ORIGINAL ARTICLES

Comparison of different marker sets for marker trajectory and principal component analysis based classification of simulated gait impairments

Josef Christian[*1], Felix Kluge[2], Björn M Eskofier[2], Hermann Schwameder[1]

[1]*Department of Sports Science and Kinesiology, University of Salzburg, Austria*
[2]*Department of Computer Science, Friedrich-Alexander University Erlangen-Nürnberg, Germany*

ABSTRACT

Objective: Many different marker sets have been used in marker trajectory based gait classification approaches. Little knowledge exists about the effects of specific marker sets on the subsequent statistical modeling. Such analysis is often based on principal component analysis. The aim of this study was to test the effect of marker set choice on marker trajectory and principal component analysis based gait classification.
Methods: This study tested the performance of principal component analysis based gait classification models with various marker sets on the basis of simulated gait impairments. Simulated gait impairments were used to enable a high level of control of the gait patterns.
Results: Classification accuracies were similar across most tested marker sets. Improved performance could be detected for some marker sets depending on the type of impairment.
Conclusion: Several potentially valid marker sets exist for a specific gait classification task even though trends could be found suggesting that optimal marker set choice is dependent on functional aspects of the movement.

Key Words: Marker set, Marker trajectory, Gait, Pattern classification, Machine learning, Principal component analysis

1. INTRODUCTION

Kinematics of walking patterns can be represented by 3D trajectories of point markers on specific body positions. Visualizations of such representations have been studied in visual perception of human gait such as gender classification[1] or recognition of individuals.[2] Troje[1] examined the application of multivariate statistics for automatic gait pattern recognition on 3D marker trajectory data. Direct application of statistics on the marker trajectory data without prior biomechanical modeling can save computational costs and is independent of mechanical model assumptions. This has been proposed as a holistic approach, resembling human perception in some respects.[3] Several studies in the gait analysis domain have applied variations of this approach and various marker sets have been used across these studies.[1, 3–12] The marker sets can be differentiated based on two major aspects. Firstly, whether the marker set covers the full body or body parts; full body marker sets were used most frequently.[1, 3–6, 11] Body part marker sets include unilateral placement on the lower extremity,[7, 10, 12] unilateral

placement on the upper and lower extremity[9] and bilateral placement on the lower body and thorax.[8] Secondly, marker sets can be differentiated based on whether the markers are placed exclusively on specific anatomical landmarks. Marker placement exclusively on anatomical landmarks was used for most associated studies and these marker sets were based on Vicon's Plug-In-Gait® model[3, 8, 11] or on customized models.[6, 7, 9, 10, 12] Federolf, Boyer[4] included marker clusters in addition to the markers on anatomical landmarks.[13] As a result of the variety of marker sets in use the number of markers ranges from 12 to 36 in the mentioned studies. Despite this, little effort has been invested in examining the effects of specific marker set choice (number and placement of the markers) on the subsequent data analysis. For clinical applications of automatic gait classification, it is important to achieve accurate classification since incorrect diagnoses can elevate costs for both, patients and clinicians. For example, accurate diagnosis is needed to decide for an adequate treatment. Therefore, it is important to use appropriate methods to solve specific gait classification problems. In marker trajectory based gait classification the choice of the marker set could be a crucial issue with uncertain implications, and exactly how the marker set choice influences the subsequent data analysis is of specific interest. In a typical pattern recognition approach, this analysis consists of feature extraction and classification,[14] therefore, the problem must investigate specific feature extraction and classification methods. The majority of the above mentioned studies applied Principal Component Analysis (PCA) for feature extraction and dimensionality reduction. Recently, self-organizing maps have also been used[7] but PCA has been favored mainly because it is particularly powerful on high redundancy data such as kinematic gait data. For classification methods, the linear Support Vector Machine (SVM) algorithm has been used in several approaches.[7–11] Other state of the art classifiers such as Ada Boost,[15] Naïve Bayes[14] or Random Forest[16] are capable of solving non-linear classification problems and are commonly considered in machine learning applications.

