DESIGN AND EXPERIMENT OF A HIGH PAYLOAD FIXED WING VTOL UAV SYSTEM FOR EMERGENCY RESPONSE

Hao Wu¹, Zhongxiang Wang¹, *, Bin Ren², Liguang Wang³, Jun Zhang², Jie Zhu¹, Zihao He³

¹ National Geomatics Center of China, No.28, Lianhuachi West Road, Haidian District, Beijing, China - (wu.hao, wangzhongxiang, zhujie)@ngcc.cn
² CHENGDU JOUV AEPENG TECH CO.,LTD, 3A-8F, Jingrong Innovation Hub, No. 200. 5th TianFu St., Hi-tech District, Chengdu, China - (renbin, wangliguang, zhangjun)@jouav.com
³ The First Geodetic Surveying Brigade of MNR, Xi’an 710054, China - 931478830@qq.com

KEY WORDS: Fixed wing, Vertical take-off and landing (VOTL), Unmanned aerial vehicle (UAV), Emergency response, orthoimage

ABSTRACT:

With the development of UAV technologies, the advantages of hybrid VTOL UAV have been realized and taken in emergency response. But, former hybrid VTOL UAV is lack of capacities on payload and endurance, which restrict the integration of multiple sensors. In this paper, a high payload fixed wing VTOL UAV, which has 20 kg payload and more than 3 hours endurance, is used to design a UAV system for emergency response. Multiple sensors including an onopronics pod, PhaseOne IXM-100 camera, high accuracy inertial navigation system and three-axis stable head are integrated with it. Based on this, specific processing software is developed to process the video data and image which could meet the requirements of emergency response in different stages. Experiment results shown that the precision of mosaic image is about 10m and the precision of orthoimage is about 1m. This work could be reference for the design and practice of UAV system with multiple sensors.

1. INTRODUCTION

Unmanned aerial vehicle (UAV) has becoming almost an everyday occurrence according to the developments of technologies and cheaper cost (Mueller et al., 2017; Yan et al., 2019). With the benefit of carrying sensors, UAV based aerial photogrammetry has been widely used in various types of emergency response activities and treated as an eye in the sky (Arnous, 2011; Min, 2013; Zhu et al., 2014; Mansourej, Mehdi, 2017; Liu, et al., 2018). While, the former UAVs used in emergency response are mainly small light ones (e.g. the payload is about 5 kg and the endurance are about 2h) (Colomina, 2014; Yang, 2017). As results, most of the remote sensing sensors with high quality and precision cannot be taken in these UAVs, and the communication signal and even aircraft may be disturbed by severe weather when disasters occurred.

In recent years, the improvement of UAVs and remote sensing sensor technologies mainly concentrate in three aspects (Cao, et al., 2018). 1) Some industrial-strength UAVs with high payload and long endurance which were used in military have been introduced to civil area (Xu, et.al, 2019; Yang, et al. 2019) 2) With the improvement of sensor types and precision, the reduced height and size make them more flexible (Hu, et al., 2015; Abdullah, et al., 2019). 3) The processing software for multiple types of remote sensing dataset could be operated in a more automatic way (Zhou, et al., 2015), and are becoming more specific to adapt to various applications (Lorenzo, 2016; Ashray, et al. 2017; Sana, et al., 2019).

According to the way of take-off and flying, these UAVs could be divided into 3 types: fixed wing, vertical take-off and landing (VTOL), and hybrid VTOL UAV (that use rotor for take-off and use fix wing for cruise). But, most of the disasters especially geological disasters occurred in mountainous regions making it difficult for fixed wing UAVs to find an appropriate place for take-off, and the endurance of VTOL UAVs is always not so long enough because most of them rely on battery power. In this situation, the hybrid VTOL UAV has significant advantages in adapting to complex environment and long time working.

This paper introduces a system for emergency response based on a high payload fixed wing VTOL UAV (named CW-100, which is produced by CHENGDU JOUV AEPENG TECH CO.,LTD) with 20 kg payload and more than 3 hours endurance. Firstly, based on the requirement analysis of emergency response, optronics pod (including visible and infrared video camera), professional aerial camera, and high precision aerial inertial navigation system are selected and integrated together with UAV and ground station. Then, specific software is developed to process the data acquired by different sensors. For example, the video obtained by optronics pod will be transformed timely from aircraft and then be produced to mosaic image for decision making in the first time. The image obtained by camera will be processed to high precision orthoimage when landed to meet the needs of further analysis. This idea proposed in this paper and the processing software specially developed for video could be reference for the design and practice of UAV system with multiple sensors.

