

# Human sex recognition based on dimensionality and uncertainty of gait motion capture data

Adam Świtoński<sup>[0000–0003–4836–6418]</sup> and Henryk Josiński<sup>[0000–0003–0196–7777]</sup>

Silesian University of Technology, Department of Computer Graphics, Vision, and Digital Systems, Akademicka 16, 44-100 Gliwice Poland  
adam.switonski@polsl.pl, henryk.josinski@polsl.pl

**Abstract.** The paper proposes a method of human sex recognition using individual gait features extracted by measures describing the dimensionality and uncertainty of non-linear dynamical systems. The correlation dimension and sample entropy are computed for time series representing angles of skeletal body joints as well as whole-body orientation and translation. Two aggregation strategies for pose parameters are used – averaging of Euler angles triplets and taking an angle of 3D rotation. In the baseline variant, the distinction between females and males is performed by thresholding the obtained measure values. Moreover, the supervised classification is carried out for the complex gait descriptors characterizing the movements of all bone segments. In the validation experiments, highly precise motion capture measurements containing data of 25 female and 30 male individuals are used. The obtained, at least promising, performance assessed by correct classification rate, the area under the receiver operating characteristic curve, and average precision, is higher than 89%, 96%, and 96%, respectively, and exceeds our expectations. Moreover, the classification accuracy based on a ranking of skeletal joints, as well as whole-body orientation and translation evaluating sex-discriminative traits incorporated in the movements of bone segments, is formed.

**Keywords:** motion capture, gait analysis, correlation dimension, sample entropy, human sex recognition

## 1 Introduction

Human sex/gender recognition means females/women and males/men are identified on the basis of their registered behaviors or appearances. It plays an essential role in numerous commercial and non-commercial applications, such as retail and marketing, security and surveillance systems, user personalization, healthcare, gaming industry, smart environments as well as social analyses. This is the reason for diverse and active research studies conducted on this challenge. There are proposed methods, among others, operating on audio data [16], face images [6], keystroke dynamics [27], eye movements [21], handwriting [8] as well as electroencephalogram (EEG) [28], electrocardiogram (ECG) [18], electromyographic

(EMG) [7], ground reaction forces [5], inertial (IMU) [17] and motion capture [22] measurements.

In this paper, gait-based sex recognition is taken into account. Thus, females and males are distinguished by the way they walk. In the case of markerless acquisition of human movements performed during gait, it allows to carry on recognition without consciousness of the identified person. There are plenty of approaches proposed for gait-based sex recognition. Primarily, they utilize feature extraction in interpretable and generic variants. In the first one, the meaning of determined features is known. For instance, in [19], eight straightforward discrete variables, such as the angle at touchdown, maximum and minimum peak angles during the stance phase, and the angle at toe-off, are used. In another proposal [12], average values and stride-to-stride standard deviations of stride, swing, stance, and double support time intervals are computed. As regards generic feature extraction, linear dimensionality reduction techniques are mostly chosen. Particularly quite common is Principal Component Analysis (PCA) [15,26,14] maximizing the variance of the resultant features. Moreover, wavelet transform [3] and convolutional neural networks [11] were applied to accomplish the task.

Encouraged by our previous work [22] in which the measures describing gait sequences of bone segment movements having statistically significant differences between populations of females and males are found, we decided to propose the gait-based recognition system and assess its performance. It computes correlation dimension or sample entropy for rotational angles as well as global orientation and translation of the human body. Then, they are used in the classification procedure in two variants. In the first one, the distinction between females and males is performed by separate thresholding the obtained measure values. Moreover, the supervised classification is carried out for the complex gait descriptors characterizing the movements of all bone segments as well as whole-body orientation and translation. Despite the fact that correlation dimension and sample entropy are well-known measures broadly applied in the biosignal and motion data analysis [2,24,10], they were used only in our initial work [22] for the problem of gait-based sex recognition.

