

## CREATING AND USING VERY HIGH DENSITY POINT CLOUDS DERIVED FROM ADS IMAGERY

**Stephan Gehrke**

**Robert Uebbing**

**Michael Downey**

North West Geomatics

Suite 212, 5438 – 11<sup>th</sup> Street NE

Calgary, Alberta, T2E 7E9, Canada

[stephan.gehrke@nwgeo.com](mailto:stephan.gehrke@nwgeo.com)

[robert.uebbing@nwgeo.com](mailto:robert.uebbing@nwgeo.com)

[michael.downey@nwgeo.com](mailto:michael.downey@nwgeo.com)

**Kris Morin**

NovAtel Inc.

1120 - 68th Avenue NE,

Calgary, Alberta, T2E 8S5, Canada

[kristian.morin@novatel.com](mailto:kristian.morin@novatel.com)

### ABSTRACT

North West Geomatics has developed a DSM generation tool for Leica ADS line-scanner data in close cooperation with Leica Geosystems. This tool has been deployed internally in production to gauge its usability and has just been released as the main new feature of the Leica XPro 5.0 software, the XPro DSM Extractor. The quality of extracted point clouds has fulfilled and, in many cases, exceeded the initial expectations. However, the SGM algorithm also delivers a very dense point cloud, the size of which surpasses any LiDAR collected data by far.

This paper outlines the challenges of managing and processing elevation data with very high resolutions. Although the main focus is on developing new strategies on how to cope with the vast amount of data, the options and limitations of currently available commercial software packages for such data sets are discussed as well. Lastly, an outlook on the advantages of using high resolution elevation data in the context of ortho-image production is given.

**KEYWORDS:** Pushbroom, Point Cloud, High Resolution, Simplification, Production

### INTRODUCTION

North West Geomatics, in conjunction with Leica Geosystems, has developed a new DSM generation engine, which is based on Semi-Global Matching (SGM). We adapted the algorithm to the unique properties of ADS line scanner imagery, which is described in detail by Gehrke et al. (2010). Since recently, important improvements to this software have been implemented, focusing on the robustness, performance increase and intelligent thinning of the resulting point clouds. It is used in North West Geomatics' production since the 2010 flying season and has recently been released by Leica Geosystems as part of the XPro 5.0 software package.

The underlying SGM approach is suited for the generation of very high resolution DSM/DEM in the order of the image ground sampling distance (GSD); it fulfills the need of high DSM resolution along with good performance. However, considering that a typical take of ADS imagery consists of up to 10 billion pixels, the dense, pixel-based SGM will easily generate point clouds containing a few billion points. Although this data set is physically split into multiple LAS output files, viewing and, in particular, editing such data volumes poses a major challenge to commercially available software packages. To address this issue, a two-step intelligent thinning algorithm has been developed, which reduces the data to a more manageable amount while preserving the most significant points.

The main deliverable products derived from the SGM-based DSM extraction are ortho-images and color-coded point clouds, referred in the following as 'image clouds'. As the image clouds represent the ground space but not necessarily the true ground elevation, the main applications are in forestry, agriculture, remote sensing for medium to small scale mapping. The future use of the DSMs for urban mapping, city modeling and large scale mapping is

still under investigation, but it can already be stated that due to the high point density there is a great potential for these applications, despite challenges in editing and modeling.

## SEMI-GLOBAL MATCHING OF ADS DATA

SGM is a new image matching approach, which originates from the computer vision community. It has been developed by Hirschmüller (2005, 2008) and adapted to the ADS line-scanner geometry by Gehrke et al. (2010). Similar to global matching, the core algorithm aggregates the matching costs under consideration of smoothness constraints. The minimum aggregated cost leads to the disparity map for a stereo pair. Based on various validations as described in next section, corresponding ground coordinates can be derived for the majority of image pixels.

### Matching Cost: Difference of Intensity Gradients

The difference of intensities of allegedly corresponding pixels (in global matching including SGM) or search windows (in local methods) is a simple but widely used cost function: The smaller the difference, the more likely is the correct match – provided that overall intensity offsets in the images of a stereo pair are removed. One solution is the Sobel filtering of the imagery as recently proposed and tested by Hirschmüller & Gehring (2009) with different matching approaches including SGM. The Sobel filter is applied solely in the direction of the epipolar lines, which corresponds to the flight direction or  $x$  in the ADS image coordinate system.