2. MAJOR DESIGN

In the process of emergency response, timely and effective image of disaster area is one of the essential bases. This refers to three points: rapid obtaining, timely transmission and high precision processing. To achieve these targets, the UAV system should be consisted of UAV aircraft, remote sensing sensors, ground station and processing software. Moreover, as the size of video is relatively smaller than image, video camera is the most appropriate way to acquire the scene of disaster area in the first time. Considering that the disaster may happened random in day or night, so that both visible and infrared video camera are needed.

* Corresponding author
Then the video data and image would be processed to meet the requirements of emergency response. Additionally, to improve the precision of processing results, position and orientation system (POS) and three-dimensional self-stable head would be integrated together with the remote sensing sensors to acquire the attitude information of aircraft and then adjust the attitude of sensors. Based on the above analysis, the components of the UAV system could be organized as Figure 1.

![Figure 1. Major Components of fixed wing VTOL UAV system](image)

Using this system, when a disaster occurred, we could design a fly route using passage planning software equipped in ground station. Then, the UAV system could take off from about 50 km away from the target region. When the aircraft has arrived at the target region (this may take about half an hour), the sensors will be opened and start working according to the predefined route. Specially, based on the long distance communication device integrated with ground station, the videos acquired by the optronics pod will be sent to ground station in real-time. Then the videos will be processed to mosaic image timely and sent to command centre for decision making by 4G or other wireless network. For the image obtained by aerial camera will be processed to high accuracy orthomage after the aircraft have landed and then used for further analysis. Specially, the ADS-B (automatic dependent surveillance-broadcast) transponder device is used in the aircraft, which could transport the radar signal to other aircrafts to avoid a crash. The workflow could be shown as Figure 2.

### 3. DESIGN OF INTEGRATION

#### 3.1 Integration of remote sensing sensors with UAV

Remote sensing sensors are the most important parts of the UAV system. Table 1 list the major indicators of selected sensors according to the capacity and payload of CW-100 UAV.

![Figure 2. Workflow of fixed wing VTOL UAV system](image)
<table>
<thead>
<tr>
<th>Sensor type</th>
<th>Model &amp; Producer</th>
<th>Major indicators</th>
</tr>
</thead>
</table>
| Optronics pod    | MG200A3, JOUA V           | ■ Visible video camera
                   | ■ Focal length: 50 mm, prime
                   | ■ Number of pixels: 1920×1080
                   | ■ Infrared video camera
                   | ■ Focal length: 50 mm, prime
                   | ■ Number of pixels: 1920×1080
                   | ■ Operating wavelength: 8-13 μm |
| Arial camera     | iXM-100, Phase One        | ■ Number of pixels: about 100 million
                   | ■ Focal length: 80 mm
                   | ■ Physical size of the pixel: 5.2 mm |
| POS              | AP20, Trimble Appianix    | ■ Performance based upon external IMU
                   | ■ Position (m): 1.5 - 3.0
                   | ■ Velocity (m/s): 0.05
                   | ■ Roll & Pitch (deg): 0.03
                   | ■ True Heading (deg): 0.10 |
| Self-stable head | RS-M150, Changguang Ruishi| ■ Max load-bearing: 20 kg |

Figure 3. The interfaces of different sensors

Figure 4. The integrated relation of different sensors

This contribution has been peer-reviewed.  
https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-1715-2020 | © Authors 2020. CC BY 4.0 License.
3.2 Integration of remote sensing sensors with Ground Station

Ground station consists of control system and linking device. Control system is used to send orders by linking device to the aircraft, which will parse it and then control the status of sensors. In the other hand, it can also receive and display the status of aircraft and sensors on the screen. The main components and their relations are shown in Figure 5.

While, to better control the sensors and keep the security of aircraft, a specific passage planning software is developed and integrated in the control system. Digital elevation model (DEM) data, digital orthograph model (DOM) data and online map service (e.g. Google map) could be inputted and managed in this software. Based on these data, a series of function are developed to define the mission area and route line. 1) Define the mission area. We could draw a polygon or input a series of coordinates to define the mission area, and then edit some of the points on the map. 2) Define the parameters for aerial photography. The parameters for aerial photography, e.g. relative flying height, spacing of photos, spacing of route line and the parameter of camera, then all the route lines and photo points will be computed and output on the screen. 3) Define multiple mission areas. If we need the aircraft to work on multiple mission areas, we could set the parameters of each area by step 2) and then join the areas together to get the whole route line. 4) Check the route line. When the whole route line is generated, we could use DEM to check the route line. If there are any security risks threatening the aircraft, the software will remind to modify it.

