УДК 612.014.4

APPLICATION OF KPCA WITH DIFFERENT KERNELS FOR HUMAN GAIT ASSESSMENT

ПРИМЕНЕНИЕ ТЕОРИИ НЕЙРОННЫХ СЕТЕЙ С РАЗЛИЧНЫМИ МАТРИЧНЫМИ КОМПОНЕНТАМИ ДЛЯ ИССЛЕДОВАНИЯ ХОДЬБЫ ЧЕЛОВЕКА

Marcin Derlatka Ph. D.

Dept. of Automatics and Diagnostics

Bialystok Technical University

The evaluation of Kernel Principal Component Analysis (KPCA) based on different kernels in human gain assessment has been made. Three types of kernel have been chosen: linear, polynomial and RBF. The normalcy index was computed to evaluate usefulness kernels in human gait assessing. The test group consists of 45 persons (156 strides) with normal or pathological (Cerebral Palsy, Spina Bifida, Anterior Cruciate Ligament and Gonarthrosis) gait.

***Key words**: multivariate statistics, human gait, kernel principal component analysis.*

Рассматривается применение теории нейронных сетей для оценки походки человека. Использовался метод анализа главных матричных компонентов (KPCA), основанный на различных матрицах. Были использованы три типа матричных компонент: линейные, многочленные и радиальные базисные функции (RBF). Для оценки целесообразности применения матриц для анализа ходьбы был рассчитан стандартный индекс. Группу исследования составила выборка из 45 человек (156 шагов), включающая как людей без патологии, так и с патологической (церебральный паралич, Spina Bifida, повреждение передней крестообразной связки и гонартроз) походкой

***Ключевые слова:** многомерная статистика, походка человека, основной принцип анализов компонента*

1. Introduction

Gait is a basic human activity. It allows us to relocate our body. Gait is also a very complex human activity. It is described by enormous number of data which consist of:

- kinematic;
- kinetic;
- anthropometric;
- electromyographic;
- other parameters. The measurement of some of those parameters is necessary to perform quantitative human gait assessment. Assessment of human gait is a very important task, e.g. for quantifying the effects of rehabilitation process or surgical intervention.

It is very difficult to analyse and assess the human gait. The main problems are:

- correlation between parameters which describe human gait;
- high inter- and intrasubject variability;
- nonlinear dependency between some gait parameters, moreover very often some relationship is true only for one subject not for all human population;
- multidimensionality of phenomena.

Nowadays, methods of multivariate statistics like principal component analysis (PCA) are very popular [1, 2, 9, 10, 11, 13]. PCA is dedicated to solve important problems, and is a good tool for human gait assessment. This method creates a new coordinate in which all axes are uncorrelated [7]. Recently the nonlinear principal component analysis (nPCA) methods have become more and more popular in all the fields where nonlinear dependencies between data exist [12, 14].

2. Method

The normalcy index has been introduced and applied in a few recent works [2, 3, 10, 13]. It was defined as an Euclidean distance between origin of coordinate system and a point describing some values of selected gait parameters. The coordinate system was the result of applying PCA for set of data describing normal gait. The NI can be used in several ways:

- To show how subject's gait deviates from an average normal profile.
- To evaluate the level of pathology;
- To evaluate the result of walking improvement method;
- To compare a subject''s gait with another subject with similar (or the same) movement deficiencies.

The kernel PCA is performed in the following way [12]:

- Dot product matrix is computed as:

$$K_{ij} = \Phi(x_i)\cdot\Phi(x_j) \qquad (1)$$

where F is an arbitraly nonlinear mapping function, which map from observation space into possibly high-dimensional dot product space F.

- The eigenvectors V^k of matrix K is computed and normalized in dot product space F;
- Projection onto the eigenvectors is made by using the following formula:

$$\epsilon = V^k \cdot \Phi(x) = \sum_{i=1}^{M} \alpha_i^k \Phi(x_i)\cdot\Phi(x) \qquad (2)$$

where: ϵ - projection of x onto eigenvectors; a^k – coefficient which normalized k^{th} eigenvalue; M – number of observations.

In order to compute dot products matrix of the form $\Phi(x)\cdot\Phi(y)$, the kernel representation of the form:

$$k(x,y)=\Phi(x)\cdot\Phi(y) \tag{3}$$

has been used.

It allows to compute the value of dot product matrix without carrying out the mapping F [14]. Crucial to kernel PCA is that there is not actually performed mapping into F, but instead all nessesary computations are mede by the use of a kernel function from equation (3) in input space. KPCA in fact is doing a standard PCA in F space. The consequence is keeping by Kernel PCA most properties of standard PCA like:

- The principal components are uncorrelated;
- The first principal components carry more variance than the others;
- The mean-squared error of approximation in representing the observations by a few first principal components are minimal;

Opposite to PCA, Kernel PCA allows the extraction of a number of principal components, which can exceed the input dimensionality.

