AUTOMATED UAV-BASED VIDEO EXPLOITATION FOR MAPPING AND SURVEILLANCE

Stephen Se [a, *], Pezhman Firoozfam [a], Norman Goldstein [a], Melanie Dutkiewicz [a], Paul Pace [b]

[a] MDA Systems Ltd., 13800 Commerce Parkway, Richmond, B.C., V6V 2J3, Canada
[b] Defence R&D Canada – Ottawa, 3701 Carling Ave., Ottawa, ON, K1A 0K2, Canada

sse@mdacorporation.com, pfiroozf@mdacorporation.com, norman@mdacorporation.com, mel@mdacorporation.com, paul.pace@drdc-rddc.gc.ca

Commission I, ICWG I/V

KEY WORDS: UAV, video exploitation, mapping, surveillance, monitoring, mosaicking, change detection, 3D reconstruction

ABSTRACT:

Airborne intelligence, surveillance and reconnaissance (ISR) can provide essential information for both military and civilian applications. Such capabilities are critical for situational awareness, mapping and monitoring, and mission planning. UAVs gather large amounts of video data, but it is extremely labour-intensive for operators to analyse the hours and hours of data collected.

At MDA, we have developed a suite of video exploitation tools to process UAV video data, including tools for mosaicking, change detection and 3D reconstruction. The main objective of this work is to improve the robustness of these tools, to integrate these tools within a standard GIS framework, to automate the tools, and to present these tools in user-friendly ways to operational users.

Our mosaicking tool produces a 2D map from the sequence of video frames contained in a video clip. For scene monitoring applications, our change detection tool identifies differences between two videos taken from two passes of the same terrain at different times. Our 3D reconstruction tools create calibrated geo-referenced photo-realistic 3D models, which are useful for situational awareness. The semi-automated approach allows the user to create surfaces interactively by marking correspondences between frames, while the automated approach generates 3D models automatically.

MDA provides UAV services to the Canadian and Australian forces, offering important ISR information to front-line soldiers. Our suite of semi and fully automated video exploitation tools can aid the operators to analyse the huge amount of UAV video data to support battle-space mapping and monitoring applications.

1. INTRODUCTION

Airborne ISR can provide essential information for both military and civilian applications. Such capabilities are critical for situational awareness, mapping and monitoring, and mission planning. Unmanned Aerial Vehicles (UAVs) are becoming the platform of choice for such operations since UAVs can fly in areas that are dangerous or contaminated, and can undertake dull, repetitive reconnaissance jobs.

Video cameras are the most common UAV sensor payload as they are relatively inexpensive and light-weight. Video cameras can collect large amounts of data, however using this data efficiently and advantageously is a challenge. How the imagery is used depends on the mission goals and at which stage of the mission it is used. Data collected from UAVs must be reviewed quickly to support real-time operations in the field, whereas it is analyzed in more depth and over longer timeframes to support mission planning and intelligence gathering.

At MDA, we have developed a suite of UAV video exploitation tools [1]. The main objective of the Advanced Mapping from Imagery (AMI) project, under the auspices of the Defence Industrial Research (DIR) program, is to investigate how to make these tools operational for non-expert users, to integrate these tools within a standard GIS framework, and to present these tools in user-friendly ways to operational users.

ESRI ArcGIS [2] was chosen as the demonstration platform for AMI development. It is a family of products for visualizing, managing, creating, and analyzing 2D and 3D geographic data. ArcGIS is being used within the Directorate of Geospatial Intelligence at both the Canadian military's Mapping and Charting Establishment (MCE) and Canadian Force Joint Imagery Centre (CFJIC). Therefore, developing on the ArcGIS platform minimizes the learning curve for these end users and ensures these tools would fit into their current work process. While the AMI GUI is built within ArcGIS, the underlying tools are not dependent on ArcGIS and could be integrated into any other selected environment.

* Corresponding author

Figure 1 shows the key modules of the AMI processing pipeline. The technology development of these AMI tools is described in this paper.

- Video plug-in – a user-friendly video viewing GUI that also shows UAV paths and video footprints on a map, for up to two videos at a time.
- Region Of Interest (ROI) selection – tool to select a region of interest (area) on the map which is of interest in processing the video imagery.
- Calibration – while this is not a product in itself, camera calibration is a critical pre-processing step to prepare the video data for exploitation: mosaicking, change detection and 3D reconstruction.
- Mosaicking – tool to create quick-look and refined mosaics from a video sequence.
- Change Detection – tool to detect changes between two videos taken from similar perspectives over the same area.
- 3D reconstruction – tool to create 3D models from a video, using both semi and fully automated approaches.

