I am the UAV: A Wearable Approach for Manipulation of Unmanned Aerial Vehicle

Yijia Lu[†], Fei Han[†], Lei Xie[†], Yafeng Yin[†], Congcong Shi[†‡] and Sanglu Lu[†]

[†]State Key Laboratory for Novel Software Technology, Nanjing University, China

[‡]Global Energy Interconnection Research Institute

Email: lyj@smail.nju.edu.cn, hfei@dislab.nju.edu.cn, lxie@nju.edu.cn, yyf@dislab.nju.edu.cn, shicongcong@geiri.sgcc.com.cn, sanglu@nju.edu.cn

***Abstract*—Nowadays, Unmanned Aerial Vehicles(UAVs) have been widely applied in our life. However, the existing approach of interacting with UAVs, i.e., using a remote controller with control sticks, is not a natural and intuitive way. In this paper, we present a novel approach for users to interact with personal UAVs using wearable devices. The basic idea of our approach is to manipulate UAVs based on human activity sensing, including motion recognition and pedestrian dead-reckoning. We have implemented the proposed approach on a DJI drone, and evaluated its performance in the real-world environment. Realistic experiment results show that our solution can replace the remote controller to manipulate the UAV.**

I. Introduction

Unmanned Aerial Vehicles(UAVs), commonly known as drones, are becoming increasingly popular and helpful in our daily life. Nowadays, UAVs have been widely used in numerous fields, such as package delivery, crop-dusting, film shooting, searching and rescuing. Nevertheless, to the best of our knowledge, most of the commercial UAVs are manipulated with remote controllers. For example, one control stick of the remote controller controls the UAV's up-and-down movements and rotations, while the other control stick changes the UAV's forward, backward, left and right pitch. This kind of manipulation may be not intuitive, or even difficult for unprofessional users to learn.

Under the circumstance, Human-Drone Interaction(HDI) has attracted much attention. The state of the art solutions in human-drone interaction mainly exploit machine vision techniques. Nagi et al. present a method using onboard video cameras for humans to interact with UAVs, with the help of face pose estimates and hand gestures[1]. Pfeil et al. explore upper body 3D spatial interaction metaphors for control and communication with UAVs, using the Microsoft Kinect[2]. FollowMe leverages a quadrocopter to follow a person and recognize simple gestures using an onboard depth camera[3]. Unfortunately, machine vision based interaction is usually limited to intensity of light and surrounding environments, which may have serious influence on the performance.

The good news is that the development of wearable devices, such as smartwatches and smartglasses, provides us with new chances to interact with UAVs conveniently and intuitively. In this paper, we propose a novel approach for human-drone interaction based on sensing the user's activity. As shown in Fig.1, with the help of wearable devices like a smartwatch and a smartglass, we are able to accurately perceive the user's arm movements and body movements, including the moving direction and distance, by leveraging accelerometers and gyroscopes. We use a smartphone as the server, which combines and optimizes the processed data from the smartwatch and the smartglass, and then sends commands to the remote controller wirelessly to manipulate the UAV.

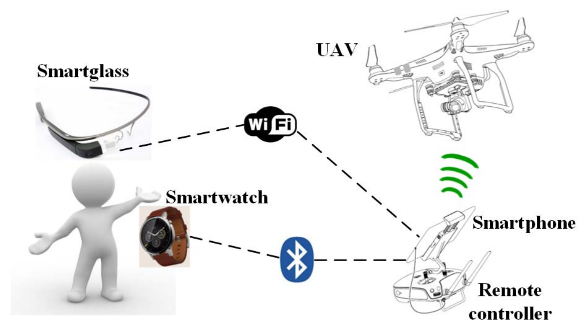


Fig. 1. Diagram of the proposed approach.

There are several challenges in designing a scheme for manipulation of UAVs based on wearable devices. The first challenge is that human-drone interaction requires higher accuracy and lower latency than other human-computer interactions, because the recognition error or intolerable delay may lead to the out of control or even crash of UAV. To address this challenge, we have particularly designed simple and discriminable motions for manipulating UAVs based on motion recognition. What's more, we have reduced the number of motion templates to save the time cost of matching. The second challenge is how to filter out unintentional motions of users. To address this challenge, we make full use of the consistency of inertial data on two wearable devices while walking, in other words, only when the smartglass and the smartwatch detect the similar movement, the manipulation will work.

We make the following main contributions in this demo paper: 1) We propose a novel approach for manipulating UAVs by using wearable devices. 2) We design a complete human-drone interaction solution which can take the place of a remote controller in most cases. 3) We implement the proposed approach on a DJI drone and two wearable devices, including one Google Glass and one Moto360 Smartwatch.

978-1-5090-6517-2/17/$31.00 ©2017 IEEE

II. System Design

A. System Overview

In our approach, we focus on how to manipulate UAVs based on the inertial sensors embedded inside the wearable devices, specifically, one smartwatch and one smartglass. Fig.2 shows the framework of our manipulation system.

