1-1-2013

Additive manufacture of titanium alloys using integrated robotic GTAW welding

John Norrish
University of Wollongong, johnn@uow.edu.au

Dominic Cuiuri
University of Wollongong, dominic@uow.edu.au

Additive manufacture of titanium alloys using integrated robotic GTAW welding

Abstract

Additive manufacturing (also known as rapid prototyping, shape metal deposition or near net manufacture) is becoming an important option for low volume production of complex parts. Much of the recent effort has been devoted to laser based processes using powder feedstock and 3D printing techniques; these techniques are well suited to producing smaller complex structures in a wide range of materials. The current paper describes a technique for cold wire GTAW deposition of titanium alloys using robotic welding and integrated post weld machining; the approach is suitable for producing much larger parts. The techniques involved and resultant weld quality will be discussed.

Keywords

manufacture, titanium, alloys, integrated, additive, robotic, gtaw, welding

Disciplines

Engineering | Science and Technology Studies

Publication Details

Norrish, J. & Cuiuri, D. (2013). Additive manufacture of titanium alloys using integrated robotic GTAW welding. International Conference on Joining Materials 17 International Institute of Welding.

Pathological Gait Detection of Parkinson's Disease using Sparse Representation

Yuyao Zhang [‡ §], Philip O. Ogunbona [‡ §], Wanqing Li [‡ §], Bridget Munro [† ¶], Gordon G. Wallace [* †]
[*]Intelligent Polymer Research Institute
[†]ARC Centre of Excellence for Electromaterials Science
[‡]School of Computer Science and Software Engineering
[§]Information and Communication Technology Research Institute
[¶]Biomechanics Research Laboratory, School of Health Sciences
University of Wollongong, Australia
Email: {yz606, philipo, wanqing, bmunro, gwallace}@uow.edu.au

***Abstract*—Gait analysis has become an attractive quantitative and non-invasive mechanism that can aid early detection and monitoring of the response of Parkinson's disease sufferers to management schedules. In this paper, we model cycles of human gait as a sparsely represented signal using over-complete dictionary. This representation forms the basis of a classification that allows the recognition of symptomatic subjects. Experiments have been conducted using signals of vertical ground reaction force (GRF) from subjects with Parkinson's disease from the publicly available gait database (physionet.org). Our method achieved a classification accuracy of 83% in recognising pathological cases and represents a significant improvement on previously published results that use a selection of the Fourier transform coefficients as features.**

I. Introduction

Gait analysis can be used in the quantification of human locomotor disorders. By quantitatively studying human walking patterns, gait analysis is able to assist in the diagnosis of neuromusculoskeletal diseases, therapy planning and rehabilitation evaluation [1]. Increasingly, gait analysis is being used to provide supplementary reports to support diagnosis or evaluate rehabilitation progress of subjects undergoing treatment or assessment in hospitals or gait laboratories [2].

When used as a diagnostic tool, it can facilitate the detection of pathological gait which may be useful in clinical decision-making and automatic disease recognition. Pathological gait is a significant indicator for many neuromusculoskeletal diseases, especially in early stages. Gait analysis can also be used to conduct quantitative studies on longitudinal data from pathological subjects. Such studies provide quantitative and objective evaluation of changes in gait before and after undergoing treatments and can assist decision making for therapy updating and rehabilitation.

The types of data captured from gait monitoring devices include kinetic, kinematic or temporal-distance and are of high dimensionality and variability. For instance, the kinetic data may include ground reaction force beneath various parts of the feet such as toes and heels. The kinematic data may involve angles of joints in frontal plane (abduction and adduction), sagittal plane (flexion and extension) and transverse plane (external and internal rotations) [2]. Once captured, the data constitutes large number of time series samples with high dimensionality. Furthermore, gait has a high intra-subject, inter-subject, inter-cycle variability. These problems exacerbate the difficulty of modelling the data and time taken to conduct the gait analysis [3].

Parkinson's disease is an example of nervous disease with symptoms of movement disorders. The degeneration of central nervous system decreases the ability of the individual to control their locomotor systems in the early stage of Parkinson's disease. The gait is affected and thus its analysis is considered a quantitative and noninvasive method for early detection of Parkinson's disease. As mentioned above, the high dimensionality and variability of gait data available for analysis make the detection task difficult.

In this paper, we use a sparse representation based on over-complete dictionary to model gait data of subjects with Parkinson's disease and identify pathological gait patterns. Sparse representation seeks a sparse linear combination of atoms from a given dictionary to faithfully represent an input signal. The reconstructed signal based on the dictionary is such that it is the best approximation using the minimum number of atoms. Compared to traditional pattern recognition and machine learning algorithms, sparse representation provides an easy interpretation for diagnostic results. The selected atoms for representing an input gait signal together with coefficients used in linear combination become evidence for diagnosis and are more descriptive. Doctors and clinicians usually prefer a descriptive method to a black box, as would be the case for neural network, for easier understanding.

The remaining parts of this paper are organised as follows. Section II presents related work in which quantitative gait analysis have been conducted to achieve pathological gait recognition. In Section III, details about the proposed algorithm as well as theoretical formulation are provided. We describe the dataset used in our experiments in Section IV. Experimental design to evaluate our proposed algorithm and results obtained on GRF signals to recognise subjects with Parkinson's disease are described in Section V. We offer some concluding remarks in Section VI.

