Depth-map Images for Automatic Mice Behavior Recognition

João P. Monteiro[1], Hélder P. Oliveira[1], Paulo Aguiar[2], Jaime S. Cardoso[1]

[1]INESC TEC (formerly INESC Porto) and Faculdade de Engenharia, Universidade do Porto, Porto, Portugal
[2]Instituto de Biologia Molecular e Celular and Centro de Matemática, Universidade do Porto, Porto, Portugal

***Abstract*— Animal behavior assessment plays an important role in basic and clinical neuroscience. Although assessing the higher functional level of the nervous system is already possible, behavioral tests are extremely complex to design and analyze. Animal's responses are often evaluated manually, making it subjective, extremely time consuming, poorly reproducible and potentially fallible. The main goal of this work is to evaluate the use of the recently available consumer depth cameras, such as the Microsoft's Kinect, for detection of behavioral patterns of mice. The hypothesis is that the simultaneous capture of video and depth information, should enable a more feasible and robust method for automatic behavior recognition, over previously available systems. We present some preliminary results regarding initial steps of the project, namely depth based mouse segmentation.**

Keywords- animal behavior; depth sensors; image segmentation

I. Introduction

A great effort has been developed by disciplines such as neurosciences and pharmacology, trying to understand the complex relationship between genes and behavior [1]. In this sense, animal experimentation remains a key instrument, and among animals used in research, mice can be recognized as one of the most important models [2]. They tend to be used for analysis of patterns of behavior of targeted and chemically induced mutations, often evaluated manually by direct observation or by analysis of video captured for this purpose.

The demand for automated methods of mice behavior analysis in laboratory arises primarily in order to resolve problems related not only to time and cost, but also reproducibility inherent in the assessment process conducted by people. Additionally, the availability of such systems, introduces the possibility to rethink behavior tests themselves. The typical testing time scale can be easily extended, and thus diversify the behaviors analyzed and the context of evaluation. Automated analysis of mice behavior constitutes a challenge due to a large number of factors, including the huge variability of the conditions of behavior tests, or the generic problem of behavior recognition itself.

In this paper we introduce an ongoing project that aims to develop a general-purpose, automated, quantitative tool for mouse behavioral assessment, which takes advantage of recent advances in computer vision systems. A wide range of behavioral tests are available for different experimental applications [3]. However, it is possible to identify open-field arena and observation of mice within their home cage, as two of the most accessible and widely used apparatus.

Regarding advances in computer vision, namely the research using depth images for object recognition, which has consistently presented several advantages over two-dimensional intensity images and the potential for greater recognition accuracy [4]. However, earlier depth sensors were expensive and difficult to use [5]. The task has been greatly simplified by RGB-D cameras, recently introduced at large scale and low cost in the market. RGB-D cameras, such as Microsoft's Kinect, are sensing systems that consist of one RGB camera and a depth sensor. The Kinect depth sensor is based on an infrared laser projector combined with a monochrome CMOS sensor, which provides depth information for each pixel at a frame rate of up to 30 fps.

II. Related Work

Automation of behavioral tests started with the use of electro-mechanical devices for experimental control when specific action-reaction or stimulus-response relations had to be quantified [6]. Though these systems can be used effectively to monitor the locomotor activity, fail to understand more complex behaviors. In this sense the visual analysis is presented as a potentially decisive and powerful supplement.

Existing vision-based methods typically use standard video images and extract motion features, adapting previous work for recognition of human actions and motion. There are several commercially available systems, which combine video-tracking technology with image analysis methods to characterize mice activities, such as PhenoTyper[1] by Noldus, HomeCageScan[2] by CleverSys or SmartCube[3] by PsychoGenics.

Besides the commercial systems, it is also worth mentioning here an open source software made available in 2010 [7]. It uses a machine learning algorithm for classifying every frame of a video sequence. Many other systems have been developed, the referred work extended simpler video-tracking based approaches (e.g. [8]), in order to allow the analysis of finer animal activities such as grooming, sniffing or rearing. It is however possible to find some limitation in the current release, such as, lack of characterization of social behavior or restrictions to camera pose, lighting conditions and mice color.

III. Kinect Based Method

Our under development system may be briefly described as follows: features are computed to convert an input video sequence composed of depth and color images, into a representation which will later be used for the automated recognition of the animal's behavior by a statistical classifier.

We started by video recording a singly housed mice, using a Kinect properly calibrated [9], from a top view of both open field test arena and home cage housing (Figure 1). Two mice behaving differently were used for these experiments. These videos, corresponding to over one hour of recording, were manually annotated labeling every frame of each video

Hélder P. Oliveira would like to thank Fundação para a Ciência e Tecnologia (FCT) – Portugal for the financial support for the PhD grant with reference SFRH/BD/43772/2008

[1] http://www.noldus.com/animal-behavior-research/products/phenotyper
[2] http://www.psychogenics.com/smartcube.html
[3] http://www.cleversysinc.com/products/software/homecagescan

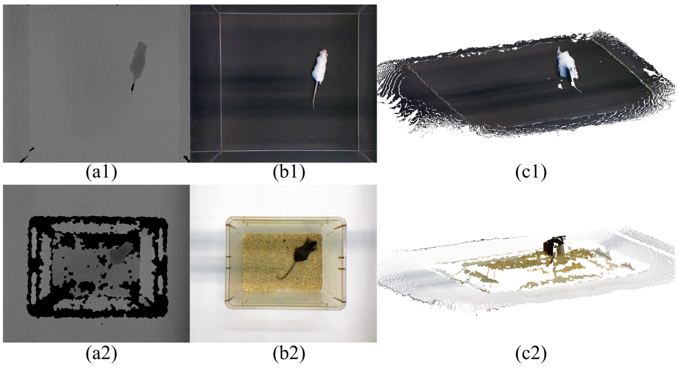

Figure 1. Example video acquisitions with kinect system of a mouse in a openfield arena (1) and home cage (2): (a) - depth map, (b) - RGB image and (c) - RGB-D fusion

sequence with a behavior of interest: walking, resting, grooming, rearing, and digging.

