Automated Recognition of Social Behavior in Rats: The Role of Feature Quality

Malte Lorbach
Universiteit Utrecht
Utrecht, The Netherlands
m.t.lorbach@uu.nl

Ronald Poppe
Universiteit Utrecht
Utrecht, The Netherlands

Elsbeth A. van Dam
Noldus Information Technology
Wageningen, The Netherlands

Lucas P.J.J. Noldus
Noldus Information Technology
Wageningen, The Netherlands

Remco C. Veltkamp
Universiteit Utrecht
Utrecht, The Netherlands

Abstract

We investigate how video-based recognition of rat social behavior is affected by two common tracking errors – animal misidentification and inaccurate localization of body parts. Our analyses show that correct identification is required to accurately recognize interactions. Precise localization of body points is beneficial for recognizing interactions that are described by a distinct pose. Including pose features only leads to improvement if the tracking algorithm can provide that data reliably.

1. Introduction

Rat social behavior is of interest for biologists who look for behavioral indicators for neurological and psychiatric disorders such as Huntington's disease. Currently, these studies involve laborious and error-prone manual coding of interactions and thus automating the coding is desired.

Video-based recognition of rat interactions typically requires three problems to be solved, namely: tracking and identifying the animals in the presence of occlusions, deriving meaningful features from these tracks, and classifying the features into interaction categories. Previous work on recognizing interactions has mainly focused on these steps in isolation, in particular by assuming perfect tracking when computing features [1, 2, 3]. The effects of mistaken identities and noisy tracking on the classification have received less attention. As a consequence, we yet lack the ability to trace back recognition errors to either tracking or classification.

With this paper we aim at unraveling the links between feature quality and recognition accuracy. We derive trajectory features from tracking data with varying degrees of common errors, and compare the performance using off-the-shelf classifiers. This work can be seen as a thorough investigation of the factors involved in automated rat social behavior analysis.

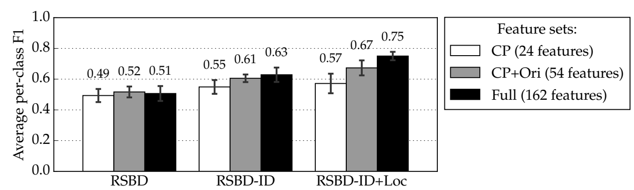

Figure 1. The average per-class F1 score using three different feature sets, tested on all three data set versions

2. Experiment

The Rat Social Behavior Data Set (RSBD), which we use throughout our analyses, comprises videos of two interacting rats from top-view perspective. Every frame of the videos is annotated with one of eight action classes.

We modify the quality of the data set in two ways. First, we incrementally correct two types of tracking errors leading to uncorrected (`RSBD`), identity-corrected (`RSBD-ID`), and location-corrected (`RSBD-ID+Loc`) tracks. Second, we derive three feature sets from those tracks capturing the rat's articulated body at varying degrees of detail.

To analyze the links between feature quality and recognition accuracy, we examine the effects of tracking errors on the classification accuracy alone (using the `Full` feature set) as well as in combination with the different feature sets. A Linear Discriminant Classifier has been selected out of six off-the-shelf classifiers as one of the best performing and yet simplest algorithms.

3. Results

Figure 1 shows the F1 scores (average over classes) of the combinations of data set versions and feature sets.

Please refer to and cite the full paper published under doi:10.1007/978-3-319-23234-8_52 in Image Analysis and Processing – ICIAP 2015.

There is an upwards trend across the data set versions irrespective of which feature set is used. In the uncorrected `RSBD` set, the F1 score remains at approximately 0.5 for all feature sets. In both identity-corrected `RSBD-ID` and location-corrected `RSBD-ID+Loc`, the F1 score increases with richer feature sets.

Not all interactions are affected by tracking errors in the same way. The accuracies are generally high for solitary actions and approaches (in which the animals are separated by definition). Contact interactions are not recognized well initially but improve gradually as errors are corrected.

