Person Recognition from Walking Dataset Using Kinect Sensor

Sriparna Saha, Amit Konar
Electronics and Telecommunication Engineering Department
Jadavpur University
Kolkata, India
sahasriparna@gmail.com, konaramit@yahoo.co.in

Ramadoss Janarthanan
Computer Science and Engineering Department
TJS Engineering College
India
srmjana_73@yahoo.com

Abstract—**A novel and convenient approach is implemented for person recognition using walking style with the help of Kinect sensor, a human motion sensing device. Kinect sensor estimates the full body skeleton using the twenty body joint co-ordinates of a person, standing in front of it. Here, we have created a walking dataset for fifteen persons and each person walks five times in front of the Kinect sensor. From each skeleton, tracked by Kinect sensor, a total of twenty features are calculated based on distances and angles formed by body joints. All the body joint coordinates are taken in 3D space by the Kinect sensor. Our proposed work is realized in 3D world for rotation invariance cases also. For the training purpose, eighty percent of the dataset is processed and remaining walking data sets are used for testing purpose. The proposed algorithm produces average recognition rate of 94.8% using support vector machine with average time required to recognize an unknown person is 1.327 sec in an Intel Pentium Dual Core processor running Matlab R011b.**

Keywords-Kinect sensor; pattern recognition; support vector machine; walking dataset.

I. Introduction

In today's world, video surveillance is a challenging field. The recognition of a person [1] from his(\her) walking style [2] can solve the security related threats to a large extent. For this purpose, Kinect sensor [3] can easily be implemented to identify a person from his(\her) walking styles [4]. We have created a walking dataset consisting of walking patterns from fifteen persons. We have achieved a high accuracy rate greater than ninety percent for recognition [5] using support vector machine (SVM) [6], a supervised learning algorithm [7].

As already discussed, our aim is to solve the problem of video surveillance [8]. The system is easily implemented for real-time applications in small offices, where a small number of persons work. Person recognition not only has applications in solving video surveillance problems, but also has importance in robotic and video gaming domains.

Human motion analysis [9] is a flourishing research area. The features obtained from motion analysis, person identification is possible [10]. Based on the analysis, running and walking [11] [12] can be classified for a person. In [13] person is identified using audio and video. Another important work is biometric person identification using face, speech and fusion [14]. Templates and feature points [15] are also implemented for person recognition.

In this proposed algorithm, we have created a walking dataset from fifteen persons. The persons walk in front of the Kinect sensor [16] in different angles [17]. The rotation invariance issue is successfully addressed in our work. The Kinect sensor [18] captures skeleton information at 30 frames/sec rate. From each skeleton of each frame [19], twenty features are extracted. The features are taken mainly based on Euclidean distances and angles between different body joints [20]. For some features, angular displacements are examined. The developed algorithm identifies an unknown person while walking, when the predefined twenty features are already stored in the dataset. Supervised machine learning algorithm [21] [22], support vector machine classifies the walking patterns and concludes about person identity. SVM implements one to all matching for person identification in this proposed work.

The main fortes of this algorithm are reliability, robustness, accurate. The proposed work is wireless, easy to implement and cost effective with overall accuracy of our algorithm is 94.8% percent with 1.327 sec in Intel Pentium Dual Core processor using Matlab R2011a. Also the system is rotation invariance, as in real time, person can work at any angle with respect to Kinect sensor.

The paper is presented as follows. The Kinect sensor is introduced in section II. Section III explains feature extraction and analysis. Experimental results are clearly explained in section IV. Section V concludes with future work.

II. Kinect Sensor

The Kinect sensor [3] consists of RGB camera, IR camera [23] , IR projector and microphone array. It basically looks like a webcam (8-bit VGA resolution of 640 × 480 pixels with a Bayer color filter) as shown in Fig. 1. Its long horizontal bar is placed with a motorized base and tilt. The Kinect sensor [24] tracks human motion [25] using a skeleton [26] with three dimensional co-ordinates within a finite range of distance (roughly 1.2 to 3.5m or 3.9 to 11ft). This is achieved using the visible IR (which is the depth sensor [27], by which z direction value for each joint is obtained) and RGB cameras [28]. The

ISBN: 978-81-923777-9-7

skeleton produced by the Kinect sensor [29] has twenty body joints as depicted in Fig. 2. The Kinect can simultaneously detect two persons skeleton [30].

