Robotics and Autonomous Systems 80 (2016) 43–54

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Visual feedback with multiple cameras in a UAVs Human–Swarm Interface

C.T. Recchiuto *, A. Sgorbissa, R. Zaccaria

DIBRIS Department, University of Genova, via all'Opera Pia 13, 16145, Italy

HIGHLIGHTS

- The implementation of a novel Human–Swarm interface for UAVs.
- Analysis of the effects of different visual feedbacks on the performances.
- A first-person POV is suitable for some typologies of simple tasks.
- If the task is complex, exocentric cameras increase the user's performance.
- Camera configurations can have an impact on the resulting performance.

ARTICLE INFO

Article history:
Received 15 April 2015
Received in revised form
27 January 2016
Accepted 18 March 2016
Available online 5 April 2016

Keywords:
UAVs
HSI
Robot swarm
Robot formation
Multirotor
Visual feedback

ABSTRACT

In all situations in which a wide area has to be monitored, a practice emerging in recent years consists in using Unmanned Aerial Vehicles (UAVs), and in particular multirotors. Even if many steps forward have been taken toward the fully autonomous control of UAVs, a human pilot is usually in charge of controlling the robots. However, teleoperating UAVs can become a hard task whenever it is necessary to deploy a swarm of robots instead of a single unit, to the end of increasing the area under observation. In this case, the organization of robots in a structured formation may reduce the effort of the operator to control the swarm.

When controlling a team of robots, the typology of visual feedback is crucial. It is known that, while overall awareness and pattern recognition are optimized by exocentric views, i.e., with cameras from above the swarm, the immediate environment is often better viewed egocentrically, i.e., with cameras on board the robots.

In this article we present the implementation of a human–robot interface for the control of a swarm of UAVs, with a focus on the analysis of the effects of different visual feedbacks on the performance of human operators.

1. Introduction

The possibility to deploy aerial robotic systems opens up new scenarios in the management of situations in which the monitoring of an area is crucial. In disaster sites, autonomous Unmanned Aerial Vehicles (UAVs), being used in the initial phase of the intervention, can anticipate the direct action of human operators in order to evaluate if the conditions are sufficiently safe and give a practical support to the rescuers. Sporting events, video surveillance, agriculture, environmental inspection are other contexts that can surely benefit from the utilization of a team of aerial robots. In these situations, in particular when it is necessary to monitor a wide area, the implementation of a high number of robots is a fundamental aspect.

Many researches are currently aiming at ensuring that the monitoring operations can be performed in a fully autonomous matter, but – at the present time – at least one human operator is required to control and coordinate the swarm of UAVs in some tasks, e.g. moving the robots to a specific location or manually monitoring a portion of the whole area.

In these contexts, letting the robots move in a structured formation allows to:

- increase the field of view, by positioning the robots in a way that a greater section of the surrounding environment can be visualized;

* Corresponding author.
E-mail address: carmine.recchiuto@dibris.unige.it (C.T. Recchiuto).

http://dx.doi.org/10.1016/j.robot.2016.03.006
0921-8890/

- speed up the learning process of the human operator, that is able to predict the robot motions in a more straightforward way and, consequently, to move the swarm faster;
- avoid collisions between robots.

On the other side, many issues arise when a human operator should deal with many robots at the same time: it is indeed not easy to find an optimal trade-off between the ease of use, the number of actions that can be implemented using the related Human–Swarm-Interface (HSI) and the resulting *situational awareness*, i.e., the humans ability to perceive the environment, comprehend the situation, project that comprehension into the near future, and determine the best action to execute [1]. In order to find an acceptable compromise between the aspects above, three fundamental issues should be particularly taken into account.

Quality and typology of the visual feedbacks. Since the robots are teleoperated, it is important to define the best way to monitor and analyze the surrounding environment. The positions of the camera, the necessity of implementing an egocentric, an exocentric view, or a hybrid combination of both, and the quality of the images in relation to the available bandwidth should be investigated, also considering the tasks to be accomplished. In particular, it is known that overall awareness and pattern recognition are optimized by exocentric view, whereas the immediate environment is often better viewed egocentrically [2]. However, a systematic analysis of the effects of different points of view and visual feedbacks in the context of a UAV swarm has not been yet carried out in the Literature.