This convolution has two effects: First, radiometric offsets are removed. Second, the  $S_x$  eliminates  $y$  gradients or, respectively, horizontal structures that might have a negative impact in matching in case of (unavoidable but small) errors in the image orientation. However, our main driver for switching the cost from Mutual Information (MI) that we initially used for ADS processing (cp. Gehrke et al., 2010) has been occasional errors; in brief: MI cost reflects the probability of a certain pair of intensities to be a valid match, and for a particular project in North West Geomatics' production the probability computation was dominated by grassland and, applied to the neighboring road (with similar intensities but slightly different pair matches due to a different BRDF), elongated outliers occurred despite the SGM smoothness constraints. Already used for large projects of various type – including all examples presented in this paper –, the Sobel-based cost has not yet shown any issues on flat structures and open areas in general. This is crucial when aiming for the derivation of a DEM.

### SGM Cost Aggregation in Brief

Any pixel-wise matching cost measure is generally ambiguous. Therefore, constraints such as the assumption of a smooth surface need to be introduced and globally aggregated along with the cost. The Sobel-based cost  $c$  is computed for all potential matches in a stereo pair, i.e., for each base image pixel  $p$  and for all disparities  $d$  that provide the link to pair image locations. This results in cost vectors for each pixel and, in a cost cuboid for the entire image. The energy function  $E(D)$  that has to be globally minimized consists of this cost and constraints on smoothness. Hirschmüller (2005, 2008) proposes to use the penalties  $P_1$  for slight variations of one disparity and  $P_2$  for any larger changes, i.e. discontinuities, in-between neighboring pixels:

$$E(D) = \sum_p e(p, d) = \sum_p \left\{ \begin{array}{l} c(p, d) + \sum_{q \in N_p} P_1 T(|d - d_q| = 1) \\ + \sum_{q \in N_p} P_2 T(|d - d_q| > 1) \end{array} \right\}$$

The  $T(\dots)$  function controls the application of penalties, i.e. the weighting between matching and smoothness constraints; it is 1 in case that its argument is true and 0 if false. SGM approximates the theoretically desired two-dimensional, global aggregation of matching cost by a number of one-dimensional cost paths for each potential disparity. A total of eight paths at every  $45^\circ$  is considered sufficient and also used for ADS processing (Gehrke et al., 2010). The overall cost vector for each base image pixel results from the summation of these eight aggregated cost vectors at that pixel. Based on that, the final disparity map is provided by the disparities corresponding to the minimum of each pixel's summed cost vector. It is generated for the base image and also for the pair image by switching the role of base and pair images. This reverse matching allows for the elimination of mismatches, which occur mainly in occluded areas.

## Considerations for ADS Data Processing

Since raw ADS data (L0) are impacted by atmospheric turbulence and will show corresponding distortions, we rectify the imagery to a plane of mean terrain height. This removes the vast majority of such distortions, resulting in a stereo-viewable product (L1). This is carried out on-the-fly by the XPro DSM Extractor.

The ADS provides three panchromatic bands in different viewing angles, in case of the ADS80: 14° backward, 2° forward and 27° forward. For SGM processing, we generally use two stereo pairs, nadir-backward and nadir-forward, with the nadir image being the common base. This allows for filling gaps from occlusions and enables consistency checks between the stereo pair results. See Gehrke et al. (2010) for further details.

SGM processing is generally carried out in image pyramid levels, mainly to reduce the theoretical range of potential disparities to the actual area of interest. In fact, the results of a certain minification level are only used to constrain the range for the subsequent processing steps. Considering the ADS imaging configuration and data properties, the power of a typical modern workstation computer (or cluster node) as well as the behavior of SGM in combination with the Sobel-based matching cost, processing is started with 16:1 minified data, which results in an initial disparity range of ~ 300 for the larger stereo angle.

The very long ADS images are divided into sub-chunks of 8,192 pixels in length and full image width (~ 12,000 L1 pixels). The resulting ~100,000,000 pixels form the input for a DSM extraction job, which is internally processed in SGM tiles of 1,024 x 1,024 pixels (plus tile overlap) and output into an LAS file (see below). Data processing is generally carried on a cluster.

## AUTOMATIC POST PROCESSING

After SGM processing including the combination of the results of ADS stereo pairs, several post-processing steps are carried out automatically. These include the final convolution using a 3x3 median filter and the segmentation-based outlier removal. Considering the huge amount of data, the SGM result is thinned in two steps: by a combination of 2x2 neighboring disparities into one as well as by intelligent point cloud thinning. It is beneficial to perform most of this work in the two-dimensional disparity space. The final output is a LAS file – for each DSM extraction job – that also provides base image intensities as well as RGB and/or FCIR color, the data for which are provided by the ADS.

### Segmentation-Based Outlier Elimination

Despite the consistency checks in the reverse matching and the redundancy from using two ADS stereo pairs in, the SGM result might still contain some outliers, e.g. in challenging areas such as water or trees. It has been proposed by Hirschmüller (2005) – and integrated into various implementations of SGM since – to assume that small isolated patches, which significantly differ in height (or disparity) from their neighborhood, are most likely errors. For the required segmentation, neighboring pixels in which disparities differ by more than 1 are considered part of different segments. Small segments are then removed from the disparity map; the threshold in the XPro DSM Extractor is 50 pixels in area.

