Proceedings of the 3rd
RSI International Conference on Robotics and Mechatronics
October 7-9, 2015, Tehran, Iran

Distinguishing Tip-Toe Walking from Normal Walking Using Skeleton Data Gathered By 3D Sensors

Mohsen Ebrahimi[1], Meysam Feghi[1], Hadi Moradi[1,2], Maryam Mirian[1,3], Hamidreza Pouretemad[4]
[1]Advanced Robotics and Intelligent Systems Laboratory, CIPCE, School of ECE, College of Engineering, University of Tehran
[2]Intelligent Systems Research Institute, SKKU, South Korea
[3]Department of Electrical and Computer Engineering, University of British Colombia, Vancouver, Canada
[4]School of Psychology, Shahid Beheshti University, Tehran, Iran
Mohsen.Ebrahimi@ut.ac.ir, Me.Feghhi@gmail.com, moradih@ut.ac.ir, MMirian@ut.ac.ir, h-pouretemad@sbu.ac.ir

***Abstract*— in this paper, we propose a markerless based approach to differentiate between two different human walking gaits, specifically, the normal walking versus tip toe walking. The markerless based approaches are suitable in many applications, such as screening of autistic children from normal ones based on their body movement patterns or detecting walking irregularities. We use a 3D Sensor to collect 3D skeletal data of subjects that represents their behavioral spatial data. Several features have been extracted and feature selection methods were used to determine the best subset of these features. The proposed classification approach has been implemented and tested on a group of 75 people reaching 86% CCR between tip toe walking and normal walking.**

***Keywords*— Gait Recognition, Motion Capture, Kinect, 3D Sensor, Tip-toe Walking;**

I. Introduction

Gait recognition and motion analysis have been studied for the past 3 decades especially for medical applications [1] [2][3][4] and bio-inspired robotics [5]. Marker-based approaches have been widely used since they can provide fairly accurate measurements using off-the shelf cameras or dedicated motion capture cameras [6][7]. However, these marker-based approaches are difficult to use to capture daily life activities in unstructured environments. Furthermore, they are hard to be used on general subjects. Consequently, the markerless-based approaches become important [8][9]. The invention of off-the-shelf 3D sensors such as Microsoft Kinect™, which can detect and capture the motion of a subject's skeleton, has provided a great opportunity to develop markerless-based approaches. There have been many studies working on motion capture and detection using these sensors [10] [11] [12].

In this work, we are motivated by a specific application of gait recognition, i.e. using gait patterns to classify subjects and determine possible health issue. One of these health issues is autism spectrum disorder (ASD) with a high progression rate in recent decades. Autism Spectrum Disorder is a group of complex disorders which is considered as a neurodevelopmental disorder ([13] [14] [15]). Autism have wide range of different disabilities in social behaviors which can result in the rejection from the community and lack of proper learning progress. Autistic children like stereotypical movement such as tip-toe walking which can be used to screen them.

Since the successful therapy of ASD depends on early detection of the disorder, developing early screening systems becomes crucial. Unfortunately, the current screening methods rely on expert evaluation of the children. The expert evaluation is labor intensive, not always available, and expensive. Consequently, the automatic approaches have been studied in the recent decade. For example in [16] EEG is taken from babies and applied machine learning techniques for screening autistic babies. In two projects, LENA [17] and [18], they focused on sound descriptors and used it to screen children with ASD. In [19], a marker-based approach is proposed to differentiate between autistic and non-autistic children performing a specific grab and walk task. In [20] [21] and [22] studies on ASD diagnosis based on Karyotype and CMA genetics tests have been proposed. The need for experts and its costs are disadvantages of this methods. Blood test [23] also has been proposed but it is at early stages of its test and needs actual blood test which may be expensive, needs specific level of expertise, and costly.

To overcome the shortcomings of the above approaches, we have developed a markerless based approach which can be used to distinguish between tip-toe walking and normal walking. We use and off-the-shelf 3D sensor which makes the approach suitable for many applications. The approach has been successfully implemented and tested on 75 normal subjects. As one of its applications, it would be tested as an autism screening tool.

