

Statistical Analysis of Gait Data to Assist Clinical Decision Making

Nigar Şen Köktaş¹ and Robert P.W. Duin²

¹Department of Mathematics and Computer Sciences,
Faculty of Arts and Sciences, Çankaya University,
06530 Balgat, Ankara, Turkey
nigarsen@cankaya.edu.tr

²Faculty of Electrical Engineering, Mathematics and Computer Sciences,
Delft University of Technology, The Netherlands
r.duin@ieee.org

Abstract. Gait analysis is used for non-automated and automated diagnosis of various neuromuskuloskeletal abnormalities. Automated systems are important in assisting physicians for diagnosis of various diseases. This study presents preliminary steps of designing a clinical decision support system for semi-automated diagnosis of knee illnesses by using temporal gait data. This study compares the gait of 111 patients with 110 age-matched normal subjects. Different feature reduction techniques, (FFT, averaging and PCA) are compared by the Mahalanobis Distance criterion and by performances of well known classifiers. The feature selection criteria used reveals that the gait measurements for different parts of the body such as knee or hip to be more effective for detection of the illnesses. Then, a set of classifiers is tested by a ten-fold cross validation approach on all datasets. It is observed that average based datasets performed better than FFT applied ones for almost all classifiers while PCA applied dataset performed better for linear classifiers. In general, nonlinear classifiers performed quite well (best error rate is about 0.035) and better than the linear ones.

Keywords: gait analysis; statistical pattern classifiers; clinical decision support systems.

1 Introduction

Gait Analysis is defined as the analysis of walking patterns of humans. A major application area is in the clinical decision-making and treatment processes for neuromuskuloskeletal diseases, among others such as security clearance systems and human identification [1]. Increasing number of Gait Laboratories in hospitals and respectively growing amounts of collected quantitative data necessitates the use of automated systems.

The use of pattern recognition techniques for analyzing gait data is studied before with varying success. Neural Networks [2-7] and Support Vector machines [7-9] are among the approaches used. Most of these works concentrates on a binary decision of the presence or non-presence of a disease. The number of subjects used for training is

generally quite low (About 10 or 15), with questionable reliability of the results. Used features are measured quantitative data such as joint angles, force plate data, cadence (distance between toes in one step) etc.

There are some significant difficulties in the design of automated systems because of the multidimensional and complex structure of the gait data. Feature reduction and selection always become an important part of the gait analysis studies. The most commonly used method for the feature extraction is based on the estimation of parameters (peak values, ranges) as descriptors of the gait patterns. In that case the classification is done according to the differences between the class averages of the training set and the parameters of the new subjects [10,11]. This method is subjective and neglects the temporal information of the gait data. There are examples of feature reduction techniques for gait data, such as Fast Fourier Transform (FFT) [3, 4, 11], Principal Component Analysis (PCA) [10-12], and wavelet transform [2]. Since all these studies differ from each other in description of the gait variables (such as subject type, measurement tools, type of variables, anatomical levels), and in construction of the classifiers, the results do not give a proper indication of the comparative performances.

Automated selection of the gait parameters is not observed in gait classification literature; medical experts select them or previous studies are taken as references. Actually, there are many medical practices, testing the variations in the gait parameters, which are caused by the related illness [12-14]. In [6] a gait classification study is performed to assist physician's decision making. Combination of MLPs is used for classification purpose, but no feature reduction or selection processes are applied. Therefore validity of the system depends on the expertise of the physician who selects features manually. Also, the judgments may vary in different experts leading the different interpretations of the classifiers. Obviously, automated selection lessens the dependence and the load on the experts and gives more freedom to the researchers.

The objective of the presented study is to propose a complete gait classification approach for automated diagnosis of the knee Osteoarthritis (OA) by comparing the convenient methods for preprocessing (feature reduction/selection) and classification. Most important contribution of the study is presenting a guide for further studies aiming classification of temporal and high dimensional medical data. It serves a base for design of clinical decision support systems by suggested automated feature reduction/selection methods and classification algorithms.

