THE JOINT RESEARCH PROJECT ANKommEn – EXPLORATION USING AUTOMATED UAV AND UGV

A. Alamouri¹, M. Gerke¹, S. Batzdorfer², M. Becker², U. Bestmann², M. Bobbe², Y. Khedar², T. Blume³, J. Schattenberg³, J. Schiemann³

¹ Technische Universität Braunschweig, Institute for Geodesy and Photogrammetry, Braunschweig, Germany
{a.alamouri, m.gerke}@tu-braunschweig.de

² Technische Universität Braunschweig, Institute of Flight Guidance, Braunschweig, Germany
{s.batzdorfer, m.becker, u.bestmann, m.bobbe, y.khedar}@tu-braunschweig.de

³ Technische Universität Braunschweig, Institute of Mobile Machines and Commercial Vehicles, Braunschweig, Germany
{t.blume, j.schattenberg, j.schiemann}@tu-braunschweig.de

KEY WORDS: Emergency, Exploration, UAV/UGV, Database management, Web application

ABSTRACT

Disasters such as floods, large fires, landslides, avalanches, or forest fires are often inevitable and cannot be fully prevented, but their impact can be minimized with sound disaster management strategies aided by the latest technological advancements. A key factor affecting these strategies is the time, where any delay can result in dramatic consequences and potentially human losses. Therefore, a quick situation report of the disaster is highly demanded, but still not an easy task because - in most cases - a priori known spatial information like map data or geo-databases, are outdated. In addition, visual and geometric information on the current situation is needed to help rescue teams and first responders. From this point of view, we came up to the main idea of the joint research project ANKommEn and its extension ANKommEn 2 (german acronym for Automated Navigation and Communication for Exploration). The project idea embodies an exploratory investigation to be smart in providing correct and timely geodata that can help in emergency cases; especially in support decision making in emergency risk management. For this purpose, automated unmanned systems, both ground (UGV) and airborne (UAV), are being developed to provide up-to-date information of rescue scenarios.

1. ANKOMMEN - SYSTEM DESIGN

The overall system for data recording consists of Unmanned Aerial Vehicle (UAV), Unmanned Ground Vehicle (UGV) and a central ground and control station. The ground station serves as a central human-machine interface to monitor and manage operation of the UAV and UGV by an operator. The system enables an exploration process of an area of interest based on initial geodata available like a priori known map. Exploration areas and tasks are updated with current information of the visual sensors such as combined RGB and IR camera within one UAV-payload and LiDAR while performing a mission. This is done by a data streaming to the ground station via data link functionalities. For data capture and preprocessing where applicable, a processing unit is integrated onboard the UAV/UGV. Georeferencing of the exploration data is done using a custom built-up GNSS/IMU system.

Project partners of the joint research projects ANKommEn and ANKommEn 2 are Institute of Flight Guidance (IFF), Institute of Mobile Machines and Commercial Vehicles (IMN) and Institute for Geodesy and Photogrammetry (IGP), all Technische Universität Braunschweig, and AirRobot GmbH & Co. KG. For supporting the development and field tests, associated project partners are included in the project consortium. These are the Fire Brigade of Braunschweig and the Lower Saxony Water Management, Coastal Defence and Nature Conservation Agency (NLWKN).

The UAV and UGV used for carrying the built up sensor sets, data acquisition and processing hardware are depicted in Figure 1 and Figure 2. As UAV a AR200 hexacopter manufactured by AirRobot has been customized concerning the upper payload compartment (black box on the front in Figure 1) for integration the basic hardware setup for positioning (GNSS/INS), data acquisition, processing and data exchange within the ANKommEn-system using different data links, “Nav-Com-Stack” shown in Figure 3. Using a proprietary connector at the bottom of the UAV different modular payloads carrying visual sensors for exploration and mapping tasks are mounted at the UAV. The GNSS antenna is mounted on the top right, antennas for data exchange using WLAN and 4G mobile data link are mounted at the landing gear and bottom side of the upper payload compartment.

Figure 1: UAV AR200 (customized for ANKommEn)

For ground based exploration tasks an UGV of type Robotnik Summit XL is used for carrying the NAV-Com-Stack (mounted inside the chassis), GNSS and data link antennas as well as visual sensors for exploration tasks, Figure 2.
The configuration shown in Figure 2 is equipped with a LiDAR (Velodyne VLP-16), a thermal camera (Flir A65) and a Pan-Tilt-Zoom (PTZ) camera integrated in the front of the chassis. The basic NAV-COM-Stack is depicted in Figure 3.

