# 1

# UAVs and UAV swarms for civilian applications: communications and image processing in the SCIADRO project

Manlio Bacco, Stefano Chessa, Marco Di Benedetto, Davide Fabbri, Michele Girolami, Alberto Gotta*, Davide Moroni, Maria Antonietta Pascali, Vincenzo Pellegrini

Manlio Bacco, Stefano Chessa, Marco Di Benedetto, Michele Girolami, Alberto Gotta, Davide Moroni and Maria Antonietta Pascali are with the Institute of Information Science and Technologies (ISTI) of National Research Council (CNR), via G. Moruzzi, 1, Pisa
Stefano Chessa is with University of Pisa, Department of Computer Science, Italy
Vincenzo Pellegrini and Davide Fabbri are with Ingegneria dei Sistemi s.p.a., via E. Calabresi, 24, Pisa

**Summary.** The use of Unmanned Aerial Vehicles (UAVs), or drones, is increasingly common in both research and industrial fields. Nowadays, the use of single UAVs is quite established and several products are already available to consumers, while UAV swarms are still subject of research and development. This position paper describes the objectives of a research project, namely *SCIADRO*[2], which deals with innovative applications and network architectures based on the use of UAVs and UAV swarms in several civilian fields.

## 1.1 Introduction

Nowadays, UAVs are attracting a lot of attention from industrial and research fields. They are suited to a large number of applications, thus making them of interest for commercial and research purposes. The use of single UAVs provides several services to the consumers and to the industry: for instance, from imagery to goods delivery, because of the low-cost and reliability that UAVs can provide. In those applications, UAVs are typically used in a Line of Sight (LoS) fashion, i.e., the pilot can Command and Control (C2) the UAV from the ground without losing sight of it. In fact, national regulamentation are quite severe because of safety reasons, and the use of UAVs for civilian

* corresponding author: alberto.gotta@isti.cnr.it

[2] SCIADRO is the acronym of *SCIAme di DROni*, which translates into *UAV swarm*.

applications must adhere to them, thus making their use more common outside of urban areas. A challenging upcoming scenario for UAVs is the use of *swarms*, or Flying Ad-Hoc Networks (FANETs): several drones, from tens to hundreds, jointly used in order to execute a given task. The joint use of multiple UAVs poses several challenges that must be met, and the SCIADRO research project deals with some specific application scenarios involving the use of swarms. As pointed out in [1, 2, 3], the number of application scenarios for UAVs is rapidly increasing, involving power line inspection, monitoring of cultural heritage sites, environmental monitoring, fire and gas detection, as well as precision agriculture. Several advantages can be brought by the use of multiple UAVs in those scenarios: for instance, it is likely that the overall cost of acquisition and maintenance of several small UAVs is lower than the overall cost of a single large UAV needed for the same task [4]. Furthermore, fault-tolerance is inherently provided by the use of swarms, because a single drone can be removed with a limited impact on the overall formation. Swarms can also provide scalability, i.e., adding or removing drones from a swarm, in order to better adapt to changing conditions or to simply replace one or more UAVs experiencing issues or battery depletion.

The rest of this paper is organized as follows: Section 1.2 provides a brief overview of the state of the art. Then, Section 1.3 describes the main application scenarios that the SCIADRO project is facing, proposing several techniques to deal with the faced issues. Eventually, the Conclusions are in Section 1.4.

## 1.2 Related works

A wide range of applications can benefit from the use of UAVs, as in [1, 2, 3], such as power lines inspection, monitoring of cultural heritage sites, environmental monitoring. In particular, precision agriculture is largely benefiting of the use of UAVs [5], due to low operational costs, high operational flexibility and high spatial resolution of imagery. The use of UAVs in this field is expected to grow faster in the next years and it has proved to be particularly effective in otherwise impervious areas, or each time their use can remove the need for expensive temporary scaffolding, such as in the case of the inspection of historical or cultural areas and buildings. In the latter scenarios, UAVs are typically equipped with the needed sensors in order to facilitate any inspections: for instance, cameras, but also short-range communication radios, in order to collect data from previously installed sensors or to deliver commands, in the case of actuators. A survey of UAV usage for imagery acquisition in the field of disaster research and management in provided in [6]. The authors focus on the inspection of pipelines, in order to quickly provide damage survey thanks to UAVs. In fact, they provide a low-cost solution for imagery collection, and the small size and maneuverability makes them a viable and low-cost option.