II. Related Work

Quantitative gait analysis normally involves discriminative feature extraction and quantitative analysis. A data set for a single subject may combine kinetic, kinematic and temporal-distance data types, depending on the available capture equipment and the intended purpose of the study. It is also known that data describing human movement are highly correlated, temporally dependent and normally contains a large number of time series samples [3]. For these reasons, it is important to employ some dimensionality reduction technique to process the data and obtain discriminative features in a reduced dimension.

The most intuitive approach is to work on the bio-mechanical features extracted from gait data since they are either highly related to certain diseases or generally shared by various neuralmuscular diseases. Senanayake et al. [4] derived the timing of gait phases from both kinetic and kinematic features. By accurately identifying each phase and testing the timing of the gait phases, pathological gait is detectable. Besides the timing of gait phases, other bio-mechanical features include peak time, intensity of GRF signals [5], minimum foot clearance (MFC) of toe displacement [6]. These bio-mechanical features significantly reduce the data dimension by only selecting values with clinical or bio-mechanical importance. Meanwhile, temporal dependence between gait cycles is neglected. The extraction of bio-mechanical features requires detection of specific gait events. However, gait events derived from an able-bodied subject may be difficult to locate in symptomatic gait data. The various methods available to define and calculate these events would further increase the subjectivity in gait analysis [7].

A more robust approach to extract features from gait data is to use analytical techniques that seek discriminative features by mathematically modelling the gait data. Such methods are able to identify local and global features automatically. Fourier transform has been used to derive features from gait signals [8]. By retaining only the first 28 coefficients from a 128-point Fourier transform of vertical GRF, the authors in [8] reduced the dimensionality of the feature of interest in the data space. The transformation and selection also preserved the top-level features between normal and abnormal gait signals. A neural network was then trained to classify the reduced gait data and a high classification rate of 95% was achieved. It is not clear whether this result represent an average over several tests or a cross-validation output. However, most pathological subjects in this study were patients with calcareous fractures and artificial limbs, who suffer heavy movement impairment. The data selected was the vertical GRF pairs of both limbs in the stance phase and did not cover long-time walking information. Furthermore, by applying Fourier transform and dropping portions of the coefficients, the information on time axis and subtle changes of the force intensity were lost. Clinically important components might also be missed. Köhle and Merkl extended the work by employing a different neuron function and involving force intensity in three dimensional space [9]. Deluzio and Astephen [10] work on gait waveforms to extract discriminative components and reduce data dimension using principal component analysis. By excluding several less important components, discriminative features are gained with lower dimension. Some other techniques that have been used to extract discriminative features and reduce the data dimension include factor analysis (FA) and multiple correspondence analysis (MCA) [3].

The applications of quantitative gait analysis mainly falls into two categories: evaluation of rehabilitation or treatment performance and pathological gait detection. Both of these applications provide important indicators for planning and refining treatments schedules. Yang et al. [11] proposed an algorithm to assess gait patterns of complex regional pain syndrome (CRPS) using multiple layer perceptron neural networks (MLP). PCA has also been used to analyse the gait data to seek the most statistically significant patterns for further evaluation [10]. Three kinds of gait signals related to knee osteoarthritis from multiple subjects were used in the study. Principal component analysis was used as, not only dimensionality reduction method but also feature extraction algorithm. The authors interpret the computed principal components in the view of bio-mechanical properties and conducted discriminative analysis to classify the data set. They achieved a result of misclassification rate of 8%.

Pathological gait detection assists in diagnosing diseases impacting locomotor systems. Efforts in this area of research have employed techniques from computational intelligence; for example, multi-variate statistical methods, support vector machine and neural networks [12]. Lai et al. [5] extracted peak values and time to those peaks in five direction from GRF. Kinematic features including maximum angle of rear foot (measured in three anatomical planes) and tibia (measured only in transverse plane) as well as time to these maximum points were put together with feature points from GRF. The resultant 30-feature samples was collected from 27 subjects. Support vector machine was used and a good classification rate was obtained by combining features from both GRF signals and kinematic signals. Some recent research also apply hidden Markov model on detecting pathological gait phase [13] and fuzzy inference systems which treat the variability within gait data as non-probabilistic uncertainties [4], [3]. There is also a trend to combine multiple algorithms to gain both high classification rate and descriptivity. For example, in [14], decision tree and neural networks were combined to rank knee osteoarthritis. In this paper, we focus on pathological gait detection of subjects with Parkinson's disease. Sparse representation is adopted in the extraction of discriminative features.

III. Pathological Gait Detection based on Sparse Representation

In our proposed method we pose pathological gait detection as a binary classification problem. A query gait signal is labelled as either normal or pathological during the classification process. In particular, we adopt sparse representation

of signals based on learned dictionary to detect pathological gait from subjects with Parkinson's disease. The dictionary learning algorithms address the problem of modelling gait signals with intra-subject and inter-subject variations. Based on sparse representation constraints and learned dictionary, atoms from the dictionary are able to faithfully recover the training signal. To motivate our approach we assume a gait signal is a combination of prototype signals with inter-subject and intra-subject variations. The prototype signals are the atoms of the learned dictionary. A query gait signal can be assigned to a certain class according to the reconstruction error [15] incurred by sparse representation or using a jointly trained classifier to work on the sparse coefficients [16], [17].

