Autonomous swarm of heterogeneous robots for surveillance operations

Georgios Orfanidis[1], Savvas Apostolidis[1,2], Athanasios Kapoutsis[1], Konstantinos Ioannidis[1], Elias Kosmatopoulos[1,2], Stefanos Vrochidis[1], and Ioannis Kompatsiaris[1]

[1] Centre for Research and Technology Hellas (CERTH)-Information Technologies Institute (ITI) {g.orfanidis, kioannid}@iti.gr
[2] Democritus University of Thrace, Greece

Abstract. The introduction of Unmanned vehicles (UxVs) in the recent years has created a new security field that can use them as both a potential threat as well as new technological weapons against those threats. Dealing with these issues from the counter-threat perspective, the proposed architecture project focuses on designing and developing a complete system which utilizes the capabilities of multiple UxVs for surveillance objectives in different operational environments. Utilizing a combination of diverse UxVs equipped with various sensors, the developed architecture involves the detection and the characterization of threats based on both visual and thermal data. The identification of objects is enriched with additional information extracted from other sensors such as radars and RF sensors to secure the efficiency of the overall system. The current prototype displays diverse interoperability concerning the multiple visual sources that feed the system with the required optical data. Novel detection models identify the necessary threats while this information is enriched with higher-level semantic representations. Finally, the operator is informed properly according to the visual identification modules and the outcomes of the UxVs operations. The system can provide optimal surveillance capacities to the relevant authorities towards an increased situational awareness.

Keywords: Unmanned Vehicles (UxVs) · Visual-based Operations · Interoperable Architecture · Surveillance Objectives

1 Introduction

The introduction of low-cost, yet, advanced technological tools in surveillance applications have assisted significantly the operational effectiveness of the relevant Law enforcement agencies (LEA) and other similar authorities. The surveillance of critical infrastructures comprises an even more complex objective due to the diverse operational scenarios and environments that the personnel has to monitor. Currently, most approaches involve a time and resource consuming combination of patrols with vehicles, guards and marine vessels which in most cases is proven to be insufficient. Considering also the diversity of the under surveillance territories and infrastructures that might be included, the problem becomes even more complex and challenging. Thus, the relevant personnel requires to be equipped with technologies that will be able to be adapted according to different operational and environmental needs, inter-operate with existing infrastructure, utilize multimodal sensing data and operate autonomously. Finally, due to the distinctiveness of the problem, it is essential to support effectively the personnel via decision support tools over the total situational awareness.

Towards providing a complete surveillance solution, proposed architecture's main objective is to develop a fully-functional autonomous surveillance system with various unmanned mobile robots [1] including Aerial (UAV), water Surface (USV), Underwater (UUV) and Ground (UGV) Vehicles, sufficiently adaptive to operate as a single instance or in a swarm and for either disperse or restricted areas depending on the monitored infrastructure. In particular, the developed framework relies on identifying objects and events based mostly on optical data acquired from multiple sources while the final outcomes are enriched with additional information from other sources. In order to enhance the operational capabilities, the framework includes multimodal processing of the visual data so that monitoring the required infrastructures could be accomplished accurately.

2 Detection services and User-Interface services of the proposed architecture

In order for a system to provide a complete overview of the monitoring area and thus, to support the personnel with valuable decisions, it is vital to analyze initially the tracked scene and extract information of higher-level for the detected objects. The identification of objects of interest within the operational environment is critical for the analysis of the situation, for the evaluation of the threats and for the estimation of future and imminent dangers. Furthermore, visual understanding modules in surveillance applications should combine data from different modalities, such as sensors operating in visual, infra-red and in thermal spectrum in order to augment the detection precision. Video sequences can further boost performance by combining images in a sequence acquired from multiple sources. Towards providing such capabilities in a surveillance system, framework of the proposed architecture involves the following detection modules based on visual data mostly.

2.1 Detection services

Detection techniques for pollution incidents. The framework integrates a dedicated module to provide the capacity of detecting oil spills over sea surface near relevant critical infrastructure like harbors and sea oil refineries. Pollution incidents can be identified by the UAVs of the system which are equipped with Synthetic Aperture Radar (SAR) sensor and after transforming the acquired data from remote sensing data to visual representations. SAR sensors were integrated and utilized due to their efficiency and operations in adverse weather conditions [2]. The entire process is optimized by deploying the swarm of the UAVs in order to increase the detection accuracy and ensure a fast response from the relevant authority. The prediction model that is integrated relies on the DeepLabv2.0 architecture ([3], [4]) and provides a semantic representation of the surveyed area based on which the criticality of the event can be evaluated properly. A simple visual result of the referred module for oil spill detection using visual semantic segmentation is provided in Fig.1.

