

TELEMETRY CASE REPORT

Open Access



Animal-borne behaviour classification for sheep (Dohne Merino) and Rhinoceros (*Ceratotherium simum* and *Diceros bicornis*)

Solomon Petrus le Roux, Jacques Marias, Riaan Wolhuter and Thomas Niesler* 

Abstract

Background: The ability to study animal behaviour is important in many fields of science, including biology, behavioural ecology and conservation. Behavioural information is usually obtained by attaching an electronic tag to the animal and later retrieving it to download the measured data. We present an animal-borne behaviour classification system, which captures and automatically classifies three-dimensional accelerometer data in real time. All computations occur on specially designed biotelemetry tags while attached to the animal. This allows the probable behaviour to be transmitted continuously, thereby providing an enhanced level of detail and immediacy.

Results: The performance of the animal-borne automatic behaviour classification system is presented for sheep and rhinoceros. For sheep, a classification accuracy of 82.40% is achieved among five behavioural classes (standing, walking, grazing, running and lying down). For rhinoceros, an accuracy of 96.10% is achieved among three behavioural classes (standing, walking and lying down). The estimated behaviour was established approximately every 5.3 s for sheep and 6.5 s for rhinoceros.

Conclusions: We demonstrate that accurate on-animal real-time behaviour classification is possible by successful design, implementation and deployed on sheep and rhinoceros. Since the bandwidth required to transmit the behaviour class is lower than that which would be required to transmit the accelerometer measurements themselves, this system is better suited to low-power and error-prone data communication channels that may be expected in the animals habitat.

Keywords: Behavioural classification, Accelerometer, GPS, Machine learning, Biotelemetry, Sheep, Rhinoceros

Background

The ability to study animal behaviour is important in many fields of science and especially biology and behavioural ecology. The latter often employs high-resolution spatial and temporal information to analyse the behaviour of animals. Depending on the sensors used, aspects such as energy expenditure, behaviour, location, speed, heart rate and temperature are monitored [1–5].

A commonly used sensor to monitor the activity level of animals is the accelerometer [6, 7]. Animal-attached tri-axial accelerometers provide a detailed picture of the

activity patterns and allow the collection of data from animals in a non-invasive manner and the subsequent analysis of their behaviour. Such animal-borne biologgers or biotelemetry tags, respectively, log or transmit raw sensor data for extended periods of time. The tags are subsequently retrieved and the acceleration signals analysed to answer key questions relating to animal behaviour. However, the manual monitoring and analysis of such accelerometer measurements can be tedious, time-consuming and not feasible for large or open-ended data volumes. Recently, researchers have begun to consider machine learning techniques to automatically classify these large datasets into behavioural classes. For example, McClune et al. classify the behaviour of badgers (*Meles meles*) as walking, trotting, snuffling and resting [8]. Automatic

*Correspondence: trn@sun.ac.za

Department of Electrical and Electronic Engineering, University of Stellenbosch, Banghoek Road, Stellenbosch, South Africa

behaviour classification systems based on tri-axial accelerometer data are trained by recording high-resolution timestamped video recordings during data collection. These recordings serve as ground truth for the raw data, which is used to manually label the data according to specific behaviours. The labelled data are subsequently used to train statistical classifiers. These classifiers can then automatically categorise large datasets according to the chosen behaviours. Automatic behaviour classification has already been applied to a range of animals using various statistical classification and machine learning techniques, including artificial neural networks [9], decision trees [8–13], discriminant function analysis [14], hidden Markov models [15], k -nearest neighbours [8, 10, 16], linear discriminant analysis [9], moving averages with thresholds [17], quadratic discriminant analysis [18], random forests [9, 19] and support vector machines [9, 15, 20]. These systems have accomplished automatic behaviour classification with high accuracy, as listed in Table 1. The table lists examples of classified behaviours for a small subset of animal species who have recently been studied using automatic behaviour classification techniques.

Currently, automatic behaviour classification is performed as a post-processing step, after data collection. This is the case for all systems described above and listed in Table 1. Such configurations provide very valuable, but historical, information. The ability to analyse animal behaviour in real time, however, has great potential for applied ecological monitoring and wildlife conservation [21]. Some studies have considered the real-time monitoring of animal behaviour using other types of sensors. For example, Cangar et al. consider the automatic real-time monitoring of locomotion and posture behaviour of pregnant cows within a confined space prior to calving using top view cameras and online image analysis [22]. The system achieved high accuracies with an average of

85% for standing and lying and 87% for eating or drinking behaviour for eight cows during the last 24 h before calving. However, video cameras consume a lot of energy and generate large amounts of data and are therefore not suitable for long-term deployment. Furthermore, image or video processing is computationally demanding, which presents challenges for the successful application in real-time nature conservation systems. Satellite- or GSM-enabled GPS transmitters, on the other hand, have a long track record of successfully monitoring animals in real time in remote locations [1, 21, 23]. Wall et al. demonstrate the importance of not only collecting but also analysing GPS data in real time, using tracking collars on African Elephants [21]. It was possible to promptly detect and avoid an elephant bull from breaking through expensive electrical fencing into neighbouring subsistence farming land in order to forage in fields of maize and to inform wildlife managers when elephant herds moved close to specific locations, such as the A2 highway on the Cape-to-Cairo route [21]. The real-time nature of the system allowed immediate alerts to be issued. The same author shows that the system was able to identify a wounded elephant using movement-rate analysis, which led to quick veterinary response and the recovery of the animal. Finally, the system was able to successfully detect elephant mortality by means of immobility analysis, which is key in their anti-poaching and real-time monitoring activities.

Wall et al. go on to point out that while accelerometer data can provide valuable information for real-time management applications, the increased volumes of data produced by the accelerometers introduce new challenges for both the tag's limited memory and battery capacities as well as the wireless communication of these data to a control room. The author concludes that some form of on-board processing is inevitable. To date, however, no study has considered the on-board classification

Table 1 Summary of some automatic animal behaviour classification systems based on the statistical classification of tri-axial accelerometer data

Animal	Behaviour	Accuracy (%)	Source
Baboon	Forage, run, rest, stand, walk	88	[19]
Badger	Walking, trotting, snuffling, resting	83–92	[8, 10]
Cattle	Walking, standing, lying down, ruminating, feeding	55–99	[11, 12, 18, 20]
Cheetah	Feeding, mobile, stationary	83–94	[15]
Elephant	Feeding, bathing, walking, swaying	87.8	[14]
Goat	Resting, eating, walking	61–82	[17]
Oystercatcher	Flying, foraging, handling prey, sitting, standing, walking	67.6–86.8	[13]
Red Fox	Foraging, leaps, trotting	95	[16]
Vulture	Eating, lying down, active flight, passive flight, running, standing, preening	80–90	[9]

The accuracies shown were calculated by the various authors in different ways. Please refer to the studies themselves to determine exactly how the individual accuracies were calculated

of tri-axial accelerometer data for real-time automatic behaviour classification. Related studies have considered the real-time monitoring of human activities based on accelerometer data [24–30]. These studies demonstrate good performance but are also less constrained in terms of battery life and communication bandwidth since users are co-operative and can be relied on to recharge the batteries and to be within reception of standard communication technologies, such as Wi-Fi and cellular networks, which can be used to transfer data. For animal monitoring purposes, these approaches are not feasible. Replacing batteries involves recapturing animals which is difficult, expensive and places additional stress on the animals. Furthermore, wild animals typically live in areas with very limited or no communication coverage except satellite or specifically designed wireless sensor networks [31]. Accordingly, reducing the required bandwidth for data communication will result in increased longevity of the tags and reduce operational costs for satellite transmitters.

