Unsupervised activity recognition for autonomous water drones

Alberto Castellini
Verona University, Italy
alberto.castellini@univr.it

Giovanni Beltrame
Verona University
beltramegiovannialberto@gmail.com

Manuele Bicego
Verona University
manuele.bicego@univr.it

Jason Blum
Verona University
jason.blum@univr.it

Matteo Denitto
Verona University
matteo.denitto@univr.it

Alessandro Farinelli
Verona University
alessandro.farinelli@univr.it

ABSTRACT

Accepted version of the manuscript. Please refer to https://dl.acm.org/citation.cfm?doid=3167132.3167396 for the final version. We propose an automatic system aimed at discovering relevant activities for aquatic drones employed in water monitoring applications. The methodology exploits unsupervised time series segmentation to pursue two main goals: *i)* to support on-line decision making of drones and operators, *ii)* to support off-line analysis of large datasets collected by drones. The main novelty of our approach consists of its unsupervised nature, which enables to analyze unlabeled data. We investigate different variants of the proposed approach and validate them using an annotated dataset having labels for activity "upstream/downstream navigation". Obtained results are encouraging in terms of clustering purity and silhouette which reach values greater than 0.94 and 0.20, respectively, in the best models.

CCS CONCEPTS

• **Computing methodologies** → **Artificial intelligence**; **Machine learning**; • **Applied computing** → **Environmental sciences**;

KEYWORDS

Activity recognition; water monitoring; autonomous drone; unsupervised learning; clustering; segmentation; time series analysis

ACM Reference Format:
Alberto Castellini, Giovanni Beltrame, Manuele Bicego, Jason Blum, Matteo Denitto, and Alessandro Farinelli. 2018. Unsupervised activity recognition for autonomous water drones. In *SAC 2018: SAC 2018: Symposium on Applied Computing, April 9–13, 2018, Pau, France.* ACM, New York, NY, USA, Article 4, 3 pages. https://doi.org/10.1145/3167132.3167396

1 INTRODUCTION

Aquatic drones are increasingly used for autonomous monitoring of catchments. In this context robotic boats must navigate rivers and lakes to acquire real-time data concerning important water parameters. While human operators are usually involved in such data collection activities, direct tele-operation of the drones is often not an option for an entire mission, hence autonomous navigation capabilities are required [3].

A promising research area in this context concerns the automatic identification of events [7], activities [1, 4] and situations [5] of interest from the analysis of large datasets collected by unmanned vehicles using artificial intelligence and statistical learning methods [8]. This paper follows this line of research and aims at developing an *unsupervised* activity recognition system for unmanned vehicles involved in water monitoring. Activities are here considered as states of the drone in the environment, such as, "the drone is navigating upstream" or "the drone is blocked". Manual data labeling is usually expensive and time consuming in this context hence automatic techniques that can extract states from unlabeled data represent crucial tools for water drone control and data analysis. The proposed approach exploits time-series segmentation methods to automatically detect time intervals in which data have similar properties: it represents a statistically grounded way to identify primitive states directly from sensor readings.

This paper provides three contributions to the state-of-the-art: *i)* a formalization of the activity recognition problem in the context of autonomous water monitoring; *ii)* a first unsupervised learning system based on Gaussian Mixture Models (GMMs) [2], Hidden Markov Models (HMMs) [9], K-Means (KM) [2] and Affinity-Propagation (AP) [6] for generating a model of water drone states from unlabeled datasets; *iii)* the successful evaluation of this model on the activity upstream/downstream navigation and a first interpretation of segments identified.

The next section introduces the system architecture and describes datasets, drone states and clustering setup. In Section 3 the methodologies are tested and evaluated. Section 4 outlines future developments.

2 MATERIAL AND METHODS

System overview. Aquatic drones, displayed in Figure 1, are equipped with sensors able to detect: *i)* GPS coordinates, *ii)* water properties, *iii)* commands to propellers, *iv)* battery voltage. Signals from different sources are integrated and synchronized. A data matrix of variables (rows) and time steps (columns) is thus generated which we aim to annotate with *state labels* (shown in the bottom of Figure 1).

Dataset. Data collection was performed in two different parts of a river. The first dataset, called ESP2, has 2831 samples (collected in 47 *mins*), the second dataset, called ESP5, has 3615 samples (collected in 60 *mins*). Sampling interval is of 1 *sec*. Both the datasets have 13 *features*, namely time, latitude, longitude, altitude, speed,

SAC 2018, April 9–13, 2018, Pau, France

ACM ISBN 978-1-4503-5191-1/18/04…$15.00
https://doi.org/10.1145/3167132.3167396

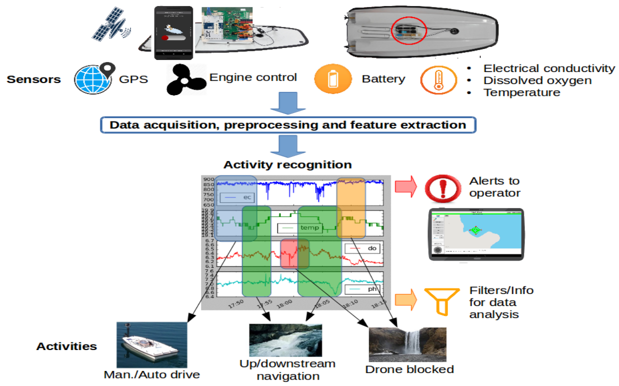

Figure 1: System overview: main elements of the proposed activity recognition system.

