PRODUCTION SUITE FOR AIRBORNE DATA

Juerg Luethy, Hilmar Ingensand
Institute of Geodesy and Photogrammetry,
Swiss Federal Institute of Technology
Email: (luethy;ingensand)@geod.baug.ethz.ch
Roland Stengele
Swissphoto AG, Regensdorf
Email: roland.stengele@swissphoto.ch

Abstract: In this paper we describe the Production Suite for Airborne Data. This toolbox is designed to increase the efficiency and the quality of airborne data processing. The core part consists of a database and the modules for project and quality management. Proven external applications like for raw data processing or terrain filtering can be integrated and they can make use of the data hub which helps to make the data more consistent and eliminates some typical error sources.
The system is designed to handle from small up to large mapping projects. The flexibility is achieved through different parameter sets and definable work flows. The quality control can be defined according to the client's requirements and offers also reporting capabilities.

1. Introduction

Within the last years airborne laser scanning (ALS) has become a standard technique for 3D data acquisition. Currently various companies develop airborne laser scanners and the number of systems in use grow continually. Besides the hardware and the core tools to control and monitor the unit there is a wide number of software application needed to process the final products. Many of the software packages for single process steps have been improved through the last years but they are typically suited for a small part in the production process and barely linked together in fact their data structure is typically proprietary. We can group theses applications in sensor-related application and processing application. The first ones are usually provided together with the sensor system and used for flight planning, data acquisition and processing the measurements to raw "products" (laser point cloud, images with georeference). The second ones are used to derive the products requested by the client and are independent from the sensor system.

During the processing of a large Laserscanning/Orthophoto project we realized that the requirements of from project and quality management are not covered well enough in the above mentioned tools. Consequently we decided to develop a production database which contains all status information about *who did when what.* This allows also to generate status/progress reports for the management. On the quality side different methods to inspect and verify elevation models were developed together with the client [1]. Finally the two parts were linked by a process manager which automatically processes data for quality inspection based on the status in the database and updates it after completion. In this paper we start with a look at the different data sets which accumulate during a project and describe how we organize the data to keep the overview. Then we present in more details some core tools of the Production Suite and the their contribution to an efficient and effective data processing.

2. Data

2.1. Data Sets

For the design of our Production Suite it is necessary to be aware of the variety of different data sets which are collected, produced or output during one project. We concentrate on the data sets which are either origin or final data or data sets which are needed over different process steps. Anyway the numbers of data sets can become quite large and therefore we group them into six different classes:

2.1.1. Primary Data (Sensor Data)

To this group fall all data sets which are collected from the airborne system during the project like GPS (reference and rover), INS, ALS measurements and Imagery. These data sets are basis for all further processing steps but they are available typically in a sensor specific binary format.

2.1.2. Secondary Data

To this category belongs all information which is not specific to one project but might be restricted in the duration of the validity. Some of them are directly used to derive the laser raw points (respectively to georeference the imagery) while other will be used later in the production or in the quality control. Examples are installation parameters like lever arm, calibration values, but also Ground Control Points Pixelmaps, break lines or vector base data

The production data can always be both intermediate and final. Only derivate products as vectors will always be stored in the DB. The DEM and DSM are stored once they are finalized and they are not the end product but will be used for other products.

2.1.3. Production Data

It contains all data sets which are relevant for further processing or are final product, like

- Georeferenced, projected and split to tiles laser data
- Digital Elevation/Surface Model (DEM/DSM), interpolated from classified points
- Digital Orthophoto (DOP)
- Further products like; contour lines, 3D building Models, forest boundaries, tree stands, power line, break lines, land usage

The production data can always be both intermediate and final.

2.1.4. Project Configuration Data

In this group fall all the requirements and for every product its specifications. These data sets are non-spatial but might be used as parameters during data processing and analysis such as point density (DEM, DSM), ground sampling distance, horizontal and vertical accuracy, projection, datum, geoid or verbal description of requirements. Based on these requirements and user can define project specific user roles, workflows (as the data processing chain) and the steps in the quality assurance/quality control (QC). This way the QC becomes adjustable depending on product and requirements but the flexibility is limited to standard products with standard QC functions.

2.1.5. Processing data

To this category belongs all spatial information for processing the data

- Project Perimeter
- Tiling schemes (hierarchical organized in region, block, sub-block and tile); associated with parameter for DEM filtering, production state of each tile, Operator, Results from QC and the processing history
- Realized flight lines/image center coordinates

The tiling scheme might be different for different products within one project. The history contains all the process steps which have been applied automatically or manually to a certain tile of each product. In the workflow the project manager can define the dependencies between the production steps (see section 3.2).

2.1.6. Meta Data

From an end user perspective Meta Data of a geo dataset shall describe the characteristics in a way that it enables the user to decide if that geo dataset fulfill his needs. The catalogue of Meta Data is typically defined by the principal. Commonly used definitions are published from ISO, 91115 [2] and also from national mapping agencies like the US Federal Geographic Data Committee [3]. These two are both available in the Production Suite but the user can also define its own catalogue.