It is in this context that we consider the feasibility of an animal-borne behaviour classification system, which captures and automatically classifies three-dimensional accelerometer data in real time and on the animal. The system is evaluated on both sheep and rhinoceros where it distinguishes in real time between behaviours such as running, walking, standing, grazing and lying down. We describe both the developed hardware and software in the interest of system reproducibility. Although the system was implemented using specially designed hardware, it could in principle be reproduced on other tri-axial accelerometer equipped biotelemetry tags to allow real-time transmission of classified behavioural information to further assist research in fields such as behavioural ecology, biosecurity, nature conservation and agriculture.

Methods

The system employs specially designed biotelemetry tags. These tags are externally attached to the animal where they sample a tri-axial accelerometer and pass the raw data through an on-board classifier, which determines the current behaviour of the animal. The updates are available in real time and are stored locally and also transmitted over a wireless communication channel. The system design involves five steps: First, the biotelemetry tag hardware is designed. Second, in a data collection phase, the tags are programmed to sample and store raw accelerometer signals while ground-truth behavioural labels are noted. Third, these recorded signals are transferred to a workstation where they are labelled and used to train and evaluate an automatic classifier.¹ Fourth, the

classification algorithm is implemented on the tags. Finally, the tags are attached to animals where they continuously capture, classify, record and transmit the behaviour of the animal. A ground-truth comparison establishes the accuracy of the on-animal classification. These steps are set out in more detail in the sections that follow.

Hardware

Five tags were developed for our study. The hardware layout is shown in Fig 1. The tags are powered by 3.7 V 1800 mAh lithium-ion batteries. The design utilised a L6932H1.2 ultra-low drop-out voltage regulator (STMicroelectronics, Geneva, Switzerland), a MSP430FR5739 low-power mixed signal microcontroller (Texas Instruments, Dallas, TX, USA), a GNS602 GPS receiver (Global Navigation Systems, North Shields, UK), two FM25V20 ferroelectric non-volatile RAM (FRAM) storage modules of 2 MB each (Cypress Semiconductor, San Jose, CA, USA), a 2 GB MicroSD card (Kingston, München, Germany) and an ADXL345 tri-axial accelerometer (Analog Devices, Norwood, MA, USA) with high-resolution (13-bit) measurement at up to ± 16 g. A low-power CC1101 sub-1 GHz RF transceiver (Texas Instruments, Dallas, Texas, USA) allows wireless data communication at 433 MHz. The tags were packaged using two different collar designs for sheep and rhinoceros. The sheep collars had a total weight of 281 g, which included a 146 mm \times 80 mm \times 65 mm polycarbonate casing (126 g), the physical electronics (32 g), a battery (33 g) and a belt (90 g) used to fit the collar around the neck of sheep. The rhinoceros collar had a total weight of 371 g, which included a 117 mm \times 77 mm \times 38 mm thick durable and waterproof casing (165 g) made of polyoxymethylene (acetal), the physical electronics (32 g), two batteries

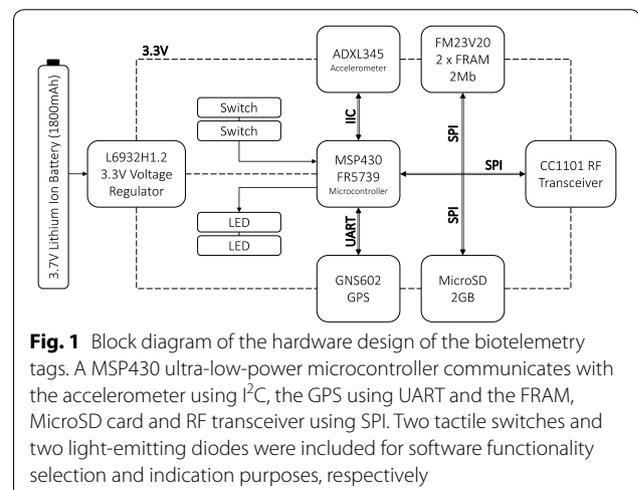


Fig. 1 Block diagram of the hardware design of the biotelemetry tags. A MSP430 ultra-low-power microcontroller communicates with the accelerometer using I²C, the GPS using UART and the FRAM, MicroSD card and RF transceiver using SPI. Two tactile switches and two light-emitting diodes were included for software functionality selection and indication purposes, respectively

¹ Data processing and classification were performed in Python using the SciPy library (version 0:14:0).

(66 g) and two belts (108 g) used to fit the collar around the legs of rhinoceros.

Initially, the hardware was configured to collect raw accelerometer measurements for both sheep and rhinoceros. In this setting, the microcontroller was programmed² to continuously sample the tri-axial accelerometer at 100 Hz for sheep and 40 Hz for rhinoceros. These measurements were timestamped and subsequently logged to the MicroSD card for later retrieval. The GPS, RF transceiver and FRAM modules were not used during this data collection step.

Data collection

Datasets for both sheep and rhinoceros were compiled as part of this study. The collected data were manually labelled with the observed behaviour to provide ground truth. The resulting annotated corpus was used to develop and evaluate statistical classifiers.

For the sheep, raw data were collected on a farm in Carnarvon, Northern Cape, South Africa, in July 2014. The tags were fitted around the necks of the sheep using collars as shown in Fig. 2a. The x , y and z axes are associated with left–right, up–down and forward–backward movements of the sheep, respectively. The data collection was performed during 3 separate days. On each of these three mornings, collars were attached to five individual sheep (Dohne Merino) randomly selected from a group of approximately twenty animals. The sheep were then led to a larger camp and left undisturbed for the duration of the day, during which the tags continuously logged the acceleration. Five common types of behaviour were identified: lying down, standing, walking, running and grazing. The behaviour of the sheep was manually documented with timestamps to serve as ground truth. At the end of each day, the sheep were collected, the collars removed and the recorded acceleration signals downloaded.

The rhinoceros dataset was collected at a private sanctuary early in 2015. Three rhinoceros were available for data collection, one male and one female White Rhinoceros (*Ceratotherium simum*) and one male Black Rhinoceros (*Diceros bicornis*). The animals were held in a camp approximately 100 ha in area and moved around freely. The collars were fitted around the left back leg of the rhinoceros as shown in Fig. 2b. The x axis was associated with up–down movement, the y axis with forward–backward movement and the z axis with left–right movement. Data collection took place over a period of 1 week. The rhinoceros were fed daily, and this provided the opportunity to fit the collars. Once the collars were fitted, the



Fig. 2 Biotelemetry collars. **a** Fitted around the necks of four sheep. The accelerometer x , y and z axes are associated with left–right, up–down and forward–backward movements of the sheep, respectively, **b** fitted around the black leg of a rhinoceros. The accelerometer x , y and z axes are associated with the up–down, forward–backward and left–right movements, respectively

rhinoceros were left undisturbed to roam freely in their camp. The collars were removed during later feeding sessions and the data downloaded, after which the collars were redeployed. Four common behaviours were identified: lying down, standing, walking and running. The rhinoceros were monitored each day at the feeding sessions during which timestamped video recordings were made. The video recordings allowed accurate labelling of the gathered data.