### 2:1 Data Reduction

The comparison between ADS-based SGM and LiDAR point clouds by Gehrke et al. (2010) has shown that the resolution (ADS < 50 points/m<sup>2</sup>, LiDAR < 10 points/m<sup>2</sup>) and the horizontal accuracy (ADS: ~ 0.5 GSD, LiDAR: 10-30 cm) are better than LiDAR for medium and high-resolution ADS data (GSD < 50 cm), while the vertical accuracy is comparable to LiDAR (~ 5 cm) only for high resolution ADS imagery (GSD ~ 5 cm) – along with the much higher point density. These differences in accuracy as well as the experience that such a huge amount of data cannot be handled properly in the interactive post-processing as described below, we decided to reduce the amount of data along with increasing the vertical accuracy.

This is achieved by joining the disparities of 2x2 neighboring base image pixels, similar to a 2:1 minification of the disparity map. The consistency of respective values is verified: disparities that differ by more than 1 from the (initial) average are disregarded; at least two disparities are required to agree for the final result. Such averaging removes noise and doubles the vertical accuracy (for results based on 4 input pixels) and reduces the data by up to 75%. The result is considered the high-resolution output of the XPro DSM Extractor. Depending on the input image GSD, it can still exceed the density of LiDAR data and theoretically achieve more comparable accuracy properties.

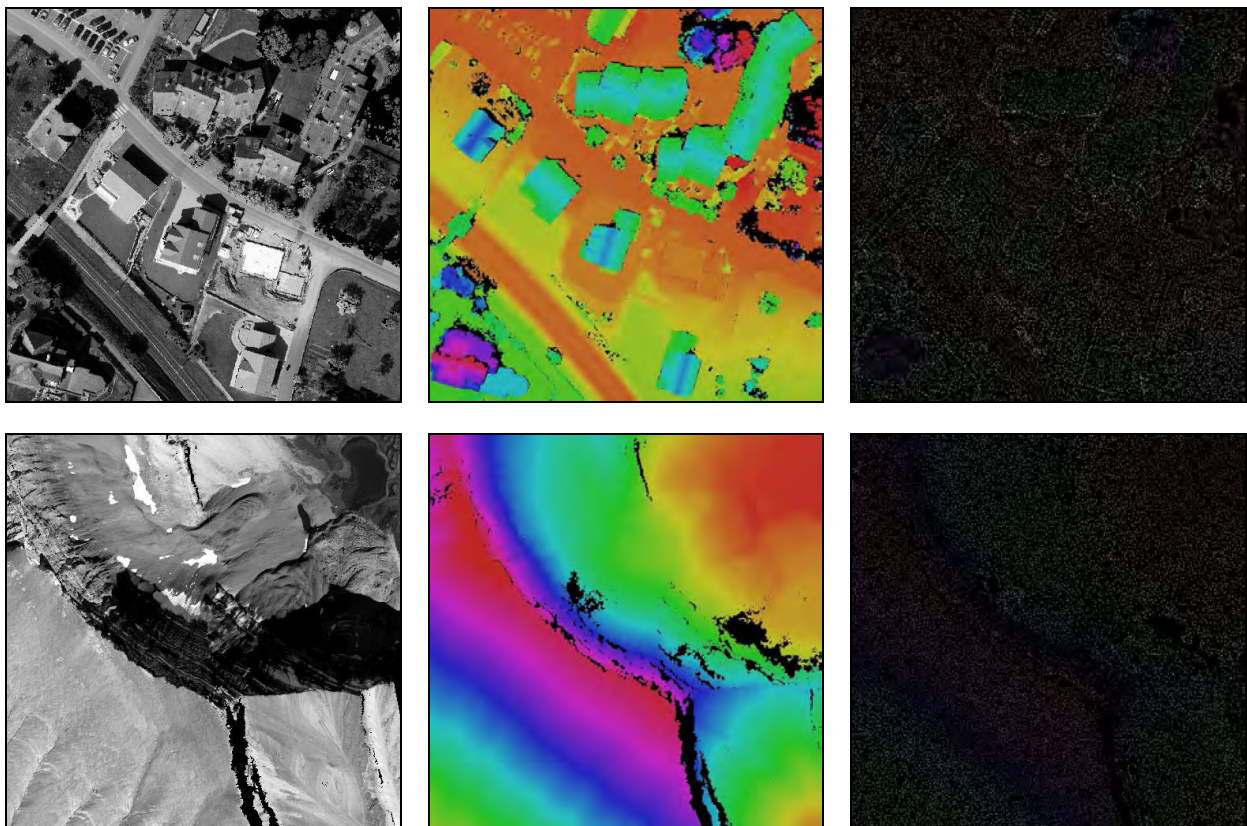
Note that the horizontal resolution of this final result is (almost) identical to an SGM output from the 2:1 pyramid level but the vertical accuracy is much better for two reasons: the use of full resolution imagery and, thus,

all available information; and the averaging of neighboring disparities. The advantage of pyramid level output is performance, a factor of (up to) 8 per level. In that regard, the XPro DSM Extractor provides the option of a ‘quick’ overview based on very fast computation using 8:1 minified imagery.

