THE FUTURE OF MEDIA PRODUCTION THROUGH MULTI-DRONES' EYES

A. Messina[1], S. Metta[1], M. Montagnuolo[1], F. Negro[1], V. Mygdalis[2], I. Pitas[2], J. Capitán[3], A. Torres[3], S. Boyle[4], D. Bull[4], F. Zhang[4]

[1]RAI Radiotelevisione Italiana, Italy, [2]Aristotle University of Thessaloniki, Greece, [3]University of Seville, Spain and [4]University of Bristol, UK

ABSTRACT

The paper presents the ongoing results within the H2020 European project MULTIDRONE (2017-2019), aimed at developing an innovative and intelligent multi-drone platform for media production. First, the paper highlights editorial requirements for a smart media production making use of multi-drone audio-visual capture systems. Then, it describes the modular end-to-end platform, consisting of a flying multi-drone team and a ground station in which advanced algorithms for automated multi-vehicle path planning, crowd avoidance and computer vision are integrated.

Particular emphasis is given to the Director's dashboard, i.e. the production tool that allows a director and her staff to plan, manage and execute multi-drone shooting missions from the editorial point of view.

Finally, the paper details preliminary findings about novel visual effects and enhanced forms of viewer engagement and interactivity that the concurrent usage of multiple drones can bring to future media production.

INTRODUCTION

Unmanned Aerial Vehicles (UAV), commonly known as drones, have become nowadays useful for a plethora of services, such as scientific data collection, agricultural applications and, quite obviously, for media production. Drones' potential in our domain is getting higher thanks to several winning features from both the motion and the shooting perspective. On the one hand, they guarantee high flexibility, time savings and lower costs, making them ideal substitutes of dollies, static cranes and, in some cases, even helicopters. On the other hand, drones enable new visual effects overcoming and extending current shooting modalities, literally allowing birds' eye views on shooting targets, thus significantly improving the viewing experience.

However, the usage of single drones does not benefit from the substantive advances that the concurrent usage of multiple coordinated drones could bring, providing customers with even more immersion into the scene.

THE MULTIDRONE PROJECT

The MULTIDRONE project[1] aims at developing an innovative and intelligent multi-drone platform for media production, mainly for outdoor event coverage. The scenario envisaged by the project is the one in which the director interfaces with the system platform communicating general and concise event coverage instructions. As a result of this Human-Computer Interaction (HCI), a mission plan is computed, consisting of feasible flight trajectories that comply with any relevant legal restrictions, and parts of the mission logistics (e.g., number of required drones, batteries, charging stations, etc.) are determined as well. Subsequently, the mission plan will be assigned to the drone team in order to (autonomously) fly and acquire the desired footage, constantly adapting to the ever-changing situations arising within the event area by exploiting their perception abilities, while optimally dealing with energy-autonomy, intra-swarm coordination and flight safety issues. All the above would only require the minimal oversight of a single flight supervisor.

Current emerging research focuses on solving simple sub-problems, e.g., outputting feasible single-UAV trajectories that allow the camera to capture desired footage conforming to basic cinematographic principles (1, 2). However, little effort has been spent towards researching multiple-UAV/swarm shooting (3), despite the obvious advantages. In contrast, the overall MULTIDRONE methodology aims at streamlining the whole drone shooting process, spanning the entire continuum from eliciting editorial requirements, media production software/hardware specifications, algorithm research and development, system design, system implementation/integration, benchmarking/validation and end-user trials. The core research efforts in the project are oriented towards multiple drone planning, control and robust communications, in order to achieve intelligent behaviour and enhanced safety, robustness and autonomy for the multiple drone team, thus, allowing the production crew to focus on the creative part of their work. In addition, it includes researching the necessary multiple drone active perception functionalities, as well as multi-view audio/video (AV) capture intelligence, targeting novel, robust techniques, capable of operating in real-world conditions. Research towards AV perception is to provide, both before and during production, a semantic world model for drone team navigation, AV production planning and execution, to track in real-time and in real-world conditions, for production purposes, individual shooting targets or target groups. Out of this work, it is expected to create a wealth of novel intelligent shooting/cinematography techniques enabling artistically meaningful multi-view capture of the event, with minimal human intervention, while following the initial instructions of the director, to analyse the multi-view data towards extracting semantic information related to humans (individuals), crowds, events, points of interest and to establish tools for the creation of novel visual effects and (immersive) media experiences. The overall objective is to deliver a completely integrated MULTIDRONE solution, validated in real conditions.

