Cronfa - Swansea University Open Access Repository

This is an author produced version of a paper published in:
Computer Graphics Forum

Cronfa URL for this paper:
http://cronfa.swan.ac.uk/Record/cronfa182

Paper:

http://dx.doi.org/10.1111/j.1467-8659.2009.01469.x

This article is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence. Authors are personally responsible for adhering to publisher restrictions or conditions. When uploading content they are required to comply with their publisher agreement and the SHERPA RoMEO database to judge whether or not it is copyright safe to add this version of the paper to this repository.
http://www.swansea.ac.uk/iss/researchsupport/cronfa-support/
Visualisation of Sensor Data from Animal Movement

Edward Grundy\textsuperscript{1,2}, Mark W. Jones\textsuperscript{1}, Robert S. Laramee\textsuperscript{1}, Rory P. Wilson\textsuperscript{2} and Emily L.C. Shepard\textsuperscript{2}

\textsuperscript{1}Visual Computing Group, Swansea University, UK
e-mail: \{csed, m.w.jones, r.s.laramee\}@swansea.ac.uk
\textsuperscript{2}Institute of Environmental Sustainability, Swansea University, UK
e-mail: \{r.p.wilson, 381601\}@swansea.ac.uk

Abstract
A new area of biological research is identifying and grouping patterns of behaviour in wild animals by analysing data obtained through the attachment of tri-axial accelerometers. As these recording devices become smaller and less expensive their use has increased. Currently acceleration data are visualised as 2D time series plots, and analyses are based on summary statistics and the application of Fourier transforms. We develop alternate visualisations of this data so as to analyse, explore and present new patterns of animal behaviour. Our visualisations include interactive spherical scatterplots, spherical histograms, clustering methods, and feature-based state diagrams of the data. We study the application of these visualisation methods to accelerometry data from animal movement. The reaction of biologists to these visualisations is also reported.

Categories and Subject Descriptors (according to ACM CCS): I.3.8 [Computer Graphics]: Applications

1. Introduction
The use of accelerometers to gather data has been investigated for both explicit and implicit user input with the aim of identifying user actions based upon the accelerometry signal resulting from the users motion. The visualisation of accelerometer data, as a subject, is in its infancy, and visualisations are generally limited to 2D time-series plots (e.g. figure 1). Recently, accelerometers have been attached to individual animals by biologists in order to monitor the behaviour of the animal over longer periods, through more environments and at a higher resolution than previously possible. The quantity and interpretive complexity of this new application area make it a prime candidate for visualisation.

Biologists at Swansea University have collected large amounts of data relating to animal movement by attaching sensors to individual subjects. These devices record a multiplicity of data, including acceleration and magnetic field strength in three axes, hydrostatic pressure, light intensity and temperature, thus recording movements of the subject animal and details of the body environment while attached [WSL08]. The small size, low weight and durability of the devices mean they can be deployed on a wide range of animals in many different environments [WSL08].

A major focus of research for the biologists is to determine animal activity and behavioural patterns from the gathered data. The tri-axial acceleration data is of specific interest as it provides quantitative data on body posture and motion. The three axes of the accelerometer are aligned to the dorso-ventral axis, the anterior-posterior axis and the lateral axis of the subject animal. These are termed (in biological parlance) heave, surge and sway respectively. These axes are analogous to the Y, Z and X axes in cartesian co-ordinates. Environmental attributes such as pressure, temperature and light level, perform a supporting role in the identification of activities. The sequences and temporal alignment of identified activities give insight into the function of behavioural patterns [SWQ\textsuperscript{*}08]. Effective visualisation of this multivariate data is a challenging proposition which would greatly assist biological analysis of the data.

2. Related Work
Although we have not found any published visualisation research, there has been research in other fields on the uses of accelerometry data. Automatic pattern recognition methods applied to accelerometry in Pervasive Computing indicate the meaning of components and descriptive statistics of the
signal. We focus on providing a tool to visualise patterns in the acceleration signal from the movement of wild animals. The closest related areas in visualisation are concerned with multi-variate data and vector fields (where we consider the acceleration signal a sequence of three-dimensional vectors).