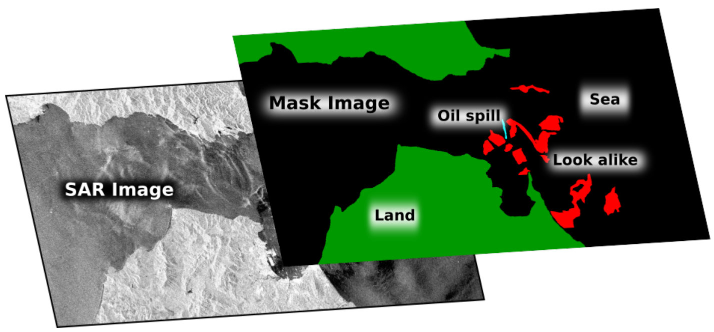

Fig. 1. SAR Image and the corresponding mask for semantic representation of the scenery

Identification and tracking of suspicious activities. The framework aims at integrating the utilization of a wide range of sensors, which typically involve visual, infrared and thermal sensors in order to identify accurately the required suspicious movements in the monitored area ([5], [6]). Due to the nature of the illegal activities and the diverse conditions that are involved in critical infrastructures, operational conditions play a significant role in the effectiveness of the detected targets and thus, it is important to alleviate such issues. Combined approaches of different sensors aids at the reduction of false positives in the detection system which can lead to false alarms. Figure 2 presents two examples of the object detection module on thermal and visual images, respectively. Initial results of the detector were evaluated using the Pascal Voc object detection metric [7] and are presented in Table 1. The detector that was integrated relies on the Faster R-CNN model [5] due to its high performance and robustness in identifying objects of smaller size. The identification of abnormal activities can enhance further the higher-level feedback that the system can provide the operator. A major advantage of the developed detection modules, based upon the [5] architecture, is the online adjustments of the UxV swarm routes towards increasing the detection accuracy and optimizing the surveillance objectives.

Low-level fusion to increase the recognition capabilities Combining multiple information from different sources comprises a significant advantage for

Object detection results - Average Precision						
Person	Car	Bus	Truck	Boat	Ship	Helicopter
0.82152	0.75726	0.57315	0.53351	0.70251	0.83586	0.71638

Table 1. Object detection results using Pascal Voc precision metric

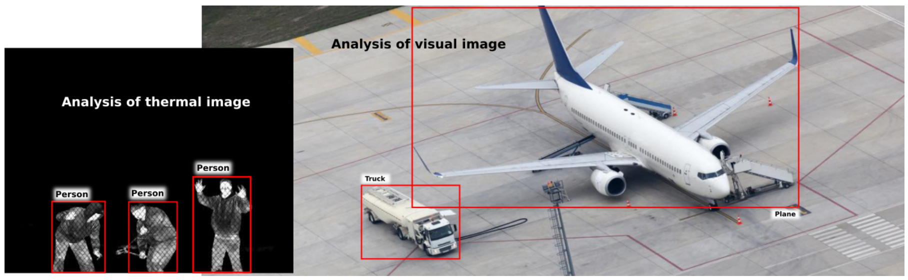

Fig. 2. Object detection applied to a thermal and a visual scene

increasing the accuracy of the recognition modules. Apart from the late data fusion modules, the framework integrates a low-level fusion approach of visual and thermal data based on [8] in order to provide alternative advantages in comparison with the latter fusion. Typically, the fused information will derive from visual and thermal cameras with the disadvantages of each camera being complemented by the merits of the other, as depicted in Figure 3. Visual cameras provide a clearer and easier to use representation in day-time while thermal cameras can capture scenes of the operational area in light restricted conditions (night-time operations and/or foggy environment).

Fig. 3. Image fusion of visual and thermal images

Additional detection module for cyber-physical attacks. The corresponding module includes a framework to increase the situational awareness of agents against cyber-physical attacks [9] and secure significantly the deployed equipment. Emphasis is given on confidentiality breaches and jamming complementing the visual identification especially, in cases where they lead to incorrect actuation or unreliable sensing due to high-jacking. In many cases, the latter can lead to system failures and inability to complete the mission successfully. Thus, the additional processing capabilities of each asset may secure both their functionalities as well as their integrity.

2.2 Navigational services

3D virtual and augmented reality interface. The system is enhanced with additional capacities in order to improve the overall experience of the operator. Towards this objective, a virtual and augmented reality [11] module was integrated and involves the use of an interaction table that allows the operators to interact with the deployed UxVs in a natural and effortless way while simultaneously, it will provide a better situation representation compared to plain view approaches.

