Region and Decision Tree-based segmentations for Multi-objects detection and classification in Outdoor Scenes

Abrar Ahmed
Dept. of Computer Science
Air University
Islamabad, Pakistan
180159@students.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*— Accurate segmentation and detection of all mixed and occluded objects in the complex indoor/outdoor scenes become a vital topic of computer vision. These above-mentioned scenarios are the significant part of important vision applications such as autonomous driving, traffic monitoring, security surveillance, humane body parts detection, objects tracking and scene recognition. It is still difficult to accurately detect all the objects in the image due to illumination changes, occlusion and different directions. Meanwhile, segmenting the image in parts helps to detect the multi objects accurately. In this paper, we designed a system having improved techniques for the accurate segmentation and detection of multi objects. Firstly, we have combined the results of two methods for accurate segmentation of multiple objects, (i) Decision trees for labeling every neighboring pixel and assigning a separate color to all object present in complex images and (ii) Region-based segmentation for significant detection of multiple regions and drawing boundaries of all objects. Secondly, detection is performed by searching the objects with previously assigned colors. Finally, we have performed labeling with class name to all objects present in the images. We have performed our experimental over two benchmarked datasets as Instance Saliency images and MSRC. Our experimental work has shown improved results with respect to other state of the art algorithms.**

Keywords*— Labeling, classification, object detection, region-based segmentation, decision tree, object recognition, preprocessing

I. Introduction

Object detection is way of identification of different real-world objects such as vehicles, animals, humans and places. Various image processing [1] techniques are used for object detection such color matching, feature extraction, neural network, template matching and correlation. Object detection is a vital spot of computer vision. It is most studied problem as it used in various applications such as security surveillance [2], medical field [3], defense, autonomous deriving [4] and maps. There is still a lot of complexities in the detection of real-world objects such as occlusion, change of illuminations and color intensities. For the optical solution of these complexities, scientists and researcher from all the world actively work for the object's detection and recognition.

Different techniques for the object detection and recognition in the complex scenes have been proposed by different researchers. For instance, a novel fusion approach is used to detect the objects proposed by T. Moh and S. Deshmukh [5]. They applied the object detection API, fusion detected edges and edge detection algorithm at original images. Their proposed approach has very high computational processing. SIFT [6] presents the nice characteristic of rotation and invariant to image but it needs accurate segmentation of objects. Different object segmentation methods [7] has proposed techniques to extract the objects from background but they are affected by environment ambiguities such as illumination change and occlusion.

In robotic application [8], local feature descriptor in real time processing is robust to rotation and scale change. A combined technique of RCNN and Bregion Regression method is proposed by L. Yu *et al*. [9]. Their proposed method shows good accuracy but needs of good replacement of candidate region size. In [10], they introduced Rol voting and multi task CNN to detect various object, however, this method is limited to occlusion. An approach 2D wavelet fractal feature is proposed by P. Zhang *et. al.* [11]. Object detection and recognition which extract the candidate regions proposed by Y. Makihara *et al*. [12]. This method achieved remarkable detection accuracy, but detection is achieved on a single place and cannot be used for real time detection.

In this paper, we have proposed an improved object detection approach that used Decision Tree-based segmentation and Region-based segmentation for the segmentation and detection of multiple object in complex scenes. The proposed system is divided into 4 phases. Firstly, images from both data sets are passed through preprocessing phase for smoothing and reducing noise from the images. Secondly, segmentation is performed using two techniques (i) Decision Tree and (ii) Region-based segmentation methods. Thirdly, combination of the results of these two techniques for improvement of multiple objects segmentation. Finally, color matching method is used for the detection and labeling of multiple objects. We have observed the performance results of proposed system over two publicly available datasets, which shows improved results with accuracy of 69.4% on Instance Saliency image and 73.4% of MSRC datasets. Fig. 1 represents the different classes of both datasets.