In each application of marker trajectory and PCA based gait classification, the question arises of which marker set should be chosen to achieve optimal performance. Therefore, the aim of this study was to investigate the impact of various marker sets on the performance of marker trajectory and PCA based gait classification models. The marker sets examined, differed in number of markers and their distribution over the body. To facilitate conclusions on the relationship between the effects of specific marker sets and the functional aspects of the gait patterns, the study was applied to artificially induced impaired gait patterns. This type of simulation allowed a high level of control while ensuring clear understanding of the functional aspects of the gait patterns.

2. METHODS

2.1 Participants

Thirty healthy male participants were recruited for this study (24±3 yrs; 180±5 cm; 77±7 kg; all values mean ± standard deviation). All participants had no gait impairments, full range of motion in their knee joints and any functional length discrepancy was smaller than 1.5 cm. The study was approved by the institutional ethics board and written informed consent was obtained from all participants.

2.2 Experimental protocol and measurements

Gait analysis was performed while participants walked on a laboratory treadmill at a constant walking speed of 5 km/h. Gait impairments were simulated by artificially induced restriction of knee extension or by artificially induced leg length discrepancy. Knee extension was restricted to 170° and 140° extension by means of an orthosis. Leg length discrepancy was induced by wearing differently soled shoes, with one shoe having a 2 cm or 4 cm thicker sole than the other shoe. The orthosis and the thicker soled shoes were applied on the right leg of all participants. Restricted knee extension (RKE) and leg length discrepancy (LLD) were treated as separate independent conditions. Thus, when the orthosis was applied, normal shoes were worn. When the modified shoes were worn, no orthosis was applied. Each participant walked under six different conditions: walking with unrestricted orthosis (RKE-0); walking with the orthosis restricted to 170° (RKE-1) and 140° (RKE-2) knee extension; walking with equally soled shoes (LLD-0) and walking with unequally soled shoes with 2 cm (LLD-1) and 4 cm (LLD-2) difference. The 3D trajectories of 46 reflective markers on the body were captured at 250 Hz by an eight camera optoelectronic device (Vicon, Oxford Metrics Ltd., Oxford, UK). Markers were placed according to Vicon's Plug-in-Gait® full body marker set. Additionally, marker clusters were placed on the shank and thigh.[4, 13]

2.3 Marker sets

The complete marker set consisting of all 46 markers contained markers on all body parts (see Figure 1). It contained markers on specific anatomical landmarks and markers placed arbitrarily on a specific segment; the former were denoted as anatomical markers and the latter as movement markers.
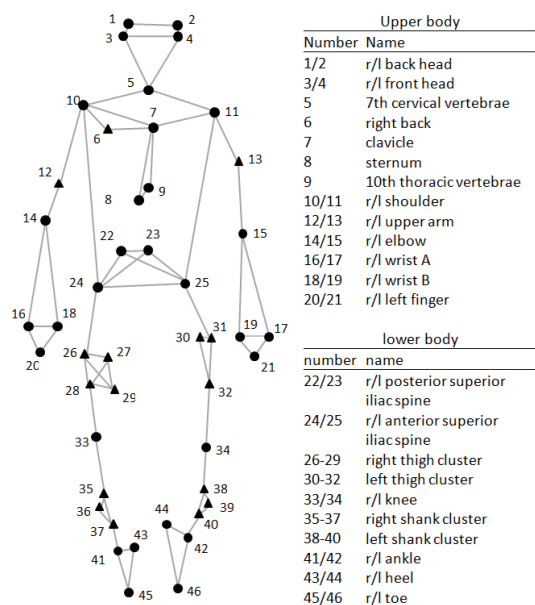
Upper body

Number	Name
1/2	r/l back head
3/4	r/l front head
5	7th cervical vertebrae
6	right back
7	clavicle
8	sternum
9	10th thoracic vertebrae
10/11	r/l shoulder
12/13	r/l upper arm
14/15	r/l elbow
16/17	r/l wrist A
18/19	r/l wrist B
20/21	r/l left finger

lower body

number	name
22/23	r/l posterior superior iliac spine
24/25	r/l anterior superior iliac spine
26-29	right thigh cluster
30-32	left thigh cluster
33/34	r/l knee
35-37	right shank cluster
38-40	left shank cluster
41/42	r/l ankle
43/44	r/l heel
45/46	r/l toe