Sparse representation seeks a sparse linear combination of atoms from a given dictionary to faithfully reconstruct an input gait signal. The dictionary is represented as a matrix whose columns are the atoms and the training signals lie in the associated column space. The sparse constraint ensures that only a few of the columns are needed for the representation. In other words the coefficient vector has few non-zero entries. We represent a database with n data samples as $\mathbf{Y} = \{y_1, y_2, \ldots, y_n\}$. Each data sample $y_i \in R^d$ may involve several features extracted from either a normal subject or a subject with Parkinson's disease. Let the sparse coefficients for the database $\mathbf{Y}$ be denoted by a matrix $\mathbf{X} = \{x_1, x_2 \ldots x_n\}$, where $\{x_i \in R^d, \quad i = 1, 2, \ldots n\}$. The problem of sparsely encoding a gait signal can be formulated as follows [18],

$$\arg\min_{x_i} \|y_i - \mathbf{D} \cdot x_i\|_2^2 \quad \text{s.t.} \quad \|x_i\|_0 < T_0, \tag{1}$$

where $\mathbf{D}$ represents the dictionary matrix with $\{d_i \in R^d, \quad i = 1, 2, \ldots, K\}$ as its atoms. $T_0 < k$ is a predefined number which constrains the maximum number of atoms allowed have corresponding non-zero entry in the coefficient vector. In Fig. 1, a gait signal and its sparse representation coefficients are shown. The signal in Fig. 1 records the ground reaction force in vertical direction for one gait cycle. The 6^{th} atom in Fig. 1 (a) gains the largest coefficient value and hence is more likely to be in the same class as the test signal.

Gait signals are of high variability between and within subjects as mentioned above. The dictionary $\mathbf{D}$ is expected to only adequately reconstruct either normal or symptomatic gait data using sparse representation in the classification task. A dictionary $\mathbf{D}$ learned from a database is supposed to have its atoms capture the structure of signals in the database. Thus, the training database with gait signals from multiple pathological subjects and more training samples than the desired number of atoms (size of dictionary) is used to train one dictionary. Another dictionary is trained using the database containing normal gait signals with the same structure as the pathological database. The two dictionaries are trained either separately by K-SVD [18] or jointly by D-KSVD [16] and LC-KSVD [17]. The performance of each algorithm will be evaluated in Section IV.

K-SVD is a dictionary learning algorithm which generalises the K-means clustering algorithm [18]. K-SVD aims to find a dictionary and a set of coefficients to minimise the reconstruction error with respect to the whole training database. The objective function is written as in (2),

$$\arg\min_{\mathbf{D}} \sum_{i=1}^{n} \|y_i - \mathbf{D} \cdot x_i\|_2^2 \quad \text{s.t.} \quad \|x_i\|_0 \leq \mathbf{T}_0. \tag{2}$$

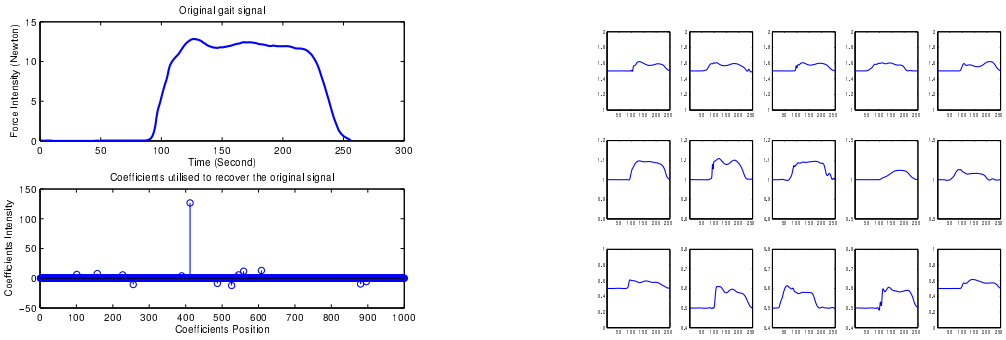
Orthogonal matching pursuit (OMP) [19] is used to iteratively solve (1) prior to optimising (2). A similar matrix is calculated using singular value decomposition (SVD) for all k atoms in the dictionary $\mathbf{D}$. The dictionary $\mathbf{D}$ is iteratively adapted to represent database $\mathbf{Y}$. Sparse representation-based classification (SRC) [15] drives the classification process using the trained dictionaries. In its original formulation SRC directly selects training signals from c different classes to construct c sub-dictionaries without an explicit feature extraction procedure. Each sub-dictionary $\mathbf{D}_i \in R^{d \times N_i}$ contains selected vector samples among all the N_i samples belonging to the ith class. In this paper, the two sub-dictionaries, $\mathbf{D}_\mathrm{c}$ and $\mathbf{D}_\mathrm{p}$, are trained for control and pathological classes respectively. A query sample is sparsely encoded using the trained sub-dictionaries either separately or jointly as one dictionary $\mathbf{D}_\mathrm{J} = \{\mathbf{D}_\mathrm{c}, \mathbf{D}_\mathrm{p}\}$. The sample is assigned to the class which can faithfully recover the input. The category of an unknown sample y_k will be identified as [15],

$$\mathrm{Label}_k = \arg\min_{k} \|y_k - \mathbf{D}_i \cdot x_{\mathbf{D}_i}\|_2 \quad i \in [\mathrm{p}, \mathrm{c}]. \tag{3}$$

