304. The suboptimal set of parameters describing human gait

M. Derlatka[1], J. Pauk
Białystok Technical University, Mechanical Department,
15-351 Białystok, Wiejska 45c
Phone ++ 48 85 7469222
[1] **E-mail:** *mder@pb.bialystok.pl*

(Received 06 June 2007, accepted 18 June 2007)

Abstract. The influence of choosing a set of parameters describing human gait for automatic gait analysis and assessment has been presented in this paper. The investigations were based on three sets of parameters and two different classificators. The conclusion is that the best set of parameters is set of coefficients which have been obtained by modelling of human gait by means of identification using the regression function.

Keywords: human gait, system of decision support, classification.

Introduction

Human locomotion is very complex phenomena. The detailed gait analysis is time-consuming task and it needs to employ an expert with his deep knowledge both engineering and medicine. So, many methods of support decision making in biomechanics have been created recently [1, 2, 3, 9, 12, 13, 15, 17]. The main properties of this kind system should be:

- high quality – the decisions made by system must be equivalent to decisions made by expert;
- easiness in an expanding system by adding new types of gait pathology or adding methods of human gait apparatus improvement;
- module to explain of decision made.

The quality of mentioned system depends on two main factors:

- choosing a set of parameters describing human gait;
- choosing a method of mapping diagnostic parameters into gait pathology or into method of human gait improvement (classification).

The main aim of this paper is analysis of decision quality made by system based on different sets of human gait parameters.

Analysed material

The authors made measurements for 40 persons (25 men and 15 women) by means of optoelectronic systems Ellite 3D or Motion Analysis System appropriate in Bioengineering Centre in Milan, Italy or in Glenrose Rehabilitation Hospital in Edmonton, Canada during scientific stays of the authors. The S.A.F.L.O. protocol has been used in both centers [5]. Obtained data present both normal and pathological gait in saggital plane. The all objects walked barefoot at their natural cadence. They represented the following type of gait:

- normal - persons who have not reported problems with gait;
- Cerebral Palsy (CP) – Spastic Diplegia;
- Spina Bifida (SB) – Myelomingocele.

Table 1. Investigated person

Group	Number of persons	Age ± SD
Normal	10	25.7 ± 4.15
Cerebral Palsy	15	15.6 ± 5.2
Spina Bifida	15	16.2 ± 10.0

Methods

Three sets of parameters describing human gait for each group were used in this paper. All of them were obtained based on the same group of people and on the same values (an instantaneous power developed by muscles around three main joints of human lower limb: the hip, the knee, the ankle).

The parameters in the first set were obtained by modeling of dynamic of human gait by means of identification based on regression function (Eq. 1-3) [10, 11]. The coefficients of human gait model were evaluated for two main phases: the support phase and the swing phase for each joint separately. So, the results of identification dynamic model of human gait by means of regression function are 18 diagnostics parameters.

$$\underline{Y}_n = \underline{U}_n \cdot \underline{a}, \tag{1}$$

for n = 1, 2, …, N
where: $\underline{Yn}$ - the output matrix of idetification equation (an instantaneous power developed by muscles around hip, knee or ankle joints in nth instant); $\underline{a}$ - martix of searched factors; $\underline{U}_n$ -the input matrix of identification equation (an instantaneous power developed by muscles around hip, knee or ankle joints in k instants before); N – number of samples.

$$\underline{U}_n = [Y_{n-1} \quad Y_{n-2} \quad \dots \quad Y_{n-k}], \tag{2}$$

$$\underline{a} = [a_1 \quad a_2 \quad \dots \quad a_k]^T \tag{3}$$

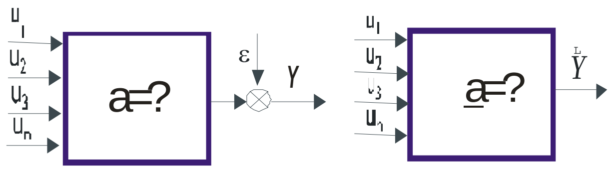

Fig. 1. Identification of human gait model

Parameters in the second set were based on so called indicies method and were created by [4, 7]:

- factor of human motion defined as:

$$k_v = 0.5 \cdot \frac{P_M}{E_K} = \frac{P_M}{m\bar{v}^2}, \tag{4}$$

where T – time of stride; $p_i(t)$ – an instantaneous power developed by muscles around i-th joint; i - indicate for the hip, the knee or the ankle joints of a human leg.

