2019 2nd International Conference on Communication, Computing and Digital systems (C-CODE)

Human Body Parts Estimation and Detection for Physical Sports Movements

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

Amir Nadeem
Dept. of Computer Science
Air University
Islamabad, Pakistan
171269@students.au.edu.pk

Satoshi Bobasu
Dept. of Intellignet system
Guangxi Normal University
Guilin, China
sat.bou@gxnu.edu.cn

***Abstract*- The main purpose of human body part detection is to estimate the size, orientation or position of the human body parts within the digital scene information. Estimation of various body parts of the human from an image is a critical step for several model-based systems and body-parts tracking. In this paper, body parts detection for pose estimation is implemented. During foreground silhouettes detection, the proposed method have used segmentation techniques to obtained salient region areas and skin tone detection. After successful silhouettes extraction, body parts estimation is applied by using body parts model. Five basic body key points was determined and in addition seven more body sub key points was estimated with the help of five basic body key points. The estimated key points of the body are then represented using circular marks on the original image. The experimental results over two challenging video datasets such as KTH-multiview football and UCF sports action datasets showed significant accuracies of 90.01% and 86.67. The proposed method performs better than the state-of the-art methods in terms of body-parts detection accuracy.**

Keywords— body parts detection, motion estimation, key points extraction, sports dataset

I. Introduction

Body parts estimation and detection from images is an important and top research in the field of image processing due to its challenging nature and widespread applications. Some of its applications are 3D games [1, 2], health exercise systems [3-5], virtual reality, activity recognition, scene analyzing, security surveillance [6] and detecting human abnormal behavior [7-9]. Recently, this research has received various directions such as extracting angle movements related to the full body, local constellation features of body parts positions, pose estimation and events detection from the single/ sequential RGB images [10]. These directions faces various difficulties, such as the articulation nature of human body, changes in lightening conditions, camera dimensions, occlusion and noisy background [11-13]. Recently, a great effort has been devoted for detecting human body parts in digital images. In spite of the facts, these procedures perform better on few body parts, however, their performance become poor [14] for elbows, wrists and other critical body-parts.

Several researches established their efforts to solve critical movements of human body parts. Like, S. Maheswari *et al.* worked on enhancing skin tone detection using heuristic thresholding [14]. Their mechanism outlines the detection of skin tone with heuristic thresholding of YC_bC_r color model and skin tone enhancement. This approach was successfully implemented for segmenting skin and non-skin regions of group of people and a single person in random images. A Jalal *et al.* [15] worked on the human activity recognition from depth video without attaching any motion sensors like optical markers to human body parts [16, 17]. H. W. Chen *et al.* [18] created a new way for body part tracking and detection with color patch segmentation and adaptive thresh holding for security surveillance cameras. An adaptive threshold technique was developed for dealing with illumination condition changes, body size changes, and changing parameter of the camera. A Jalal *et al.* designed a random forest iteration method to cope various temporal features to figure-out various actions [19]. They presented a new method using the Hidden Markov Models and recognized body parts of human silhouettes [20, 21]. W. Lee *et al.* implemented a novel hierarchical method [22] for tracking human pose in which edge-based features are used during the coarse stage.

In this paper, initially, salient region detection method is implemented to detect the visibly noticeable regions in the image. Secondly, the method for foreground segmentation by skin tone detection on the basis of skin color and luminance is implemented. These detection methods are merged to get silhouette of a person that is more accurate than any of the one method applied alone. Five basic body key points are detected by using body parts model. In addition, seven more body key points are detected with the help of five basic body key points and these points are called body sub key points. In terms of position detected of key points, the body parts estimation is measured. Finally, experimental results demonstrated the efficacy of our method after applying over KTH Multiview Football UCF Sports Action datasets.

The rest of the proposed paper is organized as follows. Section II describes the system design and illustrates our proposed system architecture. It includes pre-acquisition process which describes human segmentation from images, body part initialization and detecting twelve body key points. Section III is about experimental results in which performance of our method is described using three different tables. Section IV is the last section of our paper in which the proposed research work is concluded.

II. Proposed System Methodology

The proposed system starts with the pre-processing of single image as input and human silhouette is detected by applying salient region detection method merged with foreground detection method. Then, five basic body key points are detected with the help of silhouette of a person by applying body parts model. Our system also detects seven more body sub key points with the help of previously detected basic body key points. Fig.1. shows the system architecture of our proposed system.