On the left the combined RGB-IR camera payload is depicted, carrying and stabilizing a RGB camera (Allied Vision Manta G-917C) as well as a thermal camera (Flir A65) for simultaneous exploration within visual and IR spectrum.

For mission and task definition, controlling the UAV/UGV operations as well as exploration data processing and visualization a central ground and control station (GCS) has been designed and built up, shown in Figure 5.

Therefore an industrial computer, GNSS and a communication board are integrated within the GCS setup serving as HMI for the overall system. Using the proprietary control software, different missions and tasks can be defined by an operator and assigned to different vehicles. In order to get exploration data and status information of the UAV or UGV a second mobile GCS has been built up for supporting in-site operators, shown in Figure 6 using the same hardware setup for communication and GNSS based positioning.

2. SYSTEM COMMUNICATION

For fast and reliable data exchange, an intelligent communication structure has been developed. The system’s structure is illustrated in Figure 7. Agents within range to each other are able to communicate directly using ad-hoc WIFI, which offers high throughput (e.g. the UGV and UAV on the bottom right of Figure 7). In case of long distance communication, it’s even possible to communicate without a direct link, utilizing data hopping over arbitrary agents, e.g. the ground station communicates with the UAV, utilizing the UGV as a range extender (bottom of Figure 7). In addition to WIFI, all agents are equipped with 3G/4G modems, which enable communication via internet. Communication via internet is encrypted and secured using a central VPN server.
Figure 7: Flexible communication structure between arbitrary agents using decentral ad-hoc networking and internet

The system adaptively aggregates multiple physical communication channels between agents, when available. Furthermore different compression algorithms were implemented, to reduce bandwidth of data streams, which allows delivering more data. Additionally data channels can be prioritized by the user. This enhances flexibility and ensures that data is delivered w.r.t. the user’s preferences e.g. delivering live video feed, which is crucial for controlling the agent, in real time with superior priority, while less sensitive data with inferior priority is buffered where required and delivered when free capacity is available within the network (Quality of Service).

For convenient data access and storage, a central database server is integrated within the VPN, which delivers all gathered data to arbitrary clients.

3. MAPPING BASED ON PHOTOGRAMMETRY AND SLAM

Mapping of the desired area is done with two different methods having complementary strengths. A SLAM based approach is used to provide a real time ortho map of the scene to the operators even during the flight while conventional photogrammetry is used to generate a high-quality ortho map of the area after the mission is finished. In the following first the SLAM-based solution is described. The algorithm aims at generating a high quality orthophoto and textured 3D meshes in real-time. The proposed solution is published as an open source library called OpenREALM (Kern, 2018) on GitHub. The basic flow of the library works as follows:

- Geotagged images are first passed to a pose estimation stage. There, an interface class allows integrating state of the art visual SLAM frameworks for motion recovery (currently only ORB_SLAM2 (Mur-Artal, 2015) is used but can be extended easily with other methods).
- Extracted camera poses and the sparse cloud in the local coordinate system are subsequently transformed into a Cartesian frame using a best-fit Umeyama algorithm (Umeyama, 1991).
- Densification of the tie points is applied next. External 3D reconstruction frameworks can be integrated into the mapping process to provide high quality, dense surface information i.e. Plane Sweep Library (Haene, 2014).
- Afterwards the rectification of the image data is carried out. By using the linear pinhole model, the prior extracted elevation grid map is re-projected into the camera so the relief displacement induced by the terrain and perspective model is corrected.
- In the last stage all incremental images are finally integrated into a global mosaic and blended to correct for any roughness.

Figure 8 shows the architecture of the library. As described above, the different steps represent the various stages which make use of the underlying library layer for all its function. The transport layer is implemented in the ROS framework. More details about the project can be found from the Github page (Kern, 2018).

Figure 8: The architecture of the OpenREALM

The described SLAM-based solution relies on constant features which might not be present in some scenarios (i.e. water, sand, snow). Therefore, a fallback solution using only GNSS location and heading information was also implemented. This allows the system to generate a map even if there are not enough features. Even though the consistency and accuracy of the map obtained with the fallback solution is affected badly, it still provides a reasonable view of the mapped area. More can be seen in the Figure 9, where the various options from this solution and the conventional photogrammetry are compared. Depending on the resources available on the onboard PC i.e. CUDA compatible GPU, the OpenREALM can be executed with or without 3D reconstruction.