---

[1] *MULTIple DRONE platform for media production* - https://multidrone.eu/ (Last accessed May 2018)

REQUIREMENTS

To realize the objectives of MULTIDRONE the starting point is the collection of a comprehensive set of requirements. In fact, building something new in a domain in which procedures and practices have been established over a long time can be hard if you do not start from deeply understanding what the real needs of the professional users are. Another crucial aspect is considering what you will use the platform for, i.e. which media production scenarios you are targeting.

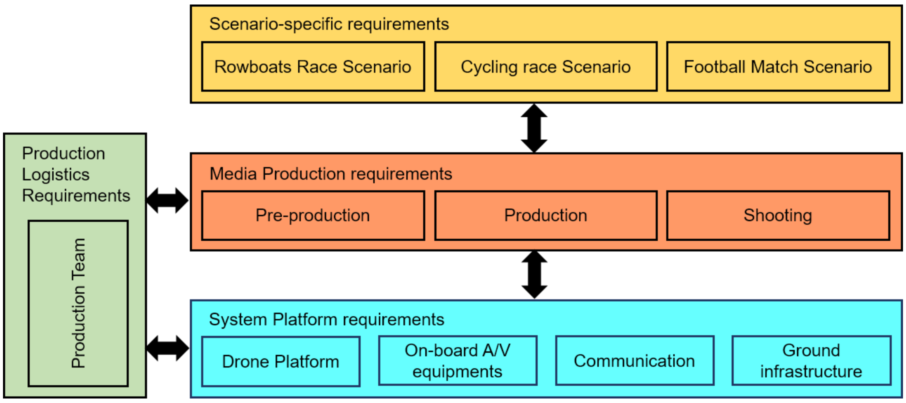

Figure 1 - Requirements analysis approach

Figure 1 shows the approach followed to define different sets of requirements, according to each specific sub-domain of a media production environment.

The first subset of requirements is about the shooting scenarios. In order to cover a relevant set of possible requirements, the project selected boat races, cycling races and football matches (4). These scenarios raise requirements that may differ from one use case to the other such as the speed of the targets (e.g., rowing boats, cyclists) and the versatility of the movement. The terrain and the range to be covered differ as well. The safety issues also vary, the most difficult being the football scenario (due to the audience in the stadium). At the same level of importance, there are some drone-oriented features such as the addressable ranges in space and time as well as target-oriented features, e.g., the availability of radio-based systems for the localisation to improve the system's target localisation precision.

The MULTIDRONE platform poses several media production requirements that concern all three experimental scenarios; in this context, drones are video sources that should be compatible with the overall media production workflow, e.g., in terms of video quality, resolution, format and camera control. Furthermore, media-related production requirements are needed both for the pre-production and for the production phase. The study of the pre-production process brought to highlight the importance of requirements mainly related to the shooting mission preparation (e.g., map production, target specification), definition and planning, together with cinematographic rule planning, training data ingestion and system training aspects. Fundamental requirements are also those about the system and flight emergency procedures given a context with many drones flying at the same time in a limited area. As of production's requirements, we identified topics related to multi-drone system control (e.g., manual/automatic control, target following, system re-planning), media event detection and handling and data/resource sharing among the system units, as well as the handling of safety and ethics (e.g., privacy). Additionally, some general cinematography rules, such as avoiding that drones shoot each other or avoiding to provide the same camera perspectives and framing, need to be respected. To this end, the characteristics of various framing types (e.g., long shot, close-ups) and drone/camera motion types have been set as specific requirements.