In classification tasks using sparse representation, the over-complete dictionary is expected to be both representative and discriminative. However, the discriminative constraints were not explicitly forced during the optimisation process in conventional dictionary learning algorithms. Only representation performance is optimised in the training process. The classification relied on the differences manifesting in the sparse nature of test samples. D-KSVD [16] is a dictionary learning method that constructs a dictionary which is both representative and discriminative. The formulated optimisation problem has been included into the framework of K-SVD and is written as,

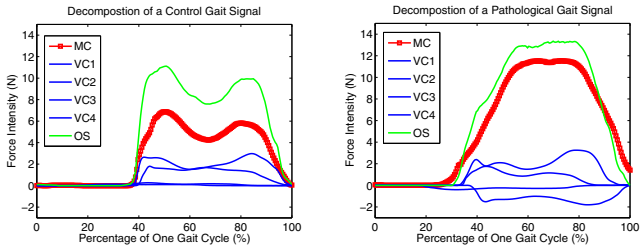
$$\arg\min_{\mathbf{W},\mathbf{X},\mathbf{D}} = \left\| \begin{pmatrix} \mathbf{Y} \\ \sqrt{\alpha} \cdot \mathbf{H} \end{pmatrix} - \begin{pmatrix} \mathbf{D} \\ \sqrt{\alpha} \cdot \mathbf{W} \end{pmatrix} \cdot \mathbf{X} \right\|_F^2 \quad \text{s.t.} \quad \|\mathrm{x}_i\|_0 < T_0, \tag{4}$$

where $\mathbf{H} = \{h_1, \ldots, h_n\}$ contains the label information, $\mathbf{W}$ is a jointly trained linear classifier and α is the parameter controlling contribution of label information. The linear classifier $\mathbf{W}$ is trained with the dictionary $\mathbf{D}$ together using K-SVD algorithm. The objective function (4) associates each training sample with a class label and combines the classification error and reconstruction error together in selecting appropriate atoms. It is possible to derive a classification scheme using the trained dictionary $\mathbf{D}$ and the linear classifier $\mathbf{W}$. Label consistent K-SVD (LC-KSVD) [17] further associates each atom in the dictionary with a class label at the initialisation



(a) (b)

Fig. 1. (a) A gait signal with its sparse coefficients and (b) Atoms utilised to recover the original signal



(a) (b)

Fig. 2. (a) Decomposed components for a normal gait cycle (b) Decomposed components for a pathological gait cycle

stage. By adding a label consistent error to the objective function of D-KSVD, the new objective function is written as,

$$\underset{\mathbf{W},\mathbf{X},\mathbf{D}}{\arg\min} = \left\| \begin{pmatrix} \mathbf{Y} \\ \sqrt{\alpha} \cdot \mathbf{H} \\ \sqrt{\beta} \cdot \mathbf{Q} \end{pmatrix} - \begin{pmatrix} \mathbf{D} \\ \sqrt{\alpha} \cdot \mathbf{W} \\ \sqrt{\beta} \cdot \mathbf{A} \end{pmatrix} \cdot \mathbf{X} \right\|_F^2 \tag{5}$$

$$\text{s.t.} \quad \|\mathrm{x}_i\|_0 < T_0.$$

The information section Q about atom labels is added in each training sample. The transition matrix A maps the sparse code matrix X into atom label space which is recognised as more discriminative space. Each atom is associated with one row of Q at the initialisation stage. For a training sample y_i, the corresponding column q_i is 1 for all the rows representing atoms from the same class as this training sample. In the K-SVD training process, the samples are forced to be decomposed using atoms which are originally from the same class as the training samples. Similar to D-KSVD, the jointly trained classifier $\mathbf{W}$ is later used as a classifier on sparse codes.

Fig. 2 shows a pair of gait signals for one cycle from both control and pathological subjects. They are sparsely encoded using a dictionary trained by D-KSVD. Each signal is decomposed into a main component (MC) and several variation components (VC1, VC2,...). The number of training samples is much larger than the number of atoms in the trained dictionary. Hence, in the training process, a limited number of main components and variation components are used to represent a much larger number of samples. Main components describing the prototype signal for control subjects are prevalent in the sub-dictionary trained with control gait signals. However, these components are rare in the sub-dictionary trained with pathological signals. This makes both the reconstruction errors and sparse codes discriminative between pathological and control signals. The main components largely determine to which class the gait signals should be assigned. The results shown in Fig. 2 are also easier for clinicians to judge and evaluate the diagnostic results; a common requirement for application in clinical gait analysis [14].

