Effective Technology Based Sports Training System Using Human Pose Model

Kannan Paulraj and Nithya Natesan
Department of Electronics and Communication Engineering, Panimalar Engineering College, India

Abstract: *This paper investigates the sports dynamics using human pose modeling from the video sequences. To implement human pose modeling, a human skeletal model is developed using thinning algorithm and the feature points of human body are extracted. The obtained feature points play an important role in analyzing the activities of a sports person. The proposed human pose model technique provides a technology based training to a sports person and performance can be gradually improved. Also the paper aimed at improving the computation time and efficiency of 2D and 3D model.*

Keywords: *Thinning algorithm, human activity, motion analysis, feature extraction.*

Received March 28, 2015; accepted September 9, 2015

1. Introduction

Sports video analysis has become a new research domain in the last decade. This new domain is an attracting area of research and can contribute to the coaches, players and even the viewers with additional statistical information [2, 4]. Also this area contributes in developing and enhancing a system for sports video processing and analysis. Human motion analysis is reviewing an increasing attention from computer vision researchers. The human body parts are segmented from an image, the movements of joints are tracked and the underlying 3D body structure [1, 3] are recovered for analysis on the area of military security, health related areas and in entertainment environments. This paper focuses on sports analysis [6].

Recently research has been improved for activity recognition, using machine learning approach to develop statistical models of human activities [7, 10]. It uses feature matching process for this implementation. Marker based human pose modeling is a simple approach but it is not possible to reconstruct all human poses in practical situation. So the marker less pose modeling is very important in the automated analysis [5, 9]. The activity analysis still has a problem in feature selection and machine learning. To recognize human activities from an image sequence, two types of approaches were addressed, approaches based on state space model and those using the template matching technique [8]. The Template Matching Method has a drawback that the process is affected by noise and time interval of movements. Due to this implementation the computational cost and complexity increases. State space model involves the recognition of features from points, lines and 2D blobs. This method is also a complex method and produces a low accurate result. This paper proposes a method to recognize the human activity and also helps on finding the distance between the feature points. The human pose model is developed using thinning algorithm.

2. Proposed Human Pose Model

Human pose modeling is be done using two prime algorithms. They are:

1. Thinning algorithm.
2. Principle Component Analysis (PCA) algorithm.

Thirteen feature points are extracted form the human body for the implementation of human pose model such as head, neck, left shoulder, right shoulder, left hand elbow, right hand elbow, abdomen, Left hand, right hand, left knee, Right knee, left leg and right leg. Through thinning, the skeleton of human body is obtained from which the pose models are implemented. After the pose modeling, positioning of the human activity is analyzed using PCA algorithm.

2.1. Thinning Algorithm

Thinning algorithm plays a vital role in finding the skeleton of human body. Thinning is morphological operation that can be used to remove the selected foreground pixels from binary images. Thinning is normally applied only to binary images and produces another binary image as an output. Thirteen feature points are identified on thinned image and the activities of human are analyzed for predefined poses. The thinning operation is performed by transforming the origin of the structuring element to each pixel in the image. Then it is compared with the corresponding image pixels. When the background and foreground pixels of the structuring element and images are matched then the origin of the structuring element is considered as background, otherwise it is left unchanged. Here, the structuring element determines the use of the thinning operation and determines the number of pixels added or deleted from the objects in an image. It is a matrix that has ones and zeros only with any arbitrary shape and size. For example, strel ('line', 6, 90) function is used as a structuring element. The flat, linear structuring element is considered for such type of application. Here, six specifies the length and 90 specify the angle (in degrees) of the line. The length is approximately the distance between the centers of the structuring element members at opposite ends of the line. Figure 1 shows the structuring elements for skeletonization by morphological thinning. At each iteration, the image is first thinned by the left hand structuring element, and then by the right hand one, and then with the remaining six 90° rotations of the two elements. The thinning operation is achieved by the hit-and-miss transform. Figure 1-a and 1-b illustrates the thinning of an image A by a structuring element B is given by Equation (1).

$$thin\ (A,B)=A-hit\ and\ miss\ (A-B) \tag{1}$$

0	0	0
	1	
1	1	1

a) Image.

	0	0
1	1	0
	1	

b) Structured element B.

Figure 1. a) and b) Mathematical approach of thinning algorithm for various images.

The proposed thinning algorithm has the following three steps.

- *Step 1*. Creation of data base with reference activity images.
- *Step 2*. Training of data base.
- *Step 3*. Recognition of test activity using PCA.