## 2 Related work

In our previous study [22], we applied correlation dimension, approximate, and sample entropies for the purpose of motion data description. They assess the dimensionality and uncertainty of the processed signal. Primarily, they were computed for rotational data describing angles between adjacent bone segments of the human body. The crucial investigation was related to the comparative analysis of extracted feature values for motion capture sequences representing gait performed by females and males. Descriptive statistics – mean values  $\pm$  standard deviation and median  $\pm$  quarter deviations – were determined, as well as non-parametric estimation and statistical hypotheses verification for the noticed differences were carried on.

The correlation dimension obtains greater values by average for females if lower limb movements are taken and smaller if shoulders are analyzed. It is mostly consistent with both entropy measures for which lower limbs behave similarly, but in place of shoulders, head movements achieve greater values for males. It can be interpreted that the system controlling females' hips and knees is more complex, which results in more sophisticated movements, and it is analogous to shoulder and head body segments.

The summary of the work [22] is depicted in Tab. 1. It contains the  $p$ -values of the Mann-Whitney-Wilcoxon test with the null hypothesis that the cumulative distribution functions  $F(x)$  (females) and  $M(x)$  (males) are the same against the right-tailed and left-tailed alternative ones, meaning  $F(x) \geq M(x)$  and  $F(x) \leq M(x)$ , respectively. It confirms that the noticed differences are statistically significant – the movements of taken body segments during gait discriminate females and males. This the reason we decided to investigate the performance of gait-based sex recognition using such extracted features. As the approximate and sample entropies are corresponding measures with quite similar discrimination properties obtained, and the first one is considered to be biased statistics, only correlation dimension and sample entropy are taken into account in this study.

Joint/Segment	Tail	CD	AppEnt	SampEnt
LeftUpLeg	right	< 0,001	< 0,001	< 0,001
RightUpLeg	right	< 0,001	< 0,001	< 0,001
LeftFoot	right	< 0,001	< 0,001	< 0,001
RightFoot	right	0,007	< 0,001	< 0,001
LeftShoulder	left	< 0,001	0,964	0,998
RightShoulder	left	< 0,001	0,999	0,891
Head	left	0,311	< 0,001	0,006

Table 1: The  $p$ -values of the Mann-Whitney-Wilcoxon test for populations of females and males described by correlation dimension (CD) as well as approximate and sample entropies (AppEnt, SampEnt) calculated for the movements of selected joints [22].

### 3 Correlation dimension and sample entropy

The system controlling human locomotion can be modeled as a non-linear dynamical system. The movements performed result in data describing poses in consecutive time instants. They are, in fact, observations – the output of the dynamical model. Thus, in the first stage, the reconstruction of the phase space is carried on. For every time instant  $i$  a vector  $x_i^m = [x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau}]$  being the time-delayed measurements is formed. The delay  $\tau$  is determined by the first local minimum of the mutual information function, and embedding dimension  $m$  is obtained using the 'False Nearest Neighbors' approach.

The Grassberger-Procaccia algorithm [9] for estimating correlation dimension is used. It determines a function called correlation sum  $CS(r)$  – the fraction of pairs of points  $(x_i^w, x_j^w)$  in the phase space whose distances are smaller than  $r$ . Then, a logarithmic scale is applied in which the linear approximation of the  $CS$  function at the beginning of the range is carried on. Ultimately, the slope coefficient of this approximation stands to be the estimate of the correlation dimension.

The sample entropy [20] is based on all possible patterns containing  $w$  consecutive points of the processed time series. The number of similar patterns whose distance is smaller than the assumed radius  $r$  is calculated for two subsequent  $w$  and  $w + 1$  values. Finally, the ratio between them is calculated and transformed by the logarithm function. The similarity radius  $r$  is a percentage (20%) of the standard deviation for the entire time series, and  $w$  is assumed to be 2.

A more detailed description of the measures taken can be found in [22].

## 4 Dataset

Highly precise motion capture measurements were used for registration purposes. The acquisition took place in the Human Motion Laboratory (HML) of the Polish-Japanese Academy of Information Technology (PJAiT) (<http://www.pja.edu.pl>) and was performed by the gold standard Vicon system with spatial accuracy below 1mm and assisted by certified staff. The collected dataset consists of 884 gait sequences of 25 females self-identifying as a woman and 30 males self-identifying as a man.