2.2. Data organization

For the data organization it must be distinguished between data sets which are stored in a database (DB) and others which are due to performance restrictions kept on a file base. Currently the DB is used to store all parameters sets and configurations, the processing data, Meta Data and the final products with the exception of Lidar points. Since the number of data sets residing outside the DB is large, it is important to follow a strict data organization.

The Primary Data is organized according to flight missions. This allows being independent from the projects and therefore acquiring data for different projects in the same flight.

The laser strips are split in tiles for further processing. The tile size is often given by the client but it must allow also an efficient processing. For organizational and performance reasons we decided to work in a hierarchical file structure with four levels, depending on the project size.

3. Production Suite

3.1. Requirements

From the practical experience with the first ALS project we compiled the requirements for a production suite. The following listing gives an overview of the high level requirements.

- Use existing tools which are already available, proved and documented
- Consistent data storage; no redundant data
- Support for statistical and visual quality control
- Traceability; *when did who what and why*
- Reporting capabilities
- Archiving of completed projects
- Flexible data organization

- Support of different data formats
- Supporting standard projections/transformation
- Automated processing

Some of the requirements can be easily fulfilled by using a Geo-Information System as a base. We decided building our application on ArcGIS from ESRI. We also agreed to rely on TerraScan [4] as the only DEM filtering tool because it is widely used and supports different requirements. On the database side we use ArcSDE and Oracle. Due to the large data volume and processing time the system architecture is based on thick clients, a database server and a file server (for Primary Data, temporary raster data sets and the laser points).

In the following sections we focus on the tools we think are significant in 3D Mapping projects but not available in other comparable software tools like GeoLAS [5] or GeoCue [6].

3.2. Module Workflow Guidance

When setting up a new project several different parameter sets can be defined. The system offers default values which are based on the experience from different projects. Yet it is worth to have a closer look at the possibility of the workflow definition offered by the tool. Especially in larger projects with often many different deliverables it is important that everyone follows the same processing steps, conducts the same quality controls and that it can be traced when these processes are carried out.

In Workflow Guidance the user can define for each project which process steps for each product must run. A set of standard process steps is offered by the tool, like filtering of laser points or QC. For other processes there are different possibilities; one time jobs do not need to be automated and are therefore defined as manual process. Some automation can be achieved by simply invoking an external script. The Model Builder from ArcGIS offers a simple way to define a geoprocessing script. For larger projects or repeated processes a new generic process step should be added programmatically.

Loops in the workflow are achieved by defining the behavior in the case a failure, this definition is required for each QC for all other steps it is optional. Branching of the process is also supported by defining the precondition for each process step.

Several checks are performed before updating the workflow definition, like a product can not be selected once Final QC passed or which product can be processed in which step.

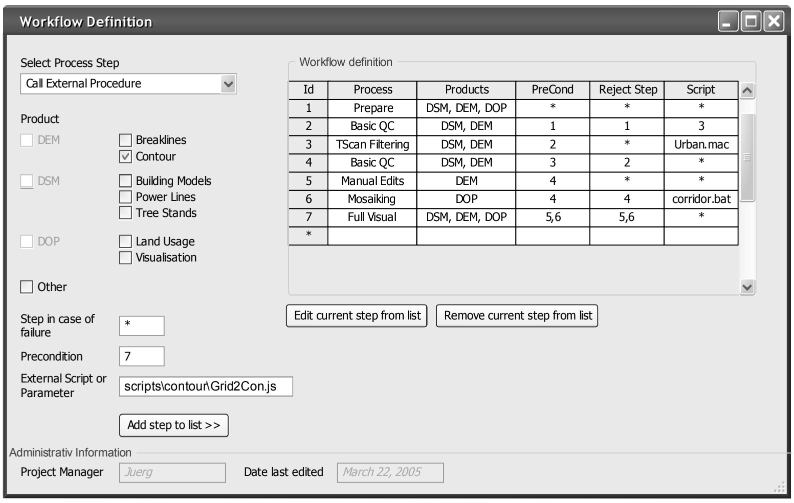

Figure 1: Workflow Definition Tool

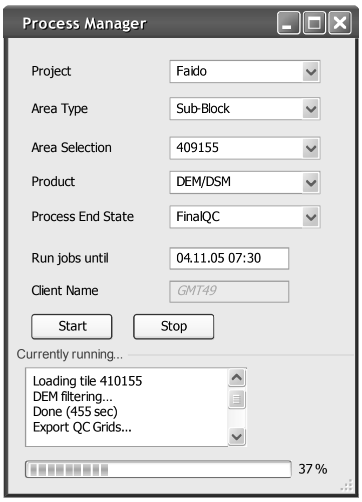

Figure 2: Process

3.3. Module Process Manager

The process manager is a small client tool which is used to run automatically processes on a client computer. Once it is started it checks on the tiling list if any tile from the selected area is ready for a tool process. This happens typically after a manual interaction like editing or Q-inspection. Based on the workflow in the Processing Data table and the current status of each tile we know which processing tool must be invoked next. At the end the process updates the tile status and the tool looks for the next tile. With each completed process the tile information is updated with important information from the process and a new status is set.