This process is repeated in a cyclic fashion until none of the thinning produces any further change.

2.2. PCA Algorithm Based Human Activity Analysis

The template matching method and state space method are used for activity analysis in the existing work. In template matching method, the data base is needed for classification whereas the learning and training phases are required in state space approach. PCA algorithm works under template matching technique and Hidden Markov Model (HMM) works based on state space approach. The template matching method has a drawback that the process is affected by noise and time interval of movements. In this HMM method, selection of states and dimensionality of the vector is complex for the recognition which leads to high computational cost and low accuracy. The Projection histogram was applied for training and matching the stored data base with the current videos, PCA and HMM methods are used to recognize the activities. Figure 2 shows the data base of various activities. In this approach, static camera is used continuously to detect the human and their activities. Figure 3 gives the simulation results of activity analysis using PCA algorithm. Tables 1 and 2 indicate the efficiency of different videos and activities.

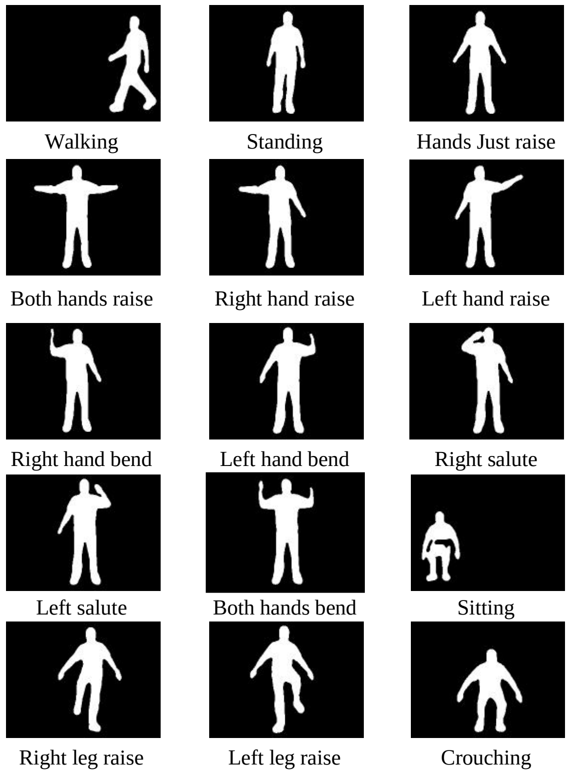
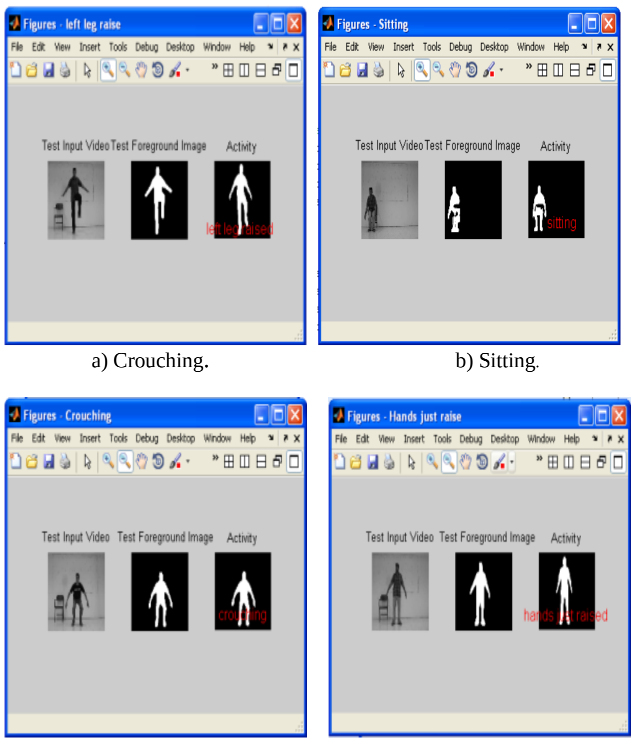


Figure 2. Data base of various activities using PCA algorithm.



a) Crouching. b) Sitting.

c) Left leg raised. d) Hand just raised.

Figure 3. Human activity recognition using PCA algorithm.

Table 1. Efficiency of PCA algorithm for different videos.