### Intelligent Thinning

Aside from the high-resolution, we decided to provide a second output that is thinned by keeping key points of the terrain. Depending on the particular data set and the thinning rate, such output is expected to be sufficient for various purposes including ortho-rectification.

There is a broad variety of methods for 3D point clouds thinning, e.g. Pauly et al. (2002), Moenning & Dodgson (2003) and approaches presented in there. However, the advantage of image-based 3D point computation is the availability of the 2D disparity map, which contains data in a regular (image) grid, i.e. in simple topology. Accordingly, image processing can be applied to identify key points or, respectively, points of significant curvature as provided by a convolution with a Laplacian. The application of a threshold for thinning is obvious but might deliver undesired results: While it will result in many points identified at discontinuities (buildings) and slopes (mountains) – and in our case around voided spots –, other regions might be underrepresented and accumulate errors, e.g. in larger areas of small curvature (plains). In that regard, Southard (1990) proposed a neighborhood ranking filter that provides a locally adapted curvature threshold: The curvature of the central point in is ranked with respect to its neighbors, and the most significant point(s) in terms of highest local ranking are selected. This ensures a globally even point distribution while keeping the locally most significant points for the thinned result. We extended this idea by considering only points within the same segment, based on the above-described segmentation. The main advantage is that thinned points will be kept on both sides of a discontinuity. This is illustrated in Figure 1, which shows points on roofs as well as right next on the ground; while points appear well distributed at first glance (i.e. globally), their local density increases around discontinuities.



**Figure 1.** Clouds of  $\sim 2,000 \times 2,000$  points, automatically derived with SGM from ADS data. From left to right: high resolution intensity, high resolution and thinned data with color-coded heights. Top: Romanshorn, Switzerland, based on 5 cm imagery. Bottom: Rocky Mountains in Jasper NP, Alberta, based on 40 cm imagery.

The thinning rate can be controlled by the size of the neighborhood ranking filter kernel in combination with the lowest rank that is accepted for the thinned result. The thinning rate can be estimated from the kernel size (squared) and the rank, e.g. the classification of highest ranked points within a 5x5 filter can be expected to reduce the amount of data down to roughly 4% ( $\sim 1/25$ ), similar to highest and second-highest ranks from a 7x7 filter ( $\sim 2/49$ ). Based on an investigation using different data sets – city, mountains, plains, and forest in different GSD –, we found that best results for a certain thinning rate are achieved by using solely highest ranked points. Respective results turned out to be closer in terms of RMS to the full resolution data. Hence, the thinning rate is controlled by altering only the neighborhood ranking filter size, with the 5x5 example considered medium thinning.

### **LAS Output: The Image Cloud**

The results of SGM including the automatic post-processing are projected into 3D space in a user-definable coordinate system. At the time of this writing, they are stored in the LAS 1.2 format as defined by the ASPRS (2008). Each DSM job will generate its own output, one file for the full resolution data that consists of 10-20 million points and an additional file for the thinned data that might still include up to about 2 million points, depending on initial coverage and thinning rate. Results can be generated in point data record formats 1 or 2. However, only some of the foreseen items, which are designed for LiDAR, are assigned from ADS/SGM results.

In point data record format 1, each object is assigned its coordinates, the intensity of the corresponding base image pixel and a default classification (ground). Since recently, we also provide the ability to store ADS color data. Those are generated in an on-the-fly rectification by projecting each object points into the red, green, blue, and near infrared color bands and assigning the interpolated, radiometric corrected intensities. Since the LAS version 1.2 only allows storing three color components, two LAS files are generated – the first file is encoded with RGB and the second file is encoded with FCIR. Both contain the identical geometric point cloud, which leads to altogether four files: High resolution RGB and FCIR and thinned RGB and FCIR. With this color information and the panchromatic intensity, the high density point cloud becomes a new type of product: the multi-band ‘image cloud’. Examples of which are shown in Figure 2.