- power indicator (an average power) defined as:

$$P_M = \sum_{i=1}^{3} \sqrt{\frac{1}{T} \int_{t_1}^{t_2} p_i(t)^2 \, dt}, \tag{5}$$

- relative power at the hip joint:

$$H_w = \frac{P_B}{P_M} \cdot 100\%, \tag{6}$$

- relative power at the knee joint:

$$K_w = \frac{P_K}{P_M} \cdot 100\%, \tag{7}$$

- relative power at the ankle joint:

$$A_w = \frac{P_S}{P_M} \cdot 100\%, \tag{8}$$

where P_B, P_K, P_S – an average power in one stride developed by muscles around the hip, the knee and the ankle joints;

- an average velocity in saggital plane.

The last set of parameters was obtained by applying Kernel Principal Component Analysis transformation to second set of parameters. This time the polynomial kernel with d=2 in equation (9) has been used [8, 14, 16].

$$k(x, y) = (x \cdot y)^d \tag{9}$$

where x and y are input vectors. The dimensionality of space F[1] has been chosen as 15. The last five variables have been rejected because their values were equal or very close to zero.

All parameters in each set were divided into learning and testing set. In all cases the same strides were in the same set.

The two little sophisticated classifiers have been used to evaluate influence of different types of parameters set to decision making system quality. We have used a decision trees and feedforward neural networks. Decision made by Center of Bioengineering or Glenrose Rehabilitation Hospital staff was treated as a model and all results were compared with it.

Results

The architecture of feed forward neural networks was different for each set of parameters. The number of inputs always was equal to number of parameters. Classes (outputs) were coded based on '1 z n' code in following way:

- 1 0 0 - normal;
- 0 1 0 - spina bifida;
- 0 0 1 - cerebral palsy.

The result architectures were following:

- 18-6-3 for identification;
- 6-4-3 for indicies method;
- 15-6-3 for indicies method after KPCA.

The data has been standardized only for indicies method. In the next two cases all input values were the similar.

Neural networks were learned by well known RPROP algorithm. The networks indicate a tendency to overrating during learning, but taking into account the dimensionality of weight vector comparing to number of patterns in learning set it is rather obvious. It is important to say that any reduction input space should give better results in generalization by neural network.

[1] F denotes the result space, which is obtained after processing KPCA.

The main property of decision tree is recurrence dividing feature space in such a way to maximize in each step goal function. The goal function is calculated based on data from learning set. The CART algorithm has been used in this paper [6]. The decision tree obtained during learning process, which has been shown on figure 2. (If the condition in node tree is true, we are going with left branch, otherwise with right one).

