Journal of Biomechanics 38 (2005) 401–408

JOURNAL OF BIOMECHANICS

www.elsevier.com/locate/jbiomech
www.JBiomech.com

A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data

R. Begg[a,*], J. Kamruzzaman[b]

[a] Biomechanics Unit, Centre for Rehabilitation, Exercise & Sport Science, City Flinders Campus, Victoria University, P.O. Box 14428, Melbourne City MC, Vic., 8001, Australia
[b] Gippsland School of Computing & IT, Monash University, Churchill Campus, Vic., 3842, Australia

Accepted 5 May 2004

Abstract

This paper investigated application of a machine learning approach (Support vector machine, SVM) for the automatic recognition of gait changes due to ageing using three types of gait measures: basic temporal/spatial, kinetic and kinematic. The gaits of 12 young and 12 elderly participants were recorded and analysed using a synchronized PEAK motion analysis system and a force platform during normal walking. Altogether, 24 gait features describing the three types of gait characteristics were extracted for developing gait recognition models and later testing of generalization performance. Test results indicated an overall accuracy of 91.7% by the SVM in its capacity to distinguish the two gait patterns. The classification ability of the SVM was found to be unaffected across six kernel functions (linear, polynomial, radial basis, exponential radial basis, multi-layer perceptron and spline). Gait recognition rate improved when features were selected from different gait data type. A feature selection algorithm demonstrated that as little as three gait features, one selected from each data type, could effectively distinguish the age groups with 100% accuracy. These results demonstrate considerable potential in applying SVMs in gait classification for many applications.

Keywords: Gait; Support vector machine; Gait classification; Elderly

1. Introduction

It is well established that ageing influences gait patterns and considerable research has documented changes during unobstructed and obstructed walking that suggest age-related declines in lower limb control (Princea et al., 1997; Begg and Sparrow, 2000). The major aim has been to identify key variables of gait degeneration in elderly individuals that might be predictors of falling behaviour. Research has shown that significant changes in gait can occur with age in temporal and distance measures such as gait velocity, stride length, and stance and swing phase times (Hageman and Blanke, 1986; Winter, 1991). In addition, foot-ground reaction force data during braking and propulsive phases (Winter, 1991; Nigg et al., 1994; Begg et al., 1998) and joint angular motion data such as the ankle, knee and hip joint angles (Judge et al., 1996; Kerrigan et al., 1998) have shown effects of aging. To date, however, the relative influence of these measures in differentiating the age groups has not been demonstrated.

Automated recognition of gait pattern changes by a machine classifier from their respective measures is expected to offer many potential advantages. For example, Maki (1997) using spatial-temporal measures of gait has shown significant changes in gait characteristics in the elderly fallers when compared to gait characteristics of elderly non-fallers. This research has particularly shown that some foot placement gait measures (e.g., step width and stride variability) displayed greater associations with falls prediction. Therefore, early identification of gait changes due to falling behaviour by a machine classifier might trigger initiation of necessary measures to prevent injurious falls such as an exercise intervention program (Lord

*Corresponding author. Tel: +61-3-9248-1116; fax: +61-3-9248-1110.
E-mail address: rezaul.begg@vu.edu.au (R. Begg).

0021-9290/$ - see front matter
doi:10.1016/j.jbiomech.2004.05.002

et al., 2003). Similar benefits could also be obtained in a clinical context via identification of abnormality in gait patterns and also by evaluating the effectiveness of treatment outcomes. In order to facilitate automated recognition of gait patterns, neural networks and fuzzy clustering techniques have been applied for classification of normal and pathological gait (Holzreiter and Kohle, 1993; O'Malley et al., 1997), and also to differentiate gait simulations, such as leg length discrepancy from joint-angle measures (Barton and Lees, 1997). However, it is well known that there are several limitations of neural network-based modelling, including: (i) dependency on a large number of parameters, e.g., network size, learning parameters and selection of initial weights, (ii) the possibility of being trapped into local minima, and (iii) over-fitting on training data resulting in poor generalization. Recently, support vector machines (SVM) have emerged as a powerful technique for general purpose pattern recognition. It has been applied to classification and regression problems with exceptionally good performance on a range of binary classification tasks (Zavaljevski et al., 2002; Ben-Yacoub et al., 1999; Chapelle et al., 1999; Ding and Dubchak, 2001; Chan et al., 2002). The primary advantage of SVM is its ability to minimize both structural and empirical risk (Gunn, 1998) leading to better generalization for new data classification.

