iBehave – application of supervised machine learning to behaviour analysis

J.A. Heward, P.A. Crook, T.C. Lukins, and J.D. Armstrong
School of Informatics, University of Edinburgh, Edinburgh, UK, douglas.armstrong@ed.ac.uk

Overview

Automatic recognition and classification of behaviour in laboratory animals is essential if behaviour research is to keep pace with other biological domains where the use of high throughput, data rich platforms are rapidly accelerating. In behaviour genetics studies and in the CNS drug discovery sector, behavioural throughput is already a rate limiting. There exists a number of approaches and in several cases implemented solutions for many common laboratory behaviours. In addition to fully automated systems there are partial solutions that maximise investigator productivity and enhance the accuracy and quantity of data returned.

The iBehave project is examining the application of supervised machine learning methods in this domain. The goal in simple terms is as follows. Given tracked data for an experiment, can we exploit the expertise of a human observer to train a computer algorithm to classify behaviour? We will present an overview of the current system along with demonstrations.

The software employs supervised machine learning methods to classify the behaviour of multiple complex objects in video footage. The principle advantage of our approach is that the system can be trained by the end-user to detect novel behaviours. This is achieved through training the system using selected exemplars identified by the end-user in pre existing footage. The software extracts coordinates and parameter values which uniquely distinguish the set of given exemplars. Thereafter, each behaviour that the system has been trained on can be automatically detected with an expression of confidence. The method can be to applied to footage containing both single and multiple articulated objects.

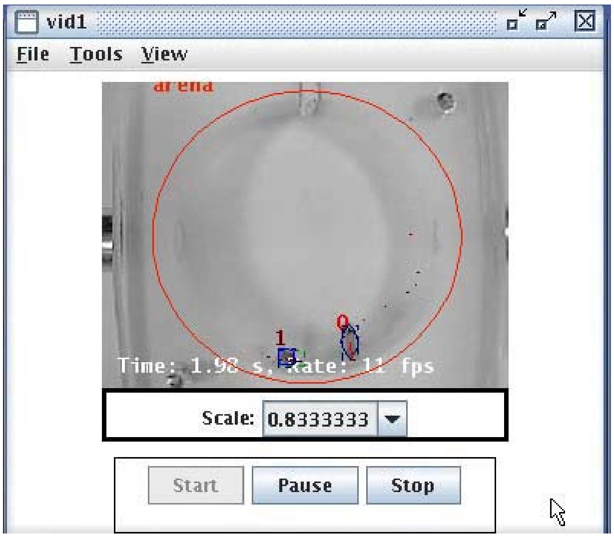

Figure 1 *Flytracker application following two adult fruit flies.*

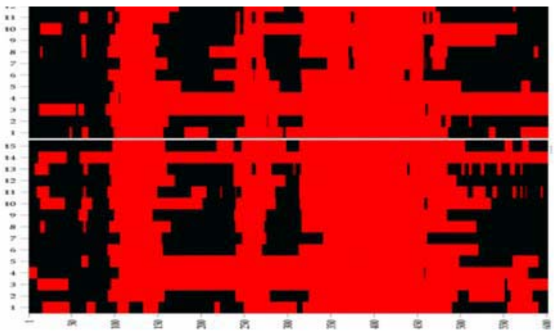

***Figure 2**. 15 experts annotating the same 600 video frames. Red indicates the behaviour occurring, black no behaviour. The top 15 rows show the first presentation the lower 15 rows the same video rotated through 180 degrees and represented later. Inter-observer variation is much higher than intra-observer variation.*

Design

The methods require three conceptual components: Parameter Extraction, Data Classification, and Annotation. Our main activity is in the Data classification and Annotation components. However, detailed parameter extraction is essential to the success of the programme.

Parameter Extraction

In the small animal behaviour measurement scenario, the software receives raw data in the form of video footage. It is first necessary to reduce the data to a set of key parameters that model the behaviour of the objects under observation. In this case, the key parameters are coordinates describing the shape, relative movement and deformation of body parts over time. Extracting these parameters reliably, across a range of experimental conditions, is a highly non-trivial challenge.

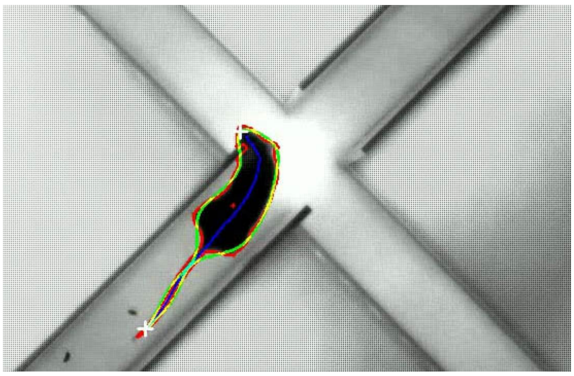

Figure 3. *Rodent behaviour requires additional parameters such as the body contours to be extracted.*

Data Classification

The operator specifies the subject behaviours by clicking start and stop to define exemplars. The key parameters pertaining to these exemplars are stored in the software and converted into classification rules describing the global boundary conditions (of the subject behaviours).

Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, August 26-29, 2008)
Eds. A.J. Spink, M.R. Ballintijn, N.D. Bogers, F. Grieco, L.W.S. Loijens, L.P.J.J. Noldus, G. Smit, and P.H. Zimmerman

Annotation

The software automatically screens new footage against the existing classification rules to annotate episodes that appear to contain the subject behaviours. Measures for statistical confidence are returned with mathematical expressions describing the classification rules employed.

Case Studies

We created a test dataset using the courtship ritual of the fruit fly *Drosophila melanogaster* for a first case study. 15 experts from a number of labs across the world were recruited and all annotated the same selection of videos which included obvious as well as difficult examples of the behaviour in question (courtship index). We then tracked the x and y co-ordinates of each individual's centre of gravity as well as the angle of orientation (see figure 1), acceleration and velocity (both linear and angular) . From these additional relative parameters are calculated.

A total of 10 minutes of footage was shown to each expert comprising nine one minute clips, of 600 frames each, taken from nine different previously recorded five minute videos were used to create a ten minute video (with a 10 frame delay between end/start of consecutive clips). Only nine clips were needed because the tenth clip was an inverted version of the fourth. The clips were of wildtype male and female *Drosophila* in a cylindrical plexiglass 20mm diameter x 5mm deep observation chamber. We compared the annotation across the experts (figure 2) which showed broad, but not exact inter-observer agreement. The video that was shown twice (albeit transformed) confirmed that intra-observer variation was lower than inter-observer variation.

80% of the annotated data was sampled to train using a decision tree with 20% held aside for validation. This was repeated for 5 random samples giving a final accuracy of 84.88% (st dev 2.54). This represents an 'average expert'. Training accuracy using single experts returns a much higher accuracy (of the order 94%).

Discussion

The iBehave method has been successfully applied to fly courtship. In particular it has highlighted the difficulties in getting multiple experts to agree on a common behavioural interpretation. We have recently extended the system to look at rodent behaviour and added additional parameters for the more complex body shape (figure 3). This work is on-going and will be presented at the meeting.