As for system platform/hardware requirements, these have been split up into four main classes: drones, on board AV equipment, communication and ground infrastructure. Regarding drones, it is fundamental to consider requirements related to physical features (e.g., drone speed, weight, flight time, stability) and perception/autonomy/safety features (e.g., obstacle detection and avoidance, emergency handling, target detection/tracking). Considering on board AV equipment the focus is mainly on needs related to cameras (e.g., video streaming, video quality, camera/gimbal control) and on-drone AV storage (e.g., memory type and size). Communication requirements deals primarily with video transmission (e.g., video coding, bitrate, resolution) and control signal transmission (e.g., control latency). Finally, with regard to the ground infrastructure, the core requirements are about general functionalities (e.g., manageability, movability, no interference to existing systems), interfacing to studio equipment (e.g., on standards, commodity computing hardware, video transportation) and HCI and dashboards (e.g., usability, AV source mixing and delivery, online mission control, management of notifications).

The last subset of requirements deals with the production team structure and logistic. To manage such a complex system, roles and responsibilities have to be very well mapped out, thus the necessity of strictly define professional profiles (e.g., line producer, director, flight supervisor, pilot, cameraman, technician) and tasks assigned to each of them.

THE PLATFORM

Figure 2 illustrates the MULTIDRONE system architecture, highlighting the main components and actors and the communications between them.

The ground station consists of five functional modules: (i) Dashboard; (ii) Supervision Station; (iii) Mission Planning and Execution; (iv) Perception and Mapping; (v) Communication. The *Dashboard* provides an interface between the Editorial Staff (director and other editorial team members) and the system. The *Supervision Station* is in charge of monitoring the system and checking the planned mission for safety and security. It also provides an interface to the Supervisor, a human operator who supervises the system and assures its safety. The *Mission Planning and Execution* module handles the planning and execution of a shooting mission, providing the interface between the ground station and the drones. It receives a shooting mission from the Dashboard in a machine-readable format and translates it to specific actions for each drone. The *Perception and Mapping* element is responsible for several functionalities such as geometric and semantic mapping, 3D target tracking and visual detection of crowds. The *Communication* component is in charge of implementing the different communication links required by the system. For instance, LTE between the

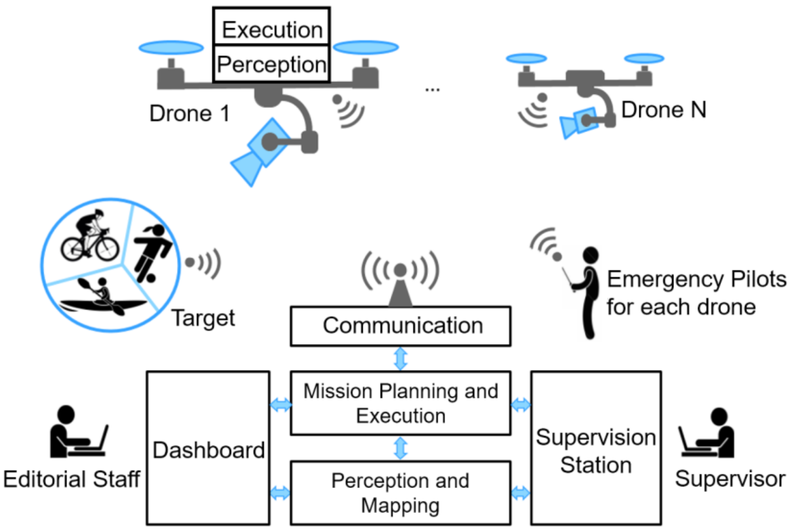

Figure 2 - MULTIDRONE system architecture

ground station and the drones, Wi-Fi for drone-to-drone communications and other radio links for communicating both the targets with the ground station and the emergency pilots with the drones.