Despite the success of SVM in other biomedical applications there has been little research on the classification ability of SVM in gait. Lee and Grimson (2002) showed that SVM achieved 94% accuracy in gender classification using gait video sequence data but there have been no reports known to the authors of SVMs applied to detect gait changes due to ageing. In the study reported here, we applied an SVM for automated recognition of young/old gait patterns using temporal and distance measures, kinetic and kinematic variables in the development of gait models to investigate their relative influence on classification.

2. Machine classifier: support vector machines

SVMs introduced by Vapnik (Vapnik, 1995) are a relatively new technique for classification and regression tasks. In a binary classification task like the one in this study, the aim is to find an optimal separating hyperplane. Fig. 1(a) shows a two-class problem with many possible hyperplanes separating the two data sets that are not necessarily optimal. In Fig. 1(b), an optimal separating hyperplane (OSH) is shown which generates the maximum margin (dashed line) between the two data sets. SVM finds this OSH by maximizing the margin between the classes. SVM first transforms input data into a higher dimensional space by means of a kernel function and then constructs a linear OSH between the two classes in the transformed space. Those data vectors nearest to the constructed line in the transformed space are called the support vectors (SV). SVM is an approximate implementation of the method of "structural risk minimization" aiming to attain low probability of generalization error (Haykin, 1999). In brief the theory of SVM is as follows (Vapnik, 1995; Kecman, 2002).

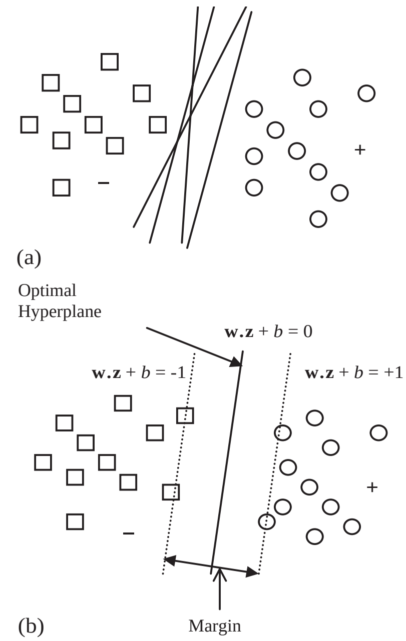

Fig. 1. An example of two-class (+&−) problem with: (a) many possible separating hyperplanes dividing the two groups, (b) optimal separating hyperplane and the maximum margin. The circles and squares represent samples of class +1 and −1, respectively.

Consider a training set $D = \{(\mathbf{x}_i, y_i)\}_{i=1}^{L}$, with each input $\mathbf{x}_i \in \Re^n$ and the associated output $y_i \in \{-1, +1\}$. Each input $\mathbf{x}$ is first mapped into a higher dimension feature space $\mathscr{F}$ by $\mathbf{z} = \phi(\mathbf{x})$ via a nonlinear mapping $\phi: \Re^n \to \mathscr{F}$. Considering the case when the data are linearly separable in $\mathscr{F}$, there exists a vector $\mathbf{w} \in \mathscr{F}$ and a scalar b that define the separating hyperplane as $\mathbf{w}\cdot\mathbf{z} + b = 0$ such that

$$y_i(\mathbf{w} \cdot \mathbf{z}_i + b) \geqslant 1, \forall i. \quad (1)$$

By maximizing the margin of separation between the classes ($2/\|\mathbf{w}\|$), SVM constructs a unique OSH as the one that minimizes $\mathbf{w} \cdot \mathbf{w}/2$ under the constraints of Eq. (1).

When the data are linearly non-separable, the above minimization problem is modified to allow classification error by introducing some non-negative variables $\xi_i \geqslant 0$, often called *slack variables*, such that

$$y_i(\mathbf{w} \cdot \mathbf{z}_i + b) \geqslant 1 - \xi_i, \forall i. \quad (2)$$

A non-zero ξ_i indicates a misclassified data point and $\sum_{i=1}^{L} \xi_i$ can be regarded as a measure of misclassification.