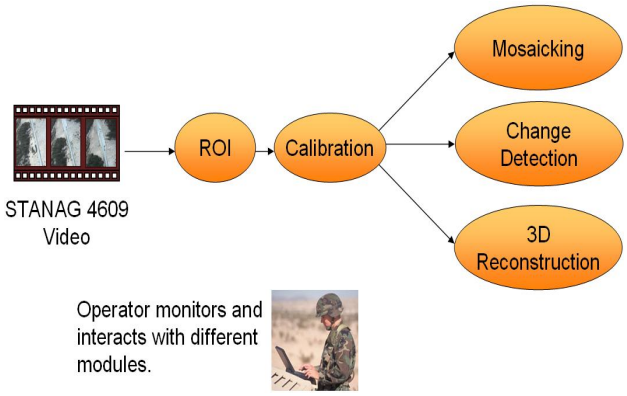

Figure 1 AMI Components

2. ROI SELECTION

Visualization is a key component of AMI, as it allows the operator to view videos, see the sensor path and the extent of the video footprint on a map, select ROIs, and create products. The ROI selection is important as it allows the user to focus the processing on selected regions, which can greatly reduce processing time, especially when not all the imaged area is of interest.

The AMI tools have been implemented as extensions of ESRI ArcGIS. Figure 2 shows the AMI GUI frontend, which plays videos and displays UAV paths and video footprints on the map using the information directly supplied in the video meta-data (GPS/INS). The user can select an ROI on the map to focus the processing on that region.

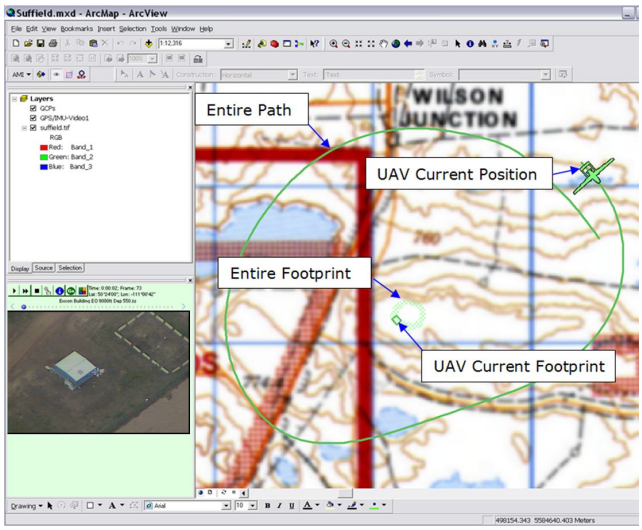

Figure 2 AMI UAV Video Player and Meta-data Display

3. CALIBRATION

Camera calibration includes estimating both intrinsic and extrinsic camera parameters from the video. The intrinsic parameters include the focal length, image centre and lens distortion, while the extrinsic parameters refer to the camera position and orientation for each video frame. We use UAV video in MPEG-2 format that is STANAG 4609 compliant. KLV encoded meta-data is embedded in the MPEG-2 Transport Stream. The GPS/INS meta-data provides rough camera position estimate for each video frame, which will be refined by the calibration process.

For calibration, we track features between frames to obtain tie-points and then perform bundle adjustment. Stable features are highly desirable to produce longer tie-point tracks that provide better constraints for bundle adjustment. Bundle adjustment [3] is done on the 3D structure and the viewing parameters (camera pose as well as the intrinsic calibration parameters). From the 2D tie-points and the GPS/INS video meta-data, the 3D coordinates of these tie-points and the camera parameters can be refined simultaneously.

Although calibration refines the relative motion between images, the geo-registration accuracy is limited by the GPS accuracy in the video meta-data. As video exploitation products have to be geo-referenced precisely to be useful for analysis, AMI provides a Ground Control Points (GCP) marking facility which allows the user to mark correspondences between the video frame and the background map or an ortho-rectified satellite image. The GCP information is optional; if provided these additional tie-points are incorporated during bundle adjustment to improve geo-registration accuracy.