978-1-7281-6625-4/19/$31.00 ©2019 IEEE
DOI 10.1109/FIT47737.2019.00047

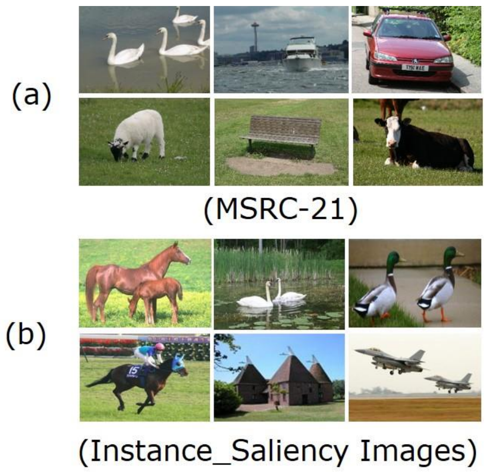

Fig. 1. Different classes of datasets (a) shows MSRC and (b) shows Instance Saliency Database

The detail of rest of the paper is as follows: Section II depict the system methodology of proposed system which contains the preprocessing, Decision Tree-based segmentation, Region-based segmentation, combination method, object detection and object labeling phase. Section III represents the analysis and comparison of the detection rate of our proposed system with other state of the art detection systems using the Instance Saliency image and MSRC datasets, respectively. Finally, Section IV shows the conclusion of paper.

II. System Methodology

We have proposed simple framework for the multiple objects detection and labeling with class names to all objects presents in the images. The proposed system architecture is divided into different phases. Initially, all the images of both datasets are passed through preprocessing step for the removal of noise and sharp edges. Secondly, segmentation of multiple objects is performed using two methods, (i) Decision Tree and (ii) Region-based. Third, results of Decision Tree and Region-based methods are combined in this phase for the improved results of segmentation. Lastly, multiple objects are detected on the base of colors assigned during the segmentation phase and class labels are assigned to all detected objects. Fig. 2 depicted the system architecture of multiple object detection.

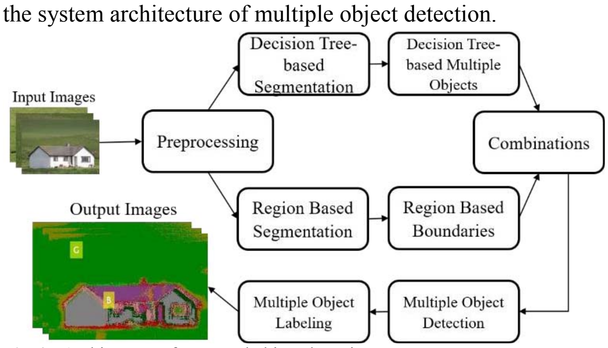

Fig. 2. Architecture of proposed object detection system

A. Image Preprocessing

During image preprocessing, color images from datasets are passed [13] through Gaussian filter for reducing the noise, sharp edge and smoothness of images.

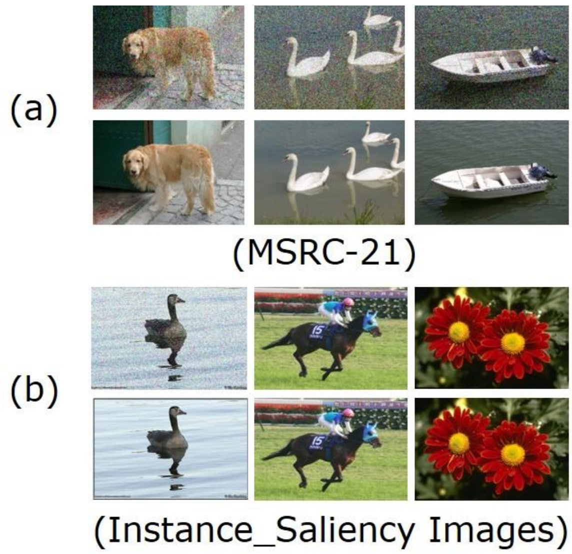

Fig. 3. Results of filtering steps (a) shows noisy images and (b) shows smooth and filtered images

Spatial calculation of gaussian filter is performed as;

$$g(\mathrm{a}, b) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\frac{a^2+b^2}{2\sigma^2}} \quad (1)$$

where a and b are the distance from the center in the horizontal axis and vertical axis respectivily. σ represents the standard deviation of Gaussain.