Data labelling

In order to train and evaluate the statistical classifiers, it is important to have accurate ground truth. All collected data for both sheep and rhinoceros were segmented and labelled manually with the corresponding behaviour. For the sheep, the ground truth took the form of manual annotations (2014 data). For the rhinoceros, the timestamped video recordings were used (2015 data). A summary of the raw sheep and rhinoceros datasets is given in Table 2.

² Microcontroller programming was performed in the C programming language using Code Composer Studio 5.5.0.

Table 2 Raw unbalanced datasets

	Lying down	Standing	Walking	Running	Grazing	Total
Sheep	17:18	0:55	6:08	1:12	6:27	32:00
Rhinoceros	2:08	3:50	0:48	0:07	–	6:53

Data collected for each behaviour class for sheep and for rhinoceros (hours:minutes)

Data collection from rhinoceros was logistically and practically challenging, and as a result the data corpus gathered was smaller than that for sheep. Running behaviour was not observed often for rhinoceros. Grazing behaviour could not be detected because the tags were attached to the back leg for the rhinoceros, and not to the neck as for sheep.

Figure 3 depicts typical measured acceleration signals for the five different behaviours in the sheep dataset. Similarly, Fig. 4 depicts typical measured acceleration signals for the three different behaviours in the rhinoceros dataset. The figures demonstrate that the acceleration waveforms of different behavioural classes exhibit different characteristics.

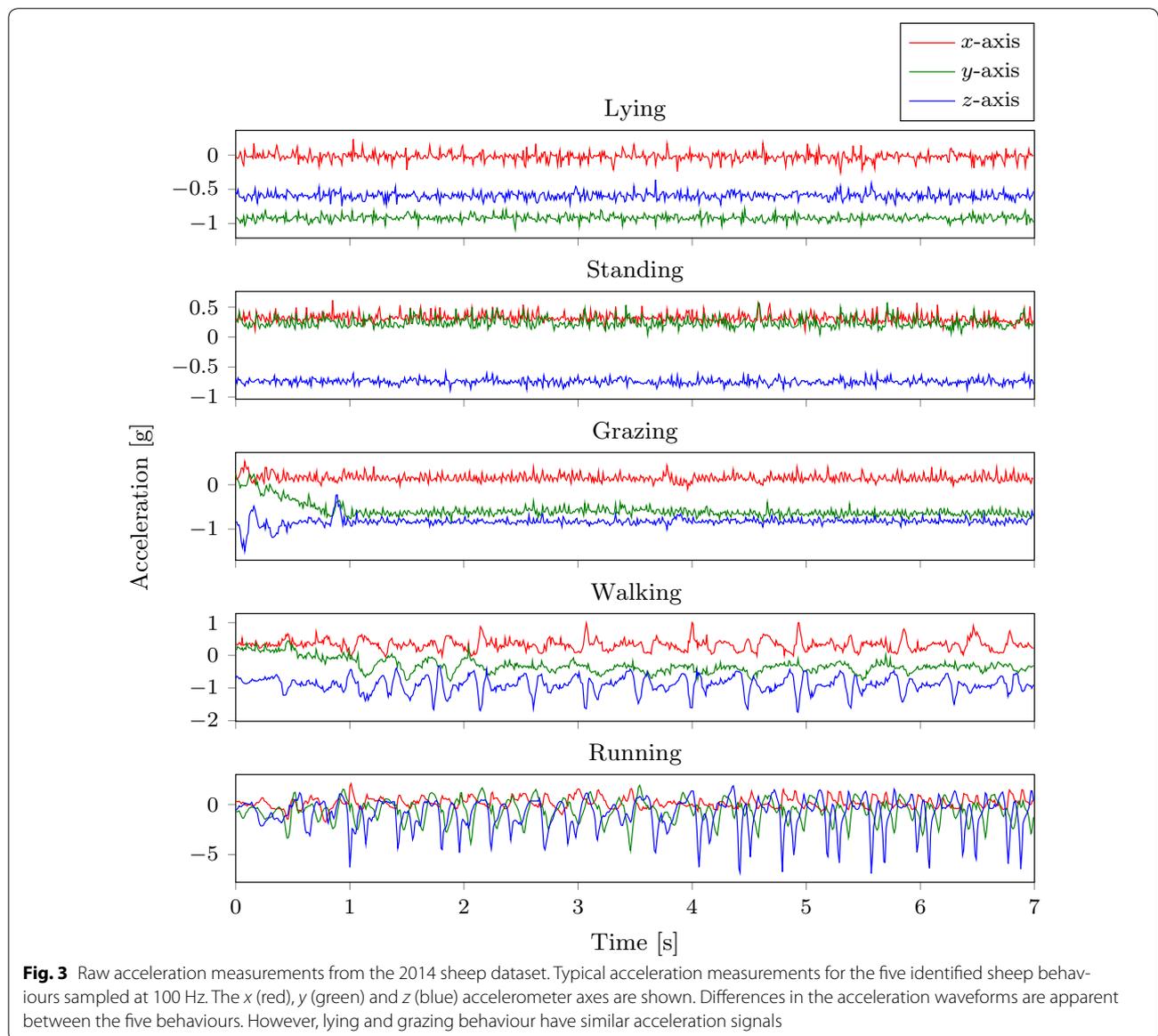


Fig. 3 Raw acceleration measurements from the 2014 sheep dataset. Typical acceleration measurements for the five identified sheep behaviours sampled at 100 Hz. The x (red), y (green) and z (blue) accelerometer axes are shown. Differences in the acceleration waveforms are apparent between the five behaviours. However, lying and grazing behaviour have similar acceleration signals

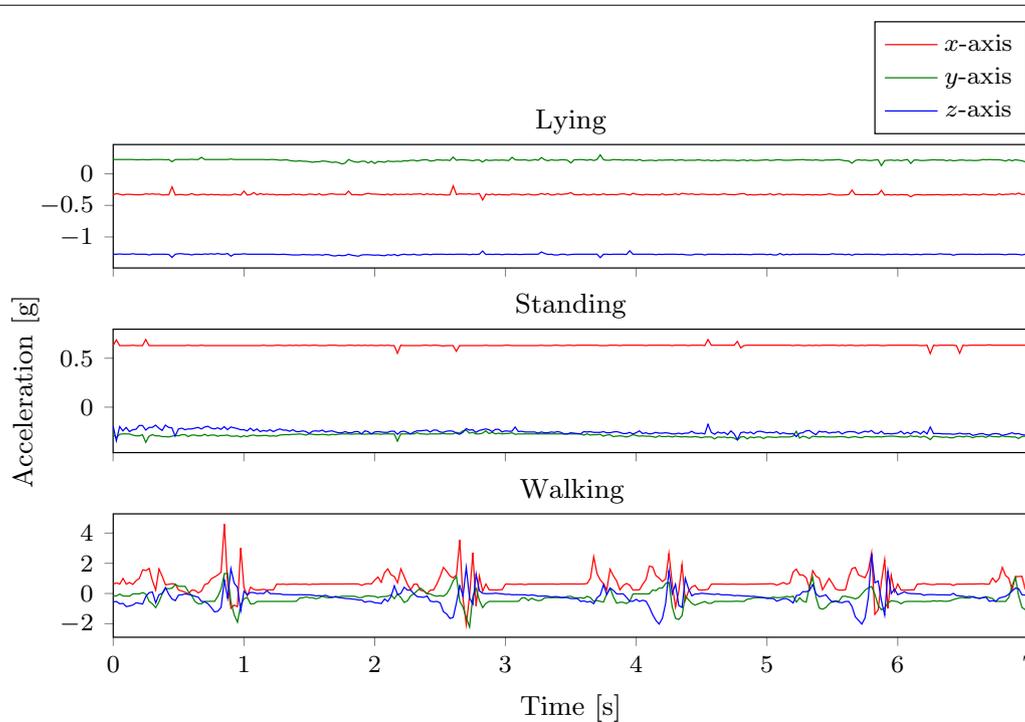


Fig. 4 Raw acceleration measurements from the 2015 rhinoceros dataset. Typical acceleration measurements for the three identified rhinoceros behaviours sampled at 40 Hz. The x (red), y (green) and z (blue) accelerometer axes are shown. Differences in the acceleration waveforms are apparent between the three behaviours