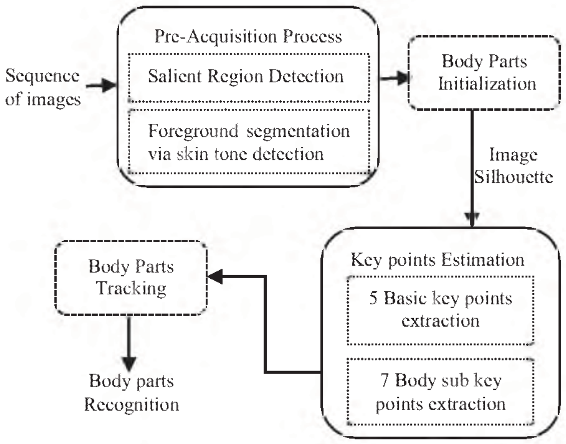


Fig. 1. Overall flow of the proposed human body parts detection method

The purpose of this paper is to enhance human body part detection, and in particular to detect body parts accurately by modeling synergies between body parts. Human body need to be segmented from each other and from the background as well. After this, we are able to detect the body parts and calculate their 3-D positions.

A. Pre-Acquisition Process:

1) Salient region Detection: To detect visually noticeable regions in an image is very useful in applications like image retrieval, foreground detection and smart image resizing [23]. These regions are called as a salient regions. In human movements detection, we detected salient regions that are visually more conspicuous due to their contrast with respect to other regions in an image [24, 25]. In addition, salient regions used contrast identifying filters that create saliency maps having "saliency values" for each pixel. This method enhanced the use of the final saliency map [26-28] in foreground detection with the help of a simple segmentation method.

This technique is efficient on a large number of images including frames of video, paintings, and images which contain noisy data etc. Saliency is calculated as the local contrast of the region of an image with comparison to its neighbor regions at various levels. This is calculated as the distance between the average feature vector [29,30] of chracteristics of the pixels of an image sub-area and the average feature vector of the pixels of its neighbor area [31]. At a given level, this contrast based saliency value $s_{i,j}$ for a pixel at position *(i, j)* in the image is calculated as the distance *d* between the average feature vectors of pixel of the inner area *r1* and that of the outer area *r2* as:

$$s_{i,j} = d\left[\left(\frac{1}{n1}\sum_{p=1}^{n1} v_p\right), \left(\frac{1}{n2}\sum_{q=1}^{n2} v_q\right)\right] \quad (1)$$

where n_1 and n_2 are the number of pixels in area r_1 and r_2, respectively.The distance *d* is a Euclidean distance; if *v* is the feature vector having correlated and uncorrelated measurements.

2) Foreground segmentation via skin tone detection:

To segment foreground via skin tone, an approach of color space transformation [32] is achieved using heuristic thresholding. It identifies skin tone regions by the appropriate YC_BC_R model. Equations (2-4) are showing how to convert RGB [33] to $Y\,C_R\,C_B$ model:

$$Y = 76 + (65.48R + 128.55G + 24.96B) \quad (2)$$

$$C_B = 128 + (-37.79R - 74.20G + 112.0B) \quad (3)$$

$$C_R = 128 + (112.0R - 93.78G - 18.21B) \quad (4)$$

While, identified skin tone regions are extracted which are enhanced via color transformation. Random thresholding is applied as a threshold [34, 35] for chrominance parts as C_R>150 and C_R<200, C_B>100 and C_B<150 which does not classify skin and non-skin region [36] correctly. The method requires the choice of a heuristic thresholding [37-39] as chrominance C_B having range between 77-127 and C_R should be between 133-173. This classified skin region without any training unlike.

We have combined the detected foreground via both methods to get a robust result. In our datasets, we are dealing with only one person having centered. Considering silhouette [40, 41] representation in Fig. 2, if we have detected some area in image Fig. 2(a) in the center then we will keep it in image Fig. 2(c) even if it is not in image Fig. 2(b). In this way image c is enhanced. Similarly, in image Fig. 2(b), if there is detected area at the sides of the image [42] and there is no area detected on the same place in image Fig. 2(a), we will ignore that area and we will not include that area in image Fig. 2(c).