Table 1
List of kernel functions that were used to develop the SVM models

Kernel function	Mathematical formula
Linear	$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i.\mathbf{x}_j$
Polynomial	$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^d$, d is the degree of polynomial
Gaussin radial basis function (RBF)	$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\lVert\mathbf{x}_i - \mathbf{x}_j\rVert^2}{2\sigma^2}\right)$, σ is the width of RBF function
Exponential radial basis function (ERBF)	$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\lvert\mathbf{x}_i - \mathbf{x}_j\rvert}{2\sigma^2}\right)$
Multi-layer perceptron (MLP)	$K(\mathbf{x}_i, x_j) = \tanh(b(x_i.x_j) - c)$, b is the slope and c is the bias
Spline	$K(\mathbf{x}_i, \mathbf{x}_j) = 1 + \mathbf{x}_i \cdot \mathbf{x}_j + \frac{1}{2}(\mathbf{x}_i \cdot \mathbf{x}_j)\min(\mathbf{x}_i, \mathbf{x}_j) - \frac{1}{6}\min(\mathbf{x}_i, \mathbf{x}_j)^3$

SVM determines OSH by maximizing the margin and minimizing the training error as a solution of the following optimization problem.

$$\begin{aligned} &\textit{minimize} && \frac{1}{2}\mathbf{w}\cdot\mathbf{w} + C\sum_{i=1}^{L}\xi_i \\ &\textit{subject to} && y_i(\mathbf{w}\cdot\mathbf{z}_i + b) \geqslant 1 - \xi_i, \forall i \\ &\text{and} && \xi_i \geqslant 0, \forall i, \end{aligned} \tag{3}$$

where C is a constant parameter, called *regularization parameter*, that determines the trade-off between the maximum margin and minimum classification error. Minimizing the first term corresponds to minimizing the Vapnik–Chervonenkis (VC) dimension of the classifier and minimizing the second term controls the empirical risk (Kim et al., 2003).

Searching the optimal hyperplane in Eq. (3) is a quadratic programming (QP) problem that can be solved by constructing a Lagrangian and transforming into the following dual problem:

$$\begin{aligned} &\textit{maximize} && W(\alpha) = \sum_{i=1}^{L}\alpha_i - \frac{1}{2}\sum_{i=1}^{L}\sum_{j=1}^{L}\alpha_i\alpha_j y_i y_j \mathbf{z}_i.\mathbf{z}_j \\ &\text{subject to} && \sum_{i=1}^{L} y_i\alpha_i = 0 \text{ and } 0 \leqslant \alpha_i \leqslant C, \forall i, \end{aligned} \tag{4}$$

where $\alpha = (\alpha_1, \ldots, \alpha_L)$ is the non-negative Lagrangian multiplier. The data points $\mathbf{x}_i$ corresponding to $\alpha_i > 0$ lie along the margins of decision boundary and are SVs.

The term $\mathbf{z_i} \cdot \mathbf{z_j}$ in Eq. (4) can be computed by using a kernel function $K(.,.)$ without having to obtain $\phi(\mathbf{x}_i)$ and $\phi(\mathbf{x}_j)$ explicitly such that $\mathbf{z_i} \cdot \mathbf{z_j} = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j)$. Having determined the optimum Lagrange multipliers, the optimum solution for the weight vector $\mathbf{w}$ is given by

$$\mathbf{w} = \sum_{i \in \mathrm{SVs}} \alpha_i y_i \mathbf{z}_i, \tag{5}$$

where SVs are the support vectors. For any test vector $\mathbf{x} \in \Re^n$, the output is then given by

$$\begin{aligned} y = f(\mathbf{x}) &= \operatorname{sign}(\mathbf{w}\cdot\mathbf{z} + b) \\ &= \operatorname{sign}\left(\sum_{i \in \mathrm{SVs}} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b\right). \end{aligned} \tag{6}$$

To construct SVMs, users must select a kernel function. So far, no analytical or empirical study has conclusively established the superiority of one kernel over another; thus the performance of SVMs in a particular task may vary with this choice. In this study, we experimented with six kernels as shown in Table 1.

3. Feature extraction

3.1. Gait data acquisition and features

Twenty-four healthy adults (12 young and 12 elderly; 50% male in both age groups) from the academic community of Victoria University and from a local senior citizen club participated in the study. All subjects undertook informed-consent procedures as approved by the Victoria University Human Research Ethics Committee. The subjects had no known injuries or abnormalities that would affect their gait. Mean ages (standard deviation in brackets) of the two groups were as follows; young 28.1($\pm$5.6) years, elderly 68.8($\pm$4.6) years.

