Real-time Unmanned Aerial Vehicle 3D Environment Exploration in a Mixed Reality Environment

Zhuming Ai, Mark A. Livingston, and Ira S. Moskowitz

Abstract—This paper presents a novel human robot interaction system that can be used for real-time 3D environment exploration with an unmanned aerial vehicle (UAV). The method creates a mixed reality environment, in which a user can interactively control a UAV and visualize the exploration data in real-time. The method uses a combination of affordable sensors, and transforms the control and viewing space from the UAV to the controller's perspective.

Different hardware and software configurations are studied so that the system can be adjusted to meet different needs and environments. A prototype system is presented and test results are discussed.

I. INTRODUCTION

In many environment mapping applications a slow processing pace is acceptable. However, real-time mapping and visualization are needed in some cases. One such case is the exploration of unknown environments in security and military applications. Such applications may require exploring 3D environments indoor or outdoor.

Unmanned aerial vehicles (UAVs) have been used in exploring unknown areas [1]. They have the advantage of surveying an area quickly. Because of the speed of a UAV and the limitation of battery life to support a UAV in the air, real-time mapping and navigation is needed for such tasks.

While much research has emphasized on developing fully autonomous UAVs [2], human interaction may be needed since it provides an option for correcting autonomous UAV mistakes while the mission is ongoing. Adjustable autonomy is a smart cooperation of human and robot that allows a robot and human to work together as a hybrid team [3]. The autonomous technology eases the connection between humans and robots, instead of replacing humans altogether.

Real-time visualization of the mapping results could provide a feedback that allows the user to make mission adjustment quickly if the UAV misses certain areas of interest or make a mistake in path planning. This may in turn speed up the process and make more efficient use of the UAV’s limited power. Thus real-time interaction and visualization is important for human robot interaction.

In this paper we will discuss the strategy we used in real-time mapping and visualization of unknown areas using a UAV. Our main contribution is to create a mixed reality environment in which humans and UAVs can work together efficiently and effectively. The remainder of this paper is organized as follows. Section II will review the related previous work; Section III will discuss the idea of viewer centered UAV control and visualization in a mixed reality environment. Our method of user centered UAV interactive control will be reviewed in Section IV. A real-time visualization method will be presented in Section V. Then we will discuss the test results and future work.

II. PREVIOUS WORK

UAVs, as a valuable source of data, have been used for mapping unknown environments. Nex et al. has published a review of such applications [4]. It is a particularly effective tool in mapping dangerous environments such as a nuclear radiation field [5].

Various degrees of real-time cooperation among people and robots can be achieved through adjustable autonomy [6]. Even a fully autonomous micro UAV (a UAV that can be carried by a person) sometimes still needs human intervention in cases such as responding to unexpected emergencies.
or real-time mission adjustments. Thus, intuitive remote control of a micro UAV and visualization may also be a necessary part of an autonomous system.

Virtual reality (VR) has been used to create ground control stations to monitor and control semi-autonomous UAVs using data from on-board cameras and geographic information [7][8]. VR has also been used in the research of simultaneous localization and mapping (SLAM) algorithms [9]. Augmented reality (AR) has recently been used to support micro UAV navigation [10]. Mixed reality (MR) is used to improve pilot situational awareness in near earth environments [11][12].

In mixed reality, we will focus on two aspects of the technology: interactive control of the UAV, and real-time information visualization.

Autonomous micro UAVs have been the focus of much research in recent years [13]. When the pose (location and orientation) of a UAV can be determined with on-board sensors, such as Global Positioning System (GPS) receivers, inertial measurement unit (IMU) sensors, and video cameras, the UAV can be controlled through waypoints (reference points in physical space used for purposes of navigation) [14]. However, in cases when pose estimation is not reliable, such as in a GPS-denied environment, direct human interactive control becomes the better option. Interactive control of the UAV from the controller’s perspective is important, or intensive training is needed to map different frameworks.

3D mapping techniques such as lidar and SLAM usually generate massive sets of 3D points in real-time. Real-time visualization of such datasets could be challenging. When the data is too large to fit at one time into a graphics card’s memory, or even the computer’s main memory, out-of-core visualization technology has been used to overcome such limits [15][16][17]. Another approach is to condense the data first, before storing and visualizing it [18]. Stereo point clouds visualization has also been proposed [19].

One of the data structures that is widely used for storing point clouds is an octree [24]. Octomap is a probabilistic 3D mapping framework that is based on octrees [25]. Algorithms have been developed for octree visualization [26][27]. Different volume rendering techniques can be used for this purpose [28]. To overcome the limitation on volume size, cluster based algorithms have been developed, they are particularly useful in medical visualization [29][30].

Much research has been done on real-time UAV navigation and exploration [20]. SLAM algorithms have been studied in many applications [21][22][23]. Although there are still many issues with these methods, the fact that SLAM does not require a GPS signal makes it attractive in GPS-denied environments.

