Monitoring Animal Behavior in the Smart Vivarium

Serge Belongie, Kristin Branson, Piotr Dollár and Vincent Rabaud
Department of Computer Science & Engineering, University of California, San Diego, USA

Abstract
In the course of modern medical research, it is common for a research facility to house thousands of caged mice, rats, rabbits, and other mammals in rooms known as *vivaria*. In any experiment involving a group of animals it is necessary to perform environmental and physiological monitoring to determine the effects of the procedure and the health of the animals involved. Such monitoring is currently performed by human observers, and for practical reasons, only a small subset of cages can be inspected for limited amounts of time. This short paper outlines the computer vision and machine learning technology behind the *Smart Vivarium*, a system for automated, continuous animal behavior monitoring. The Smart Vivarium will serve as an invaluable tool for medical researchers as it will make better use of fewer animals. Early discovery of sick animals will prevent diseases from spreading, and in general will lead to more efficient caretaking of animals.
Additionally, the proposed technology can serve as a powerful tool for monitoring sentinel cages in potential bioterrorism targets and chemical agent research facilities.

Keywords
mice, vivarium, tracking, behavior recognition

1 Introduction
A single vivarium can contain thousands of cages of mice, making close monitoring of individual mice impossible. Automated behavior analysis of individual mice will allow for earlier detection of abnormal behavior, and thus an improved level of animal care, as well as more detailed and exact data collection which will improve the efficiency of medical experiments.

Video surveillance of mice has the important characteristic of being non-intrusive; no modification to the environment is necessary. It is now feasible because of the recent availability of low-cost video cameras. Because of the huge number of medical experiments conducted on caged mice, this feasibility has led to a large amount of research on this problem; see for example the proceedings of *Measuring Behavior,* 1996 – present. To our knowledge, all current approaches (for example, [1]) require overhead mounted cameras. This simplifies the problem because the amount of occlusion is reduced. However, this kind of surveillance requires a specially designed cage, since in a standard mouse cage the overhead view is obstructed by the feeder and cage top (see Figure 1). In this short paper, we describe recent advances from our research group at the University of California, San Diego, on the problems of non-intrusive mouse tracking and behavior recognition from a side cage view.

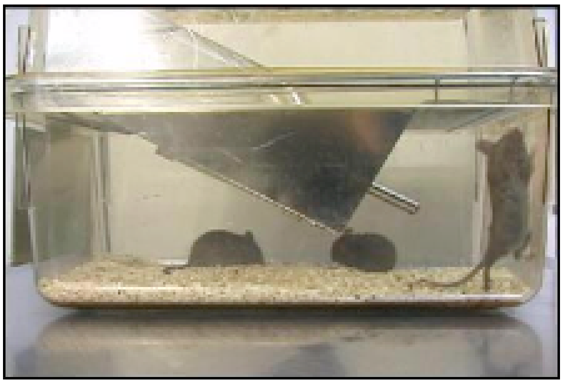

Figure 1. *Still frame captured from a video sequence of three mice (240 by 360 pixels). The metal container at the top of the cage holds food pellets and a water bottle. It also prevents the use of an overhead mounted camera. The bedding on the floor of the cage is the only dynamic part of the background, other than reflections.*

2 Tracking Multiple Mouse Contours

Our tracking research focuses on the problem of tracking the contours of multiple identical mice from video of the side of their cage; see Figure 1 for an example frame. Although existing tracking algorithms may work well from an overhead view of the cage, the majority of vivaria are set up in a way that prohibits this view.

This problem is uniquely difficult from a computer vision standpoint. Because mice are highly deformable 3D objects with unconstrained motion, an accurate contour model is necessarily complex. Because mouse motion is erratic, the distribution of the current mouse positions given their past trajectories has high variance. The biggest challenge to tracking mice from a side view is that the mice occlude one another severely and often. Tracking the mice independently would inevitably result in two trackers following the same mouse. Instead, we need a multitarget algorithm that tracks the mice in concert. As the number of parameters that must be simultaneously estimated increases linearly with K, the number of mice, the search space size increases exponentially with K [4]. Thus, using existing approaches to directly search the contour space for all mice at once is prohibitively expensive.

In addition, tracking individual mouse identities is difficult because the mice are indistinguishable. We cannot rely on object-specific identity models (e.g., [3]) and must instead accurately track the mice *during* occlusions. This is challenging because mice have few if any trackable features, their behavior is erratic, and edges (particularly between two mice) are hard to detect. Other features of the mouse tracking problem that make it difficult are clutter (the cage bedding, scratches on the cage, and the mice's tails), inconsistent lighting throughout the cage, and moving reflections and shadows cast by the mice.