DSL-based mission specification. To effectively command the swarm of the UxVs, a dedicated command language is utilized for the description of the missions in order to command and navigate each asset. The language can cover all the aspects of a mission and includes metadata, operation commands, event oriented commands and post-operation commands (i.e. processing collected data). Based on the hierarchical architecture, the language as well as the higher-level commands are adaptable so that, the navigation of the assets could be modified according to the detection outcomes.

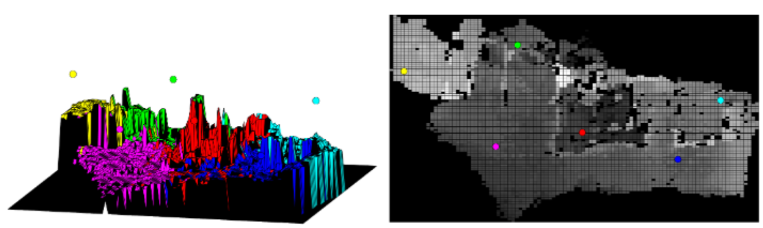

Fig. 4. Real-time surveillance of unknown/dynamically-changing environments

Resource Controller. The integrated module provides an easy-to-operate, remote-control platform, to remotely control the swarm of autonomous, heterogeneous agents (ground, underwater or aerial robots). The platform receives high level objectives and requirements from the human operator, and translates them to real-time, remote-action commands, solely provided to each member of the swarm in order to achieve the mission's objectives. For example, Fig. 4 presents a simulation instance where six UxVs were deployed having a twofold objective (which forms a trade-off). On the one hand, the part of the 3D terrain that is monitored (i.e., visible) by the robots has to be maximized and, on the other hand, for each visible point of the terrain, the closest robot has to be as close as possible to that point. Furthermore, Resource Controller supports plug-n-play additions and subtractions of members to the swarm, always with regards to the mission's requirements [12]. Another crucial factor is that the controller is highly adjustable and capable of handling changes in the missions on the fly. For the implementation of the resource controller, state-of-the-art algorithms on the field of multi-robot path planning [13], have been used and adapted to the specific framework.

CISE-compliant common representation model and risk models. A CISE-compliant common representation framework [14] was developed in order to cover the streaming data from different sensors and other external resources and supports the representation and mapping of data on semantic knowledge structures, valuable in enriching the extracted information. At the same time, intelligent knowledge-driven reasoning schemes are being used for the contextualized aggregation and interpretation of information, derived from different modalities, with final purpose to enrich the low-level fusion capabilities with implicit knowledge derived from hidden relations. The final purpose is the development of a "high level fusion" methodology which centrally combines data from the different geographically dispersed and heterogeneous sensors coupled with contextual, geospatial and temporal information, so as to provide a complete information mapping. In addition, the system integrates a software tool that assists in the short-term prediction of the spatial evolution of hazardous oil pollution incidents and a number of specific illegal activities. This module allows for early identification of the ongoing incidents, as for example is the case of the trajectory of an unauthorized trespasser. An adaptation of [15] strategy is selected in order to process data from diverse sensors and constantly provide better estimations to the decision support module as an extension of the visual components.

Visual analytics and decision support. A decision support tool is included and takes into account data coming from higher level aggregation of sensor data and risk models in order to facilitate the operator to take the right decisions. In this aspect, a visual analytics [16] module was also deployed to visualize information from various fusion nodes and sensor streams in accordance with the necessary UI (User Interface) and HCI (Human-Computer Interaction) [17] requirements.

3 Architecture and pilot use cases

3.1 System Architecture

The developed system provides a fully functional platform for the autonomous deployment of a single, or a fleet of, unmanned vehicle(s), including UAV, USV, UUV and UGV. The system is designed to be used in search-n-rescue and surveillance operations, and incorporates multimodal sensors as part of an inter-operable network, to detect, assess and respond to hazardous situations in remote-areas and border missions and tasks.

The complete network of sensors includes static networked sensors such as surveillance radars, as well as mobile sensors like cameras mounted on robotic platforms. Interoperability is supported by a common framework for the definition and interaction of agents (UxVs and/or static networked sensors) characteristics and monitored phenomena. In order to improve the detection, tracking and imaging capabilities, photonics-based technologies, as well as a passive radar, adapted for use on-board a UAV and other UxVs are used, extending the capabilities of already existing radars. In addition, a radio-frequency communication signal sensor for UAVs based on SDR (software defined radio) [10] has been customized and mounted on different unmanned mobile platforms, in order to intercept emission sources that are present in the area and enrich the overall situational awareness picture with this information, allowing the identification of unauthorized communications and an estimation of the emitters source location.

