Enabling context aware data analysis for long-duration repetitive stooped work through human activity recognition in sheep shearing

Mark Robinson[1], Lei Lu[1], Ying Tan[1], Kusal Goonewardena[3], Denny Oetomo[1], Chris Manzie[2]

***Abstract*— There is evidence to suggest that changes in kinematics and neuromuscular control in activities that take place over long periods of time lead to increased injury risk. The collection of biometric data over long time periods could provide insight into these injuries. However, it is difficult to analyse long period biometric data for occupations as the analysis depends on the activity being performed, and it is not practical to manually label the amount of data required. A sufficiently accurate human activity recognition algorithm can provide a means to segment the activities and allow this analysis, but the classification must be robust to the inter-individual differences, as well as the intra-individual variations in movement over time that are the target of analysis. This work presents a person-independent human activity recognition algorithm for sheep shearing using a Hidden Markov Model with physical features that are identified to be relevant to spinal movement quality. The classifier achieved an F1 score of 96.47% in identifying the shearing task.**

I. INTRODUCTION

Repetitive and prolonged spinal flexion (stooping) remains a task common to many occupations and sporting activities. It is a continuing ergonomic challenge, having been identified as an important risk factor in lower back injuries, a condition estimated to impact 80% of people during their life [1]. Stooped work is especially common in industries where workplace modifications are difficult, such as in agriculture and construction; and it is estimated that 100s of millions of workers are at risk of lower back injuries from stooped work globally [2]. In sheep shearing, where workers spend upwards of 6 hours in a stooped working posture each day, injury rates are severe [3] and lower back injuries account for 50% of the cost of all injuries [4].

Research in sheep shearing indicates that injury risk increases throughout the day, and 68% more injuries occur in the last two hours of work than in the first two hours of work [4]. As the task remains the same, any changes in kinematics or neuromuscular control that occur over the day are potentially important factors in injury risk. It is accepted that muscle fatigue increases injury risk, and results in altered kinematics [5], and changes in neuromuscular control [6]. There is evidence to suggest that prolonged and repetitive spinal flexion also changes neuromuscular control [7] and movement [8], and leads to lower back injury and pain [9], [10]. In the recent research relating to sheep shearing back injuries [11]–[14], there is no study that collected kinematic data from workers shearing more than 5 sheep (per case). The collection and analysis of data over an entire working day or longer could provide a significant insight into the factors leading to injury.

While improvements in sensor technology enable the collection of longer term biometric data outside of the lab, the processing of such data is challenging for real working scenarios. One reason for this is in complex activities, different sub-activities may require different analyses. Other methods require cyclic activities [1], [15]; in a complex task this requires identifying the start and end points of each repetitive activity. Sheep shearing is repetitive work, containing two main parts, (1) the catch and drag—which consists of a significant manual handling effort where a sheep (typically weighing 70 kg) is dragged to the shearing stand, and (2) shearing the sheep—stooped work with complex 3D spinal movement with comparatively much lower forces on the spine. This cycle is repeated upward of 200 times per day for an experienced shearer [3].

In order to more easily label the large amount of collected data, we need an activity recognition algorithm that can distinguish these two sub-tasks by identifying the start and end points of shearing cycles. To do this across many subjects without labelling data for each person, the model should be person-independent.

The most important aspect of human activity recognition might be the selection of appropriate features from the data [16]. For a person-independent model [17] suggests that joint-angle-based features outperform features derived from raw sensor data. As different people are likely to differ in the way that they perform tasks, features that are consistent between people are necessary [16]. In forward bending, differences in lumbo-pelvic rhythm are known to contribute to movement differences between people [1]. It is expected that selecting features to minimise this difference will be successful in a person-independent classifier.

In this work, 50 hours of motion capture data is collected from working shearers and a person-independent Hidden Markov Model (HMM) activity recognition method was selected As it can model temporal dependencies between activities with the consideration of stochastic human variations, it is suitable for this application [18]. The results show that HMM is effective with a small number of joint-angle-based features to classify the distinct parts of the shearing task, enabling the analysis of longer term data. Other machine learning methods could also provide avenues for improvement. We will investigate various machine learning techniques in our future work.