The rest of this paper is organized as follows: In part II we introduce our approach to this problem. In part III we introduce

978-1-4673-7234-3/15/$31.00 ©2015 IEEE

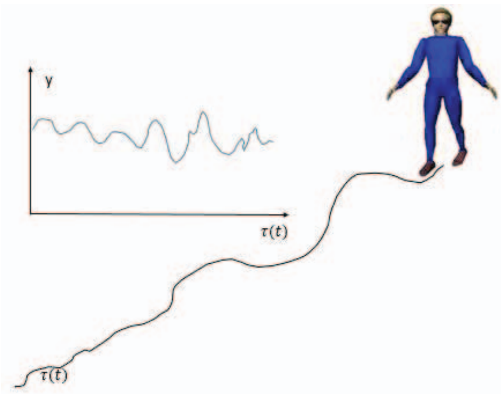

Fig. 1: The height of the hip (y) and the path of a person

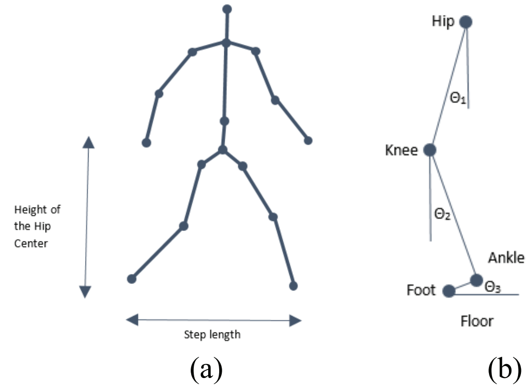

Fig. 2: a) the walking configuration of a subject, b) the 3 angles for a leg.

our test structure and system in addition to results. In part IV we inference on obtained results.

II. The Proposed Approach

The intuition behind gait recognition using skeleton data is inspired from human gait recognition in which a human differentiates between tip toe walking and normal walking by observing the body and leg movements. For instance, Fig. 1 shows a human walking and his hip movement along vertical direction. We expect to have different patterns for tip toe and normal walking. In the following the feature extraction and selection processes have been explained.

A. Feature Extraction

The feature extraction has been performed using expert knowledge to propose possible features. The expert assisted feature extraction was based on watching normal and tip-toe walkers and analyzing their motion patterns in time and frequency domains.

Fig. 2(a) shows the height of the hip center and the step length while Fig. 2(b) shows the 3 joint angles for a leg. These variables are selected by an expert to extract the features from. The extracted features are divided into time and frequency domain categories. It should be mentioned that the step length is considered as the distance between the left ankle and the right ankle at their maximum distance along $\tau(t)$.

1) Time-Domain Features:

Table I shows 136 features extracted in time domain, in which HH stands for Height of the Hip. Each row shows the general title for a feature in which the mean, max, standard deviation and energy of the position, velocity, and acceleration of the feature are considered.

It should be mentioned that many features would depend on the height of the subject. For instance, although the height of the hip of a subject would stay higher in tip-toe walking compared to normal walking, however, its value depends on the height of the subject. Consequently, to normalize it, the ratio of the height of the hip to the height of the subject is considered. This has been done for feature sets TD1, TD2, TD7, TD8, and TD9. In TD1, c is a constant determined empirically using experimental data.

Two other visible differences between tip-toe walking and normal walking are in the height of the heel and in the step length. In tip-toe walking, the height of the heel is higher and the length of the step is shorter [26]. This fact has been considered in TD2 and TD3.

TABLE I. List of Time-Domain features

Feature Subset ID	*Feature Description*	*# of features*
TD1	| HH| / (shoulder-to-ankle distance + c)	12
TD2	difference between ankles and foots/height	12
TD3	difference between ankles and foots/stalk size	12
TD4	The average of each of hip, knee, and ankle angles' mean, max, STD, and energy over a gait cycle [27].	48=4*4*3
TD5	3 right leg angles when a gait occurred	12=4*3
TD6	3 right leg angles in all over of the path	12=4*3
TD7	HH/head-to-ankle distance	12
TD8	HH/hip-to-ankle distance	12
TD9	HH/step length	4

Finally, in TD4 the mean, max, STD, and energy of the each joint angle's mean, max, STD, and energy, in a given gait cycle, is considered as a feature. In these features the velocity and acceleration of the joints have not been considered. It should be mentioned that we have assumed similar pattern between the right and left legs. Consequently, the features of one leg would be enough for feature extraction analysis. We use mean, max, STD and energy because these statistics represent unique behavior of the considered parameters.