The correction of identity swaps leads to two major improvements. Firstly, the confusion of *following* with *moving away* is largely resolved although some confusion persists. Secondly, virtually all *nape attacks* that had been mistaken as *following* are now corrected. Consequently, the recall of *nape attacking* improves although precision stays at a low level. Additionally correcting the body point locations increases the precision of *nape attacking*, and the recall of *pinning*. Confusions remain between these two classes and also between *following* and *approaching*. A number of small improvements across all classes eventually leads to higher average F1 scores at both frame and class level.

4. Discussion

On the overall performance level, we have seen that eliminating tracking errors leads to better classification. This pattern occurred for all tested classifiers, which suggests that the effect is indeed inherent to the underlying data and not to the classifier. We further showed that orientation and pose features are important for the recognition. If those features are correct, they lead to better classification. If they are not, that is, if the tracking algorithm fails to provide stable pose information, we induce the risk of overfitting to the noise in the features. As a consequence, the classification accuracy stagnates or even decreases. A potential way to overcome this limitation is to include more training data, which are particularly expensive to obtain. Moreover, when we trained the classifier with corrected data but applied it to uncorrected data, we failed to achieve competitive performance ($\mu_{\text{classes}} = 0.42$, $\sigma = 0.05$, 5-fold cross-validation). For that reason, we do not benefit from corrected, clean features as long as we cannot guarantee that we can generate them without expensive, manual intervention.

On the class level, we observed that the classes are affected differently by tracking errors and the choice of features. By which type of tracking error an interaction is most affected is determined by its characteristics. Since most of our interactions are sensitive to the correct role assignment, we see large gains in F1 score after correcting identity swaps. Clearly, maintaining the correct identities is a necessity for social behavior recognition.

Another characteristic of the interactions is how important the relative pose is for the recognition. We expect that the more an interaction is defined by the pose, the better it should be recognized if correct pose features are provided. We find supporting evidence in the results. *Nape attacking*, *pinning*, and *following* benefit most from the correction of body part locations and thus pose. Accordingly, adding uncorrected orientation and pose features results in only a small improvement (`RSBD-ID`: `CP` $\rightarrow$ `Full` $= +0.08$). We conclude that the accuracy of social behavior recognition can be improved by incorporating reliable orientation and pose features.

5. Conclusion

In this paper we investigated the effects of feature quality on video-based recognition of rat social behavior. We looked at the impact of two types of tracking errors – misidentification and inaccurate localization – as well as the type of features that are derived from the tracking data.

From the analysis of the classification accuracy across interaction classes, we observed that although correcting tracking errors improves the classification, each class is affected differently. Correctly identifying the animals is required to recognize virtually all interactions, whereas correctly tracking body parts has a larger impact on classes that are defined by a distinct relative pose. Hence, including orientation and pose features is advantageous under the condition that the tracking algorithm can provide them reliably.

We have further found that perfect tracking alone is insufficient for recognizing difficult, ambiguous behavior. Exploiting temporal context and reaction patterns alongside with features that go beyond 2D trajectories are directions that seem worth pursuing in the future.

Acknowledgments This work was funded by the European Commission under the 7$^{\text{th}}$ Framework Programme (FP7) as part of the Marie Curie Initial Training Network "PhenoRat" (GA no. 317259).

References

[1] T. Arakawa, A. Tanave, S. Ikeuchi, A. Takahashi, S. Kakihara, S. Kimura, H. Sugimoto, N. Asada, T. Shiroishi, K. Tomihara, T. Tsuchiya, and T. Koide. A male-specific QTL for social interaction behavior in mice mapped with automated pattern detection by a hidden Markov model incorporated into newly developed freeware. *J. Neurosci. Meth.*, 234:127–134, 2014.

[2] X. P. Burgos-Artizzu, P. Dollár, D. Lin, D. J. Anderson, and P. Perona. Social behavior recognition in continuous video. In *Proc. CVPR*, pages 1322–1329, 2012.

[3] L. Giancardo, D. Sona, H. Huang, S. Sannino, F. Managò, D. Scheggia, F. Papaleo, and V. Murino. Automatic visual tracking and social behaviour analysis with multiple mice. *PLoS ONE*, 8(9):e74557, 2013.

Please refer to and cite the full paper published under doi:10.1007/978-3-319-23234-8_52 in Image Analysis and Processing – ICIAP 2015.