![Figure 5. The main components and their relations of ground station](image)

4. DESIGN OF PROCESSING SOFTWARE

4.1 Software for Video Processing

Video is a common media, with the characteristics of intuitive, accurate, dynamic, and real-time. It contains rich spatial information and can truly express geographic spatial scenes. However, due to the lack of geographic information, traditional aerial video is only used as visual information in military fields such as enemy reconnaissance and target tracking, or in civil fields such as power line patrol and emergency monitoring.

The proposed video processing software encodes the video on geographic coordinates, and integrates it with geographic information system (Video GIS), which can realize the spatial positioning analysis of video data, facilitate the management and retrieval of video data, and improve the effect of GIS visualization. The video processing software mainly includes three parts:

1) Video geocoding module, which encodes video data and geographic information data, establishes the corresponding relationship between video frame and geographic information, and provides guarantee for the subsequent geographic positioning of image.

2) Video real-time splicing module, which makes use of the matching relationship between video frames, extracts video key frames, and calculates the relative position relationship between video frames.

3) Geolocation module, which makes use of the matching relationship between video key frames and geographic location information to realize the geographic positioning of each pixel on the video key frame splicing image.

4.1.1 Video geocoding module

Video geocoding requires synchronous return of video data and geographic information. Only in this way can the accuracy of geocoding of video frames be guaranteed. However, in the ordinary ways, video data is transmitted through a real time line image transmission system (that is, image transmission equipment); geographic information is transmitted back to the ground through a digital transmission station. The two transmission sets are not synchronized. Thus, this paper presents a method to synchronously package video data and corresponding geographic information into video data, ensuring the synchronization of video data and geographic information. The main technical process of video geocoding is shown in Figure 6.

![Figure 6. The flowchart of video geocoding](image)
a) Video capture device collects video data and encodes it as H.264 (or H.265);
b) Using serial communication, the acquisition device obtains the IMU and GPS data of flight controller, and IMU data of the stable platform which are corresponding to the current video data. That is, the geographic information corresponding to the current video frame is obtained, and it is packaged as POS data according to the protocol;
c) POS data is encoded into H.264 (or H.265) format;
d) The H.264 (or H.265) encoded data of the video stream and the H.264 (or H.265) encoded data of the POS data are fused into a H.264 (or H.265) data stream;
e) The fused H.264 (or H.265) data stream is encapsulated into video stream with TS format;
f) The encapsulated TS format video stream is sent to the ground equipment by image transmitter.

The data receiving terminal is placed on the ground, and after establishing communication with the transmitting terminal, the video stream including the geocoding can be received in real time. The video geocoding module receives the video data, and distributes it and the POS data according to the video protocol. Then, the one-to-one correspondence between the video frame and the POS data is obtained, that is, the video geographic decoding is completed.

4.1.2 Real-time video splicing module

Video images have the disadvantages of small field of view, narrow image width, small coverage area, etc., and cannot intuitively reflect the whole view of the measurement area. Therefore, video frames are often spliced in applications. The real-time video splicing proposed in this paper is to select video frames (key frames) that meet certain conditions and have overlapping areas from the real-time video stream, and then complete the conversion of each key frame to navigation coordinates, according to the transformation relationship between them. Because of the overlapping relationship between each key frame, it is realized that each key frame is spliced into the entire image.

Among them, the transformation relationship between the key frame and the navigation coordinate system is shown in Equation 1. In the whole transformation, the internal matrix $K$, rotation matrix $R$ and translation vector $t$ of are involved. According to the transformation relationship, each frame of image can be transformed into the navigation coordinate system.

It can also be seen from Equation 1 that the internal and rotation parameter matrices are the key to achieving key frame stitching with high accuracy. The internal matrix $K$ of the camera can be obtained by the camera calibration with a initial value. The initial values of $R$ and $T$ can be obtained through the matching relationship of the feature points of adjacent video frames, or from the IMU + GPS data. However, the accuracy of these initial values is not high. If these values are used directly for splicing, the error will be large. The essential problem of splicing is also about how to improve the accuracy of $K, R$, and $T$.