electrical conductivity, dissolved oxygen, temperature, battery voltage, heading, acceleration, command to propeller 1 and 2. From each dataset we generated two matrices used for model training: the matrix of *raw* data (RAW), containing all variables except time, latitude and longitude (namely, 10 variables in total) and the matrix of *processed* data (PRO) containing both the moving means and standard deviations of the variables in the raw matrix over a sliding window of 10 seconds (20 variables in total). Both normalized (NORM) and unnormalized (UNORM) versions of these matrices were used for model generation. Normalization was performed by scaling each variable to the range $[0, 1]$.

Activities. Aquatic drones perform different activities during their missions. Manual labeling was performed for five activities. Here we focus on upstream/downstream navigation (UDN) which is crucial in water monitoring because it influences sensor readings and therefore provides important contextual information for both decision making and data analysis. Labeling was performed in a partial way, namely, experts analyzed georeferenced path images and videos and they labeled time intervals in which specific states clearly occurred. They left unlabeled (i.e., label "-" in Figure 2) time intervals in which the state of the drone was not completely clear.

Clustering/segmentation setup. Sensor time series were processed via four clustering/segmentation techniques to determine groups of points with coherent behavior. **GMMs** [2] generated models having from 2 to 8 clusters. The algorithm was re-initialized 300 times and the model with maximal log-likelihood was used. Initial component means were generated by the k-means algorithm, initial mixing proportions were set to uniform, initial covariance matrices defined diagonal. Parameter learning was performed by the Expectation-Maximization (EM) algorithm (< 100 iterations). **HMMs** [2] generated models having from 2 to 8 hidden states. Observation models were set to single component multivariate Gaussian distributions (with one dimension for each observed variable). The initial state distribution was set to uniform, the initial transition matrix was set to random stochastic, initial means and covariance matrices were computed by k-means. The model was trained by the EM algorithm (< 20 iterations) and the Viterbi algorithm [2] was used to generate the most likely sequence of hidden states (i.e., drone states) given the observed sequence of sensor readings. **K-means** [2] used Euclidean distance $\|\cdot\|^2$, number of clusters between 2 and 8 and it was re-initialized 300 times. **AP** [6] used preference parameter from 30 to 180 (step 30) times the value of the median of the similarity matrix.

Performance evaluation. To assess the performance of our framework we employed two measures, *purity* and *silhouette*. Purity is a measure of the extent to which clusters contain a single class, and it is computed by formula $P(C) = \frac{1}{N}\sum_{k\in K}\max_{d\in D}|k\cap d|$, where C is a clustering, N is the total number of points, K is the set of clusters and D is the set of classes. Purity values close to 1 identify clusterings having almost one label for each cluster. Silhouette is an internal measure that contrasts the average distance to elements in the same cluster with the average distance to elements in other clusters. Given the i-th data point, it is computed as $S(i) = \frac{b(i)-a(i)}{\max\{a(i),b(i)\}}$, where $a(i)$ is the average dissimilarity of point i with all other data within the same cluster and $b(i)$ is the lowest average dissimilarity of point i to any other cluster, of which i is not a member. Values close to 1 indicate points belonging to perfectly compact and separated clusters.

3 RESULTS

We generated clustering models according to five dimensions of analysis, namely, *i)* clustering methods GMM, HMM, KM, AP, *ii)* datasets ESP2, ESP5, *iii)* RAW and PRO data, *iv)* NORM and UNNORM data, *v)* number of clusters from 2 to 8 for GMM, HMM and KM, and preference coefficient from 30 to 180 (step 30) for AP. A total of 324 models were generated and evaluated on purity (related to the detection of upstream/downstream navigation) and silhouette.

For each experiment we selected the four models, one for each clustering method, having the best performance in terms of purity. Table 1 shows the performance of selected clusterings and Figure 2 the best segmentations, in terms of both purity and silhouette, for each experiment. In fact, we used the mean silhouette of the clustering to select the most significant clusterings among those having highest purity. The clusterings having best performance are C4 for ESP2 and C7 for ESP5 (see bold values in Table 1). Their purities/silhouettes are 0.94/0.20 and 0.98/0.21, respectively.

	Method	Data	Norm.	# Cl.	Pur.	Sil.
			ESP2			
C1	**GMM**	RAW	UNORM	8	0.97	0.01
C2	**HMM**	PRO	UNORM	2	1.00	0.00
C3	**KM**	RAW	NORM	7	0.95	0.16
C4	**AP**	RAW	NORM	6	**0.94**	**0.20**
			ESP5			
C5	**GMM**	RAW	UNORM	5	0.86	0.04
C6	**HMM**	PRO	NORM	8	0.98	0.11
C7	**KM**	PRO	UNORM	7	**0.98**	**0.21**
C8	**AP**	PRO	UNORM	8	0.98	0.12

Table 1: Performance of the best purity models for each method (i.e., GMM, HMM, KM, AP) on upstream/downstream navigation (UDN).