Data pre-processing and feature extraction

Analysis of the gathered datasets showed very little spectral content above 10 Hz. The three acceleration axes were therefore low-pass-filtered using an 8th-order Butterworth filter with a cut-off frequency of 10 Hz before segmenting the signals into consecutive frames and applying a Hamming data window. Frames spanned 512 and 256 accelerometer measurements, respectively, for sheep and rhinoceros to achieve similar frame durations, due to different sampling frequencies of 100 Hz for sheep and 40 Hz for rhinoceros. These frames were allowed to overlap by 50%. The features extracted from each frame were based on those employed in [20, 32] and included the 11 features listed in Table 3. For each feature, except average signal magnitude, three values were computed, one for each accelerometer axis, leading to a total of 31 features per frame. The same features were extracted for both the sheep and the rhinoceros datasets.

Dataset balancing

The two datasets were balanced after feature extraction, in order to avoid possible bias between classes or animals. Since the sheep dataset did not identify individual

Table 3 Features extracted from the compiled datasets

Feature	Equation
Average signal magnitude	$\frac{1}{N} \sum_{i=1}^N \sqrt{x_i^2 + y_i^2 + z_i^2}$
Maximum value	$\max(\mathbf{x})$
Minimum value	$\min(\mathbf{x})$
Mean (\bar{x})	$\frac{1}{N} \sum_{i=1}^N x_i$
Standard deviation (σ_x)	$\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$
Variance	σ_x^2
Skewness	$\frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^3}{\sigma_x^3}$
Kurtosis	$\frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^4}{\sigma_x^4}$
Energy	$\frac{1}{N} \sum_{i=1}^N x_i ^2$
Spectral entropy	$\sum_{i=1}^N P(x_i) \log \frac{1}{P(x_i)}$
Pairwise correlation between the axes	$\frac{\text{COV}(\mathbf{x}, \mathbf{y})}{\sigma_x \sigma_y}$

Each frame consists of N sequential samples, and \mathbf{x} denotes a vector of these samples for each accelerometer axis, x , y and z . The FFT of \mathbf{x} is denoted by \mathbf{X} and the normalised power spectrum of \mathbf{x} by $P(\mathbf{x})$. Cross-correlation is calculated for each axis pair (x, y), (x, z), and (y, z). All features except average signal magnitude provide three values: one per axis

animals, the classes were balanced over all the sheep. This was achieved by random selection from the aggregated data to limit the number of samples for each class to the number available for the least frequent class. Hence, the resulting dataset is evenly drawn from all animals while maintaining the same number of total samples from each class. In the case of the rhinoceros dataset, collection took place from three distinct and identifiable animals. Therefore, this dataset could be balanced over behavioural classes as well as animals. This was achieved by limiting the data available for each class from each animal to that available for the least frequent class among all animals. The resulting dataset had an equal number of samples for each class and for each rhinoceros. We decided to omit the small amount of running data from the rhinoceros dataset for cross-animal testing purposes, since this behaviour was observed for only one animal. The total number of feature vectors available after balancing the datasets is given in Table 4.

Linear discriminant analysis

Linear discriminant analysis (LDA) is a classification method, which assigns a d -dimensional input feature vector $\mathbf{x} = \{x_1, \dots, x_d\}$ to one of K classes (C_k) using a linear decision boundary [33]. This is accomplished after projecting the data to a lower $K - 1$ dimensional space using a transformation matrix \mathbf{W} , as shown in Eq. 1.

$$\mathbf{y} = \mathbf{W}^T \mathbf{x} \tag{1}$$

Here, $\mathbf{y} = \{y_1, \dots, y_{K-1}\}$ is the $K - 1$ dimensional projected vector. Class separation is maximised in the lower-dimensional space by choosing the matrix \mathbf{W} to maximise the between-class variability while minimising the within-class variability. This is achieved by maximising the generalised Fisher criterion, as given in Eq. 2.

$$J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}} \tag{2}$$

In Eq. 2, \mathbf{S}_B is the sample covariance of the class means:

$$\mathbf{S}_B = \sum_{k=1}^K N_k (\boldsymbol{\mu}_k - \boldsymbol{\mu})(\boldsymbol{\mu}_k - \boldsymbol{\mu})^T \tag{3}$$

where the number of samples in class k is given by N_k . \mathbf{S}_W is the within-class covariance matrix:

$$\mathbf{S}_W = \sum_{k=1}^K \sum_{\mathbf{x} \in C_k} (\mathbf{x} - \boldsymbol{\mu}_k)(\mathbf{x} - \boldsymbol{\mu}_k)^T \tag{4}$$

The mean for each class by $\boldsymbol{\mu}_k$ and the overall mean by $\boldsymbol{\mu}$. The decision boundary is found by solving the generalised eigenvalue problem in Eq. 5 [34]:

$$\mathbf{S}_B \mathbf{W} = \lambda \mathbf{S}_W \mathbf{W} \tag{5}$$

This results in $K - 1$ eigenvectors, which are used to discriminate between the K classes. Classification is accomplished by determining the nearest class mean in the $K - 1$ dimensional space.

$$k = \underset{k}{\operatorname{argmin}} d(\mathbf{W}^T \mathbf{x}, \mathbf{W}^T \boldsymbol{\mu}_k) \tag{6}$$

where $d(\mathbf{a}, \mathbf{b})$ is in our case the Euclidean distance between vectors \mathbf{a} and \mathbf{b} .

An underlying assumption of LDA is that each class density is a multivariate Gaussian and that the classes share the same covariance matrix [35]. We were motivated to use the LDA technique since the computations required for classification are simple and could be implemented on our low-power hardware, and because informal analyses of the extracted features indicated their densities to be approximately Gaussian. However, other classifiers could be considered in future.

