Pain Level Recognition using Kinematics and Muscle Activity for Physical Rehabilitation in Chronic Pain

Temitayo A. Olugbade, Nadia Bianchi-Berthouze,
Nicolai Marquardt

UCL Interaction Centre
University College London
London, United Kingdom

Amanda C. Williams

Dept. Clinical Psychology
University College London
London, United Kingdom

Abstract—**People with chronic musculoskeletal pain would benefit from technology that provides run-time personalized feedback and help adjust their physical exercise plan. However, increased pain during physical exercise, or anxiety about anticipated pain increase, may lead to setback and intensified sensitivity to pain. Our study investigates the possibility of detecting pain levels from the quality of body movement during two functional physical exercises. By analyzing recordings of kinematics and muscle activity, our feature optimization algorithms and machine learning techniques can automatically discriminate between people with low level pain and high level pain and control participants while exercising. Best results were obtained from feature set optimization algorithms: 94% and 80% for the *full trunk flexion* and *sit-to-stand* movements respectively using Support Vector Machines. As depression can affect pain experience, we included participants' depression scores on a standard questionnaire and this improved discrimination between the control participants and the people with pain when Random Forests were used.**

Keywords—pain; automatic recognition; body movement; muscle activity; physical activity; depression

I. INTRODUCTION

Chronic musculoskeletal pain (CP) is a prevalent medical condition of persistent pain without tissue damage [1-3]. CP is a result of changes in the nervous system that lead to the amplification of pain. About 19% and 30% of adults in Europe and the United States, respectively, suffer from CP [1-2]. The cost to the national health services of the United Kingdom amounts to about £12 billion per annum [4]. The affective experiences in CP, including pain, fear of pain, and depression, lend to reduced mobility, and lowered quality of life [5-8].

Physical rehabilitation is critical in CP to improve mobility, muscle fitness and confidence in movement [9-11] – ideally as self-management rather than prolonged treatment [12]. Self-directed exercise in CP requires knowledge and confidence that many people lack. Tailoring physical rehabilitation to the dynamic psychological needs and fluctuating pain levels of the individual with CP improves adherence and efficacy [8, 13]. Tailoring to pain level is important as increased pain during physical activity increases sensitivity to pain and discourages movement [6, 14]. This also promotes movement behaviours that aim to protect from pain but undermine the efficacy of physical rehabilitation and can lead to further pain, anxiety and decreased fitness in the long term [8].

Studies of technology tailored to physical performance and/or level of engagement for self-managed physical rehabilitation [15-17] have focussed on physical progress and overlooked issues of pain levels or distress due to pain. In this paper, we investigate automatic recognition of levels of pain during everyday functional movements at the core of physical rehabilitation in CP. We focus on people with chronic low back musculoskeletal pain as one of the most prevalent forms of CP [1]. Automatic discrimination between levels of pain allows personalized support for exercise [13], building physical and psychological capabilities (e.g. confidence in moving) by reducing negative experiences, strain, and consequent setback.

In the remainder of this paper, we discuss literature on body movement as a modality to express and recognize pain experience and related behaviour. Next, we review previous work on the automatic recognition of pain and pain-related behaviour from body expressions. We then describe our approach and the corpus used, the extraction of kinematic and muscle activity features, and the relationship found between reported pain and depression scores in the datasets used. We present the evaluation of our system for the automatic recognition of pain levels for two functional activities – *sit-to-stand* and *full trunk flexion* – respectively. We conclude with a discussion of our findings and future directions.

II. BACKGROUND

Facial and vocal expressions dominate the literature on expressions of pain, but in chronic pain the body is an important expressive modality [18-19]. Aung et al. [20] showed that in CP, bodily behaviours associated with pain were more frequent than facial expressions and provided insight on how people experienced movement in relation to pain. In the following subsections, we discuss previous work on pain-related behaviour and its automatic detection.

A. Pain and Body Movement

People with CP (hereafter referred to as "PCP") adopt specific behaviours in response to pain or anxiety about pain to protect the body [18-19]. Typical pain behaviours include guarding, hesitation, bracing, and rubbing [18]. Pain behaviours, including the facial expressions of pain, also communicate pain, possibly to solicit support or understanding from others [21].

Analysis of protective behaviour has shown kinematic feature descriptions to differentiate pain groups. One such

description is the amount of movement, significantly different between PCPs and non-PCPs and among PCPs anticipating different levels of pain [22-23]. Muscle activity measured by electromyography (EMG) also differentiates between PCPs and non-PCPs. For example, previous studies [23] showed the absence of the flexion relaxation phenomenon (a natural relaxation in paraspinal muscle activity between 40 and 70 degrees of trunk flexion) in PCPs during full trunk flexion. Furthermore, there was less muscle activation in PCPs than non-PCPs during forward trunk flexion. It is important to note that in [23], no differences in muscle activation patterns were found between pain groups when not moving. These studies suggest the possibility of creating systems that automatically discriminate between PCPs and non-PCPs during physical activity.

B. Automatic Recognition of Pain using Body Movement

Most studies on automatic recognition of pain experience have focussed on facial expression across contexts [24-26], facilitated by recently available datasets (e.g. [27]). Attempts to recognize pain experience from body cues typically create recognition models separately for each exercise given.