The following section discusses the details of gait analysis and data collection process. Section two gives more details about the used method and experiments, then discussion of the results are presented. Finally conclusion is presented.

1.1 Data Collection

Gait analysis enables the clinicians to differentiate gait deviations objectively. It serves not only as a measure of treatment outcome, but also as a useful tool in planning ongoing care of various neuromuskuloskeletal disorders such as cerebral palsy, stroke, OA, as an alternative or supportive to other approaches such as X-rays, MRI, chemical tests etc [13, 14]. Gait process is realized in a 'Gait Laboratory' by the use of computer-interfaced video cameras to measure the patient's walking motion. Electrodes called markers placed on the skin to follow the muscle activity by using the

infrared domain of the cameras. In addition, force platforms embedded in a walkway monitor the forces and torques produced between the patient and the ground.

In this study, data are collected in Ankara University Faculty of Medicine Department of Physical Medicine and Rehabilitation Gait Laboratory by the gait analysis experts, from 110 normal and 111 OA patients. In this laboratory, there are standard gait laboratory equipments, which are supported by “VICON” a commercial system for gait analysis. A subject is asked to walk on the platform and in one cycle of gait, temporal changes of joint angles, joint moments, joint powers, force ratios and time-distance parameters from four anatomical levels (pelvis, hip, knee and ankle) and three motion planes (sagittal, coronal and traversal) of the subjects are gathered. Resultant data (such as knee angle/time) is tabulated in graphic/numerical forms. Figure 1 shows examples to temporal gait parameters.

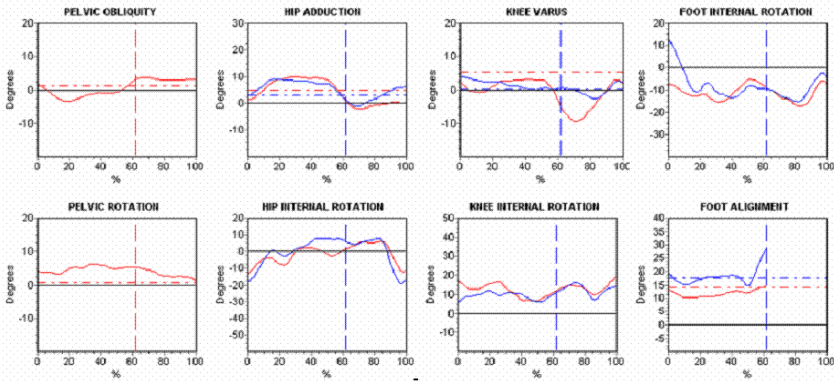


Fig. 1. Temporal gait parameters (data set B): The data is reported in 2-D charts with the abscissa defined as the percentage of the gait cycle and the ordinate displaying the gait parameter

Each of these parameters is represented by a graph that contains N samples taken in equally spaced intervals for one gait cycle. So the parameters for a given subject can be arranged as a M -dimensional vector X as below:

$$X = [X^{(1)}, X^{(2)}, \dots, X^{(M)}] \text{ where}$$

$$X^{(i)} = [X^{(i)}_1, X^{(i)}_2, \dots, X^{(i)}_N]$$

2 Experiments

The software in gait laboratory produces a dataset for each subject, including 33 gait parameters; each has 51 sample points in time. Therefore a $[33 \text{ (parameters)} \times 51 \text{ (time samples)}]$ dimensional array is created for each subject. Since the total number of the features is too large relative to the number of subjects, most of the commonly used classifiers will suffer from the curse of dimensionality [15, 16]. Different reduction techniques are applied for comparison. Averaging consecutive time samples creates six datasets having different dimensions. Moreover, FFT is applied to each waveform and one, five, ten or 25 FFT coefficients rather than 51 time samples represent each parameter. A PCA applied dataset is also added to this list as third reduction method. PCA is a

procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components (*PCs*). PCA is applied to feature matrix that is created by combining all gait parameters and a subset of principal components is selected for classification. Since first 65 *PCs* contain 95% of the total variance explained by all *PCs*, they are included in final dataset namely *dPCA*.

