Discriminating motion patterns of ACL reconstructed patients from healthy individuals

Worapan Kusakunniran, Nattaporn Dirakbussarakom,
Nantawat Prachasri, Duangkamol Yangchaem
Faculty of Information and Communication Technology
Mahidol University, Thailand
worapan.kun@mahidol.ac.th, nattaporn.dir@student.mahidol.ac.th,
nantawat.pra@student.mahidol.ac.th, duangkamol.yan@student.mahidol.ac.th

Jos Vanrenterghem, Mark Robinson
Research Institute for Sport and Exercise Sciences
Liverpool John Moores University, UK
j.vanrenterghem@ljmu.ac.uk, m.a.robinson@ljmu.ac.uk

Abstract

Injury to the Anterior Cruciate Ligament (ACL) can lead to inadequate movement during sport and daily life activities, leading to increased risk of re-injury or dropouts from any form of physical activity. Thus, it is important to detect such movement problems so that they can be prevented through focused rehabilitation programmes. This paper proposes a method to seek out differences of movement patterns between an ACL reconstructed group and a healthy control group. Principal Component Analysis (PCA) is applied to movement data in a training dataset. Then, Cohen's d is used to select such principle components (PCs) that can efficiently distinguish movement patterns of ACL reconstructed patients from healthy individuals. In our experiment, 10 subjects are used to evaluate the proposed method. Each subject contains nine observed variables of movement information. The proposed method can achieve a promising performance of above 90% accuracy to discriminating motion patterns of ACL reconstructed patients from healthy individuals. Also, vector loads of the selected PCs are plotted and visualized. Four variables significantly discriminated the ACL reconstructed group from the healthy control group, which are: 1) ground reaction force, 2) hip joint moment, 3) knee joint moment, and 3) ankle joint moment. Some of which have been identified as key predictors of ACL injury risk.

1 Introduction

Reconstruction of the Anterior Cruciate Ligament (ACL) after rupture has an impact on sporting activities, with few athletes returning to their previous level of play and many suffering further injury [1]. Understanding how the motion patterns of ACL reconstructed patients are different from healthy individuals may help explain these issues. This can be complicated since the motion patterns can be described using a large amount of variables. This paper is the first to apply machine leaning techniques to address such problem. However, there are some related research works [2][3] which focus on the relating problem domain of discriminating Osteoarthritis (OA) patients from healthy participants.

In this paper, to help reduce the complexity of the motion patterns, Principal Component Analysis (PCA) was applied. It reduces the data dimensionality, and more importantly, PCA can help seek out discriminant components from movement patterns in order to differentiate between the motion patterns of ACL reconstructed and healthy individuals. Then, such discriminant components are filtered over again by using Cohen's d [4]. Later, vector loads of the selected discriminant components are visualized to explain how motion patterns of the ACL reconstructed group are different from motion patterns of the healthy control group.

The motion patterns used in this paper were collected from 10 subjects (i.e. 5 subjects in the ACL reconstructed group and 5 subjects in the healthy control group) by using the 3D motion capture/camera system in the biomechanical laboratory, School of Sport and Exercise Science, Liverpool John Moores University. Each subject's movement was tracked by using reflective markers. The tracked positions of such markers were recorded by a 3D camera system and this combined with ground reaction forces were used to calculate 9 joint loading characteristics in a dynamic change of direction, including: 1) ground reaction force, 2) hip joint moment, 3) knee joint moment, 4) ankle joint moment, 5) hip angle, 6) knee angle, 7) ankle angle, 8) center of mass, and 9) center of mass velocity. The joint moment represents the torque acting around a joint at an instance time during a gait cycle [5]. They were used in our study. The aim was to see if PCA could discriminate between the two subject groups based on these variables.

The rest of this paper is organized as follows. The method proposed for discriminating motion patterns of ACL reconstructed patients from healthy individuals is explained in section 2. Experimental results are shown in section 3, and conclusions are drawn in section 4.

2 The Proposed Method

Figure 1 shows the framework of the proposed method. In the training process, given two sets of movement data $S_A = \{A_i\}_{i=1}^{N_A}$ and $S_B = \{B_i\}_{i=1}^{N_B}$ where A_i is the sample i in the healthy control group, B_i is the sample i in the ACL reconstructed group, and