The process manager controls processes overnight or over the weekends and therefore makes the processing much more efficient. The configuration is robust versus failure of single processing nodes and the overall reliability of the system is therefore given.

3.4 Module Quality Control (QC)

Quality management plays a central role in all kind of projects but especially in the ones where airborne data acquisition is involved [7]. The costs for airborne acquisition are significant part of a project budget. Therefore quality must be taken into account from the project start. We distinguish in the workflow between a Basic QC and the Full Visual inspection, where Full Visual Inspection requires approval from two analysts for passing. The Quality Control module is open for configuration. The user can select from a list of potential checks the ones which are appropriate for the current project and defines also the parameters where possible (see Figure 3).

The tool generates for each quality inspection a collection of data sets which are loaded into the GIS. The collection consists typically of the hill-shaded DEM and DSM, several other grids which are derived from the two products, contours, no data areas, comparisons of the DEM with ground control points and already existing base data, as vector (like building footprints) or raster data sets (like pixelmap, orthophoto). For each project and QC type a map template can be defined where Layer order, legends and additional, project specific datasets can be pre-loaded. Where available, statistical information form point density or GPS are derived during the data preparation and the tool writes the statistics to the database.

The analyst is partially guided through the visual inspection to ensure that he search for known issues. The user has also the possibility to query the database for statistical information. When a problem is encountered the user can load additional datasets like the trajectory to trace the problem back to the roots. The linking of all data sets in the context of the visual controls saves a lot of time.

Findings are graphically marked from the analyst and are made available in TerraScan to correct the problem manually. In the next QC iteration the analyst has to check only the earlier marked areas and not again the entire tile. In the independent QC (dual control) the error shape gives an indication where the operator was working and if a problem might not be improvable like low DEM density in a forest.

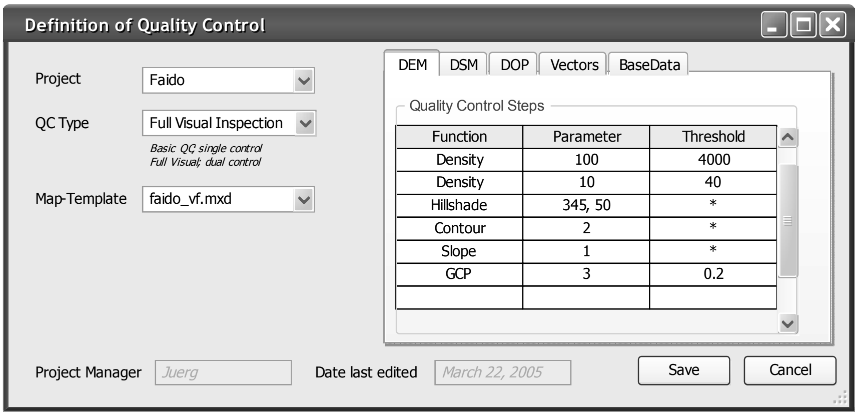

Figure 3: Definition of the quality control steps

4. Benefits

From the experience in several different projects we realized that the requirements for such a tool are often contradictory; the project manager needs the flexibility to support different projects with different specifications but at the same time he needs to ensure, that every operator follows strictly the workflow to guarantee the quality level. The architecture and design of the Production Suite for Airborne Data covers therefore most of these requirements. Another important benefit from the tool is the tracking of the process steps where the history helps detecting problems in the processes and thus supports continuous improvement.

The client can take advantage of this tool in two ways; on the one side he can be sure that the quality is achieved by stable processes and an excellent and proved quality control. Additionally some reporting tools are made available over a web interface allowing the client to get real-time information about the project progress.

5. References

[1] Artuso, 2003. Pratical Methods for the Verification of countrywide Terrain and Surface Models, In 3-D reconstruction from airborne laserscanner and InSAR data, Dresden, Maas, Vosselman and Streilein (ed.); ISPRS Archives XXXIV, Part 3/W13

[2] ISO International Organization for Standardization, 2003, ISO 19115:2003: Geographic information - Metadata. Committe: ISO, Geneva,

[3] FGDC, 1998, Content standards for digital geospatial metadata Version 2 -June 1998. Washington

[4] TerraScan User Manual. http://www.terrasolid.fi (last accessed March 18 2005)

[5] Hug, 2004, Advanced Lidar Data Processing with LAStools, In International Society for Photogrammetry and Remote Sensing XXth Congress, Istanbul, Altan (ed.); ISPRS Archives XXXV, Part B2

[6] GeoCue User Manual. http://www.niirs10.com (last accessed April 19, 2005)

[7] Luethy, 2004. How to evaluate the Quality of Airborne Laser-Scanning Data, In Laser-Scanners for Forest and Landscape Assessment, Freiburg i.Br., Thies, Koch, Spiecker, Weinacker (ed); ISPRS Archives XXXVI, PART 8/W2