TABLE I
NUMBER OF SUBJECTS IN THE DATASET

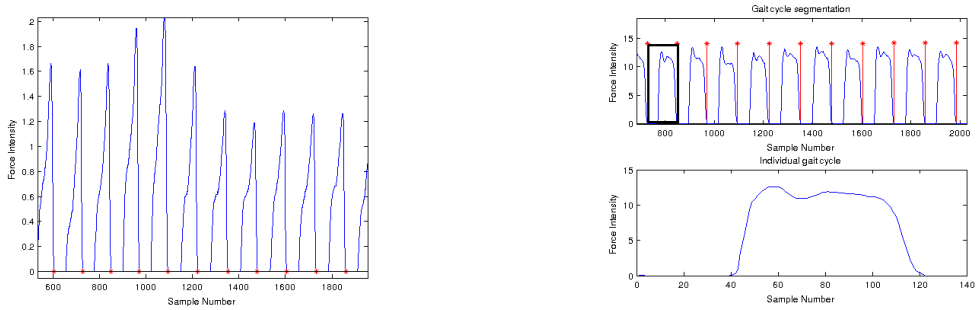
Group	Number of Patients	Number of Controls
Ga [20]	25	16
Ju [21]	29	25
Si [22]	34	28

IV. EXPERIMENTAL DATA

Experiments were conducted on natural gait data sets. The data set contains both pathological and control subjects of matched ages. Pre-processing procedures were undertaken to segment the data into gait cycles for each subject.

A. Gait data measured from subjects

The data set records vertical GRF when subjects walked at self-selected speeds. The data set contains vertical force intensity signals from subjects with Parkinson's disease and controls from physionet.org. For each subject, eight force sensitive sensors were inserted beneath both left and right feet. Force intensity for each individual sensor as well as the total force intensity for each limb are available in this data set. All the information included in this data set is from three previous studies using the same measurement equipment. The number of subjects from each study is shown in Table I. Among all the selected 157 subjects, 94 subjects are male while 63 subjects are female. The subjects with Parkinson's disease are of matched age with control subjects.



(a) (b)

Fig. 3. Gait cycle segmentation using toe-off events

B. Pre-processing on gait data

All the force intensity waveforms were sampled at 100 samples per second for around 2 minutes. Each of the long term signals in this data set involves several gait cycles. Among all 8 gait signals measured from different positions of feet, force information on heels and toes is usually utilised to separate each gait cycle from the whole waveform in terms of the toe-off and heel-strike events. In this paper, the toe-off events were utilised to separate gait signals into multiple cycles. By tracking the moment when the force intensity at the toe part decreases to zero ("incoming zero"), the toe-off event can be identified to differentiate two cycles as shown in Fig. 3 (a) and (b). In Fig. 3 (a), each incoming zero has been marked by an asterisk. The upper plot in Fig. 3 (b) shows a truncated gait signal involving several cycles. Cycles of this gait signal are separated by red lines drawn from the asterisk-marked points in Fig. 3 (a). The lower figure in Fig. 3 (b) presents a separated cycle which is the first complete cycle enclosed by a black frame. Segmented cycles are data samples in this study.

V. Experimental Design

Features were extracted from vertical GRF data described in Section IV-A from heels, toes as well as the entire foot. A gait sample which describes one gait cycle may consists of features derived from different parts of the feet. Samples were classified using sparse representation-based classification algorithms including SRC with K-SVD trained dictionaries (SRC†), D-KSVD and LC-KSVD. The resultant classification rates were also compared to those obtained using support vector machine (SVM) on the same features.

A. Feature Extraction

Because the durations of gait cycles vary within and across subjects, features were extracted using linear interpolation to align them to the same range. Thus the features incorporated in the classifiers are of identical length. Gait cycles were normalised against time and weight of subjects individually so that force values sampled at specific time points are aligned to the correspondence percentage of the whole cycle. This also enabled comparisons between subjects with different weights. In this paper, features combine interpolated vertical GRF beneath heels & toes of both limbs (LRHT) and the entire foot of a single limb (VGRF). Each feature sample extracted from a gait cycle is of the same dimension.

Fourier transformation (FT) was also used to remove the inter-cycle variations and extract discriminative features. Furthermore, by selecting only the significant coefficients provided by FT, the dimension of the gait data was reduced. FT was conducted on GRF beneath the entire foot of a single limb (VGRF_{FT}). We chose 25 real and 25 imaginary FT coefficients because they preserved over 99% of the power of the original signal.

In classification algorithms using sparse representation, sparse features were further extracted. Above-mentioned features from test samples were further projected onto the trained dictionaries. The resultant sparse codes of feature samples became new feature samples which were used with either the dictionary or the jointly trained classifier to make a decision of the class labels.

B. Experiments using SVM

In the experiments using SVM, the open-source support vector machine library, "libsvm" [23] was employed. Parameters of SVM with a specific kernel were adjusted, and using each of the features mentioned above, to achieve the best performance. In this paper, we tested the linear and radial basis function kernels. For each kernel and feature, the SVM model was trained with relevant parameters varying from 2^{-10} to 2^{11}. The best classification rate with the optimised parameters was chosen as the final result for that kernel-feature. The SVM models were firstly trained using feature samples and label information of each sample and then tested by classifying unseen samples.