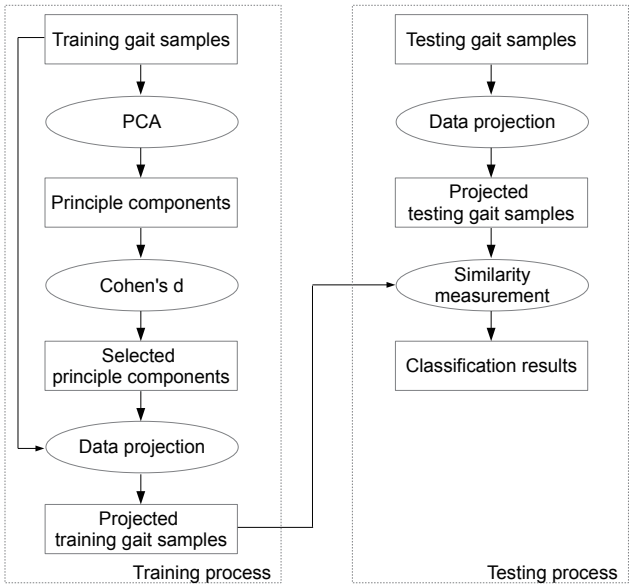

Figure 1. Framework of the proposed method.

N_A and N_B are total numbers of samples in S_1 and S_2 respectively, there are three main steps.

First, Principal Component Analysis (PCA) [6][7] is applied to acquire an optimized version of the data by reducing its dimension and finding its dominant variables. Given $A_i = \{a_{i,j}\}_{j=1}^{M}$ and $B_i = \{b_{i,j}\}_{j=1}^{M}$, where $a_{i,j}$ is the element j of A_i, $b_{i,j}$ is the element j of B_i, and M is the dimension of each movement data, the matrix U is constructed as below.

$$U = \begin{bmatrix} a_{1,1} & a_{1,2} & ... & a_{1,M-1} & a_{1,M} \\ a_{2,1} & a_{2,2} & ... & a_{2,M-1} & a_{2,M} \\ & & . & & \\ & & . & & \\ & & . & & \\ a_{N_A,1} & a_{N_A,2} & ... & a_{N_A,M-1} & a_{N_A,M} \\ b_{1,1} & b_{1,2} & ... & b_{1,M-1} & b_{1,M} \\ b_{2,1} & b_{2,2} & ... & b_{2,M-1} & b_{2,M} \\ & & . & & \\ & & . & & \\ & & . & & \\ b_{N_B,1} & b_{N_B,2} & ... & b_{N_B,M-1} & b_{N_B,M} \end{bmatrix}_{N \times M} \quad (1)$$

where $N = N_A + N_B$. Then, each element in U is standardized [8] in a column manner as:

$$\widetilde{U} = \{\frac{u_{i,j} - \mu_j}{\sigma_j}\}_{N \times M} \quad (2)$$

where $u_{i,j}$ is the element in U at the row i and column j, and μ_j and σ_j are the mean and standard deviation of elements in the column j of U, respectively.

$$\begin{aligned} \mu_j &= \frac{\sum_{i=1}^{N} u_{i,j}}{N} \\ \sigma_j &= \sqrt{\frac{\sum_{i=1}^{N} (u_{i,j} - \mu_j)^2}{N}} \end{aligned} \quad (3)$$

Then, the covariance matrix (C) of the training movement data is constructed as in the equation (4).

$$C = \frac{\widetilde{U}^T \times \widetilde{U}}{N} \quad (4)$$

where $\widetilde{U}^T$ is the transpose of the matrix $\widetilde{U}$. A process of eigen-decomposition is applied to the matrix C to generate eigenvectors and their corresponding eigenvalues. To reduce the data dimension, k eigenvectors corresponding to the k largest eigenvalues are selected. Usually, k is much less than M. In this paper, k is selected where the top k eigenvectors can cover about 99% of the variation of the whole training dataset, as shown in the equation (5).

$$k = \underset{1 \le k \le M}{\operatorname{argmin}} \frac{\sum_{i=1}^{k} |v_i|}{\sum_{i=1}^{M} |v_i|} > 0.99 \quad (5)$$

where v_i is the eigenvalue which corresponds to the eigenvector i. In this paper, the k selected eigenvectors are called Principal Components (PCs).

Second, Cohen's d [4][9] is applied to filter PCs by analyzing their discriminabilities i.e. Cohen's d values. PCs with high Cohen's d values are finally selected. The Cohen's d value (d_i) for the i^{th} PC (PC_i) is calculated as below.

$$d_i = \frac{\widetilde{\mu}_{A,i} - \widetilde{\mu}_{B,i}}{\sqrt{\frac{\widetilde{\sigma}_{A,i}^2 + \widetilde{\sigma}_{B,i}^2}{2}}} \quad (6)$$

where $\widetilde{\mu}_{A,i}$ and $\widetilde{\mu}_{B,i}$ are means of projected values by PC_i of training samples in the healthy control group and the ACL reconstructed group respectively, and $\widetilde{\sigma}_{A,i}$ and $\widetilde{\sigma}_{B,i}$ are standard deviations of projected values by PC_i of training samples in the healthy control group and the ACL reconstructed group respectively. These calculations are shown in the equation (7).