In the area of Pervasive Computing, automatic recognition of user input recorded with an accelerometer is seen as a way to allow greater freedom of input, or obtain information about user activity. Kela et al. [KKM+06] and Schlömer et al [SPHB08] both describe the use of an accelerometer as an input device. These methods have been studied in an attempt to recognise general activities performed by humans and provide behavioural context information for devices worn by users [DSGP03]. Generally, methods improve upon an early description of feature extraction proposed by Van Laerhoven and Cakmakci [VLC00], which involves creating descriptive statistics for windowed regions of the acceleration signal.

The unbiased study of the wild animal behaviour has always been problematic because (i) animals may behave differently in the presence of humans [WCS+89], (ii) they generally cannot be observed all the time (and sometimes hardly observed at all) and because, (iii) even if the animal is visible, it is often difficult to quantify behaviour, particularly with respect to intensity [DSA+90]. Remarkable progress in solid-state technology, however, has led to a recent proliferation in animal-attached technologies which record may aspects of animal biology even when there is no visible contact between animal and researcher [RCW05]. An apparently attractive solution to solving the problem of recognising and quantifying behaviour in this domain uses accelerometers [YSN+99]. Since accelerometers give information on static acceleration [SWQ+08], which is derived from animal posture with respect to the gravitational field [YNS+01], and dynamic acceleration, which is derived from animal movement [SNK+02], examination of accelerometer data should enable workers to determine behaviour [SWQ+08]. However, the process of allocating accelerometer signals to behaviours is complex because the recording frequency has to be high enough to provide at least 5 data points per repetitive behaviour cycle (e.g. per wing-beat or stride [RCW05]) and this should ideally be collected in all three dimensional axes [WSL08] so that large amounts of data are collected over a very short time.

The difficulties in determining behaviours by simple visual inspection of three time-lines in acceleration where diverse and temporally-variable patterns may occur over a wide breadth of scales (ranging from single repetitive movements in e.g. a limb beat, to many hours of immobility during rest) have led workers to propose complex statistical approaches [WBS+06] that are difficult to grasp and do not enhance intuition because they are not linked to any visual representation of the signal which might help identify informative patterns.

The favoured method to present data and results in pervasive computing and biology publications is typically based on a separate line graph for each component of the signal, as in figure 1, or the more stylised presentation given in figure 2, (commonly used to convey the principles to non-experts). An intensity versus time plot effectively describes the temporal relation between “behaviour signatures”, recognised as different intensity readings, and identification of the dominant axis (which represents orientation). However, this visualisation does not provide easy insight into the tri-axial nature of the data, as correlation between the axes has to be consciously observed. Both the pervasive computing and biology applications rely on statistics derived from the acceleration data. Many derived values are calculated over windowed regions and become difficult to comprehend because of the different temporal scaling. In biological research, comparing directly measured attributes (such as pressure or temperature) as well as derived attributes is also valuable. Currently, this is achieved by producing a new 2D time-series graph for the attribute plot, where the x-axes of all the graphs are aligned. The many related line graphs re-
3. Acceleration Data

The acceleration data contains static and dynamic components. A stationary device records its orientation in the gravitational field (the static component). When the device is moving, local acceleration forces distort the reference axis according to the acceleration vector (the dynamic component). Due to the Principle of Equivalence, acceleration forces and gravitational forces are indistinguishable, and therefore recorded by the same sensor. Accurately separating the two components is an active area of research. Assuming the static component is low-frequency and the dynamic is high-frequency, the static component can be approximated by applying smoothing filters to the raw data, and thus the dynamic component can be inferred.