Software and implementation

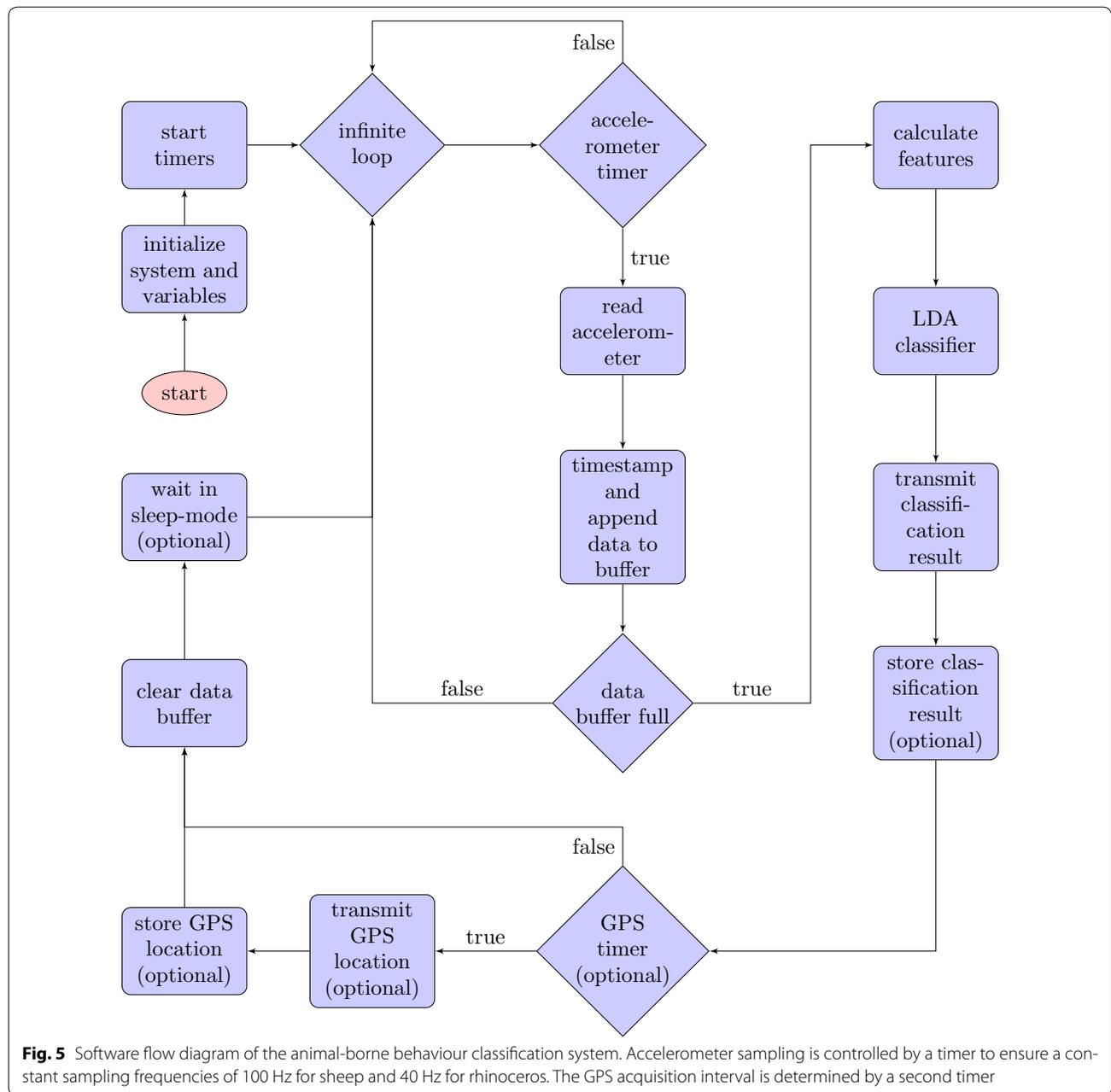
Animal-borne behaviour classification is achieved by implementing the trained LDA classifier on the tag itself. The tag executes the software routine shown in Fig. 5.

The routine begins by initialising the system and setting all variables to appropriate values. Next, two timers are started. The first corresponds to the desired accelerometer sampling frequency in our case 100 Hz for sheep and 40 Hz for rhinoceros. The second corresponds to the desired GPS sampling interval. The routine then enters the main loop, in which it waits for the read-accelerometer timer to trigger an interrupt. The corresponding interrupt service routine samples the accelerometer and stores the raw x , y , and z acceleration values in a data buffer. This process is repeated until a pre-defined number of samples have been obtained (512 for sheep and 256 for rhinoceros). Feature extraction is subsequently

Table 4 Balanced datasets

	Lying	Standing	Walking	Running	Grazing	Total
Sheep	769	768	766	768	768	3839
Rhinoceros	297	297	297	–	–	891

Indicating the number of feature vectors extracted for each behaviour class for the sheep and rhinoceros data presented in Table 2



performed on the frame stored in the data buffer. The feature vector is subsequently passed to the LDA classifier, and the classification result transmitted and optionally also stored on-board. At this stage, if the GPS timer was triggered, the GPS location can be obtained, transmitted and stored. The loop completes by clearing the data buffer. An additional delay can be introduced at this point during which the system enters a low-power state for a set duration. This is useful when behavioural updates are required less frequently. However, we configured the system for continuous behavioural updates; hence,

no delay was introduced. With this software implementation, behavioural updates are available approximately every 5.3 s for sheep and 6.5 s for rhinoceros, since the sampling is taken at constant intervals (100 Hz for sheep and 40 Hz for rhinoceros), and a set number of samples are stored in the data buffer (512 samples for sheep and 256 samples for rhinoceros) before classification and transmission.

The animal-borne behaviour classification system was implemented and field-tested to evaluate its performance. In April 2015, the tags were configured with the

classifier trained on the 2014 sheep dataset and tested at the same farm in Carnarvon on two randomly chosen sheep over a period of 2 days. Video recordings were made to serve as ground truth with which the classification accuracy could be determined. In April 2016, the tags were configured with the classifier trained on the 2015 rhinoceros dataset and tested at the same location on three different rhinoceros over a period of 4 days. Video recordings were again made to provide the ground truth with which to evaluate classification accuracies.

Results

The first goal of the experiments was to determine the utility of the features in Table 3 for behaviour classification in sheep and rhinoceros. All classifiers are trained and evaluated on the datasets in Table 4 within a cross-validation framework. For the sheep dataset, tenfold cross-validation was used. Since the identities of the animals were not unambiguously known for this dataset, the tenfold split was executed by random selection. However, in the case of the rhinoceros dataset, inter-animal classification performance could be evaluated. This was achieved by threefold cross-validation, where training data were always taken from two animals and testing data from the third. This is a leave-one-animal-out cross-validation scheme. Sequential forward selection (SFS), a simple greedy search technique, was used to determine the best combination of features to use for classification [36]. SFS incrementally adds features to the classifier while optimising cross-validated classification accuracy. For both sheep and rhinoceros datasets, SFS was performed on a development set that was split from each training partition during cross-validation. This ensured that the held-out cross-validation partition on which final evaluation was performed was completely independent of all data used during training and system optimisation. Figure 6 shows behaviour classification accuracies measured during cross-validation on the held-out data.

For sheep, the classifier was trained to distinguish between five behaviours: lying down, standing, walking, running and grazing. The best classifier employed 27 features and achieved an accuracy of 85.70%. Table 5 shows the corresponding confusion matrix. It is clear that lying down, standing and running can be accurately classified. Walking behaviour was confused with standing and with grazing in 5.74% and in 8.21% of cases, respectively. Grazing behaviour was confused with lying down in 38.54% of all cases. This is due to the similarities of the measured acceleration signals for the two behaviours, as shown in Fig. 3.

