Image exploitation algorithms for reconnaissance and surveillance with UAV

N. Heinze, M. Esswein, W. Krüger, G. Saur
Fraunhofer IOSB, Fraunhofer Str 1, 76131 Karlsruhe, Germany

ABSTRACT

UAV have a growing importance for reconnaissance and surveillance. Due to improved technical capability also small UAVs have an endurance of about 6 hours, but less sophisticated sensors due to strong weight limitations. This puts a high strain and workload on the small teams usually deployed with such systems. To lessen the strain for photo interpreters and to improve the capability of such systems we have developed and integrated automatic image exploitation algorithms. An import aspect is the detection of moving objects to give the photo interpreter (PI) hints were such objects are. Mosaiking of imagery helps to gain better oversight over the scene. By computing stereo-mosaics from mono-ocular video-data also 3-d-models can be derived from tactical UAV-data in a further processing step. A special instrument of gaining oversight is to use multi-temporal and multifocal images of video-sensors with different resolution of the platform and to fusion them into one image. This results in a good situation awareness of the scene with a light-weight sensor-platform and a standard video link.

Keywords: Video exploitation, LUNA, UAV, Mosaiking, MiSAR, video stabilization, super-resolution, multi-resolution, MTI

INTRODUCTION

Increasing airframe capabilities, improved data link capabilities and new sensors like improved IR- and small SAR-sensors make the application of small UAVs more and more interesting. Low weights give easy handling characteristics and operational costs. Flight-time extends to about 6 hours which is positive in terms of capabilities but on the other side puts a high strain on the personal doing the image exploitation. UAVs like the German LUNA – see Figure 1- are stronger influenced by wind turbulences as large systems operating at high altitudes because of their lower flight profile and small mass. Sensors and stability of the sensor platform is inferior to those of large and cost intensive platforms. These drawbacks demand for compensation by image exploitation algorithms. These algorithms like image stabilization, mosaicking, multi-resolution, super-resolution and moving object detection are not limited to small UAV application, but deliver their maximum benefit in this application.

Figure 1: Small UAVs like the German LUNA from EMT, Penzberg, Germany have long endurance (i.e. 6 hours) and define in combination with sophisticated video exploitation algorithms a very cost efficient platform for surveillance.

We will address following algorithms:

1. **Multi-resolution**: Images of different focal length are processed to generate data-fusion-results.
2. **Super-resolution with emphasis on SAR sequences**: Image-sequences are processed to generate high-resolution products with less noise.
3. **Image stabilization**: Processing the video stream, the displayed output is stabilized in real-time, enabling the operator to achieve optimal results.
4. **Real-time detection of moving objects with moving sensor**: On moving platforms moving targets are automatically detected, though the sensor platform is in motion itself. The underlying algorithm performs with a very low false alarm rate (FAR) despite its high sensitivity.
5. **Geo-coded mosaicking**: Processing the video stream, the video images are stitched to panorama images in real-time, which provide excellent overview.
6. **Stereo Images/3-d-Models**: The video stream is automatically processed in order to provide three-dimensional views.
7. **Experimental video-exploitation system: ABUL**: System for testing algorithms especially under realistic conditions, to gain insight in the demands.

**MULTIRESOLUTION:**

A main problem of video imagery is that it gives very view oversight, particularly if still-image results are generated from the video source for the purpose of documentation. To overcome this situation, often several still images with different focal length are generated. In combination these images provide all information but in a quite bulky way. Another approach, which we describe later is to use mosaiking. This results in an image with a width similar to the video-frame and a large extension in height, when generated on the fly along the flight-path. 2-dimensional-mosaiking has to be planned during flight. Our multi-resolution-approach can be performed on data of the whole flight, so that also multi-resolution images can be generated for interesting targets directly after passing the object.

The sensor is a lightweight system with 4 fixed-focus optical sensors, optionally equipped with a further IR or zoom-sensor – see Figure 3. This is the standard sensor-configuration for LUNA. Images are transferred via radio link to a ground station. Due to bandwidth limitations, images from most sensors can only be transmitted with low frame-rate. This results in a collection of images of

- different time and
- position
The objective is to reference the images so that a multi-resolution image (cf. Figure 5) can be generated. This is done for 3 to 4 resolution levels with about 10 images, so that particularly in high resolution levels more than one image can be processed.