2) Frequency-Domain Features:

The periodic nature of gaits suggests that frequency domain features may help in the discrimination between tip-toe walking and normal walking.

For instance, we expect tip-toe walkers show different dominant frequency during walking. Thus we extracted 220 frequency features as listed in Table II.

TABLE II. LIST OF FREQUENCY-DOMAIN FEATURES

Feature Subset ID	***Feature Description***	***# of features***
FD1	The top 5 highest and lowest frequencies of position, velocity, and acceleration of HH	30=10*3
FD2	The top 5 highest and lowest frequencies of position, velocity, and acceleration of height of the ankles	60=10*3*2
FD3	The top 5 highest and lowest frequencies of position, velocity, and acceleration of height of the knees	60=10*3*2
FD4	The top 5 highest and lowest frequencies of step length	10=5+5
FD5	The maximum and minimum 5 frequencies of 6 angles of foots(3 for right and 3 for left)	60=10*3*2

B. Feature Selection

Since there is a large number of features and small number of samples, we first evaluated each feature set's discrimination power. Then the best feature sets are selected and feature conditioning, i.e. feature selection and feature combination, is applied to achieve better results. Finally, the best combination of features is used for classification.

C. Classification

For the classification purpose we tested several methods. We selected linear SVM as a simple well-suited method. MLP is chosen as a general function approximator (GFA) and Random Forest as one of the best ensemble solutions. It has been shown that linear Support Vector Machine (SVM) with stochastic gradient descent learning algorithm works the best. To train classifiers we use 10-fold cross-validation. We also used MLP with BFGS [24] quasi newton back propagation, to benefit from its global optimal solution, and Random Forest [25] supervised learning method, which is one of the best ensemble methods.

III. IMPLEMENTATION AND RESULTS

The novelty of this work is in determining features and developing a classification method that can use a markerless based, off-the-shelf, and low cost 3D sensor. Thus we selected Microsoft® Kinect® from among good alternatives such as Asus Xtion since it has been widely used and there have been good APIs, such as OpenNI, for research. It should be noted that our proposed approach is not limited to this solution and can be implemented on any 3D sensor. For the purpose of skeleton data collection we used Microsoft® Kinect® SDK on windows over OpenNI because of its accuracy of skeleton estimation. Due to many restrictions, such as the very large dataset and privacy issues, we have developed an in-house application to save data collected by Kinect, cover the face of the subjects, and keep a compressed video (up to 97% compression compared to the original Kinect video) for expert evaluation (Fig. 3).

A. The experimental setup

Fig. 4 shows the experimental setup in which the subjects are asked to walk in front of the Kinect diagonally. Each subject had four set of data, two were walking from left to right and two were walking from right to left, with respect to the Kinect's X axis. The walk on each side consisted of one tip-toe walking and one normal walking.

If a subject walks toward the Kinect, i.e. along the Z axis of the Kinect which is considered the normal configuration using Kinect for applications such as games, then detecting the tip-toe walking becomes hard since the ankle is not easily detectable. On the other hand, if a subject walks along the X direction, i.e. parallel to the Kinect, then the API can become confused and make mistakes in correctly detecting the left and right foot from each other. Thus the data can be fairly unreliable. Consequently, walking diagonally in XZ plane is the best configuration to detect the needed features for tip-toe walking classification. It