The use of UAV swarms, on the other side, is less established than the use of single UAVs. This depends on the issues that the contemporary use of multiple drones poses to C2, for instance, in order to avoid any collisions among the members of the swarm, or the need for intelligent algorithms to exchange data, coordinate the swarm, and process collected/generated data. From the network point of view, an UAV swarm is typically referred to as FANET. A FANET has distinctive features w.r.t. other networks: a high mobility degree, a flight formation, and average and peak movement speeds of single drones and of the whole swarm, that must be carefully taken into account. For instance, aerial mobility experiences less constraints than the terrestrial one (3D movements instead of 2D), with different speeds. Nonetheless, their use is attracting the attention of both the industrial and the scientific world in several fields [1, 6, 7], and this work proposes several civilian applications that can largely benefit of the use of UAV swarms.

## 1.3 The SCIADRO project

Co-funded by the regional government of Tuscany, Italy, the SCIADRO project aims at developing the enabling technologies which are key to accomplishing a rather rich and diverse span of missions through the use of a coordinated drone swarm within a civilian environment. Goals of the foreseen missions include environmental monitoring, first response to natural disasters, monitoring of social events, safety inspection of public utility grids or other critical infrastructures. The project team includes seven partners, being either research institutions or Small and Medium-sized Enterprises (SME) with permanent operations in Tuscany. More in detail, the project aims at: (*i*) developing sensors to monitor the presence of potential pollutants within surveyed environments; (*ii*) achieving computer vision techniques and algorithms which can detect complex objects and extract information on local anomalies which might affect them; (*iii*) developing suitable logics and algorithms which can effectively organize and guide the overall swarm motion and actions during a mission; (*iv*) studying, developing and demonstrating network architectures and protocols which can allow communication among multiple drones within a swarm while also increasing communication reliability towards the ground segment and reducing the *Size, Weight and Power* (SWaP) requirements of in-flight radio-communication equipment.

In the following, we present an overview of the activities on communication and image processing currently carried out within the SCIADRO framework.

### 1.3.1 Inspection of aerial power lines

The inspection of aerial power lines using UAVs is a service offered by several companies, in order to acquire high quality videos both in the visible light spectrum and in the infrared spectrum. The SCIADRO project aims at

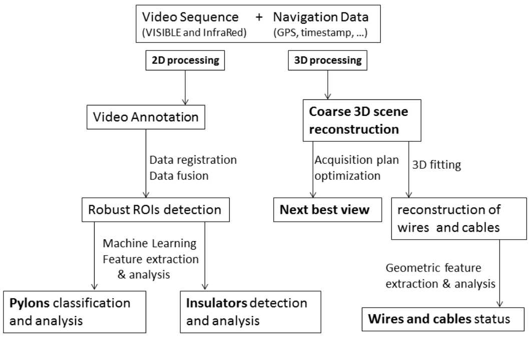


Fig. 1.1: The processing pipeline for 2D and 3D video processing.

automatizing the monitoring process by using a single UAV or a swarm, in a cost-effective manner. The complete automation of such a task requires to empower scene understanding capabilities on board of the UAV, and to enable each drone to share the most relevant information with other drones in order to check, refine and complement the shared piece of information. The acquired sequence of images undergoes both 2D and 3D processing, as summarized in Figure 1.1. Further details on this are in Sections 1.3.2 and 1.3.3.

### 1.3.2 Image understanding and recognition

Many applications of computer vision entered everyone's life, e.g. face detection software embedded in digital cameras, or the Optical character recognition (OCR) software for standard scanners; on the other hand, the recent literature shows that SCIADRO faces a challenge in innovation. The objective of enabling an UAV to perform very complex tasks such as object detection, recognition, and analysis in an unknown environment requires very fast and robust algorithms, and there are still no standard approaches in the literature. Moreover, fast and robust methods of image analysis are intended to be automatic and specifically designed for the collaborative setting of a swarm of UAVs. Even if the collaborative setting poses a number of issues (e.g. regarding the information sharing and processing), it could be seen as a strength, if a data fusion step is properly implemented on the different data flows: this kind of processing may efficiently increase in quantity and quality the information extracted by the diverse sensors hosted by the UAVs.