Figure 1. Marker placement. Anatomical markers are depicted as circles, movement markers as triangles

Six marker subsets were defined with respect to body parts and inclusion of movement markers. Marker sets with only anatomical markers were denoted as anatomical (A) marker sets. Marker sets including anatomical and movement markers were denoted as combined (C) marker sets. The marker sets are listed in Table 1 and illustrated in Figure 2.

Table 1. Six different marker sets were defined with respect to body parts and the inclusion of movement markers

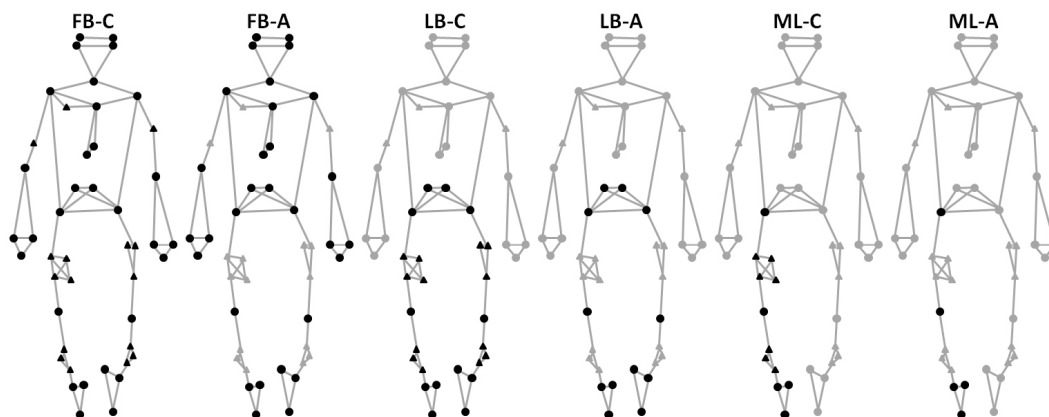| Body parts	Type	Name	N°
Full body	Combined	FB-C	46
Full body	Anatomical	FB-A	30
Lower body	Combined	LB-C	25
Lower body	Anatomical	LB-A	12
Modified leg	Combined	ML-C	12
Modified leg	Anatomical	ML-A	5

Figure 2. Six different marker sets. Included markers are marked black, excluded markers are marked gray

2.4 Data analysis

2.4.1 *Preprocessing*

The data analysis was implemented in Matlab® version R2015b using the statistics and machine learning toolbox (The MathWorks Inc., Natic, MA, USA). The raw marker coordinate data were low pass filtered at 6 Hz. For each trial ten individual consecutive stride cycles from one heel strike to the next heel strike of the left leg were extracted. To compensate for anatomical differences and different positions of the participants on the treadmill, the waveforms of each marker and spatial direction were centered on their mean value over the entire stride. Waveforms of each individual stride were time normalized to 101 samples per stride and averaged over the ten strides for each participant and condition. Subsequently, the data of each participant and condition

were arranged in a row vector to construct a spatio-temporal representation of the gait pattern in a vector space.[4, 8, 10–12] The waveforms of each marker and spatial direction were concatenated to form the gait pattern vectors with m × 303 dimensions (m markers × 3 spatial directions × 101 points in time). For further analysis, the gait pattern vectors (row vectors) were vertically concatenated to construct an input matrix. Thus, each column of an input matrix corresponded to a dimension and each row to a participant walking at a certain condition. One input matrix was constructed for each marker set and classification task. Four binary classification tasks were considered: RKE-0 vs. RKE-1, RKE-0 vs. RNE-2, LLD-0 vs. LLD-1 and LLD-0 vs. LLD-2. This resulted in a total of 24 input matrices (6 marker sets × 4 classification tasks). The subsequent statistical modelling

was conducted on each input matrix independently based on a typical pattern recognition scheme including feature extraction, classification and model evaluation.[14] For simplicity, the feature extraction and classification methods are described independently from the model evaluation. Thus, the feature extraction and classification sections describe how the model was built on a set of training data. The evaluation section describes how the generalization performance of the model was estimated.