The default Vicon Blade skeleton, containing 22 bone segments as visualized in Fig. 1, was applied. This means that pose space is described by 72 parameters – 22 3D rotations represented by Euler angles triplets as well as whole-body orientation and translation. The parameters specify the human body relative to the reference T-Pose (Fig. 1b).

The gait was performed alongside a straight, five-meter-long route (see Fig. 2) with two interpreted individually paces – the preferred natural and increased ones. The applied frequency of registration was 100Hz.

It is exactly the same dataset that was used in [22].

## 5 Experimental setup

In the preprocessing stage, the gait main cycle is extracted. It contains two adjacent steps performed by the left and right lower limbs. The extraction algorithm is based on tracking the extremes of distances between ankles as described in [23]. Then, the measures – correlation dimension and sample entropy – are computed for the time series representing every pose parameter and used in the recognition.

Data registration took place over a long period of several years across numerous Vicon updates. Moreover, the exact pose estimation depends on the calibration procedure involving the range of movements performed by participants.

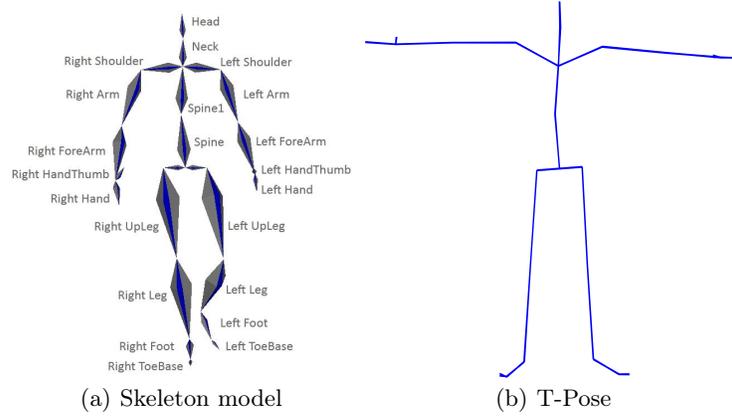


Fig. 1: Applied skeleton model and T-Pose.

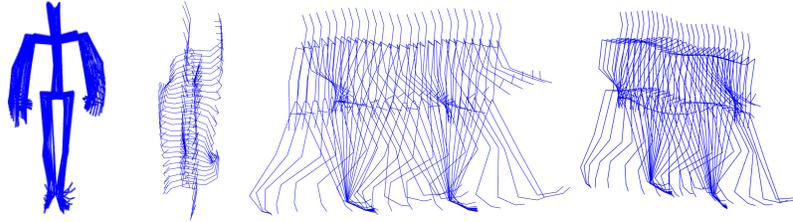


Fig. 2: Example gait instance – front, top, side, and perspective views.

It results in varying pose models with the same skeleton, but different meanings of the local coordinates systems and Euler angles triplets. It is visualized for the RightForeArm segment in Fig. 3 representing histograms of correlation dimension values. This is a segment with only one degree of freedom, but for most female recordings, it is incorporated within RX angle while to males – RZ. Thus, recognition taking every Euler angle separately utilizes acquisition-specific features instead of individual ones and causes its performance to be overevaluated.

This is the reason why two aggregation strategies are proposed. In the first one, as we expected, the identical relationships (greater or smaller) between populations of females and males for parameters describing the movements of the same joint or related to whole-body orientation and translation, the average, across these parameters, is computed and taken in further analysis. In the second proposal, before measure computation, rotations are transformed from Euler angle into axis-angle representation and the angle is chosen.