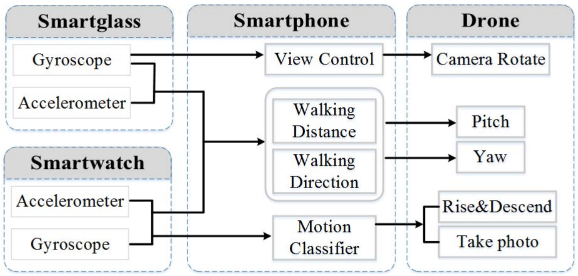


Fig. 2. Framework of the proposed approach.

The components of our system are as follows:

1) Smartglass: The smartglass senses the movement of the user's head via an accelerometer and a gyroscope. On one hand, it detects the user's steps and turns while walking. On the other hand, it detects the variation of the user's horizon. The results will be sent to the smartphone through WiFi.

2) Smartwatch: The smartwatch senses the periodic movement of the user's arm and detects steps and turns while walking, based on the inertial sensor readings. Besides, it senses the user's appointed motions by matching the current motion with trained templates. The results will be sent to the smartphone via Bluetooth.

3) Smartphone: The smartphone works as a server, linking the two wearable devices and the UAV. It receives the recognition results from the smartglass and the smartwatch, and then determines whether to generate a flight command according to the consistency of the recognition results.

4) UAV: The UAV gets commands from the smartphone, and then follows the commands to pitch, roll, yaw, rise, descend, etc.

B. Motion Recognition

We design four motions to control the flight of the UAV, including taking off/rising, descending, stoping rising or descending and taking photos. Table I shows the specific motions and the corresponding flight commands. The motions we designed are natural and simple, especially the motions for rising and descending, which make the user feel like flying with wings.

When we perform the same motion twice, the amount of inertial data in each motion period is different. However, the variation trend of the data is consistent. Fig.3 displays the variation of acceleration on one axis when performing the "rise" motion twice. We can find the two lines have different lengths but keep the same variation trend. Therefore, we cannot directly calculate the distance between the received sensor data and the trained motion templates. In regard to the motion recognition methods using machine learning, it often uses the training data to generate a classifier, leading to high

TABLE I
MOTIONS AND CORRESPONDING COMMANDS

Motion	Command
Lateral raise twice quickly	Take off Rise(if has taken off)
Lateral raise once slowly	Descend
Lift the arm to the chest	Stop rising(if rising) Stop descending(if descending)
Swing the arm forward and backward twice	Take photo

training cost. Fortunately, Dynamic Time Warping(DTW) is able to solve this problem perfectly. In time series analysis, DTW is an algorithm for measuring the similarity between two temporal sequences which vary in speed. DTW algorithm utilizes dynamic programming to find the optimal match of two temporal sequences by calculating the distance between them.

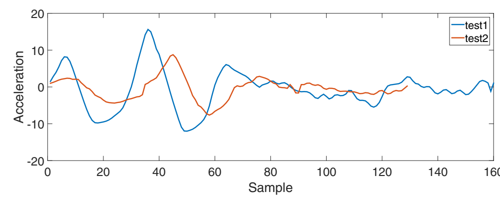


Fig. 3. Different sequences of the same motion vary in speed.

As shown in Table I, there are four motions in our system. To use DTW, we first establish the template of each motion. Each motion correlates with six kinds of sensor data, i.e., 3-axis acceleration readings a_x, a_y, a_z and 3-axis angular velocity readings w_x, w_y, w_z. Consequently, the template of each motion consists of the six kinds of sensor data. For simplicity, we only use one template for each motion.

For the received sensor data of a potential motion M_i, we calculate the DTW distance between M_i and a motion template $T_j, j \in [1, 4]$, as shown in Eq. (1). Here, d_{ax}, d_{ay}, d_{az}, d_{wx}, d_{wy}, d_{wz} mean the DTW distances in x-axis, y-axis, z-axis of acceleration and gyroscope data, respectively.

$$D_{DTW} = d_{ax} + d_{ay} + d_{az} + d_{wx} + d_{wy} + d_{wz} \quad (1)$$

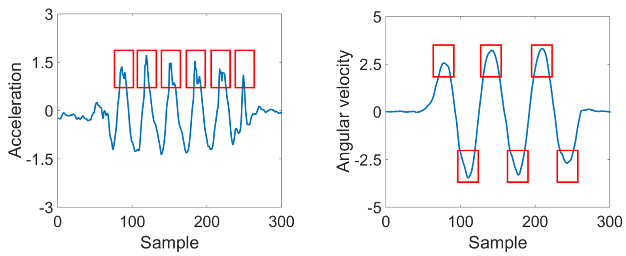
After calculating the DTW distance between the potential motion M_i and each template $T_j, j \in [1, 4]$, we respectively get four DTW distances D_{DTW1}, D_{DTW2}, D_{DTW3}, D_{DTW4}. Then, we compare the four results and select the motion with smallest DTW distance as the recognition result of the potential motion M_i.