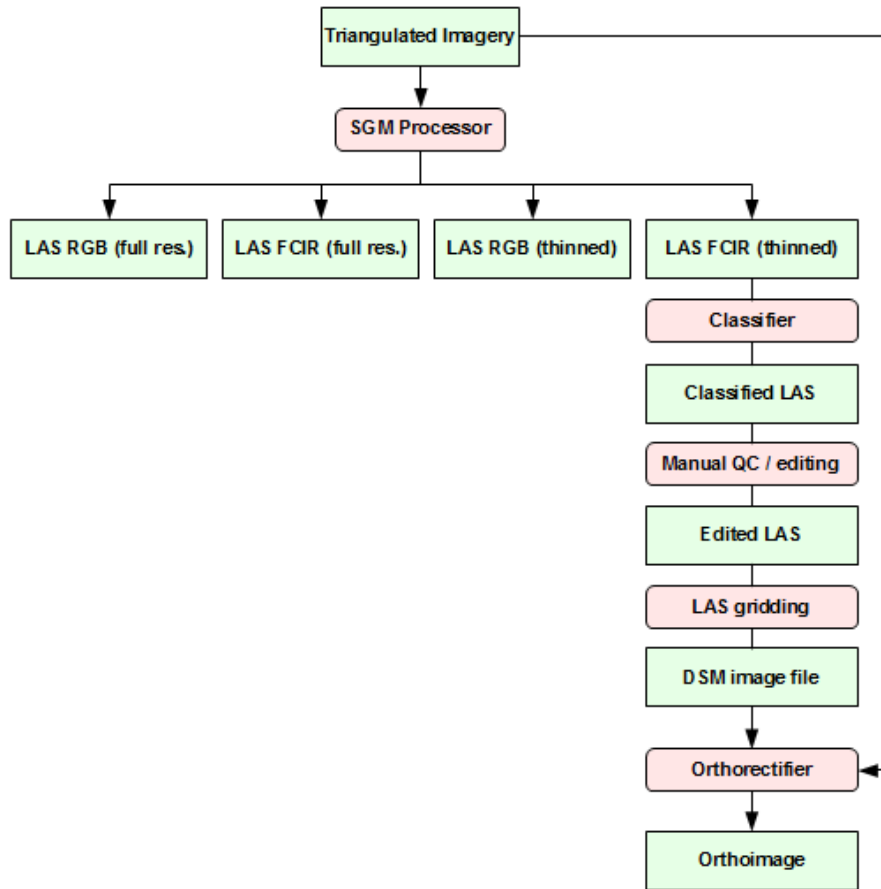
**Figure 2.** TIN representations of RGB image clouds from San Antonio, Texas; automatically derived from 15 cm ADS imagery without no manual post-processing involved.

## **PRODUCTION WORK FLOW**

In the following, the production work flow at North West Geomatics is discussed in the context of processing non-urban areas at medium to low resolutions, i.e. GSDs between 30 cm and 1 m. The work flow that is required to process imagery containing buildings and other man-made structures, especially at smaller GSD, is still under review and its discussion is outside of the scope of this paper.

As outlined in Figure 3, the DSM Extraction application, which is part of the Leica XPro software package, generates four LAS 1.2 files containing the point clouds from the input triangulated imagery: full resolution and thinned data, each in a version with RGB and FCIR encoding. Although the full resolution LAS files for one take of ADS imagery is already reduced to 2:1, they typically contain about a billion of points. Unfortunately, most

commercially available software packages for LAS editing and viewing do not support LAS containing such data volumes. Therefore production at North West Geomatics exclusively works with thinned point clouds. The default setting for thinning is 'medium', which leaves approximately 4% of the number of points compared to the full resolution. Even at this thinning rate the number of points per image strip can easily reach in the tens of millions of points.



**Figure 3.** North West Geomatics DSM processing work flow.

The final goal of the work flow is to generate ortho products from the triangulated data and its derived surface model. The main challenge is the creation of a DSM that contains well defined water bodies and does not have any erroneous points or outliers. The SGM process and the automatic post processing as described above do an excellent job at removing outliers in regular image content, but in some cases water bodies as well as shadow areas do require manual clean up. In none-urban areas shadows rarely appear, mostly when flying at low light conditions in areas with strong relief, e.g. rock formations in mountainous regions. At North West Geomatics a classifier is run to determine water bodies based on the Normalized Difference Vegetation Index (NDVI), which is derived from the intensities (DNs) of the red and near infrared (NIR) ADS bands:

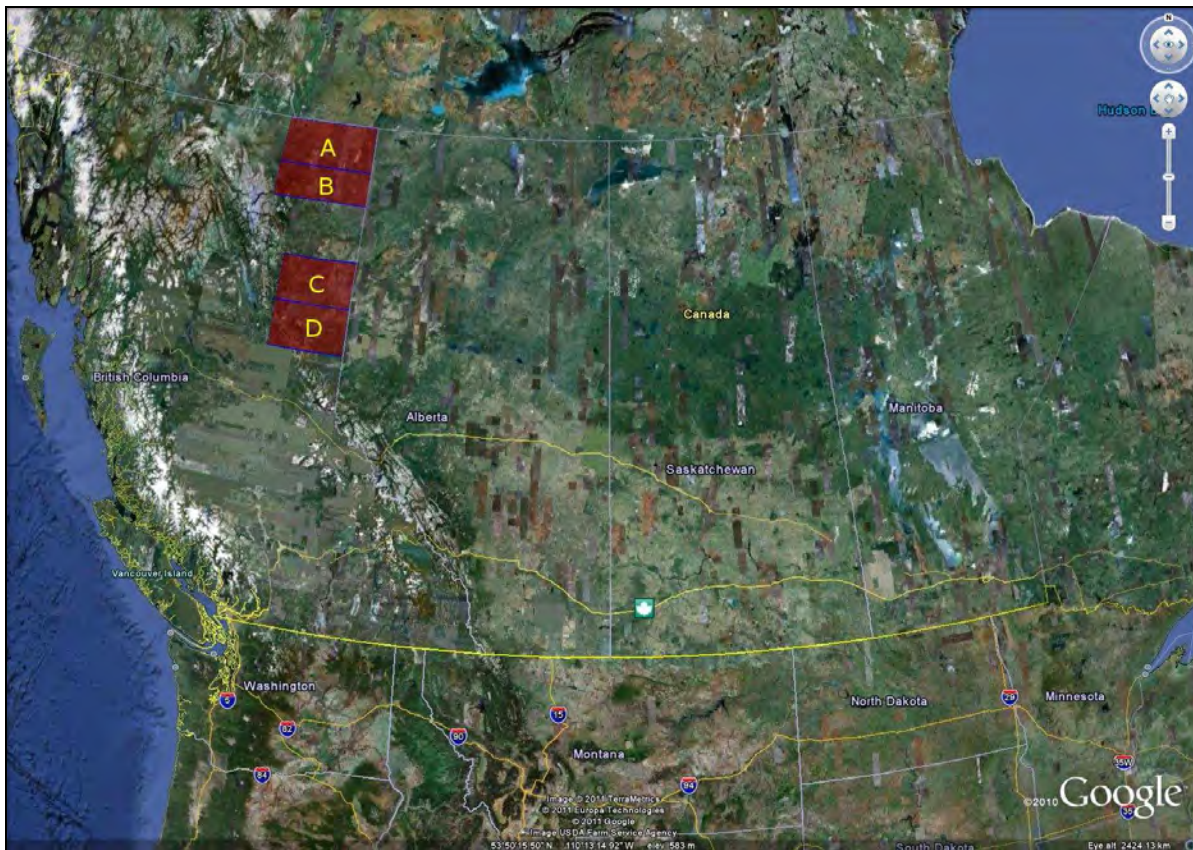
$$NDVI = \frac{DN_{NIR} - DN_{RED}}{DN_{NIR} + DN_{RED}}$$

The NDVI threshold of -0.3 has been determined as the most successful to classify water bodies in ADS imagery. The classification step is followed by a manual quality control and editing step, for which off the shelf commercial LiDAR processing software, such as Terrasolid's Terrascan, is utilized. The output is an image cloud that is then gridded to a raster file so it can be used by XPro's ortho-rectifier. The grid posting is typically 10x the image GSD. In some cases the raster file may also be smoothed with a low pass convolution filter, especially if the emphasis on the final product is aimed at the forestry market to void blurring of forest canopies.

It should be noted that the image cloud itself is also a deliverable product. An increasing interest in such an image cloud can be detected, but unfortunately the commercially available tools for viewing and editing such data are not quite up to the challenge that the volume of data present, at least for the full resolution image clouds.

## PRODUCTION EXAMPLE

The example project presented in this paper is a typical production use case. The project area in Northern British Columbia covers approximately 85,000 sq km (33,000 sq miles) and was processed to 40cm (15.7 inches) GSD. The project layout is illustrated in Figure 4 and shows the four blocks A, B, C and D. All blocks were acquired fairly late in the flying season – mid September through beginning of October – at low solar elevations in the range of 30 degrees, leading to long shadows in the imagery.



**Figure 4.** Location of the blocks A, B, C, and D in Northern British Columbia.

In the following Figure 5 a small patch in the middle section of Block D are displayed. Note the dense and even distribution of points, even in the thinned LAS file which contains only approximately 3% of the full resolution LAS file. Figure 6 displays the final result of the all the image clouds gridded to a 5 m DSM. Note that the grid files are only copied together for this display and not truly mosaicked, because as the rectification process will utilize one DSM grid file per ADS image take.

