Direct View Manipulation for Drone Photography

Yi-Ling Chen* Wei-Tse Lee* Liwei Chan† Rong-Hao Liang‡ Bing-Yu Chen‡
*‡National Taiwan University †Keio University

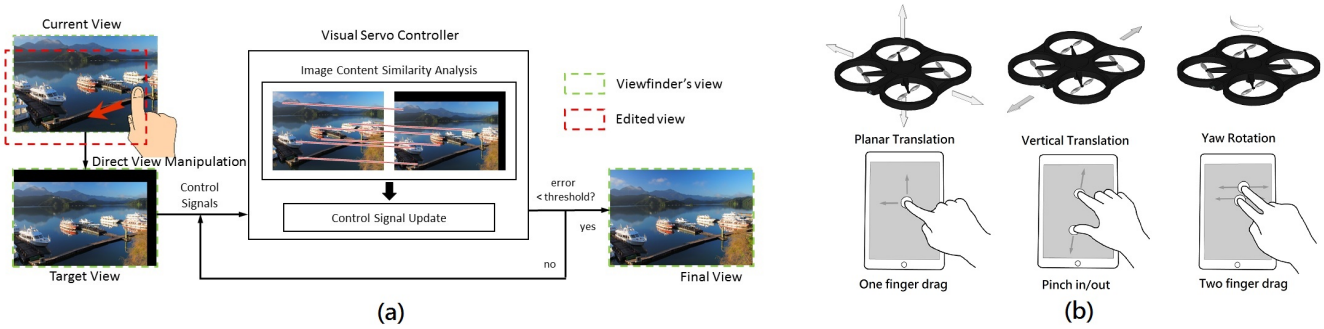

Figure 1: *(a) System workflow of the proposed interaction technique for drone photography. (b) Multi-touch gesture set supported by the proposed interface, which is designed based on the metaphor of photo manipulation.*

1 Introduction

For a long time, photographers hold and move their cameras, and consider how to frame a good shot all at the same time. With the emergence of drones, people start to let the flying carriers to hold their cameras in order to take more compelling pictures. However, the viewports between the photographer and device become decoupled and every single movement needs to be explicitly instructed via a remote controller. Even with the first-person view video streaming, users still have to be very skillful to fluently pilot the drone without causing distraction to photo composition. Inspired by the concept of *viewfinder editing* [Baek et al. 2013], we propose a more intuitive interface to control the flying camera (*i.e.*, the drone) by *direct view manipulation* embodied with multi-touch gestures, which allows the users to directly alter and rearrange the visual elements in the picture *prior* to image capturing. In our proof-of-concept implementation, the viewfinder of a flying camera is mapped to the screen of a mobile device. The physical camera movements are encoded by common photo manipulation operations, such as translation and scaling, with multi-touch gestures.

2 Direct View Manipulation

The design principle of the proposed interaction technique is to enable users to consider what a picture looks like instead of how to move the camera when piloting a drone. To this end, we choose to encode the physical camera movements (*e.g.* moving left/right and forward/backward) by a series of simple photo manipulation operations, such as translating or resizing a picture. In addition, multiple operations can be combined to indicate a more complex 3D movement of drone.

In practice, a user holds a mobile device with a touchscreen, which serves as a "viewfinder" displaying the live streaming from the camera mounted on a quadcopter. Once a touch event occurs, the current view becomes manipulatable and can be altered through typical multi-touch gestures. For example, an *one-finger drag* may translate the content within the viewfinder both horizontally and vertically. In addition, *pinch in/out* gestures may resize the image and reposition it with respect to the viewfinder. Though not related to photo manipulation, *two-finger drag* is used to perform yaw rotation to supplement the functionality of modifying drone's head orientation. The full gesture set is illustrated in Figure 1(b). It is worth noting that the target view not only indicates the user's attempt of drone movement, but also provides visual cues to guide the drone to better approach the desired view, which is accomplished by a *visual servo controller*.

We introduce an additional visual servo feedback loop to drive the drone to better match the user-specified view by image content similarity analysis, as demonstrated in Figure 1(a). Specifically, we employ SURF feature matching [Bay et al. 2006] to evaluate the *registration* error between the current and target view by computing the displacement of the matched keypoints. This information is interpreted into control signals to drive the drone to approach the target view. On the touchscreen, the user will have the impression that the live streaming is *attracted* to the target view. To let the user clearly understand the progress of visual servoing, we superimpose the gradient map revealing the salient structures of the target view as a semi-transparent layer to provide visual feedback of how the current view is aligned with the target view.

We implemented the proposed interaction scheme on an unmodified consumer quadcopter, *i.e.* AR Drone[1], which provides open SDK and rich resources of third party libraries. A user study was conducted on 10 participants with no or little experiences of piloting a drone. In general, the participants gave positive feedback about our system and recognized it as an intuitive interface that fits to their daily experience of manipulating photos on mobile devices.

References

BAEK, J., PAJAK, D., KIM, K., PULLI, K., AND LEVOY, M. 2013. WYSIWYG computational photography via viewfinder editing. *ACM Trans. Graph. 32*, 6, 198:1–198:10.

BAY, H., TUYTELAARS, T., AND GOOL, L. V. 2006. SURF: Speeded up robust features. In *Proc. ECCV '06*, 404–417.

*E-mail: {yiling,wlee}@cmlab.csie.ntu.edu.tw
†E-mail: liwei@kmd.keio.ac.jp
‡E-mail: {rhliang,robin}@ntu.edu.tw

[1] http://cdn.ardrone2.parrot.com/