The background, dress color and lighting of the room are irrelevant for skeleton detection using Kinect sensor. As it can recognized human motion in a very wide range of surrounding physical conditions [31].

Figure 1. Kinect sensor

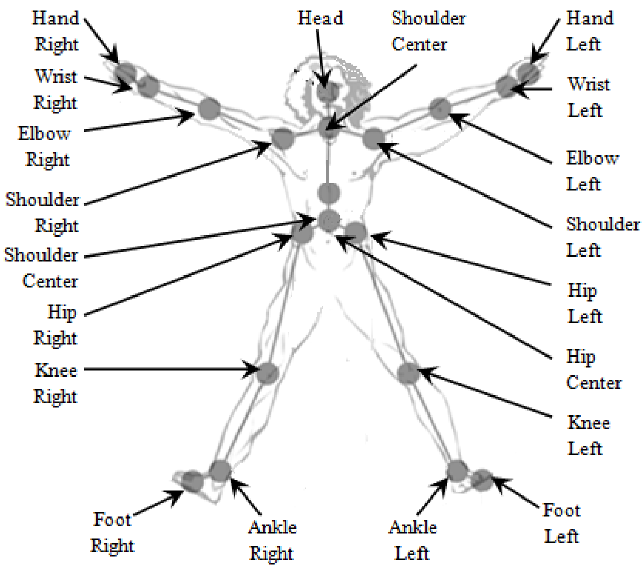

Figure 2. Twenty bosy joints of a skeleton

III. Feature Extraction and Analysis

In our proposed algorithm, total twenty features are extracted for each person from the skeleton information obtained using Kinect sensor. Among these features, twelve features are calculated for each frame [32] and all the other features are extracted between two consecutive frames. Fig. 3 explains the walking skeleton movement for a single person. The person starts from a standing posture before started walking in front of the Kinect sensor. In this figure the person walks parallel to the Kinect sensor. All the features extracted for this novel work is based on the concept of rotation invariance.

A. Calculation of Feature 1 Using Euclidean Distances and Angle

Height of the person [33] for each frame is measured and as the rotation of the person with respect to Kinect sensor varies widely for this proposed walk, so there is a change in height of the person in each frame. This height is not the actual height of the person but the approximated height calculated using Kinect sensor. For height calculation, six distances and one angle value is taken into account. During walking, the person bends and hence cumulating distance is the measure for the height of the person.

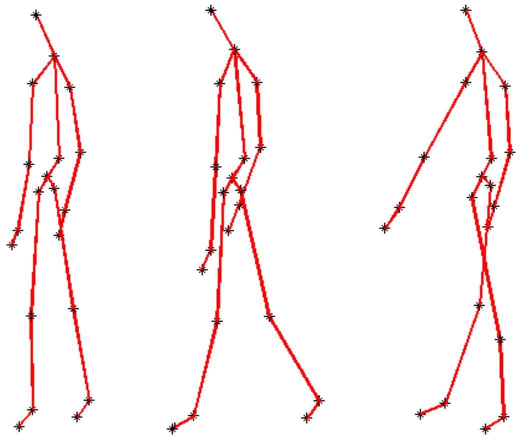

Figure 3. Example of skeleton formed during walking of a person

At first, the Euclidean distances between head (H) and shoulder centre (SC), shoulder centre and spine (S), spine and hip centre (HC), hip centre and hip right (HR), hip right and knee right (KR), knee right and ankle right (AR) are computed for a frame no i using (1-6). The angle between the two vectors form by the initial and final body joints using spine, hip centre and hip centre, hip right is also figured out by (7). The distance between ankle right and foot right (FR) is very small, thus it is neglected.