Gait recordings (both kinematics and foot ground reaction forces) were performed during comfortable walking on a 15 m laboratory walkway. All subjects in this study completed 3 gait trials and the features were calculated using the mean of 3 trials. Mean($\pm$standard deviation) walking velocities of the two groups were: young 1.27($\pm$0.35) m/s, elderly 1.05($\pm$0.33) m/s. The intra-subject variability (standard deviation) in walking speeds among the gait trials varied from 0.01 to 0.07 across the subjects. Three types of gait parameters (basic spatial/temporal, kinetic and kinematic) were recorded during normal walking on the laboratory walkway. Foot-ground reaction forces in the vertical and anterior–posterior directions were recorded using two force-sensing platforms (AMTI, USA). Peak forces during heel-strike, mid-stance, and push-off phases were extracted and normalized to body weight (Giakas and Baltzopoulos, 1997). Fig. 2 illustrates sub-division of the stance phase and the features extracted from these graphs. Movement of the lower limb was recorded using a 3D PEAK (Peak Performance Inc., USA) Motion analysis system via reflective markers attached to lower limb joints and segments (hip, knee, ankle, heel and toe).

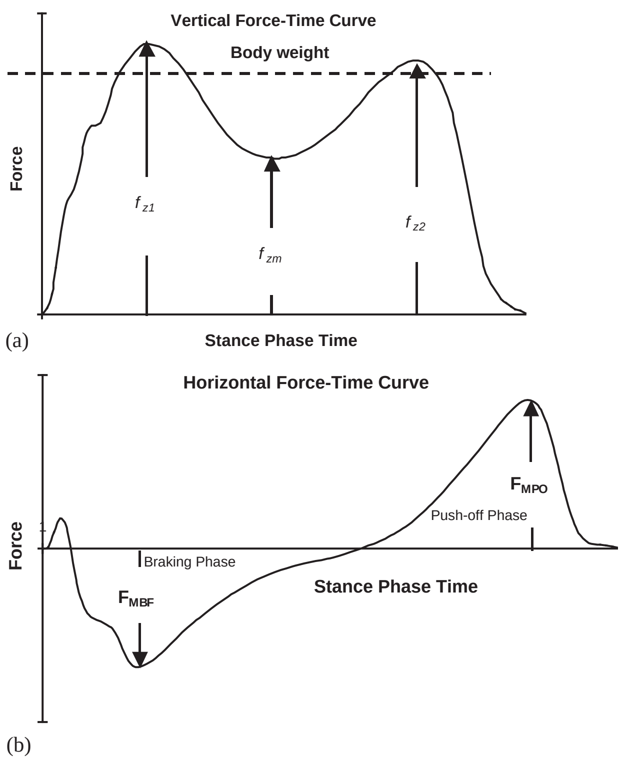

Fig. 2. Typical foot-ground reaction forces during gait showing the key kinetic features extracted and stance sub-phases: (a) Vertical force-time graph (f_{z1}-maximum vertical force during heel-strike, f_{zm}-minimum vertical force during mid-stance and f_{z2}-maximum vertical force during push-off), (b) Horizontal force–time graph (F_{MBF}-maximum heel strike force during the braking phase, F_{MPO}-maximum push-off force during push-off phase).

Joint angles at heel contact and toe-off and angular range of motion (ROM) during the stance and swing phases were calculated. Altogether, 24 gait features describing basic, kinetic and kinematic aspects of gait data were extracted and are listed in Table 2.

3.2. Training and testing the SVM

All 24 gait features were normalized by calculating their z-scores (i.e., $(x-\mu)/\sigma$, where μ is the mean and σ is the standard deviation for the gait feature) before applying them to the classifiers. A six-fold cross-validation scheme was adopted to evaluate the generalization ability of the classifier. Cross-validation procedures have been used in a number of classification evaluations, particularly for limited data sets (Barton and Lees, 1997). In this scheme, the data set was uniformly divided into six subsets with one used for testing and the other five used to train and construct the SVM decision surface. This was repeated for other subsets so that all subsets were used as the testing sample.

The following three measures of accuracy, sensitivity and specificity were used to assess the performance of the SVM classifier (Chan et al., 2002; Pang et al., 2003).

$$\text{Accuracy} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{FP}+\mathrm{TN}+\mathrm{FN}} \times 100\%,$$

$$\text{Sensitivity} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \times 100\%,$$

$$\text{Specificity} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}} \times 100\%,$$

where TP is the number of true positives, i.e., the SVM identifies an elderly gait that was labelled as elderly; TN is the number of true negatives, i.e., SVM identifies a young gait that was labeled as young; FP is false elderly identifications; and FN is false young identifications. Accuracy indicates overall detection accuracy, sensitivity is defined as the ability of the classifier to accurately recognize an elderly gait pattern whereas specificity would indicate the classifier's ability not to generate a false detection (normal young gait).