Processing steps:
1. A color histogram adaptation is done, to reduce difference in the image histogram.
2. A feature based image registration with respect to a predefined anchor-image is computed.
3. Computation of an output image using low resolution images as background and high resolution images as foreground.

Generation of multi-resolution images takes approximately 5 seconds on standard hardware, mainly depending on the warping of the output image. Our approach has the advantage that also with a light UAV without a heavy sensor-system-
equipment large still images can be generated for the whole flight path. These images can be annotated and handed over to the user.

**SUPER-RESOLUTION:**

Resolution is always a crucial part of video-systems. In video-systems many low resolution frames are available. By combining the information of a video-sequence we can

1. Reduce noise
2. Improve resolution

Depending on sensor and application emphasis lays more on the first or the second

Input is a sequence of low-resolution images. Result is a single high resolution image with reduced noise. The principle algorithm is depicted in Figure 6. Our implementation combines our robust image registration from [1] with the super-resolution approach from [2].

![Figure 6: Principle steps in our super-resolution algorithm.](image-url)
Super-resolution on optical imagery is normally of less interest because in many cases it is possible to increase the focal length of the sensor, to make a large step to get higher resolution. This step can not be made so easily with IR-systems, SAR systems and optical systems integrated in a flying system where changing the sensor is not easily possible. Our algorithm has also been successfully tested on laser gating images from Obzerv (Quebec, Canada) – see Figure 7 – where speckle noise results in severe degradation of image quality. In UAV-applications also improvements for video and IR-sensors have been shown. Video-sequences are processed to generate high resolution products.

![Figure 7: Laser-gating-Camera (Obzerv): 50 images of a moving camera processed: reading of name becomes possible, railing and antennas can be identified. Images: © Obzerv (Quebec, Canada)](image)

1.1 Super-resolution for SAR on small UAV

Large benefits can be identified when processing MiSAR images with our super-resolution algorithm. EADS Germany has developed a miniature SAR-system MiSAR – see Figure 8 - for lightweight UAVs like the LUNA-System which weights only 4 kg. MiSAR adds to tactical UAV-systems the all-weather reconnaissance capability, which is missing until now. Unlike other SAR sensors, that produce large strip maps at update rates of several seconds, MiSAR generates sequences of SAR images with approximately 1 Hz frame rate [3] [4]. This has the advantage, that also agile flight-paths can be used.
We have developed image processing- and exploitation-algorithms for such SAR-image sequences. A main component is the generation of image sequence mosaics to get more oversight. Another algorithm is super-resolution for SAR which is different from multi-look processing because it works completely in the image-domain. Typically 9 images are processed to produce results as in shown in Figure 9. The combination of an innovative SAR-sensor in combination with super-resolution makes high quality SAR-imagery available also on a small platform.

**IMAGE STABILIZATION:**

By processing the video stream the video output is stabilized in real-time. So information in the image which can only be detected with difficulties - due to strong camera movements - is presented in a stable manner. This helps the PI (photo interpreter) to detect also small objects and movements.

By computing the transformation between the images of the sequence and applying a “low pass filter” on the computed motion-parameters, we can produce different types of motion stabilization. The “low-path filter” is a Kalman filter applied to the motion. Presenting the different parameters of the motion filter directly to the PI’s would result in maximal flexibility but requires a complicated parameterization from the user. So we decided to define a number of predefined parameter settings which can easily be used and named properly. This parameter settings include
- total stabilization -all movement is compensated -,
- medium movement – sensors moves in typical manner and
- fast movement, which is appropriate for bad weather conditions and long focal length.

Due to the sensor movement not all parts of the scene in the stabilized image are covered by the actual video image. Image parts which are not longer observed would normally be without image information. These areas are reconstructed by mosaicking from earlier images.

**REALTIME DETECTION OF MOVING OBJECTS WITH MOVING SENSOR**

Moving targets are automatically detected, though the sensor platform is in motion itself. The underlying algorithm performs with a very low false alarm rate (FAR) despite its high sensitivity.

Moving objects are of major interest during surveillance missions. But detecting them during a mission-time of several hours is an exhausting task. This task can be supported by using an algorithm to detect moving objects. Due to the movement of the sensor on the UAV, changing illumination and aspect angles, it is a challenging task to develop a robust algorithm with few false alarms. Currently, we use two algorithms. The first one uses motion-compensated difference images [1]. The second one is directly based on the point features used by the image registration module. Moving objects are detected at locations, where local image motion does not coincide with the global image transformation used for registration. The detected object-hypotheses are marked on-line in the image to give the PI a hint for moving objects.