Seeking accuracy and reduction the chance of mismatching features, some work was done on image segmentation. It is essentially this effort, which is discussed from this point forward. Its goal was to robustly be able to extract masks surrounding the mice against the involving arena. In order to test several segmentation methods, was performed an additional manual annotation of mice spatial localization in 41 depth map images selected from different videos in different conditions of acquisition. As a gold standard of global thresholding, Otsu's method [10] was used. Other used method was the Gaussian mixture background model (GMBM) [11], which takes into account the pixel's recent history. As a simpler approach it was assumed that the background is static (Fixed Background Model - FBM). In that way, every new frame is subtracted from the background model previously obtained, and the resulted difference values of each pixel give the information for segmentation. It was also considered local background segmentation approach used in [12] (Local). It proposes that a density image is defined by transforming the depth information on the XY plane to the XZ plane. In the new image each column in is the histogram of the corresponding column in the original image. The threshold curve is then computed as the shortest path from one horizontal margin to another of the density image, where the cost of each pixel is its frequency value.

While the true positives (TP) give the number of correctly detected foreground pixels, the true negatives (TN) give the number of correctly identified background pixels. In contrast the false negatives (FN) are pixels that are falsely marked as background, whereas false positives (FP) are falsely detected as foreground. Hereupon, the measures true positive rate (TPR), given by TPR = TP/(TP + FN), and the false positive rate (FPR), given by FPR = FP/(FP + TN), were computed.

Table 1 shows that the best results were verified for FBM method due to the existence of a general stability of the background in the images analyzed. The Local method showed issues dealing with complex backgrounds and uniformity of cost criteria for different situations as open field and home cage arenas. Since in our samples, the mice area was much smaller than the background, the presence of different objects at different depths (substrate materials covering the bottom of the cage) caused wrong classifications by Otsu method. Occasional occurrences of long periods of immobility lead to the failure segmentations by GMBM method.

TABLE I. RESULTS FOR DEPTH BASED MOUSE SEGMENTATION

	Segmentation methods			
	GMBM	***Otsu***	***Local***	***FBM***
True positive rate	0.30	1.00	0.56	0.68
False positive rate	0.00	0.95	0.02	0.00

IV. CONCLUSIONS

Though the use of automated approaches has been documented for mice behavior recognition, such systems are not widely used, present limitations and may be cost prohibitive.

Segmentation should be further tweaked although some interesting preliminary results. It may also be noted that depth map images from Kinect are typically noisy and incomplete, mostly due to occlusion, relative surface angle and material, making us to consider Kinect RGB images an essential complement to depth maps (Figure 1).

Further work must be done in order to testify the importance of using additional information provided by depth maps, for the behavior recognition task. Namely, velocity and position features computation directly from depth map segmentation mask.

ACKNOWLEDGMENT

The authors are grateful to Instituto de Biologia Molecular e Celular for allowing us to use their facilities and mice to generate the training video database.

REFERENCES

[1] E. S. Lander, "Initial impact of the sequencing of the human genome," *Nature*, vol. 470, no. 7333, pp. 187–197, 2011.

[2] J. F. Cryan and A. Holmes, "The ascent of mouse: advances in modelling human depression and anxiety," *Nat Rev Drug Discov*, vol. 4, no. 9, pp. 775–790, 2005.

[3] S. Brooks and S. Dunnett, "Tests to assess motor phenotype in mice:a user's guide," *Nat Rev Neurosci*, vol. 10, no. 7, pp. 519–529, 2009.

[4] K. W. Bowyer, K. Chang, and P. Flynn, "A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition," *Computer Vision and Image Understanding*, vol. 101, no. 1, pp. 1–15, 2006.

[5] E. Stoykova, A. Alatan, P. Benzie, and et al, "3-D time-varying scene capture technologies," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 11, pp. 1568–1586, 2007.

[6] L. H. Tecott and E. J. Nestler, "Neurobehavioral assessment in the information age," *Nat Neurosci*, vol. 7, no. 5, pp. 462–466, 2004.

[7] H. Jhuang, E. Garrote, X. Yu, and et al, "Automated home-cage behavioural phenotyping of mice," *Nat Commun*, vol. 1, 2010.

[8] P. Aguiar, L. Mendonça, and V. Galhardo, "OpenControl: A free open-source software for video tracking and automated control of behavioral mazes," *J. Neurosci. Methods*, vol. 166, no. 1, pp. 66–72, 2007.

[9] D. Herrera C., J. Kannala, and J. Heikkilä, "Accurate and Practical Calibration of a Depth and Color Camera Pair," in *Conference on Computer Analysis of Images and Patterns*, 2011, vol. 6855, pp. 437–445.

[10] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans on Syst., Man and Cybernet.*, vol. 9, no. 1, pp. 62–66, 1979.

[11] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *IEEE Conference on Computer Vision and Pattern Recognition*, 1999, vol. 2, pp. 637–663.

[12] H. Oliveira, J. Cardoso, A. Magalhães, and et al, "Simultaneous Detection of Prominent Points On Breast Cancer Conservative Treatment Images," in *19th IEEE Conference on Image Processing*, 2012.