An early study [28] for the discrimination between PCPs and non-PCPs in flexion exercises used the forward flexion/extension lumbar paraspinal muscle activation ratio, the mean lumbar paraspinal muscle activation, and the amount of flexion to achieve an accuracy of 0.86. More recently, Gioftsos and Grieve [29] developed an automatic recognition system for pain classes in the sit-to-stand-to-sit movement. The system discriminates between non-PCPs, PCPs, and non-PCPs simulating CP. They found displacements of knee, hip, and trunk with measures of the centre of foot pressures separated the three classes with an accuracy of 0.86. These studies do not attempt to discriminate between levels of pain.

Dickey et al. [30] examined pain recognition in PCPs performing movements including flexions and sit-to-stand. Features of intravertebral deformation and intervertebral motion were obtained by tracking the movement of pedicle screws inserted into the spine during surgery. They predicted pain intensity on a 0-10 scale with an accuracy of 0.75. Unfortunately, this method is not feasible outside this surgical population. Using the Emo-Pain dataset [20], we recently investigated the detection of guarding behaviour due to anticipated pain [31]. Physiotherapists labelled the guarding behaviour as ground truth. Presence or absence of guarding in sit-to-stand and balance exercises were recognized with 0.8 and 0.77 accuracies respectively using angles and energies of anatomical joints and mean muscle activation. Olugbade et al. [32] used the same dataset to explore the possibility of discriminating between non-PCPs, PCPs with low level pain, and PCPs with high level pain in forward trunk flexion (i.e. reaching forward as much as possible, simulating reaching a book on a shelf). They reached an accuracy of 0.86 using a model that relied on measures of upper arm steadiness, trunk flexion, neck displacement, and muscle activation change point in the upper trapezius and lumbar paraspinal muscle groups.

In contrast to rehabilitation sessions in hospital settings, chronic pain rehabilitation is integrated in everyday activity where comprehensive sensing is not possible [13, 33-34]. Thus, it is important to investigate feature transfer across anxiety-inducing functional movements to enable the design of pain-aware wearable technology design. In this paper, we therefore investigate how the features discriminatory of pain [32], and other features derived from tracking the same anatomical segments, may transfer to two other movement types, *sit-to-stand* and *full trunk flexion* (i.e. reaching down to touch the toes). Given the prevalence of depression in chronic pain [7, 35-36], we added depression scores to further improve discrimination between pain levels in PCPs.

III. METHODS

In this section, we describe the corpus used, the pain classes we aim to automatically recognize, and the approach we used in modelling these classes.

A. Pain Corpus

The datasets used for this study are part of the Emo-Pain corpus [20], a multimodal corpus consisting of motion capture, EMG, video, and audio data from 23 PCPs and 30 non-PCPs. The participants were recorded while performing five functional exercises representing everyday movements that PCPs find challenging and often avoid for fear of increasing pain. The exercises were organized in two routines at different difficulty levels. In the normal level routine, the exercises were performed according to the preference of the subject (e.g. balancing on the preferred leg). In the difficult level, subjects were given specific instructions (e.g. "try to avoid using hand support"). We focus on motion capture and EMG data for *sit-to-stand* and *full trunk flexion* exercises.

The Emo-Pain corpus is labelled with various sets of ratings. In our study, we use pain scores before and after exercises and depression scores. Before beginning any of the activities, each subject completed the Hospital Anxiety and Depression Scale (HADS). The depression subscale (between 0 and 21) of the HADS (HADS-D) was designed to measure anhedonic depression in non-psychiatric populations [37]. The mean and standard deviation of HADS-D scores for the Emo-Pain corpus were 8.71 and 4.85 respectively. In addition, each subject was asked on arrival to indicate their usual level of pain, from 0 (no pain at all) to 10 (extreme pain). After each exercise, the subject was asked to score pain on the same scale. The mean (standard deviation) of pain on arrival, after *sit-to-stand* exercises, and *full trunk flexion* exercise were 3.74 (2.95), 4.68 (3.31), and 4.24 (3.24) respectively. Our aim was to predict the level of pain experienced during the exercises.

B. Motion Capture and Electromyography Data

The motion capture data (gathered with *IGI Animazoo* motion capture system) in the Emo-Pain corpus comprises frames of Euler angles for eighteen anatomical segments. These were used to derive three-dimensional positions of twenty-six anatomical nodes, which are shown in Fig. 1-left. The EMG data (gathered with *BTS FreeEMG* surface EMG kit) consists of four channels of muscle activity data recorded bilaterally from the upper trapezius muscle group (labelled 3 and 4 in Fig. 1-right) and also bilaterally from the lumbar

paraspinal muscle group (labelled 1 and 2 in Fig. 1-right). For each channel, the rectified upper envelope was taken as the muscle activity signal.

In the Emo-Pain recording, each participant was asked to perform three repetitions of the *sit-to-stand* for each difficulty level. Since the *full trunk flexion* is a more tasking movement for PCPs, participants performed only one instance of this exercise per routine. Unfortunately, there were instances where either the data was noisy or the subject did not report the level of pain experience. We could not use those instances in this study and this resulted in 104 instances of *sit-to-stand* and 18 instances of *full flexion*. Of these instances, only for 98 instances of sit-to-stand (from fourteen PCPs and eight non-PCPs) and 16 instances of full flexion (from twelve PCPs and four non-PCPs) were there corresponding depression scores.