This work is supported by Australian Wool Innovation, Mark Robinson is supported by an Australian Government RTP scholarship

mrobinson2@student.unimelb.edu.au

[1]Department of Mechanical Engineering, The University of Melbourne, Melbourne, Australia

[2]Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne, Australia

[3]Elite Akademy Sports Medicine, Melbourne, Australia

Fig. 1. Sheep shearer with XSens portable motion capture

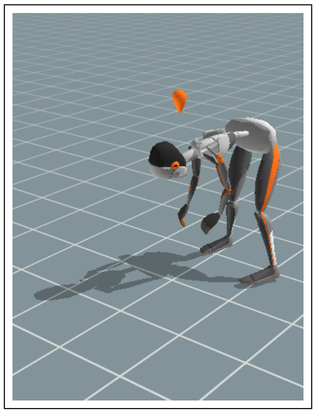

Fig. 2. Motion capture output used for data labelling

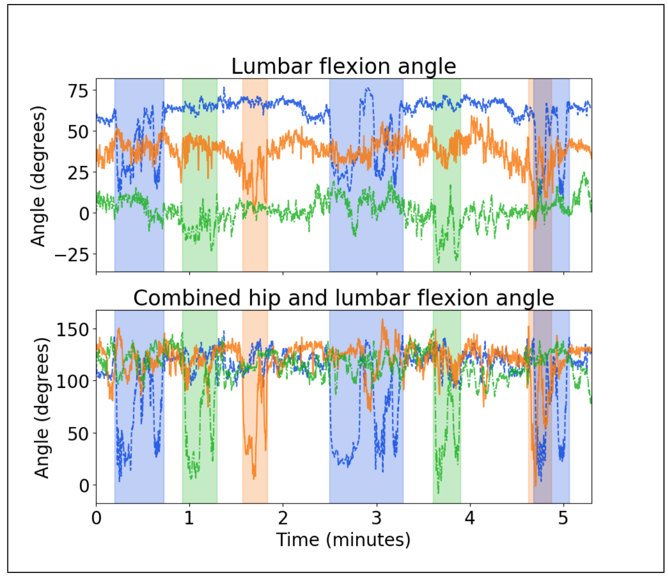

Fig. 3. A segment of data from three subjects showing lumbar flexion, and combined lumbar and hip flexion with catch and drag sections shaded

II. METHODS

A. Participants

Six male sheep shearers aged between 21 and 54 (mean 29.3 ± 11.75) years were recruited for the study. Informed consent was obtained from each participant and the experiment was approved by the University of Melbourne human ethics advisory group (Ethics ID 1853436). This represents a wide-ranging selection of shearers with varying levels of skill, from two weeks experience to more than 30 years. One shearer was also recorded in two different sheds, and this data is included as an extra subject.

B. Experimental set-up

A portable wireless inertial measurement unit (IMU) based motion capture system (Xsens Awinda) was used to collect kinematic data. The system consists of 17 IMU sensors. Through sensor fusion of the accelerometer, gyroscope, and magnetometer channels in each IMU, combined with a scaled skeletal model, joint kinematics were produced by the Xsens MVN Analyze software, sampled at 60Hz. The experimental setup can be seen in Figure 1.

The subjects' body-segment lengths were measured, and the IMUs were then placed using velcro straps and tight-fitting shirt; with lower body sensors placed under the shearers' clothes, and secured with tape. The IMUs were placed according to the Xsens guidelines, with the exception of the lower leg sensors. The recommended placement on the shin-bone was not suitable for sheep shearers, as the inside of both legs are required to hold and manoevure sheep during the work. This would expose the IMUs to large external forces and cause significant discomfort for the shearer. These IMUs were re-positioned to the outside of the leg.

C. Protocol

The experiment involved instrumenting shearers performing their regular work activities. The data was therefore collected in alignment with the typical rest-work schedule for sheep shearing. Shearing takes place in four two-hour sessions (shearing "runs"), with start times at 7:30am, 10:00am, 1:00pm, 3:30pm, with a one hour lunch-break after run 2, and 30 minute breaks after run 1 and run 3. One shearer was instrumented each day, and data was collected for the entire working day. Shearers were required to be at their station 30 minutes early at the start of the day, during which time the sensors were attached and calibrated. For the other runs, shearers were required 10 minutes early for calibration.

A total of 50 hours of data was collected across seven days, as occasional disruptions occurred during shearing. Two shearing runs, one from the start and end of the day for each subject were exported to video as seen in Figure 2 and manually labelled with custom software as being in either the catch and drag, or the shearing phase of the task.

D. Feature selection and activity recognition

For discriminating between stooped work and other activities, features based on spinal flexion are likely to be successful. However, while stooping, rotations of both the lumbar spine and the pelvis contribute to trunk motion [19]. The relative contribution of each, and the co-ordination between the two segments are known to change between individuals, as well as over time with muscle fatigue [1], and between cohorts symptomatic and non-symptomatic with lower back pain [20], which is common among shearers.