(a) (b) (c)

Fig. 2. Image representation after detecting foreground. (a) foreground detected via skin tone (b) foreground detection via salient region segmentation, (c) combined features.

B. Body Parts Initilization:

After extracting human silhouette, we initialized each silhouette by proposing body parts model algorithm. In body parts models [43], we can detect basic body key points and basic body sub-key points with the help of its shape and body parts angles.

C. Body parts Models:

In basic body key points, we have calculated the distance from top area of silhouette to the bottom area of the silhouette [44]. To examine the height and width of the person, we normalized human head width as 1/8.5th times of the height. In this way, we determined the head width by calculating the height of the silhouette by considering pixel-wise searching. To detect the head position, we used following formulation as;

$$Y_H^f \leftarrow Y_H^{f-1} + \Delta Y_H^{f-1} \quad (5)$$

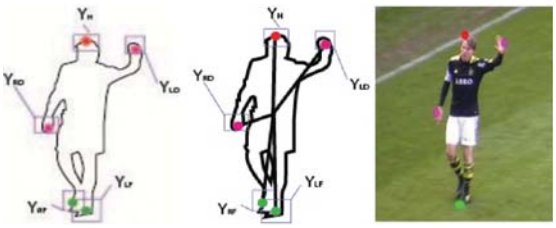
where Y_H^f is head position at any given frame f. While, it is obtained by finding the difference between sequences of frames.

However, to track limbs to estimate the feet and hips position, we used a series of joint points having attributes of forward kinematics technique [45]. This technique operates i^{th} joint space of limbs area surrounded by feature data. These joints spaces are calculated as

$$Y_i^f = Y_{i-1}^f + \left(R_{i-1}^{f-1} \ldots\ldots R_0^{f-1}\right).\left(Y_i^{f-1} - Y_{i-1}^{f-1}\right) \quad (6)$$

Y_i^f is i^{th} joint position of limb Y in the certain frame f. $(R_{i-1}^{f-1} \ldots R_0^{f-1})$ are the concatenation of the rotation matrix [46] of the basic joints. To detect feet, we have taken the lowest points on the lower limbs. These limbs are searched in both directions upto the certain height to find the feet. If we find the point on left side of the other foot than we call first foot the right foot and another as left foot and vice versa. If we are unable to find the foot on both sides then second foot will be on the same place and there is occlusion.

To detect left and right hands, the detection of the upper arm is then used to guide the detection [47-49] for the lower arms. From the estimated lower arm positions, multiple search directions are identified based on foreground and skin color mechanism. While, lower arm rectangles are initialized and converged to local maximums via gradient descent. These lower arms rectangles are used to find the hand, which is modeled in Figure 3.



(a) (b) (c)

Fig. 3. Basic body key points technique. (a) Five points with body parts model, (b) Five points extracted using a body part model. (c) key points embedded in original image.

D. Body Sub Keypoints Extraction:

After five basic body key point's detection, we can detect seven more points on the body. Firstly, we determine a torso point [50], therefore, we calculated centroid of the two feet and upper head. To adjust the position of the torso, we formulated as;

$$Y_T^f \leftarrow Y_T^{f-1} + \Delta Y_T^{f-1} \quad (7)$$

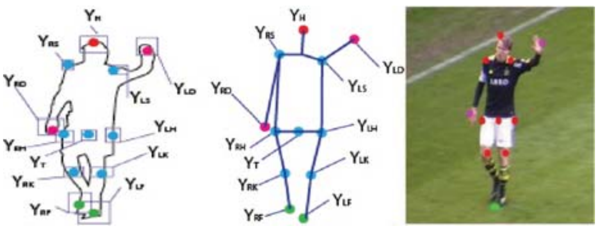
While, Y_T^f is torso point position at any given frame f. It is determined by estimating the difference between sequences of frames. After detecting the point, we calculated two more points which are left and right hip points [51, 52]. To find the points of both hips, the aspect ratio information of both torso height and width are fixed to their surround. While, each of these points have specific distribution which is averaged together with other points to produce torso points. These points are used to compute expected length and joint angle [53, 54] during initialization parameters.

The two shoulders points are identified by converging an appropriate function. The function domain is an isosceles triangle whose base vertices are the two shoulder points and its main vertex is the neck point. When the triangle base is maximized, the function is minimized and vice versa.