During the study the author has chosen three types of kernel:

- linear kernel (PCA);
- polynomial kernel defined as:

$$k(x,y)=(x\cdot y)^d \tag{4}$$

KPCA is reduced to standard PCA with d=1;

- the RBF (radial basis functions) kernel as:

$$k(x,y)=\exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \tag{5}$$

The normalcy index using KPCA for three types of kernel has been built.

The following parameters have been used to build the normalcy index [5]:

- factor of human motion defined as:

$$k_v=0.5\cdot\frac{P_M}{\bar{E}_K}=\frac{P_M}{mv^2} \tag{6}$$

where $\bar{E}_K$ - an average kinetic energy of the investigated person; P_M – so called power indicator (an average power) defined as:

$$P_M=\sum_{i=1}^{3}\frac{1}{T}\int_{t_1}^{t_2}p_i(t)^2\,dt \tag{7}$$

where $T=t_2-t_1$ – time of stride; $p_i(t)$ – an instantaneous power developed by muscles around i-th joint; i - indicate for hip, knee or ankle joints of a human leg.

- relative power at the hip joint:

$$H_w=\frac{P_B}{P_M}\cdot 100\% \tag{8}$$

- relative power at the knee joint:

$$K_w=\frac{P_K}{P_M}\cdot 100\% \tag{9}$$

- relative power at the ankle joint:

$$A_w=\frac{P_S}{P_M}\cdot 100\% \tag{10}$$

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

- an average velocity in saggital plane;
- time of stride;

It is easy to notice that almost all parameters are based on average power developed around one or all three joints of a human leg. Selection of power was caused by its properties. The power is a variable which combine kinetics and kinematics information. Moreover some earlier works show that parameters given by equations (6) – (10) carry important diagnostic information [4, 5].

The algorithm of normalcy index calculation is the following:

- All data are standardized to avoid influence of the difference units in parameters;
- The data have been mapped according to kernel;
- The new coordinates have been created based on 'normal' data only;
- The rest of data have been transformed to the new coordinates;
- The normalcy index is defined as a square of Euclidean distance from the origin of new coordinates.

3. Material and Results

More than 45 persons, men and women, took part in the investigation. 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;
- Anterior Cruciate Ligament (ACL);
- Gonarthrosis.

The data have been collected by means of Elite–3D system. The investigation has been made in the Center of Bioengineering in Milan, Italy using S. A. F. L. O. protocol [6]. All persons walked barefoot at their natural cadence. The gait of each person has been recorded several times. Over 150 stride cycles have been collected (Table 1). All the recorded data have been used for further investigations without any selection (e. g. based on statistics properties).

Tab. 1. Number of investigated persons and recorded strides

	Normal	Cerebral Palsy	Spina Bifida	ACL	Gonarthrosis
Number of persons	15	8	9	7	7
Number of strides	78	21	20	22	15

4. Discussion

The results obtained with PCA and KPCA with polynomial and RBF kernels are shown in the tables 2 - 4 respectively. Of course, the absolute values of the normalcy index are not important too much and depend on e.g. the number of variables in space F, type of used kernel. So, it cannot be used to compare results obtained with different algorithms (e. g. with different kernels). The value of normalcy index rises together with level of pathology in gait in tables 2 and 3.It doesn't have such dependence in case of RBF KPCA. The results from table 4 are useless in human gait assessment because they don't differentiate groups of walking. It is a little surprising, because radial basis function is usually a good tool in biomedical applications [15]. RBF KPCA won't be taken into consideration in further discussion.

Tab. 2. The results obtained with standard PCA

Group	mean NI	Range min-max	Percentiles $25^{th}-75^{th}$	Percentiles $10^{th}-90^{th}$
Normal	6.91	0.24 - 22.62	2.40 - 10.46	1.25 - 14.19
ACL	10.37	1.99 – 42.82	3.76 – 12.56	2.52 – 18.12
Gonarthrosis	32.01	15.55 – 65.76	18.58 – 40.30	16.67 – 52.61
Spina Bifida	100.70	39.01 – 230.71	55.79 - 122.45	43.23 – 201.77
Cerebral Palsy	273.64	64.30 - 1499.43	129.89 – 224.99	73.43 – 583.87

Tab. 3. The results obtained with KPCA with polynomial kernel, 24 variables, argument d= 2

Group	mean NI	Range min-max	Percentiles $25^{th}-75^{th}$	Percentiles $10^{th}-90^{th}$
Normal	1.52	0.51 – 3.98	0.83 – 1.99	0.71 - 2.82
ACL	6.45	1.35 – 31.25	2.32 – 8.69	1.74 – 11.27
Gonarthrosis	18.15	4.03 – 46.02	10.24 – 19.94	8.61 – 32.80
Spina Bifida	74.19	13.98 – 188.71	36.85 - 87.29	27.04 – 155.43
Cerebral Palsy	221.09	41.95 – 1218.88	101.21 – 185.03	56.45 – 504.41