It was expected that by combining these two angles, the inter- and intraindividual differences over time would be reduced, and this can be seen in Figure 3. The sum of the right hip and lumbar flexion angles and velocities were selected as features. The lumbar and hip joint angles were extracted from the motion capture, and downsampled to 20 Hz to speed up training. The joint angles were smoothed with a third-order Savitzky-Golay (S-G) filter with a window size of approximately 7.5 seconds, corresponding to half the length of a typical catch and drag. The S-G filter was also used to numerically differentiate the joint angles to obtain the corresponding velocities, without overly amplifying noise. The extracted features and labels for each subject were separated into a training set to parameterise the HMM, and

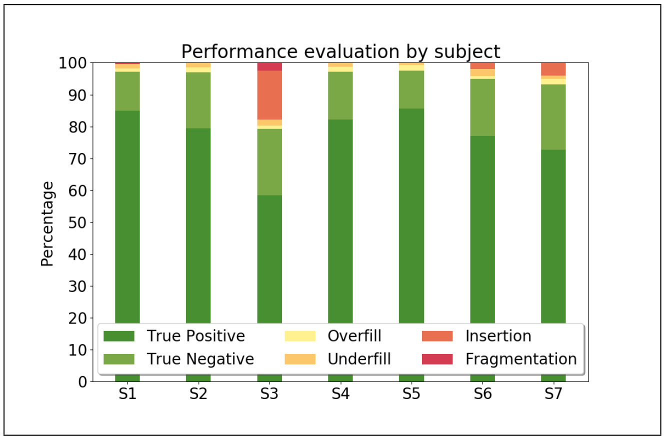

Fig. 4. Activity recognition performance evaluation for subjects (S1-S7) using test data with metrics proposed in [23]

testing set with 75% and 25% of the data respectively.

HMMs are the simplest dynamic bayesian network, and a good introduction to HMMs can be found in [21], among others. A two-state HMM, with multivariate Gaussian state emission probability distributions, was used with shearing, and catch and drag states. Initial state distributions were calculated from the labelled data without temporal information, with a uniform transition matrix. The model was trained using a supervised expectation maximisation algorithm with the labelled training data with a machine learning library in Python [22]. A NULL class was not used for this application, as all the data was taken from the shearing runs where it was assumed all activities would be relevant.

E. Analysis

The performance of the classifier was calculated for each participant from the test data set using standard precision and recall metrics [16]. Precision is the ratio of true positives to all positives returned by the classifier, and recall is the ratio of true positives to all positives. The F_1 score, which is the harmonic mean of precision and recall scores is also reported. The errors are further investigated using the method presented in [23], which categorises the errors to provide additional insight. In order to evaluate the generalisation of the classifier to an unseen user, leave-one-person-out cross validation is also performed [16]. Where the HMM is re-trained on data from all but one subject and evaluated on the remaining subject.

III. RESULTS

The precision, recall, and F_1 scores are calculated for each subject, and can be seen in Table I.

Further detail is presented for each subject in the test data set in Figure 4, with error types classified frame-by-frame using the method in [23]. This shows small amounts of underfill, overfill, and fragmentation errors, as well as significant insertion errors in some subjects.

Precision, recall, and F_1 scores for each subject for the leave-one-person-out cross validation can be seen in Table II.

TABLE I
TEST DATA PERFORMANCE EVALUATION

	Shearing			Catch & Drag		
	Prec.	Recall	F_1	Prec.	Recall	F_1
S-1	98.87	97.89	98.38	87.08	92.71	89.80
S-2	97.95	98.23	98.09	92.47	91.37	91.92
S-3	77.97	91.86	84.35	80.36	56.22	66.15
S-4	98.05	98.52	98.22	92.39	90.15	91.25
S-5	98.00	99.01	98.50	93.28	87.16	90.12
S-6	96.53	95.82	96.17	84.39	86.77	85.56
S-7	92.53	98.64	95.49	95.32	77.73	85.63
Total	95.26	97.71	96.47	89.95	80.86	85.16

TABLE II
LEAVE-ONE-PERSON-OUT CROSS VALIDATION

	Shearing			Catch & Drag		
	Prec.	Recall	F_1	Prec.	Recall	F_1
S-1	98.51	97.33	97.92	85.33	91.36	88.24
S-2	96.96	98.31	97.63	93.19	88.26	90.66
S-3	92.83	91.00	91.91	70.09	75.00	72.46
S-4	98.49	98.22	98.35	94.01	94.87	94.44
S-5	97.13	90.80	93.86	62.75	85.24	72.29
S-6	96.96	94.62	95.78	82.30	89.40	85.70
S-7	95.80	98.73	97.24	94.52	83.50	88.67
Mean	96.67	95.57	96.10	83.17	86.80	84.64