In terms of communication, the developed platform adopts a cloudlet-based approach with open interfaces that allows the integration of a multitude of heterogeneous sensors and robots from different providers. According to the specific needs of a particular mission, the platform allows the authorities/operators to rapidly deploy the most suitable sensors and team of robots to remotely operate and collect the needed evidence. High-level, low cognitive-load tele-operation is also supported by a Domain Specific Language (DSL) that allows the operator to specify missions and invoke assets. Last, a virtual, augmented reality based multimodal interaction table allows users to interact with robots in a natural way with touch screen and other hand gestures in 3D, allowing them to understand the situation better than in plain view.

From the user point of view, the main functionalities that the platform supports are:

- Control the fleet of UAVs, UGVs, USVs and UUSVs as a team without the need for the control operators to get involved into tedious tele-operation activities.
- Monitor and control the networked border security radars.
- Monitor and interact with robots through a virtual and/or augmented reality based multimodal interaction table.
- View the results provided by each sensor on the robots positioned on the map in real time.
- View the results from the decision support module in real time for rapid assessment and intervention tailored to the situation.

- Automatic recognition of high-importance, hazardous or suspicious activities, based on multiple sensors.
- Automatic alerts when such abnormal activities are detected.
- Automatic adaptation of robot swarm patrol and radar aim according to the identified incidents.

The developed system has taken into account both maritime and remote-land-area operation scenarios [Fig. 5].

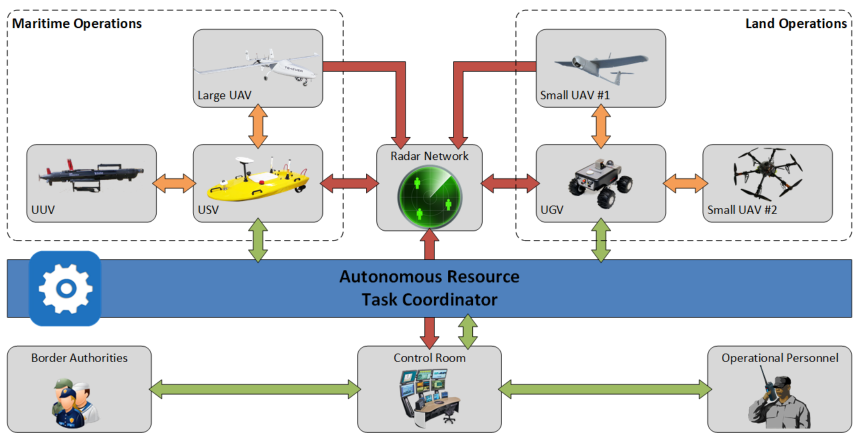

Fig. 5. Higher-level depiction of the system's architecture.

A typical operation of the system in a maritime scenario, usually includes monitoring and detection of aerial, surface or underwater incidents and threats. In such a scenario, the involved assets may be a coastal radar network, large fixed wing UAVs, USVs and UUVs. The communications of the UAV and the USV with the resource controller is direct, through RF links. Since the UUV cannot communicate directly with the resource controller, a USV will coordinate with it and act as communication relayer. Depending on the specific mission and on available assets, employing more than one unit of each type of unmanned vehicle brings proportional increase in detection capability, especially for large surveillance areas.

On the other hand, a typical operation of the system in a remote-land-area scenario, includes monitoring and detection of aerial or surface incidents and threats. Land operations may involve assets like a fixed radar network, large fixed wing UAVs, a small fixed-wing UAV, a small multi-copter UAV, UGVs and a small tethered multi-copter UAV. Depending on the use case at hand and on available assets, employing more than one unit of each type of unmanned vehicle brings proportional increase in detection capability, especially for large surveillance areas.

3.2 Use Cases

In order to assess the developed architecture, the final system will be evaluated in three real scenarios:

Identifying pollution incidents in harbors and sea oil refineries . The system will demonstrate the capability to track pollutants spilled at sea and to determine key environmental conditions needed for defining the response and for forecasting the fate of the pollutants within the territory of the maritime critical infrastructure. This capability will be tested using a natural phenomenon as a proxy of pollutant spill: a river plume within a port will be tracked by vehicles and environmental conditions measured.

Unauthorized trespassing in a maritime infrastructure . The use case involves the monitoring of large sea territories under the responsibility of the relevant authorities. The role of the data mule is assigned to heterogeneous autonomous vehicles equipped with a plethora of sensors like optical and thermal cameras. The mobile devices interact with static infrastructure enabling the commander to determine whether an alarming situation is developing.