Videos	No. of frames	No. of poses correctly detected	Efficiency (%)
Video 1	1006	880	87.5
Video 2	1147	990	86.3
Video 3	1382	1230	89.0
Video 4	864	790	91.4
Video 5	1502	1300	86.6
Video 6	1257	1050	83.5
Video 7	1286	1100	85.5
Video 8	671	600	89.4
Video 9	1600	1400	87.5
Video 10	1826	1590	87.1
			87.4

Table 2. Efficiency of human poses for a single video.

Activity	No. of poses	No. of poses correctly detected	Efficiency (%)	Time (Sec) (for single frame)
A	93	89	95.6	3.27
B	110	90	81.8	3.15
C	98	83	84.7	2.86
D	98	92	93.8	2.42
E	95	80	84.2	2.00
F	100	84	84.0	3.86
G	89	79	88.8	3.45
H	89	82	92.1	2.68
I	98	83	84.7	1.93
J	53	46	86.8	2.71
K	36	32	88.9	1.84
L	47	40	85.1	2.36
	1006	880	87.5	2.71

To overcome the problem of template and state space approaches, feature points based human activity analysis is presented in this paper. The feature points are derived from the human pose models. The thirteen feature points are taken for developing pose models. To analyze the activities, only six points are considered such as head, right hand, left hand, right leg, left leg, and neck. To generalize the analysis, the distances between the points are calculated using the Equation (2). It works well for all persons.

$$d = \sqrt{\left((x_2 - x_1)^2 + (y_2 - y_1)\right)^2} \qquad (2)$$

Algorithm of PCA for Human activity analysis includes,

- *Step 1.* Obtain six points from human body model.
- *Step 2.* Determination of distances between points.
- *Step 3.* Compare the distances found in step1 with height of human body.
- *Step 4.* Recognition of activities.

The distance based activity recognition is so important and more generalized as the feature points are tracked in every frame of sequences. The various distances are measured between body parts that are shown in Figure 4 and Table 3.

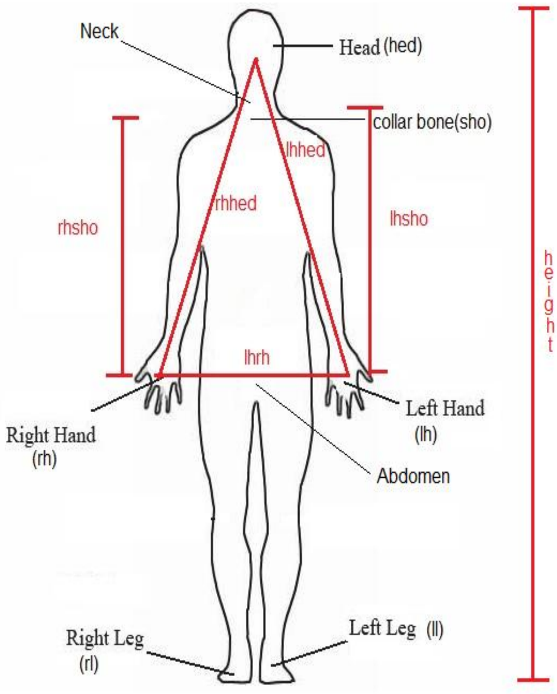


Figure 4. Measurement of distances between parts.

Table 3. Measurement of various distances on body model.

S.No	Parameter	Distance
1	height	head~ bottom end
2	rhhed	right hand~ head
3	lhhed	left hand~ head
4	lhrh	left hand ~ right hand
5	rhsho	right hand~ neck (collarbone(sho))
6	lhsho	left hand~ neck (collarbone(sho))
7	rlbend	right leg~ bottom end
8	llbend	left leg~ bottom end

In the first step of algorithm, the distances between the various points are measured. The parameter 'height' refers the distance between the head to bottom end. 'rhhed' is the distance between right hand to head. The distance between left hand to head is calculated as 'lhhed'. 'lhrh' means the distance between lefthand to right hand. 'rhsho' represents the distance between right hand to neck. The distance between left hand to neck is represented as 'lhsho'. Also, right leg to bottom end and left leg to bottom end can be measured by 'rlbend' and 'llbend'.

Then, the distance values for various activities can be computed to derive the general conditions. Here, two videos (V1andV2) are considered to find the values of parameters which are indicated in Table 3 namely 1-8. The 'min' and 'max' represent the initial position and final position of corresponding activity in the frame. Tables 4-12 shows the distance of activities- right hand raise, right hand up, right hand salute, left hand raise, left hand up, left hand salute, right leg lift, left leg lift, crouching.

Table 4. Distance values for the activity "Right hand raise".