As a case study we analyze model C7 which is the best clustering for activity UDN. It was generated by k-means using processed

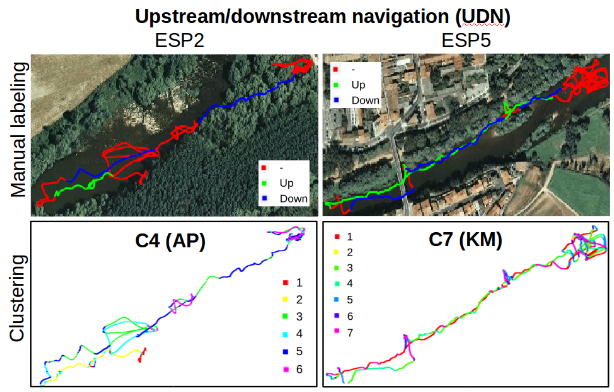

Figure 2: Best clusterings for experiments ESP2 and ESP5. In manual labeling, label '-' (red) means "no label available".

(PRO) and un-normalized (UNORM) data, and it has 7 clusters. The cluster which best matches downstream navigation is cluster 3 (with 58% of coverage of the path manually labeled as downstream navigation) while the cluster which best matches upstream navigation is cluster 1 (with 67% of coverage of the path manually labeled as upstream navigation).

We compare these two clusters to identify the properties that characterize upstream and downstream navigation in experiment ESP5. The variables that show different means between cluster 1 and 3, according to Student's t-test (p-value <0.05) are: mean heading $\bar{h}$, mean values of commands to propellers $\bar{m}_1$ and $\bar{m}_0$, standard deviation of battery voltage $\hat{v}$, mean electrical conductivity $\bar{ec}$, standard deviation of dissolved oxygen $\hat{do}$ and standard deviation of heading $\hat{h}$.

Differences between mean headings of cluster 1 and 3 have an intuitive interpretation since upward and downward navigation have opposite directions. Different mean values of commands to propellers point out that full power was provided during upstream navigation to contrast the water flow (cluster 1), while low power was provided during downstream navigation, when the boat was propelled also by the water flow. Standard deviation of battery voltage was higher in upstream navigation (cluster 1) than in downstream navigation (cluster 3) because battery voltage decreases more sharply when the boat moves upstream than when it moves downstream. The mean electrical conductivity had lower values in upstream navigation than in downstream navigation, probably because the relative movement between the boat and the water influences electrical conductivity sensor reading. A similar behavior was observed for the standard deviation of dissolved oxygen which is higher in upstream navigation than in downstream navigation, probably because of increased turbulences produced by the boat during upstream navigation.

4 CONCLUSIONS AND FUTURE WORK

This work can be extended in several directions. As for the selection of the most significant clusters we aim at sorting all clusters (generated by different methods and parameter settings) according to their silhouette and merge them in a hierarchical structure. The capability of the activity recognition system to interact with humans should be also enhanced in two ways: *i)* by enabling the system to suggest new activities and acquire from humans new knowledge about them (*human-in-the-loop*), *ii)* by enabling humans to understand the knowledge contained in the system (*eXplainable AI*).

ACKNOWLEDGMENTS

This work is partially funded by the European Union's 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

[1] R. Barták and M. Vomlelová. 2017. Using Machine Learning to Identify Activities of a Flying Drone from Sensor Readings. In *Proceedings of FLAIRS 2017*. AAAI Press, 436–441. https://aaai.org/ocs/index.php/FLAIRS/FLAIRS17/paper/view/15488

[2] C. M. Bishop. 2006. *Pattern Recognition and Machine Learning*. Springer-Verlag New York.

[3] L. Bottarelli, M. Bicego, J. Blum, and A. Farinelli. 2016. Skeleton-Based Orienteering for Level Set Estimation. In *Proceedings of ECAI 2016*. 1256–1264.

[4] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu. 2012. Sensor-Based Activity Recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C* 42, 6 (2012), 790–808. https://doi.org/10.1109/TSMCC.2012.2198883

[5] M. R. Endsley. 2000. Theoretical underpinnings of situation awareness: a critical review. In *Situation Awareness Analysis and Measurement*. Lawrence Erlbaum Associates.

[6] B. J. Frey and D. Dueck. 2007. Clustering by Passing Messages Between Data Points. *Science* 315, 5814 (2007), 972–976. https://doi.org/10.1126/science.1136800

[7] R. P. Higgins. 2005. Automatic event recognition for enhanced situational awareness in UAV video. In *IEEE Military Communications Conference (MILCOM 2005)*. 1706–1711 Vol. 3. https://doi.org/10.1109/MILCOM.2005.1605920

[8] R. Laxhammar. 2007. *Artificial Intelligence for Situation Assessment*. Master's thesis. Royal Institute of Technology, KTH CSC, Sweden.

[9] L. R. Rabiner. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* 77, 2 (1989), 257–286. https://doi.org/10.1109/5.18626