For rhinoceros, the classifier was trained to distinguish between three behaviours: lying down, standing and walking. The best classifier used 15 features and achieved

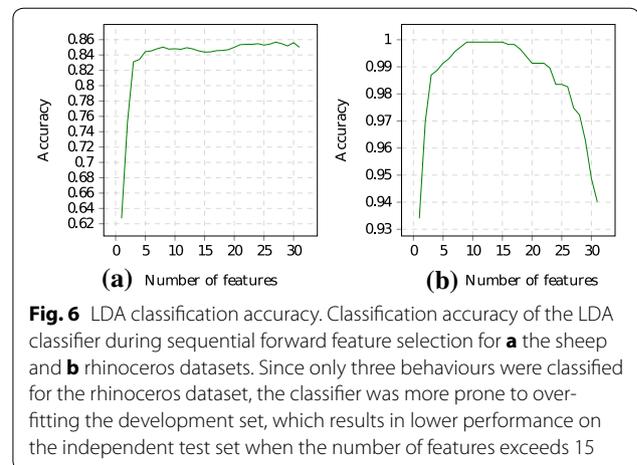


Fig. 6 LDA classification accuracy. Classification accuracy of the LDA classifier during sequential forward feature selection for **a** the sheep and **b** rhinoceros datasets. Since only three behaviours were classified for the rhinoceros dataset, the classifier was more prone to overfitting the development set, which results in lower performance on the independent test set when the number of features exceeds 15

an almost perfect classification accuracy of 99.78%. Table 6 shows the corresponding confusion matrix. It is clear that both walking and lying down were classified very accurately. Standing behaviour was confused with lying down in 0.67% of cases. Furthermore, Fig. 6 shows that, for both sheep and rhinoceros, most performance gain is afforded by approximately the first 5 of the possible 31 features. For sheep, these were standard deviation (z axis), minimum value (x and y axes), variance (z axis) and spectral entropy (z axis). For rhinoceros, the top five features determined by SFS were maximum value (x axis), minimum value (x and y axes), mean value (x axis) and variance (z axis). The results of SFS therefore allow the dimensionality of the feature vector to be reduced while maintaining the best possible performance.

The second goal was to determine whether a LDA classifier, using a subset of features with low computational demands, could be used to accurately classify the behaviour of sheep and rhinoceros in real time, given the limited resources of the animal-borne system. We found that our system was able to accomplish real-time on-animal classification using a 12-dimensional feature vector consisting of the maximum value, minimum value, mean value and standard deviation for all three axes. The system achieved classification accuracies of 82.40% for sheep and 96.10% for rhinoceros and delivered updates every 5.3 and 6.5 s for sheep and rhinoceros, respectively. Although further investigation is needed to determine the trade-offs between feature complexity, system performance and power consumption, the chosen features agree to a large extent with those chosen by SFS.

Discussion

The accuracies achieved by our on-animal classifiers were slightly below the best accuracies achieved during laboratory development (Fig. 6). To some extent, this is due to the partial mismatch between the features used by the

Table 5 Confusion matrix showing the accuracies achieved by the LDA classifier trained on the 2014 sheep dataset

	True behaviour				
	Lying down (%)	Standing (%)	Walking (%)	Running (%)	Grazing (%)
<i>Predicted behaviour</i>					
Lying down	90.51	6.76	0	0	2.73
Standing	0.65	96.74	2.47	0	0.14
Walking	0.13	5.74	85.63	0.26	8.21
Running	0	0	0.78	99.09	0.13
Grazing	38.54	2.6	2.47	0	56.39

on-animal system, and those identified to be most useful by SFS. It is therefore expected that on-animal classification can be further improved by employing the feature sets identified during SFS. Further research is needed to determine the trade-offs between feature complexity, system accuracy and power consumption for real-time implementations. Nevertheless, the on-animal accuracies achieved are sufficient to be useful in field tests, it is not far below the best laboratory results and it compares well with the accuracies listed in Table 1. This indicates that the behaviour classes considered are distinguishable on the basis of accelerometer measurements for both sheep and rhinoceros and that this can be achieved on-animal and in real time. As also anticipated by Wall et al., we believe such real-time behavioural updates have the potential to greatly aid nature conservation efforts [21]. We also believe that animal-borne behaviour classification can provide additional value to the movement-rate analysis and immobility detection algorithms used by the same authors. Similar analytical techniques can be applied to determine normal daily thresholds for each behaviour and thereby identify abnormal behaviour. For example, excessive lying behaviour and below-normal walking behaviour might indicate animal injury or illness [21, 37]. This is also of particular interest in precision agriculture applications and might compliment current techniques used to detect lameness in cattle. Darr et al. confirm that excessive lying behaviour is linked to lameness in cattle, which can be treated if behavioural changes

Table 6 Confusion matrix showing the accuracies achieved by the LDA classifier trained on the 2015 rhinoceros dataset

	True behaviour		
	Walking (%)	Standing (%)	Lying down (%)
<i>Predicted behaviour</i>			
Walking	100	0	0
Standing	0.67	99.33	0
Lying down	0	0	100

are promptly detected [37]. Therefore, they implemented a system that stores timestamped average values of each accelerometer axis every 15 s. A base station can later retrieve the information over a wireless link, whereafter the data are compared to an 0.5 g threshold to determine animal lying time. Furthermore, animal immobility might be more quickly identified by detecting excessive lying behaviour. Since immobility detection is critical to nature conservation efforts [21], we plan in future to add immobility as a separate class and train the classifiers accordingly. This might allow quick response in an attempt to catch poachers red handed. At present, our specific focus is the monitoring of rhinoceros to assist conservation efforts, since these animals are currently subject to an unprecedented and sustained poaching onslaught [38]. Our field tests indicate that both the behaviour and GPS location can be monitored in real time as illustrated in Fig. 7, which shows the location and behaviour of a rhinoceros over 1 day. One can easily identify the different behaviours and use this information to learn the animal's spatiotemporal preferences over longer periods of time in order to identify abnormal behaviour.

Since our system only transmits the classification result (6 bytes) as opposed to the raw data (2816 bytes) or extracted features (53 bytes), it is better suited to the low-power and error-prone data channels that prevail when considering wireless communication in the animal's habitat. In addition, animal-borne behaviour classification shows potential to reduce the overall power consumption and increase the resolution of GPS loggers/transmitters by dynamically updating the rate at which GPS locations are obtained [39]. Brown et al. calculate the statistical variance of a single accelerometer axis to identify animal activity levels as resting, medium- or high intensity and set the GPS schedule accordingly [39]. Their system reduced the overall power consumption while achieving a greater number of locations per day with a lower mean search time per location attempt [39]. Our system has potential to compliment such strategies by dynamically changing the GPS schedule based upon the animal's behaviour. Furthermore, implementing behaviour-specific

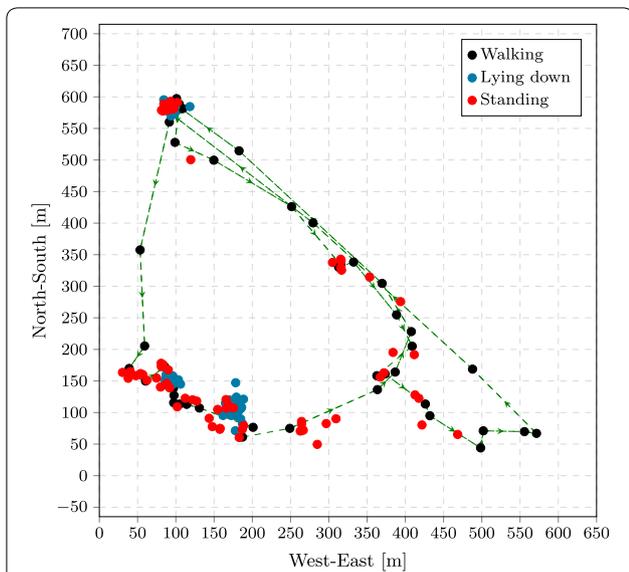


Fig. 7 Rhinoceros behaviour and movement collected in real time. Since the behaviour is classified on the biotelemetry tag, it can be combined with GPS location to produce information on what, where and when the animal is exhibiting specific behaviours. In this figure, we can see where the rhinoceros was lying down (blue), where it stood (red) and where it walked (black)

