Automated distinguishing of mouse behavior in new environment and under amphetamine using decision trees

A. Konushin[1], E. Lomakina-Rumyantseva[1], D. Kropotov[2], D. Vetrov[1], A. Cherepov[3], and K. Anokhin[3]
[1]*Graphics and multimedia lab, Moscow State University, Moscow, Russia. {ktosh, vetrovd}@graphics.cs.msu.ru, lr2kate@gmail.com*
[2]*Dorodnicyn Computing Center of the Russian Academy of Sciences, Moscow, Russia, dkropotov@yandex.ru*
[3]Department of Systemogenesis, *P.K. Anokhin Institute of Normal Physiology, Moscow, Russia, k_anokhin@yahoo.com*

Abstract

Traditional activity measures do not provide a clear discrimination between mouse behavior in novel environment or under various psychomotor stimulants like *d*-amphetamine [1]. We propose a new approach based on machine learning. A decision tree classifier is trained on a set of mouse trajectories. Only x- and y- coordinates are currently used. After classifier is trained, the classification can be applied in real-time with respect to the output of video tracking system during experiment. The developed approach has shown promising results on distinguishing effects of new environment and d-amphetamine according to mouse behaviors.

Keywords

Video tracking, decision trees, novel environment, amphetamine, machine-learning.

Introduction

In recent years automated home cage systems have been introduced as a solution to the need of high-throughput behavior screening procedures [2]. Many of such systems are equipped with video tracking modules. However, traditional activity measures cannot automatically distinguish the effects of different drugs and environment, leaving the task to the human observer. This significantly limits the throughput capacity of such systems, because automated cage systems provide information 24 hours a day, generating an enormous volume of data.

Several new activity-based measures were recently proposed, e.g. [3], but they are based on pure observation and heuristically constructed. We propose a novel approach based on machine learning. This approach gives us a general framework for automatic development of classifiers for distinguishing the effects of various drugs and stimuli.

Proposed method and experiments

We had three types of recorded mouse behavior: mouse in a new environment (class 1), under effect of d-amphetamine (5 mg/kg) (class 2) and in familiar environment (class 3). Totally, we had 15 mice and 15 minutes (25000 time samples) behavior record of each, 6 mice from class 1 and class 2, and 3 mice from class 3. In order to increase data set we split each path into four equal parts thus obtaining enough representative samples (60 in total).

To design classification algorithm we computed a number of features according to mouse locomotor motion in the cage. After several tests the following three key features of path were identified: mean speed (v_mean), speed standard deviation (v_std), and correlation coefficient between speed and acceleration. The latter was computed in the following way:

$$\text{Corr\_coeff} = \frac{1}{n-4}\sum_{i=3}^{n-1}(v_i - Ev)(a_i - Ea) \Big/ \sqrt{DvDa},$$

where $v_i = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}$ - speed at the i-th time point, $a_i = v_i - v_{i-1}$ – acceleration at the *i*-th time point, $Ev = \frac{1}{n-1}\sum_{i=2}^{n} v_i, Ea = \frac{1}{n-2}\sum_{i=3}^{n} a_i, Dv = \frac{1}{n-2}\sum_{i=2}^{n}(v_i - Ev)^2, Da = \frac{1}{n-3}\sum_{i=3}^{n}(a_i - Ea)^2$.

We used decision trees CART algorithm [4] which gave us the following decision rule:

> **If** ((corr_coeff <= 0.51425) **and** (v_mean <= 0.00245))
> **Then** prediction = class 3 (probability = 1.000000)
> **Else if** ((corr_coeff <= 0.51425) **and** (v_mean > 0.00245) **and** (v_std <= 0.00805))
> **Then** prediction = class 1 (probability = 1.000000)
> **Else if** ((corr_coeff <= 0.51425) **and** (v_mean > 0.00245) **and** (v_std > 0.00805))
> **Then** prediction = class 2 (probability = 1.000000)
> **Else if** (corr_coeff > 0.51425)
> **Then** prediction = class 2 (probability = 0.869565)

The right classification in leave-one-out mode was achieved in 95% cases (57 of 60). This result proves that the behavioral features mentioned above discriminate reasonably well these states of mouse.

Note that the features used can be computed in real-time by using only part of mouse trajectory up to the current moment. This makes it possible designing on-line version of classifier capable to estimate the probability of each class at every time moment. Let x_j be the number of time samples at which the mouse was classified to class *j*. Then the probability that mouse belongs to given class *k* can be computed as $p_k = (x_k + c)/(x_1 + x_2 + x_3 + 3c)$. Here the constant c serves for preventing large fluctuations of probability in the beginning of trajectory. In our experiments we put $c = 50$. Figure 1 shows the typical changes of probabilities which are computed for each time sample. Using diagrams similar to figure 1, we can observe classifier's performance and make conclusion in a real-time mode.

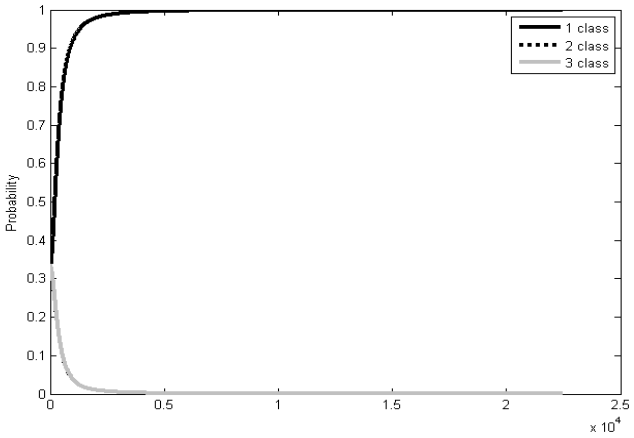

Figure 1*. Probabilities of different classes for particular mouse for all time moments.*

Conclusion and future work

Our proposed method has shown promising results for on-line classification of distinguishing the effects of amphetamine and new environment. We plan to introduce new features based on nose and tail base point tracking, and automatically labeled behavioral acts. Also on-line classification will be tested on long experiments for recognition of different doses of d-amphetamine and vanishing of drug or other stimuli effect.

References

1. Cherepov A.B., Lukashev A.O., Novosyolov I.A., Rayevsky K.S., Anokhin K.V. (2002) Relationships between novelty-versus d-amphetamine- induced behavior and brain c-fos expression in two inbred mouse strains: a functional systems analysis, *Measuring Behavior* 2002.
2. Spruijt B.M., DeVisser L.(2006) Advanced behavioral screening: automated home cage ethology. *Drug Discovery Today: Technologies*., **3(2)**, 231-237.
3. Kafkafi.N, Elmer G.I. (2005) Activity density in the open field: a measure for differentiating the effect of psychostimulants. *Pharmacology, Biochemistry and Behavior*, **80**, 239–249.
4. Breiman L., Friedman J., Olshen R., Stone C. (1984) Classification and Regression Trees. *Belmont, California*.