Following the reduction step, Mahalanobis distance (*MD*) is used to extract informative and discriminative features for effective classification. *MD* is a distance measure based on correlations between variables by which different patterns can be identified and analyzed. It is mostly used for determining similarity of an unknown sample set to a known one. In this study, the feature set creating the larger distance for two classes is selected as most discriminative feature set. Individual performances of the each gait parameters to discriminate two classes are compared. Table 1 shows the properties of these datasets created by combining the selected *best* features and the Mahalanobis Distance values produced by them.

Table 1. Datasets after feature reduction and selection processes

Composed dataset	# time samples	# selected gait parameters	Dimension of the dataset	<i>MD</i>
best51d	51	1	51	14.57
best25d	25	2	50	18.27
best10d	10	5	50	23.55
best5d	5	10	50	19.80
best2d	2	25	50	20.39
best1d	1	33	33	14.58
bestfft25	25	2	50	13.24
bestfft10	10	5	50	15.59
bestfft5	5	10	50	18.03
bestfft1	1	33	33	10.63
dPCA	-	-	65	31.93

Table 2. List of used classifiers

Classifier	Description [17]
Logistic Linear Classifier (loglc)	Computation of the linear classifier for the dataset by maximizing the likelihood criterion using the logistic (sigmoid) function.
Support vector classifier (svc)	Optimizes a support vector classifier for the dataset by quadratic programming. The default classifier is linear.
Linear Bayes Normal Classifier (ldc)	Computation of the linear classifier between the classes of the dataset by assuming normal densities with equal covariance matrices
Quadratic Bayes Normal Classifier (qdc)	Computation of the quadratic classifier between the classes of the dataset assuming normal densities.
Feed-forward neural net classifier (bpxnc)	A feed-forward neural network classifier with 1 hidden layer with 5 units in layer. Training is stopped after 100 epochs. Standard Matlab initialization is used. Training algorithm: "bpxnc" for back-propagation, "lmnc" for Levenberg-Marquardt
Feed-forward neural net classifier (lmnc)	
Automatic radial basis SVM (rbsvc)	A support vector classifier having a radial basis kernel
Parzen Classifier (parzenc)	Scaled datasets and log functions are used for better results. Computation of the optimum smoothing parameter for the Parzen classifier between the classes in the dataset.
Parzen density based classifier (parzencd)	For each of the classes in the dataset a Parzen density is estimated. For each class, feature normalization on variance is included in the procedure

In classification stage, a set of linear and nonlinear classifiers is tested by a ten-fold cross validation method. PRTools [17] is used for classifier construction. Total of nine classifiers are used by default parameters of the toolbox. Some adjustments are applied for better classification accuracy. Table 2 shows descriptions and adjustments of these classifiers, for more information see also [15-18].

Since density based classifiers (ldc, qdc, parzenc) suffer from a low numeric accuracy in the tails of the distributions, ‘log’ function is used to compute log- densities of them. This is needed for over trained or high dimensional classifiers. Almost zero-density estimates may otherwise arise for many test samples, resulting in a bad performance due to numerical problems. Classifiers are tested by cross-validation which is a methods for estimating generalization error based on *resampling*. Table 3 shows the classification error rates observed as a result of ten fold cross validation for all datasets presented in Table 1.