C. Turn Detection

Although the information of the user's body turn can be extracted from the magnetometer, considering that most wearable devices have no magnetometer, we detect turns using the gyroscope data in our approach. Turn detection is based on the fact that the rotation axis of the human body during a turn is always along the direction of the gravity. However, the gyroscope measures the angular velocities of rotation on each axis of the device's body-frame. So it's necessary to transform the angular velocity from body-frame to earth-frame. We get the turn angle by calculating the integral of angular velocity around the gravitational direction on the smartglass and the smartwatch, respectively.

D. Step Detection

First, we use a low-pass filter to remove high-frequency noise in the inertial data. Then we utilize a sliding window to detect steps. For the smartglass, the acceleration along the gravitational direction has obvious peaks and each peak represents one step, as shown in Fig.4(a). While for the smartwatch, it rotates around the shoulder joint when the arm swings forward and backward. So both of the peak and the valley in data represent one step, as shown in Fig.4(b).



(a) Smartglass (b) Smartwatch

Fig. 4. Step detection.

E. Consistency Judgement

The two wearable devices detect user's turns and steps seperately, and then send the result to the server immediately. Consequently, it is important to judge consistency in our approach. We have two metrics for consistency judgement: 1) Magnitude. The angle of the turn received from two wearable devices must be similar in magnitude. If the difference between the two magnitudes is larger than the threshold, we consider they are not consistent. 2) Timestamp. When the server receives the result from the wearable device, it will record the magnitude as well as the timastamp. In addition to magnitude, the difference in timestamp should also be under a threshold.

III. Implementation and Evaluation

We have implemented the proposed approach using commercial devices[4]. The UAV is DJI Phantom3 Professional, which allows developers to use the mobile SDK to create a customized mobile app. The mobile SDK provides interfaces to flight and camera control of the UAV[5]. The smart devices are Google glass and Moto360 smartwatch, both running on Android platform. We use a Huawei MT7-CL00 smartphone running Android 6.0 as the system server. The wearable devices' sample frequency is set as 50Hz.

A. Accuracy of motion recognition

Fig.5 plots the confusion matrix for four motions. Each row represents the actual motions performed by the user and each column represents the recognized motion. Each element in the matrix corresponds to the probability of the motion in the row that is recognized as the motion in the column.

	Rise	Descend	Stop	Take photo
Rise	0.98	0.02	0.00	0.00
Descend	0.00	1.00	0.00	0.00
Stop	0.00	0.00	1.00	0.00
Take photo	0.00	0.00	0.00	1.00

Fig. 5. Confusion matrix for four motions.

B. Accuracy of turn and step detection

Experiment results show that the accuracy of step detection of our approach is above 92%. As for the accuracy of turn detecion, the average error of calculated turning angles is within 5°. The error of the UAV's turning angles is caused by two factors. First, when we send a specific value of angle to the remote controller of the UAV, the UAV will not turn the actual value we give. Second, there exists some deviation between the actual turning angle and the calculated turning angle, which is derived by integrating the angular velocities.

IV. Conclusion

In this paper, we present a novel wearable approach for convenient manipulation of UAVs. We use two common wearable devices embedded with inertial sensors to estimate the moving distance and the direction of the user. Besides, we sense the user's arm motions with an accelerometer and a gyroscope based on template matching. The realistic experiments show our approach is able to manipulate the UAV successfully.

Acknowledgment

This work is supported in part by National Natural Science Foundation of China under Grant Nos. 61472185, 61373129, 61321491, 61502224; JiangSu Natural Science Foundation under Grant No. BK20151390. This work is partially supported by Collaborative Innovation Center of Novel Software Technology and Industrialization. This work is partially supported by 2016 Program A for outstanding PhD candidate of Nanjing University under Grant No. 201601A008. Lei Xie is the corresponding author.

References

[1] Nagi, Jawad, et al. "Human Control of UAVs using Face Pose Estimates and Hand Gestures." *ACM/IEEE International Conference on Human-Robot Interaction*, 2014.
[2] Pfeil, Kevin, L. K. Seng, and J. Laviola. "Exploring 3d gesture metaphors for interaction with unmanned aerial vehicles." *International Conference on Intelligent User Interfaces*, 2013.
[3] Naseer, Tayyab, J. Sturm, and D. Cremers. "FollowMe: Person following and gesture recognition with a quadrocopter." *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013.
[4] Demo of "I am the UAV", http://cs.nju.edu.cn/lxie/demo/uav.mp4.
[5] DJI Developer, https://developer.dji.com/mobile-sdk/.