Thus, a co-visual relationship is built between key frames, then initial $K, R$ and $T$ values are optimized. The optimization is adopted the beam adjustment. Which adopts a graph optimization model, where the vertices are composed of the camera POS and the point X in three-dimensional space, and the edge constraint is the error between the observed data and the estimated data.

$$V_p = \text{obtained by elevation iteration using the flight control module. The positioning process from image point to object point is shown in Figure 7.}$$

In the target position calculation, the latitude and longitude coordinates of the target in the WGS84 geodetic coordinate system are calculated, and compared with the latitude and longitude height of the target on google earth, to measure the target position calculation error. According to the camera’s attitude information in the WGS84 geodetic coordinate system, it is divided into two cases: ortho and tilt.

2) Geolocation of splicing image

In real-time video splicing, each video key frame is synchronized with POS data. First, the POS data of each video key frame is extracted, the geographic location of each video key frame is calculated. Then the geographic ground location of splicing image is obtained using the pixel-to-object positioning scheme (as described above), that is, Longitude, latitude, and height of the upper left and the bottom right corner pixels are obtained. Combined which and the spatial resolution, the longitude and latitude corresponding to the pixel can be calculated by the following formula.

$$
\Delta\text{Lon} = \frac{\text{RightBottom\, lon} - \text{LeftTop\, lon}}{W}\\
\Delta\text{Lat} = -\frac{\text{RightBottom\, lat} - \text{LeftTop\, lat}}{H}
$$

$$V_p$$ is obtained by elevation iteration using the flight control module. The positioning process from image point to object point is shown in Figure 7.
4.2 Software for image Processing

The optical area array image processing module includes a quality inspection module and an image fast processing module.

1) Quality inspection module, firstly it conducts a quick quality check for the acquired images, to check whether there are missing or multiple shots, and whether the image range covers the entire area. Then it carries out high-precision one-click fast splicing function, super-large image data processing and multiple sorties processing.

2) Fast image processing module, combined with UAV image fast matching, block adjustment without GCP (ground control point), POS data aided aerial triangulation technology, the mass image data fast processing is completed.

4.2.1 Quality inspection module

After the flight is completed, the aerial photos and POS data are imported into the software, and the photos are displayed on the graphical interface according to the POS information. The quality of the photos and the flight status are intuitively checked, and the leakage of POSs is detected. Then the range line is imported to check whether the flight covers the entire area.

After the POS data editing is completed, the overlap of the pictures on neighboring image strips or on the same strip are automatically calculated with matching points. The longitudinal overlap and side overlap are checked whether meets the requirements. When tree or water is full of water on the photo, or when the features are not obvious, the automatic calculation of the overlap will result in errors due to the lack of matching points. Then hand pricking point is needed for calculation of overlap and image rotation angle.

4.2.2 Image fast processing module

UAV image fast matching, block adjustment without GCP, POS data aided aerial triangulation, UAV images fast splicing, algorithm of dodging uniform colour of UAV images, UAV images parallel processing and other technologies are used to complete massive images fast processing, forming a platform for rapid acquisition and processing of aviation remote sensing data, providing timely and efficient surveying and mapping support for geological disaster prevention and emergency rescue, and improving the scientific prevention and control capabilities of geological disasters and the efficiency of emergency response. The specific technical route is as follows:

1) Automatic matching of multi-type UAV image data

When a single image matching algorithm can no longer meet the requirements of UAV image processing, a variety of matching strategies are used to establish rough geometric constraints, under the constraints of initial feature matching, and to improve the matching accuracy through repeated constraints of feature and geometry.

2) Relative orientation solution

The motion parameter estimation method is used to estimate the connection matrix in the coplanar constraints with cognominal points, thereby restoring the relative position, avoiding the problem of the initial value required in the traditional solution process, and making the solution more rigorous.

3) Storage and solution of normal equation

For the processing of large amounts of data, an index mechanism that can save storage space and facilitate the storage of sparse matrix elements is designed, and the computer's parallelization strategy can be used.

4) Block adjustment of irregular strip overlap

Facing with the characteristics of long strip and irregular side overlap of UAV images, the traditional nonlinear correction and block adjustment method are improved. Considering all strips, automatically generated connection points and GCP points are used for adjusting non-linear correction block adjustment, so as to reduce the error between strips.

5) Gross error detection of macroscopic measurement data

Considering that the least squares method is very sensitive to gross errors, and the least squares is the most basic method. The least squares correction number is used as a basis to find a way to detect gross errors, reduce the weight of gross error, and reduce the impact of gross errors on the final solution results.