B. Decision Tree-based Segmentation

After filtering process, Decision Tree-based (DT) segmentation is applied on all images of datasets [14]. DT segmentation contains two steps (i) learning phase and (ii) predicting phase. In learning phase, DT splits image in small trees and every tree give labeling at leaf nodes. Splitting function is used for splitting the images in small patches. DT make and learn the *T* trees of decision on small patches and then gather the prediction. Textural and contextual information [15] of each tree is constructed based on histogram of image patches and region prior of bag of semantic patches (BOST) [16]. Each leaf node *l* of SF tree is associated with learned class distribution *P(class/l)*. While, decision tree *T* works at a given pixel *I* until reached at leaf node l_t. Class distribution [17] is produced by SF for all trees *T* over the leaf nodes $Leaf_i = (l_1, I. l_T)$ by averaging the learned class distribution.

$$P(class|Leaf_i) = \frac{1}{T}\sum_{t=1}^{T} P(class|l_t) \quad (2)$$

At each pixel *I*, the prediction is applied for the classification of all pixels. In this way, all classified pixels are obtained. Misclassified pixels become holes, which are cause of insufficient visual similarity [18]. Combining the DT results with similar regions [19] is a proposed solution to this problem (See Fig. 4).

Fig. 4. Results of Decision Tree-based segmentation

C. Region-based Segmentation

The J-segmentation algorithm [20] is used in Region-based Segmentation [21] for detecting the similar regions in images. In similar regions, all pixels are homogeneous in color. Different neighboring homogeneous regions [22] can be detected by extracting the colors of images. With the Region-based segmentation, we easily detect all pixels with its similar region and regions boundaries [23]. This method has two steps (i) color quantization [24] step and (ii) spatial segmentation [25] step. This segmentation method needs good color quantization.

In our proposed technique, we have implemented a color quantization algorithm. In color quantization step, color pixels of image are replaced by color class labels and color class map [26] is generated. After quantization, color information is stored in color map. All pixels in a class map have same color are assigned with a label. In spatial segmentation step, *J* values [27] is calculated over each pixel for producing the *J*-image. *J*-image pixel values are the *J*-values that calculated over the local window of these pixels. Size of image regions are determined by size of local window. Boundaries are detected by large size window and intensities localized by small size window. *J*-image is a grayscale image and is calculated as;

$$J = D_A/D_B = (D_M - D_B)/D_B \quad (3)$$

where D_A is distance between different classes, D_B is the distance between members within each class and D_M is distance between pixel and mean of all pixels. D_M is calculated as;

$$D_M = \sum_{z\in Z} |z - m|^2 \quad (4)$$

where Z is a set of all N pixels, $z \in Z$ *and* m is the mean calculated as;

$$m = \frac{1}{N}\sum_{z\in Z} z \quad (5)$$

Finally, Multiscale region growing [28] method is used for the segmentation of J-image. Region merging is performed on the base of color similarity of regions.

Fig. 5 depicts the region-based segmentation. Threshold parameter Th_J is used for region merging (See Fig. 5). Homogeneousness of similar regions is controlled by Th_J is calculated as;

$$Th_J = m_J + as_J \quad (6)$$

where m_J is mean, s_J is standard deviation and a is a parameter, whose value is (- 0.6, ...0.4).

Fig. 5. Boundaries detection using Region-based segmentation.