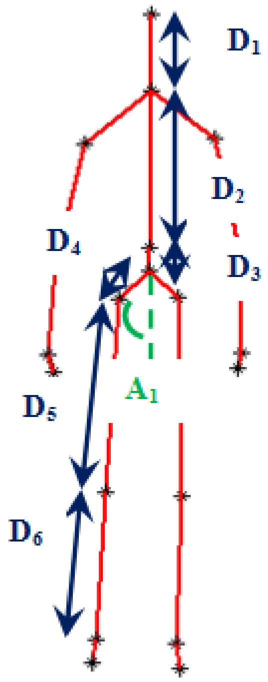

Figure 4. Explanation of height calculation

ISBN: 978-81-923777-9-7

$$D_1 = \|H_i - SC_i\| \quad (1)$$

$$D_2 = \|SC_i - S_i\| \quad (2)$$

$$D_3 = \|S_i - HC_i\| \quad (3)$$

$$D_4 = \|HC_i - HR_i\| \quad (4)$$

$$D_5 = \|HR_i - KR_i\| \quad (5)$$

$$D_6 = \|KR_i - AR_i\| \quad (6)$$

$$A_1 = angle\langle (S_i, HC_i), (HC_i, HR_i) \rangle \quad (7)$$

where angle calculates the angle in degree between the two vectors specified within the first brackets.

$$Feature_{1,i} = D_1 + D_2 + D_3 + (D_4 \times A_1) + D_5 + D_6 \quad (8)$$

As the height of the person is crucial information for identifying a person, so we have computed the approximated height of the person for each frame.

The angle calculation between two vectors is explained below. Let, the co-ordinates of first three vertices be (a_1, b_1,c_1), (a_2,b_2,c_2) and (a_3,b_3,c_3) respectively. The vectors *v1* and *v2* formed by these three vertices are shown in Fig. 3.

$$v1 = (a_1 - a_2)\vec{i} + (b_1 - b_2)\vec{j} + (c_1 - c_2)\vec{k} \quad (9)$$

$$v2 = (a_3 - a_2)\vec{i} + (b_3 - b_2)\vec{j} + (c_3 - c_2)\vec{k} \quad (10)$$

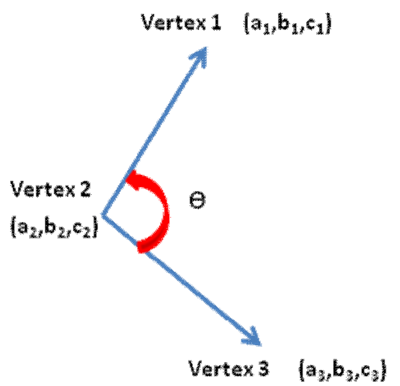

Figure 5. Calculation of angle between vertices

The angle between two vectors is calculated by (11).

$$\text{angle} = \frac{\text{atan2(norm(cross(v1,v2),dot(v1,v2)))}}{\text{pi}} \times 180^0 \quad (11)$$

where norm function returns a matrix.

Cross product or vector product is defined by

$$\vec{M} \times \vec{N} = mn \sin \theta \vec{P} \quad (12)$$

where m and n are the magnitudes of $\vec{M}$ and $\vec{N}$ vectors correspondingly, θ is the angle between the two vectors and $\vec{P}$ be the unit vector normal to the plane $\vec{M}$ and $\vec{N}$.

Dot product or scalar product is defined by (13).

$$\vec{M} \bullet \vec{N} = mn \cos \theta \quad (13)$$

atan2 is the arctangent angle whose range is $\left(-\pi/2, +\pi/2\right)$.

$$\operatorname{atan2}(y,x) = \begin{cases} \arctan\left(\frac{y}{x}\right) & x > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & y \geq 0, x < 0 \\ \arctan\left(\frac{y}{x}\right) - \pi & y < 0, x < 0 \\ +\pi/2 & y > 0, x = 0 \\ -\pi/2 & y < 0, x = 0 \\ undefined & y = 0, x = 0 \end{cases} \quad (14)$$