Biologists have proposed that dominant behavioural patterns may be identified from acceleration data using a hierarchical approach [SWQ08]. The static component is first used to identify animal posture, a step that immediately reduces the number of behaviours the animal could have been performing. Patterns in the dynamic component provide further insight. The main axis in which movement is recorded describes the type of motion (e.g., bird flight is characterised predominantly by dynamic changes in the heave axis, whereas a swimming fish produces regular changes in sway acceleration), whereas the frequency and amplitude of changes describe the “vigour”. An example of these observations are shown in figure 2.

Biologists explore the data by viewing three line graphs (one for each channel of the tri-axial sensors) using environmental data, such as pressure and temperature, to support the decision making process. After a long search process, periods of data can be manually identified and labelled as a specific activity. In practice, interpretation of even basic signals can be time consuming, difficult, and error-prone due to (a) three components of two acceleration sources present in a single signal, (b) the signal being presented in the temporal domain (which can be long), (c) no simple visualisation of body orientation to assist the deductive process, (d) signal noise, and (e) slight variances in animal behaviour. We propose methods to visualise accelerometry data which effectively portray postures and variances, reducing the time required to infer behaviours. We develop these visualisations to include environmental and derived attributes.

4. Visualisations

This section presents visualisations of Daily Dairy data [WSL08] for a single device attached to an Imperial Cormorant in South America during December 2006. The data was captured at 8Hz for 8 hours and 40 minutes resulting in 249,988 measurements. Each measurement contains tri-axial acceleration, tri-axial local magnetic field intensity, hydrostatic pressure, infra-red levels, and temperature, resulting in nine channels of data. The data are recorded at 22-bit resolution onto 128Mb flash RA memory attached to the device. The accelerometer has an accuracy of ±0.06g, with a sensitivity range of ±2.21g. Note that recorded data pertain only to motions undergone by the device, and may be offset by a fixed angle due to the attachment to the subject animal. The static component of a datum is computed by normalising a windowed mean of surrounding data. The size of this window is a user-option, but typically accounts for 32 data.

4.1. Spherical Scatterplots

Established vector field visualisations [LHD04] generally rely on the topology of the underlying sample grid (usually two or three dimensional) and encode the vector value as a colour or glyph in a two or three dimensional space. The one dimensional space afforded by sampling over time is a limits the applicability of these methods to these data.

By discarding the time-series and considering each vector as an offset from the origin, a three dimensional scatterplot of the acceleration vectors can be produced (figures 3(a) and 3(b)). For normalised three dimensional vectors this method is equivalent to projecting the points onto the surface of a sphere. When the static component is visualised (figure 3(c)), this performs an implicit conversion to spherical co-ordinates, (φ, θ, r) (where φ and θ are rotations about the central axes, and r is the radial distance from the origin).

A scatterplot makes geometric distribution of the data more apparent, providing an intuitive summary of the acceleration signal. In this case, “features” are deemed any sequence of interest to the biologists. However, this loose definition can be made more robust by noting that biologists...
infer a great deal from the body posture of an animal (which translates to the static component of the signal). Postures can be associated with certain behaviours (as demonstrated in figures 1 and 2), and in this respect the scatterplot visualisation effectively depicts the major postures which occur in the data. The strong trends visible in figure 3 are due to the primary postures of the animal being captured by the acceleration sensor and visualised effectively. The location of the dense point clusters, relative to the reference sphere, provide an intuitive representation of features previously identified by an arduous manual search of line-graphs. Furthermore, the size of a feature is indicative of the variance around a posture, which relates to the dynamic component of the signal caused by behaviour. Activities with a large amount of variance, or strong dynamic components, were the most difficult to recognise manually, but are easily discernable in the scatterplot as less dense regions.

A drawback of this visualisation is the lack of temporal information. Colour may be used to represent time in a fourth dimension, as shown in figure 3(d) where time is presented via the “redness” of the data. However, it is not easy to discern small periods of time when the sequence is long. Even with a large colour space, the perceptibility of small colour gradients remain restrictive.