D. Combinig Technique

In this phase, we describe the method to combine [29] the Decision Tree and Region-based segmentation for improving the results. Recall that the results of DT at each pixel are the prediction and classification of pixels. Some pixel may have misclassified due to incorrect prediction. These misclassified pixels create holes in the regions. While, in Region-based segmentation, when threshold [30] $\boldsymbol{Th}$ is small, it produces the regions similar in color texture. So, the results of both methods are combined to generate the more accurate the similar regions, filling the holes [31] in the region and drawing the accurate boundaries of objects. Fig. 6 depicts the results of combining technique.

Fig. 6. Combination of Decision Tree and Region-based Segmentation.

E. Multiple Objects Detection

After the accurate segmentation of multiple objects, we observed the color similarity for the detection of multiple objects [32-35] present in the complex scenes. We matched the color labels of all classes assigned during the training of Decision Trees with images of datasets. After detecting [36] multiple objects accurately, bounding box [37-40] is drawn around the objects. Detected objects with bounding box around them is shown in Fig. 7.

Fig. 7. Detected objects using color matching.

F. Multiple Objects Labeling

After accurately detection of objects [41-44], multiple objects [45-49] are labeled [50-54] with the class names [55-57] mentioned in the datasets. Fig. 8 shows the multiple labeled objects.

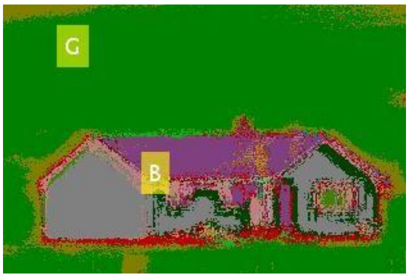

Fig. 8. Multiple labeled objects, B shows Building class and G shows grass class.

III. Performance Evaluation

Experimental results are evaluated on two publicly available benchmarked datasets (i) MSRC-21 [45] and (ii) Instance Saliency image database [46]. MSRC-21 dataset contains 591 RGB images and all classes are labeled such as *cow*, *car*, *sheep*, *building*, *plane*, *face*, *horse*, *tree*, *bicycle* and *grass*. Experiments are performed on 10 classes. Ground-truths of all images are available. The dimension of images is 213×320. All images are different from each other with different colors and different backgrounds. In addition, overall MSRC dataset is really challenging due to complex background and sharp color changes.

Instance Saliency image database is used for the evaluation of proposed method. Experiments are performed on 10 classes. Instance Saliency images database contains 999 RGB images and all classes are labeled such as *sea*, *bike*, *elephant*, *humane*, *cat*, *dog*, *duck*, *fish*, *ball* and *spoon*. Ground-truths of all images are available. All images are different from each other with different colors and different backgrounds. The dimension of images is 400×267. For the segmentation of images, we used Decision Tree and Region-based and color matching technique for the detection of multiple objects.

A. Experimental Settings and Results

We have performed experiments on MATLAB platform with a PC of Intell Core-i3 CPU 2.5 GHz having a RAM of 6 GB. In Table I, we present the mean accuracy of object segmentation on MSRC dataset which extract the multiple objects present in the images. While in Table II, we show the segmentation accuracy over Instance Saliency image database. We performed improved segmentation by combining the results two techniques. Table III depicts the detection accuracy over MSRC where we plot a bounding box around the detected objects. Table IV shows the experimental results of detection over Instance Saliency image database. While, Table V represents the comparison of results to other state of the art methods.

TABLE I. OBJECT SEGMENTATION ACCURACY OVER MSRC DATASET

Classes	CW	CR	SP	BG	PE	FR
Accuracy (%)	81.1	72.6	78.4	80.2	82.7	70
Classes	**FE**	**HE**	**TE**	**BE**	**GS**	**SN**
Accuracy (%)	71.3	78.1	75.3	69.5	70.3	77.2
Classes	**BK**	**CH**	**CT**	**DG**	**BY**	**BT**
Accuracy (%)	76.4	78.2	80	81	71.5	82
Mean Segmentation Accuracy = 73.4%						

where CW =Cow; CR = Car; SP = Sheep; BG = Building; PE = Plane; FE =Face; HE =Horse; TE =Tree; BE =Bicycle; FR=Flower; SN=Sign; BK=Book; CH=Chair; CT=Cat; DG=Dog; BY=Body; BT=Boat; and GS =Grass.