### Performance

The B.C. blocks have been processed on computing nodes with 8 core 2.93 GHz CPUs and 8 GB RAM. Each SGM jobs is processed entirely on one node. Also, each job creates one set of LAS files (full resolution RGB, full resolution FCIR, thinned RGB and thinned FCIR).

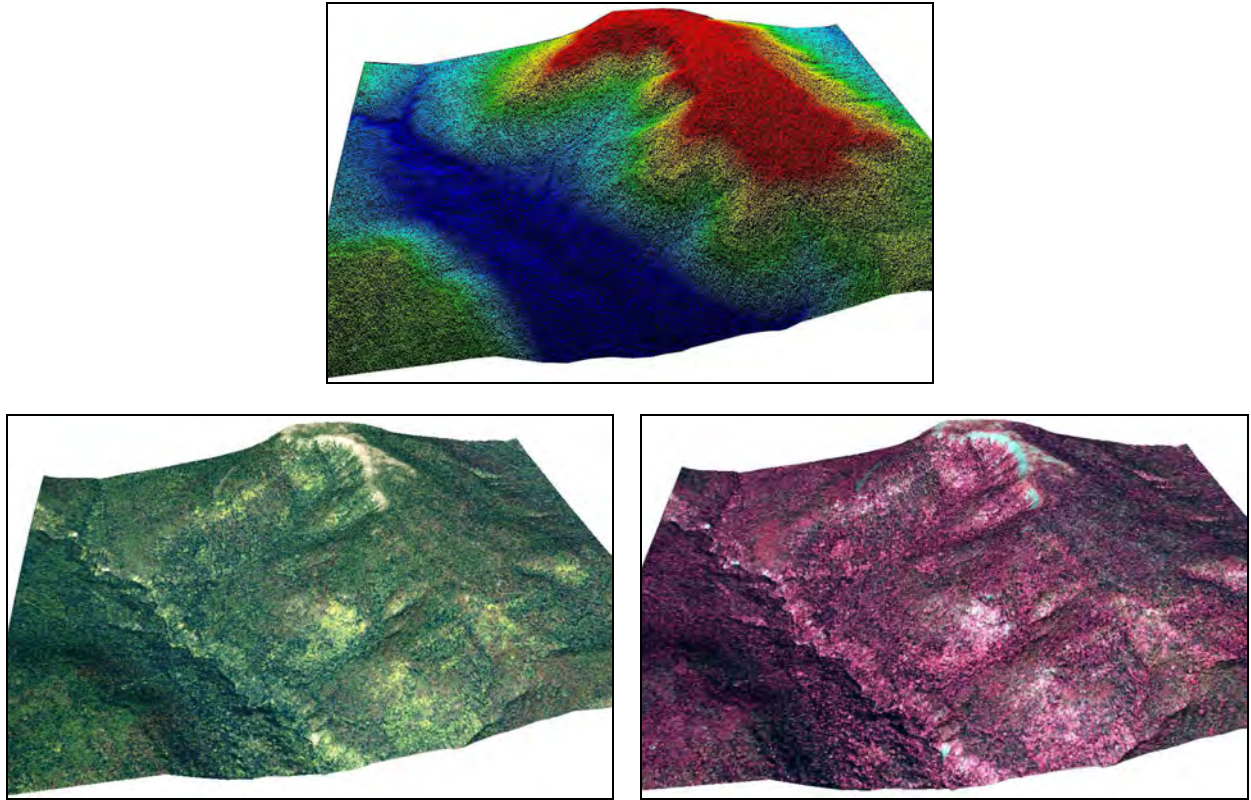


Figure 5. Encoded thinned LAS files shown as 3D elevation image (top), 3D RGB and 3D FCIR.

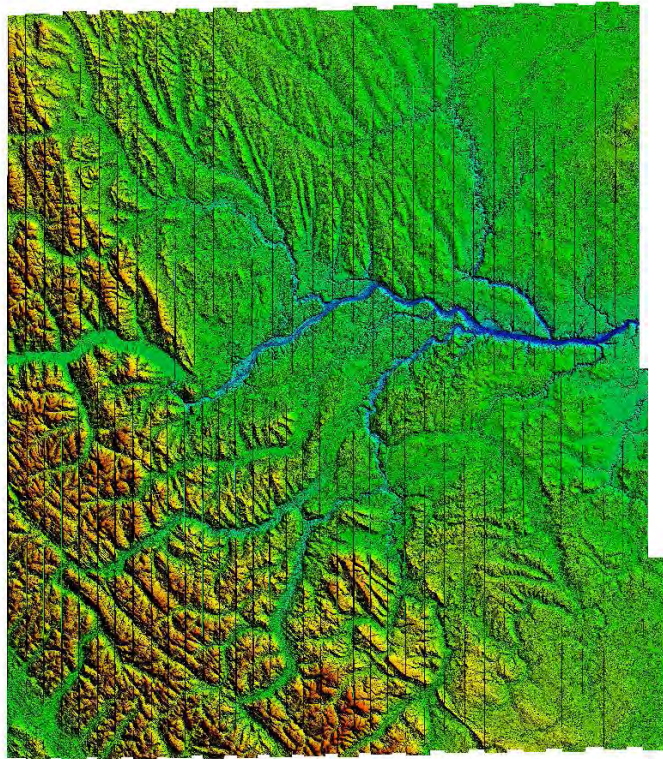


Figure 6. Blocks C and D, gridded DSM overview with color-coded elevation.