6) System error model

UAV photogrammetry uses non-measurement cameras, and the acquired data has more obvious system errors than traditional aerial remote sensing data. Therefore, the system errors need to be compensated before adjustment, and the commonly used error correction model is not complete. Thus, pay attention to the error rule of the UAV image system, select a reasonable system error model and parameters to compensate for the system error in the UAV image data, and it will be effectively improved adjustment accuracy.

7) POS data aided aerial triangulation

Using the POS data to assist the aerial triangulation, so as to achieve the rapid positioning of the UAV image and the efficient production of image maps.

5. EXPERIMENTS

The experimental area is located near 104°25′46″E and 31°36′10″N, Beichuan County, Mianyang City, Sichuan
Province. It covers an area of about 5 square kilometres, with a rectangular shape and a height of about 100 meters, hilly terrain. The plane coordinate system is WGS84, the coordinate projection is Gauss Kruger, the central meridian is 105 degrees, and the elevation datum is WGS84 geodetic height. The fast splicing image are produced in the mode of no GCP. See Table 2 for the basic information:

<table>
<thead>
<tr>
<th>Relative row height (m)</th>
<th>637</th>
<th>Ground resolution (cm)</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aerial range (km×km)</td>
<td>2×2.5</td>
<td>Flight time (min)</td>
<td>41</td>
</tr>
<tr>
<td>Focal length of optical camera (mm)</td>
<td>80</td>
<td>Focal length of photoelectric pod (mm)</td>
<td>4.2</td>
</tr>
<tr>
<td>Video resolution</td>
<td>1080P</td>
<td>Number of acquired images</td>
<td>554</td>
</tr>
<tr>
<td>Amount of acquired video data (GB)</td>
<td>1.27</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Firstly, automatic and manual extraction of video key frame image is conducted. Then, geocoding of key frame image, and real-time and fast splicing of key frame image are conducted. The result is shown in Figure 10.

3) Image data processing
Based on the existing image and the embedded POS external orientation elements of each image, the key frame image is corrected to achieve effective registration with the existing emergency data.

Combined with the quick quality inspection function of the system, it checks whether the image is missing, whether each image can be read and displayed normally, whether the image range is consistent with the planning range, and whether there is cloud coverage, etc. The results are shown in Figure 11.

Combined with the multi sorties acquisition image processing function of the same sensor of the system, the 10GB image data is conducted with "one key fast splicing" (Figure 12). The results show that the time used is less than 24min (table x below), which is about 25min less than the 50min specified in the technical requirements.
In order to test the accuracy of the processing results, CORS RTK mode is used to collect 38 feature detection points, and the statistical plane mean square error is selected as the plane accuracy evaluation index. The results (Table 3) show that when 500 frames are extracted, the accuracy of video splicing can reach about 10 meters, the plane accuracy of optical camera image splicing can reach about 1 meter, and the processing time is about 23 minutes.

<table>
<thead>
<tr>
<th>Result Type</th>
<th>Inspection category</th>
<th>Inspection result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video fast splicing image</td>
<td>Time performance</td>
<td>23min19s</td>
</tr>
<tr>
<td>Aerial optical camera fast splicing image</td>
<td>Plane accuracy</td>
<td>±10.346m</td>
</tr>
<tr>
<td>Aerial optical camera fast splicing image</td>
<td>Plane accuracy</td>
<td>±10.93m</td>
</tr>
</tbody>
</table>

Inspection points number: 38, Computer performance: Win10 64 CPU i7-9750@2.6 GHz 32G

### 6. CONCLUSION

In recent year, hybrid VTOL UAV has achievement a rapid progress, and became an important method to acquire the information in emergency response. But, the lack of capacities of payload and endurance of former UAVs making it is difficult to integrate multiple sensors. In this paper, a high payload fixed wing VTOL UAV, which has 20 kg payload and more than 3 hours endurance, is used to build a system for emergency response. The remote sensing sensors including an optronics pod, PhaseOne IXM-100 camera, high accuracy inertial navigation system and three-axis stable head are used and integrated with it. Moreover, specific processing software is developed to process the video data and image which could meet the requirements of emergency response in different stages. Especially, the video data acquired by optronics pod could be transmitted and processed timely when the aircraft is flying. This could significantly improve the time for obtain the scene of target area. This idea and software proposed in this paper could be reference for the design and practice of UAV system with multiple sensors.

### 7. REFERENCES


Geomatics and Information Science of Wuhan University, 43(12), 2250-2257. DOI: 10.13203/j.whugis20180309. (In Chinese)