B. Calculation of Feature 2 to 9 Using Angular Displacements

These features are the angular displacements of eight body joints, namely elbow right (ER), elbow left (EL), wrist right (WR) and wrist left (WL) from the two hands and knee right (KR), knee left (KL), ankle right (AR) and ankle left (AL) from two legs of the person. The Kinect sensor captures the skeleton at 30 frames/sec rate, thus the timing difference between any two consecutive i and (i+1) frames is given in (15).

$$dt = 1/30 \quad (15)$$

$$Feature_{2,i} = \|ER_{i+1} - ER_i\| / dt \quad (16)$$

$$Feature_{3,i} = \|EL_{i+1} - EL_i\| / dt \quad (17)$$

$$Feature_{4,i} = \|WR_{i+1} - WR_i\| / dt \quad (18)$$

$$Feature_{5,i} = \|WL_{i+1} - WL_i\| / dt \quad (19)$$

$$Feature_{6,i} = \|KR_{i+1} - KR_i\| / dt \quad (20)$$

$$Feature_{7,i} = \|KL_{i+1} - KL_i\| / dt \quad (21)$$

ISBN: 978-81-923777-9-7

$$Feature_{8,i} = \|AR_{i+1} - AR_i\| / dt \tag{22}$$

$$Feature_{9,i} = \|AL_{i+1} - AL_i\| / dt \tag{23}$$

During walking, the angular displacements between those eight body joint changes from frame to frame, thus these eight features lead to a promising pattern for each person.

C. Calculation of Features 10 and 11 Using Displacements

Every person walks according to his (/her) own speeds if not in normal day to day situation. The distance between two ankle right (AR), ankle left (AL) and knee right (KR), knee left (KL) for a footstep (i.e., for each frame) contributes a significant information about person recognition.

$$Feature_{10,i} = \|AR - AL\| \tag{24}$$

$$Feature_{11,i} = \|KR - KL\| \tag{25}$$

D. Calculation of Features 12 to 20 Using Angles

During walking, the angle of different body joints of the person form angles, these angles are extracted for production of last nine features. Features 12 to 16 are computed with respect to the spine of the person.

$$Feature_{12,i} = angle\langle (SR_i, ER_i), (SC_i, S_i) \rangle \tag{26}$$

$$Feature_{13,i} = angle\langle (SL_i, EL_i), (SC_i, S_i) \rangle \tag{27}$$

$$Feature_{14,i} = angle\langle (ER_i, WR_i), (SC_i, S_i) \rangle \tag{28}$$

$$Feature_{15,i} = angle\langle (EL_i, WL_i), (SC_i, S_i) \rangle \tag{29}$$

$$Feature_{16,i} = angle\langle (KR_i, HR_i), (KR_i, AR_i) \rangle \tag{30}$$

$$Feature_{17,i} = angle\langle (KL_i, HL_i), (KL_i, AL_i) \rangle \tag{31}$$

$$Feature_{18,i} = angle\langle (HC_i, KR_i), (HC_i, KL_i) \rangle \tag{32}$$

$$Feature_{19,i} = angle\langle (SC_i, H_i), (SC_i, S_i) \rangle \tag{33}$$

$$Feature_{20,i} = angle\langle (SC_i, H_i), (SC_i, SR_i) \rangle \tag{34}$$

Feature 19 and 20 are calculated to track whether the person nods his (\her) head while walking.

IV. Experimental Results

For this research work, we have created a walking dataset from fifteen persons and for each person five different walking dataset along different angles with respect to Kinect sensor is taken. Thus total information present in the walking dataset is 15 × 5. Each walking data is collected for 5 sec at 30 frames/sec rate [34]. So for each walking pattern for a specific person, we have to calculate twenty features for 5 × 30. Total number of features obtained for a walking data is 5 × 30 × 20. In the classification stage, the training to testing ratio is taken as 4:1 for multi class SVM. The recognition rate found is 94.8% experimentally with average time of 1.327 sec in an Intel Pentium Dual Core processor running Matlab R011b [35]. One sample output obtained for subject 1 is depicted in Table 1 for frame no 30.