IV. DISCUSSION

The motivation for this activity recognition is to enable longer term biometric data analysis through the automatic labelling of data. The primary goal is to accurately identify the shearing cycle in the data. Doing this can allow for cyclic analyses, for example, calculating an ensemble average and comparing across the day. For this reason, the results for the shearing category are considered more important. The classifier achieved an F_1 score of 96.47% for the shearing class, with a person-independent classifier. While the catch and drag class is less important, it performs worse with an F_1 score of 85.16%. The leave-one-person-out cross validation results are similar with a shearing F_1 score of 96.10%, indicating that the person-independent classifier should also generalise well to new data.

There are still problems with the classifier, and these errors are categorised using the method in [23], shown in Figure 4, to provide further insight and direct improvements. Four types of errors are identified: underfill, overfill, insertions, and deletions. Under- and overfill errors relate to the timing of the state transition itself. Underfill indicates that the transition occurs too early, and overfill that it occurs too late. These errors less severe, because timing errors are often present in the labelling of the data. In this case, the data was labelled manually from the motion capture, and it is likely that timing of the labels are slightly offset from the true transition point. If these are neglected then the classifier can be seen as 100% accurate for three of the seven subjects.

There are only two other types of errors generated by the classifier presented here: fragmentations, and insertions. A fragmentation error occurs where a false negative segment is surrounded by two true positives; serving to split the correctly identified segment in two. For this application, even a small error here would have increased impact analysing the data as it would introduce two incorrect cycles.

An insertion error occurs where a false positive is returned that has no overlap with an actual positive [23]. It indicates that the classifier is identifying an activity that isn't happening at all. The insertion errors seen in Figure 4 all occur within the catch and drag part of the task, and this is the likely reason for the lower classifier performance in that category. Subject 3 appears to have significantly worse performance than the others. This is perhaps misleading, as upon further inspection of the insertion errors, in all cases they can be seen to occur while the shearer takes a break during a shearing run and is sitting down and leaning forward with a poor posture. As this occurs consistently among different subjects, this indicates a problem with the activity recognition to discriminate between sitting in high spinal flexion, and the shearing task—rather than poor performance in a particular subject. These periods of rest should more accurately be considered part of a NULL class, which was excluded in the classifier design. This is a problem that should be addressed in future iterations and could be improved by including additional features targeted at discriminating between stationary and non-stationary activities, such as features derived from raw sensor acceleration [24].

No other types of errors were found in the evaluation. The success of the use of combined hip and lumbar flexion as a feature indicates that while there are differences in forward bending between individuals and in a single individual over the course of a day, these differences are likely a result of altered lumbopelvic rhythm.

While this data was collected using a full body motion capture system, the physical features selected only required data from two joints. The small number of selected features helps to reduce the chance of over-fitting the data, as well as making the algorithm more practical. Because only two joint angles and velocities have been used, it is possible that this work could be replicated using only two IMUs placed on the pelvis and thoracic spine, making it practical for a wearable sensor application. IMUs placed on the pelvis and thoracic spine have also been used in other work to quantitatively assess the quality of spinal movement [1], [15]. The classifier presented here could therefore be incorporated to add context to a quantitative movement assessment over a long period.

V. CONCLUSIONS

This paper presents an HMM based person-independent human activity recognition method to classify kinematic data from sheep shearers into the separate sub-tasks in shearing, and identify the start and end points of the shearing cycle. This is useful to allow biometric data analyses to be extended to long term data collected from sheep shearers in real working conditions. It also provides evidence that kinematic differences (both between individuals and after fatigue) in forward bending in shearing is a result of altered contributions from hip and lumbar flexion.

REFERENCES

[1] B. Hu and X. Ning, "The influence of lumbar extensor muscle fatigue on lumbar–pelvic coordination during weightlifting," *Ergonomics*, vol. 58, no. 8, pp. 1424–1432, 2015.

[2] Z. Luo and Y. Yu, "Wearable stooping-assist device in reducing risk of low back disorders during stooped work," *2013 IEEE Int Conf Mechatron Autom*, no. Ld, pp. 230–236, 2013.

[3] S. Milosavljevic, A. B. Carman, A. G. Schneiders, P. D. Milburn, and B. D. Wilson, "Three-dimensional spinal motion and risk of low back injury during sheep shearing," *Appl. Ergon.*, vol. 38, no. 3, pp. 299–306, 2007.