Such prepared gait descriptors are utilized in the classification conducted in two variants. In the first one, every segment, as well as whole-body orientation and translation, are classified separately by the thresholding – the females are identified in case of discovered relations greater or smaller. As the opti-

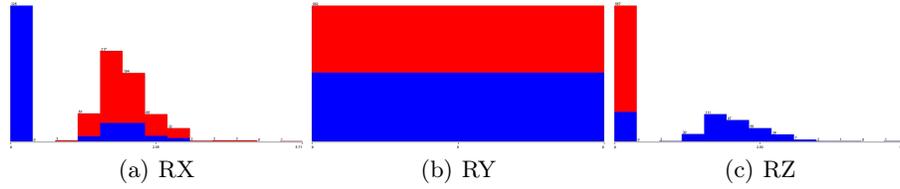


Fig. 3: Histograms of correlation dimension calculated for Euler angles of Right-ForeArm segment. The colors blue and red denote female and male samples, respectively.

mal threshold value is not obvious and difficult to predict, the performance is assessed by the area under the receiver operating curve (ROC-AUC) and the average precision (AP). It means true (TPR) and false (FPR) positive rates, as well as precision and recall, being relative numbers of actual females predicted as females or males are determined iteratively for successive thresholds and aggregated by the calculated area under obtained in such a way TPR-FPR and precision-recall curves.

In the complex variant, all aggregated measure values for bone segments, as well as whole-body orientation and translation, are classified by the selected supervised techniques. They are namely:

- Naive Bayes with parametric and non-parametric estimation [13],
- k nearest neighbors (kNN) with normalizable Euclidean distance metric [1],
- Random Forest [4] with 100 unconstrained depth trees,
- Multilayer perception (MLP) [29] with a single hidden layer containing a number of units equal to the number of input features.

For this variant, besides ROC-AUC and AP, the CCR ratio is calculated as well. It is the percent of correctly classified samples of the testing set and most intuitively corresponds to the recognition performance.

In the complex variant, a cross-validation algorithm was used to split the data into training and testing parts. Due to correlation dimension and sample entropy may be individual features as well, the custom division into the folds, instead of the default random procedure, was carried out. The folds contain data from single participants only. It means the training set does not have data of the same individuals as the testing one and sex is recognized on the basis of discovered properties for different persons. The number of folds is 55 – 25 females and 30 males.

In the case of the recognition based on single segments or whole-body orientation or translation (the first variant), there is no training stage – the aggregated/raw measure values directly point to female and male classes. Thus, only the testing set is used, and it contains all samples of the collected dataset.

## 6 Results

The results obtained for rotations of bone segments, as well as whole-body orientation and translation, are presented in Tab. 2. The recognition exceeding 80% of ROC-AUC and AP is pretty efficient, particularly taking in mind that it utilizes only a narrow subspace of pose parameters. It is clear that sample entropy is much more robust in the discrimination of females and males than correlation dimension. The top-ranked bone segments are related to the toe base, neck, spine, foot, and forearm movements. Moreover, the analysis of whole-body translation gives high accuracy of sex prediction. Both aggregation strategies perform similarly – averaging is a bit more efficient for sample entropy and angle of rotation for correlation dimension.

Segment/global	Average aggregation				Angle of rotation			
	CD		SE		CD		SE	
	RA	AP	RA	AP	RA	AP	RA	AP
Head	52.99	49.58	52.87	52.87	55.83	55.83	52.88	52.88
LeftArm	64.79	64.79	55.16	55.16	65.07	65.07	54.91	54.91
LeftFoot	57.11	57.11	73.29	76.75	57.70	57.70	71.99	75.93
LeftForeArm	53.73	53.73	53.66	56.08	53.73	53.73	53.66	56.08
LeftHand	60.07	60.07	61.91	61.91	61.98	61.98	60.04	60.04
LeftLeg	50.25	50.52	54.78	55.15	50.25	50.52	54.78	55.15
LeftShoulder	55.89	52.22	66.86	73.33	54.99	54.99	66.80	73.39
LeftToeBase	56.30	49.43	86.88	87.18	56.30	49.43	86.88	87.18
LeftUpLeg	52.51	51.56	61.80	63.57	51.59	50.71	58.53	61.00
Neck	53.20	49.84	77.36	77.36	54.88	54.88	78.66	78.66
RightArm	62.62	62.62	56.12	56.12	62.86	62.86	55.80	55.80
RightFoot	52.51	50.47	71.49	71.04	54.66	54.66	70.68	70.39
RightForeArm	53.12	53.12	48.88	51.70	53.12	53.12	48.88	51.70
RightHand	58.44	58.44	57.52	57.52	61.08	61.08	56.58	56.58
RightLeg	53.43	53.43	58.23	57.41	53.43	53.43	58.23	57.41
RightShoulder	51.70	49.81	56.85	64.39	55.60	55.60	55.95	63.19
RightToeBase	60.30	60.30	84.80	85.76	60.30	60.30	84.80	85.76
RightUpLeg	55.27	52.66	66.76	66.39	52.75	51.04	63.54	64.42
Spine	60.74	60.88	78.07	77.98	59.11	58.13	75.64	76.12
Spine1	59.16	59.34	59.36	62.76	58.27	58.57	52.62	57.44
Global rotation	58.64	58.64	66.29	66.29	58.84	58.84	64.65	65.33
Global translation	59.21	60.05	84.67	82.32	-	-	-	-