TABLE I. Sample Features for Subject 1 for Frame No 30

Feature No	Feature Value
1	1.48872 (in meter)
2	1.49709 (in meter)
3	1.41508 (in meter)
4	1.33307 (in meter)
5	0.554023 (in meter)
6	0.630403 (in meter)
7	0.158728 (in meter)
8	1.23169 (in meter)
9	0.707106 (in meter)
10	0.648353 (in meter)
11	0.352995 (in meter)
12	11.0206 (in degree)
13	20.5092 (in degree)
14	12.5126 (in degree)
15	20.5092 (in degree)
16	13.6156 (in degree)
17	7.89635 (in degree)
18	32.3477 (in degree)
19	4.07476 (in degree)
20	55.0363 (in degree)

V. Conclusion and Future Work

The efficient and easy to implement approach of this research work describes person recognition from walking patterns using support vector machine. This system is robust and is implemented for rotation invariance purposes also. The proposed work produces high recognition rate of 94.8% and the time required to identify the person is 1.327 sec. Thus real time application is possible virtue of this work.

One disadvantage of the process is that during recognition, the walking speed for a person is nearly constant. The variation of speed can produce erroneous results. In future, we will try to deal with this problem.

Acknowledgment

We would like to thank University Grant Commission, India, University of Potential Excellence Program (Phase II) in Cognitive Science, Jadavpur University.

ISBN: 978-81-923777-9-7

REFERENCES

[1] A. M. Burton and V. Bruce, “Naming faces and naming names: Exploring an interactive activation model of person recognition,” *Memory*, vol. 1, no. 4, pp. 457–480, 1993.

[2] J. K. Aggarwal and Q. Cai, “Human motion analysis: A review,” in *Nonrigid and Articulated Motion Workshop, 1997. Proceedings., IEEE*, 1997, pp. 90–102.

[3] Z. Zhang, “Microsoft kinect sensor and its effect,” *Multimedia, IEEE*, vol. 19, no. 2, pp. 4–10, 2012.

[4] M. P. Murray, R. C. Kory, and B. H. Clarkson, “Walking patterns in healthy old men,” *Journal of Gerontology*, vol. 24, no. 2, pp. 169–178, 1969.

[5] H.-D. Yang, A.-Y. Park, and S.-W. Lee, “Gesture spotting and recognition for human–robot interaction,” *Robotics, IEEE Transactions on*, vol. 23, no. 2, pp. 256–270, 2007.

[6] C. Cortes and V. Vapnik, “Support vector machine,” *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[7] R. P. W. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. De Ridder, D. M. J. Tax, and S. Verzakov, “A matlab toolbox for pattern recognition,” *PRTools version*, vol. 3, 2000.

[8] M.-F. Balcan, A. Blum, P. P. Choi, J. D. Lafferty, B. Pantano, M. R. Rwebangira, and X. Zhu, “Person identification in webcam images: An application of semi-supervised learning,” 2005.

[9] L. Wang, W. Hu, and T. Tan, “Recent developments in human motion analysis,” *Pattern recognition*, vol. 36, no. 3, pp. 585–601, 2003.

[10] R. Brunelli and D. Falavigna, “Person identification using multiple cues,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 17, no. 10, pp. 955–966, 1995.

[11] C. Yam, M. S. Nixon, and J. N. Carter, “Automated person recognition by walking and running via model-based approaches,” *Pattern Recognition*, vol. 37, no. 5, pp. 1057–1072, 2004.

[12] C. Yam, M. S. Nixon, and J. N. Carter, “On the relationship of human walking and running: Automatic person identification by gait,” in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, 2002, vol. 1, pp. 287–290.

[13] T. Choudhury, B. Clarkson, T. Jebara, and A. Pentland, “Multimodal person recognition using unconstrained audio and video,” in *Proceedings, International Conference on Audio-and Video-Based Person Authentication*, 1999, pp. 176–181.

[14] C. Sanderson, *Biometric person recognition: Face, speech and fusion*. VDM Publishing, 2008.