[4] L. J. Fragar, R. C. Franklin, and A. Lower, "Occupational Health and Safety Risk Associated with Sheep and Wool Production in Australia," ACAHS & RIRDC, Tech. Rep., 2001.

[5] R. R. Bini, F. Diefenthaeler, and M. C. Bolli, "Fatigue effects on the coordinative pattern during cycling: Kinetics and kinematics evaluation," *J Electromyogr Kinesiol*, vol. 20, no. 1, pp. 102–107, 2010.

[6] T. Paillard, "Effects of general and local fatigue on postural control: A review," *Neurosci. Biobehav. Rev.*, vol. 166, no. 3758, p. 38, 2007.

[7] M. Solomonow, "Neuromuscular manifestations of viscoelastic tissue degradation following high and low risk repetitive lumbar flexion," *J Electromyogr Kinesiol*, vol. 22, no. 2, pp. 155–175, 2012.

[8] X. Zhu and G. Shin, "Kinematics and muscle activities of the lumbar spine during and after working in stooped postures," *J Electromyogr Kinesiol*, vol. 23, no. 4, pp. 801–806, 2013.

[9] M. Solomonow, B. He Zhou, R. V. Baratta, Y. Lu, M. Zhu, and M. Harris, "Biexponential recovery model of lumbar viscoelastic laxity and reflexive muscular activity after prolonged cyclic loading," *Clin Biomech*, vol. 15, no. 3, pp. 167–175, 2000.

[10] G. Shin and C. D'Souza, "EMG activity of low back extensor muscles during cyclic flexion/extension," *J Electromyogr and Kinesiol*, vol. 20, no. 4, pp. 742–749, 2010.

[11] P. Pal, D. Gregory, S. Milosavljevic, A. Carman, and J. Callaghan, "A review of risk factors associated with low back injury in wool harvesting," *J. Occup. Health Saf. Aust. N. Z.*, vol. 24, no. 5, pp. 435–453, 2008.

[12] D. E. Gregory, C. Laughton, A. Carman, S. Milosavljevic, and J. P. Callaghan, "Trunk postures and peak and cumulative low back kinetics during upright posture sheep shearing," *Ergonomics*, vol. 52, no. 12, pp. 1576–1583, dec 2009.

[13] P. Pal, S. Milosavljevic, D. Gregory, A. Carman, and J. Callaghan, "The influence of skill and low back pain on trunk postures and low back loads of shearers," *Ergonomics*, vol. 53, no. 1, pp. 65–73, 2010.

[14] M. Robinson, R. Mayer, Y. Tan, D. Oetomo, and C. Manzie, "Effects of varying the rest period on the onset angle of lumbar flexion-relaxation in simulated sheep shearing: a preliminary study," in *IEEE Int Conf Rehabil Robot*, 2019, pp. 83–88.

[15] D. J. Larson, B. L. Pinto, and S. H. Brown, "Differential effects of muscle fatigue on dynamic spine stability: Implications for injury risk," *J Electromyogr Kinesiol*, vol. 43, pp. 209–216, 2018.

[16] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Computing Surveys*, vol. 46, no. 3, pp. 1–33, 2014.

[17] A. Zinnen, U. Blanke, and B. Schiele, "An analysis of sensor-oriented vs. model-based activity recognition," *Proceedings - International Symposium on Wearable Computers, ISWC*, pp. 93–100, 2009.

[18] J. Yamato, J. Ohya, and K. Ishii, "Recognising Human Actions in Time-sequential Images using HMMs," pp. 379–385, 1992.

[19] J. Zhou, X. Ning, and F. Fathallah, "Differences in lumbopelvic rhythm between trunk flexion and extension," *Clin Biomech*, vol. 32, pp. 274–279, 2016.

[20] S. A. Scholtes, S. P. Gombatto, and L. R. Van Dillen, "Differences in lumbopelvic motion between people with and people without low back pain during two lower limb movement tests," *Clin Biomech*, vol. 24, no. 1, pp. 7–12, 2009.

[21] L. R. Rabiner and B. H. Juang, "An Introduction to Hidden Markov Models," *IEEE ASSP Magazine*, vol. 3, no. 1, pp. 4–16, 1986.

[22] J. Schreiber, "pomegranate: Fast and flexible probabilistic modeling in python," *J Mach Learn Res*, vol. 18, pp. 1–6, 2018.

[23] J. Ward and H. Gellersen, "Performance metrics for activity recognition," *ACM Trans Intell Syst Technol*, vol. 2, no. 1, pp. 1–23, 2011.

[24] Ó. Lara and M. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.