$$\begin{aligned} \widetilde{\mu}_{A,i} &= \frac{\sum_{j=1}^{N_A} p_j}{N_A} \\ \widetilde{\mu}_{B,i} &= \frac{\sum_{j=1}^{N_B} q_j}{N_B} \\ \widetilde{\sigma}_{A,i} &= \sqrt{\frac{\sum_{j=1}^{N_A} (p_j - \widetilde{\mu}_{A,i})^2}{N_A}} \\ \widetilde{\sigma}_{B,i} &= \sqrt{\frac{\sum_{j=1}^{N_B} (q_j - \widetilde{\mu}_{B,i})^2}{N_B}} \end{aligned} \quad (7)$$

where p_j and q_j are projected values by PC_i of training samples j in the healthy control group and the ACL reconstructed group respectively, as below.

$$\begin{aligned} p_j &= \widetilde{U}_j \times PC_i \\ q_j &= \widetilde{U}_{j+N_A} \times PC_i \end{aligned} \quad (8)$$

where $\widetilde{U}_j$ is the row j of $\widetilde{U}$ which corresponds to the standardized training sample j in the healthy control group, and $\widetilde{U}_{j+N_A}$ is the row $j + N_A$ of $\widetilde{U}$ which corresponds to the standardized training sample j in the ACL reconstructed group.

Third, assume that there are l PCs ($e = \{e_1, e_2, ..., e_l\}$) passing the Cohen's d filtering process,

Figure 2. Vector loads of the principal component for 9 joint loading characteristics in a dynamic change of direction. (a) Ground reaction force. (b) Hip joint moment. (c) Knee joint moment. (d) Ankle joint moment. (e) Hip angle. (f) Knee angle. (g) Ankle angle. (h) Center of mass. (i) Center of mass velocity.

the training samples are then projected into this feature space as below.

$$P = \widetilde{U} \times e^T \tag{9}$$

Thus, the representative features for the healthy control group (f_A) and the ACL reconstructed group (f_B) are calculated as below.

$$\begin{aligned} f_A &= \sum_{i=1}^{N_A} P_i \\ f_B &= \sum_{i=N_A+1}^{N} P_i \end{aligned} \tag{10}$$

In the testing process, given a sample of movement data $G = \{g_j\}_{j=1}^{M}$, it is first standardized based on the calculated mean and standard deviation from the equation (3).

$$\widetilde{G} = \{\frac{g_j - \mu_j}{\sigma_j}\}_{j=1}^{M} \tag{11}$$

Then, it is projected into the trained feature space, as below.

$$\widehat{G} = \widetilde{G} \times e^T \tag{12}$$

Finally, the class (c) of this testing sample is determined as in the following equation.

$$c = \underset{A,B}{\operatorname{argmin}}(s(f_A, \widehat{G}), s(f_B, \widehat{G})) \tag{13}$$

where c is either A (i.e. healthy individuals) or B (i.e. ACL reconstructed patients), and s is a chosen similarity measurement function e.g. Euclidean distance [10].

3 Experiment

The human motion patterns were collected from 10 different subjects by using the reflective markers through the 3D camera system and the ground reaction force on the force plate, where 5 subjects were in the ACL reconstructed group and the other 5 subjects were in the healthy control group. In our experiment, 5 trials were recorded for each subject which was asked to perform a side-cutting task. In each trial, the motion data consists of 9 variables (ground reaction force, hip joint moment, knee joint moment, ankle joint moment, hip angle, knee angle, ankle angle, center of mass, and center of mass velocity) in 101 normalized time points with 3 degrees of freedom (x, y and z directions). Thus, the original dimension of the data is $9 \times 101 \times 3 = 2727$.

This dataset is divided into 2 parts as: 1) 2 subjects in the ACL reconstructed group and 2 subjects in the healthy control group are used in the training process, and 2) the other 3 subjects in the ACL reconstructed group and the other 3 subjects in the healthy control group are used in the testing process.

In our training phase, 30 principal components were found to cover 99% of the variation of the whole training dataset. The maximum Cohen's d value was 1.25, corresponding to the 5^{th} principal component which represented about 7.96% of the variation, as shown in

Table 1. The top 30 principal components (PCs) sorted by their eigenvalues, with their corresponding Cohen's d values (column 2) and percentages representing the variation on the training dataset (column 3).