Table 3. Generalization error rates of classifiers

Datasets	Classifiers								
	loglc	svc	ldclog	qdclog	bpxnc	lmnc	rbscv	parzenlog	parzendc
best1d	0,083	0,080	0,066	0,086	0,071	0,073	0,047	0,098	0,093
best2d	0,097	0,088	0,062	0,110	0,069	0,070	0,035	0,070	0,062
best5d	0,088	0,083	0,050	0,121	0,063	0,068	0,044	0,055	0,076
best10d	0,084	0,086	0,046	0,130	0,056	0,070	0,085	0,086	0,077
best25d	0,115	0,079	0,068	0,158	0,065	0,070	0,116	0,089	0,103
best51d	0,207	0,138	0,091	0,206	0,111	0,155	0,143	0,142	0,152
best1ffd	0,110	0,098	0,082	0,104	0,095	0,110	0,074	0,096	0,096
best5fftd	0,152	0,124	0,083	0,113	0,100	0,117	0,102	0,076	0,095
best10fftd	0,127	0,119	0,082	0,132	0,100	0,125	0,095	0,112	0,164
best25fftd	0,148	0,098	0,110	0,132	0,115	0,143	0,103	0,129	0,106
dPCA	0,072	0,074	0,049	0,157	0,057	0,065	0,054	0,051	0,205

3 Discussion of Results

Comparing three reduction techniques using MD values, it is observed that averaged datasets perform better than the corresponding FFT based dataset. PCA applied dataset created largest distance value for classes, expectedly. Because, both PCA and MD methods are performed on covariance matrix of the input set. PCA is a well-known method to reduce feature dimensionality but not suitable for explicit feature selection like in current study. Because the contributions of the gait parameters to mapped feature set is not understandable. So it would be better to select features by some other methods and then to apply PCA. MD values are used to score the discriminatory ability of gait parameters, too. The number of appearance times of the parameters in datasets is listed below:

- KFlex (Knee Flexion): 7
- HMAbd (*Hip* Abduction Moment): 7

- KMFlex (Knee Flexion Moment): 5
- KRot (Knee Rotation): 4
- KMVal (Knee Valgus Moment): 4
- KPTot (Knee Power Total): 3

Four knee-related gait parameters are selected by an expert physician for classification in a previous study implemented by the same data [6]. These parameters have high scores in above list (KFlex, KMFlex, KMVal, KPTot). This supports the validation of automated feature selection process used in this study. Another mentionable result of feature selection process is that a hip related parameter has a score as high as knee related ones. The high discriminatory ability of this feature shows that knee OA causes high variation in hip abduction moments as much as in knee flexion moments of the patients.

Comparing the performances of the classifiers on the basis of the current number of subjects, it may be concluded that nonlinear classifiers performed quite well and better than the linear ones. We have also observed that high regularization prevents linear classifiers learning from more data. If we evaluate the success rates of the datasets by a standard baseline method, such as Linear Discriminant classifier (*ldclog* in Table 3), we see that PCA and averaging applied datasets performed better than the FFT applied ones. Actually among the linear classifiers *ldclog* performed as good as non-linear ones. Radial basis support vector classifier produced best generalization rate for almost all datasets.

4 Conclusion

The results of the experiments performed in this study are important for defining further studies about automatic diagnosis of gait disorders. Starting from the beginning of the study three different feature reduction techniques are compared first by the MD criterion and then by performances of the classifiers. It is observed that datasets created by FFT techniques produce worse results than the others. PCA applied dataset produces best result by almost all classifiers. Obviously, non-linear classifiers perform better than the linear ones with current number of subjects, but it can be suggested that considering the training costs of the algorithms, linear classifiers with a convenient regularization rate may be included in the further studies with *more* data.

It is also concluded that, temporal information of the waveforms is not so significant for the classifier performances. The severe difference between the performances of the datasets, *best1d* (all gait parameters with one time sample), and *best51d* (one gait parameter with all time samples) shows that including more gait parameter is more informative than including more time samples.

We detect a high match between currently selected features and the ones suggested by gait analysis expert. It can be concluded that our selection criterion approximates the expert knowledge and so contributes to the validation of the approach. The interpretation of the results of selection process reveals considerable information about effects of the illness. For example, expert physician concludes that motion of the hip angle is distorted as much as that of knee angle for sick subjects. Automatic feature selection may be preferred to find variation in all levels and motion planes of the subjects. But still, the feature selection process may be improved by using more

general criteria such as 1-Nearest Neighbor performance and also by using different search strategies than individual selection such as forward or backward selection.