C. Pain Classes

Three pain classes were considered in this paper: non-PCP, low level pain PCP (who reported pain of less than 5), and high level pain PCP (with pain of 5 or more). The reason for aggregating the 11-point pain scale to the two levels is because subjects have difficulty reliably discriminating between non-extreme adjacent points on a scale [38]. It was also important for us to distinguish between PCPs with no pain and non-PCPs as these two groups are fundamentally different. Non-PCPs do not suffer persistent pain and the psychological states that come with it such as fear of movement. For the *sit-to-stand* exercise, there were 33 instances of non-PCP, 30 of low level pain PCP, and 35 of high level pain PCP. For the *full trunk flexion* exercise, there were 4 instances of non-PCP, 7 for low level pain PCP, and 5 for high level pain PCP.

D. Classification Algorithms

We used Random Forest (RF) and Support Vector Machine (SVM) algorithms in this study as they have been shown to be powerful in affect recognition [32, 39-41]. An SVM [42] works by finding the hyperplane that separates observations into two classes with a maximum margin between the two, generally a soft margin. A soft margin uses a regularization parameter that penalizes misclassification. We considered polynomial, radial-basis function, and hyperbolic-tangent kernels. For pain recognition, we used a two-level SVM hierarchical architecture. The SVM at the first level of the hierarchy (*SVM1*), discriminates between PCPs and non-PCPs. On encountering a PCP observation, the second SVM (*SVM2*), is triggered to classify it as either low level pain or high level pain. We expected that the hierarchical architecture would lead to better results as the feature subset optimal for differentiating healthy participants from those with pain may differ from those discriminating between pain levels. For each of the two SVMs, a grid search was done to find the optimal kernel and regularization parameter. For the *full trunk flexion* exercise, both SVMs were optimal as linear SVMs with a regularization parameter of 10 for SVM1 and 0.01 for SVM2. In the *sit-to-stand*, the optimal SVM1 is linear with regularization parameter of 1, while SVM2 is optimal with a radial basis function kernel with width of 3.01 and regularization of 10.

The RF is an ensemble of decision trees [43]. Each tree is built by random sampling of the training data and a random
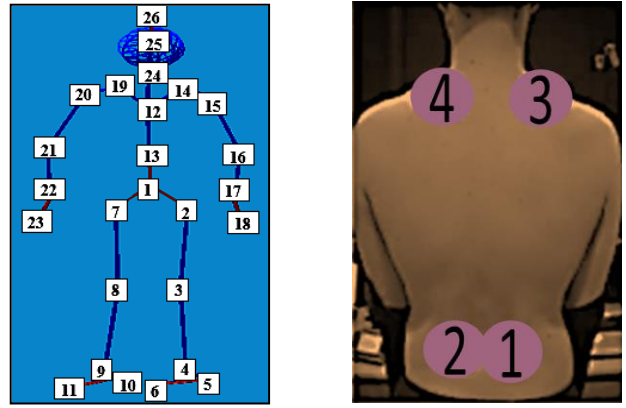

Fig. 1. Labelled stick figure (left) shows the 26 tracked anatomical nodes and the labelled torso posterior (right) shows the 4 tracked muscle groups

subset of the features is used to expand each node of the tree. For a classification task, each tree makes a prediction and a vote is taken over these. The majority vote is taken as the prediction of the RF. We did not reproduce the hierarchical architecture used in the SVM here because we relied on the supplementary feature selection implicit to the RF. In our study, all of the training data was used to build each tree of the RF. The optimal number of trees and number of features to expand each node were computed from a grid search over a number of values. For the *full trunk flexion* exercise, the optimal RF has 500 trees and selects one feature at random to split each node. For the *sit-to-stand* exercise, the optimal RF has 1000 trees and it uses a subset of features the size of its feature set selected at random to split each node.

In both of the SVM and RF models used in our study, the features were normalized to zero mean and unit variance as is standard in reducing the effect of variations in feature variances on distance-based inductions algorithms (like SVM).

E. Kinematic and EMG Feature Set Optimization

A variation of the Branch and Bound algorithm (B&B) proposed in [44] was used for the optimization of the feature selection for the two classification algorithms used. The B&B is a breadth-first tree search using backward elimination. The root node of the tree is the original feature set and the successors of any node are derived by eliminating one feature at a time from the feature subset of that node. A node is expanded only if the accuracy of the classification model that uses the feature subset represented by that node is greater than a set bound. In the variation used here, the bound at any node is the value of its parent node with the bound at the root node set to 0. Although the B&B is a computationally intensive algorithm, it was manageable in our study as the original feature sets we used are small-sized. To reduce the complexity, we pruned each search tree at a fraction of its height to guide the B&B to deeper search along only the most promising paths.

IV. FEATURE SELECTION

A. Feature Set for the Full Trunk Flexion Exercise

In the recognition of the three pain classes for the forward trunk flexion in [32], five kinematic and eight muscle activity features had proven discriminatory. Given the similarity

between the forward trunk flexion and *full trunk flexion*, we wanted to investigate if the feature set used in [32] would be effective in the *full trunk flexion*. Both movements involve flexion of the trunk and extension of the arms. It should be noted that the two similar movement types have differences. For example, the *full trunk flexion* requires larger flexion of the trunk and the extension of the arms are in a different direction. The *full trunk flexion* is also different in that there is an expected relaxation of the lower back muscles at the fully-flexed position, which will not be seen in the forward trunk flexion. As mentioned earlier, the merit of using the same feature set for both exercise types is the minimization of the number of body features that need to be tracked. This is especially important when portable motion sensing is a necessary consideration as has been found for chronic pain rehabilitation [13, 33-34].

