A Stereo Vision-Based Support System for Tele-Operation of Unmanned Vehicle

Ming-Chang Wen[a], Cheng-Hsuan Yang[a], Yie Chen[a], Er-Xuan Sung[a] and Shih-Chung Kang[a]
National Taiwan University, Taiwan

[a]Department of Civil Engineering, National Taiwan University, Taiwan
E-mail: raywen@caece.net, jason610155@caece.net, yie@caece.net, stuart@caece.net, sckang@ntu.edu.tw

ABSTRACT

This research provides a stereoscopic vision system on an unmanned vehicle system (UVS) for disaster inspection. Taiwan is a disaster-prone area where earthquakes, typhoon, floods and debris flow happen frequently. These disasters usually cause terrain transformation and building collapse that raise the risk for human inspection. However, authorities need instant onsite data right after disaster happened to deploy appropriate rescue missions and distribute resources. Instead of sending bunch of sensor into the field with unmanned vehicles and retrieve a huge amount of numeric data for time-consuming post analysis, an instinct and cognitive method for quick understanding and reconstruction of onsite situation is needed. Therefore, we developed a stereoscopic vision system which can be easily integrated with all UVS. In order to enhance the cognition of the operator, we applied two methods: (1) stereoscopic vision for head mount display, (2) head tracking data for gimbal control. We used two camera sources to simulate human vision and optimized the perspective to suit human eyes' field-of-view (FOV). By displaying the optimized image on the head mount display, the operator would have a realistic first-person view of the UVS. We also developed control logic for 3-axis gimbals that can interpret the head tracking data into gimbal motion so operator can move camera's direction by turning heads. The integration of these two methods provides a enhanced visual aid that can be adapted to almost every UVS nowadays. In the future research, we will deploy the proposed system on common UVS to validate the improvement of human cognition.

Keywords –
stereoscopic vision; NUI; UVS; inspection

1 Tele-operation

Tele-operation interfaces for controlling remote unmanned vehicle have remained largely unchanged during the past fifty years [1]. Current methods mainly use visual imagery covering a restricted field-of-view (FOV) for inspection. Sensory cues are significantly lost which include ambient visual information, depth information, kinaesthetic/vestibular input, and sound [2]. Real UAV applications usually take place in environments that are unstructured, uncertain, and not precisely known a priori. Therefore, fully autonomous control of UVS is typically infeasible/impossible in such cases. Imposing human's intelligence on the task to cope with such uncertainty, some tele-operation of their behavior is desired [3]. The state-of art technologies of 3D input/output are the opportunities to push current methods into the next level. In this research, we are focusing on two types of 3D input/output: stereoscopic vision and head tracking.

1.1 Stereoscopic Vision

Stereoscopic vision is the most suitable and feasible solution for simulating human eye vision. The imagery currently used is usually served as a visual aid for supervisory control. Data acquired from camera and other sensors are re-rendered into plots, charts, or tables for showing as much information as possible. Therefore, a problem of supervisory control is that operators usually confronted with overwhelming complexity of unmanned system interfaces. These non-intuitive visual representations can easily lead to human error and failures [4]. We like to take advantage of stereoscopic machine vision and 3D head mount display for duplicating the experience of human eyes. The stereoscopic vision can be realized using a dual camera set that each camera provides imagery of specific angle-of-view (AOV) for each eye. To make the system highly compatible, the dual camera sources are then processed and streamed out in the form of side-by-side (SBS) 3D format in real-time. Most of the current commercial head mount display products have the ability to play SBS 3D format.

1.2 Head Tracking

In order to improve the control method of camera angle-of-view (AOV), we introduced head tracking devices to synchronize head and gimbal movements. Camera gimbal on UVS is mainly controlled using potentiometers such as sliders and switches as command input interface. Thus, for a 3-axis gimbal, it's a 3 degree-of-freedom (DOF) mechanism. Apparently, using any three physical potentiometers is far from instinct. From the aspect of natural user interface (NUI), a more intuitive method is to simulate the human head movement. We put a 3-axis gyro sensor on head mount display on operator to track the head movement. Then, several mathematical algorithms are needed for transforming the head tracking data into understandable numeric data like Euler angles to allow further signal interpretation for controlling gimbal actuators. The most common gimbal control method is using pulse width modulation (PWM). Since, a PWM signal encoder is needed for communicating with gimbal controller.

