Rodent Behavior Analysis

Thomas C. Henderson
Vision-Based Behavior Analysis Seminar
Universität Karlsruhe (TH)
12 November 2003

1 How papers fit in survey framework

Rodent behavior analysis has as its goal the determination and temporal sequencing of rodent behaviors based on video image analysis. This fits in well with the structure of the human motion surveys, although it is applied to animals and not humans.

In terms of the first survey (Moeslund [2]), which divides the analysis into initialization of models, tracking (segmentation), pose estimation, and recognition, the work described here proposes a technique for the first three. In fact, the paper concludes that this approach does not provide a good mechanism for behavior recognition. Moreover, the assumptions of the paper match those of the survey in that there are restrictions on the camera and object, and the appearance of the object is modeled.

In terms of the details of the survey, the main results reported here will be on the paper by Twining et al. [4]. In that paper, the following features related to the survey are found:

- **Initialization**: a model of the initial pose is given in terms of a threshold (15 % brightest pixels), and shape features from thresholded image as matched to similar features from shape models.
- **Tracking**: relation established through landmark points; figure and ground separated based on thresholds; information reduction achieved through shape model; no frame to frame method is given
- **Pose Estimation**: direct model approach in sense of survey in that the contour (as a correlated set of feature points) provide for predict, match and update cycle.

2 What problem is posed?

The authors would like to develop:

1. a video image analysis system
2. that can handle deformable shape models

3. using a compact description

4. learned from examples

5. which can distinguish behaviors via postures

6. by performing automatic, robust segmentation of the object

7. and use posture and position to infer behavior.

The claim is that behavior can indeed be determined from posture, and that simple measures (e.g., location) do not correlate unambiguously with quantities of interest. To date, human observers are required to annotate video sequences and this has several drawbacks:

- labor intensive
- subjective
- biased
- gives label, but does not provide for quantitative details of behavior.

3 What solution is proposed

The authors propose to model points on the contour of the rodent with an active shape model; the reference values for the shape model will be learned from image data both for the shape as well as for local appearance (gray level) models. The model will be applied to images in a dynamic way which permits the feature points of the contour to migrate to the best locations.

4 What methods are used

The particular method used is an active shape model that uses the coordinates of the landmark features to produce a reference shape model vector set. These vectors are themselves subjected to a principal components analysis (rotated to a frame which minimizes the correlation between the vector elements), and only the top t eigenvectors are kept as the model (where t ensures that 98 % of the shape variability is modeled).

In addition, for each landmark point, the boundary normal at that location is determined and a vector is built from a gray level sample in the boundary normal direction. An appearance vector model (similar to the shape model) is built from this information.

In order to use the model, an initial location estimate is assumed, and every landmark point is migrated to a better gray level fit location until the process converges.

5 What contributions are made

The particular approach to modeling the shape of the rodent is novel, as is the use of the appearance model. The system's advantages are that it:

- can train on singular monochrome images
- was correct on all tested images
- can reliably extract the outline of the rodent
- provides a compact description
- provides an estimate of the goodness of the model fit.

In the experiments, 106 images were tested – all successfully.

6 What's missing, incomplete or unconvincing

There are several problems with the approach:

- cannot extract and classify posture
- the linear Active Shape Model does not always constrain the search space appropriately (some classes are not linear)
- computational complexity is high (there are 20 modes, that is, eigenvectors, kept which combined with the nonlinearity of the shape space means a lower frame rate).

Several issues are not directly discussed but may pose significant problems in any real application:

- The set of training images are most likely all from one setup and run; across training set issues need to be explored in which it may be determined if there existed some advantageous properties for the training data.
- How many landmark feature points are adequate? Good? Perfect?
- Why not use some landmark points that are not on the contour?

- During the search phase, there is a loop until convergence to the best fit; is there any guarantee that this convergence will occur? Can there be an oscillation between two nearby poses? Will it ever converge to a local minimum that is not the best fit?
- Why should the appearance model hold? The gray levels can change drastically over different regions of the image; why not use more stable cues (e.g., edges)? why not use color images and transform to YIQ or some other color space?
- Why are 20 modes selected? Maybe covering 75 % of the variability of the shape would be good enough?
- The method works least well on grooming postures; why not give details so that the success rate can be compared to other techniques (e.g., the paper by Rousseau [3] or Heeren [1])?

7 What is quality of work

The work seems to be of very good quality; this is indicated by the use of advanced shape modeling techniques (Procrutean shape analysis!), as well as the clear exposition of the use of PCA. In addition, the experiments are good and well-described, and the major drawbacks of the method are given - this indicates high intellectual honesty.

8 What are next questions to be studied

The authors mention that they intend to develop an approach to nonlinear models, but the other issues raised above would also need to be addressed. Since there was no followup paper that I could find, it seems to me that other approaches may be more successful (e.g., the Bayesian analysis particle filters, or Hidden Markov Models).

References

[1] D.J. Heeren and A.R. Cools. Classifying postures of freely moving rodents with the help of fourier descriptors and a neural network. *Behavior Research Methods, Instruments and Computers*, 32(1):56–62, 2000.

[2] T.B. Moeslund and E. Granum. A survey of computer vision based human motion analysis. *Computer Vission and Image Understanding*, 81:231–268, 2001.

[3] J.B. Rousseau, P.B. van Lochem, W. Melder, C. Costa-Florencio, W.H. Gispen, and B.M. Spruijt. Classification of rat behavior by a neural network. In *Measuring Behavior Conference*, The Netherlands, August 1998.

[4] C.J. Twining, C.J. Taylor, and P. Courtney. Robust tracking and posture description for laboratory rodents using active shape models. *Behavior Research Methods, Instruments and Computers*, 33(3):381–391, 2001.