TABLE I. FULL TRUNK FLEXION: KINEMATIC AND MUSCLE ACTIVITY FEATURES. 'x' INDICATES THE FEATURES THAT ARE PART OF THE OPTIMIZED SET FOR THE MODEL INDICATED BY THE COLUMN

Exercise: Full Trunk Flexion					
ID			SVM1	SVM2	RF
Kinematic Features					
1	Upper arm displacement profile	number of peaks			
2		range of distance (in time) between peaks normalized to the length of the profile	x		
3		mean amplitude of peaks normalized to maximum amplitude of the profile			
4	Neck displace-ment profile	range	x		x
5	Trunk flexion	range	x		
Muscle Activity Features					
6	Time of the change point normalized to the duration of the exercise	right and left lower back muscles		x	x
		right and left upper back muscles	x		x x
7	Amplitude difference at the change point normalized to the maximum amplitude of the signal	right and left lower back muscles	x	x x	x x
		right and left upper back muscles	x x	x	x

Thus, the kinematic features used in [32] were extracted as listed in Table I. Each feature considered was normalized as necessary to account for the variation in anatomical structure of different subjects [32]. This is in addition to the normalization done for the sake of the induction algorithms mentioned earlier. Similarly, muscle activity features were extracted as in [32] based on the change point in muscle activation from high activity to low activity phase expected in a healthy subject completing the trunk flexion [23]. For each of the four muscle activity channels recorded (Fig. 1-right), we extracted the time when a change point occurred (from high activity phase to low activity phase) normalized by the duration of the exercise and the amplitude difference about the change point normalized to the maximum amplitude (features are specified in Table I).

B. Feature Set for the Sit-to-Stand Exercise

For the *sit-to-stand* exercise, in addition to the features inspired by the review in Section II, we explored features of movement behaviours with an experienced physiotherapist to better understand which features they might use. We extracted nine kinematic features (listed in Table II).

As expected from [45], visual exploration of muscle activation patterns across the three pain classes showed that most of the non-PCPs exhibited a common pattern as in Fig. 2-left. In this pattern, there is high muscle activity on starting the *sit-to-stand* exercise and this continues up to the point when weight has been transferred from the hips to the feet. The PCPs, on the other hand, showed a variety of patterns possibly due to the variety of strategies people use for fear of pain. Most of these muscle activity patterns deviated from the ones shown by non-PCP. Fig. 2-middle, for example, shows Subject P014 avoiding the use of her right side in the *sit-to-stand* movement. The plots of the subject's lower back muscle activities in Fig. 2-right show higher and prolonged activity (even at the end of the exercise) in her left lower back muscles.

TABLE II. SIT-TO-STAND: KINEMATIC AND MUSCLE ACTIVITY FEATURES. 'x' INDICATES THE FEATURES THAT ARE PART OF THE OPTIMIZED SET FOR THE MODEL INDICATED BY THE COLUMN.

Exercise: Sit-to-Stand				
ID		SVM1	SVM2	RF
Kinematic Features				
1	Duration of the exercise	x		x
2	Vertical velocity of the pelvis			
3	Vertical displacement of the neck	x	x	x
4	Amount of trunk flexion relative to ground		x	x
5	Amount of trunk flexion, relative to the body, before the buttocks lift from the seat	x	x	
6	Left & right hip angles at the point of buttocks lift	x		x
7	Left & right knee angles at the point of buttocks lift		x	x x
Muscle Activity Features				
8	Time of the change point normalized to the duration of the exercise — right and left lower back muscles			
	right and left upper back muscles		x	x
9	Amplitude range of the signal — right and left lower back muscles		x	x
	right and left upper back muscles	x x	x x	

For each of the four EMG channels, two features were extracted. One of the features is the ratio of the muscle activity change point to the duration of the whole movement. Like in the *full trunk flexion* exercise, we use the term "change point" to refer to the point when muscle activity changes from high activity phase to low. This feature was computed using a method derived from [46] where two equal-sized windows are slid through each signal and the change point is taken to be midway between the two windows where the difference in mean amplitude is the greatest (each window is 12 frames long with one frame between). The second feature extracted is the amplitude range of the muscle activity signal. An overview of these features is shown in Table II.