III. MIXED REALITY: HUMAN-CENTERED UAV CONTROL AND VISUALIZATION

We consider two possible scenarios when the methods presented in this paper can be used: the UAV is close enough and within the controller’s line of sight, or the UAV is outside of the line of sight because of distance or obstacles.

In the first scenario, the exploration target area is in a close range within the line of sight. Controlling the UAV from the cockpit’s coordinate framework directly conflicts with the view the controller is seeing in real-time. A novice operator needs extensive training to comfortably control a UAV. This is inefficient and could be dangerous in certain missions. It would be best if the controller can control the UAV from his/her own perspective. Real-time interaction requires viewer centered UAV control. An augmented reality environment can be used in this case. However, a traditional AR may not work directly in this situation. In a traditional AR application, a video or optical see-through display is used and information is superimposed on top of the image of the real world. In the UAV applications, even if the target area is within the line of sight, it may still be too far to use the AR directly. It would be very difficult to see the detail of the environment because of the distance. Thus, we propose an AR display mode that is similar to seeing through a binocular.

The second scenario is when the target area of interest is far away. In this scenario to understand the location and orientation of the UAV under remote control is even more important for the mission. In this case, VR is more suitable for the task.
In both scenarios, viewer centered interactive remote control and real-time visualization are very useful. In a system, VR and AR do not have to be mutually exclusive; our method lies somewhere in between. The user can reconfigure the system to meet their needs.

The traditional method of remote control of a UAV within line of sight is to control its roll, pitch, yaw, and thrust through a remote control device such as a joystick. All the control parameters are defined from the UAV’s point of view. The user (controller) must imagine that he/she is sitting in the UAV’s cockpit and control roll, pitch, and yaw from that perspective. This mental transformation of coordinate frameworks needs a lot of training. When the UAV is far away from the controller and it is difficult for the controller to see the orientation of the vehicle, he/she has to control the vehicle through the video feed from the on-board camera. This makes the transformation through coordinate frameworks even more difficult. Study has been done to determine the influence of the situation awareness when additional frames of reference are involved in controlling UAVs [31]. They found clear evidence of the influence of motion feedback on the situation awareness of the moving UAV operator.

The system we are developing is a mixed reality system that consists of a visualization module, a control module, and a communication module, as shown in Figure 1. The visualization module has a video or optical see-through head-mounted display (HMD) that can be used both as a AR and a VR display when it is in the see-through blocked mode. A joystick or any mobile device that can simulate a joystick is used as the control module to send commands to the UAV. Orientation sensors are mounted on both the HMD and the joystick to track the orientation of both the user's head and hands. To make the system simpler, a single orientation sensor can be used to track user’s head or hand, assuming they are always pointing roughly to the same direction. A laptop is used to communicate with the UAV, the HMD, and the joystick.

At times, it is beneficial to mix these two modes together. For example, while in the binocular AR mode, the user may turn around to examine the environment that has already been mapped behind him/her. This motion will make the system switch to the VR mode.

IV. USER-CENTERED UAV CONTROL

We have developed a method to remotely control a UAV from the controller’s perspective [32]. The method achieves intuitive interactive remote control of a micro UAV within line of sight. The method uses a combination of affordable sensors and transforms the control space to the controllers perspective from the ground.

In a mixed reality environment, the controller’s head position and orientation can be tracked. In a simplified version of the system, we can assume that when the controller is remotely operating a micro UAV, naturally, he/she is facing the micro UAV while manipulating the control device. This can be treated as a orientation sensor that provides relative orientation information between the controller and the micro UAV. Thus, we can assume that the orientation of the device is always following the controller’s orientation.

When a micro UAV is within the controller’s line of sight, the most convenient way of controlling it is to direct the UAV from the controller’s perspective. Flying directions should be in the ground controller’s perspective, which is different from the UAV’s perspective, as shown in Figure 2.

V. VISUALIZATION IN MIXED REALITY

A. Data Representation

The direct results of mapping algorithms are usually point clouds. It is sufficient to store the
point clouds directly if the area is small and the exploration mission is short. However, for large areas and when the mission takes a long time, the size of the point clouds grows very rapidly and eventually it becomes impractical to store them directly.

There are different ways to store the mapping data; octree is one that is widely used since an octree offers control of the level of detail, and this is helpful in applying it to many application domains. It does have its limitations, though. The maximum size is limited and fixed, so it has problems if the area mapped is out of bound and the tree needs to grow. The maximum resolution is also limited and related to the size of the dataset. Nevertheless, this data structure is useful in many applications. We will use it as the base of our data structure, but expand to multiple layers to have octrees inside octree so that it can expand to a much large area.

B. Visualization

A octree can be visualized by drawing each leaf as a cube or a polygon. The drawback is that when the area has a lot of detail, the number of leaves will grow, to a point that the speed is too slow for interactive visualization. We, instead, treat each octree as a volume, and use volume rendering technology to visualize it. Graphics processing unit (GPU) assisted volume rendering is highly effective, and the speed is constant with the increase of the details (number of leaves). A similar approach has been used for rendering large volumetric data by representing texture with octree [33].