After bundle adjustment, the calibrated video is ready for subsequent video exploitation, including mosaicking, change detection and 3D reconstruction.

4. MOSAICKING

Individual video frames in a video sequence provide only a "soda-straw" view of a scene. Even if the user views the entire video data set by playing the video from start to end, the user needs to rely on memory of the scene content, and to connect this with the knowledge of the UAV trajectory and sensor pointing angle, to form a mental picture of the area that was surveyed by the sensor. Thus, it is very difficult to obtain situational awareness from video data if the only viewing mechanism is a video player. The video mosaicking tool helps to solve this problem by producing a high-resolution 2D map from the sequence of video frames contained in a video clip, effectively showing the information from all the frames simultaneously, and thus providing better geographical situational awareness.

The AMI mosaicking tool offers both quick-look and refined mosaicking options. For the quick-look mosaic, the video frames are stitched together based on the un-calibrated video meta-data only, and therefore mis-alignment is normal, due to the fact that the meta-data only provides rough camera position estimates. For the refined mosaic, the video meta-data is first improved by calibration, after which the video frames are aligned and blended into a seamless mosaic.

Figure 3 shows an example of an AMI refined mosaic generated from an infra-red (IR) video, overlaid on top of a reference map to provide a high resolution up-to-date 2D map to the operator.

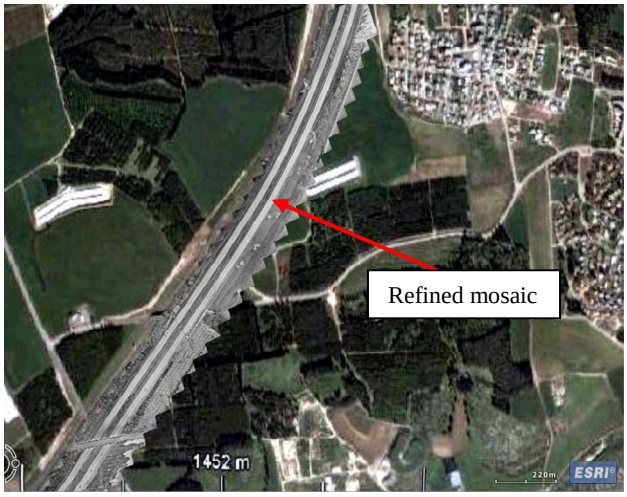

Figure 3 Refined Mosaic Overlaid on Background Map

.

5. CHANGE DETECTION

For scene monitoring applications it is important to be able to identify changes and anomalies. However, it is tedious and difficult for humans to compare hours and hours of video clips looking for differences. The AMI change detection tool automatically finds differences between two videos taken from two passes of the same terrain at different times. Changes between the two video files are identified and flagged on the video frames, so that operator attention can be utilized more effectively by focusing only on the detected changes.

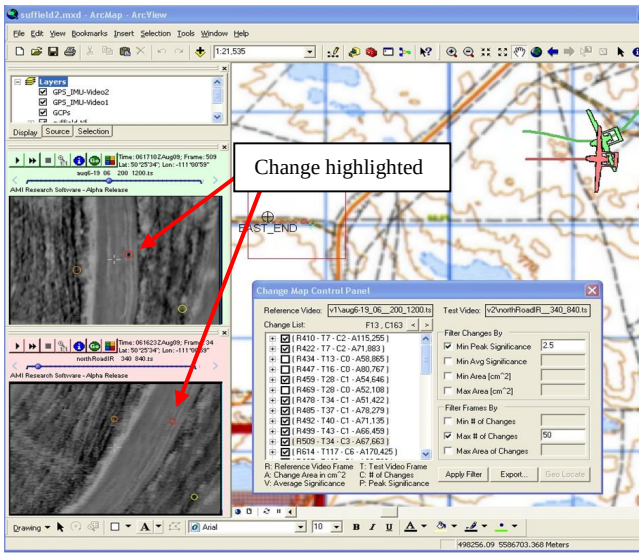

Figure 4 AMI Change Detection Tool

AMI offers a suite of filters to provide the operator with control over the type of change to detect, in order to both avoid overwhelming the operator with too many changes and to help the operator hone in to the attributes of the changes that are significant for that scene. For example, the operator can control the intensity or area thresholds of the changes to be detected. Figure 4 shows the AMI change detection tool, where two IR videos with the detected changes are displayed. The change highlighted by the red circle corresponds to an unknown object on a road: it is observed in the first video but not in the second video.