C. Depression score as a feature

To explore any association between pain and depression in our datasets, we correlated HADS-D scores with reported pain. There was no significant association (Pearson Correlation = -.268, p>0.1) before exercise, but significant
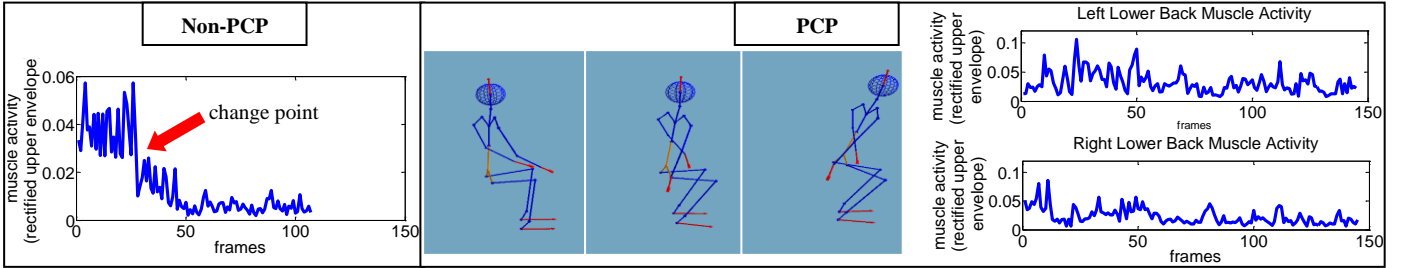
Fig. 2. (**Left**) Typical muscle activity in a non-PCP performing a sit-to-stand movement with the red arrow pointing to the change point from high activity to low. (**Middle**) Subject P014 is avoiding the use of her right side in standing up, and (**Right**) her left and right lower back muscles deviate from the expected activity pattern seen for non-PCPs and with less activity on the right.

correlations were found at the end of each of the exercise types: *full trunk flexion* .657 (p<.01) and .719 (p<.01) for the normal and difficult levels respectively; for *the sit-to-stand* Pearson Correlation was .688 (p<.01) for the difficult level.

TABLE III. ACCURACY & F1 SCORES FOR OUR CLASSIFICATION MODELS: BF = BODY FEATURES, O = FEATURE OPTIMIZATION, D = DEPRESSION SCORE

Full Trunk Flexion - SVM Model				
	BF	*BF+D*	*O(BF)*	*O(BF)+D*
Accuracy	0.56	0.50	**0.94**	**0.94**
F1 Control	0.55	0.55	**0.86**	**0.86**
F1 Low level pain	0.44	0.40	**1**	**1**
F1 High level pain	0.67	0.55	**0.91**	**0.91**
Average F1	0.55	0.50	**0.92**	**0.92**
Full Trunk Flexion - RF Model				
	BF	*BF+D*	*O(BF)*	*O(BF)+D*
Accuracy	0.75	0.81	**0.88**	**0.88**
F1 Control	0.60	**0.75**	**0.75**	**0.75**
F1 Low level pain	0.73	0.86	**0.92**	**0.92**
F1 High level pain	**0.91**	0.80	**0.91**	**0.91**
Average F1	0.75	0.80	**0.86**	**0.86**
Sit-to-Stand - SVM Model				
	BF	*BF+D*	*O(BF)*	*O(BF)+D*
Accuracy	0.64	0.64	**0.80**	0.76
F1 Control	**0.78**	**0.78**	0.75	0.75
F1 Low level pain	0.48	0.47	**0.72**	0.65
F1 High level pain	0.67	0.67	**0.90**	0.84
Average F1	0.64	0.64	**0.79**	0.75
Sit-to-Stand - RF Model				
	BF	*BF+D*	*O(BF)*	*O(BF)+D*
Accuracy	0.65	0.67	0.59	**0.69**
F1 Control	0.67	0.79	0.53	**0.83**
F1 Low level pain	**0.58**	0.55	0.52	0.56
F1 High level pain	**0.71**	0.66	0.73	0.67
Average F1	0.65	0.67	0.59	**0.69**

Given these correlations and the existing research evidence, we included depression scores as a feature for pain recognition together with the kinematic and EMG features. However, in the case of the two-level SVM architecture, depression was included as a feature only for SVM2 because we expect depression to affect pain level in PCPs but not in non-PCPs. It should also be noted that we used the depression scores recorded through standard questionnaire before the exercise sessions. This is feasible as people with CP routinely keep track of their mood on diary or smartphone applications. However, recent work shows the possibility to automatically infer depression level from behaviour over longer observation [47-49]. In the future, a person may be tracked over long periods allowing updated depression scores to be periodically available to our system.

V. RESULTS

We discuss the performances of the pain classification all based on leave-one-subject-out cross-validation (which is standard when dealing with human subjects). Table III shows the accuracy and F1 scores for the three pain classes for the two-level SVM architectures (*SVM1* and *SVM2*) and RF models of the *full trunk flexion* and *sit-to-stand* with and without feature set optimization, and with and without depression score as feature.

TABLE IV. FULL FLEXION: CONFUSION MATRICES: BF = BODY FEATURES, O = FEATURE OPTIMIZATION, D = DEPRESSION SCORE

		RF Models: O(BF)		
		Control	*Low level pain*	*High level pain*
Ground Truth	*Control*	3 (75%)	0 (0%)	1 (25%)
	Low level pain	**3 (42.9%)**	4 (57.1%)	0 (0%)
	High level pain	0 (0%)	0 (0%)	5 (100%)
		RF Model: O(BF) + D		
		Control	*Low level pain*	*High level pain*
Ground Truth	*Control*	3 (75%)	0 (0%)	1 (25%)
	Low level pain	**1 (14.3%)**	6 (85.7%)	0 (0%)
	High level pain	0 (0%)	1 (20%)	4 (80%)