Tab. 4. The results obtained with KPCA with RBF kernel, 72 variables

Group	mean NI	Range min-max	Percentiles $25^{th}-75^{th}$	Percentiles $10^{th}-90^{th}$
Normal	0.0812	0.0014 - 0.2226	0.0125 - 0.1442	0.0057 – 0.1813
ACL	0.1189	0.0232 – 0.3175	0.0360 – 0.1803	0.0318 - 0.2558
Gonarthrosis	0.0336	0.0314 - 0.0563	0.0319 - 0.0321	0.0316 – 0.0323
Spina Bifida	0.0321	0.0321 – 0.0321	0.0321 - 0.0321	0.0321 – 0.0321
Cerebral Palsy	0.0321	0.0321 – 0.0321	0.0321 – 0.0321	0.0321 – 0.0321

During analysis results one can easily notice that mean values (in tables 2 and 3) are significantly different in each of the groups (the appropriate statistical tests were made; $P<0.05$). The better results give us a polynomial kernel where the range of 25^{th}-75^{th} percentiles creates easily separated areas with a relatively big margin. Unfortunately, areas for groups of different kinds of walking are still overlapped in range of all data and 10^{th}-90^{th} percentiles in this case too. It is important that when we take into considerations all collected strides there are only few misclassification subjects (or decision is uncertain). Most of the improper decisions are the result of insufficient number of recorded strides of an investigated person. The much better quality in classification is main superiority polynomial nonlinear PCA on standard PCA (linear kernel) using normalcy index in human gait assessment.

The biggest problem is the distinguishing between ACL and normal data. The reason is a similarity in selected gait parameters describing those two groups and the quality of so called normal data. The most probably, there is no way to separate ACL and normal gait with proposed gait parameters and used set of data. Besides this case, the gait parameters, which have been used, give important diagnostics information.

5. Conclusion

- The evaluation of three kernels in human gait assessment has been described.
- It seems that the best is polynomial kernel.
- Author believes that there are possibilities in finding better set of diagnostics parameters using e.g. EMG signals.
- It is also important to preselect data set to obtain good representation of the selected kind of walking with high quality.
- The nonlinearity and increasing dimensionality of KPCA gives more possibilities especially in such complex phenomena like human locomotion.

References

1. Deluzio K. J., Wyss U. P., Costigan P. A., Sorbie Ch., Zee B. Gait assessment in unicompartmental knee arthroplasty patients: Principal component modeling of gait waveforms and clinical status// Human Movement Science 18 -1999 -№ 18 -P.701-711.
2. Derlatka M. Application of PCA in human gait assessment// Acta of Bioengineering and Biomechanics -2004 - Vol 6, Suppl.1.-P.329-337
3. Derlatka M. Dimensionality reduction in human gait assessing// Journal of Vibroengineering-2004 - Vol. 6, №1 - P. 19-21.
4. Derlatka M. Method of analysis and energetic assessment of human gait. (Metoda analizy i oceny energetycznej chodu cziowieka.) (in Polish) Ph. D. thesis. IBIB PAN. Warsaw, 2002.
5. Derlatka M. Neural networks in human gait assessment. Mokslinio-praktinio seminario BIOMDLORE'03(4). Vilnius 2004, pp. 9-15.
6. 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 - 1998.- Vol.36 P. 179-185.
7. Kwiatkowski W. Methods of automatics pattern recognition (Metody automatycznego rozpoznawania wzorcyw). Wojskowa Akademia Techniczna, Warsaw, 2001 -P. 142-150.
8. Muller K. R., Mika S., Ratsch G., Tsuda K., Scholkopf B. An Introduction to Kernel Based Algorithms // IEEE Transactions on Neural Networks - 2001 -Vol. 12, No. 2 - P. 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.- No78(8).- P.814–828.
10. Romei M. Galli M., Motta F., Schwartz M., Crivellini M. Use of the normalcy index for the evaluation of gait pathology// Gait and Posture.-2004.-No19.-P.85-90.
11. Sadeghi H. Local or global asymmetry in gait of people without impairments// Gait and Posture -2003 -No17 - P.197-204
12. Scholkopf B., Smola A., Muller K. R. Nonlinear Component Analysis as a Kernel Eigenvalue Problem// Neural Computation.-1998.-No 10.-P.1299-1319. Technical Report No. 44, 1996. Max Planck Institut fьr biologische Kybernetik, Tьbingen.
13. 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 -2000 -No 11.-P.25-31
14. Smola A.: Learning with Kernels. PhD thesis. Technische Universitдt Berlin, 1998.
15. Ibwiercz M. Application neural networks in modelling of selected biomedical systems. (Zastosowanie sieci neuronowych do modelowania wybranych systemyw biomedycznych).- Bialystok. 1997 - P.128

This work is supported by grant W/WM/3/03