4.2. Multi-variate Visualisation of Sensor Attributes

As the orientation vector can be reduced to two polar dimensions (even if implicitly) a scalar attribute (mapped to $[0, 1]$) can be added to the radial distance to find a new distance from the origin, thereby elevating data points from the surface of the sphere and replacing the magnitude of acceleration with another attribute of the data. As the direction of acceleration remains, features of the orientation data are preserved with respect to angular distance. We term this method an overlay of the scalar attribute onto the orientation vector data. Overlays of the pressure attributes and derived signal energy attributes are shown in figures 4(b) and 4(c), where the underlaying acceleration data is shown in figure 4(a). The intention of this method is to demonstrate that other components of the data can be used to further differentiate and identify behaviour, using the acceleration data as a common reference point for the many attributes.

The benefit of spatially separating data points according to pressure is best illustrated by figure 4(b). The visual depiction of diving behaviour in the overlay presents a new set of visual features. Ascents and descents of dives are represented as lines near the poles connecting the surface of the
sphere to the elevated data. The large mass elevated from the sphere and below the equator represent underwater activity with the major arcs representing changes in body orientation as the animal moves from one major posture to another.

Correlations with attributes can be found by mapping different attributes to colour and overlay. For example, correlations between pressure and signal energy can be observed during descent, ascent and swimming phases of the dives by overlaying the pressure data and mapping signal energy to colour (figure 4(d)). Signal energy is higher as the animal descends the water column (south pole), while ascent (north pole) is relatively passive. This is explained by the fact that cormorants have to actively descend due to air trapped in their feathers, while ascent is assisted by this buoyancy. In this visualisation, four dimensions of the data (acceleration direction, pressure and signal energy) are presented while remaining coherent, intelligible and informative.

Individual occurrences of behaviours can be isolated by reducing the length of the sequence. A single dive, isolated from the many dives performed by the cormorant (shown in figure 4(b) and figure 4(d)), is illustrated by figure 5. The 2D time-series plot of this sequence is also given for comparison. Detailed views of the overlay show the orientation to be presented without distortion, allowing the method to be used for analysis as well as exploration. In figure 5, high frequency components of the orientation signal during descent (caused by flipper beats as the bird actively descends the water column) are preserved in the overlay. The complexity of the underwater swimming period is also more completely realised in the spherical scatterplot than the 2D plot. An indication of the speed of a movement is given by rendering the location of measured data points along with connecting line segments. However, temporal relations of the measurements are only presented by the connectivity of the line segments.

4.3. Visualisation of Derived Data

The distribution of the orientation data can be visualised by constructing a histogram over a sub-division of the sphere. In the spherical histogram, the orientation data density is represented by the height of the bins, the “footprint” of which is controlled by the user (figures 6(b-d)). The height of each bin visually indicates time spent in a particular orientation.

In evaluating the scatterplot (fig. 3) we have confirmed that features presented to the user correspond to postures manually identified by the expert. Applying a density based clustering method allows the discovery of frequently adopted orientations (see Jain et al. [JMF99] and Xu and Wunsch [XW05] for a review of data clustering methods). Given K clusters, fuzzy c-means results in each datum $p_i$ having an associated vector $u_i = \{u_{i1}, u_{i2}, ..., u_{ik}\}$, which describe the degrees of membership to each cluster [Bez81].
The closest (or major) cluster \( m_i \) of a point is \( p_i \) is given by:

\[
m_i = \arg \max_{k=1}^K (u_{ik})
\]

Clusters can be visualised by colouring each point, using the major cluster \( m_i \) to determine hue (i.e. hue \( = \frac{m_i}{K} \times 360 \)) and indicating membership degree \( u_{im} \) via saturation (figure 7(a)).