TABLE II. OBJECT SEGMENTATION ACCURACY OVER INSTANCE SALIENCY IMAGE DATASET

Classes	SA	CT	DG	HN	DC
Accuracy (%)	81.1	72.6	78.4	80.2	82.7
Object Classes	**FS**	**DR**	**HM**	**HR**	**BS**
Accuracy (%)	71.3	78.1	75.3	69.5	70.3
Mean Segmentation Accuracy = 69.4%					

where SA =SEA; CT = CAT; DG = DOG; HN = HEN; DC = DUCK; FS =FISH; DR =DEER; BS =BUS; HM =HUMANE; and HR =HORSE.

TABLE III. OBJECT DETECTION ACCURACY OVER MSRC DATASET

Classes	CW	CR	SP	BG	PE	FR
Accuracy (%)	81.1	72.6	78.4	80.2	82.7	74.1
Classes	**FE**	**HE**	**TE**	**BE**	**GS**	**SN**
Accuracy (%)	71.3	78.1	75.3	69.5	70.3	80.2
Classes	**BK**	**CH**	**CT**	**DG**	**BY**	**BT**
Accuracy (%)	79	70	82	77.7	82.3	77
Mean Object Detection Accuracy = 73.4%						

TABLE IV. OBJECT DETECTION ACCURACY OVER INSTANCE SALIENCY IMAGE DATABASE

Classes	CW	CR	SP	BG	PE
Accuracy (%)	81.1	72.6	78.4	80.2	82.7
Object Classes	**FE**	**HE**	**TE**	**BE**	**GS**
Accuracy (%)	71.3	78.1	75.3	69.5	70.3
Mean Object Detection Accuracy = 66.4%					

TABLE V. COMPARISON OF PROPOSED APPROACH WITH STATE OF THE ART OBJECT DETECTION METHODS.

Method	*Detection Accuracy (%)*
CNN [49]	76.9
Efficient scale space [47]	79
Harmony potential [48]	71.3
Proposed approach	**86.1**

IV. Conclusion

In this paper, we have proposed an effective technique for the accurate segmentation and detection of multiple objects. Our proposed approach generates the desired results after passing the images through three phases. First, for the segmentation phase, we have used two methods as Decision Tree and Region-based segmentation. Decision Tree-based method estimates the probability distribution of each pixel and Region-based method extracts the boundaries of similar regions. Second, this phase includes combination of Decision Tree and Region-based segmentation. Finally, we detect the multiple objects by color matching and label all objects with their class name. Our experiments have been validated over two publicly available benchmark datasets as Instance Saliency images and MSRC.

In future work, we would like to use results for scene understanding. We are also interested in different scenes recognition such as sports scene, traffic scene, hilly scene, laboratory scene and indoor scenes.

ACKNOWLEDGMENT

This research was partially supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2018R1D1A1A02085645). Also, it was supported by a grant (19CTAP-C152247-01) from Technology Advancement Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

REFERENCES

[1] H. Lin , J. Si , Glen and P. Abousleman, "Orthogonal Rotation-Invariant Moments for digital image processing," IEEE transactions on image processing, vol. 17, no. 3, pp. 272-282, 2008.

[2] T. Kim, A. Jalal, H. Han, H. Jeon and J. Kim, "Real-Time Life Logging via Depth Imaging-based Human Activity Recognition towards Smart Homes Services," *in Proceedings International Symposium on Renewable energy sources and healthy buildings*, pp. 63, 2013.