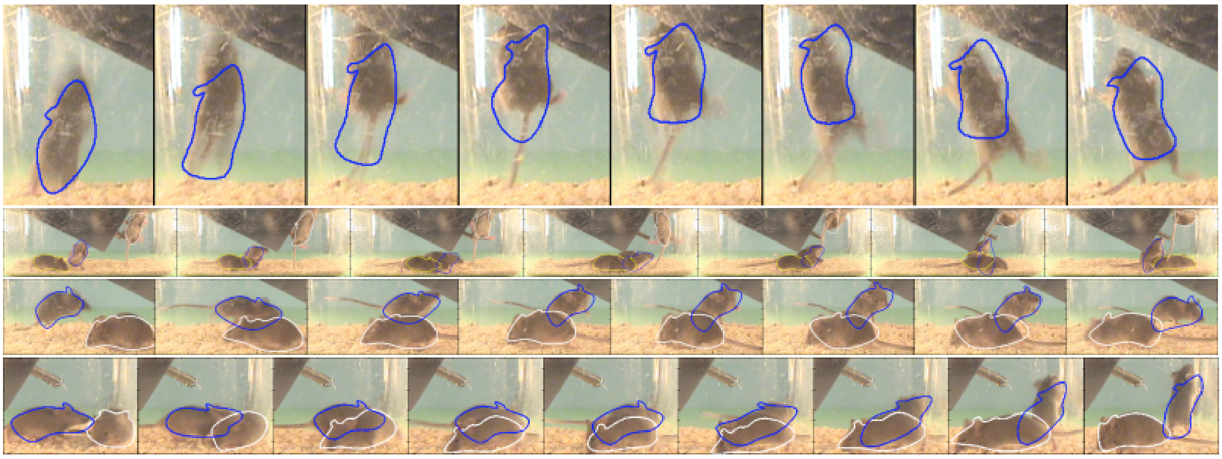

Figure 2. *Still frame summary of the successes of our algorithm. We plot the average affine transformation applied to the contour with the most total weight.*

We propose a solution that combines existing blob and contour tracking algorithms. However, just combining these algorithms in the obvious way does not effectively solve the difficulties discussed above. We propose a novel combination of these algorithms that accentuates the strengths of each individual algorithm. In addition, we capitalize on the independence assumptions of our model to perform most of the search independently for each mouse. This reduces the size and complexity of the search space exponentially, and allows our Monte Carlo sampling algorithm to search the complex state parameter space with a reasonable number of samples. Our algorithm works with a detailed representation of a mouse contour to achieve encouraging results.

We evaluated our blob and contour tracking algorithm on a video sequence of three identical mice exploring a cage, available at http://smartvivarium.calit2.net. This sequence contained 11 occlusions of varying difficulty. Summary still frames are shown in Figure 2. These results demonstrate the following strengths of our algorithm:

• Our contour tracking algorithm is robust to erratic mouse behavior – we never lose a mouse. For instance, we follow mice that jump, drop from the ceiling, and make quick turns and accelerations that are not fit by our simple dynamics model.
• Two contours never fit the same mouse.
• Our algorithm is rarely distracted by background clutter. This implies that our feature extraction methods and the blob and contour combination provide robust observation likelihoods. The only exceptions are when *both* algorithms make mistakes: when the blob tracker mistakes shaded bedding for foreground and the contour tracker fits to the edge of a tail.
• Perhaps the most impressive result is that our algorithm accurately tracks the mice through 7 out of 11 occlusions and partway through the other 4. This is because of the detailed fit provided by the contour tracking algorithm and its ability to use features available during occlusions.
• In general, our algorithm usually found very good contour fits outside of occlusions, much better than those obtained using contour tracking alone.

More information about our algorithm can be found in [2].

3 Behavior Recognition via Sparse Spatio-Temporal Features

After tracking and thus determining the positions and identities of each mouse, we focus our attention on recognizing their behavior. The method we have developed so far is successful in determining five basic behaviors: sleeping, drinking, exploring, grooming, and eating. Its design is very general and easily portable to other activities (like scratching or nesting that were omitted simply for lack of training footage).

Most behavior recognition methods developed in the computer vision literature are focused on human activity recognition (for a survey see [6]), but these methods are generally inapplicable to rodent behavior for the following reasons:

• With the exception of the eyes and ears, there are very few distinguishable features on the body of a rodent
• Rodent limbs are almost imperceptible in a single frame (except perhaps for the tail)
• Relevant activities can happen in a burst (e.g. less than a second)

Furthermore, many of the traditional approaches for human activity recognition assume simplifications that we cannot make, including:

• Simple backgrounds and little or no occlusion
• Small or no variation in the behaviors and posture of the subjects
• High resolution, clean data

Many of the problems described above have counterparts in object recognition. The inspiration for our approach comes from approaches to object recognition that rely on sparsely detected features in a particular arrangement to characterize an object, e.g. [8] and [7]. Such approaches tend to be robust to image clutter, occlusion, object variation, and the imprecise nature of the feature detectors. In short they can provide a robust descriptor for objects without relying on too many assumptions.