**Table 1.** Northern British Columbia performance numbers

<b>Block</b>	<b>Image strips</b>	<b>Number of Jobs</b>	<b>Secs per job (average)</b>	<b>Points/sec</b>	<b>Time in hours</b>
A	38	856	2103	43,881	500
B	39	672	2127	43,387	397
C	34	750	2597	35,534	541
D	45	1067	3300	27,964	978
<b>Total</b>	<b>156</b>	<b>3345</b>	<b>2600</b>	<b>35,487</b>	<b>2416</b>

Table 1 illustrates the relationship between terrain and processing performance. As seen in figure 5, Block D contains much more relief than Block C which results in significantly longer processing times per job when comparing these two blocks. For the whole project the 14 node cluster deployed at North West Geomatics required approximately one week to generate all the point clouds. The manual quality control and editing of shadows and water bodies required four man weeks.

Finally, the ortho-rectification processing time for blocks A and B totaled 102 hours, or in other words, for each hour of ortho-rectification time approximately 10 hours of DSM processing time, excluding manual editing, is required.

## CONCLUSIONS

This paper has outlined the use of Leica's XPro DSM extraction software package in a production environment at North West Geomatics. It has been shown that using the SGM approach produces high density image clouds that can be used for ortho-rectification and can stand by themselves as products. Medium resolution non urban regions are ideal for the current DSM extraction implementation. Large areas, as shown in the British Columbia project, can be processed within weeks and are very cost competitive compared to LiDAR acquisition. There are still challenges editing and viewing these data sets due to the sheer data volume. The information value of image clouds is very high, but the currently available tools have not yet reached the ability to provide smooth and effective means for processing and visualization. A wealth of information is contained in the high resolution image clouds, but currently only the thinned output can actually be utilized.

In the future, the use of higher resolution imagery, especially for urban regions, will be investigated for production use. The results from these investigations will flow back into the DSM extraction software. The upcoming LAS 2.0 specification will allow storing more than 3 values for encoding thus eliminating the need to create individual files for RGB and FCIR encoding. Potentially, also other properties such as the NDVI values could be stored in the LAS 2.0 file. Software vendors providing solutions to handle image clouds will need to accommodate the need to process and visualize highly dense data sets which are now not only coming from advanced LiDAR systems, but also from ADS imagery.

## REFERENCES

- ASPRS, 2008: LAS Specification 1.2. [www.asprs.org/society/committees/standards/asprs\\_las\\_format\\_v12.pdf](http://www.asprs.org/society/committees/standards/asprs_las_format_v12.pdf), accessed Feb 21, 2011.
- Gehrke, S., K. Morin, M. Downey, N. Boehrer, and T. Fuchs, 2010: Semi-Global Matching: An Alternative to LiDAR for DSM Generation? *Int. Arch. Phot. & Rem. Sens., Calgary, AB, 38(B1)*.
- Hirschmüller, H., 2005: Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. *Proc. IEEE Conference on CVPR, New York, NY*.
- Hirschmüller, H., 2008. Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2).
- Hirschmüller, H., and S. Gehring, 2009: Stereo Matching in the Presence of Sub-Pixel Calibration Errors. *IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL*.

- Mascardi, V, 1998: Extraction of Significant Terrain Features from RSG and TIN: A Survey. *Assignment for the Ph. D. Course of Computer Graphics*, [www.disi.unige.it/person/MascardiV/Papers/VivianaPublications.html#GRAPHICS98](http://www.disi.unige.it/person/MascardiV/Papers/VivianaPublications.html#GRAPHICS98), accessed Feb 21, 2011.
- Moenning, C, and N. A. Dodgson, 2003: A new point cloud simplification algorithm. *Proc. 3rd Int. Conf. on Visualization, Imaging and Image Processing: 1027–1033*.
- Pauly, M., M. Gross, and L. P. Kobbelt, 2002: Efficient simplification of point-sampled surfaces. *Proc. 13th IEEE Visual.: 163–170*.
- Southard, D. A., 1990: Piecewise Planar Surface Models from Sampled Data. *Scientific Visualization of Physical Phenomena*, Springer, New York, NY: 667-680. (Citation: Mascardi, 1998).