Figure 9: A comparison of results achieved using visual mapping approaches

This contribution has been peer-reviewed. https://doi.org/10.5194/isprs-archives-XLI-2-W13-165-2019 | © Authors 2019. CC BY 4.0 License.
The other solution uses an offline photogrammetry pipeline based on Agisoft Photoscan’s Python API (Agisoft, 2019) was implemented to enable high quality results in variety of formats. The process is automatically triggered after the last image of a survey is saved by the drone. The method first aligns all the available images using the features from the individual image. The GPS info available in the images stored as ExifTags is further used to aid in the alignment process. Once the images are aligned, a dense cloud is generated. The images are ortho-rectified and an ortho-mosaic is generated. The process exports the results in various different formats including the dense cloud and ortho map which is later used in the web services to visualize (Figure 10).

As can be seen from the Figure 9 qualitatively, the ortho-map from the Photoscan is the most consistent in comparison to the SLAM based solution. This advantage comes at the cost of computational resources required to for the processing. SLAM solution which uses the 3D reconstruction on the other hand is slightly worse but has the advantage that the map is available for the user to see in real time which can be critical in some scenarios. The relative integrity of the Slam based solution can be quantified by taking Agisoft ortho-map as the ground truth. The deviations of all the object locations from Figure 10 are then averaged and calculated for all the methods. The results are presented in Table 1. It’s clear that the planar assumption results in the highest error and when the elevation is considered by taking into account the sparse or dense cloud, the relative integrity is much better.

Table 1: Relative orthophoto integrity errors for all approaches

<table>
<thead>
<tr>
<th>Approach</th>
<th>Assumption</th>
<th>Relative Error (m)</th>
<th>Average</th>
<th>Std.dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNSS + Heading</td>
<td>Planar</td>
<td>2.60</td>
<td>1.60</td>
<td></td>
</tr>
<tr>
<td>Visual SLAM</td>
<td>Planar</td>
<td>5.10</td>
<td>6.30</td>
<td></td>
</tr>
<tr>
<td>Sparse Interpolation</td>
<td>Elevated</td>
<td>0.86</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>3D-Reconstruction</td>
<td>Elevated</td>
<td>0.89</td>
<td>0.85</td>
<td></td>
</tr>
</tbody>
</table>

4. POINT CLOUD PROCESSING AND CLASSIFICATION

Point cloud data is collected using Velodyne VLP 16 sensors (see Figures 2 and 4). On both, ground and aerial vehicles, the data processing is done onboard (see Figure 3). First step in data processing pipeline is registration of the actual sensor data within a common frame of reference. In a first step, the actual reading from the sensor is pre-aligned using GNSS/IMU. Since the obtained position and orientation (hereafter a combination of translation and orientation is referred to as pose) may be corrupted due to noise or shading effects, a fine registration step is applied afterwards.

Incoming scans are matched to a local reference map. An earth fixed reference is obtained using Graph-SLAM approach, which relates a locally constructed pose graph with a global one (Schmiemann, 2017). Interconnecting both, mainly based on time information, and minimizing the residual in least mean squares sense, leads to an optimal alignment and eventual a reference pose can be obtained, which transforms the local Pose-Graph into a common frame of reference (e.g. UTM).

The mapping approach illustrated above, works well for single vehicles, but tends to non-consistency, if multiple maps are generated by independent agents. To generate more consistent maps, matching algorithms were developed and implemented. Main challenges in this domain include handling of unorganized point cloud data, the huge amount of points per map and generation of meaningful features to match with little overlap and huge difference in perspective.

To achieve accurate matching performance, semi supervised learning algorithms were utilized. Feature learning is performed directly on point clouds using PointNet (Qi, 2016), which originally was designed for classification tasks. The structure can be subdivided into two components.

For each point within the map a representation is learned, which then is pooled, to derive a global feature for the complete map. Those map-wise features are then fed to a classifier, which can be trained e.g. to approximate a probability distribution over class affiliation.

We use the feature abstraction layer only and train with triplet loss, which aims to minimize feature distance to positive matches whilst maximizing feature distance to negative matches (Arandjelovic, 2015). A match can then be found by querying the nearest neighbors within a given set of database maps and thresholding. Figure 11 exemplary shows four different maps of overlapping areas, which shall be merged, whereby color encodes height.