The drones are equipped with two main modules apart from communications: (i) Execution and (ii) Perception. The *Execution* module is responsible for receiving, scheduling and finally executing the specific actions required by the ground station. It controls the movement of the drone, the angles of the gimbal and the camera parameters (e.g., focal length). In case of emergency, this module can plan a safe path to an emergency landing zone. The *Perception* module is responsible of analysing filmed frames to perform visual target tracking, keeping the drone localized using the on-board sensors and detecting obstacles along the drone pathway.

Besides the main functional modules, Figure 2 also shows the main actors in the system. The *Editorial Staff* (e.g., the Director, the Cameraman) is the user of the system. The *Supervisor* is responsible for the whole system safety and security by monitoring the mission execution and eventually starting some security protocols. The *Targets* are the sportspeople who are going to be visually tracked by the drones and, in case, by a GPS module to enhance accuracy. Several protocols are in place for autonomous landing in safe zone in case of error or emergency. However, the *Emergency Pilots* are always ready to take manual control of the drones if they consider that there is a safety problem.

THE DIRECTOR'S DASHBOARD

The Director's dashboard (dashboard henceforth) is the set of tools used by the director and her team to govern the system from the editorial point of view both in the pre-production and in the production phases. The guiding principle is the concept of "Event", defined as any real world occurrence that has a (possibly partly determined) spatial and temporal localisation, and a set of actors playing different roles in the action. In MULTIDRONE, the generic Event conceptualisation specialises in the "Event of editorial interest", i.e. any Event that has an editorial relevance in a media production process. An Event of editorial interest, or simply Event of Interest (EOI), can be structured into sub-events inside the dashboard. The key point is that a Shooting Mission is always linked to an EOI and represents the (possibly planned) reaction to its occurrence.

Figure 3 illustrates the use case breakdown of the dashboard whose main aim is that of supporting the director and her team in the production of one or more editorial Shooting Missions (*Mission* in the Figure). The higher-level process upstream of any Mission Production is that of Event Management, i.e. the process by which

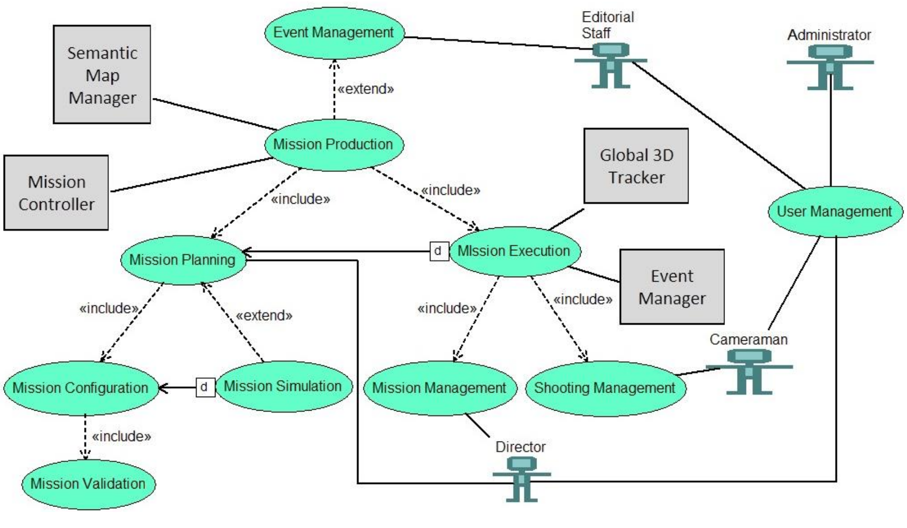

Figure 3 – Use case diagram of the Director's dashboard

EOIs are organised hierarchically and associated to Shooting Missions to be executed in reaction to them.

The dashboard manages information such as geometric relationships between drones (formation) or about how the formation should evolve along time while shooting. The dashboard takes care of the parameters governing the relative motion of the whole formation w.r.t. the reference target as well as the motion of each drone in the formation w.r.t. the formation centre. This high-level description is mapped into one or more executable shooting missions by the ground station software.