The inspection of the aerial power lines naturally splits in the following set of actions: (*i*) detection of wires and cables; (*ii*) analysis of the wires and cables; (*iii*) detection and classification of the electric towers; (*iv*) analysis of the tower components (insulators, hanging points). In more details, the detection of wires and cables is inspired by Candamo et al. [8], whose method builds a feature map on the basis of a pixel motion estimation, the morphological properties and the linear patterns, computed via a multi-window Hough space. In our setting, the detection will also benefit of the thermal imaging and the 3D fitting, in order to increase the robustness against scene complexity, occlusion, noise, and environmental clutter. Once wires and cables are correctly detected, the thermal data and the high resolution images are used to assess their health status, i.e., look for defects, loose connections, or excess vegetation. Few related works can be found about detection and classification of electric towers, apart from some preliminary results in [9]. The most recent approaches [10] consider this task as a supervised learning problem. In particular, we will use a multi-layer perceptron neural network both to predict whether the region inside an image is *background* or *tower*, hence giving a single region of interest for the tower detection; and to distinguish the tower type on the basis of a dataset of training. A rich dataset of images to train the neural networks is needed to get good results. The correct classification of the tower would improve the performance of the automatic detection and analysis of the tower sub-components (i.e., insulators and hanging points), which will use a region-based segmentation and template matching, as in [11].

### 1.3.3 Scene Acquisition and 3D Reconstruction

The degradation status of aerial power lines depends on a wide variety of factors, each requiring its own method of measurement and evaluation. For example, as noted in Section 1.3.2, the physical status of the surface of wires and insulators can be evaluated by relying on Computer Vision techniques.

However, there are parameters and peculiar scenarios that require accurate geometric measurements, such as checking for proper hanging of the wire by modeling it as a mathematical function (i.e., the *catenary curve*), and ensuring that the space surrounding the wire is clear of obstructions (e.g., vegetation). SCIADRO will use Computer Graphics techniques, to accurately acquire a representation of the 3D scene, which can be analyzed to detect anomalies and to assess the general health status of aerial power lines. Yet, it can be visualized in real-time, to ease the decision-making process. In fact, a 3D representation of the aerial power lines and of the surrounding environment are generated, and analyzed by *photogrammetry* algorithms, or through active acquisition devices (e.g., UAV-mountable 3D scanners, LiDAR devices). The target area is *sampled* by several UAVs, and each sample produces a 3D georeferenced point, by collectively creating a so called *point cloud*.

By analyzing the point density in the cloud and combining it with the path of UAVs, we can compute the next the next route waypoints they should match, allowing for the reconstruction of so-far unknown areas of the scene. It is clear that the integration of such an algorithm, known as *next best view* selection, will produce, as time passes, a better and better reconstruction of the scene. Once enough data has been collected, a combination of Computer Vision and Computer Graphics techniques will be used to detect the aerial power lines. At first, points are classified based on their confidence in belonging to an elongate structure (i.e, a wire) by analyzing the anisotropy of the spatial distribution of the neighboring points w ith the eigenvalues of the covariance matrix of the 3D coordinates. As a sample implementation, an experimental plug-in was developed in MeshLab [12] to harness the algorithmic correctness (see Figure 1.2). By observing that a catenary seen from above is a straight line, a Hough Transform [13] is applied to the projection of candidate points onto the ground plane; each detected segment is then projected in its containing vertical plane and a RANSAC fitting procedure is applied to compute the parameters of the catenary equation that will be used to test whether the cable is hanging properly. But, even if the wire catenary itself complies to construction constraints, obstructions such as tree branches should trigger a maintenance action. To cope with these scenarios, 3D analysis will exploit technologies such as voxel coloring [14] to check whether power lines have a surrounding adequately clear of obstructions such as vegetation branches. Even if an automatic analysis is possible, it is often very useful to make the