![Figure 10: The relative integrity of the maps generated from SLAM in reference to the PhotoScan after registration](image)

![Figure 11: Four maps generated by different agents in urban scenario](image)
Figure 12 provides bird’s-eye view of all four maps (see Figure 11), merged within a common frame of reference in overlay with satellite imagery. Additionally the trajectories of the agents collecting the map data is drawn in light red, yellow, dark red and purple respectively.

Figure 12: Merged maps of Figure 11 overlaid with satellite imagery. Trajectories of agents are drawn in color. Trajectory for map (a) is shown in light red, trajectory for map (b) is shown in yellow, trajectory for map (c) is shown in dark red and trajectory for map (d) is shown in purple.

Further data processing is done to increase the comprehensibility of generated maps. Therefore the raw point cloud data, which only contains positional information and an additional channel for intensity, shall be enriched with semantic information. The enrichment process can be subdivided into three crucial tasks, which are abstraction of geometrical features, clustering and predicting semantics.

For each point, a covariance matrix of the local neighborhood is generated. The ratios of the dominant eigenvalues indicate spatial relations, e.g. if one eigenvalue is orders of magnitude greater than all the others, one can assume a one dimensional neighborhood in which all points form a somewhat linear shape. On the other hand e.g. if all three dominant eigenvalues are equally ranged, one can assume the local neighborhood to form some kind of ellipsoidal shape (Guinard, 2017).

With those feature abstractions, clustering algorithms can be used, to group geometrically homogenous areas. Results of the clustering process, using algorithms introduced by (Landrieu, 2016), are shown in the center of Figure 13. One can see that the clustering process produces geometrically homogeneous areas on different scales, e.g. complete facades of buildings containing multiple ten thousands of points are grouped to a single cluster, whilst small groups of less than hundred points, e.g. representing a bush, are grouped in single clusters, as well. The result of the last step is shown in the bottom of Figure 13. With the clusters as nodes and their geometric relations as edges, a graph is constructed. Using graph convolutions (Landrieu, 2017), a model is trained to recognize structures and classify them to semantic classes respectively.

Figure 13: Multiple views of the same map (top left / light red map in Figure 11). Top: Feature response with red channel representing linearity, green channel represents planarity and blue channel represents verticality. Center: Result of the clustering process, whereby clusters are colored randomly. Bottom: Prediction result for different classes. Red: Building, Light Green: High Vegetation, Dark Green: Natural Terrain, Yellow: Low Vegetation, Gray: Man-Made Terrain, Purple: Artefacts

5. GEODATABASE MANAGEMENT AND GEODATA DEPLOYMENT

5.1. Geodatabase: concept and workflow

In every moment during emergencies, the situation can rapidly change and therefore a large amount of geo-datasets containing great number of files and object data should be respected, stored and analyzed with such a multi sensor – multi platform system as implemented in the research project. In such environment, a key challenge is to provide the needed spatial and also non-spatial data in an easy way based on accurate and simple solutions for an identification and retrieval of, often distributed and heterogeneous, geodata. From this aspect, thinking in a creation of geodatabases is essential, but the design of a geodatabase is a critical stage in ensuring its effectiveness and proficiency. If a geodatabase concept is designed incorrectly, the data provision will not function to its full potential, essentially creating problems for the users. Therefore we concentrate on the design of geodatabases with respect to pertaining technical functionalities, as well as with a focusing
on Postgres\(^1\) and PostGIS\(^2\). The potential is that Postgres enables on the one hand an easy access of the data available and on the other hand there is an efficient compatibility with other software that deals with data processing, retrieving and viewing like QGIS, Geoserver, etc. Moreover, Postgres comes with a quite liberal licence policy. PostGIS is a spatial database extender for Postgres object-relational database. It adds support for geographic objects allowing location queries to be run in SQL.