The Mission Production process is made of two sub-processes: Mission Planning and Mission Execution. The latter depends on the first one. Logically, this means that when producing any Shooting Mission, there will always be a Mission Planning process followed by a Mission Execution process. Mission Planning is the phase in which all aspects related to a Shooting Mission can be defined through the dashboard's user interface. The Mission Planning process is broken down into a Mission Configuration process, optionally supported by a Mission Simulation process. A fundamental sub-process of Mission Planning is Mission Validation, i.e. the process by which an editorial Shooting Mission is analysed and approved or rejected by the system. Mission Execution is broken down into two distinct sub-processes: Mission Management and Shooting Management. Mission Management is the process by which the Director takes final decisions about which among the several Shooting Missions associated to an EOI and which among the planned Shooting Actions composing the selected Shooting Mission have to be executed. Shooting Management is performed by the cameraman who acts on camera parameters to optimise the quality of each drone's shooting.

Figure 3 also presents the roles involved in the overall process. The Director is taking part in the Mission Planning and the Mission Management processes. The Cameramen is involved in the Shooting Management process. A generic Editorial Staff role is employed in the Event Management and the User Management processes. The Administrator role has the function to manage user roles of the dashboard application, e.g., to include new users and associate those with roles and build editorial teams for missions.

Once the design phase was completed, a wireframe was created to reflect the desired features and workflows and to act as a guide for the actual development of the dashboard. Figure 4 shows some examples of the wireframe's pages.

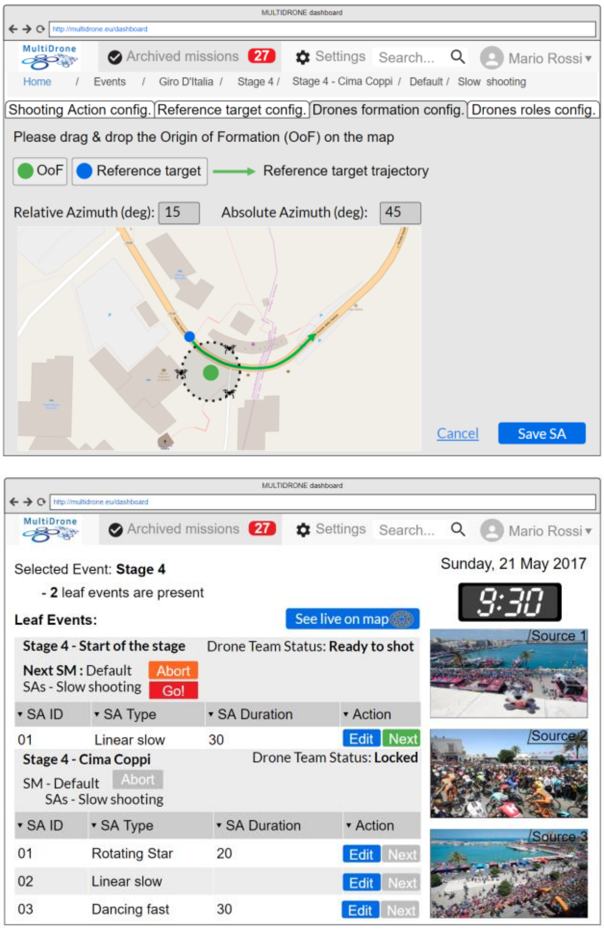

Figure 4 – Examples of wireframe's pages

VIEWING EXPERIENCES WITH DRONE VIDEOS

During an aerial video captured by a drone platform, shooting and flying parameters can be set dynamically. Therefore, the camera parameters and the relative motion between a camera and a specific target can be properly coordinated thus returning to the viewer an innovative and immersive experience. Nevertheless, the relationship between various scenarios, shot types and drone parameters needs to be investigated. It is important to characterise the optimal drone parameters (or their optimal operating envelopes) for specific scenarios and shot types through a subjective evaluation. These results can help to understand the perceptual influence of these parameters in drone cinematography, to constrain flight planning within acceptable limits in the director's Dashboard, to recommend optimum shot parameters and to design innovative shooting techniques and shot transitions.