GPS schedules can enable researchers and conservationists to focus on specific behaviours while maximising the tag's lifetime, since unwanted behavioural updates can be avoided. Wilson et al. demonstrate the need for such behaviour-specific updates in their study of the locomotion dynamics of hunting in wild cheetahs [40]. Their system dynamically adapted the rate at which both the GPS and inertial measurement unit (IMU) were sampled, based upon the animal activity level (measured using the accelerometer) and the time of day (measured using the GPS). This implementation resulted in a reduced average power consumption and higher-resolution data for specific hunting events [40]. Although the system successfully recorded 367 runs over a period of 17 months, a further 530 runs were identified in the activity data, but not recorded, because the collar did not trigger on every run owing to the time of day and conservative trigger thresholds. We anticipate that similar studies might benefit from the use of animal-borne behaviour classification rather than fixed thresholds. Animal-borne behaviour classification might also be used in GPS drift-corrected dead reckoning applications as described in Dewhirst et al. [41]. The study points out that animal dynamic body acceleration has an approximately linear relationship with animal speed, which is a key metric used along with calibrated magnetometer measurements to achieve GPS drift-corrected dead reckoning. This improves the accuracy of

animal path and travel distance estimates [41]. In future work, our technique might provide additional information that can be used in conjunction with the vector of the dynamic body acceleration (VeDBA), magnetometer measurements and GPS data to achieve this goal.

It is clear that biological sensor tags can play a critical role in animal conservation [42]. We hope that further research in the presented and related directions can serve to optimise system longevity and data resolution, thereby enabling advances in applied ecological monitoring and wildlife conservation by means of sensor-equipped tags and real-time on-board processing techniques.

Summary and conclusion

An animal-borne behaviour classification system was successfully designed, implemented, deployed and tested on sheep and rhinoceros. The specially designed biotelemetry tags were able to achieve real-time on-animal classification of behaviour using statistical classification of tri-axial accelerometer measurements. When distinguishing between the five behaviour classes (standing, walking, grazing, running and lying down) for sheep, the best laboratory system achieved a cross-validated classification accuracy of 85.70% using 27 features, while an on-animal classifier using 12 features achieved 82.40% in independent tests. When discriminating between three behaviour classes (standing, walking and lying down) for rhinoceros, the best laboratory system achieved a cross-validated classification accuracy of 99.78% using 15 features, while an on-animal classifier using 12 features achieved 96.10% in independent tests. The animal-borne system produced live behavioural updates approximately every 5.3 s for sheep and 6.5 s for rhinoceros. We conclude that the behaviour of sheep and rhinoceros can be determined with high accuracy from measured tri-axial acceleration data. Furthermore, we conclude that it is possible to achieve this on the animal to provide real-time behavioural information. The development of our system is continuing with the extension of the data sets, the application to other types of animal and the assessment of other classification techniques. In the longer term, we aim to use our animal-borne classifiers to model animal behaviour for conservation applications.

Authors' contributions

SPIR, RW and TRN conceived the design, while JM contributed feature selection, classifier optimisation and data interpretation. SPIR designed and built the tags, acquired the data and analysed the experimental results. SPIR and TRN were major contributors in writing the manuscript. TRN directed the project. All authors read and approved the final manuscript.

Acknowledgements

The authors gratefully acknowledge support by the National Research Foundation of the Republic of South Africa, by Telkom South Africa, and by Innovus of Stellenbosch University.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Consent for publication

Not applicable.

Ethics approval

This research complies with accepted standards for the use of animals in research and teaching as reflected in the South African National Standards 10386:2008. The ethical clearance permit number is SU-ACUM14-00016.

Funding

Funding was provided by the National Research Foundation of the Republic of South Africa, by Telkom South Africa, and by Innovus of Stellenbosch University.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 12 May 2017 Accepted: 7 November 2017