10-fold cross-validation was taken as the validation method to evaluate the generalisation performance (classification rate). Feature samples from 157 subjects were grouped into 10 folds. Each fold involves samples extracted from both patients and controls. Features samples extracted from controls were all marked as normal while those extracted from abnormal persons were marked as pathological. All samples collected from the same person were assigned to the same fold. Each

TABLE II
CLASSIFICATION OF FEATURES USING DIFFERENT SVM KERNEL FUNCTIONS

	Linear Kernel	Radial Basis Kernel
VGRF	64.97%	68.79%
VGRF_{FT}	65.61%	65.61%
LRHT	80.25%	81.53%

TABLE III
CONFUSION MATRIX FOR RADIAL BASIS KERNEL WITH RESPECT TO LRHT

	Number of Controls	Number of Patients
True	51	77
False	18	11

fold thus held 15 to 17 subjects. At each cross-validation iteration, feature samples in one fold were reserved as test fold while the other folds were used to train the classifiers. The iteration continued until all folds were once used as the test fold. The cross-validation performance was tested on all the three features using both the proposed algorithms and SVM.

A subject can be recognised as a gait-intact subject when majority of feature cycles extracted from this gait signal are classified as normal. We aim to apply the previous verified classifier and features to our current data set and compare the classification results with those using sparse representation based algorithms.

Table II shows the best classification rate with SVM on different features using two kernel functions with optimised parameters. The radial basis kernel recommended by [23] works best across all the features with the highest classification rate of 81.53% for LRHT. The confusion matrix for this result is shown in Table III. The cost parameter c and degree value in the kernel function are 2^5 and 2^{-4} respectively. According to Table II, LRHT containing force information from heels and toes of both limbs shows the highest discrimination among control and pathological gait.

C. Experiments using SRC†

Performance of the joint dictionary and individual dictionaries was also tested on the above-mentioned features. For the experiments using SRC†, features from pathological and control subjects were used to train two dictionaries separately. The test samples stacked in column vectors were sparsely encoded either by two dictionaries separately or by a joint dictionary. Experiments on individual dictionaries (SRC_I) and the joint dictionary (SRC_J) were conducted for various levels of sparsity. Classification rates using one joint dictionary and two individual dictionaries are compared in Section V-E. The dictionary sizes for the feature VGRF, VGRF_{FT} and LRHT are 350, 75, 900 respectively.

As shown in Table. IV, with a joint dictionary, the highest classification rate of 76.64% was achieved by SRC_I with feature LRHT. For sparse coding using individual dictionaries, a relatively lower classification rate of 75.16% was reached, which is rather close the that of SRC_J. and training samples. However, SRC_J with a joint dictionary is more robust to the

TABLE IV
EXPERIMENTAL RESULTS FOR SRC†

Features	Classification Rate		Sparsity	
	SRC_I	SRC_J	SRC_I	SRC_J
VGRF	58.60%	63.69%	5	25
VGRF_{FT}	59.24%	60.51%	25	15
LRHT	75.16%	76.64%	20	10

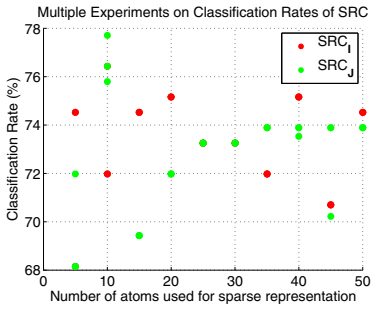


Fig. 4. Experiment results with sparsity from 5 to 50

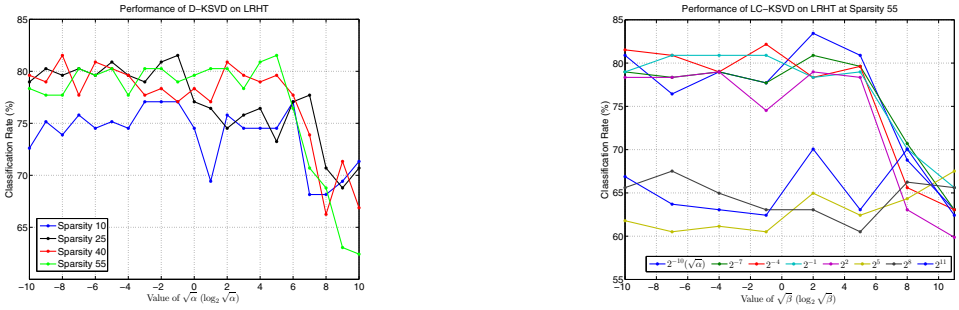
changes of sparsity as is shown in Fig. 4. Both SRC_I and SRC_J were conducted triple times on LRHT. The classification results against each sparsity value are shown in Fig. 4. With the sparsity increasing, the classification rates of SRC_J trends to converge to 74% while those for SRC_I are still fluctuating.

D. Experiments using D-KSVD and LC-KSVD

The parameter tuning procedure was also applied on D-KSVD and LC-KSVD. The classification rates obtained with the optimised parameters were later used to evaluate features and algorithms. The training and test data groups for LC-KSVD and D-KSVD follow the same procedure as those for SRC† in Section V-C. Control and pathological samples were jointly used to train a unique dictionary with potential sub-dictionaries embedded. Two classifiers were trained with these two algorithms jointly. The test gait samples were firstly sparsely encoded using a dictionary and the sparse codes were then classified using the linear classifier trained together with the dictionaries.