As an alternative to acquire real footage, simulation engines were used here to generate animated test videos for experimentation. Simulation facilitates the generation of different test scenarios and provides flexibility over the choice of environment, target(s) and actions. Furthermore, camera and drone parameters can be carefully designed and easily changed, providing a much lower cost solution to generate large amounts of data compared to live shooting.

In order to give simulated video good realism, we adopted Unreal Engine 4 (UE4) from Epic Games.[2] This platform is widely used in the development of gaming engines, it is relatively easy to learn and offers extensive community support.

In initial experimentation, six test scenarios were designed using UE4 to simulate outdoor sport events including car and bicycle racing (sample frames of two test scenarios are shown in Figure 5). Their optimal parameters (the height and speed of the drone) were evaluated in a two-stage experiment (height first and then speed) involving forty participants (20 females and 20 males) with an average age of 33. Each trial consisted of the participant viewing a 3 seconds mid-level grey screen, before viewing a randomly chosen sequence. The subjects were then asked to rate their viewing experience from 5 (Excellent) to 1 (Bad), following a single stimulus discrete quality evaluation method. After the whole test session, each participant was informally interviewed about her viewing experience and scoring criteria. After all subjective data has been collected, mean opinion scores were calculated for each test sequence by taking the average opinion score for all the participants, along with the 95% confidence intervals.

Figure 5 – Sample frames for the Cycle-Follow and Car-FlyBy scenarios

[2] https://www.unrealengine.com (Last accessed May 2018)

The results showed that the preference of drone parameters is, to some extent, content dependent and, as might be expected, there is variability between individuals tested. The range of "optimal" drone parameters however does remain fairly consistent across the different test scenarios for the same event, which was observed to be related to the size of target objects, target framing and shot types. During informal interviews with individuals, conducted after each subjective test, we found out that the scoring criteria of the participants varied across ages and genders and also that short video clip durations may create the so-called "Wow! Factor" and a positive experience, whereas longer shot durations of the same type may cause some perceptual discomfort.

From these first experimentations, we got the strong feeling that simulation can be really useful in drone video production for exploring the operating envelopes for different shot types and scenarios. Furthermore, they provide a valuable training environment for directors, pilots and camera operators prior to a live event.

CONCLUSIONS AND FUTURE WORK

This paper presented the ongoing results within the European project MULTIDRONE (2017-2019), whose aim is to develop an innovative and intelligent multi-drone platform for media production. The project started from a comprehensive analysis of requirements which guided the research in fields like multiple drone planning, control and robust communications, intelligent behaviour and enhanced safety, robustness and autonomy. From the point of view of media production, a major result being developed is the director's dashboard, a software console used by the director and her team to govern the system from the editorial point of view both in the pre-production and in the production phases. Naturally, the last word about worthiness of the whole approach stands with customers. Experiments with users showed that although the preference of drone parameters is, to some extent, content dependent, the range of system parameters however does remain fairly consistent across the different test scenarios. The project is now integrating the various software and hardware modules into the first end-to-end prototype, with which to conduct full-fledged experimental productions in 2019.

REFERENCES

1. Coaguila, R., Sukthankar, G. R., Sukthankar, R., 2016. Selecting Vantage Points for an Autonomous Quadcopter Videographer. FLAIRS Conference. pp. 386–391.
2. Joubert, N., Jane, L. E., Goldman, D. B., Berthouzoz, F., Roberts, M., Landay, J. A., Hanrahan, P., 2016. Towards a Drone Cinematographer: Guiding Quadrotor Cameras using Visual Composition Principles. CoRR. abs/1610.01691.
3. Saeed, A., Abdelkader, A., Khan, M., Neishaboori, A., Harras, K. A., Mohamed, A., 2017. On Realistic Target Coverage by Autonomous Drones. CoRR. abs/1702.03456.
4. MULTIDRONE consortium, 2017. Multidrone media production requirements. Deliverable D2.1.

ACKNOWLEDGEMENTS

This work was partly funded by the European Union's Horizon 2020 programme (grant No 731667, MULTIDRONE).