Published online: 21 November 2017

References

- Kays R, Crofoot MC, Jetz W, Wikelski M. Terrestrial animal tracking as an eye on life and planet. *Science*. 2015. <https://doi.org/10.1126/science.aaa2478>.
- Bidder OR, Walker JS, Jones MW, Holton MD, Urge P, Scantlebury DM, Marks NJ, Magowan EA, Maguire IE, Wilson RP. Step by step: reconstruction of terrestrial animal movement paths by dead-reckoning. *Mov Ecol*. 2015;3(1):23. <https://doi.org/10.1186/s40462-015-0055-4>.
- Kumar A, Hancke GP. A zigbee-based animal health monitoring system. *IEEE Sens J*. 2015;15(1):610–7. <https://doi.org/10.1109/JSEN.2014.2349073>.
- Williams HJ, Holton MD, Shepard ELC, Largey N, Norman B, Ryan PG, Duriez O, Scantlebury M, Quintana F, Magowan EA, Marks NJ, Alagaili AN, Bennett NC, Wilson RP. Identification of animal movement patterns using tri-axial magnetometry. *Mov Ecol*. 2017;5(1):6. <https://doi.org/10.1186/s40462-017-0097-x>.
- Fourati H, Manamanni N, Afilal L, Handrich Y. Posture and body acceleration tracking by inertial and magnetic sensing: application in behavioral analysis of free-ranging animals. *Biomed Signal Process Control*. 2011;6(1):94–104. <https://doi.org/10.1016/j.bspc.2010.06.004>.
- Brown DD, Kays R, Wikelski M, Wilson R, Klimley AP. Observing the unwatchable through acceleration logging of animal behaviour. *Anim Biotelem*. 2013;1(1):20.
- Shepard EL, Wilson RP, Quintana F, Laich AG, Liebsch N, Albareda DA, Halsey LG, Gleiss A, Morgan DT, Myers AE, et al. Identification of animal movement patterns using tri-axial accelerometry. *Endanger Species Res*. 2008;10:47–60.
- McClune DW, Marks NJ, Wilson RP, Houghton JD, Montgomery IW, McGowan NE, Gormley E, Scantlebury M. Tri-axial accelerometers quantify behaviour in the Eurasian badger (*Meles meles*): towards an automated interpretation of field data. *Anim Biotelem*. 2014;2(1):5.
- Nathan R, Spiegel O, Fortmann-Roe S, Harel R, Wikelski M, Getz WM. Using tri-axial acceleration data to identify behavioural modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *J Exp Biol*. 2012;215(6):986–96.
- McClune DW, Marks NJ, Delahay RJ, Montgomery WI, Scantlebury DM. Behaviour-time budget and functional habitat use of a free-ranging European badger (*Meles meles*). *Anim Biotelem*. 2015;3(1):7. <https://doi.org/10.1186/s40317-015-0025-z>.
- Vázquez Diosdado JA, Barker ZE, Hodges HR, Amory JR, Croft DP, Bell NJ, Codling EA. Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system. *Anim Biotelem*. 2015;3(1):15. <https://doi.org/10.1186/s40317-015-0045-8>.
- Robert B, White B, Renter D, Larson R. Evaluation of three-dimensional accelerometers to monitor and classify behaviour patterns in cattle. *Comput Electron Agric*. 2009;67(1):80–4.
- Shamoun-Baranes J, Bom R, van Loon EE, Ens BJ, Oosterbeek K, Bouten W. From sensor data to animal behaviour: an oystercatcher example. *PLoS ONE*. 2012;7(5):37997.
- Soltis J, Wilson RP, Douglas-Hamilton I, Vollrath F, King LE, Savage A. Accelerometers in collars identify behavioural states in captive African elephants *Loxodonta africana*. *Endanger Species Res*. 2012;18(3):255–63.
- Grünewälder S, Broekhuis F, Macdonald DW, Wilson AM, McNutt JW, Shawe-Taylor J, Hailes S. Movement activity based classification of animal behaviour with an application to data from cheetah (*Acinonyx jubatus*). *PLoS ONE*. 2012;7(11):49120.
- Painter MS, Blanco JA, Malkemper EP, Anderson C, Sweeney DC, Hewgley CW, Červený J, Hart V, Topinka V, Belotti E, Burda H, Phillips JB. Use of bio-loggers to characterize red fox behavior with implications for studies of magnetic alignment responses in free-roaming animals. *Anim Biotelem*. 2016;4(1):20. <https://doi.org/10.1186/s40317-016-0113-8>.
- Moreau M, Siebert S, Buerkert A, Schlecht E. Use of a tri-axial accelerometer for automated recording and classification of goats grazing behaviour. *Appl Anim Behav Sci*. 2009;119(3):158–70.
- Watanabe N, Sakanoue S, Kawamura K, Kozakai T. Development of an automatic classification system for eating, ruminating and resting behaviour of cattle using an accelerometer. *Grass Sci*. 2008;54(4):231–7.
- Fehlmann G, O'Riain MJ, Hopkins PW, O'Sullivan J, Holton MD, Shepard ELC, King AJ. Identification of behaviours from accelerometer data in a wild social primate. *Anim Biotelem*. 2017;5(1):6. <https://doi.org/10.1186/s40317-017-0121-3>.
- Martiskainen P, Järvinen M, Skön J-P, Tiirikainen J, Kolehmainen M, Mononen J. Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Appl Anim Behav Sci*. 2009;119(1):32–8.
- Wall J, Wittemyer G, Klinkenberg B, Douglas-Hamilton I. Novel opportunities for wildlife conservation and research with real-time monitoring. *Ecol Appl*. 2014;24(4):593–601. <https://doi.org/10.1890/13-1971.1>.
- Cangar A, Leroy T, Guarino M, Vranken E, Fallon R, Lenehan J, Mee J, Berckmans D. Automatic real-time monitoring of locomotion and posture behaviour of pregnant cows prior to calving using online image analysis. *Comput Electron Agric*. 2008;64(1):53–60. <https://doi.org/10.1016/j.compag.2008.05.014>.
- Cagnacci F, Boitani L, Powell RA, Boyce MS. Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. *Philos Trans R Soc Lond B Biol Sci*. 2010;365(1550):2157–62. <https://doi.org/10.1098/rstb.2010.0107>.
- Curone D, Bertolotti GM, Cristiani A, Secco EL, Magenes G. A real-time and self-calibrating algorithm based on triaxial accelerometer signals for the detection of human posture and activity. *IEEE Trans Inf Technol Biomed*. 2010;14(4):1098–105. <https://doi.org/10.1109/TITB.2010.2050696>.
- Purwar A, Jeong DU, Chung WY. Activity monitoring from real-time triaxial accelerometer data using sensor network. In: 2007 International conference on control, automation and systems. 2007. p. 2402–406. <https://doi.org/10.1109/ICCAS.2007.4406764>.
- Saponas S, Lester J, Froehlich JE, Fogarty JA, Landay JA. Ilearn on the iphone: real-time human activity classification on commodity mobile phones. Technical report. University of Washington. 2002.
- Karantonis DM, Narayanan MR, Mathie M, Lovell NH, Celler BG. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE Trans Inf Technol Biomed*. 2006;10(1):156–67. <https://doi.org/10.1109/TITB.2005.856864>.
- He J, Li H, Tan J. Real-time daily activity classification with wireless sensor networks using hidden Markov model. In: 2007 29th Annual international conference of the IEEE engineering in medicine and biology society. 2007. p. 192–195. <https://doi.org/10.1109/IEMBS.2007.4353008>.
- Tapia EM, Intille SS, Haskell W, Larson K, Wright J, King A, Friedman R. Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In: 2007 11th IEEE international symposium on wearable computers. 2007. p. 37–40. <https://doi.org/10.1109/ISWC.2007.4373774>.
- Kumari P, Mathew L, Syal P. Increasing trend of wearables and multimodal interface for human activity monitoring: a review. *Biosens. Bioelectron*. 2017;90:298–307. <https://doi.org/10.1016/j.bios.2016.12.001>.

31. Bracciale L, Catini A, Gentile G, Loreti P. Delay tolerant wireless sensor network for animal monitoring: the pink iguana case. In: De Gloria A, editors. Applications in electronics pervading industry, environment and society. ApplePies 2016. Lecture Notes in Electrical Engineering, vol 429. Cham: Springer; 2018. p. 18–26.
32. Bao L, Intille SS. Activity recognition from user-annotated acceleration data. In: International conference on pervasive computing. Berlin: Springer; 2004. p. 1–17.
33. Bishop C. Pattern recognition and machine learning. New York: Springer; 2007.
34. Li T, Zhu S, Ogiwara M. Using discriminant analysis for multi-class classification: an experimental investigation. *Knowl Inf Syst.* 2006;10(4):453–72.
35. Mika S, Ratsch G, Weston J, Scholkopf B, Mullers K-R. Fisher discriminant analysis with kernels. In: Neural networks for signal processing IX. Proceedings of the 1999 IEEE signal processing society workshop. IEEE; 1999. p. 41–48.
36. Guyon I, Elisseeff A. An introduction to variable and feature selection. *J Mach Learn Res.* 2003;3(Mar):1157–82.
37. Darr M, Epperson W. Embedded sensor technology for real time determination of animal lying time. *Comput Electron Agric.* 2009;66(1):106–11. <https://doi.org/10.1016/j.compag.2009.01.004>.
38. Haas TC, Ferreira SM. Conservation risks: when will rhinos be extinct? *IEEE Trans Cybern.* 2016;46(8):1721–34. <https://doi.org/10.1109/TCYB.2015.2470520>.
39. Brown DD, LaPoint S, Kays R, Heidrich W, Kümmeth F, Wikelski M. Accelerometer-informed GPS telemetry: reducing the trade-off between resolution and longevity. *Wildl Soc Bull.* 2012;36(1):139–46. <https://doi.org/10.1002/wsb.111>.
40. Wilson AM, Lowe JC, Roskilly K, Hudson PE, Golabek KA, McNutt JW. Locomotion dynamics of hunting in wild cheetahs. *Nature.* 2013;498:185–9.
41. Dewhirst OP, Evans HK, Roskilly K, Harvey RJ, Hubel TY, Wilson AM. Improving the accuracy of estimates of animal path and travel distance using GPS drift-corrected dead reckoning. *Ecol Evol.* 2016;6(17):6210–22. <https://doi.org/10.1002/ece3.2359>.
42. Wilson ADM, Wikelski M, Wilson RP, Cooke SJ. Utility of biological sensor tags in animal conservation. *Conserv Biol.* 2015;29(4):1065–75. <https://doi.org/10.1111/cobi.12486>.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