2.4.2 *Feature extraction*

Before applying PCA, each input dimension was z-transformed to have zero mean and unit variance. The principal components were ranked in descending order according to the fraction of explained data variance. The principal components that contain together 95% of the variance were selected as features for classification. The remaining components were supposed to contain mainly noise and were therefore discarded.[5]

2.4.3 *Classification*

Classification was conducted using four supervised machine learning algorithms: Linear SVM, Naïve Bayes, AdaBoost (AdaBoost.M1) and Random Forest. The parameters of the algorithms were tuned by means of a leave-one-subject out cross validation (more details below) within a range of preset values. For the SVM, the penalization parameter C was tuned ($C = 10^x$; $x = 4, 3, \ldots, 4$). For AdaBoost and Random Forest, the number n of weak learners was tuned (n = 10, 20, 40, 80, 160, 320). The weak learners for AdaBoost were chosen to be decision stumps.[17] All other parameters were set to their default values.

2.4.4 *Evaluation*

A leave-one-subject out cross validation was used for model evaluation,[17] and the input matrix was vertically split into folds, where each fold contained only data of a single participant (two data items per participant; one from each class). Each fold was held out iteratively for testing while the remaining folds were used for training the model. Parameters for the z-transformation and the principal components were computed exclusively on the training data and later used to transform the test data. The data were arranged and split in the same order for each model to enable pairwise statistics for model comparison. A confusion matrix was built for each classification task. Algorithmically, the model evaluation resulted in a nested cross validation design with an internal loop for parameter optimization and an external loop for performance estimation.

2.4.5 *Statistics*

The confusion matrices were used to compute the posterior distribution of the accuracy of each model.[18] The mean and the corresponding 95% credible interval of the posterior distribution of the accuracy were calculated. The one-sided probability to have a posterior accuracy greater than chance level (50%) was computed for each model. A paired samples Wald test was used to compare the performance of the models.[19] The tests were performed to compare each of the marker subsets with the marker set including all available markers (FB-C) and to compare the anatomical marker sets (FB-A, LB-A, ML-A) with the corresponding combined marker sets (FB-C, LB-C, ML-C). The tests were performed for each classifier independently. The significance level was set to $\alpha = 0.05$.

3. RESULTS

3.1 Comparisons with chance level classifier

The results of the evaluation are shown in Figure 3. For task RKE-0 vs. RKE-1 all classifiers performed approximately at chance level for all marker sets except ML-A. With the ML-A marker set all classifiers performed significantly above chance level (Naïve Bayes: $p < .01$, AdaBoost: $p = .01$, Random Forest: $p = .02$, SVM: $p = .04$) (see Figure 3A). On the other classification tasks, all classifiers performed significantly above chance level with each marker set ($p < .01$) (see Figure 3B-3D).

3.2 Comparisons with FB-C marker set

For the RKE tasks, all significant differences indicated an increased accuracy with a smaller marker set than with the FB-C marker set. At least one significant increase could be found for each classifier (see Figure 3A-3B). For the LLD tasks all significant differences indicated a decreased accuracy with a smaller marker set than with the FB-C marker set (see Figure 3C-3D). At least one significant decrease could be found for AdaBoost and SVM.