Using equation 6; we have already determined the position of the limbs [55]. So to find the exact location of the shoulder point, we have taken in account the position [56] of upper limbs as well. At the end, we identified knee by just considering middle of the hip and the foot. Finally, estimation [57-59] of knees points positions are formulated as;

$$Y_K^f = (Y_F^f - Y_{Hip}^f)/2 \quad (8)$$

Y_K^f is the position of the knee, Y_F^f is the position [60] of the foot and Y_{Hip}^f is the position of the hips.



(a) (b) (c)

Fig. 4. Body parts models (a) All twelve point with body parts model are present, (b) is showing lines connecting these points, and (c) is representing that how our real image looks after detecting points.

III. EXPERIMENTAL RESULTS

In our experiment; we used two data sets. One is KTH Multiview Football dataset and other is UCF Sports Actions dataset. We check the efficiency of our method by estimating distance of detected point from ground truth.

A. Dataset Description

KTH Multiview Football dataset [61] contains 5907 images. Each image has 2D ground truth pose and 14 annotated joints. There are 3 different players. This data is recorded during the football game.

Fig. 5. Some samples from KTH Football Multiview dataset. Upper row images are real images and down row images are detected points.

UCF Sports dataset [62] consists of a collection of actions from many different sports which are normally displayed on broadcast television channels like ESPN and BBC. The dataset includes many actions which involve *Diving*, *golf swing*, *kicking*, *lifting*, *riding horse*, *running*, *skateboarding*, *swing bench* and *swing-side* along with the annotations of the humans.

Fig. 6. Some samples from UCF Sports Action dataset. Upper row images show real images and down side images are detected points

B. Measuring the efficiency:

To check the efficency of our method; we first calculated euclidean distance of each detected body parts from ground truth [63] and our proposed method. To calculate the euclidean distance, we used;

$$E_k = \sqrt{\sum_{k=1}^{K}\left(\frac{x_k}{s_k}-\frac{y_k}{s_k}\right)^2} \tag{9}$$

where x_k is the ground truth, y_k is detected point and E is the euclidean distance. To examine our detected point method with respect to the ground truth, we evaluate the threshold of 15 euclidean distance between ground truth and proposed model by considering eq. (10) as;

$$A = \frac{100}{k}\left[\sum_{k=1}^{K}\begin{cases}1, & if\ E \le 15\\ 0, & if\ E > 15\end{cases}\right] \tag{10}$$

$\boldsymbol{A}$ is the accuracy of the k body part. If distance of detected point is greater than 15, then that point is ignored. While, if its range is 15 or less then that point is taken into account. For all points between 1 to K; this procedure is repeated and numbers are added as well.

In TABLE I, we evaluated detection accuracy results with respect to KTH Football Multiview dataset. On KTH Football dataset; our proposed method shows significant detection accuracy of 90.01%.

TABLE I. HUMAN BODY PART DETECTION RESULTS BASED ON PROPOSED METHOD USING KTH FOOTBALL DATASET

Body Parts	Distance from ground truth	Detection Accuracy (%)
Upper head	9.4	90
Left Shoulder	12.6	86
Right Shoulder	12.3	93
Left hand	10.0	92
Right Hand	10.7	93
Left Hip	11.1	90
Right Hip	12.3	89
Left Knee	13.3	86
Right Knee	12.0	83
Left foot	8.5	94
Right foot	9.7	95
Mean Detection Accuracy rate =90.01 %		

In TABLE II, experimental results with respect to UCF Sports Action dataset is presented. Our method achieved detection accuracy as 86.67%.

TABLE II. HUMAN BODY PART DETECTION RESULTS BASED ON PROPOSED METHOD USING UCF SPORTS ACTION DATASET

Body Parts	Distance from ground truth	Detection Accuracy (%)
Upper head	10.5	88
Left Shoulder	12.7	91
Right Shoulder	14.8	95
Left hand	10.3	91
Right Hand	10.1	89
Left Hip	11.2	79
Right Hip	11.3	88
Left Knee	14.1	83
Right Knee	12.2	81
Left foot	10.3	92
Right foot	9.9	96
Mean Detection Accuracy rate =86.67 %		

While, in TABLE III, performance of our method is compared with state of the art methods. Results shows that our performance on both the datasets is far better than existing methods shown in TABLE III. In S Park et al. method [64], framework of Markov random field is used to merge the pixels into connected blobs and to register inter-blob connections. G. Blumrosen *et al.* [65] used RSSI measurement to detect body parts. In S. Li *et al.* [66], they used conventional neural network for human pose estimation. While, H. W. Chen *et al* [7] used color patch morphological segmentation and adaptive thesholding.