4.4. Posture States and Transitions

It is attractive to consider a datum with a high \( u_{im} \) value as indicating the animal was “in” posture \( m_i \) at that time. From this, we can introduce a threshold \( \tau \) to act as a discriminant between “in-posture” and “non-posture” measurements:

\[
INPOST(p_i, \tau) = \begin{cases} 
\text{true} & \text{if } u_{im} > \tau \\
\text{false} & \text{otherwise}
\end{cases}
\]

Observe that when \( INPOST(p_i, \tau) = \text{false} \), \( p_i \) will be outside all postures in \( K \) due to:

\[
\left( \max_{k=1}^K (u_{ik}) < \tau \right) \rightarrow (\forall v \in u_i (v < \tau))
\]

In the time-series, sequences of non-posture data will be delimited by either (a) the end of the sample, or (b) a datum which is in a discovered posture. A non-posture sequence can be considered a transition when it is terminated by in-posture data. Automatic discovery of postures, and transitions, allow us to describe the sample with a state machine, where each state corresponds to a posture, and sequences of non-posture data correspond to transitions. Allowing for some fuzziness in the classification, the behaviour of many animals can be discretely classified, as demonstrated by the labelling of the data shown in figure 1.

Each identified posture can be visualised in the scatter-plot space by drawing a sphere located at the centroid of the cluster, where a sphere at \( \mu_k \) encloses all points with \( u_{im} > \tau \). Connectivity between those states represent transitions which correspond to sequences of non-posture data, as defined earlier (figure 7(b)).

Behavioural tortuosity is a measure defined as the Euclidean distance between two points in the accelerometer time-series:

\[
BHVT(t, t') = \sum_{i=1}^{t'} \| f(i+1) - f(i) \|_2
\]

where \( f(i) \) gives the heave, surge and sway readings for the measurement at time \( i \). This can be interpreted in a similar way to geographical tortuosity as it relates to the complexity of the motion. The tortuosity of a transition is the length of the transition between two states, \( k \) and \( k' \). With \( n \) transitions of a known length from \( k \) to \( k' \), the average transition can be found by summing the length of all the paths from \( k \) to \( k' \), and dividing by \( n \). To visualise this statistic, we introduce a helix around the edge in the posture graph, where the periodicity of the helix indicates the tortuosity, (figure 7(c)).

The number of transitions can also be visualised using this method. In both cases, the higher period of the helix indicates regions of interest to the biologists, as a large number of transitions indicates a common movement, and or a high tortuosity indicates greater energy output (a relation between energy expenditure and accelerometry readings was shown by Halsey et al. [HSH*08] and further investigated by Wilson et al. [WWQ*06]).

Actual transitions in the data can also be viewed by brushing the posture graph to display transitions to and from a state. These transitions can be viewed with attribute overlays or colouring, as illustrated by figure 7(d) where the graph is combined with the pressure overlay (e.g. figure 4(b)).

5. Domain Expert Review

From a biologist’s perspective, the scatterplot serves the critical function of presenting multidimensional data in an intuitive manner. Typically for biologists, analyses of multiple parallel-logged transducer data (figure 1) require complex statistics where trends are identified via numbers which need detailed systematic inspection.

Acceleration data are particularly intractable in this re-
gard since axes covary and patterns are exceedingly intricate. One recent suggestion for the identification of behaviours using tri-axial accelerometry data was to adopt a hierarchical approach [SWQ+08]. However, these authors made no suggestion as to how this should be done and others following this approach appear to have simply worked their way manually through the data (e.g. Gomez-Laich et al. [LWQS08]). In fact, the scatterplot abides by exactly the suggestion of Shepard et al. [SWQ+08] but actually shows the frequency and extent of the posture groupings (either by the histogram function, figure 6, or by point density, figure 3) and is able to do so as an obvious function of other recorded parameters (e.g. figure 4) for any selected subsets of the data incorporating hugely different time periods (figure 4(b), figure 5).