3.3 Comparisons of combined and anatomical marker sets

For the RKE tasks, two significant differences could be found that indicate an increased accuracy with the anatomical marker set compared with the corresponding combined marker set (see Figure 3A-3B). For the LLD tasks, two significant differences could be found, of which one indicated an increased accuracy and one a decreased accuracy with the anatomical marker set compared with the corresponding combined marker set (see Figure 3C-3D).
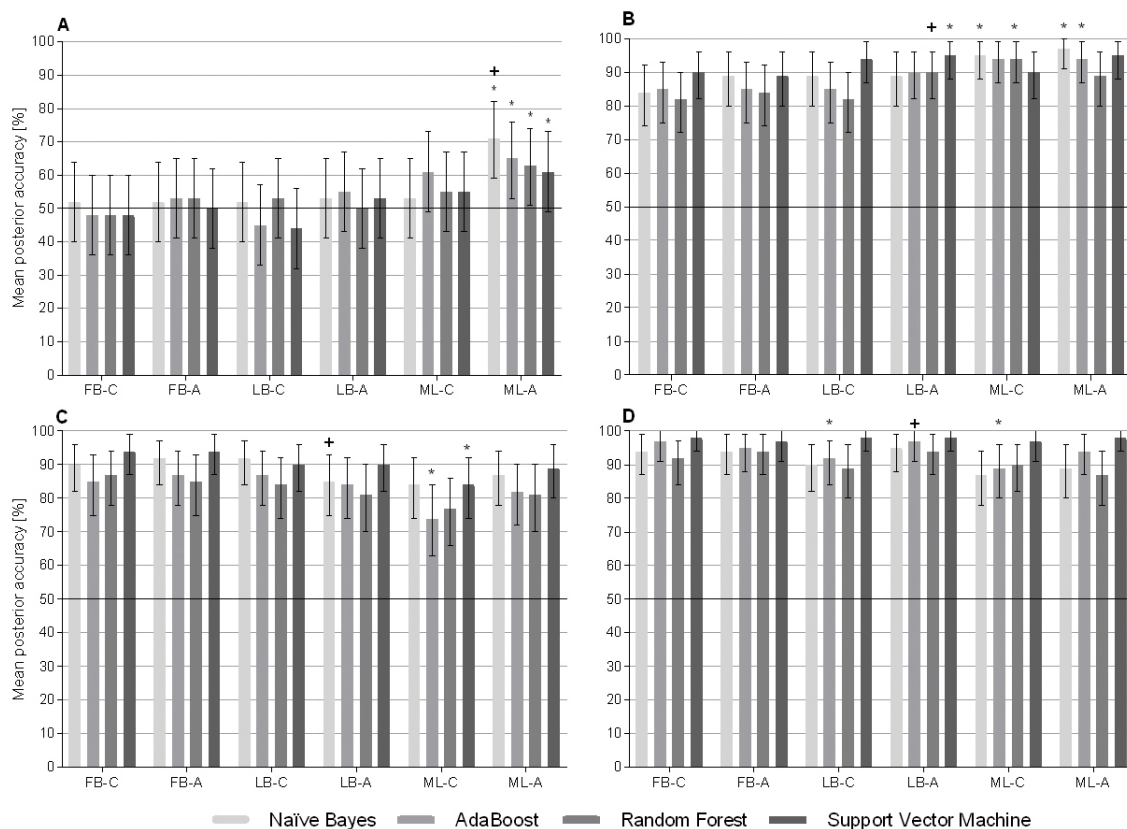
Figure 3. Mean posterior accuracies and 95% credible intervals for six marker sets and four classifiers. Each classification task is depicted in a separate chart. A: RKE-0 vs. RKE-1; B: RKE-0 vs. RKE-2; C: LLD-0 vs. LLD-1; D: LLD-0 vs. LLD-2.

* refers to a significant difference to the full body marker set ($p < .05$);
+refers to a significant difference to the corresponding combined marker set ($p < .05$)

3.4 Number of principal components

Table 2 contains the number of principal components used for classification. The full data set was used to compute these results.

Table 2. Number of principal components containing 95% of the data variance computed on the full data set

	FB-C	FB-A	LB-C	LB-A	ML-C	ML-A
RKE-0 vs. RKE-1	30	29	27	26	21	20
RKE-0 vs. RKE-2	31	30	27	26	20	19
LLD-0 vs. LLD-1	30	29	27	26	21	20
LLD-0 vs. LLD-2	30	29	27	25	21	20

Note. Rows: classification tasks. Columns: marker sets.