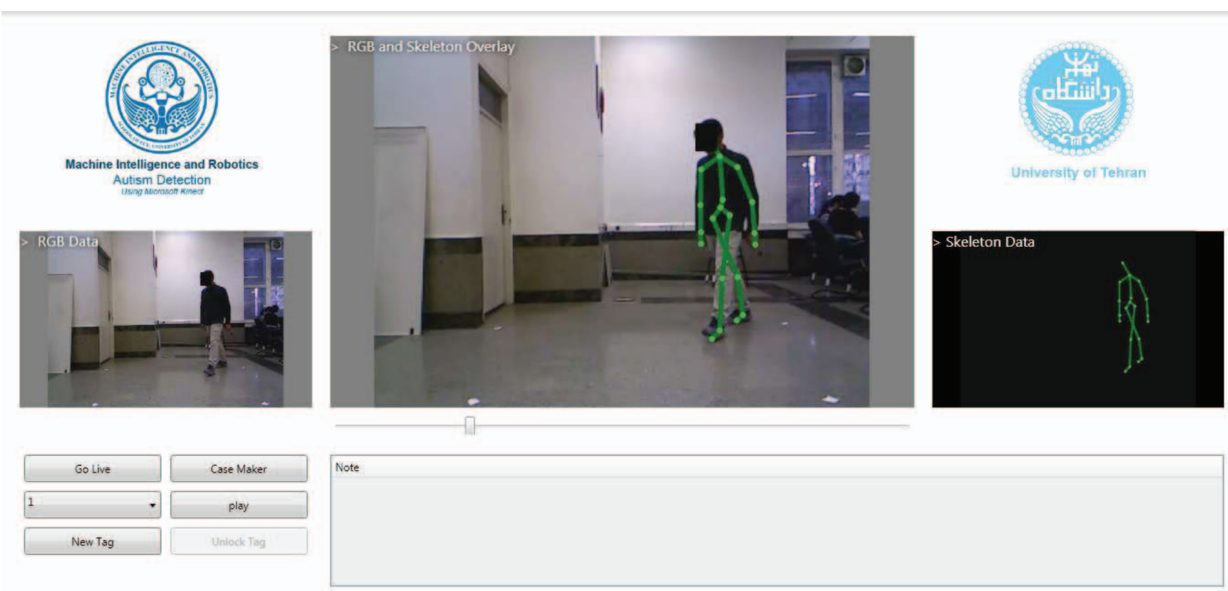

Fig. 3: The system developed in the lab which hides the face of subjects to avoid privacy issues (the left and middle video stream). The left video stream is the raw RGB data while the middle one shows the skeleton overlapped on the raw image. The right video stream shows the extracted skeleton only. It is possible to extract specific part of the video streams and comment them (the case maker and new tag buttons).

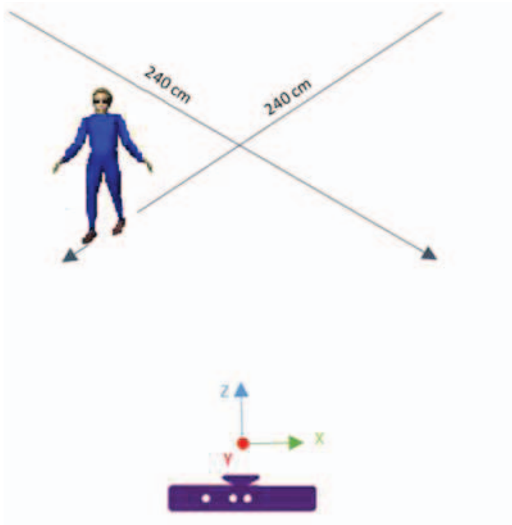

Fig. 4: The test setup in which the subjects walk on front of the 3D sensor

should also be noted that in a natural setup, the subjects may move in any direction and it is better to not limit them to a specific direction. In such cases, the data which is closer to diagonal walking can be used for classification.

In general, the trajectory $\tau(t)$ (Eq. 1) in XZ plane is considered and the motion in X and Z axes are mapped into this trajectory (Fig. 1). In other words, there would be a two dimensional data, i.e. y and $\tau(t)$, to extract the features from.

$$\tau: t \in [0, T] \rightarrow \tau(t) = \sqrt{x(t)^2 + z(t)^2} \qquad (1)$$

B. The subjects

To collect data, 75 students aged between 18-28 years old were asked to walk on front of the sensor. In cases that the subjects did not correctly performed tip-toe walking, which were hard to be detected even by humans and we call them low tip-toe walkers, their data were excluded from the analysis. Also, the subjects who had bulky pants that interfered with the skeleton detection algorithm, which we call them mini skeleton samples, were eliminated from the analysis. Consequently, the total of 207 tip-toe walking samples were collected. The number of normal walking samples were higher than 207 which was enough for our analysis.