Once appropriate filters and parameters have been selected, a change detection report in HTML format, as illustrated in Figure 5, can be generated automatically. Apart from showing the video frames with the detected changes circled, the latitude and longitude of the change is also reported, which is useful to help the operator locate such changes in a geographic context.

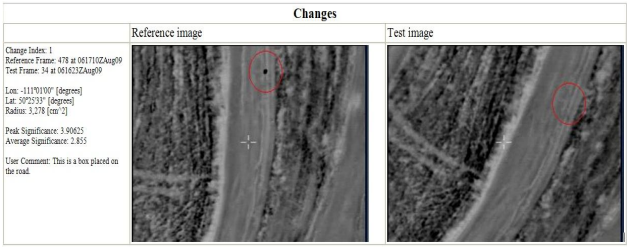

Figure 5 Change Detection Report Example

6. 3D RECONSTRUCTION

For situational awareness, it is easier to understand the scene by visualizing it in 3D. Both semi-automated and automated approaches have been developed in AMI. The AMI semi-automated tool allows the user to create surfaces interactively, by marking correspondences between frames. It requires some operator assistance but generates relatively clean 3D models. This approach is suitable for simple 3D objects that can be modelled with relatively few surfaces.

The AMI 3D automated tool can generate 3D models automatically, provided that the video frames contain sufficient texture. It does not require operator assistance, but the resulting 3D models may be noisier. This approach is suitable for natural terrains or complex 3D objects that need to be modelled with a large number of surfaces.

Both of the AMI 3D reconstruction tools create calibrated geo-referenced photo-realistic 3D models, which can be viewed with 3D GIS viewers such as ESRI ArcGlobe, ESRI ArcGIS Explorer and Google Earth. The models can be viewed from different perspectives, and allow distance measurements and line-of-sight analysis.

6.1 Semi-Automated 3D Reconstruction

A Literature Survey on semi-automated 3D reconstruction showed that the typical approach is domain-driven. For example, if the objective is to reconstruct 3D models of buildings, one approach is to provide a set of primitives, such as a particular type of roof or building [4]. In this case the user selects such a primitive and the system then estimates the parameters for the primitive, such as the height and width of the building, based on the image information.

For AMI, our selected design for the semi-automated approach is for the user to reconstruct polygon surfaces interactively, instead of using 3D primitives. The reconstruction is therefore not restricted to buildings, but is flexible to allow modeling of other man-made 3D objects, such as vehicles.

We have developed the GUI frontend as part of the 3D Analyst extension to ArcGIS, as shown in Figure 6. The upper portion of the GUI is the 2D key frames viewer, while the lower portion is the 3D viewer. The camera positions for each frame can be visualized in 3D, which facilitates the selection of image pairs for pair-wise tie-point marking. The user can scale the camera positions along the optical line of sight so that both the camera positions and the 3D object can be viewed together, while preserving the relative geometry.

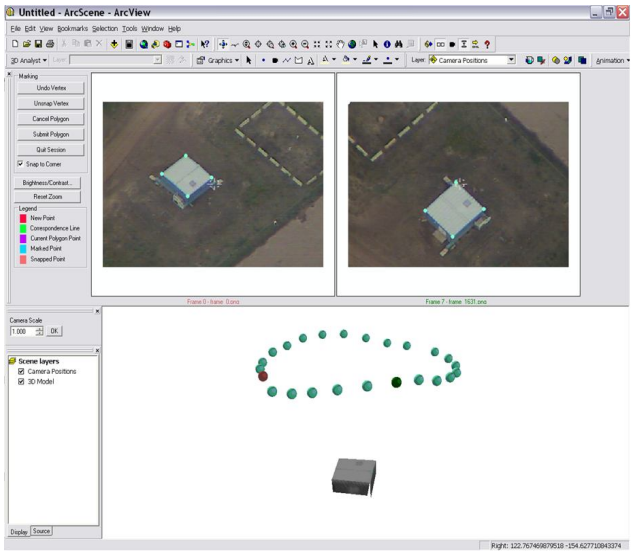

Figure 6 AMI Semi-Automated 3D Reconstruction Tool. The camera positions have been moved along the optical line of sight towards the 3D object.