Table 2: Area under receiver operating characteristics curve [%] (RC) and average precision [%] (AP) in the distinction of females and males bone segments movements as well as whole-body orientation and translations during gait based on descriptors containing correlation dimension (CD) and sample entropy (SE) values for average across Euler angles and angle of rotation aggregation strategies.

The results for complete descriptors with measure values of all bone segments as well as whole-body orientation and translation are depicted in Tab. 3 and Tab. 4. In the case of the angle of rotation aggregation strategy, which is unworkable for non-rotational pose parameters, averaging is carried out for whole-body translation. The best performance is obtained by sample entropy features, Random Forest classifier and angle of rotation aggregation strategy. It is specified by 89.57%, 96.61%, 96.91% of CCR, ROC-AUC and AP measures, respectively. However, the 5NN classifier with the averaging aggregation strategy is only slightly worse and it has 88.78% CCR. The same general conclusion as previously stated can be drawn – sample entropy is more robust, and both aggregation strategies perform very similarly.

Classifier	Correlation dimension			Sample entropy		
	CCR	ROC-AUC	AP	CCR	ROC-AUC	AP
NaiveBayes normal	65.65	70.42	69.32	86.17	91.96	92.35
NaiveBayes kernel	65.65	70.32	69.91	86.73	92.64	92.51
1NN	57.71	64.48	58.21	85.49	83.36	74.43
3NN	58.50	63.59	60.04	85.71	91.17	85.93
5NN	57.14	64.66	62.21	88.78	93.34	90.28
RandomForest	65.42	72.87	75.47	89.23	96.38	96.66
MLP	57.82	60.67	62.21	86.28	93.90	95.05

Table 3: Performance [%] of females and males gait classification based on measure extracted for time series of every pose parameter and averaging aggregation strategy.

Classifier	Correlation dimension			Sample entropy		
	CCR	ROC-AUC	AP	CCR	ROC-AUC	AP
NaiveBayes normal	67.23	72.89	75.96	86.96	92.32	93.63
NaiveBayes kernel	68.25	73.17	75.43	87.19	92.92	92.61
1NN	61.34	65.07	58.62	86.17	83.67	74.44
3NN	62.59	66.50	62.51	84.47	91.29	86.12
5NN	62.81	70.18	68.52	87.64	93.10	89.64
RandomForest	67.01	72.56	75.43	89.57	96.61	96.91
MLP	62.36	65.27	68.45	83.22	91.74	93.75

Table 4: Performance [%] of females and males gait classification based on measure extracted for time series of every pose parameter and angle of rotation aggregation strategy.

In Tab. 5, confusion matrices with the number of true and false positives as well as true and false negatives are presented. The percentages of misrecognition of females predicted as males and males as females are very similar.

		Predicted				Predicted				Predicted	
		Female	Male			Female	Male			Female	Male
Actual	Female	401	45	Actual	Female	403	43	Actual	Female	390	56
	Male	47	389	Actual	Male	61	375	Actual	Male	57	379
			Random Forest, AR				5NN, AVG				Naive Bayes, kernel, AR

Table 5: Confusion matrices for females and males gait classification based on sample entropy extracted for time series of every pose parameter with averaging (AVG) and angle of rotation (AR) aggregation strategy.