Robots behavior. Given the SOMU (Single Operator Multi UAVs) context, the robots should be endowed with a certain grade of autonomy, in order to decrease the workload of the human operator. Indeed, when the size of the swarm grows, it is practically difficult, if not impossible, to specify, control and monitor the state of all the agents individually and in a centralized way. Therefore, a common way to control a swarm of UAVs is to let them move in a structured formation, that increases the interaction abilities with the environment and simplify the number of commands that the user should give [3]. If the robots are able to autonomously detect obstacles and other robots, the logic underlying obstacle avoidance and inter-robot collision avoidance can also be internally handled.

User commands. In relation to the robot behaviors and to the tasks to be executed, the level of interaction of the user with the swarm should be established. In case of simple tasks and of decentralized robot behavior a minimal level of user interaction can be required (i.e., if the robots are able to keep the position in the formation and autonomously avoid obstacle, the user may only select direction and orientation). On the contrary, in case of a centralized behavior, a highest level of user interaction is required (i.e., switch between robots, select a robot subset, etc.), with a more complex and error-prone strategy.

The presented work mainly focusses on the first of these three issues, with an analysis of the effects of different visual feedbacks on human performance when dealing with a swarm of quadrotors. For the aim of the study, a custom HSI has been designed with a decentralized strategy, where the robot are able to keep their position in formation, avoiding obstacles and inter-robots collision, while following simple user inputs given with a joypad with two analog sticks: direction (forward or backward) and rotation (clockwise and counterclockwise).

The paper is therefore structured as follows: Section 2 surveys previous works related to Human–Swarm Interfaces, with a specific focus on the typology of visual feedbacks, robot behavior and user interface; Section 3 describes the interface used for the experimental phase, with a detailed description of the algorithm underlying the robots behavior, the interaction with the human operator, the implementation of the system in a 3D simulated environment and the protocol of the implemented tests. Finally, Sections 4 and 5 describe the results and discuss them, and Section 6 presents conclusions.

2. Human–Swarm Interfaces for UAVs: related works

The study of Human–Swarm Interfaces (HSI) is a relatively new area of research, aimed at investigating methods and techniques for the interaction between human operators and multi-robot systems. In the following, an analysis of related works dealing with the three fundamental aspects underlined in the Introduction is presented.

2.1. Multi-robot display and visualization

The development of techniques and strategies aimed at tele-operating robots is continuously growing [4,5]. However, exploiting the visual feedback being collected and returned poses many issues, in particular with multiple cameras [6].

Indeed, when dealing with cameras for controlling robots, two main problems arise, both related to the need to merge information originating from multiple viewpoints to provide situational awareness:

- filtering data to prevent operator overload, since displays from multiple robots can easily overwhelm an operator;
- reconciling the different views of a same element of the scene by finding correspondences between different cameras.

These issues have usually been analyzed dealing with ground robots and more recently with UAVs and a variety of display techniques have been investigated for visualizing and controlling multiple tele-operated robots.

In an immersed egocentric view, the human operator views the surrounding environment by means of a camera mounted on top of the robot, while with exocentric views the camera is mounted on an external robot or it is fixed in the environment [7]. In any case, usually the field of view provided by a video feed is much narrower than human vision, causing a perceptual impairment that leaves the operator prone to operation errors. Experimental tests showed that using egocentric cameras may cause failures in recognizing hazards, degradation of situational awareness, disorientation, in a greater extent than using exocentric cameras [8]. A similar conclusion was drawn following a different approach, based on the evaluation of how humans estimate distances in an immersive virtual environment [9]. In particular, it was investigated how users were able to apply error corrective feedback to improve the accuracy of judgements of egocentric and exocentric distances. The results showed how exocentric distances (i.e., those between two objects) are generally better perceived than egocentric distances (those from an observer to an external object), indirectly pointing out that using an exocentric point of view can help during robotic tasks.

Another interesting point is the possibility of independently controlling the camera position and orientation. Hughes and Lewis [10] suggested that providing a camera that is controlled independently from the orientation of the vehicle may yield significant benefits.

The effects of video stream quality have been also subject of research. It has been demonstrated that, even in optimal conditions, spatial comprehension from a video stream is an intensely difficult task requiring complete concentration, conditions that typically do not occur in reality [11].

The possibility to integrate visual feedback with other information is also extremely interesting. Techniques of virtual environment simulation were applied to address some of the challenges of remote manipulation of teleoperated systems in unstructured environments, with a particular interest on remote excavation [7], while Nielsen et al. [12] compared the contribution of video and map information in navigation tasks and investigated how human