

A PPCA-BAYESIAN APPROACH FOR KNEE OSTEOARTHRITIS CLASSIFICATION USING GROUND REACTION FORCES

Margarita Kotti^{1,2}, Lynsey D. Duffell¹, Aldo A. Faisal^{2,3}, Alison H. McGregor¹

¹Musculoskeletal (MSK) Laboratory, Imperial College London; ²Brain Behaviour Laboratory, Department of Bioengineering, Department of Computing, Imperial College London; ³MRC Clinical Sciences Centre, Imperial College London

Introduction

Osteoarthritis (OA) is the most common form of arthritis in the UK. A challenge in analysing knee OA kinematics, a form of behaviour, is its variability [Faisal, 2008]. In this paper, a novel fully automated framework is proposed which classifies subjects as normal or suffering from knee OA using ground reaction force (GRF) measurements over the Z-axis (GRF-Z). In order to build statistical models of subjects belonging to the normal or OA category probabilistic principal component analysis was employed (PPCA) along with a Bayes classifier [Faisal, 2010]. 163 subjects were used to learn the PPCA parameters and the remaining 18 for testing by means of a Bayesian classifier. An accuracy of 78.95% is achieved on data of subjects taking 3 steps only.

Methods

One hundred eighty one subjects participated in the study. Forty six of them suffered from knee OA. All subjects walked at normal speed along walkways with 2 Kistler Type 9286 (Winterthur, Switzerland) force plates. 3 trials per subject were captured at 1000 Hz. Trials with no clean force plate strike were excluded. Data for complete gait cycles were time normalized by linear interpolation.

PPCA [Tipping, 1999] is appropriate to extract a small number of artificial variables (principal components - PC) that account for most of the variance in the observed variables. Specifically, PPCA is employed on the behavioural data [Belić, 2011] in order to extract the mean vector and the covariance matrix for each category (normal vs OA sufferer) separately. The experimental protocol is subject-independent, i.e. if a subject is included in the training set, then he/she is not used in the test set.

Results

As a pre-processing step, the GRF-Z data were sub-sampled by a factor of 10 to avoid model starvation and min-max normalization was

applied. Fig. 2 shows the percentage of variance of the GRF-Z data explained versus the number of PCs used. The first PC explains about 57% of variance in the data, whereas over 97% of the variance is explained by 11 PCs or more. The accuracy utilizing 11 components equals 78.95%.

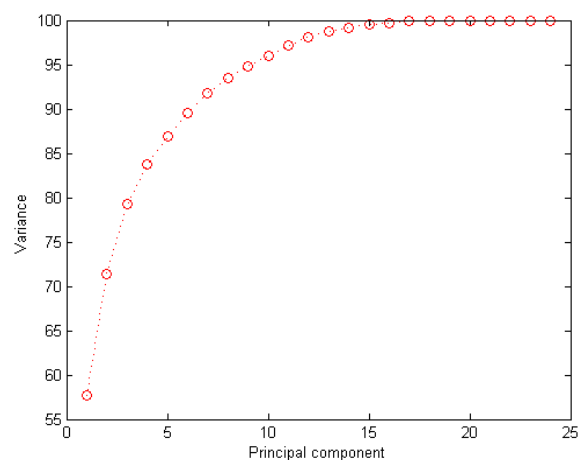


Figure 2: The number of principal components able to explain the variance of the GRF-Z.

Discussion

Our bioinformatics of behaviour approach yields an effective assessment of whether the subject suffers from knee OA. The PPCA approach reveals that the variability shown by the ground reaction forces of subjects talking only a few steps is able to classify knee OA.

Acknowledgement

The authors acknowledge support from the Medical Engineering Solutions in Osteoarthritis Centre of Excellence funded by the Wellcome Trust and the EPSRC.

References

- Belić & Faisal, Neuroscience 12 (Suppl 1): P102, 2011.
- Faisal et al, Nature Reviews Neuroscience 9:292-303, 2008.
- Faisal et al, PLoS ONE 5 (11): e13718, 2010.
- Tipping & Bishop, J Royal Statistical Society 61: 611-622, 1999.