[15] K. Fukui and O. Yamaguchi, “Person identification apparatus and method using concentric templates and feature point candidates.” Google Patents, 09-Nov-1999.

[16] T. Leyvand, C. Meekhof, Y.-C. Wei, J. Sun, and B. Guo, “Kinect identity: Technology and experience,” *Computer*, vol. 44, no. 4, pp. 94–96, 2011.

[17] T. B. Moeslund and E. Granum, “A survey of computer vision-based human motion capture,” *Computer Vision and Image Understanding*, vol. 81, no. 3, pp. 231–268, 2001.

[18] J. Solaro, “The Kinect Digital Out-of-Box Experience,” *Computer*, pp. 97–99, 2011.

[19] L. Snidaro, G. L. Foresti, and L. Chittaro, “Tracking human motion from monocular sequences,” *International Journal of Image and Graphics*, vol. 8, no. 03, pp. 455–471, 2008.

[20] C. Herrera and J. Kannala, “Joint depth and color camera calibration with distortion correction,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 10, pp. 2058–2064, 2012.

[21] D. M. J. Tax and R. P. W. Duin, “Support vector domain description,” *Pattern recognition letters*, vol. 20, no. 11, pp. 1191–1199, 1999.

[22] J. A. K. Suykens and J. Vandewalle, “Least squares support vector machine classifiers,” *Neural processing letters*, vol. 9, no. 3, pp. 293–300, 1999.

[23] H. Zhuang, B. Zhao, Z. Ahmad, S. Chen, and K. S. Low, “3D depth camera based human posture detection and recognition Using PCNN circuits and learning-based hierarchical classifier,” in *Neural Networks (IJCNN), The 2012 International Joint Conference on*, 2012, pp. 1–5.

[24] J. Tong, J. Zhou, L. Liu, Z. Pan, and H. Yan, “Scanning 3d full human bodies using kinects,” *Visualization and Computer Graphics, IEEE Transactions on*, vol. 18, no. 4, pp. 643–650, 2012.

[25] C. Nakajima, M. Pontil, B. Heisele, and T. Poggio, “Full-body person recognition system,” *Pattern recognition*, vol. 36, no. 9, pp. 1997–2006, 2003.

[26] J. Clark, “Object digitization for everyone,” *Computer*, pp. 81–83, 2011.

[27] J.-H. Cho, K. H. Lee, and K. Aizawa, “Enhancement of Depth Maps With Alpha Channel Estimation for 3-D Video,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 6, no. 5, pp. 483–494, 2012.

[28] L. Spinello and K. O. Arras, “People detection in RGB-D data,” in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, 2011, pp. 3838–3843.

[29] F. Ryden, “Tech to the Future: Making a,” *Potentials, IEEE*, vol. 31, no. 3, pp. 34–36, 2012.

[30] K. Yun, J. Honorio, D. Chattopadhyay, T. L. Berg, and D. Samaras, “Two-person interaction detection using body-pose features and multiple instance learning,” in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, 2012, pp. 28–35.

[31] C. Dal Mutto, P. Zanuttigh, and G. M. Cortelazzo, “Fusion of Geometry and Color Information for Scene Segmentation,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 6, no. 5, pp. 505–521, 2012.

[32] S. Wachter and H.-H. Nagel, “Tracking of persons in monocular image sequences,” in *Nonrigid and Articulated Motion Workshop, 1997. Proceedings., IEEE*, 1997, pp. 2–9.

[33] C. BenAbdelkader, R. Cutler, and L. Davis, “Person identification using automatic height and stride estimation,” in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, 2002, vol. 4, pp. 377–380.

[34] A. Agarwal and B. Triggs, “Recovering 3D human pose from monocular images,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 1, pp. 44–58, 2006.

[35] S. Theodoridis, A. Pikrakis, K. Koutroumbas, and D. Cavouras, *Introduction to Pattern Recognition: A Matlab Approach: A Matlab Approach*. Academic Press, 2010.

ISBN: 978-81-923777-9-7