A more detailed analysis of selected classifier workings is presented in Fig. 4, which contains dependencies between TPR and FPR as well as precision and recall measures for two selected cases. If the Random Forest classifier is taken, almost all females are recognized (TPR $\approx$ 98%), causing only about 30% false detection (FPR). Moreover, for 60% recall, there is a precision greater than 99%.

As the recognition performance achieved for toe base segments is surprisingly high – their range of movement seems to be insignificant during gait –, some acquisition-specific issues may have an influence on the results. This is the reason we decided to investigate the classification efficiency in the complex variant without taking into account just toe base movements. Additionally, we did the same with whole-body orientation and translation. The results are depicted in Tab. 6. The accuracy of sex detection is only a bit worse – the percentages of correctly classified gait samples (CCR) are 87.30%, 88.66%, and 85.15% in the case of discarding in the analysis toe base movements, whole-body parameters as well both of them, respectively.

## 7 Summary and conclusions

In the paper, the problem of female and male gait distinction is faced. It is based on the interpretable measures describing properties – uncertainty and dimensionality – of nonlinear dynamical systems. Sample entropy and correlation dimension are extracted for time series representing parameters of the human body skeleton model. Two aggregation strategies for the values related to the same segments and whole-body descriptors are proposed. The final recognition is carried out by the thresholding the aggregated measure values and by the selected supervised techniques.

The work is motivated by our results presented in [22] in which statistically significant differences in the correlation dimension and sample entropy values of

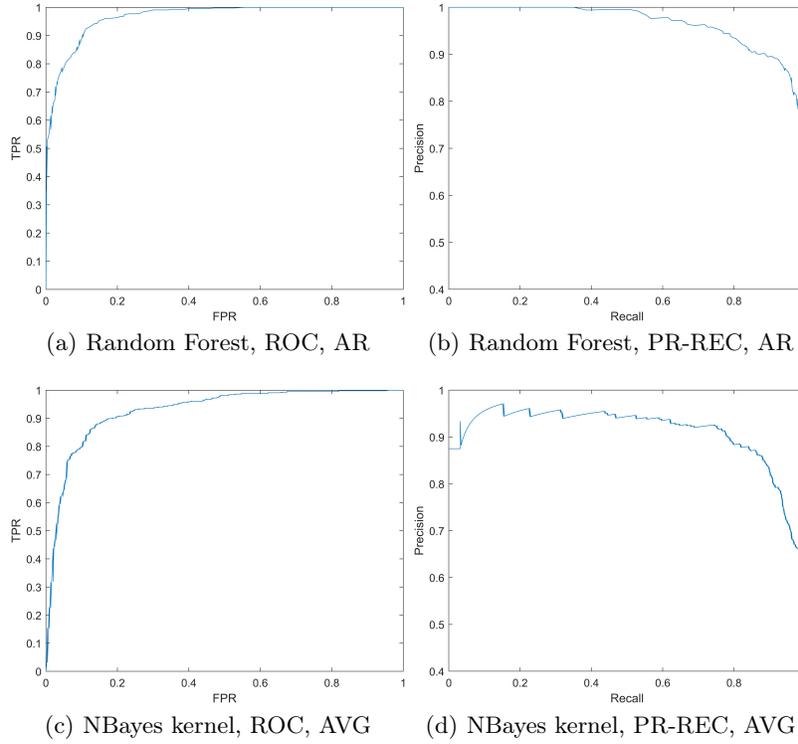


Fig. 4: Receiver operating characteristics and precision-recall curves for females and males gait distinction based on sample entropy calculated for time series of every pose parameter with averaging (AVG) and angles of rotations (AR) aggregation strategies.