Table 2
Gait features used to train and test the SVMs. Temporal data were normalized to gait cycle; force data were normalized to body weight, HC—heel contact, TO-toe off, ROM-range of motion. Features extracted from ground reaction forces include peak forces in the vertical and horizontal (anterior–posterior) directions (refer to text and Fig. 2 for more description). All data were normalized to their equivalent z-scores before applying to the SVM

Basic features (9)	Kinetic features (5)	Kinematic features (10)
• Absolute/normalized stance swing and double support time • Stride length (m) • Walking speed (m/s) • Cadence (steps/min)	• Max vertical HC force (f_{z1}) • Vertical min mid-stance force (f_{zm}) • Max vertical push-off force (f_{z2}) • Max horizontal HC force during braking phase (F_{MBF}) • Max horizontal push-off force (F_{MPO})	• Ankle angle at HC & TO • Knee angle at HC & TO • Ankle ROM during stance, swing & stance-to-swing phases • Knee ROM during stance, swing & stance-to-swing

Table 3
Overall accuracy (Acc), sensitivity (Sen) and specificity (Spe) of gait detection using different kernel functions and types of gait features in six-fold experiment

Kernel function and performance measures		Gait variables				
		Basic	Kinetic	Kinematic	Kinetic + Kinematic	All variables
Linear	Acc	62.5	83.3	87.5	91.7	91.7
	Sen	58.3	91.7	75.0	91.7	91.7
	Spe	66.7	75.0	100.0	91.7	91.7
Polynomial	Acc	58.3	83.3	91.7	91.7	91.7
	Sen	57.1	91.7	83.3	91.7	91.7
	Spe	60.0	75.0	100.0	91.7	91.7
Gaussian Radial Basis Function	Acc	62.5	83.3	83.3	91.7	91.7
	Sen	66.7	83.3	75.0	91.7	100.0
	Spe	58.3	83.3	91.7	91.7	83.3
Exponential Radial Basis Function	Acc	62.5	87.5	83.3	91.7	91.7
	Sen	66.7	91.7	75.0	91.7	91.7
	Spe	58.3	83.3	91.7	91.7	91.7
Multi-layer Perceptron	Acc	62.5	83.3	87.5	87.5	91.7
	Sen	58.3	83.3	75.0	83.3	91.7
	Spe	66.7	83.3	100.0	91.7	91.7
Spline	Acc	62.5	83.3	87.5	91.7	87.5
	Sen	66.7	83.3	83.3	91.7	91.7
	Spe	58.3	83.3	91.7	91.7	83.3

Results of six cross-validation tests were combined to obtain an average result for the three measures of accuracy. Tests were also conducted to examine performance of the SVMs for different kernel functions (see Table 1) and different regularization parameters '*C*'.

4. Experimental results

Overall accuracy, sensitivity, and specificity results for young and elderly gait pattern identification are summarized in Table 3. Accuracy was at best 62.5% when all nine basic gait variables were used in the SVM inputs, however, the accuracy rate reached 91.7% (in polynomial kernel) when kinematic data were used to train the SVM. In general, combining features showed improved performance compared to individual data types. There were some differences in performance among the six kernels when applied to individual data types but when all 24 features were combined the kernels provided similar classification performance (91.7%) except the Spline that provided slightly reduced accuracy (87.5%).

Sensitivity and specificity results showed low to moderate values for the basic gait data (range: 57.1–66.7%). Kinetic data showed higher sensitivity across all the six kernels (83.3–91.7%) compared to specificity (75.0–83.3%), whereas kinematic data demonstrated the opposite trend with consistently higher specificity (91.7–100.0%) across all the kernels relative to their sensitivity results (75.0–83.3%). When both data types were combined, the classifier performance was found to be the same (91.7%) on both sensitivity and specificity measures in all but the MLP kernel.

Table 4
CPU time required to construct support vectors using different kernels (No. of features = 24)

Kernel Function	Average CPU time (ms)
Linear	0.1085
Polynomial	0.1153
Gaussian radial basis function	0.2605
Exponential radial basis function	0.1185
Multi-layer perceptron	0.1187
Spline	0.1318

Table 4 presents the average CPU time (ms) needed to construct SVM classifier on a 1.6 GHz processor PC for each kernel function. The time depends on the computational cost to calculate the kernel matrix and solve the optimization problem. Linear kernel is simple and computationally the fastest (0.1085 ms) whereas Gaussian RBF kernel proved to be the most expensive computationally ($\sim$2.5 times slower than linear kernel).