PCs	Cohen's d values	% of the variation
1	-0.90	18.61
2	0.42	15.23
3	0.41	13.94
4	0.31	11.60
5	1.25	7.96
6	-0.72	5.88
7	0.17	4.92
8	0.00	3.85
9	0.67	3.30
10	0.59	2.61
11	0.01	1.67
12	-0.15	1.28
13	0.18	1.18
14	0.04	0.99
15	0.10	0.80
16	-0.07	0.74
17	0.01	0.67
18	-0.12	0.58
19	-0.11	0.46
20	0.08	0.44
21	0.28	0.41
22	-0.08	0.35
23	0.33	0.30
24	-0.07	0.27
25	-0.05	0.23
26	-0.08	0.19
27	0.11	0.18
28	-0.26	0.16
29	-0.04	0.12
30	0.04	0.11

Table 1. The accuracy of the classification between the ACL reconstructed group and the healthy control group was about 90%.

In the column 3 of Table 1, a percentage representing the variation on the training dataset of PC i was calculated by dividing the absolute value of the eigenvalue of the PC i by the sum of the absolute values of all eigenvalues then multiplied by 100.

Then, vector loads of the 5^{th} PC for all 9 variables were visualized in Figure 2. They are plotted to explain which joint loading characteristics are different between the ACL reconstructed group and the healthy control group. The larger value of the vector load represents the higher impact on such motion patterns discrimination. The y-axis represents the vector loads of the principal component, while the x-axis represents the 101 normalized time points. Based on the vector loads of the 5^{th} PC, the dominant variables which discriminated between the ACL reconstructed group and the healthy control group were ground reaction force, hip joint moment, knee joint moment, and ankle joint moment. Among the three degrees of freedom (x, y, z), x-component was dominant in the ground reaction force, and z-component was dominant in the hip joint moment, the knee joint moment and the ankle joint moment, to discriminate between the two subject groups.

4 Conclusion

It is difficult to identify the differences in motion patterns between ACL reconstructed and healthy individuals from all data because of the large variation among different subjects in the same group. This study has shown a reduced number of variables to identify such differences under the PCA-projection space, allowing for a more focused evaluation of motion pattern differences. Further work will need to verify whether these reduced dimensions can be of particular importance to biomechanists and ultimately clinicians who try to help ACL reconstructed patients return to their sport more effectively. Increased number of subjects might give more power. The benefits that could be achieved are 1) the constructed PC-based model can be used to predict whether a newly given gait pattern (i.e. the 9 joint loading characteristics) belong to the ACL reconstructed group or the healthy control group; and 2) the carefully analysis of the vector loads can be used to indicate specific problems of the ACL reconstructed patients.

References

[1] J. H. Daruwalla, J. W. Xerogeanes, P. E. Greis, R. E. Hancock, L. D. Kaplan, G. F. Hatch, K. P. Spindler, D. L. Johnson, E. C. McCarty, and C. T. Moorman, "Rates and determinants of return to play after anterior cruciate ligament reconstruction in division 1 college football athletes: a study of the acc, sec, and pac-12," *Orthopaedic Journal of Sports Medicine*, vol. 2, no. 3, 2014.

[2] K. J. Deluzio and J. L. Astephen, "Biomechanical features of gait waveform data associated with knee osteoarthritis an application of principal component analysis," *Gait & Posture*, vol. 25, pp. 86–93, 2007.

[3] P. A. Federolf, K. A. Boyer, and T. P. Andriacchi, "Application of principal component analysis in clinical gait research: identification of systematic differences between healthy and medial knee-osteoarthritic gait," *Journal of Biomechanics*, vol. 46, pp. 2173–2178, 2013.

[4] J. Cohen, *Statistical power analysis for the behavioral sciences*. Routledge Academic, 2013.

[5] H. M. Clayton, E. Hodson, J. L. Lanovaz, and G. R. Colborne, "The hindlimb in walking horses: 2. net joint moments and joint powers," *Equine Veterinary Journal*, vol. 33, no. 1, pp. 44–48, 2001.

[6] L. I. Smith, "A tutorial on principal components analysis," *Cornell University, USA*, 2002.

[7] J. Shlens, "A tutorial on principal component analysis," *Google Research, USA*, 2014.

[8] A. A. Khan, D. D. Iliescu, R. J. Sneath, C. E. Hutchinson, and A. A. Shah, "Principal component and factor analysis to study variations in the aging lumbar spine," *IEEE Journal of Biomedical and Health Informatics*, 2014.

[9] K. Kelley and K. J. Preacher, "On effect size," *Psychological Methods*, vol. 17, no. 2, pp. 137–152, 2012.

[10] M. M. Deza and E. Deza, *Encyclopedia of distances*. Springer, 2009.