2 System Design

We designed a subsystem for unmanned vehicles that aims to enhance the visual experience of tele-operation. To achieve the maximum compatibility, we defined our system to be a plugin module for all sorts of unmanned vehicles. It can be used under situations of both standing alone or integrated.

The system has three major parts: (1) an optimized dual camera set for stereo vision, (2) a 3-axis gimbal mechanism for camera angle adjustment, (3) a remote user interface contains a head mount display. Figure 1.

The dual camera set is made to acquire 3D imagery by simulating human eyes. The 3-axis gimbal can control the rotation angle of camera in all 3D direction. A remote user interface has to be capable of rendering realtime 3D imagery output captured from dual camera and tracking user head movements.

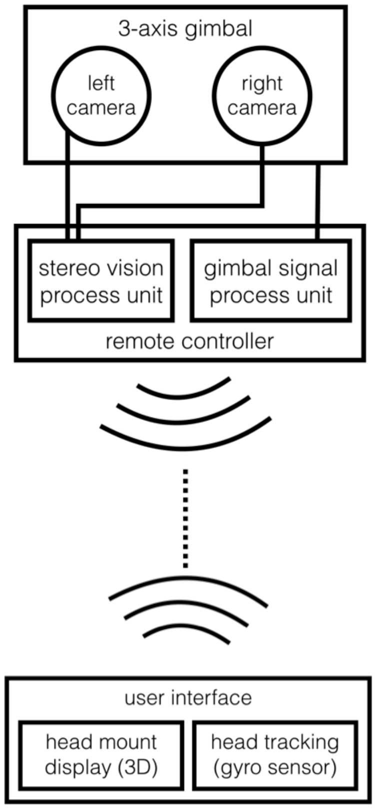

Figure 1. The system structure layout

3 Implementation

For hardware, we integrated dual camera, servo gimbal, and radio receivers on a Arduino based platform for remote robot. Besides, a head mount display and multiple potentiometers is used to prototype the user interface. For software, we developed and combined all function and computing algorithms on Unity 3D which is a platform with high compatibility and extendibility on multiple platforms including mobile device, desktop, and web.

3.1 Hardware

To implement stereoscopic vision, we use two camera modules. Each of them has capability of resolution up to 1920*1080, 30 fps frame rate, 90 degrees FOV, and USB video class (UVC) protocol compliance. Figure 2. We place two cameras with a interpupillary distance (IPD) which is 64mm in between according to the data base of Anthropometric Survey of US Army Personnel [5]. Both cameras are focus to infinity to ensure processed 3D imagery staying in focus. We made a mounting plate for cameras on gimbal that is finely aligned and firmly secured to avoid any possible displacements during the operation of gimbal and UVS. Figure 3.

Figure 2. A camera module used in the system

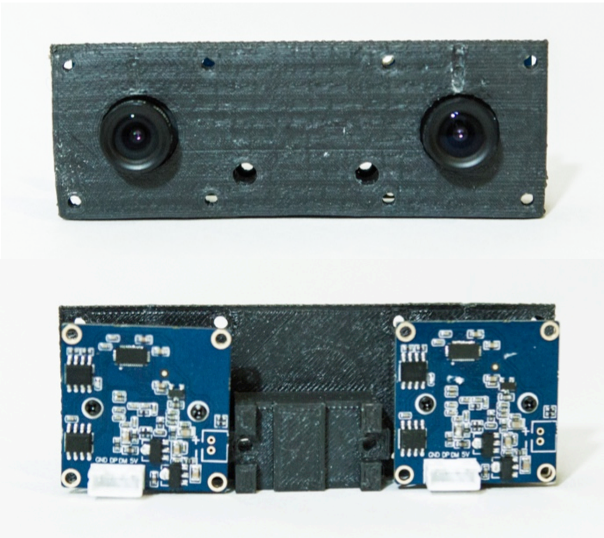

Figure 3. Designed dual camera mount for simulating human eyes

We use three 9g class servos as actuators of a 3-axis gimbal. Figure 4. Each servo controls an axis with 180 degrees range freedom. They are made for simulating the head movement (axis rotation). These servos are driven by PWM signals sending from an Arduino-based controller that is preprogramed for interpreting digital data into PWM signals. We then assemble the dual camera set and gimbal together. Figure 5.