Compared to SRC†, the additional parameter to be adjusted in D-KSVD is α which controls the contribution of classification errors in the objective function. We verified values of $\sqrt{\alpha}$ in (4) from 2^{-10} to 2^{11} as for SVM and SRC†. LC-KSVD further adds the parameter β that adjusts the penalty of using atoms with different labels than the input sample. This parameter, $\sqrt{\beta}$ in (5), was also tested with the value varying from 2^{-10} to 2^{11}. Both α and β were tested with the sparsity values 10, 25, 40, 55. The best performance for each sparsity level is shown in Table V and Table VI.

The classification rates using D-KSVD and LC-KSVD on the LRHT features based on gait cycles are shown in Fig. 5. As shown in Fig. 5 (a), when sparsity reaches a certain threshold, allocating more atoms for reconstructing the input signal did not positively impact the classification rate but the simulation speed of the experiments deteriorates. For both D-KSVD and LC-KSVD with a constant sparsity, a large-valued parameter



(a) (b)

Fig. 5. Experiment Results for D-KSVD and LC-KSVD with Varying Parameters

TABLE V
EXPERIMENTAL RESULTS FOR D-KSVD ON LRHT

Sparsity	Classification Rate	Parameters $\sqrt{\alpha}$
10	77.07%	2^{-1}
25	81.53%	2^{-1}
40	80.89%	2^{-4}
55	81.53%	2^{5}

TABLE VI
EXPERIMENTAL RESULTS FOR LC-KSVD ON LRHT

Sparsity	Classification Rate	Parameters	
		$\sqrt{\alpha}$	$\sqrt{\beta}$
10	78.98%	2^{-1}	2^{-4}
25	81.53%	2^{-7}	2^{-7}
40	81.53%	2^{-7}	2^{-7}
55	83.44%	2^{-10}	2^{2}

α or β will overfit the classification model. The classification rate decreased when the value of $\sqrt{\alpha}$ exceeded 2^6. In these cases, classification errors dominate the reconstruction error. The added label information will dominate the input signals, which leads to a high training classification rate while the general performance becomes poor.

E. Discussion

Among all the classification algorithms, LC-KSVD reached the maximum classification rate of 83.44%. SVM and D-KSVD converged at 81.53% with SRC† being at 76.64%. All the highest rates were achieved on LRHT. Although the highest classification rates of SVM and sparse representation-based algorithms are quiet close (only three more persons were correctly classified for the case with 83.44%), the latter methods are more descriptive and interpretable.

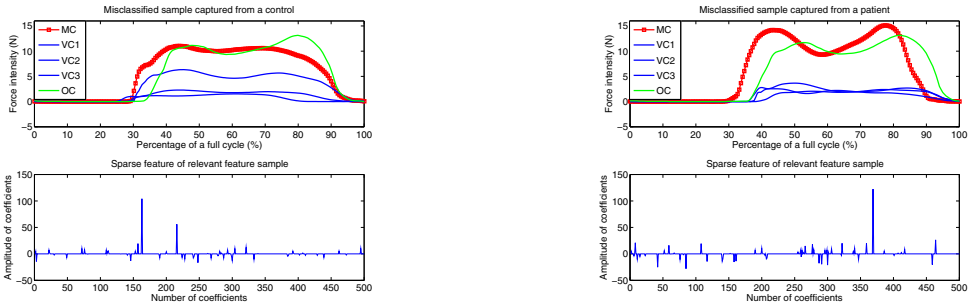
According to results quoted above, LRHT turns out to be the most discriminative feature between pathological and control subjects. GRF changes beneath heels and toes are generally used to differentiate gait phases during walking [2]. LRHT contains interpolated GRF cycles from heels and toes of both limbs, which provides not only the durations and orders of gait phases but also conditions of walking balance. Pathological subjects in this data set are in Stage 2 or over of the Hoehn and Yahr Stages of Parkinson's disease, wherein stage symptoms start to impact both limbs. Thus, LRHT provides extensive information about symptoms manifesting in gait.

In Fig. 6, the original VGRF cycle (OC) and main component (MC) are shown together with three variation components (VC) with the largest amplitudes. Both samples were projected onto the same dictionary trained by LC-KSVD. Main component of the misclassified pathological sample in Fig. 6 (b) gains features of a control sample when compared to MC in Fig. 2 (a). Misclassified control samples undergoes similar variation (Fig. 6 (a)). The manner in which samples are sparsely represented determines the results of the classification. Because OMP was used to decompose input samples, similarities between the samples and atoms dominate the representation process. In our paper, feature samples are extracted from quasi-periodic signals which are of high similarity, and this distorts the classification algorithms.

Due to correlations among gait cycles, it is necessary to improve the logic of determining class labels of subjects based on cycle labels. In future work we may explore data fusion techniques, so decisions on each gait cycle may be combined using some of the classifier combination techniques [24], [25]. Meanwhile, all gait cycles are currently treated as a single cycle in the sparse coding stage. In fact, feature samples extracted from the same person are supposed to share a similar main component. Constraints on the similarity of sparse codes for a certain subject may also assist in classifying subjects according to gait cycles. As the stage of Parkinson's disease increases, the gait deteriorates when compared to the normal. One would expect that the accuracy of detection increases with the stage of the disease. However, based on the data set we used in our experiments, we have not noticed a clear correlation between the accuracy of detection and the stage of the disease. This will be further investigated when we collect a large data set.