4. DISCUSSION

The results of the study show that for marker trajectory and PCA based gait classification, a variety of marker sets can yield similar accuracies on the same classification task independently of the used classifier. For an individual clas-

sification task and classifier, the accuracies were within a range of approximately 10% for most marker sets. This is a remarkable result as the number of markers ranged from five in the smallest marker set (ML-A) to 46 in the largest one (FB-C). This finding explains to some extent, the great variety of marker sets used in various studies since the choice of the marker set might not be a crucial aspect in the methods. This suggests that the information carried by some markers is highly redundant. The interdependency of the marker movements is caused by the internal and external constraints (e.g. mechanical, neurophysiological) that govern the human motor system in order to control its degrees of freedom.[4] For that reason, PCA is justified as it is specifically powerful reducing the dimensionality of highly correlated data. Therefore, PCA is capable of experimentally determining the degrees of freedom of gait patterns.[20]

Even though the classification accuracies were generally similar, several significant differences were detected and these indicated two major tendencies depending on the type of classification task. Firstly, the RKE tasks accuracies tended to

increase when using a smaller marker set. This is suggested by the observation that all significant differences indicated an increased accuracy with a smaller marker set than with the FB-C marker set. Furthermore, for the RKE-0 vs. RKE-1 task each classifier achieved a significantly higher classification rate than 50% only with the smallest (ML-A) marker set for each classifier. For the RKE-0 vs. RKE-2 task all classifiers achieved higher accuracies with the ML-A marker set, even though only significant for Naïve Bayes and AdaBoost. Thus, the same tendency can be observed in both RKE tasks but the accuracies for the RKE-0 vs. RKE-2 task were generally much higher. This is reasonable as the effect of the 140° restriction on the gait pattern is expected to be larger than of the 170° restriction (see Figure 3A-3B). Secondly, regarding the LLD tasks accuracies tended to decrease when using a smaller marker set. This is suggested by the observation that all significant differences indicated a decreased accuracy with a smaller marker set than with the FB-C marker set. Furthermore, in both LLD tasks all classifiers had generally decreased accuracies with the ML-C and the ML-A marker set compared with the FB-C marker set, even though not significant in all cases. To explain these two contrary tendencies, the different functional aspects of the RKE and LLD tasks can be considered. Restricting the knee extension has a large impact on a well localized area of the body, in particular the knee joint. It can be assumed that the whole body movement is affected by this type of restriction but that the dominant effects can be observed directly at and around the restricted joint. In this case, the movement of the markers on the restricted leg might be affected most. This means that these markers carry more discriminative information, resulting in a higher signal to noise ratio of the trajectory data compared to markers on other body parts. Excluding markers with lower discriminative power reduces the number of dimensions with low signal to noise ratio. The exclusion of noisy dimensions can have a significant impact on the performance of PCA.[21] This weakness of PCA may be overcome by using other feature extraction methods. Sources of noise are the natural inter- and intra-subject gait variability. The intra-subject gait variability was reduced by averaging the trajectories over ten stride cycles. By contrast, the effects of the LLD on the marker trajectories cannot be localized so clearly. It is not obvious which markers carry more discriminative information than others. The tendency of decreased accuracies with the smaller marker sets might be caused by excluding markers that contain valuable discriminative information. These two contrary effects are thus functionally reasonable. For practical applications, functional aspects could be considered for optimizing the choice of the marker set. In cases where large effects are expected in a well-defined area of the body and peripheral effects are not of particular interest, a specifically chosen body part marker set should be preferred over the full body marker set. Conversely, in cases, were the effects are allocated over the body or little knowledge exists about the localization, a full body marker set should be preferred. To yield more detailed recommendations on how to optimize the marker set choice in clinical applications further research on specific gait classification problems is required.