[3] 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 *Proceedings 12th IEEE International Conference on Ubiquitous Robots and Ambient Intelligence*, pp. 294-299, 2015.

[4] A. Jalal, S. Kamal and D. Kim, "Individual Detection-Tracking-Recognition using depth activity images," in *Proceedings 12th IEEE International Conference on Ubiquitous Robots and Ambient Intelligence*, pp. 450-455, 2015.

[5] N. J. Goodall, "Can you program ethics into a self-driving car?," IEEE Spectrum, vol. 53, no. 6, pp. 28-58, 2016.

[6] E. Lo, "Hyperspectral anomaly detection based on a non-uniform partition of the pixel," in proc. of WHISPERS, pp. 1-4, 2014.

[7] 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, pp. 11735-11759, 2014.

[8] S. Deshmukh, and T.-S. Moh, "Fine object detection in automated solar panel layout generation", in proc. of ICMLA, pp. 1402-1407, 2018.

[9] L. Yu, X. Chen and Sansan Zhou, "Research of image main objects detection algorithm based on deep learning", in proc. of ICIVC, pp. 70-75, 2018.

[10] K. Kuan, G. Manek, J. Lin, Y. Fang and V. Chandrasekhar, "Region average pooling for context-aware object detection," in proc. of ICIP, pp. 1347-1351, 2017.

[11] Y. Zhao and A. Cai., "A novel relative orientation feature for shape-based object recognition", in proc. of NIDC, pp. 686-689, 2009.

[12] P. Zhang, T. D. Bui, and C. Y. Suen, "Recognition of similar objects using 2-D wavelet-fractal feature extraction," in proc. of ORUI, pp. 316-319, 2002.

[13] S. Kamal, A. Jalal and D. Kim, "Depth Images-based Human Detection, Tracking and Activity Recognition Using Spatiotemporal Features and Modified HMM, *Journal of Electrical Engineering and Technology,* pp. 1921-1926, 2016.

[14] 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.

[15] 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), pp. 1657-1662, 2017.

[16] A. Sable and K. Jondhale, "Modified double bilateral filter for sharpness enhancement and noise removal", in proc. of ACE, pp. 295-297, 2010.

[17] A. Jalal, M. Mahmood and K. Kim, "Vision-based Human Activity Recognition System Using Depth Silhouettes: A Smart Home System for Monitoring the Residents," Journal of Electrical Engineering & Technology, 2019.

[18] A. Jalal, M. A. K. Quaid and K. Kim, "A Wrist Worn Acceleration Based Human Motion Analysis and Classification for Ambient Smart Home System," Journal of Electrical Engineering & Technology, 2019.

[19] K. Kim, A. Jalal and M. Mahmood, "Vision-based Human Activity recognition system using depth silhouettes: A Smart home system for monitoring the residents, *Journal of Electrical Engineering and Technology,* **2019.**

[20] M. Rahman and T. Chow, "Image retrieval using an improved similarity measure: SRIC similarity with region importance and consistency", in proc. of ICCIT, pp. 336-343, 2014.

[21] 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.

[22] A. Jalal, S. Kamal, and D. Kim, "Depth Map-based Human Activity Tracking and Recognition Using Body Joints Features and Self-Organized Map," in *Proceedings of the IEEE International Conference on computing, communication and networking technologies*, pp. 1-6, 2014.

[23] Mouazma Batool, A. Jalal, and Kibum Kim, "Sensors Technologies for Human Activity Analysis Based on SVM Optimized by PSO Algorithm", IEEE ICAEM conference, 2019.

[24] 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.

[25] A. Jalal and M. Mahmood, "Students' Behavior Mining in E-learning Environment Using Cognitive Processes with Information Technologies," Education and Information Technologies, Springer, 2019.

[26] 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.

[27] M. Doller and H. Kosch, "Approximating the selectivity of multimedia range queries", in proc. of ICME, pp. 4-8, 2005.

[28] 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.