Further, the approach shows precisely the importance of animal posture in different behaviours and forces workers to consider body attitude more thoughtfully. This has consequences for biological insights that go beyond simple recognition of particular behaviours such as ‘flying’, ‘sitting on the water surface’, ‘diving’ (see figure 7). For example, figures 4(b) & 4(d), which treat Imperial Cormorant diving behaviour, immediately show the consistency of the descent angle over multiple dives (the ‘rod’ extending from the sphere’s South pole) and the sinusoidal variance in the return-to-the-surface angle (the ‘diffuse cloud’ emanating from the sphere’s North pole), suggesting the birds may be slowing their ascent to the surface, possibly to reduce the likelihood of the bends (directly answering the open research question posed by Sato et al. [SNK+02]). The pattern is made clear by consideration of a single dive in figure 5, which also highlights how the body angle of the bird during the bottom phase of the dive (dark red region to the left of the sphere) is much more variable and not as horizontal as it is at the surface (small blue area on the left hand side of the sphere’s equator). This variance in body posture during the bottom of the dive presumably reflects the complexities of prey search and pursuit. Here, the sphere describes posture changes over time precisely, which immediately encourages an analysis of the tortuosity of the ‘track’ to see how patterns may relate to prey search and capture. To our knowledge, this has never even been considered before in the wealth of literature of body acceleration in diving animals (e.g. Yoda et al. [YSN+99], Lovvorn et al. [LWK+04], Sato et al. [SWT+07]).

It is thus with great excitement that these visualisations have been adopted by the Swansea University Smart Tag research group, and we fully expect fundamental new insights into the way animals operate that we could not have conceived before.

6. Conclusion

This work has presented and explored a new approach for the treatment of animal-acquired tri-axial acceleration data that encodes all three acceleration lines into a single point whose position in space moves according to acceleration. Other, simultaneously-recorded or derived parameters can be incorporated into the visualisation by colour mapping or by increasing the distance of the normalised acceleration vector from the sphere. Previously, visualisation was generally limited to multiple 2D time-series plots of the data. Our methods add increasing ability to first detect postures (spherical scatterplot), then behaviours (overlays) and finally, statistics of behaviours (spherical histogram and posture state graph), directly mirroring the methodology of the biologists and greatly simplifying data exploration, analysis and hypothesis generation through effective visualisation. Additional examples are shown in figure 8 and an accompanying video. These methods afford biologists a facility to examine up to six dimensions simultaneously and provides a powerful and intuitive representation that has already revealed hitherto unconsidered insight into animal behaviour. Furthermore, we have incorporated very positive domain expert feedback.

Despite this rapid progression we feel this application of accelerometry has far more to offer and we have only scratched the surface of understanding the challenges of its visualisation. The future direction is to try out more visualisation techniques and continue improving animal behaviour identification with a goal to achieving a higher level description of animal motion.

Acknowledgments Thanks to Niko Liebsch for 8(a). Ed Grundy was initially funded by an ESF grant number 061021WA1 arranged by IT Wales, and later by a donation to the IES by Wildlife Computing. Emily Shepard is funded by a Natural Environment Research Council (NERC) studentship (award no. NER/S/A/2005/13416A). Device development was supported by a Rolex Award for Enterprise awarded to Rory Wilson.

References


Figure 8: Visualisations of accelerometry from other animals. (a) A Leatherback Turtle with a Daily Diary device attached to the carapace, (b) Four hours of the turtles' behaviour visualised with our method. Two minutes of rolling behaviour can be seen as deviations from the vertical column. (c) Nineteen hours of albatross data showing soaring behaviour: the bird rolls its' body to fly efficiently for long periods, resulting in an arc across the top of the sphere. (d) Pressure overlay and colouring of 16 hours of data from other animals. (a) A Leatherback Turtle with a Daily Diary device attached to the carapace, (b) Four hours of the turtles' behaviour visualised with our method. Two minutes of rolling behaviour can be seen as deviations from the vertical column. (c) Nineteen hours of albatross data showing soaring behaviour: the bird rolls its' body to fly efficiently for long periods, resulting in an arc across the top of the sphere. (d) Pressure overlay and colouring of 16 hours of data from other animals.