Our algorithm has several steps:

1. Each point cloud is inserted into the octree when the data is received from the sensor on-board the UAV. This turns out to be the most time consuming step in the process. To make sure that the volume can be updated in real-time, parallel processing techniques such as multithreading are used. Still, a small percentage of randomly chosen point clouds have to be thrown away.

2. The octree is converted to a format that the volume rendering algorithm can handle. In our case it is a 3D matrix.

3. A volume rendering algorithm using GPU is used to render the volume.

The octree has a preset size and resolution. When the mapped area is beyond the boundary, the scene will be rendered as multiple volumes. Figure 3 shows an indoor scene that is rendered using the software we developed. The data is captured using a depth (RGB-D) camera system and a SLAM algorithm.

Although the occupied areas in the space under exploration may be sparse, free areas actually also contain information, such as the probability of occupancy. One advantage of using volume rendering for environment visualization is that this information in the environment can be rendered without additional cost. Figure 4 is one such example. The intensity and transparency of each voxel are rendered proportional to the occupancy probability of the node. The data comes from LSD-
VI. Tests and Results

A Parrot AR.Drone 2.0 made by Parrot SA is used in this study. This is a quad-copter that can be controlled by a computer through a WiFi connection. A laptop computer running a Linux operating system is used to communicate with the UAV via WiFi. A Logitech joystick (F310 Gamepad) is connected to the laptop. An InterSense InertiaCube2 is connected to the laptop via an RS232 to USB converter; it is attached to the Logitech joystick to track the orientation of the joystick.

The Robot Operating System (ROS) (http://www.ros.org) is used as the development platform. The control program is developed as ROS control nodes.

We tested both the scenarios described before. For UAV pose estimation, the AR.Drone’s onboard camera and orientation sensor are used to determine the UAV’s pose. The SLAM algorithm developed by Engel et al. [34][35] was used in our test to determine the UAV’s position.

When SLAM is not reliable, the InertiaCube2 is used instead. A ROS node is developed to communicate with the InertiaCube2 sensor, and to receive the orientation information. The InertiaCube2 sensor data is retrieved through the USB serial port with the help of the InterSense software development kit. The data is converted to ROS sensor_msgs::Imu format and published to a topic that other ROS nodes can subscribe to. Thus any ROS node, such as the UAV control node, can receive the orientation tracking information from the InertiaCube2 sensor.

The initialization process has the controller point the remote control device at the UAV at the time when it is in takeoff.

Comparing the two scenarios (with or without the micro UAV’s position information), the availability of the position information does make the control more accurate. However, using a remote control with embedded orientation also has its advantages. Such sensors are already available in many remote control devices. When smart phones and tablets are used as the remote control device, the embedded orientation sensor is very convenient for this purpose.

In both scenarios, the controller can easily control the UAV with a simple instruction on how to use the remote control. The test shows that with this remote control, a novice user can control the UAV comfortably with very little training.

To test the mixed reality environment, a Microsoft Kinect is used to simulate on-board sensors. Research has been done to mount these sensors on a UAV [36]. RGBD-SLAM [37] is used to send the point clouds to our software system. The result is shown in Figure 3. The colors come from the RGB camera. There are obvious holes and noise, and these are typical artifacts of SLAM algorithms.

We also tested with a different setting where onboard camera is used (Figure 5). AR.Drone’s video is streamed to LSD-SLAM package [38] to generate point clouds. The point clouds are sent to Octomap package [39] to store the data into an octree. VTK [28] is used to visualize the data. Our algorithm converts the octree to the data structure that VTK can handle. VTK’s smart volume rendering algorithm is used to render the volume data. An InterSense InertiaCube2 is used to track the orientation of the user. MR display combines the view from the user and the point clouds generated from the SLAM algorithm using the video data from the UAV.
VII. DISCUSSION

In this preliminary study, only qualitative tests have been performed. A controlled user study is needed to further quantify the effectiveness of the method, and it will be carried out in the future.

The system described in this paper is still under development. Each module in the system is functioning; however, we still need to put all the modules together to be a functional system.

VIII. CONCLUSIONS

The main contribution of this paper is that it presents a novel human robot interaction system that can be used for real-time 3D environment exploration with an UAV. The method creates a mixed reality environment, in which a user can interactively control a UAV and visualize the exploration data in real-time. This allows the user to quickly make mission adjustments when the UAV misses certain areas of interest or make a mistake in path planning.

The method uses a combination of affordable sensors and transforms the control and viewing space from the UAV to the controller’s perspective. The traditional method of remote control of a micro UAV within line of sight is to control its roll, pitch, yaw, and thrust through a remote control device from the UAV’s point of view. This mental transformation of coordinate frameworks is difficult and needs a lot of training. The intuitive remote control method presented in this paper allows a novice user to control a micro UAV within line of sight without the need for intensive training. The method can be easily ported to a smart phone embedded with an orientation sensor.

IX. ACKNOWLEDGMENT

This project is supported by the U.S. Naval Research Laboratory Base Program “Intelligent Microflyer”.

REFERENCES