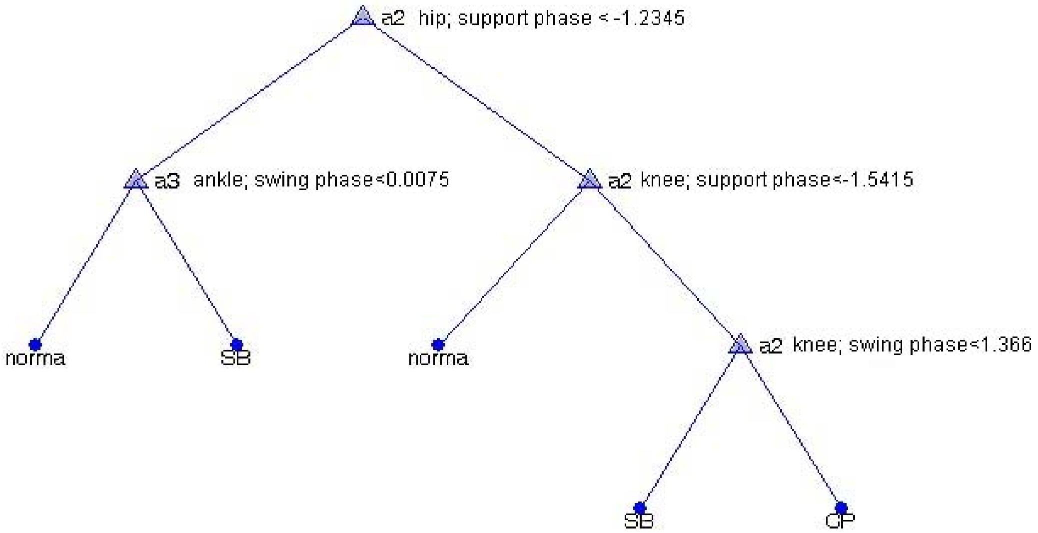

Fig. 2. Decision tree with obtained decision rules after learning based on data from identification

The obtained decision trees have only a few levels (3-4) and pruning tree gives no better results.

The percent of correct classifications made by neural network and decision trees depended on chosen set of diagnostic parameters are shown on the figures 3-5. It is important to say that all results were calculated on testing set.

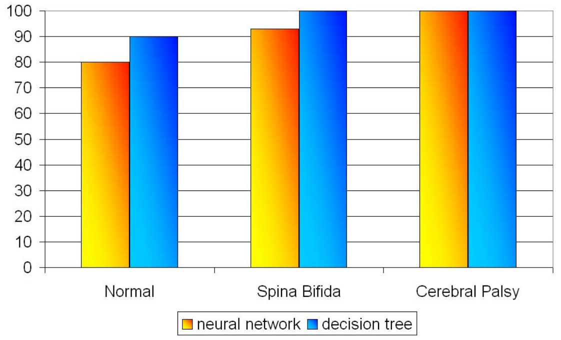

Fig. 3. The percent of correct classifications made by neural network and decision tree based on data from identification process

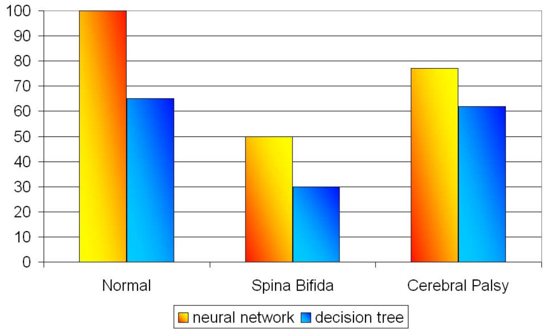

Fig. 4. The percent of correct classifications made by neural network and decision tree based on indicators defined by equations (4) – (8) and average velocity

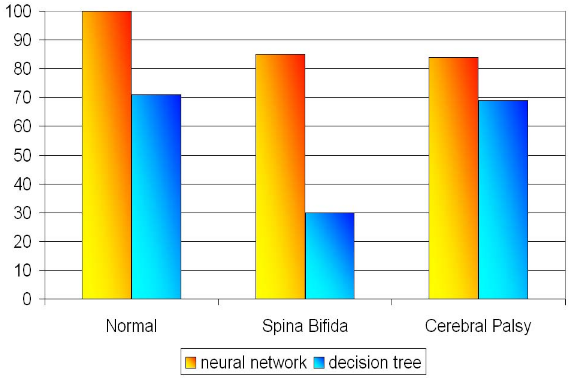

Fig. 5. The percent of correct classifications made by neural network and decision tree based on parameters obtained from KPCA.

Discussion

The best results were obtained by means of classificators based on data from identification method. The worst result (80% of correct classifications) was given by neural network for normal subjects. It is important to say that the best results were obtained by decision tree based on only 4 from 18 coefficients! One can noticed that those 4 coefficients describe all main joint of human legs, and a two of them are connected with the support phase and next two are connected with the swing phase.

The indicies method is the worst for automated gait analysis. The percent of correct classification is very insufficient. It is easy to notice that the results obtained by means of neural network are better than results obtained from decision tree in this case.