Performance of the classifier (linear kernel) as a function of the regularization parameter 'C' (Fig. 3) revealed maximum performance (91.7%) within a

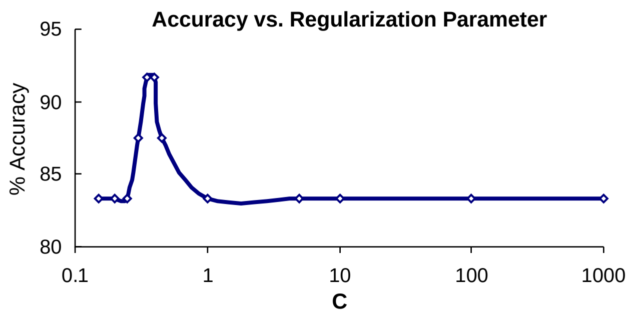

Fig. 3. Dependence of % Classification accuracy on 'C' (linear kernel function).

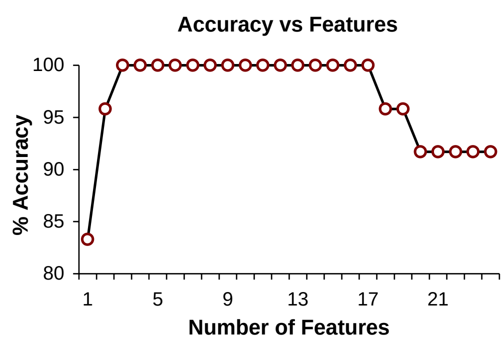

Fig. 4. Dependence of % classification accuracy on the number of features selected by the forward selection algorithm (linear kernel). Best three selected features that provided maximum accuracy: knee range of motion (K_{ROM}), horizontal peak push-off force (F_{MPO}) and normalized double support time (%DS).

narrow range of C (0.2–0.45), outside this range performance deteriorated (83.3%).

In order to test the effect of number of features on classification performance, a forward feature selection algorithm was used in which a feature was sequentially added one at a time that most increased or least decreased the classification accuracy (Chan et al., 2002). Fig. 4 plots accuracy as a function of features and shows that with only 3 selected features perfect classification (100.0%) of the two age groups can be achieved. The graph also highlights that for these gait data, after 17 selected features the overall performance of the classifier deteriorated with additional features.

5. Discussion

Results of this study suggested that SVMs can discriminate between young and elderly walking. The test results also demonstrated that the SVM models were able to map the underlying data structure relating to young and ageing populations. Such discriminative quality has many applications in clinical and rehabilitation contexts, for example, identification of abnormality in gait patterns and evaluating the effectiveness of interventions.

One aim was to test which type of gait feature would be most effective in the young/old classification task. SVM models were developed using three commonly used types of gait data (basic spatial/temporal, kinetic and kinematic). Table 3 suggests that basic spatial/temporal and kinetic gait features alone did not offer maximum classification results. Kinematic data, however, provided the maximum accuracy (91.7%) using a third-order polynomial kernel. Overall, the recognition rate improved for all the kernels when kinetic and kinematic data types were combined to develop the SVM models, suggesting that the SVM classification performance could be enhanced by improving the information contents of the recognition system similar to the one used in the data fusion technique (Ben-Yacoub et al., 1999).

Sensitivity measures classifier's ability to detect elderly gait patterns whereas specificity represents detecting young gait patterns. The results revealed improved sensitivities for kinetic features such as maximum and minimum forces, which appear to provide vital information for detecting characteristics of ageing gait. Joint angle measures during stance, swing, and stance-to-swing transitions are useful in detecting young gait characteristics as evident by their consistently higher specificity values. The results in Table 3 also suggested that kinetic and kinematic features complement each other when detecting young-old gait patterns as evidenced by similar sensitivity and specificity outcomes.