IV. RESULTS

To select the best features and train our classifier, we apply the z-score normalization and 10-fold cross-validation on the train data that made up by 80 percent of full data. The procedure repeated 100 times and in each iteration we randomize the sample data. Table III shows the classification results, i.e. CCR's mean and STD, using SVM, MLP, and Random Forest (RF) classifiers using each feature set alone. Overall SVM performs better than MLP and Random Forest. It can be seen that feature sets TD4, TD5, and TD9 perform better than the others which suggest that we have to consider these three feature sets for feature selection phase.

TABLE III. CCR FOR EACH SET OF FEATURES

Correct Classification Rate				
Feature set ID	*Number of Features*	*Mean(STD)*		
		Linear SVM	*MLP*	*RF*
TD1	12	61.6(8.2)	55.8(9.1)	57.7(8.2)
TD2	12	63.5(8.6)	59.1(10.2)	59.8(9.2)
TD3	12	65.0(8.3)	60.6(8.6)	60.1(9.4)
TD4	48	81.2(6.6)	81.5(5.7)	79.2(6.8)
TD5	12	79.7(7.1)	77.6(8.1)	74.1(6.7)
TD6	12	63.1(9.2)	62.1(9.2)	66.3(8.6)
TD7	12	60.7(8.1)	57.2(8.8)	55.9(9.8)
TD8	12	48.1(10.6)	48.0(9.0)	49.3(9.1)
TD9	4	70.7(7.4)	67.8(8.6)	64.6(9.8)
FD1	60	53.1(9.6)	53.7(9.5)	58.6(8.3)
FD2	60	49.7(8.8)	50.1(8.8)	55.7(7.1)
FD3	30	53.0(9.1)	51.8(9.3)	56.8(8.7)
FD4	60	50.9(9.1)	51.3(9.4)	54.3(8.7)
FD5	10	51.4(9.0)	51.2(9.3)	56.7(8.5)

After feature set evaluation, these three set of features are further evaluated using feature conditioning concepts, i.e. feature selection and feature combination, to achieve better results. First we apply feature selection on the top 3 feature sets and determined the most informative features in the current set using best first feature selection method. Afterward, we use feature combination of these feature sets. Table IV shows the result of feature selection, i.e. rows 1 to 3, in which the best features in each feature set is selected. The results of the feature combination is shown in rows 4 to 7.

TABLE IV. THE FEATURE SELECTION AND FEATURE COMBINATION RESULTS

Correct Classification Rate				
Feature Subset ID	*Number of Features*	*Mean(STD)*		
		Linear SVM	*MLP*	*RF*
TD4'	5	80.7(5.5)	78.9(6.1)	81.1(6.1)
TD5'	4	82.3(6.1)	81.6(6.6)	77.9(7.3)
TD9'	3	69.3(6.8)	67.4(6.9)	64.6(8.7)
TD4'+ TD5'	9	86.8(5.7)	84.1(6.3)	84.8(4.9)
TD4'+ TD9'	8	78.8(5.9)	76.9(7.4)	81.3(6.2)
TD5'+ TD9'	7	80.9(7.1)	78.5(8.1)	76.0(6.7)
TD4'+TD5'+TD9'	12	83.4(5.7)	80.7(6.2)	81.9(5.8)

V. DISCUSSION

As it can be seen from Table III, the frequency-domain features do not have very discriminant power. The most important feature sets belong to the leg angles (TD4 and TD5) and the ratio of the height of the hip over step length (HH/step length). Table V shows the list of final features selected in the feature conditioning process.

TABLE V. THE FEATURE SELECTED IN THE FEATURE CONDITIONING PROCESS

Selected TD4 features	*Selected TD5 features*	*Selected TD9 features*
Max of the means of θ_2	Mean of θ_2	Mean of HH/step length
STD of the max of θ_2	Mean of θ_3	STD of HH/Step length
Energy of the mean of θ_2	Max of θ_3	Energy of HH/Step length
STD of energy of θ_3	STD of θ_1	
Energy of STD of θ_3	STD of θ_2	

These results matches human intuition since the best features are about the knee and ankle angles, i.e. θ_2 and θ_3 respectively. For instance, max of the means of θ_2 shows the maximum of the mean of knee angle over several steps. In tip-toe walking, the knee angle should be much less than the knee angle in normal walking.