These experiments showed us that statistical pattern recognition algorithms produce promising results for automated analysis of the gait data. The comparison of available statistical approaches for gait classification is expected to guide researchers working on this area for further studies.

Acknowledgements

This study is supported by The Scientific and Technological Research Council of Turkey (TUBITAK).

References

1. Simon, S.R.: Quantification of human motion: gait analysis—benefits and limitations to its application to clinical problems. *Journal of Biomechanics* 37, 1869–1880 (2004)
2. Chau, T.: A review of analytical techniques for gait data, Part 2: Neural Networks and Wavelet Methods. *Gait Posture* 13, 102–120 (2001)
3. Kohle, M., Merkl, D., Kastner, J.: Clinical gait analysis by neural networks: issues and experiences. In: *Proceedings of the 10th IEEE Symposium on Computer-Based Medical Systems*, p. 138 (1997)
4. Barton, J.G., Lees, A.: An application of neural networks for distinguishing gait patterns on the basis of hip-knee joint angle diagrams. *Gait & Posture* 5, 28–33 (1997)
5. Sen Koktas, N., Yalabik, N., Yavuzer, G.: An Intelligent Clinical Decision Support System for Analyzing Neuromuskuloskeletal Disorders. In: *International Workshop on Pattern Recognition in Information Systems*, pp. 29–37 (2008)
6. Sen Koktas, N., Yalabik, N., Yavuzer, G.: Ensemble Classifiers for Medical Diagnosis of Knee Osteoarthritis Using Gait Data. In: *Proceeding of IEEE International Conference on Machine Learning and Applications* (2006)
7. Begg, R., Kamruzzaman, J.: A Comparison of Neural Networks and Support Vector Machines for Recognizing Young-Old Gait Patterns. In: *Proceeding of IEEE TENCON Conference* (2003)
8. Begg, R., Palaniswami, M., Owen, B.: Support Vector Machines for Automated Gait Classification. *IEEE Transactions on Biomedical Engineering* 1, 52–65 (2005)
9. Salazar, A.J., De Castro, O.C., Bravo, R.J.: Novel approach for spastic hemiplegia classification through the use of support vector machines. In: *Proceedings of the 26th Annual International Conference of the Engineering in Medicine and Biology Society* (2004)
10. Dobson, F., Morris, M.E., Baker, R., Graham, H.K.: Gait classification in children with cerebral palsy: A systematic review. *Gait and Posture* 25, 140–152 (2007)
11. Chau, T.: A review of analytical techniques for gait data, Part 1: Fuzzy, statistical and fractal methods. *Gait Posture* 13, 49–66 (2001)
12. Deluzio, K.J., Astephen, J.L.: Biomechanical features of gait waveform data associated with knee Osteoarthritis: An application of principal component analysis. *Gait and Posture* 25, 86–93 (2007)
13. Gök, H., Ergin, S., Yavuzer, G.: Kinetic and kinematic characteristics of gait in patients with medial knee arthrosis. *Acta Orthop Scand* 2002 73(6), 647–652 (2002)

14. Kaufman, K., Hughes, C., Morrey, B., An, K.: Gait characteristics of patients with knee Osteoarthritis. *Journal of Biomechanics* 34, 907–915 (2001)
15. Jain, A.K., Duin, R.P.W., Mao, J.: Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(1), 4–37 (2000)
16. Duda, R.O., Hart, P.E., Stork, D.G.: *Pattern Classification*. John Wiley and Sons, New York (2001)
17. Duin, R.P.W.: PRTOOLS (version 4). A Matlab toolbox for pattern recognition. Pattern Recognition Group, Delft University of Technology (February 2004)
18. Kuncheva, L.I.: *Combining Pattern Classifiers: Methods and Algorithms*. Wiley-Interscience, Hoboken (2004)