It appears that not all gait features are good contributors to separation between the two age groups because it was found that (Fig. 4) with only 3 features it was possible to achieve greater accuracy than using all 24 features. After 17 features, inclusion of further features caused overall classification accuracy to deteriorate. This suggests some gait features are redundant, in not providing additional discriminatory information. Similar findings have been reported in glaucoma diagnosis (Chan et al., 2002), and also for discriminating movement patterns from brain–computer-interface data where maximum classification performance was found with only 20 selected features out of 1000 features (Yom-Tov and Inbar, 2002). The forward selection algorithm irrespective of model parameters consistently selected two features: (1) knee angular ROM (K_{ROM}) and (2) maximum horizontal peak push-off force (F_{MPO}). K_{ROM} and F_{MPO} alone provided a maximum separation of 95.8% showing that they yielded the most discriminatory information in the development of separating hyperplanes by the SVM. A 2D scatter-plot of K_{ROM} and F_{MPO} (Fig. 5) confirms this. Statistical tests using multivariate analysis of variance (MANOVA,

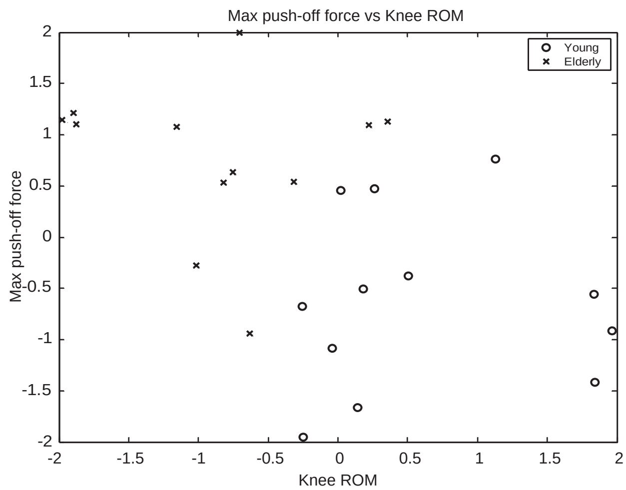

Fig. 5. Scatter plots showing distribution of two groups' data as a function of the two best features (maximum horizontal push-off force (F_{MPO}) and knee range of motion (K_{ROM}) during the swing phase) selected by the forward feature selection algorithm. Features are plotted using their respective *z*-score data.

Spss Inc.) were performed, with K_{ROM} and F_{MPO} as dependent variables and ageing (young/old) as a factor, to test differences between the two age groups. The results of the statistical testing revealed that both of these gait features were significantly different between young and elderly (K_{ROM}: $F(1,22) = 20.31$, $p < 0.0001$; F_{MPO}: $F(1,22) = 17.56$, $p < 0.0001$), thus supporting the performance of the SVM method. Winter (1991) also reported F_{MPO} to be significantly affected by ageing. It is interesting to note that the third feature selected to attain 100% classification was normalized double support time, which is a basic gait parameter. Feature selection appears to be useful for identifying the important discriminatory features and also to eliminate redundant features. Using a subset of the original features for developing predictive models has other benefits too, including model simplicity and reduced training time (Muller et al., 2003). Other feature selection algorithms have been proposed in the literature, such as backward elimination techniques (Chan et al., 2002) and genetic algorithms (Yom-Tov and Inbar, 2002).

Regularization parameter C provides a balance between classification violation and margin maximization (Eq. (3)). A high C can minimize training error but will also compromise margin separation. As illustrated in Fig. 3, the classifier performance has an optimal range, outside of which performance may decrease, therefore, it is important that the C is carefully selected.

In a gender classification task by SVM from gait video sequence data, Lee and Grimson (2002) reported a comparable success rate (91%) with a second-order polynomial kernel and a slightly better performance (94%) with linear kernel. In recognizing ageing effects, our results indicate similar recognition rate with 24 gait features, however shows 100% accuracy with properly selected fewer gait features. Overall, the results of this experiment suggest that SVMs are an effective tool for recognizing gait pattern changes with aging and it holds considerable potential for future applications involving gait pattern detection and classification.

References

Barton, J.G., Lees, A., 1997. An application of neural networks for distinguishing gait patterns on the basis of hip-knee joint angle diagrams. Gait and Posture 5, 28–33.

Begg, R.K., Sparrow, W.A., Lythgo, N.D., 1998. Time-domain analysis of foot-ground reaction forces in negotiating obstacles. Gait and Posture 7, 99–109.

Begg, R.K., Sparrow, W.A., 2000. Gait characteristics of young and older individuals negotiating a raised surface: implications for the prevention of falls. Journal of Gerontology: Medical Sciences 55A, 147–154.

Ben-Yacoub, S., Abdeljaoued, Y., Mayoraz, E., 1999. Fusion of face and speech data for person identity verification. IEEE Transactions on Neural Networks 10, 1065–1074.