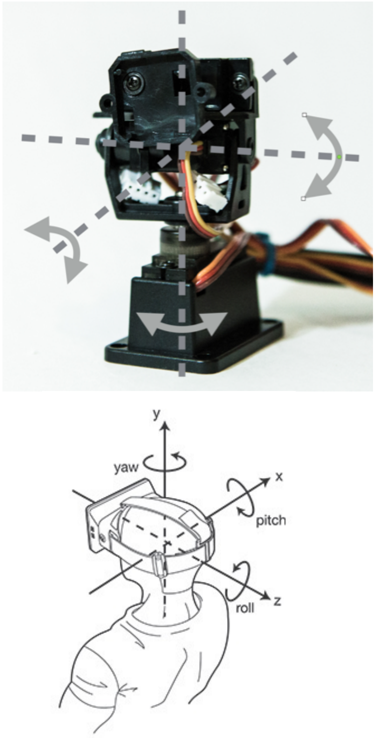

Figure 4. The 3-axis gimbal mechanism and the corresponding head rotation axis

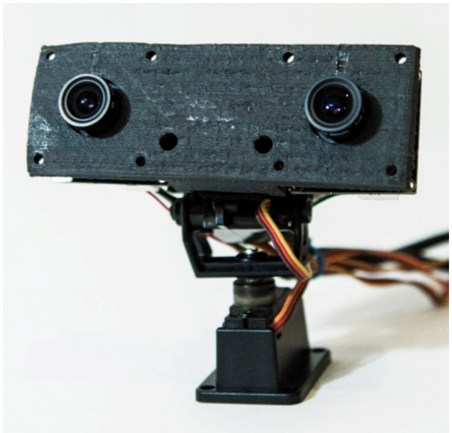

Figure 5. The final appearance of our system

On the user end, we choose Oculus Rift DK2 head mount display as our 3D imagery output device since it already has a built-in gyro sensor and multi platform SDK. Figure 6.

Figure 6. Oculus Rift Development Kit 2 head mount display

3.2 Software

We develop all functions and algorithms on Unity 3D platform. The software accepts separate camera imagery and gyro sensor inputs. We include Oculus SDK libraries and process the dual camera source into cropped SBS 3D format [6]. Figure 7. In order to best represent the human eyes' FOV, we need to determined the distance of imagery plain in 3D environment by converting to the equivalent focal length of human eyes [7]. Figure 8.

