Activity recognition for autonomous water drones based on unsupervised learning methods

Alberto Castellini^{*}, Giovanni Beltrame, Manuele Bicego, Domenico Bloisi, Jason Blum, Matteo Denitto, and Alessandro Farinelli

Verona University, Department of Computer Science, Strada Le Grazie 15, 37134 Verona, Italy, {alberto.castellini,manuele.bicego,domenico.bloisi,jason.blum, matteo.denitto,alessandro.farinelli}@univr.it, giovannialberto.beltrame@studenti.univr.it *Corresponding author

1 Introduction

Water monitoring involves the acquisition of data from rivers, lakes, and other bodies of water to gather knowledge about water properties. New technologies have recently started a transition from manual and passive monitoring to automatic and active monitoring, in which water drones autonomously move into catchments according to strategies that maximize the acquisition of information [2]. These intelligent systems collect large amount of data which need advanced computational methods to be analyzed, interpreted, and exploited for decision making [1].

A promising research area in this context concerns the automatic identification of events [9], activities [3,14], and situations [5,6,12,7] of interest from the analysis of large datasets collected by unmanned vehicles using artificial intelligence and statistical learning methods [13,10]. Here we follow this line of research and aim at developing an *activity recognition* system for unmanned vehicles involved in water monitoring. The main data sources for activity recognition can be grouped into two macro-categories, namely *vision-based* and *sensor-based* sources [3]. In this work we use only the second kind of sources. Moreover, while activity recognition is a concept often applied to human activities [3], this is a key features also for autonomous robots when the activities of others can not be communicated [14] or when such activities can not be perceived from the robots (e.g., because there are no available sensors to directly perceive such activity) or because sensors can be too expensive (e.g., detecting whether the water drone is navigating upstream or downstream in a river).

Specifically, our aim in this work is twofold: *i*) to support drone operations (*on-line decision making*), *ii*) to support the extraction of knowledge from the large sets of data collected by drones (*off-line data analysis*). The main contributions we present here are: i) the application of unsupervised statistical methods to activity recognition in the context of autonomous water monitoring, and ii) the empirical evaluation of a system for detecting a specific activity: "up-stream/downstream navigation" on real data. Our empirical evaluation provides promising results for the identification of this activity.

2 Material and Methods

System overview. This work is part of the H2020 EU project INTCATCH¹ which aims at developing innovative strategies for integrated water monitoring of catchments. Our low-cost aquatic drones, displayed in Figure 1, are equipped with sensors able to detect: i) GPS coordinates, i.e., latitude and longitude, ii) water properties, such as temperature, dissolved oxygen and electrical conductivity, *iii*) commands to propellers, *iv*) battery voltage. In the data acquisition phase, signals coming from different sources are integrated and synchronized, and new variables are possibly generated by fusing information from these sources. A data matrix of n variables and m time steps is thus generated which we aim to annotate with *activity labels*, such as, "the boat is navigating upstream". Each activity instance should be represented by the time interval in which the activity was detected and the subset of variables that provided information to identify the activity. Therefore, an activity instance should be represented by a sub-matrix, possibly after row rearrangements, containing specific information patterns. From this point of view our problem could be seen as a multivariate time-series segmentation [11,4] or a clustering problem [8].



Fig. 1. System overview: acquatic drones and available data.

Dataset. Data collection was performed in different environments and locations including rivers, fish ponds an lakes. Every dataset is a matrix of 13 *features*, namely *time*, *latitude*, *longitude*, *altitude*, *speed*, *electrical conductivity*, *dissolved* oxygen, temperature, battery voltage, heading, acceleration, command to propeller 1 and 2, and a certain number of observations depending on the duration of the data acquisition process. Different sensors have different sampling intervals, hence interpolation is performed during data preprocessing to merge all signals and obtain a common sampling interval of 1 second.