A. Results for of Pain Classification in Full Trunk Flexion

In the *full trunk flexion* exercise, optimization with only body movement features improves pain recognition remarkably (see Table III) from accuracy of 0.56 to a classification of 0.94 for the SVM. Similarly, optimized body movement features give a considerable improvement (see Table III) in pain recognition for the RF from an accuracy of 0.75 without optimization to 0.88. By adding depression as a further feature, we observe an improvement in accuracy in RF model (with the un-optimized feature set) reaching accuracy of 0.81. The improvement seems to be the result of better discrimination of the low level pain group as can be seen in the confusion matrices in Table IV. On the other hand, we observe a slightly reduction in performances when adding depression as a feature for the SVM with the un-optimized feature set. There was no change in performance when depression was added for both the RF and SVM models using the corresponding optimized feature sets. The RF does better than the SVM using un-optimized feature sets with and without depression as additional feature, while the SVM performs better using optimized feature sets. We observed that the high level pain class was the class most consistently

recalled in both the SVM and RF using the optimized feature sets. This follows intuition that people experiencing low level pain and in the absence of fear of pain may behave as healthy participants – making it more difficult to discriminate between these two classes. In fact, we may not want to differentiate PCPs that behave like healthy persons because this may mean that they have learned to cope with their conditions and they are able to function in everyday life, which is the main aim of physical rehabilitation.

TABLE V. Sit-to-Stand: Confusion Matrices:
BF = Body Features, O = Feature Optimization, D = Depression Score

		Control	Low level pain	High level pain
	RF Models: BF			
		Control	*Low level pain*	*High level pain*
Ground Truth	*Control*	**22 (66.7**%)	7 (21.2%)	4 (12.1%)
	Low level pain	8 (26.7%)	19 (63.3%)	3 (10%)
	High level pain	3 (8.6%)	9 (25.7%)	23 (65.7%)
	RF Model: BF + D			
		Control	*Low level pain*	*High level pain*
Ground Truth	*Control*	**30 (90.9**%)	3 (9.1%)	0 (0%)
	Low level pain	9 (30%)	17 (56.7%)	4 (13.3%)
	High level pain	4 (11.4%)	12 (34.3%)	19 (52.3%)
	RF Models: O(BF)			
		Control	*Low level pain*	*High level pain*
Ground Truth	*Control*	*17 (51.5*%)	12 (36.3%)	4 (12.1%)
	Low level pain	11 (36.7%)	18 (60%)	1 (3.3%)
	High level pain	3 (%)	9 (%)	23 (%)
	RF Model: O(BF) + D			
		Control	*Low level pain*	*High level pain*
Ground Truth	*Control*	*31 (93.9*%)	2 (6.1%)	0 (0%)
	Low level pain	8 (26.7%)	17 (56.7%)	5 (16.7%)
	High level pain	3 (8.6%)	12 (34.3%)	20 (57.1%)

B. Result of Pain Classification in Sit-to-Stand

In the *sit-to-stand* exercise, the optimization of the body features improves accuracy (see Table III) from 0.64 to 0.80 for the SVM. However, using the optimized feature set for the RF unexpectedly reduces accuracy from 0.65 to 0.59 for the RF. This decrease in accuracy is likely due to the fact that the optimized feature set used were obtained using a fuller dataset (with the 104 instances) than the one (with the 98 instances) used in the cross-validation reported. As mentioned earlier, we had to drop instances where the depression score of the participant was not available. This was necessary to allow comparison of the models that included depression as a feature with the ones that did not. Similar to the *full flexion*, adding depression as a feature slightly improves accuracy in the RF from 0.65 to 0.67 using the un-optimized feature set. It also leads to improvement from 0.59 to 0.69 using the optimized feature set. These improvements seem to be as a result of better precision and recall of the control class. The confusion matrices in Table V, in fact, show at least 24.2% increase in the recognition accuracy of the control group. Again, however, for the SVM, there is a decrease in performance when depression was included with the un-optimized body feature set. This decrease becomes remarkable when the optimized feature set was used.

VI. Discussion and Conclusion

This paper investigated automatic classification of non-PCPs and pain levels in PCPs in *full trunk flexion* and *sit-to-*

stand exercises using kinematic and EMG features. We achieved near-perfect classification in the *full trunk flexion,* with only one instance misclassified. Pain classification in *full trunk flexion* may be easier because it requires lower back movement avoided by many PCPs. On the other hand, *sit-to-stand* cannot be avoided in normal life, so PCPs adopt varied protective strategies making pain recognition a harder task [31]. Still, our recognition models in the *sit-to-stand* performed about as well as models in previous studies [29] in distinguishing between PCPs and non-PCPs. In addition, with respect to previous literature [28-29], we were also able to discriminate well above chance level between two levels of pain in PCPs.

In addition to the movement features, we investigated the use of depression scores for the classification. In the best case for the SVM, depression had no effect on pain recognition in both *full trunk flexion* and *sit-to-stand*. Having more training instances might allow for better classification of anomalies such as high level pain with relatively low depression scores. In addition, an alternative approach to account for depression is selective use of depression scores (low depression scores more informative) in the modelling. However, when using RF, the classification improved with depression in the best cases for both exercise types. In general, although the SVM and RF always performed substantially better than chance, as expected with the use of the hierarchical architecture in the SVM we found that the SVM performed better than the RF. The low number of training instances may be responsible for this relatively lower performance of the RF.