Classifier	NTB		NWBT		NWB		NTBWB	
	CCR	ROC	CCR	ROC	CCR	ROC	CCR	ROC
NaiveBayes normal	80.39	88.36	86.28	91.31	85.94	91.37	78.00	86.40
NaiveBayes kernel	81.18	88.56	85.26	91.95	86.62	92.03	78.12	86.54
1NN	80.27	77.51	84.24	81.66	83.11	79.84	77.78	73.48
3NN	80.50	85.81	83.11	88.85	83.56	89.02	76.53	82.36
5NN	82.31	89.39	86.73	91.99	85.37	91.30	78.34	86.35
RandomForest	87.30	94.54	87.87	96.11	88.66	96.08	85.15	92.31
MLP	81.97	89.91	84.35	92.70	84.58	93.26	81.63	89.91

Table 6: Performance [%] of females and males gait classification based on angle of rotation aggregation strategy and sample entropy extracted for time series of pose parameters except those describing toe base movements (NTB), whole-body translation (NWBT), whole-body orientation and translation (NWB) as well as toe base movements, whole-body orientation and translation (NTBWB).

the time series describing movements of bone segments for the populations of females and males are discovered. However, this is a new research report, which is a natural continuation of the previous one. It goes further and assesses the performance of the recognition – statically significant differences do not mean the efficient classification is feasible and how it should be arranged. Moreover, it verifies the variant in which all Euler angles and translation data are involved.

We state the obtained results to be satisfactory. It is workable to predict females and males on the basis of their gait with almost 90% accuracy (CCR). Moreover, it is feasible to precisely control the balance between properly and improperly recognized females and males depending on expectations according to ROC and precision-recall curves. In addition, discriminative traits in the whole-body and successive bone segment movements are evaluated. Especially lower limb data are highly efficient in sex distinction. The extracted measure values are obviously robust in the faced recognition problem, but they very likely contain individual traits as well. It makes the classification problem to be more difficult and the obtained results even more valuable.

The main limitations of the work are similar to the ones mentioned in [22] and are related to short time series lengths. Moreover, despite the collected dataset being impressive as regards the number of recordings and participants, it has limited diversity. It mainly contains samples of young individuals coming from the same demographic region and which were taken in very similar conditions. An extended investigation with a more diverse dataset depends on recordings availability.

Furthermore, we expect the chaotic properties of the signal may give a valuable description of women’s and men’s gait discrimination. Due to insufficient time series lengths for computing the Lyapunov exponent, we are going to train a neural network using time courses of known dynamical models as presented in [25] and then apply it for feature extraction of mocap sequences.

**Acknowledgements** This publication was supported by the Department of Computer Graphics, Vision and Digital Systems, under the statutory research project (Rau6, 2024), Silesian University of Technology (Gliwice, Poland)

## References

1. Aha, D., Kibler, D.: Instance-based learning algorithms. *Machine Learning* **6**, 37–66 (1991)
2. Ahmadi, B., Amirfattahi, R., et al.: Comparison of correlation dimension and fractal dimension in estimating bis index. *Wireless Sensor Network* **2**(01), 67–73 (2010)
3. Arai, K., Asmara, R.: Human gait gender classification using 3d discrete wavelet transform feature extraction. *Int. J. Adv. Res. Artif. Intell* **3**(2) (2014)
4. Breiman, L.: Random forests. *Machine Learning* **45**(1), 5–32 (2001)
5. Chatain, C., Ramdani, S., Vallier, J.M., Gruet, M.: Recurrence quantification analysis of force signals to assess neuromuscular fatigue in men and women. *Biomedical Signal Processing and Control* **68**, 102593 (2021)