[29] A. Ahmed, A. Jalal, and A. A. Rafique, "Salient Segmentation based Object Detection and Recognition using Hybrid Genetic Transform", IEEE ICAEM conference, 2019.

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

[31] A. Jalal, Maria Mahmood, and M. A. Sidduqi, "Robust spatio-temporal features for human interaction recognition via artificial neural network, IEEE conference on International Conference on Frontiers of information technology, 2018.

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

[33] L.-S Liu, Y. Zhang, H.-Q. Lu, L. Zhao and B.-G. Wang, "A Fast sub-pixel edge detection method for images," Journal of Optoelectronics Laser, 2005.

[34] 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, Journal of Electrical Engineering & Technology, 2018.

[35] N. Yokoya and A. Iwasaki, "Object detection based on sparse representation and Hough voting for optical remote sensing imagery," in proc. of STAEORS, pp. 2053-2062, 2015.

[36] A. Jalal, S. Kamal and D.-S. 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.

[37] M. Mahmood, A. Jalal, and H. A. Evans, "Facial Expression Recognition in Image Sequences Using 1D Transform and Gabor Wavelet Transform,

IEEE conference on International Conference on Applied and Engineering Mathematics, 2018.

[38] 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.

[39] A. S. Tamboli and M. A. Shah, “A generic structure of object classification using genetic programming”, in proc. of CSNT, pp. 723-728, 2011.

[40] 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.

[41] 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.

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

[43] 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.

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

[45] A. Jalal and S. Kamal, “Real-Time Life Logging via a Depth Silhouette-based Human Activity Recognition System for Smart Home Services,” in *Proceedings of the IEEE International Conference on Advanced Video and Signal-based Surveillance*, pp. 74-80, 2014.

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

[47] 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 *Proceedings IEEE International Conference on Informatics, electronics and vision*, pp: 1-6, 2015.

[48] 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 *Proceedings IEEE International Conference on Informatics, electronics and vision*, 2015.

[49] J. Wang, S. Wang, Q. Cui, and Q. Wang, “Local-based active classification of test report to assist crowdsourced testing,” in Proceedings of IEEE International Conference on Automated Software Engineering, pp. 190-201, 2016.

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

[51] 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.

[52] 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 Transactions on internet and information systems, vol. 9(5), pp. 1856-1869, 2015.

[53] P. Simari, D. Nowrouzezahrai, E. Kalogerakis and K. Singh, “Multi-objective shape segmentation and labeling,” in Proceedings of International symposium on Geometry Processing, pp. 1415-1425, July 2009.

[54] 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.

[55] A. Jalal, S. Kamal, and D. Kim, “Shape and motion features approach for activity tracking and recognition from Kinect video camera,” in *Proceedings 29th International Conference on Advanced Information Networking and Applications Workshops*, pp. 445-450, 2015.

[56] 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, ISSN: 0098-3063, vol. 58, Issue.3, pp. 863-871, 2012.

[57] 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 Conference on computing, communication and networking technologies*, pp. 1-6, 2014.

[58] A. Adnan, S. Gul, M. Ali and A. Dar , “Content based image retrieval using geometrical-shape of objects in image”, in proc. of ICET, pp. 222-225, 2007.

[59] J. Yao, S. Fidler and R. Urtasun, “Describing the scene as a whole: joint object detection,” in proc. of CVPR, 2012.

[60] J. Jiayan and Z. Tu, “Efficient scale space auto-context for image segmentation and labeling”, in proc. of CVPR, pp. 1810-1817, 2009.

[61] G. Josep, X. Boix, J. Weijer, A. Bagdanov, J. Serrat and J. Gonzalez, “Harmony potentials for joint classification and segmentation”, in proc. of CSCCVPR, pp. 3280-3287, 2010.

[62] J. Bappy and A. R.-Chowdhury. "CNN based region proposals for efficient object detection." in proc. of ICIP, pp. 3658-3662, 2016.