TABLE III. COMPARISON OF HUMAN BODY PART DETECTION ACCURACIES BETWEEN STATE OF THE ARTS METHODS AND PROPOSED METHOD

Methods	KTH Multiview Football dataset	UCF Sports Actions dataset
S Park et al. [63]	69.1	64.33
G Blumrosen et al. [64]	77.8	81.1
S Li et al.[65]	82.9	78.2
H W Chen et al. [18]	84.01	79.0
Proposed Method	**90.01**	**86.67**

IV. CONCLUSION

In this paper, we presented a method of detecting robust human silhouette via salient regions and skin tone segmentation methods. We proposed a novel body parts model scenario having basic body key points and body sub key points mechanisms. Our results shows significant improvements from conventional systems. In the future, we have planned to incorporate the complexity of occlusion in between body parts.

ACKNOWLEDGMENT

Research was sponsored by the Cooperative Society of Innovation laboratory and was accomplished under Contract Agreement Number EW127MS-18-4-0052.

REFERENCES

[1] J. E. Lim, O. Choi, H. S. Na and D. K. Baik, "A Context-Aware Fitness Guide System for Exercise Optimization in U-Health," IEEE Transactions on Information Technology in Biomedicine, vol. 13, pp. 370-379, May 2009.

[2] K. Buys, C. Cagniart, A. Baksheev, T.-D. Laet, J. D. Schutter and C. Pantofaru, "An adaptable system for RGB-D based human body detection and pose estimation," Journal of visual communication and image representation, vol. 25, pp. 39-52, Jan 2014.

[3] A. Jalal and S. Kim, "Global security using human face understanding under vision ubiquitous architecture system," World Academy of Science, Engineering, and Technology, Vol. 13, pp. 7-11, 2006.

[4] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman and A. Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation," IEEE Transaction on pattern analysis and machine intelligence, vol. 24, no. 7, pp. 881-892, July 2002.

[5] A. Jalal and A. Shahzad, "Multiple facial feature detection using vertex-modeling structure," in Proc. IEEE ICL, pp. 1-7, 2007.

[6] A. Jalal and Y. Rasheed, "Collaboration achievement along with performance maintenance in video streaming," in Proc. IEEE ICL, pp. 1-8, 2007.

[7] X. Yang and Y. Tian, "Eigenjoints-based action recognition using naive-bayes-neartest-neighbor," in Proceedings of CVPR Conference, Providence, RI, pp.14-19, June 2012.

[8] 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.

[9] D. Arifoglu and A. Bouchachia, "Activity recognition and abnormal behaviour detection with recurrent neural networks," Procedia Computer Science, vol. 110, pp. 86-93, 2017.

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

[11] A. Jalal, S. Kim and B. J. Yun, "Assembled algorithm in the real-time H.263 codec for advanced performance," " in Proc. IEEE workshop on ENCHI, pp.295-298, 2005.

[12] A. Jalal, M. Z. Uddin, J. T. Kim, and T.-S. Kim, "Recognition of human home activities via depth silhouettes and R transformation for smart homes," Indoor and Built Environment, vol. 21, 2012.

[13] A. Jalal, and M. A. Zeb, "Security enhancement for e-learning portal," IJCSNS, vol. 8(3), pp. 41-45, 2008.

[14] S. Maheswari and R. Korah, "Enhanced skin tone detection using heuristic thresholding," Biomedical Research, vol. 28, pp. 4147-4153, 2017.

[15] A. Jalal, J. T. Kim, and T.-S Kim, "Development of a life logging system via depth imaging-based human activity recognition for smart homes," in Proceedings of the ISHB, pp. 91-95, 2012.

[16] A. Jalal and I. Uddin, " Security architecture for third generation (3G) using GMHS cellular network," in IEEE ICT, pp. 74-79, 2007.

[17] A. Jalal, S. Lee, J. Kim, and T. Kim, "Human activity recognition via the features of labeled depth body parts," in Proc ICOST, 2012.