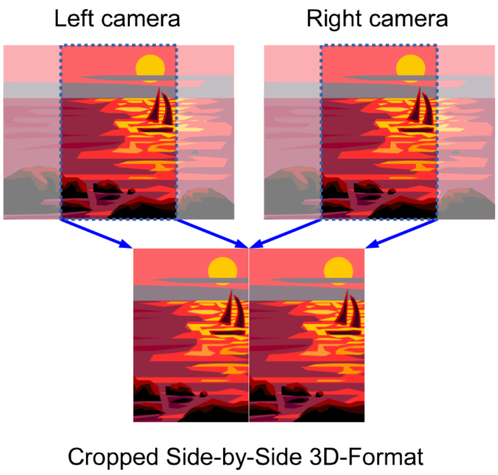

Figure 7. Cropped SBS 3D concept

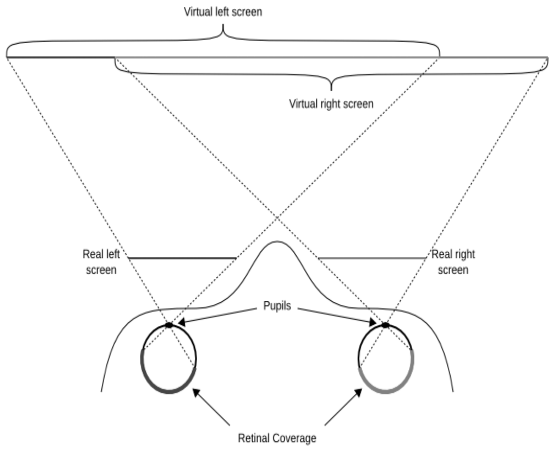

Figure 8. The position setup of screens placed in the 3D environment

An algorithm for converting and mapping gyro sensor data into gimbal servo angle is developed inside the program. Then, the software outputs realtime 3D imagery to head mount display and servo angle to gimbal controller.

4 Field Test

We have integrated our system on an unmanned ground vehicle (UGV) and perform a ground search task. Figure 9. We have tested our system on 5 experienced UGV operators. We have predefined a scenario that operators need to control the UGV to go through a close course that multiple obstacles are randomly placed in the middle of the way. We record their path and used time before and after equipping our system.

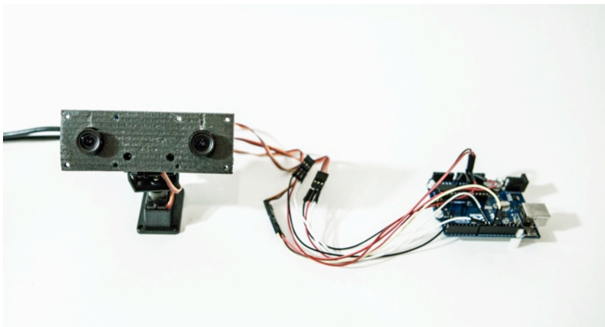

Figure 9. Modularized system that can be integrated to any UVS.

5 Conclusion and Future Work

The proposed system generally raised the performance of operator by providing human-like vision and easing the operation of gimbal. The 3D imagery provides depth information that allows operator to differentiate close objects and having kinesthetic visual feedback just like natural human eyes.

We developed the system aims to act as a plugin module for all types of UVS. So we would like to test how it improves the operation process on unmanned aerial, underwater vehicles. Our future works such as long-range wireless imagery streaming solution and system integration would be the critical issue for further application. Also, we will design different experiment environments and have real UVS operators participated to verify our system by analyzing the behavior change of UVS operators.

References

[1] Fong T., Conti F., Grange S. and Baur C. Novel interfaces for remote driving: gesture, haptic and PDA. *Mobile Robots XV and Telemanipulator and Telepresence Technologies VII, Proceedings of SPIE*, 4195:300-311, 2001.

[2] McCarley J. S. and Wickens C. D. Human Factors Concerns in UAV Flight. Research report associate with Human Factors Research and Engineering Group, Federal Aviation Administration, USA, 2004.

[3] Rauschert A., Meitinger C. and Schulte A. Experimentally discovered operator assistance needs in the guidance of cognitive and cooperative UAVs. In *Proceedings of International Conference on Humans Operating Unmanned systems*, Brest, France, 2008.

[4] Lee D., Franchi A., Son H. I., Ha C., Bu ̈lthoff H. H. and Giordano P. R. (2013). Semiautonomous Haptic Teleoperation Control Architecture of Multiple Unmanned Aerial Vehicles. *IEEE/ASME TRANSACTIONS ON MECHATRONICS*, 18(4), 1334-1345.

[5] Gordon C. C., Bradtmiller B., Clauser C.E., Churchill T., McConville J.T., Tebbetts I., and Walker R.A. *1987–1988 Anthropometric Survey of U.S. Army Personnel: Methods and Summary Statistics.* TR-89-044. Natick MA: U.S. Army Natick Research, Development and Engineering Center, 1989.

[6] The Missing Gear. NerdCam3D Mk1. Online: http://www.themissinggear.eu/app/download/9593010623/NerdCam3D_InstructionManual.pdf?t=1399832672, Accessed: 22/12/2014

[7] Doc-Ok.org. Will the Oculus Rift make you sick? Online: http://doc-ok.org/?p=18, Accessed: 16/03/2014.