The features investigated in our paper led to very good performance possibly due to the thorough analyses discussed in IV. Overall, we found that the extracted features–except the vertical pelvis velocity–were included in at least one optimal subset (not all reported in this paper due to space limitation), but there is still room for improvement. For example, [29] showed that the center of foot pressure helps discriminate between PCPs and non-PCPs. It may also help discriminate between pain levels in CP even though it was not possible to compute this feature from our dataset. Although in [29] these features were captured by force plates in front of the subject's seat, the increasing popularity of wearable and on-body sensing devices (e.g. in [50-51]) will allow these features be captured in more ubiquitous settings. More intensive feature extraction techniques could also lead to improvement in recognition performance although our aim was to identify a reduced feature set across functional movements. Aside from trying new features, pooling the classification of pain in the different exercises using multi-task machine learning methods may also improve recognition [26]. Multi-task learning may additionally provide the possibility of a common model independent of exercise type. Finally, reducing the number of anatomical segments tracked for recognition through careful feature crafting and optimization provides an opportunity for portable motion sensing for physical rehabilitation in ubiquitous settings, which is critical in CP [13, 34].

References

[1] H. Breivik, B. Collett, V. Ventafridda, R. Cohen, and D. Gallacher, "Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment", European Journal of Pain, vol. 10(4), pp. 287–333, 2006.

[2] C. B. Johannes, T. K. Le, X. Zhou, J. A. Johnston, and R. H. Dworkin, "The prevalence of chronic pain in united states adult: results of an internet-based survey", J. Pain, vol. 11(11), pp. 1230–1239, 2010.

[3] International Association for the Study of Pain, "Introduction", Pain, vol. 24(Supplement 1), S3–S8, 1986.

[4] The Health and Social Care Information Centre, Chronic Pain [Online] http://tinyurl.com/oyn6tyq Accessed: 15 April 2015. 2011.

[5] IASP Taxonomy Working Group, "Part III: Pain terms, a current list with definitions and notes on usage", in Classification of Chronic Pain, IASP Press, 2011, pp. 209–214.

[6] J. Vlaeyen and S. Linton, "Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art", Pain, vol. 85(3), pp. 317–332, 2000.

[7] R. J. Gatchel, Y. B. Peng, M. L. Peters, P. N. Fuchs, and D. C. Turk, "The biopsychosocial approach to chronic pain: scientific advances and future directions", Psychological Bulletin, vol. 133(4), p. 581, 2007.

[8] D. Turk and A. Okifuji, "Psychological factors in chronic pain: evolution and revolution", J Cons & Clin Psychol, vol. 70, p. 678, 2002.

[9] M. Friedrich et al. "Combined exercise and motivation program: effect on the compliance and level of disability of patients with chronic low back pain: a randomized controlled trial", Archives of Physical Medicine and Rehabilitation, vol. 79(5), pp. 475–487, 1998.

[10] S. Taimela, C. Diederich, M. Hubsch, and M. Heinricy, "The role of physical exercise and inactivity in pain recurrence and absenteeism from work after active outpatient rehabilitation for recurrent or chronic low back pain: a follow-up study", Spine, vol. 25(14), pp. 1809–1816, 2000.

[11] S. D. Liddle, G. D. Baxter, and J. H. Gracey, "Exercise and chronic low back pain: what works?", Pain, vol. 107(1), pp. 176–190, 2004.

[12] S. May, "Self-management of chronic low back pain and osteoarthritis", Nature Reviews Rheumatol, vol. 6(4), p. 199, 2010.

[13] A. Singh et al, "Motivating people with chronic pain to do physical activity: opportunities for technology design", CHI, p. 2803, 2014.

[14] K. Wiech and I. Tracey, "Pain, decisions, and actions: a motivational perspective", Frontiers in Neuroscience, vol. 7, 2013.

[15] E. Wade, A. Parnandi, and M. Mataric, "Using socially assistive robotics to augment motor task performance in individuals post-stroke", IROS, pp. 2403–2408, 2011.

[16] C. Schönauer et al. "Chronic pain rehabilitation with a serious game using multimodal input", Proc. ICVR, pp. 1–8, 2011.

[17] A. Koenig et al, "Controlling patient participation during robot-assisted gait training", J Neuroeng Rehabil, vol. 8(14.10), p. 1186, 2011.

[18] F. J. Keefe and A. R. Block, "Development of an observation method for assessing pain behavior in chronic low back pain patients", Behavior Therapy, vol. 13(4), pp. 363-375, 1982.

[19] P. O'Sullivan, "Diagnosis and classification of chronic low back pain disorders: maladaptive movement and motor control impairments as underlying mechanism", Manual Therapy, vol. 10(4), p. 242, 2005.

[20] M. Aung et al. "The automatic detection of chronic pain-related expression: requirements, challenges and a multimodal dataset", IEEE Trans in Affective Computing, 2015.

[21] T. Hadjistavropoulos et al, "A biopsychosocial formulation of pain communication", Psychol Bulletin, vol. 137(6), p. 910, 2011.

[22] J. Rainville et al, "Exercise as a treatment for chronic low back pain", The Spine Journal, vol. 4(1), pp. 106-115, 2004.

[23] P. J. Watson et al. "Surface electromyography in the identification of chronic low back pain patients: the development of the flexion relaxation ratio", Clinical Biomechanics, vol. 12(3), pp. 165-171, 1997.

[24] Z. Hammal and J. F. Cohn, "Automatic detection of pain intensity", ICMI, pp. 47–52, 2012.