Following the marking of frame correspondences in a pair-wise fashion, textured 3D surfaces are reconstructed. The partial 3D models can be visualized immediately to provide feedback to the user. The final 3D model can be exported in KML format and the geo-referenced 3D model can be visualized using various 3D GIS viewers.

6.2 Automated 3D Reconstruction

Following the taxonomy described in [5], multi-view 3D reconstruction algorithms can be categorized into four classes:

1. 3D volumetric approaches
2. Surface evolution techniques
3. Algorithms that compute and merge depth maps
4. Techniques that grow regions/surfaces from a set of extracted feature or seed points

Based on the AMI context and requirements, the third category in the list above was determined to be the most suitable and efficient approach for this project.

The AMI automated 3D reconstruction tool processes the input frames in a pair-wise fashion and generates geo-referenced 3D models in KML format. A video clip captured while a UAV flew around a building of interest was processed: Figure 7 shows the resulting photo-realistic 3D model loaded in Google Earth. Figure 8 shows another 3D model result loaded into ESRI ArcGIS Explorer; this demonstrates that the tool also works well with IR data, which is valuable for night operations.

Figure 7 A Photo-realistic 3D Model Generated by Automated 3D Reconstruction Tool Viewed in Google Earth

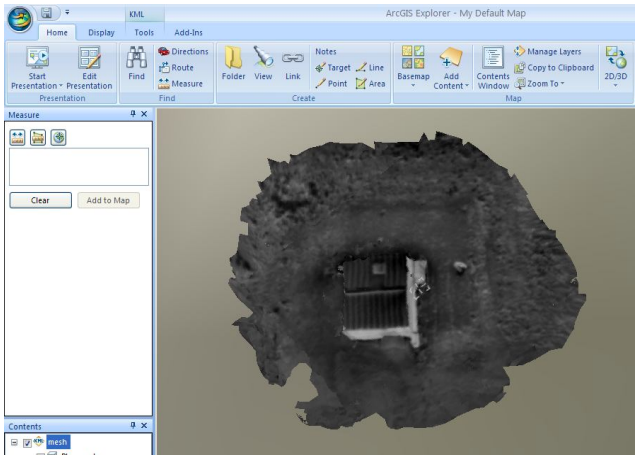

Figure 8 An IR 3D Model Generated by Automated 3D Reconstruction Tool Viewed in ArcGIS Explorer

7. CONCLUSIONS

MDA provides UAV services to the Canadian and Australian forces with the Heron, a Medium Altitude Long Endurance (MALE) UAV system [6]. The on-going flight operations service provides important ISR information to commanders and front-line soldiers. Our suite of semi and fully automated video exploitation tools can aid operators in analysing the large amounts of UAV video data collected to support battle-space mapping and monitoring applications.

The AMI project has developed a suite of UAV video exploitation tools for non-expert users. The video data exploitation functionality has been integrated into ArcGIS to demonstrate a user-friendly tool-set for video data exploitation that complements the GIS functionality provided by ArcGIS.

AMI video exploitation products include camera calibration, mosaic generation, change detection and 3D reconstruction. Test results show that these video exploitation products are generated successfully from Heron UAV video data. The AMI tools are being evaluated by the Canadian Forces Directorate of Geospatial Intelligence including MCE and CFJIC.

ACKNOWLEDGEMENTS

We would like to thank Dr. Paul Pace, DRDC, for technical guidance; D Geo Int for military support and guidance, in particular Mr. JPL (Pierre) Naud; and the DIR program for financial support.

REFERENCES

[1] S. Se, P. Firoozfam, N. Goldstein, L. Wu, M. Dutkiewicz, P. Pace, JPL Naud, Automated UAV-based mapping for airborne reconnaissance and video exploitation, *Proceedings of SPIE Vol. 7307*, Orlando, Florida, 2009.

[2] ESRI ArcGIS, http://www.esri.com

[3] B. Triggs, P. McLauchlan, R. Harley, A. Fitzgibbon, Bundle adjustment – a modern synthesis, *Proceedings of International Workshop on Vision Algorithms*, Corfu, Greece, September 1999.

[4] E. Gulch, H. Muller, T. Labe. Integration of automatic processes into semi-automatic building extraction, *International Archives of Photogrammetry and Remote Sensing (IAPRS)*, Vol. 32, Munich, Germany, 1999.

[5] S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2006.

[6] MDA Unmanned Systems, http://www.uavs.ca