For the comparisons between combined and anatomical marker sets no clear statement can be made on which type is preferable. The few significant differences observed did not indicate a definite trend and from a practical perspective it is notable that the use of additional movement markers does not seem to be a crucial point when choosing the marker set. It is obvious that movement markers augment redundant information as they are placed on the same segments as anatomical markers. The additional non-redundant information consists of two different parts. Firstly, information about rotations around the longitudinal segment axis, if there are only two anatomical markers placed on the endpoints of the segment. Secondly, information about non-rigid segment movements mainly caused by skin displacement.[13] These sources of information are essential for adequately fitting a mechanical model to the data. For marker trajectory based gait classification the rotation aspect might play a secondary role since information about rotations around the longitudinal segment axis can also be mediated by markers on adjacent segments in some cases. For example, a rotation of the shank will necessarily result in a movement of the foot. In non-rigid segment movements, the movement markers have the potential of supporting the noise filtering effect of the PCA. Since the marker trajectory relationships on an ideally rigid segment would be linear, PCA is best suited to separate the highly correlated information from rigid segment movement and the random noise resulting from non-rigid segment movement.[5] On the other hand, movement markers can increase the interindividual variance of the marker trajectories because they are placed on approximate locations. This could make it more difficult to identify systematic group differences, although, these mechanisms do not have a significant impact in the current study, but may become more relevant in other gait classification tasks.

For all classification tasks the number of principal components used for classification was the smallest with the smallest marker set and increased with the number of markers in the set (see Table 2). This increase of about 10 principal components is relatively small compared to the increase of the input dimension from the smallest ($5 \times 303 = 1515$) to the largest ($46 \times 303 = 13938$) marker set. This result indicates the high correlation between the marker trajectories.

For future studies it could be of interest to investigate the effects of even larger or smaller markers sets.

While four classification algorithms were tested in this study it should be emphasized that it was not the purpose to compare the performance of the algorithms with each other for classification accuracy on the same task. Rather, the aim was to consider how each individual classifier reacted on the different marker sets. The results showed that all classifiers reacted in a similar way to the various marker sets. This indicates that the observed results were not dependent on the specific classification algorithm and allowed a more general interpretation.

5. CONCLUSIONS

The present study investigated the effects of various marker sets on the performance of marker trajectory and PCA based gait classification models. It was shown that various poten-tially valid marker sets exist for a certain gait classification task. This explains the variety of marker sets used in previous approaches, since the differences between the marker sets were rather small. However, notable tendencies could be observed. The well-controlled simulated impairments allowed linking of these tendencies with functional aspects of the gait patterns. If the effects of the impairment can be well-localized, it can be beneficial to select a smaller marker set which only includes the most informative markers. A larger or full body marker set can be beneficial, if the effects of the impairment cannot be clearly localized. These findings can be considered when arranging a marker set for future studies or clinical applications of marker trajectory and PCA based gait classification.

REFERENCES

[1] Troje NF. Decomposing biological motion: a framework for analysis and synthesis of human gait patterns. Journal of Vision. 2002; 2(5): 371-87. PMid:12678652. http://dx.doi.org/10.1167/2.5.2

[2] Cutting JE, Kozlowski LT. Recognizing friends by their walk: Gait perception without familiarity cues. Bulletin of the Psychonomic Society. 1977; 9(5): 353-6. http://dx.doi.org/10.3758/bf03337021

[3] Federolf P, Tecante K, Nigg B. A holistic approach to study the temporal variability in gait. Journal of Biomechanics. 2012; 45(7): 1127-32. PMid:22387120. http://dx.doi.org/10.1016/j.jbiomech.2012.02.008

[4] Federolf PA, Boyer KA, Andriacchi TP. Application of principal component analysis in clinical gait research: identification of systematic differences between healthy and medial knee-osteoarthritic gait. Journal of Biomechanics. 2013; 46(13): 2173-8. PMid:23910389. http://dx.doi.org/10.1016/j.jbiomech.2013.06.032

[5] Daffertshofer A, Lamoth CJ, Meijer OG, et al. PCA in studying co-ordination and variability: a tutorial. Clinical Biomechanics. 2004; 19(4): 415-28. PMid:15109763. http://dx.doi.org/10.1016/j.clinbiomech.2004.01.005

[6] Dillmann U, Holzhoffer C, Johann Y, et al. Principal Component Analysis of gait in Parkinson's disease: Relevance of gait velocity. Gait & Posture. 2014; 39(3): 882-7. PMid:24374062. http://dx.doi.org/10.1016/j.gaitpost.2013.11.021