Videos		Parameters (from Table 1)							
		1	2	3	4	5	6	7	8
V1	Min	127.8	77.4	75.8	-3	-25	79	0	0
	Max	126.2	68.3	72.8	19	-26	88	0	0
V2	Min	116.9	72.2	74.3	-2	-36	89	0	1
	Max	117.6	64.1	73.7	14	-38	89	0	1

Table 5. Distance values for the activity "Right hand up".

Videos		Parameters							
		1	2	3	4	5	6	7	8
V1	Min	127.2	49	73.8	25	-42	65	0	0
	Max	129.0	38.1	74.9	24	-28	72	0	0
V2	Min	116	48	74.5	25	-43	76	0	1
	Max	117.4	25	68.6	26	-46	53	0	0

Table 6. Distance values for the activity "Right hand salute".

Videos		Parameters							
		1	2	3	4	5	6	7	8
V1	Min	128.9	16.8	77.6	14	-32	45	0	1
	Max	127.4	30.5	74.9	24	-36	43	0	0
V2	Min	118	13	74.5	12	-43	76	0	0
	Max	116.4	28	69.4	20	-46	53	0	0

Table 7. Distance values for the activity "Left hand raise".

Videos		Parameters							
		1	2	3	4	5	6	7	8
V1	Min	127.0	71.5	70.8	-41	-9	83	0	0
	Max	127.4	73.8	67.9	-26	30	98	0	0
V2	Min	116.9	71.2	63	-20	-6	89	0	1
	Max	117.6	64.1	73.7	-24	18	89	0	1

Table 8. Distance values for the activity "Left hand up".

Videos		Parameters							
		1	2	3	4	5	6	7	8
V1	Min	128.2	72	45.3	-42	21	77	0	2
	Max	127.0	70	37	-39	19	81	0	1
V2	Min	117	62	28	-29	23	72	0	0
	Max	116.9	63.3	24.2	-28	21	74	0	0

Table 9. Distance values for the activity "Left hand salute".

Videos		Parameters							
		1	2	3	4	5	6	7	8
V1	Min	125	78	15.2	-27	26	46	2	0
	Max	128	76.2	19.1	-29	25	44	2	0
V2	Min	118	70	14	-36	13	59	0	1
	Max	117	68.2	25	-25	16	58	0	0

Table 10. Distance values for the activity "Right leg rise".

Videos		Parameters							
		1	2	3	4	5	6	7	8
V1	Min	127.2	67.6	65.7	1	-5	90	14	0
	Max	126.3	66	65.9	16	10	120	8	0
V2	Min	115	67.3	591	12	1	114	8	0
	Max	116.1	63.4	57.7	18	-2	113	25	0

Table 11. Distance values for the activity "Left leg rise".

Videos		Parameters							
		1	2	3	4	5	6	7	8
V1	Min	129	67.7	69.8	11	8	120	0	6
	Max	127	69.6	66.3	7	19	133	0	19
V2	Min	116	60.2	60	6	17	115	0	8
	Max	114	57.7	60.0	2	25	112	0	26

Table 12. Distance values for the activity "Crouching".

Videos		Parameters							
		1	2	3	4	5	6	7	8
V1	Min	115	76	76	17	11	149	0	2
	Max	117	76	75	5	2	148	0	1
V2	Min	89	59	67	2	3	113	0	2
	Max	82	61	65	4	-2	116	0	2

Table 13. Conditions for detecting activities.

S. No	Activity	Condition
1	Right hand rise (rhr)	((abs(rhsho)<(height/10)) && (rhhed>(height/3)))
2	Left hand rise (lhr)	((abs(lhsho)<(height/10)) && (lhhed>(height/3)))
3	Right hand up (rhu)	((rhhed<(height/2.5)) &&(rhhed>(height/5)))
4	Left hand up (lhu)	((lhhed<(height/2.5)) &&(lhhed>(height/5)))
5	Right leg rise (rlr)	(abs(rlbend)>(height/16))
6	Left leg rise (llr)	(abs(llbend)>(height/16))
7	Crouching (crou)	(lhrh>(1.2*height))
8	Right hand salute (rhs)	(rhhed<(height/5))
9	Left hand salute (lhs)	(lhhed<(height/5))

Table 14. Analysis of twelve human activities.