Within this project, the main datasets include images and point clouds. From a logical point of view, and to be smart in the data provision, a classification of the datasets based on data type has been achieved. For example, RGB images, thermal photos, orthophotos, DSM, UAV’s/UGV’s Lidar data, etc. Each type should be stored in a specific folder labelled with the corresponding ID of the actual mission (Figure 14). One can store a dataset within Postgres/PostGIS in a spatial database allowing a smart and easy data access and retrieval, but storing images directly as raster layer in a PostGIS-database is complex and still not easy task because the Postgres/PostGIS concept does not support this issue yet. To overcome this challenge, a Python based mechanism has been developed to store the geographic location of the image center points and other image metadata into Postgres database, and linking them to the original images. The mentioned mechanism can be structured in three stages as follow (Figure 14):

- CSV-files related to images will be created in the first stage. These files include metadata of images like: image name, the geographic coordinates of the image center points, etc.
- The second stage is CSV-files conversion into shape files that to be implemented as layers in QGIS.
- Finally, new database tables (sql-tables) will be created based on the shape files (converted in the second step) and loaded into the target database where the sql-table should be stored. These sql-tables have a specific spatial index enables an improved performance in terms of data archiving and retrieving.

Based on a potential linkage between Postgres and QGIS (Figure 15, left), the images of interest can be retrieved and viewed. In QGIS, one can connect a specific database and select the table of interest (namely: sql-table in Postgres database) and import the corresponding shape file as layer showing the main image points (Figure 15, middle). Each point has its own image linkage based on the path to dataset, and therefore an image view in QGIS is possible using the “Pop-up windows” through a click event on the point of interest (Figure 15, right). At the same time, orthophotos and DSMs can be directly imported and viewed as georeferenced raster layers in QGIS independent from pop-ups functionality. The advantage is that the users and operators can edit and process the orthophotos and DSMs in an easy way.

Concerning on the point clouds provided from ANKommEn capturing system, on the one hand from sensors mounted on UAV/UGV system, and on the other hand based on photogrammetric processing using RGB images, we classify the point clouds available based on sensor type used as well as labelled with flight mission ID. For e.g. UAV-Point Cloud, UGV-Point Cloud and RGB-Point Cloud. In the same manner for image data storage into Postgres databases, point clouds available can be also imported into a target database using a python based script (Figure 16).

---

\(^1\) https://www.postgresql.org
\(^2\) https://postgis.net/
5.2. Data deployment

Once having a complete geodatabase, it should address needs for data provision to end users in an easy way like web application that enables users/operators getting their work done without having to learn a lot of skills or technical issues. Therefore, in this section we introduce an initial framework for geodata deployment and geo-web based application.

Beside databases, geo-web applications need other components like: a geoserver for data hosting, a digital basemap to provide the geographical contexts, operational layers for editing layers, queries, etc. The Figure (17) illustrates the principle of distributing geospatial information stored in ANKommEn databases.

Here, the local geodata content (images, point clouds, etc.) is implemented into a remote web map service for end users. The images and the derived data such as orthophotos, DSMs, etc. are published based on Geoserver as georeferenced raster layers. These layers could be retrieved using their own Uniform Resource Locator (URLs) that to be used for the creation of friendly interactive maps. For this purpose, Leaflet is the leading open-source JavaScript library that enables creating interactive maps using Hypertext Markup Language Mark-up Language (HTML) standards and Cascading Style Sheets (CSS). The Figure (18) shows a leaflet based web map including an orthophoto projected into basemap “Open Street Map”.

In contrast, the point clouds are converted directly into the html file format using Potree Converter\(^1\) and can be retrieved as layers using their URLs (Figure 19).

---

\(^1\) Potree is a free open-source WebGL (Web Graphic Library) based point cloud renderer for large point clouds, developed at the Institute of Computer Graphics and Algorithms, TU Wien.
6. CONCLUSION

Within this paper, the current work accomplished in the joint research project ANKommEn and its successor ANKommEn 2 has been introduced. The project idea embodies an exploratory investigation to provide correct and timely geodata that can help in emergency cases; especially in support decision making in emergency risk management. Starting by description of the system design and its components, then an intelligent communication structure implemented in the system has been addressed. Furthermore, two mapping methods of the desired area are used, SLAM and photogrammetry based methods. Finally, a geodatabase design has been presented so that all of datasets available such as images, point clouds, etc. can be connected and retrieved. A future task is to suggest additional ideas i.e. future design considerations that can help improving system communication, as well as to enhance the geodatabase model designed; especially in the terms of advanced relationships between geospatial data and database tables.

ACKNOWLEDGEMENTS

The presented work was done within the joint research projects ANKommEn and ANKommEn 2 funded by the German Federal Ministry of Economic Affairs and Energy administrated by the Space Administration of the DLR (ANKommEn - funding code: 50NA1518, ANKommEn 2 - funding code 50NA1807).

REFERENCES