Chan, K., Lee, T.W., Sample, P.A., Goldbaum, M.H., Weinreb, R.N., Sejnowski, T.J., 2002. Comparison of machine learning and traditional classifiers in glaucoma diagnosis. IEEE Transactions on Biomedical Engineering 49, 963–974.

Chapelle, O., Haffner, P., Vapnik, V.N., 1999. Support vector machines for histogram-based classification. IEEE Transactions on Neural Networks 10, 1055–1064.

Ding, C.H.Q., Dubchak, I., 2001. Multi-class protein fold recognition using support vector machines and neural networks. Bioinformatics 17, 349–358.

Giakas, G., Baltzopoulos, V., 1997. Time and frequency domain analysis of ground reaction forces during walking: an investigation of variability and symmetry. Gait and Posture 5, 189–197.

Gunn, S., 1998. Support vector machines for classification and regression. ISIS Technical Report, University of Southampton, UK.

Hageman, P.A., Blanke, D.J., 1986. Comparison of gait of young women and elderly women. Physical Therapy 66, 1382–1387.

Haykin, S., 1999. Neural Networks—A comprehensive Foundation. Prentice-Hall, Englewood Cliffs, NJ.

Holzreiter, S.H., Kohle, M.E., 1993. Assessment of gait pattern using neural networks. Journal of Biomechanics 26, 645–651.

Judge, J.O., Davis, R.B., Õunpuu, S., 1996. Step length reductions in advanced age: the role of ankle and hip kinetics. Journal of Gerontology: Medical Sciences 51A, 303–312.

Kerrigan, D.C., Todd, M.K., Croce, U.D., Lipsitz, L.A., Collins, J.J., 1998. Biomechanical gait alterations independent of speed in the healthy elderly: evidence for specific limiting impairments. Archives of Physical Medicine and Rehabilitation 79, 317–322.

Kecman, V., 2002. Learning and Soft Computing: Support Vector Machines, Neural Networks and Fuzzy Logic Models. IEEE MIT Press, USA, NJ.

Kim, H., Pang, S., Je, H., Kim, D., Bang, S.Y., 2003. Constructing support vector machine ensemble. Pattern Recognition 36, 2757–2767.

Lee, L., Grimson, W.E.L., 2002. Gait analysis for recognition and classification. In: Proceedings of the Fifth International Conference on Automatic Face and Gesture Recognition, IEEE Computer Society, Silver Spring, MD.

Lord, S.R., Castell, S., Corcoran, J., Dayhew, J., Matters, B., Shan, A., Williams, P., 2003. The effect of group exercise on physical functioning and falls in frail older people living in retirement villages: a randomized, controlled trial. Journal of American Geriatric Society 51, 1685–1692.

Maki, B.E., 1997. Gait changes in older adults: predictors of falls or indicators of fears. Journal of American Geriatric Society 45, 313–320.

Muller, K., Anderson, C.W., Birch, G.E., 2003. Linear and nonlinear methods for brain computer interfaces. IEEE Transactions on Neural System and Rehabilitation Engineering 11, 165–169.

Nigg, B.M., Fisher, V., Ronsky, J.L., 1994. Gait characteristics as a function of age and gender. Gait and Posture 2, 213–220.

O'Malley, M.J., Abel, M.F., Damiano, D.L., Vaughan, C.L., 1997. Fuzzy clustering of children with cerebral palsy based on temporal-distance gait parameters. IEEE Transactions on Rehabilitation Engineering 5, 300–309.

Pang, C.C.C., Upton, A.R.M., Shine, G., Kamath, M.V., 2003. A comparison of algorithms for detection of spikes in the electroencephalogram. IEEE Transactions on Biomedical Engineering 50, 521–526.

Princea, F., Corriveaua, H., Héberta, R., Winter, D.A., 1997. Gait in the elderly. Gait and Posture 5, 128–135.

Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. Springer, New York.

Winter, D., 1991. The Biomechanics and Motor Control of Human Gait: Normal, Elderly and Pathological. University of Waterloo Press, Waterloo.

Yom-Tov, E., Inbar, G.F., 2002. Feature selection for the classification of movements from single movement-related potentials. IEEE Transactions on Neural Systems and Rehabilitation Engineering 10, 170–177.

Zavaljevski, N., Stevens, F.J., Reifman, J., 2002. Support vector machines with selective kernel scaling for protein classification and identification of key amino acid positions. Bioinformatics 18, 689–696.