Also, it is expected that the ratio of the height of the hip to the step length would be a discriminative feature since in tip-toe walking the hip moves more vertically, which is represented in the height of the hip, than horizontally, which is represented in the step length. In contrast, in normal walking the height of the hip moves less vertically than horizontally. Consequently, it's mean, STD, and energy become important features for tip-toe walking classification.

It should be noted that overall the combination of the joint angle features (subsets of TD4 and TD5) are more discriminative than the hip movement features.

VI. CONCLUSION

In this paper we introduced a markerless gait recognition approach in which the tip-toe walking can be discriminated from the normal walking. The approach has been implemented on an off-the-shelf non-intrusive 3D sensor, to distinguish normal walking from tip-toe walking which can be used in applications such as healthcare. The system has been tested on 75 normal people walked in two different gait patterns on front of the 3D sensor and their movement patterns have been extracted. We performed feature extraction and feature selection to determine the best subset of features that can perform the classification. We have currently reached 86.84% CCR using 9 features.

In the next step, to improve the classification rate and reduce STD, we would perform further feature selection, classification, and partitioning approaches to determine the most suitable feature set.

Also, we would test the approach on autistic children. We expect to have good results due to the symmetric gaits of children with ASD [28]. Finally, the in-house system developed for this purpose would be improved to be more intelligent to ignore low tip-toe and mini-skeleton walking. Finally, the other stereotypical patterns such as flapping and turning.

ACKNOWLEDGMENT

This research is partially funded by Cognitive Sciences and Technologies Council (COGC) of Iran. We should thanks from Dr Hamid Reza Pour Etemad and the Centre for the Treatment of Autistic Disorders (CTAD) that help us to fulfil tests and observations without any hesitations. Special thanks to Dr. Babak Nadjar Araabi for his great consultant for the test bed design. We want to thanks all the members of the Advanced Robotics and Intelligent Systems (ARIS) Lab specially Pegah Soleiman and Hossein Khandani for their generous help throughout the project.

REFERENCES

[1] D. Gouwanda and S. M. N. A. Senanayake, "Emerging Trends of Body-Mounted Sensors in Sports and Human Gait Analysis," in 4th Kuala Lumpur International Conference on Biomedical Engineering 2008, vol. 21, N. A. Abu Osman, F. Ibrahim, W. A. B. Wan Abas, H. S. Abdul Rahman, and H.-N. Ting, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 715–718.

[2] M. Yeasin and S. Chaudhuri, "Development of an Automated Image Processing System for Kinematic Analysis of Human Gait," Real-Time Imaging, vol. 6, no. 1, pp. 55–67, Feb. 2000.

[3] S. R. Simon, "Quantification of human motion: gait analysis—benefits and limitations to its application to clinical problems," J. Biomech., vol. 37, no. 12, pp. 1869–1880, Dec. 2004.

[4] W. Zijlstra and K. Aminian, "Mobility assessment in older people: new possibilities and challenges," Eur. J. Ageing, vol. 4, no. 1, pp. 3–12, Feb. 2007.

[5] A. T. SprÃ¶witz, M. Ajallooeian, A. Tuleu, and A. J. Ijspeert, "Kinematic primitives for walking and trotting gaits of a quadruped robot with compliant legs," Front. Comput. Neurosci., vol. 8, Mar. 2014.

[6] M. Field, D. Stirling, F. Naghdy, and Z. Pan, "Motion capture in robotics review," 2009, pp. 1697–1702.

[7] T. B. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-based human motion capture and analysis," Comput. Vis. Image Underst., vol. 104, no. 2–3, pp. 90–126, Nov. 2006.

[8] W. I. Sellers and E. Hirasaki, "Markerless 3D motion capture for animal locomotion studies," Biol. Open, vol. 3, no. 7, pp. 656–668, Jul. 2014.