Unauthorized trespassing in land territories . The autonomous system will allow to patrol hardly accessible territories within the area of a critical infrastructure leading to an optimized surveillance and control situation system with a maximum coverage. The exploited surveillance units will be the source for directing the patrols and tracking the illegal activities in order to mitigate personal risks and increase monitoring capabilities.

4 Conclusion

The platform presented in this paper provides the relevant monitoring authorities an efficient tool to perform surveillance operations in both remote land and maritime scenarios that includes critical infrastructures. The proposed hierarchical architecture can integrate capacities that focus on the exploitation of multi-modal cameras and identify potential threats within the operational scope. State of the art technologies have been utilized, in order to exploit legacy systems and combine them with innovative solutions, providing a user friendly operational and decision support tool, that will maximize the operational efficiency and minimize the required time to take action in critical incidents.

Acknowledgement

This work was supported by ROBORDER project funded by the European Commission under grant agreement No 740593. The authors would like to thank the ROBORDER consortium for their valuable overall contribution.

References

1. Haddal, C. C., Gertler, J. :"Homeland security: Unmanned aerial vehicles and border surveillance". In: Library of Congress Washington DC Congressional Research Service 2010.
2. Fingas, M., Brown C.: "Review of oil spill remote sensing." In: Spill Science & Technology Bulletin 4.4 (1997), pp. 199–208.
3. Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: "Encoder-decoder with atrous separable convolution for semantic image segmentation".In: In Proceedings of the European Conference on Computer Vision (ECCV) 2018, pp. 801–818.
4. Badrinarayanan, V., Kendall, A., Cipolla, R. :"Segnet: A deep convolutional encoder-decoder architecture for image segmentation".In: IEEE transactions on pattern analysis and machine intelligence 2017, vol. 39(12), pp. 2481–2495.
5. Ren, S., He, K., Girshick, R., Sun, J. .: "Faster r-cnn: Towards real-time object detection with region proposal networks".In: In Advances in neural information processing systems 2015, pp. 91–99.
6. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., Berg, A. C. : "Ssd: Single shot multibox detector".In: In European conference on computer vision 2016, pp. 21–37.
7. Everingham, M., Van Gool, L., Williams, C. K., Winn, J., Zisserman, A.: "The pascal visual object classes (voc) challenge".In: International journal of computer vision 2010, 88(2), pp. 303–338.
8. Bhuiyan, S. M., Adhami, R. R., Khan, J. F. : "Fast and adaptive bidimensional empirical mode decomposition using order-statistics filter based envelope estimation". In: EURASIP Journal on Advances in Signal Processing, 2008, vol. 1, 728356.
9. Pasqualetti, F., Drfler, F., Bullo, F. : "Attack detection and identification in cyber-physical systems". In: IEEE transactions on automatic control 2013, vol. 58(11), pp. 2715–2729.
10. Tuttlebee, W. H. : Software defined radio: enabling technologies. John Wiley & Sons, 2003.
11. Billinghurst, M., Clark, A., Lee, G. :" A survey of augmented reality." In: Foundations and Trends in HumanComputer Interaction, vol. 8(2-3), pp. 73–272.
12. A. Ch. Kapoutsis, S. A. Chatzichristofis, E. B. Kosmatopoulos, "A distributed, plug-n-play algorithm for multi-robot applications with a priori non-computable objective functions", In: The International Journal of Robotics Research, SAGE, Volume 38, Issue 7, June 2019, pp 813-832.
13. A. Ch. Kapoutsis, S. A. Chatzichristofis, E. B. Kosmatopoulos, DARP: Divide Areas Algorithm for Optimal Multi-Robot Coverage Path Planning, Journal of Intelligent & Robotic Systems, Springer, Volume 86, Issue 3, June 2017, pp 663680.
14. Tikanmki, I., Ruoslahti, H. : "Increasing Cooperation between the European Maritime Domain Authorities", 2017.
15. Darema, F. : Dynamic data driven applications systems: A new paradigm for application simulations and measurements". In: In International Conference on Computational Science 2004, pp. 662–669. Springer, Berlin, Heidelberg.
16. Keim, D., Andrienko, G., Fekete, J. D., Grg, C., Kohlhammer, J., Melanon, G. : "Visual analytics: Definition, process, and challenges". In: In Information visualization, pp. 154–175. Springer, Berlin, Heidelberg 2008.
17. Card, S. K. :"The psychology of human-computer interaction". In: CRC Press, 2018.