Our approach is based on describing a behavior in terms of local regions of interesting motion. Figure 3 shows example frames of two clips of a mouse grooming where the global appearance and motion are quite different but local regions of motion are quite similar.

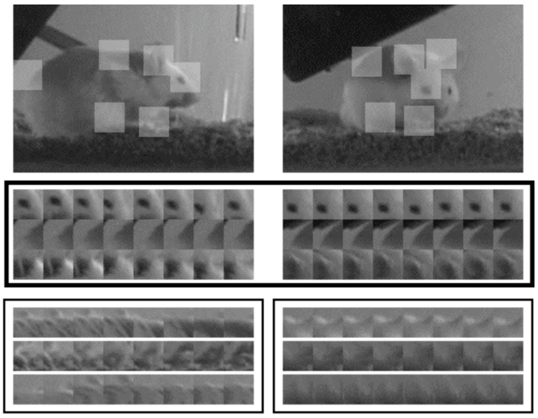

Figure 3. *Highlighted regions of local motion in example frames of two from videos of mice grooming. Six prominent areas of motion were extracted from each behavior. Note that although the posture of the mouse is quite different in the two cases, three of the six regions (shown in the top three rows) for each mouse are quite similar. The other three have no obvious correspondence although it is very hard to perceive what these are without motion.*

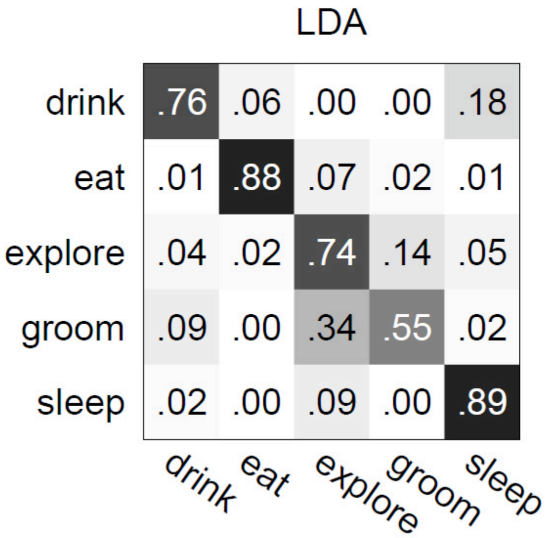

Figure 4. *The confusion matrix on test data shows the efficiency of our approach. Each row represents how our algorithm classified a given activity (using Linear Discriminant Analysis). The most confusion occurs when a mouse is grooming and the algorithm incorrectly classifies it as exploring.*

Although our method is still in development, it is already very reliable (cf. figure 4). These results were obtained with only a small amount of training data. As we increase the amount of training data, and incorporate a more robust model for activity, the accuracy and number of activities we can detect shall continue to increase.

4 Conclusion

We have described our recent progress in two areas of the *Smart Vivarium* project at UC San Diego: multiple mouse tracking and behavior recognition in the home cage environment. In our ongoing work we are collaborating with medical researchers to apply this technology for purposes of behavioral phenotyping in inbred mouse strains.

This work is funded by NASA GSRP grant NGT5-50479, the UCSD Division of Calit2 – the California Institute for Telecommunications and Information Technology – under the Smart Vivarium project, NSF Career grant #0448615, subcontracts B542001 and B547328 under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under contract No. W-7405-ENG-48, and the Alfred P. Sloan Fellowship.

References:

1. Twining, C; Taylor, C.; Courtney, P. (2001). Robust tracking and posture description for laboratory rodents using active shape models. In *Behavior Research Methods, Instruments and Computers, Measuring Behavior Special Issue*.
2. Branson, K.; Belongie, S. (2005). Tracking multiple mouse contours (without too many samples). *To appear, IEEE International Conference on Computer Vision and Pattern Recognition*, San Diego, CA.
3. Comaniciu, D.; Ramesh, V.; Meer, P. (2003). Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5):564-575.
4. MacCormick, J.; Blake, A (2002). A probabilistic exclusion principle for tracking multiple objects. *International Journal of Computer Vision*, 39(1):57-71.
5. Dollár, P.; Rabaud, V.; Cottrell, G.; S. Belongie. (2005, *submitted*). Behavior recognition via sparse spatio-temporal features. *IEEE International Conference on Computer Vision*, Beijing, China.
6. Gavrila, D. M. (1999). The visual analysis of human movement: A survey. *Computer Vision and Image Understanding*, 73(1):82–98.
7. Agarwal, S.; Awan, A.; Roth, D. (2004). Learning to detect objects in images via a sparse, part-based representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(11):1475– 1490.
8. Fergus, R.; Perona, P.; Zisserman, A. (2003). Object class recognition by unsupervised scale-invariant learning. *IEEE International Conference on Computer Vision and Pattern Recognition*, Madison, WI.