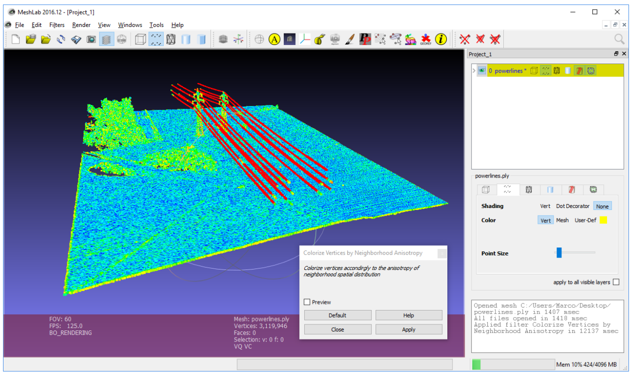

Fig. 1.2: Powerline 3D points classification, MeshLab experimental plug-in.

surveyor or the operator aware of the swarm configuration and the actual condition of the scene of interest. Hence, the need for a visualization tool that is able to cope with multiple video stream and present them to the user in an effective way. The problem of assembling and presenting multiple views of the same scene has been tackled by several important works, both for still images [15] and for video streams [16, 17], with the goal of generate smooth transitions from one point of view to the other. Building on these ideas, in the context of SCIADRO, we will create a visualization tool that helps the user understand the current spatial configuration of the swarm and assist her in inspecting the scene with a constrained navigation metaphor.

### 1.3.4 Sensing the crowd with UAV swarms

Even if the current regulation limits using UAVs in urban and populated areas, we foresee that such limitations will have to be mitigated in the near future, opening to some possible application scenarios, where UAV swarms could be employed as well. According to this view, we investigate in SCIADRO how UAV swarms can be employed to achieve crowdsensing tasks. The term *crowdsensing*, is referred here to the possibility of collecting data produced by people devices in urban or rural areas. In fact, the great majority of today's

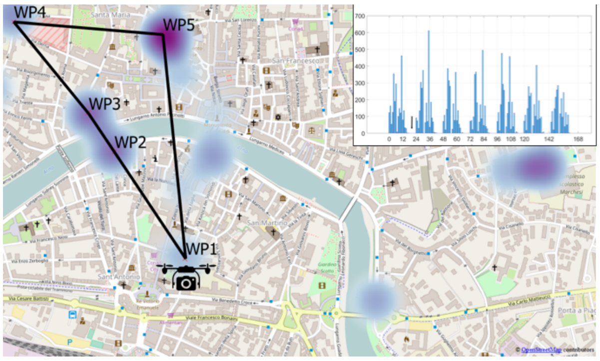


Fig. 1.3: The crowdsensing scenario under consideration.

smart devices provide advanced sensing and computational capabilities that can be exploited to collect data generated directly from the crowd. We cite the ParticipAct living lab [18] among the successful crowdsensing campaigns, which was designed to gather environmental and multimedia data generated by smartphones or smartwatches (i.e., temperature, noise intensity, user surveys, pictures or movie clips). We plan to employ an UAV swarm as mobile agents that collect specific kinds of data from people's devices, by flying over the region of interest. Such devices already have the possibility of directly interacting, giving rise to the so-called Mobile Social Networks (MSN) [19]. UAVs could be analogously seen as peers of a MSN. In fact, they can be equipped with short range communication interfaces in order to join the MSN formed on the ground, similarly to other devices. Under this context, UAVs are mobile collectors of data produced on the ground. Design optimized trajectories is a key element that allows the UAV swarm to efficiently overfly the most crowded locations. Therefore, understanding the way the people move and the way the people interact is crucial for an efficient use of UAVs in crowdsensing scenarios. Figure 1.3 shows the map of Pisa with an overlay heatmap of the most crowded areas. The UAV overflies a number of waypoints (WP1 to WP5), since we learned from history of user mobility that those areas become crowded at specific time slots. In order to measure the effectives of the trajectory proposed, we also assess a methodology to measure how many data an UAV can collect from the ground. To this purpose, we study some metrics useful to determine: ($i$) the quality of the interactions among the users; ($ii$) the fraction of the users *sensed* by the UAV. In particular, we are interested in measuring: ($i$) the number of interactions of the UAV swarm with other devices (also referred to as contact number); ($ii$) the duration of such interactions (also referred to as contact duration). The inset in the right side of Figure 1.3 shows the average number of hourly contacts of the UAV with other ground devices on a hourly basis. We can compare such values to those obtained with UAV trajectories obtained without any notions of the human sociality. We refer to such last ones as *social-oblivious* paths. Preliminary results show a remarkable increase on the number of UAV swarm contacts through social-aware trajectories, by reducing the flight time.