Activities. Aquatic drones face different activities during their missions, where an activity can be defined, in general, as a specific action that the drone is performing in a given state of the environment. Manual labeling was performed for five activities, namely, acquiring data in/out of water, upstream/downstream navigation, manual/automatic driving, navigation facing waves/no waves, blocked boat. Labeling was performed in a partial way, namely, experts analyzed georeferenced path images, videos of the acquisition phases and manual notes, and they labeled time intervals in which specific activities surely occurred. They left

¹ See the project website for further info http://www.intcatch.eu/

unlabeled (i.e., label 0 in Figure 2) time intervals in which activity occurrence was not completely sure. In this work we aim at assessing a specific activity, namely if the drone is moving upstream or downstream in a river.

Unsupervised learning methods for activity recognition. We performed clustering of sensors observations using Gaussian Mixture Models (GMMs) and Hidden Markov Models (HMMs). In both cases the Expectation-Maximization (EM) algorithm was used to learn the model. Inference was performed using the junction tree algorithm for GMMs and Viterbi algorithm for HMMs. Clustering performance was computed by purity. Purity is an external method of the extent to which clusters contain a single class, and it is computed as $P(C) = \frac{1}{N} \sum_{k \in K} \max_{d \in D} |k \cap d|$, where C represents a clustering, N is the total number of points, K is the set of clusters and D is the set of classes.

3 Results

GMMs and HMMs were trained on a dataset collected on a river, with a total of 3615 observations (multiple sensor readings) and 9 features, namely, speed, electrical conductivity, dissolved oxygen, temperature, battery voltage, heading, acceleration, command to propeller 1 and 2. During the entire mission, which lasted about 1 hour, the aquatic drone moved both upstream and downstream, as shown in Figure 2 (the stream direction is from left to right in the map). Moreover, sensor readings include data acquired before the drone was inserted into the water and after it was removed from the water.

We set the number of clusters (hidden states) to 3, in order to enable the detection of the target activity (i.e., upstream and downstream navigation) considering also the possibility of having measurements that show no specific water flow (e.g., when the drone was out of the water). Clusterings generated by GMMs and HMMs are graphically shown in points (b) and (c), respectively, of Figure 2. The value of purity for GMMs is 0.54 while for HMM we have 0.95. GMMs clearly detect the intervals in which the boat was out of the water, corresponding to cluster 3 (blue) and 1 (red) in the picture (low level of electrical conductivity identify such an activity). However, this model groups together all observations collected in water and it cannot distinguish upstream from downstream navigation. This behavior is well represented by the low value of purity. On the other hand, HMMs achieve good performance in the detection of both situations out of water (cluster 2/green points in Figure 2) and upstream/downstream navigation (cluster 1/red points and cluster 3/blue points, respectively). The high purity achieved by HMM confirms the good performance of this clustering.

In order to understand what characterizes upstream and downstream navigation in our dataset we analyzed model parameters. In particular we compared the distributions of each variable of the HMM model between upstream and downstream clusters (see box plots in the right-hand side of Figure 2) and we identified statistically significant differences in means using Student's *t*-test. We found that variables *heading*, *commands to propellers 1 and 2*, *dissolved oxygen* and *speed* were significantly different using a threshold of 0.05 for p-values. Differences in *heading* have intuitive interpretation since the boat had opposite directions while navigating upstream and downstream in the river. The *commands to propellers 1 and 2* were higher in upstream navigation (cluster 1) than in downstream navigation. This is because the boat needed full power to move against the water flow. The *speed* of the boat was also higher during upstream navigation. An interesting element of the analysis is the unexpected behavior of *dissolved oxygen* which had higher level during upstream navigation (i.e., $7.53 \ \mu g/L$) than during downstream navigation (i.e., $7.48 \ \mu g/L$). Since the water moved along the same path (only in different directions) during upstream and downstream navigation we suppose that the slight but statistically significant difference is due to the increased turbulence generated by the boat along the upstream path.

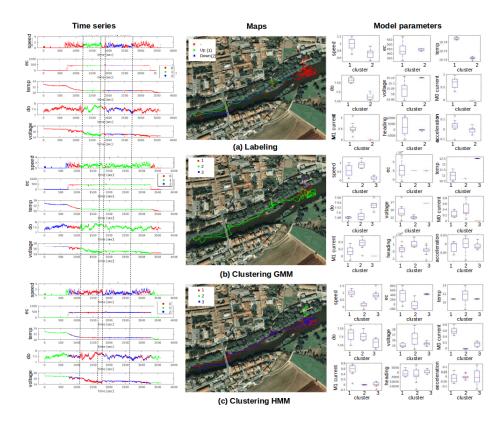


Fig. 2. Results for clustering based on GMMs (purity: 0.54) and HMMs (purity: 0.95) in a river. Left: segmented time series. Center: segmented paths in the map. Right: box plots of model parameters. Best viewed in colors.