6. Dhomne, A., Kumar, R., Bhan, V.: Gender recognition through face using deep learning. *Procedia computer science* **132**, 2–10 (2018)
7. Di Nardo, F., Mengarelli, A., Maranesi, E., Burattini, L., Fioretti, S.: Gender differences in the myoelectric activity of lower limb muscles in young healthy subjects during walking. *Biomedical Signal Processing and Control* **19**, 14–22 (2015)
8. Gattal, A., Djeddi, C., Bensefia, A., Ennaji, A.: Handwriting based gender classification using cold and hinge features. In: *Image and Signal Processing: 9th International Conference, ICISP 2020, Marrakesh, Morocco, June 4–6, 2020, Proceedings* 9. pp. 233–242. Springer (2020)
9. Grassberger, P., Procaccia, I.: Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena* **9**(1-2), 189–208 (1983)
10. Harezlak, K., Kasproski, P.: Application of time-scale decomposition of entropy for eye movement analysis. *Entropy* **22**(2), 168 (2020)
11. Horst, F., Slijepcevic, D., Zeppelzauer, M., Raberger, A., Lapuschkin, S., Samek, W., Schöllhorn, W., Breiteneder, C., Horsak, B.: Explaining automated gender classification of human gait. *Gait & Posture* **81**, 159–160 (2020)
12. Hughes-Oliver, C., Srinivasan, D., Schmitt, D., Queen, R.: Gender and limb differences in temporal gait parameters and gait variability in ankle osteoarthritis. *Gait & Posture* **65**, 228–233 (2018)
13. John, G.H., Langley, P.: Estimating continuous distributions in bayesian classifiers. In: *Eleventh Conference on Uncertainty in Artificial Intelligence*. pp. 338–345. Morgan Kaufmann, San Mateo (1995)
14. Kastaniotis, D., Theodorakopoulos, I., Economou, G., Fotopoulos, S.: Gait-based gender recognition using pose information for real time applications. In: *2013 18th International Conference on Digital Signal Processing (DSP)*. pp. 1–6. IEEE (2013)
15. Kobayashi, Y., Hobara, H., Heldoorn, T.A., Kouchi, M., Mochimaru, M.: Age-independent and age-dependent sex differences in gait pattern determined by principal component analysis. *Gait & Posture* **46**, 11–17 (2016)
16. Kumari, M., Talukdar, N., Ali, I.: A new gender detection algorithm considering the non-stationarity of speech signal. In: *2016 2nd International Conference on Communication Control and Intelligent Systems (CCIS)*. pp. 141–146. IEEE (2016)
17. Mostafa, A., Barghash, T.O., Assaf, A.A.S., Gomaa, W.: Multi-sensor gait analysis for gender recognition. In: *ICINCO*. pp. 629–636 (2020)
18. Okin, P.M., Kligfield, P.: Gender-specific criteria and performance of the exercise electrocardiogram. *Circulation* **92**(5), 1209–1216 (1995)
19. Phinyomark, A., Osis, S.T., Hetingga, B.A., Kobsar, D., Ferber, R.: Gender differences in gait kinematics for patients with knee osteoarthritis. *BMC Musculoskeletal Disorders* **17**(1), 1–12 (2016)
20. Richman, J.S., Moorman, J.R.: Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology – Heart and Circulatory Physiology* **278**(6), H2039–H2049 (2000)
21. Sargezeh, B.A., Tavakoli, N., Daliri, M.R.: Gender-based eye movement differences in passive indoor picture viewing: An eye-tracking study. *Physiology & behavior* **206**, 43–50 (2019)
22. Świtoński, A., Josiński, H., Polański, A., Wojciechowski, K.: Correlation dimension and entropy in the assessment of sex differences based on human gait data. *Frontiers in Human Neuroscience* **17** (2023)
23. Świtoński, A., Josiński, H., Wojciechowski, K.: Dynamic time warping in classification and selection of motion capture data. *Multidimensional Systems and Signal Processing* **30**(3), 1437–1468 (2019)

24. Szczesna, A.: Quaternion entropy for analysis of gait data. *Entropy* **21**(1), 79 (2019)
25. Szczesna, A., Augustyn, D., Harezlak, K., Josinski, H., Switonski, A., Kasprowski, P.: Datasets for learning of unknown characteristics of dynamical systems. *Scientific Data* **10**(1), 79 (2023)
26. Troje, N.F.: Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision* **2**(5), 371–387 (2002)
27. Tsimperidis, I., Arampatzis, A., Karakos, A.: Keystroke dynamics features for gender recognition. *Digital Investigation* **24**, 4–10 (2018)
28. Wang, P., Hu, J.: A hybrid model for eeg-based gender recognition. *Cognitive neurodynamics* **13**, 541–554 (2019)
29. Witten, I.H., Frank, E., Hall, M.A.: *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 3rd edn. (2011)