### 1.3.5 MP-RTP-based multimedia data transmission from UAVs to ground

SCIADRO also addresses using multipath techniques to deliver quasi-real-time multimedia data from one or more UAVs toward a terrestrial Ground Control Station (GCS). The data link is assumed separated from the C2 link for safety reasons. We focus on the more challenging case of a live video rather than still frames, which require less bandwidth. An UAV can experience bad channel conditions towards a GCS due to obstacles or channel fading; furthermore some Doppler effect must be taken into account, in case of UAV high speed. Moreover bandwidth requirements must be tackled since multimedia flows are bandwidth-eager, but broadband links can be very expensive. Here, we propose the use of MPRTP (MultiPath Real-time Transport Protocol)[3] to support high-throughput (HT) multimedia applications. Several physical links are aggregated by MPRTP, in order to provide an efficient, reliable and cost-effective HT logical channel. Having more than one channel available at any time, the better performing channel (or channels subset) can be used, excluding (or limiting the use of) bad performing ones. The use of MPRTP provides a way to *schedule* the transmission of a multimedia flow over multiple links, according to the implemented scheduling policy. Figure 1.4 sketches the use of MPRTP between an UAV and its GCS in the proposed scenario.

A preliminary implementation of MPRTP protocol is available online at https://github.com/multipath-rtp/gst-mprtp, and is built upon the open source multimedia framework GStreamer. The scheduler is the core module, highlighted in Figure 1.4 as part of the MPRTP implementation. The scheduler core functionalities can be summarized as: (*sender side*) splitting of the multimedia flow in $n$ sub-flows, and transmission of the $i$-th subflow on the $i$-th physical link; (*receiver side*) aggregation of the $n$ subflows to reconstruct the original multimedia flow, and transmission of feedback data (RTCP). The current scheduler implementation estimates the overall available bandwidth, and splits the multimedia flow, accordingly. The source video bitrate is then set, by providing the maximum achievable video quality, according to the available bandwidth. However, the current implementation does not provide a way

[3] https://tools.ietf.org/html/draft-ietf-avtcore-mprtp-03

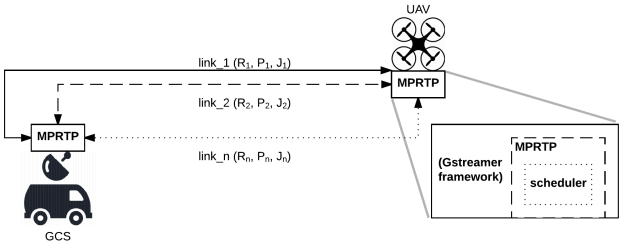


Fig. 1.4: An exemplary scenario involving the use of MPRTP for multimedia transmissions from an UAV to a GCS.

to reduce the load on lossy and/or high-delay sub-links, thus causing corrupted video at the receiver. We are currently implementing a modified version of the scheduler, which reacts in real-time to time-varying channel statistics, and that is able to properly choose subset of the physical links. Figure 1.4 shows how our scheduler reacts to Round-Trip Time (RTT) $R_i$, packet loss rate $P_i$, and jitter $J_i$ variations on the $i$-th link. A preliminary demonstration of what our implementation provides is available at wnlab.isti.cnr.it/ncmprtp.

## 1.4 Conclusions

In this work, we present an overview of the activities currently ongoing in the SCIADRO project. The use of UAVs and UAV swarms is under investigation, and several applications have being taken into account, as exemplary use cases of interest for the market. While the use of single drones is quite established, swarms still require some investigations, and the SCIADRO project will shed some lights on still open issues and possible solutions.

## Acknowledgment

This work has been partially supported by the Tuscany region in the framework of SCIADRO project (FAR-FAS 2014).