[18] H. W. Chen and M. McGurr, "Moving human full body and body parts detection, tracking, and applications on human activity estimation," SPIE Defense + Security, May 2016.

[19] A. Jalal, N. Sharif, J. T. Kim and T. S. Kim, "Human Activity Recognition via Recognized Body Parts of Human Depth Silhouettes for Residents Monitoring Services at Smart Home," Indoor and Built Environment, vol. 22, pp. 271-279, Feb 2013.

[20] A. Jalal and S. Kim, S, "Advanced performance achievement using multi-algorithmic approach of video transcoder for low bit rate wireless communication," ICGST, vol. 5(9), pp. 27-32, 2005.

[21] 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.

[22] W. Lee and R. Nevatia, "Body Part Detection for Human Pose Estimation and Tracking," IEEE Workshop on Motion and Video Computing, USA, pp. 23, Feb. 2007.

[23] 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.

[24] 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.

[25] 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.

[26] P. Milanfar, "A tour of modern image filtering: New imsights and methods, both practical and theoretical," IEEE signal processing magazine, vol. 30, pp. 106-128, Dec 2012.

[27] 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.

[28] 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.

[29] A. Jalal, "IM-DailyDepthActivity dataset," imlab.postech.ac.kr/ databases.htm, 2015, [Online: accessed 14 April- 2015].

[30] 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.

[31] 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.

[32] D. N. Anh, "Detection of lesion region in skin images by moment of patch," IEEE Internation Conference on Computing and Communication Technologies, Nov. 2016.

[33] A. Jalal, S. Kamal, and D. Kim, "A depth video sensor-base lifelogging human activity recognition system for elderly care in smart indoor environments," Sensors, vol. 14(7), pp. 11735-11759, 2014.

[34] 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.

[35] A. Jalal, Majid 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.

[36] D. Buller, A. Buller, P. R. Innocent and W. Pawal, "Determining and classifying the region of interest in ultrasonic images of the breast using neural networks," Artificial Intelligence in Medicine, vol. 8, pp.53-66, Feb. 1996.

[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. Ahmadi et al. , "3D Human Gait Reconstruction and Monitoring Using Body-Worn Inertial Sensors and Kinematic Modeling," IEEE Sensors Journal, IEEE Sensors Journal, vol.16, pp. 8823-8831, Jul. 2016.

[46] T. Otani, H. Matsuda, H. Touyama and T. Nakata, "Gaze Estimation Using Human Joint Rotation Angel," IEEE International Conference on Cyberworlds, Sweden, Feb. 2016.

[47] 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.

[48] 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.

[49] 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.

[50] A. Dargazany and M. Nicolescu, "Human Body Parts Tracking Using Torso Tracking: Applications to Activity Recognition," IEEE Ninth International Conference on Information Technology, Jun. 2012.

[51] 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.

[52] 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.

[53] 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.

[54] 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.

[55] Y. Luo, H. Yang and Z. Hu, "Human limb motion real-time tracking based on CamShift for intelligent rehabilitation system," IEEE International Conference on Robotics and Biomimetics, China, Feb. 2010.

[56] 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.

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

[58] 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.

[59] 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.

[60] 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.

[61] M. Burenius,J. Sullivan and S. Carlsson, "3D Pictorial Structures for Multiple View Articulated Pose Estimation," In Proceedings of CVPR, 2013.

[62] M. D. Rodriguez, J. Ahmed and M. Shah, "Action MACH: A Spatio-temporal maximum average correlation height filter for action recognition," Computer Vision and Pattern Recognition, 2008.

[63] L. Wang, Y. Zhang and J. Feng, "On the Euclidean distance of images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, pp. 1334-1339, Jun. 2005.

[64] S. Park and J. K. Aggarwal, "Segmentation and Tracking of interacting human body parts under occlusion and shadowing," IEEE Proceeding of the workshop on motion and video computing, USA, 2002.

[65] G. Blumrosen and A. Luttwak, "Human body parts tracking and kinematic features assessment based on rssi and inertial sensor measurements," MDPI, Sep. 2018.

[66] S. Li and A. B. Chan, "3D Human Pose Estimation from Monocular Images with Deep Convolutional Neural Network," Asian Conference on computer vision, Singapore, 2014.