![Detection of moving objects in UAV-scene (moving sensor)](image)

**GEOCODED MOSAICING:**

While processing the video stream, the video frames are stitched to panorama images in real-time, which provide an excellent overview. Due to limitations in cost and weight, the precision of collateral-data from small UAV is not very high. On the other hand, short-term image mosaicking is very precise and generates almost no visual artefacts. But stitching hundreds or thousands of images will result in a drift of parameters. Therefore, additional effort is needed to create large accurate geo-coded mosaics, which can be displayed in common GIS-environments – see [5] [6]. We use mainly geo-coded satellite imagery as reference images. By using a specially developed tool “abulref” some few tie points can be defined.

Using these correspondences to compensate the parameter-drift, we are able to generate mosaics which fit to the reference image like an elastic ribbon – see Figure 12. The resulting large geo-coded mosaics can be used together with other geo-products, thereby integrating video imagery in the normal workflow dealing with satellite imagery and maps, with the advantage of having actual high resolution coverage of important areas. On basis of these results, change detection and multisensory data-fusion tasks can be solved.
STEREO IMAGES/3D-MODELS:

By combining the mosaicking and the stereo techniques, we are able to generate large stereo mosaics, which have similarity to stereo-images generated from push broom sensors. We can make use of larger stereo-bases than in the single image mode, because the mosaics consist of only small image slices from every image pair and the overlap between the images used for the stereo mosaic is much smaller than those used in single frame stereo.

These large high-quality stereo mosaics give the PI the opportunity to inspect visually the 3-d-structure of a scene after less than a minute of processing time. This is especially of interest when exploiting industrial complexes or examples of camouflaged objects.

Stereo mosaics are very useful for an interactive exploitation of scenes, but the generation of the 3-d-information is done by the brain of the human observer. In this case, the 3-d-information is not available for further processing or fusion with other products. This would require the explicit extraction of 3-d-information.

The challenge of explicit 3-d processing lies in the geometrical situation:

- only small disparities of a few pixels are typical,
- the sensor is not calibrated,
- irregular sensor-movements and distortions due to the video transmission.

For the depth-map extraction from stereo mosaics a “dense pixel stereo algorithm” is used, in a simple and fast variant ([7], [8]). The products are 3-d-models generated from stereo imagery – Figure 12.
We have developed a video exploitation system ABUL - Figure 13 - , mainly meeting the demands of small UAVs like the LUNA system. Several image exploitation algorithms described in this paper are implemented and can easily be used by photo interpreters. This gives us the chance to evaluate the system in close collaboration with relevant users. So discussion and feedback with users can take place resulting in new ideas for algorithms and tools [9]. Test systems are supplied to several departments of the Bundeswehr and systems are tested in Kosovo and Afghanistan - Figure 14-, so that the performance of these experimental systems can gradually be improved by incorporating feed-backs and actual scientific results.

Figure 13: ABUL: left: screen one with selection for exploitation algorithms on the right and the six video streams from LUNA-sensors, right: ABUL screen two with mosaic und navigation-panel
SUMMERY

Capabilities of UAV-Systems can be improved by image-exploitation and image-processing algorithms. We described several examples developed in our institute. Examples are multi-resolution-algorithms, super-resolution with emphasis on SAR sequences, Image stabilization, real-time detection of moving objects with a moving sensor, geo-coded mosaics, and stereo-mosaics/3-d-models.

Developing algorithms is only part of the undertaking to improve image exploitation. Necessarily also the integration into the work-flow of interactive image exploitation has to be done. Therefore we developed the experimental video-exploitation system ABUL as a system for testing algorithms, to gain insight in the demands, and the performance under real-world conditions. In this system also capabilities for interoperable data exchange are integrated, so that all STANAG 4609 Video data from other UAV-systems can be processed. This was shown in 2008 Common Shield and MAJEX 2008.

ACKNOWLEDGMENTS

We would like to thank WTD 81 (Greding), BWB-L1.1 (Koblenz), KSA/SGA (Gelsdorf), EADS Germany (Ulm/Friedrichshafen), and EMT (Penzberg) for supporting our work.

REFERENCES