The result of making KPCA on indicies method parameters set has a bigger dimensionality of feature vector. The percent of correct classifications are much higher than in previous case. Unfortunately results obtained by means of decision tree are still unacceptable. The results improve especially in subjects with Spina Bifida (recognized by neural networks). The result improved from 50% to 85% correct classifications. Overall results given by neural networks based on KPCA parameters are comparable with results given by decision tree based on identifications. The difference in those two cases are rather small and adding another pathologies or adding more subjects into analyzed material could probably gives reverse results.

The authors are conscious of some limitations of above studies. The results are limited to chosen group of investigation persons. The same method could give a little different result but we are sure that still the same methods will give best results.

Conclusions

The best set of parameters is set of coefficients which have been obtained by identification with the regression function. It is suboptimal set of parameters describing human gait in the range of used data. Adding any new pathologies or other material (subjects, EMG charts) could change results.

KPCA allows achieving much better results even for the set of parameters with low diagnostics information.

References

[1] **Armand S., Watelain E., Mercier M., Lensel G., Lepoutre F.** Identification and classification of toe-walkers based on ankle kinematics using a data-mining method. Gait & Posture 23, 2006, pp. 240-248.

[2] **Chau T.:** A review of analytical techniques for gait data. Part 1: fuzzy, statistical and fractal methods. Gait and Posture 13, 2001, pp. 49-66.

[3] **Derlatka M.** Application of Kernel principal component analysis in human gait. Journal of Vibroengineering Vol. 7 nr 3, 2005, pp. 27-30.

[4] **Derlatka M.** Application of PCA in human gait assessment. Acta of Bioengineering and Biomechanics. Vol.6, Suppl.1, 2004, pp. 329-337.

[5] **Frigo C., Rabuffetti M., Kerrigan C. D., Deming L. C., Pedotti A.** Functionally oriented and clinically feasible quantitative gait analysis method. Medical & Biological Engineering & Computing, Vol.36 (2), 1998, pp. 179-185.

[6] **Hand D., Mannila H., Smith P.** Principles of data mining. Polish edition. WNT, Warsaw 2005

[7] **Jaworek K.** Indicies method of assessing human gait and run. (in Polish). IBIB PAN. No. 32. 1992 Warsaw.

[8] **Muller K. R., Mika S., Ratsch G., Tsuda K., Scholkopf B.** An Introduction to Kernel Based Algorithms. IEEE Transactions on Neural Networks Vol. 12, No. 2, 2001, pp. 181-202.

[9] **Olney S. J., Griffin M. P., McBride I. D.** Multivariate examination of data from gait analysis of persons with stroke. Phys Ther 1998;78(8):814–828.

[10] **Pauk J.** Identification of human locomotion dynamics. (Identyfikacja dynamiki lokomocji dwunożnej człowieka.) (in Polish). Ph. D. thesis. IBIB PAN, Warsaw 2005.

[11] **Pauk J., Jaworek K.:** Parametric identification of lower limbs during walking of a man. Design & Nature, WIT Press, Southampton, 2002, pp. 361-366

[12] **Romei M. Galli M., Motta F., Schwartz M., Crivellini M.** Use of the normalcy index for the evaluation of gait pathology. Gait and Posture 19, 2004 pp.85-90.

[13] **Sadeghi H.** Local or global asymmetry in gait of people without impairments. Gait and Posture 17, 2003 pp. 197-204.

[14] **Scholkopf B., Smola A., Muller K. R.** Nonlinear Component Analysis as a Kernel Eigenvalue Problem. Neural Computation, 10:1299-1319, 1998. Technical Report No. 44, 1996, Max Planck Institut für biologische Kybernetik, Tübingen.

[15] **Schutte L. M. Narayanan U., Stout J. L., Selber P., Gage J. R., Schwartz M. H.:** An index for quantifying deviations from normal gait. Gait and Posture 11, 2000 pp.25-31.

[16] **Smola A.:** Learning with Kernels. PhD thesis. Technische Universität Berlin, 1998.

[17] **Wolf S., Loose T., Schablowski M., Doderlein L., Rupp R., Gerner H. J., Bretthauer g., Mikut R.** Automated feature assesment in instrumented gait analysis. Gait & Posture 23, 2006, pp. 331-338.

Paper supported by grant no W/WM/1/06 (M. Derlatka) and the Polish Foundation for Polish Science (J. Pauk)