VI. CONCLUSION

In this paper, we model gait as sparsely represented signal based on learned dictionary to address the variability of gait data between and among persons. Dictionaries were designed to possess both representative and discriminative properties. Our representation forms the basis of features extracted and



(a) (b)

Fig. 6. Examples of misclassified samples from both pathological and control subjects

used in a classification scheme that detects pathological gaits of persons suffering Parkinson's disease. We provide rigorous experiments to evaluate and validate our proposed scheme. Results indicate that our proposed scheme outperforms comparable schemes reported in the literature.

In our continuing work we are exploring the construction of dictionaries based on sparse and low-rank approximation. This will allow us to incorporate constraints that combine sparseness, compactness, representation and discriminability. Such dictionaries hold promise to increase the detection accuracy of our method.

REFERENCES

[1] D. Lai, R. Begg, and M. Palaniswami, "Computational intelligence in gait research: A perspective on current applications and future challenges," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 687–702, 2009.

[2] M. W. Whittle, *Gait analysis:an introduction*, 4th ed. Philadelphia, PA : Butterworth-Heinemann Elsevier, 2007.

[3] T. Chau, "A review of analytical techniques for gait data. part 1: fuzzy, statistical and fractal methods," *Gait & Posture*, vol. 13, no. 1, pp. 49–66, 2001.

[4] C. M. Senanayake and S. M. N. A. Senanayske, "Computational intelligent gait-phase detection system to identify pathological gait," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 5, pp. 1173–1179, 2010.

[5] D. Lai, P. Levinger, R. Begg, W. Gilleard, and M. Palaniswami, "Automatic recognition of gait patterns exhibiting patellofemoral pain syndrome using a support vector machine approach," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 810–817, 2009.

[6] R. Begg, M. Palaniswami, and B. Owen, "Support vector machines for automated gait classification," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 5, pp. 828–838, 2005.

[7] S. R. Simon, "Quantification of human motion: gait analysis–benefits and limitations to its application to clinical problems," *Journal of Biomechanics*, vol. 37, no. 12, pp. 1869–1880, 2004.

[8] S. H. Holzreiter and M. E. Köhle, "Assessment of gait patterns using neural networks," *Journal of Biomechanics*, vol. 26, no. 6, pp. 645–651, 1993.

[9] M. Köhle and D. Merkl, "Analyzing human gait patterns for malfunction detection," in *Proceedings of the 2000 ACM symposium on Applied computing - Volume 1*, ser. SAC '00. New York, NY, USA: ACM, 2000, pp. 41–45.

[10] K. Deluzio and J. Astephen, "Biomechanical features of gait waveform data associated with knee osteoarthritis: An application of principal component analysis," *Gait & Posture*, vol. 25, no. 1, pp. 86–93, 2007.

[11] M. Yang, H. Zheng, H. Wang, S. McClean, J. Hall, and N. Harris, "A machine learning approach to assessing gait patterns for complex regional pain syndrome," *Medical Engineering & Physics*, vol. 34, no. 6, pp. 740–746, 2012.

[12] J. Kamruzzaman and R. Begg, "Support vector machines and other pattern recognition approaches to the diagnosis of cerebral palsy gait," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2479–2490, 2006.

[13] J. Bae and M. Tomizuka, "Gait phase analysis based on a hidden markov model," *Mechatronics*, vol. 21, no. 6, pp. 961–970, 2011.

[14] N. Ş. Köktaş, N. Yalabik, G. Yavuzer, and R. P. Duin, "A multi-classifier for grading knee osteoarthritis using gait analysis," *Pattern Recognition Letters*, vol. 31, no. 9, pp. 898–904, 2010.

[15] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.

[16] Q. Zhang and B. Li, "Discriminative K-SVD for dictionary learning in face recognition," in *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 2691–2698.

[17] Z. Jiang, Z. Lin, and L. Davis, "Learning a discriminative dictionary for sparse coding via label consistent K-SVD," in *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 1697–1704.

[18] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[19] Y. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *1993 Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, vol. 1, 1993, pp. 40–44.

[20] G. Yogev, N. Giladi, C. Peretz, S. Springer, E. S. Simon, and J. M. Hausdorff, "Dual tasking, gait rhythmicity, and parkinson's disease: Which aspects of gait are attention demanding?" *European Journal of Neuroscience*, vol. 22, no. 5, pp. 1248–1256, 2005.

[21] J. M. Hausdorff, J. Lowenthal, T. Herman, L. Gruendlinger, C. Peretz, and N. Giladi, "Rhythmic auditory stimulation modulates gait variability in parkinson's disease," *European Journal of Neuroscience*, vol. 26, no. 8, pp. 2369–2375, 2007.

[22] S. Frenkel-Toledo, N. Giladi, C. Peretz, T. Herman, L. Gruendlinger, and J. M. Hausdorff, "Treadmill walking as an external pacemaker to improve gait rhythm and stability in parkinson's disease," *Movement Disorders*, pp. 1109–1114, 2005.

[23] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 27:1–27:27, May 2011.

[24] J. Kittler, M. Hatef, R. P. Duin, and J. Matas, "On combining classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, March 1998.

[25] L. I. Kuncheva, "Switching between selection and fusion in combining classifiers: An experiment," *IEEE Transactions on Systems Man and Cybernetics, Part B: CYBERNETICS*, vol. 32, no. 2, pp. 146–156, April 2002.