4 Conclusions

This paper takes a first important step towards activity recognition for autonomous water drones. The main idea behind our approach is to employ statistical methods for multivariate time-series segmentation to perform activity recognition in this specific context. Our approach achieves promising results by correctly interpreting real data collected during data acquisition missions. Moreover, by observing the statistical distribution provided by the model it is also possible to identify the most representative variables for a given activity. Future work in this direction includes a more detailed analysis of our approach for different activities/datasets and an investigation to further analyse the interpretability of the models by human operators.

Acknowledgments

This work is partially funded by the European Unions Horizon 2020 research and innovation programme under grant agreement No 689341. This work reflects only the authors view and the EASME is not responsible for any use that may be made of the information it contains.

References

- U. Ahsan and A. Bais. A review on big data analysis and internet of things. In 2016 IEEE 13th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), pages 325–330, 2016.
- L. Bottarelli, M. Bicego, J. Blum, and A. Farinelli. Skeleton-based orienteering for level set estimation. In ECAI 2016 - 22nd European Conference on Artificial Intelligence, 29 August-2 September 2016, The Hague, The Netherlands - Including Prestigious Applications of Artificial Intelligence (PAIS 2016), pages 1256–1264, 2016.
- L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu. Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6):790–808, 2012.
- F. Duchêne, C. Garbay, and V. Rialle. Learning recurrent behaviors from heterogeneous multivariate time-series. Artificial Intelligence in Medicine, 39(1):25–47, 2007.
- M. R. Endsley. Toward a theory of situation awareness in dynamic systems. Human Factors, 37(1):32–64, 1995.
- M. R. Endsley. Theoretical underpinnings of situation awareness: a critical review. In M. R. Endsley and D. J. Garland, editors, *Situation Awareness Analysis and Measurement*. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2000.
- A. Farinelli, D. Nardi, R. Pigliacampo, M. Rossi, and G. P. Settembre. Cooperative situation assessment in a maritime scenario. *International Journal of Intelligent* Systems, 27(5):477–501, 2012.
- D. Hallac, S. Vare, S. Boyd, and J. Leskovec. Toeplitz inverse covariance-based clustering of multivariate time series data. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '17 (August 13-17), pages 215–223, New York, NY, USA, 2017.

- R. P. Higgins. Automatic event recognition for enhanced situational awareness in UAV video. In *MILCOM 2005 - 2005 IEEE Military Communications Conference*, pages 1706–1711 Vol. 3, 2005.
- A. Holst, B. Bjurling, J. Ekman, Å. Rudström, K. Wallenius, M. Björkman, F. Fooladvandi, R. Laxhammar, and J. Trönninger. A joint statistical and symbolic anomaly detection system: Increasing performance in maritime surveillance. In 15th International Conference on Information Fusion, pages 1919–1926, 2012.
- E. Keogh, S. Chu, D. Hart, and M. Pazzani. Segmenting time series: A survey and novel approach. In In an Edited Volume, Data mining in Time Series Databases. Published by World Scientific, pages 1–22. Publishing Company, 1993.
- 12. D. A. Lambert. Situations for situation awareness. In Proceedings of the 4th International Conference on Information Fusion, 2001.
- R. Laxhammar. Artificial intelligence for situation assessment. Master's thesis, School of Computer Science and Engineering, Royal Institute of Technology, KTH CSC, Sweden, 2007.
- D. L. Vail, M. M. Veloso, and J. D. Lafferty. Conditional random fields for activity recognition. In *Proceedings of the 6th International Joint Conference on Au*tonomous Agents and Multiagent Systems, AAMAS '07, pages 235:1–235:8, New York, NY, USA, 2007. ACM.