S.No	rhr	lhr	rhu	lhu	rlr	llr	crou	rhs	lhs
A	-	-	-	-	-	-	-	-	-
B	1	0	-	0	0	0	0	-	0
C	0	1	0	-	0	0	0	0	-
D	1	1	-	-	0	0	0	-	-
E	-	0	1	0	0	0	0	-	0
F	0	-	0	1	0	0	0	0	-
G	-	-	1	1	0	0	0	-	-
H	-	-	-	-	1	-	0	-	-
I	-	-	-	-	-	1	0	-	-
J	-	0	-	0	0	0	0	1	0
K	0	-	0	-	0	0	0	0	1
L	-	-	-	-	-	-	1	-	-

From the obtained distance values, the conditions can be framed to identify nine poses and it is shown in Table 13 the conditions are developed by the experimentation with various videos. After finding the values of rhr, lhr, rhu, lhu, rlr, llr, crou, rhs and lhs, a suitable combination has been formed to analyze the activities. For example, if rhr=1 and crou=rlr=llr=lhr=lhu=lhs=0, then it is recognized as right hand activity. Similarly all the activities are recognized. Table 14 shows the analysis of twelve human activities.

3. Simulation Results

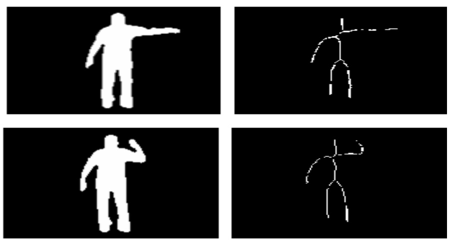


Figure 5. Results of thinning algorithm.

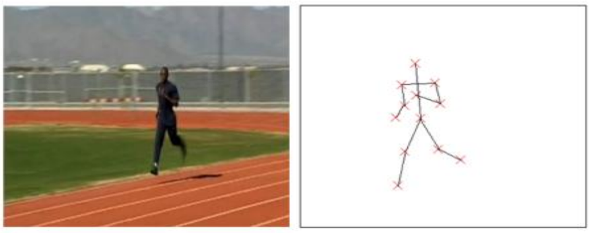
Figure 5 shows the results of the thinning algorithm for different poses. The next step was to develop the pose models for the human postures. After the human body parts were detected, the human pose models were constructed using the feature points in two dimensional (2D) views. The feature point was a point on the human body which was used to represent the body segments. The feature points that were derived from the human pose models play an important role in analyzing the activities.

The video sequences were acquired at the rate of 30 frames /second with the frame size of 320×240 pixels resolution. Figure 6 shows the simulation results of proposed work in which Column I indicates the original input video frame whereas Column II shows the two dimensional model of sports person.

1. Frame no: 1100.

a) Original video frame. b) Human body model.

2. Frame no: 1240.



3. Frame no: 1546.

4. Frame no: 1678.

5. Frame no: 1700.

Figure 6. Results of human body modelling.

Table 15 . Efficiency of 2D model for different videos.

Videos	No. of frames	No. of poses correct in 2D Modelling Using Proposed Thinning Algorithm			No. of poses correct in 2D Modelling Using Existing Delaunay Triangulation Modelling (DTM)	
		Stick Fig.	Cylinder	Efficiency (%)	No. of poses correct	Efficiency (%)
Video 1	1006	880	880	87.5	833	82.8
Video 2	1147	1026	1026	89.5	1103	88.1
Video 3	1382	1275	1275	92.3	1300	91.1
Video 4	864	810	810	94.0	800	92.6
Video 5	1502	1400	1400	93.2	1300	89.6
Video 6	1257	1050	1050	83.5	1202	82.6
Video 7	1286	1100	1100	85.5	1200	83.3
Video8	671	650	650	96.8	640	95.3
Video 9	1600	1350	1350	84.4	1421	85.8
Video 10	1826	1590	1590	87.1	1600	87.6
Average				89.4	Average	87.9

Table 16. Efficiency of 3D model for different videos.

Videos	No. of frames	No. of poses correct in 3D Modelling Using Proposed Thinning Algorithm		No. of poses correct in 3D Modelling Using Existing Delaunay Triangulation Modelling (DTM)	
		No. of poses correct	Efficiency (%)	No. of poses correct	Efficiency (%)
Video 1	1006	860	85.5	860	84.7
Video 2	1147	940	81.9	940	80.1
Video 3	1382	1200	86.8	1200	82.3
Video 4	864	740	85.6	740	81.7
Video 5	1502	1250	83.2	1250	82.2
Video 6	1257	1000	79.6	1000	75.7
Video 7	1286	1100	85.5	1100	83.8
Video8	671	600	89.4	600	88.9
Video 9	1600	1380	86.2	1380	84.2
Video 10	1826	1570	85.9	1570	83.5
Average			85.0	Average	82.7