[7] Hoerzer S, von Tscharner V, Jacob C, et al. Defining functional groups based on running kinematics using Self-Organizing Maps and Support Vector Machines. Journal of Biomechanics. 2015; 48(10): 2072-9. PMid:25869722. http://dx.doi.org/10.1016/j.jbiomech.2015.03.017

[8] Christian J, Kröll J, Strutzenberger G, et al. Computer aided analysis of gait patterns in patients with acute anterior cruciate ligament injury. Clinical Biomechanics. 2016; 33: 55-60. PMid:26945722. http://dx.doi.org/10.1016/j.clinbiomech.2016.02.008

[9] Maurer C, Federolf P, von Tscharner V, et al. Discrimination of gender-, speed-, and shoe-dependent movement patterns in runners using full-body kinematics. Gait & Posture. 2012; 36(1): 40-5. PMid:22304784. http://dx.doi.org/10.1016/j.gaitpost.2011.12.023

[10] von Tscharner V, Enders H, Maurer C. Subspace identification and classification of healthy human gait. PloS One. 2013; 8(7): e65063. PMid:23861736. http://dx.doi.org/10.1371/journal.pone.0065063

[11] Eskofier BM, Federolf P, Kugler PF, et al. Marker-based classification of young–elderly gait pattern differences via direct PCA feature extraction and SVMs. Computer Methods in Biomechanics and Biomedical Engineering. 2013; 16(4): 435-42. PMid:22149087. http://dx.doi.org/10.1080/10255842.2011.624515

[12] Nigg BM, Baltich J, Maurer C, et al. Shoe midsole hardness, sex and age effects on lower extremity kinematics during running. Journal of Biomechanics. 2012; 45(9): 1692-7. PMid:22507350. http://dx.doi.org/10.1016/j.jbiomech.2012.03.027

[13] Andriacchi TP, Alexander EJ, Toney MK, et al. A Point Cluster Method for In Vivo Motion Analysis: Applied to a Study of Knee Kinematics. Journal of Biomechanical Engineering. 1998; 120(6): 743-9. PMid:10412458. http://dx.doi.org/10.1115/1.2834888

[14] Theodoridis S, Koutroumbas K. Pattern Recognition: Elsevier Science; 2008.

[15] Freund Y, Schapire RE. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. Journal of Computer and System Sciences. 1997; 55(1): 119-39. http://dx.doi.org/10.1006/jcss.1997.1504

[16] Breiman L. Random Forests. Mach Learn. 2001; 45(1): 5-32. http://dx.doi.org/10.1023/a:1010933404324

[17] Eskofier BM, Kraus M, Worobets JT, et al. Pattern classification of kinematic and kinetic running data to distinguish gender, shod/barefoot and injury groups with feature ranking. Computer Methods in Biomechanics and Biomedical Engineering. 2012; 15(5):

467-74. PMid:21294006. `http://dx.doi.org/10.1080/10255` `842.2010.542153`

[18] Brodersen KH, Cheng Soon O, Stephan KE, Buhmann JM, editors. The Balanced Accuracy and Its Posterior Distribution. 20th International Conference on Pattern Recognition (ICPR); 2010.

[19] Brodersen KH, Deserno L, Schlagenhauf F, et al. Dissecting psychiatric spectrum disorders by generative embedding. NeuroImage Clinical. 2014; 4: 98-111. PMid:24363992. `http://dx.doi.org` `/10.1016/j.nicl.2013.11.002`

[20] Witte K, Ganter N, Baumgart C, et al. Applying a principal component analysis to movement coordination in sport. Mathematical and Computer Modelling of Dynamical Systems. 2010; 16(5): 477-88. `http://dx.doi.org/10.1080/13873954.2010.507079`

[21] Halouska S, Powers R. Negative impact of noise on the principal component analysis of NMR data. Journal of Magnetic Resonance. 2006; 178(1): 88-95. PMid:16198132. `http://dx.doi.org/10.` `1016/j.jmr.2005.08.016`