of ICIRS*, pp. 821, 2011.

[59] B. Liefeng, K. Lai, X. Ren, and D. Fox, “Object Recognition with Hierarchical Kernel Descriptors,” *in Proc. of CVPR*, 2011.