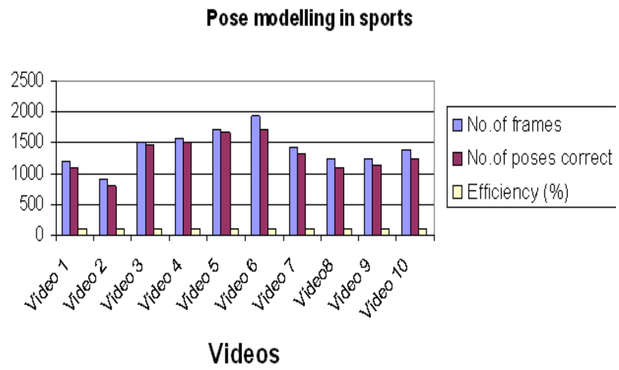


Figure 7. Human pose analysis of different videos.

Tables 15 and 16 and Figure 7 show the human pose analysis of various videos. Ten videos were considered for the analysis. The efficiency can be calculated using number of poses correctly detected with the help of the proposed algorithm. Also the table gives the comparison of the efficiency between existing and proposed algorithm.

4. Conclusions

The proposed human pose model technique could be used for training a sports person, due to this injuries could be prevented and performance can be gradually improved when compared with the performance before training without human pose model technique. Here thirteen feature points were considered form the human body parts to implement the human pose model. An efficiency of 89.4% for 2D models and 85% for 3D models are achieved, which is superior than the existing algorithm and the comparison is performed. Computation time of 2.71sec for 2D model and 1.43 sec for 3D models is also achieved. The limitation of this paper is that only the frontal poses of human were considered using single static camera and in future, this work can be extended to solve problems such as multiple person modelling and side pose of human with multiple cameras.

References

[1] Agarwal A. and Triggs B., "Recovering 3D Human Pose From Monocular Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 44-58, 2006.

[2] Aggarwal J. and Cai Q., "Human Motion Analysis: A Review," *Computer Vision and Image Understanding*, vol. 73, no. 3, pp. 428-440, 1999.

[3] Aggarwal K. and Park S., "Human Motion: Modelling and Recognition of Actions and Interactions," *in Proceedings of 2nd International Symposium on 3D Data Processing Visualization and Transmission*, Thessaloniki, pp. 640-647, 2004.

[4] Eicher M. and Ferrari V., "Human Pose Co-Estimation and Applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2282-2288, 2012.

[5] Huo F., Hendriks E., Paclik P., and Oomes A., "Markerless Human Motion Capture and Pose Recognition," *in Proceedings of 10th Image Analysis for Multimedia Interactive Services Workshop*, London, pp. 13-16, 2009.

[6] Kannan P. and Ramakrishnan R., "Development of Human Pose Models for Sports Dynamics Analysis Using Video Image Processing Techniques," *International Journal of Sports Science and Engineering*, vol. 6, no. 4, pp. 232-238, 2012.

[7] Lee M. and Cohen I., "A Model Based Approach For Estimating Human 3D Poses in Static Images," *IEEE Transactions on Pattern Analysis And Machine Intelligence*, vol. 28, no. 6, pp. 905-916, 2006.

[8] Lin Z. and Davis L., "Shape-Based Human Detection And Segmentation Via Hierarchical Part-Template Matching," *IEEE Transactions on Pattern Analysis And Machine Intelligence*, vol. 32, no. 4, pp. 604-618, 2010.

[9] Wei X. and Chai J., "Intuitive Interactive Human-Character Posing with Millions of Example Poses," *IEEE Computer Graphics and Applications*, vol. 31, no. 4, pp. 78-88, 2011.

[10] Yu-jie L., Feng B., Zong-Min L., and Hua L., "3D Model Retrieval Based on 3D Fractional Fourier Transform," *The International Arab Journal of Information Technology*, vol. 10, no. 5, pp. 421-427, 2013.

Kannan Paulraj received his M.E degree from Anna University, India in 2001 and Ph.D degree in Sports Technology from Tamilnadu Physical Education and Sports University, India in 2014. Currently he is a Professor and Head of Electronics and Communication Engineering, Panimalar Engineering College, India. His research interests include Image Processing and Optical communication.

Nithya Natesan received her M.E degree from Anna University, India in 2005. Currently she is an Assistant Professor at Panimalar Engineering College, India. Her research interests include Compression and Decompression techniques of image processing.