## References

1. P. Daponte, L. De Vito, G. Mazzilli, F. Picariello, S. Rapuano, and M. Riccio, "Metrology for drone and drone for metrology: Measurement systems on small civilian drones," in *IEEE Metrology for Aerospace.* IEEE, 2015, pp. 306–311.
2. M. Bacco, L. Caviglione, and A. Gotta, "Satellites, UAVs, Vehicles, and Sensors for an Integrated Delay Tolerant Ad Hoc Network," in *6th International Conference on Personal Satellite Services (PSATS): Genoa, Italy, July 28-29, 2014.* Springer, 2016, pp. 114–122.

3. M. Bacco, E. Ferro, and A. Gotta, “UAVs in WSNs for agricultural applications: An analysis of the two-ray radio propagation model,” in *IEEE SENSORS*. IEEE, 2014, pp. 130–133.
4. I. Bekmezci, O. K. Sahingoz, and Ş. Temel, “Flying ad-hoc networks (FANETs): A survey,” *Ad Hoc Networks*, vol. 11, no. 3, pp. 1254–1270, 2013.
5. A. Matese, P. Toscano, S. F. Di Gennaro, L. Genesio, F. P. Vaccari, J. Primicerio, C. Belli, A. Zaldei, R. Bianconi, and B. Gioli, “Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture,” *Remote Sensing*, vol. 7, no. 3, pp. 2971–2990, 2015.
6. S. M. Adams and C. J. Friedland, *A survey of unmanned aerial vehicle (UAV) usage for imagery collection in disaster research and management*, 2011.
7. H. Menouar, I. Guvenc, K. Akkaya, A. S. Uluagac, A. Kadri, and A. Tuncer, “UAV-Enabled Intelligent Transportation Systems for the Smart City: Applications and Challenges,” *IEEE Communications Magazine*, vol. 55, no. 3, pp. 22–28, 2017.
8. J. Candamo, R. Kasturi, D. Goldgof, and S. Sarkar, “Detection of thin lines using low-quality video from low-altitude aircraft in urban settings,” *IEEE Transactions on aerospace and electronic systems*, vol. 45, no. 3, pp. 937–949, 2009.
9. J. Tilawat, N. Theera-Umpon, and S. Auephanwiriyakul, “Automatic detection of electricity pylons in aerial video sequences,” in *International Conference on Electronics and Information Engineering*, vol. 1, Aug 2010, pp. 342–346.
10. C. Sampedro, C. Martinez, A. Chauhan, and P. Campoy, “A supervised approach to electric tower detection and classification for power line inspection,” in *Int. Joint Conf. on Neural Networks*, July 2014, pp. 1970–1977.
11. A. Wronkowicz, “Vision diagnostics of power transmission lines: Approach to recognition of insulators,” in *9th International Conference on Computer Recognition Systems*. Springer, 2016, pp. 431–440.
12. P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia, “MeshLab: an Open-Source Mesh Processing Tool,” in *Eurographics Italian Chapter Conference*. The Eurographics Association, 2008.
13. R. O. Duda and P. E. Hart, “Use of the hough transformation to detect lines and curves in pictures,” *Commun. ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972.
14. S. M. Seitz and C. R. Dyer, “Photorealistic scene reconstruction by voxel coloring,” *Int. J. Comput. Vision*, vol. 35, no. 2, pp. 151–173, Nov. 1999.
15. P. Brivio, L. Benedetti, M. Tarini, F. Ponchio, P. Cignoni, and R. Scopigno, “Photocloud: interactive remote exploration of large 2d-3d datasets,” pp. 86–96, mar-apr 2013.
16. M. Waschbsch, S. Wrmlin, and M. Gross, “3d video billboard clouds,” *Computer Graphics Forum*, vol. 26, no. 3, pp. 561–569, 2007.
17. C. Lipski, F. Klose, and M. Magnor, “Correspondence and depth-image based rendering: a hybrid approach for free-viewpoint video,” *IEEE Trans. Circuits and Systems for Video Technology*, vol. 24, no. 6, pp. 942–951, Jun 2014.
18. G. Cardone, A. Cirri, A. Corradi, and L. Foschini, “The participact mobile crowd sensing living lab: The testbed for smart cities,” *IEEE Communications Magazine*, vol. 52, no. 10, pp. 78–85, October 2014.
19. M. Girolami, S. Chessa, and A. Caruso, “On service discovery in mobile social networks: Survey and perspectives,” *Computer Networks*, vol. 88, pp. 51 – 71, 2015.