[25] S. Kaltwang et al. "Continuous pain intensity estimation from facial expressions", Adv. in Visual Comp., pp. 368-377, 2012.

[26] B. Romera-Paredes, H. Aung, N. Bianchi-Berthouze, and M. Pontil, "Multilinear multitask learning", Machine Learning, p. 1444, 2013.

[27] P. Lucey, J. F. Cohn, K. M. Prkachin, P. E. Solomon, and I. Matthews, "Painful data: The UNBC-McMaster shoulder pain expression archive database", Automatic Face & Gesture Recognition, pp. 57-64, 2011.

[28] D. K. Ahern et al. "Comparison of lumbar paravertebral EMG patterns in chronic low back pain patients and non-patient controls", Pain, vol. 34(2), pp. 153-160, 1988.

[29] G. Gioftsos and D. Grieve, "The use of artificial neural networks to identify patients with chronic low-back pain conditions from patterns of sit-to-stand manoeuvres", Clin Biomech, vol. 11(5), pp. 275-280, 1996.

[30] J. P. Dickey, M. R. Pierrynowski, D. A. Bednar, and S. X. Yang, "Relationship between pain and vertebral motion in chronic low-back pain subjects", Clinical Biomechanics, vol. 17(5), pp. 345-352, 2002.

[31] M. S. Aung, N. Bianchi-Berthouze, P. Watson, and A. C. de C Williams, "Automatic recognition of fear-avoidance behavior in chronic pain physical rehabilitation", Pervasive Comp Tech for Healthcare, 2014.

[32] T. A. Olugbade, M. S. H. Aung, N. Marquardt, A. C. de C. Williams, and N. Bianchi-Berthouze, "Bi-modal detection of painful reaching for chronic pain rehabilitation systems", ICMI, pp. 455-458, 2014.

[33] S. Felipe, C. Bradley, A.Singh, A. Williams, and N. Bianchi-Berthouze, "Roles for personal informatics in chronic pain", PervasiveHealth, 2015

[34] A. Singh, S. Piana, D. Pollarolo, G. Volpe, G. Varni, A. Tajadura-Jiménez, A. Williams, A. Camurri, and N. Bianchi-Berthouze, "Go-with-the-flow: Tracking, Analysis and Sonification of Movement and Breathing to Build Confidence in Activity Despite Chronic Pain", HCI, in press

[35] S. Banks and R. Kerns, "Explaining high rates of depression in chronic pain: a diathesis-stress framework", Psychol Bull, vol. 119, p. 95, 1996.

[36] T. Pincus and S. Morley, "Cognitive-processing bias in chronic pain: a review and integration", Psychol Bulletin, vol. 127, pp. 599-617, 2001.

[37] A. S. Zigmond and R. P. Snaith, "The Hospital Anxiety and Depression Scale", Acta Psychiatrica Scandinavica, vol. 67(6), pp. 361-370, 1983.

[38] A. E. Reading, "A comparison of pain rating scales", Journal of Psychosomatic Research, vol. 24(3), pp. 119-124, 1980.

[39] M. Luštrek and B. Kaluža, "Fall detection and activity recognition with machine learning", Informatica, vol. 33(2), pp. 197-204, 2009.

[40] L. I. Kuncheva, T. Christy, I. Pierce, and P. M. Sa'ad, "Multi-modal biometric emotion recognition using classifier ensembles", in Modern Approaches in Applied Intelligence, Springer, 2011, pp. 317-326.

[41] H. Griffin et al, "Laughter type recognition from whole body motion", ACII, pp. 349-355, 2013.

[42] V. Vapnik, "Universal learning technology: Support Vector Machines", NEC J. Advanced Technology, vol. 2(2), pp. 137-144, 2005.

[43] L. Breiman, "Random forests", Machine Learning, vol. 45, p. 5, 2001.

[44] P. Narendra and K. Fukunaga, "A Branch and Bound algorithm for feature subset selection", IEEE Trans on Comp, vol. 100, p. 917, 1977.

[45] G. L. Shum, J. Crosbie, and R. Y. Lee, "Three-dimensional kinetics of the lumbar spine and hips in low back pain patients during sit-to-stand and stand-to-sit", Spine, vol. 32(7), pp. E211-E219, 2007.

[46] D. E. Marple-Horvat and S. L. Gilbey, "A method for automatic identification of periods of muscular activity from EMG recordings", J Neurosci Mtds, vol. 42(3), pp. 163-167, 1992.

[47] H. Meng, A. Kleinsmith, and N. Bianchi-Berthouze, "Multi-score learning for affect recognition", ACII, p. 225, 2011.

[48] J. Joshi, A. Dhall, R. Goecke, and J. F. Cohn, "Relative body parts movement for automatic depression analysis", Proc. Conference on Affective Computing and Intelligent Interaction, pp. 492–497, 2013.

[49] S. Alghowinem, R. Goecke, M. Wagner, G. Parker, and M. Breakspear, "Head pose and movement analysis as an indicator of depression", ACII, p. 283, 2013.

[50] A. Tajadura-Jiménez et al, "As Light as your Footsteps: Altering Walking Sounds to Change Perceived Body Weight, Emotional State and Gait", CHI, pp. 2943-2952, 2015.

[51] M. Weigel et al, "iSkin: Flexible, Stretchable and Visually Customizable On-Body Touch Sensors for Mobile Computing", CHI, 2015.