[9] A. Sinha, K. Chakravarty, and B. Bhowmick, "Person identification using skeleton information from kinect," presented at the Proc. Intl. Conf. on Advances in Computer-Human Interactions.

[10] J. Rosado, F. Silva, V. Santos, and Z. Lu, "Reproduction of human arm movements using Kinect-based motion capture data," 2013, pp. 885–890.

[11] J. Rosado, F. Silva, and V. Santos, "A Kinect-Based Motion Capture System for Robotic Gesture Imitation," in ROBOT2013: First Iberian Robotics Conference, vol. 252, M. A. Armada, A. Sanfeliu, and M. Ferre, Eds. Cham: Springer International Publishing, 2014, pp. 585–595.

[12] Q. Zhang, X. Song, X. Shao, R. Shibasaki, and H. Zhao, "Unsupervised skeleton extraction and motion capture from 3D deformable matching," Neurocomputing, vol. 100, pp. 170–182, Jan. 2013.

[13] G. Baird, "Diagnosis of autism," BMJ, vol. 327, no. 7413, pp. 488–493, Aug. 2003.

[14] E. L. Hill and U. Frith, "Understanding autism: insights from mind and brain," Philos. Trans. R. Soc. B Biol. Sci., vol. 358, no. 1430, pp. 281–289, Feb. 2003.

[15] Newschaffer, Craig J., and L. Kresch Curran, "Autism: an emerging public health problem.," Public Health Rep., vol. 118.5, pp. 393–399, Oct. 2003.

[16] W. Bosl, A. Tierney, H. Tager-Flusberg, and C. Nelson, "EEG complexity as a biomarker for autism spectrum disorder risk," BMC Med., vol. 9, no. 1, p. 18, 2011.

[17] Dongxin Xu, J. Gilkerson, J. Richards, U. Yapanel, and S. Gray, “Child vocalization composition as discriminant information for automatic autism detection,” 2009, pp. 2518–2522.

[18] S. H. R. E. Motlagh, H. Moradi, and H. Pouretemad, “Using general sound descriptors for early autism detection,” 2013, pp. 1–5.

[19] P. Perego, S. Forti, A. Crippa, A. Valli, and G. Reni, “Reach and throw movement analysis with support vector machines in early diagnosis of autism,” 2009, pp. 2555–2558.

[20] Y. Shen, et. al. , “Clinical Genetic Testing for Patients With Autism Spectrum Disorders,” PEDIATRICS, vol. 125, no. 4, pp. e727–e735, Apr. 2010.

[21] K. Wang, et.al. “Common genetic variants on 5p14.1 associate with autism spectrum disorders,” Nature, vol. 459, no. 7246, pp. 528–533, May 2009.

[22] J. A. S. Vorstman, W. G. Staal, E. van Daalen, H. van Engeland, P. F. R. Hochstenbach, and L. Franke, “Identification of novel autism candidate regions through analysis of reported cytogenetic abnormalities associated with autism,” Mol. Psychiatry, vol. 11, no. 1, pp. 18–28, Jan. 2006.

[23] N. Momeni, J. Bergquist, L. Brudin, F. Behnia, B. Sivberg, M. T. Joghataei, and B. L. Persson, “A novel blood-based biomarker for detection of autism spectrum disorders,” Transl. Psychiatry, vol. 2, no. 3, p. e91, Mar. 2012.

[24] R. Battiti and F. Masulli, “BFGS Optimization for Faster and Automated Supervised Learning,” in International Neural Network Conference, Dordrecht: Springer Netherlands, 1990, pp. 757–760.

[25] L. Breiman, “Random Forests,” Mach. Learn., vol. 45, no. 1, pp. 5–32, Oct. 2001.

[26] P. Jacquelin, S. Thorofare, and J. R. Davids, “Gait Analysis: Normal and Pathological Function.,” J. Pediatr. Orthop., vol. 12, no. 6, p. 815, Dec. 1992.

[27] A. Ball, D. Rye, F. Ramos, and M. Velonaki, “Unsupervised clustering of people from ‘skeleton’ data,” 2012, p. 225.

[28] V. L. Chester and M. Calhoun, “Gait Symmetry in Children with Autism,” Autism Res. Treat., vol. 2012, pp. 1–5, 2012.