OPTIMIZING THE USE OF DIGITAL AIRBORNE IMAGES FOR 2.5D VISUALIZATIONS

Aleksandra Sima and Simon Kay

Joint Research Centre, Institute for Protection and Security of the Citizen, Agriculture Unit, Ispra, Italy; {aleksandra.sima/simon.kay}(at)jrc.it

ABSTRACT

Monoscopic virtual representations of 3D geometries are rapidly becoming important products of many databases and software applications. Many GIS tools – even freeware, such as Google Earth – permit the visualization of city planning models as well as landscapes derived from 3D geometries (digital surface models draped with imagery, called 2.5D visualization). These applications also are steadily becoming less qualitative, and more metric, as they are integrated into GIS environments. Up until now, such image rendering has usually been made with non-photogrammetric sensors, and has not been based upon state-of-the-art air survey systems. In the photogrammetry domain, the orthogonally projected image remains the paradigm. This approach however neglects imagery that may better represent the surfaces of objects such as building facades. We propose that off-nadir parts of vertical imagery – typically ignored after the orthorectification process – provide us systematically with much data that can be used to optimize the 2.5D rendering process.

Keywords: DSM, rendering, 2.5D visualization.

INTRODUCTION

Virtual views of rural or urban landscapes presented in 2.5D viewing – that is, monoscopic rendering of imagery draped over a detailed surface model – have came of age in 2005. This represents a major shift from the 2D (orthogonal projection) or 3D (stereoscopic, usually also vertical viewing) viewing modes. On-line services such as Google Earth have placed some level of 2.5D capability on desktop computers, albeit still at a fairly generalized level of texture detail, with free software and standard Earth observation imagery (satellite, airborne).

However, all state of the art software and service providers of this nature use Earth observation data that has been orthorectified, i.e. geometrically corrected to some projection system and sampled so as to choose the most nadir viewing image available. This approach derives directly from the photogrammetric paradigm of 2D cartography predominant during the 20th century. Nevertheless, more and more imagery is acquired with multiple views – sometimes as many as 6 or more – that are rarely used in the final orthoimage product or 2.5D rendered viewing. These data offer a potentially rich source for the optimization of textures used in rendering of any image products, either orthoimages or 2.5D rendered views.

Moreover, interest in using such visualization in metric applications is increasing, as GIS tools become more able to permit direct measurements in such 2.5D scenes and visualization scales become larger. Data sources and computer power now combine to permit building frames to be approximated reasonably well, giving rise to digital surface models derived (for example from LIDAR) with a 1 m grid spacing being a fairly common product in many cities.

Nevertheless, the focus in recent years has been on the production of orthoimage products which eliminate all significant parallax (1). The photogrammetric requirements of so called "fully rectified" products are high resolution imagery and a DSM with a commensurate density (usually not more than 4 times the resolution of the imagery). High spatial resolution is usually achieved by flying low; the inevitable result is increased parallax in the image. The airborne imagery is therefore usually acquired with a high degree of sidelap and forward overlap, ensuring that sufficient geometry information is available for DSM production, as well as providing a near-nadir viewing pixel for every point in the scene. The procedure is somewhat cumbersome with traditional film cameras, but three-line airborne scanners and digital frame cameras (2) are well suited to the high frequency of image exposures (often as low as 1Hz) and subsequent high data volumes from such a flight. The products – 2D orthoimages with near-perfect geometry, and 2.5D visualizations using the highly detailed DSMs – are impressive, and have quickly become economically competitive data sources for large scale urban applications such as land registry updating.

Fully orthorectified imagery flight plans, however, produce an enormous amount of redundant data. Flights with 80% forward and 60% sidelap may produce as many as 15 pixels for a given point on the ground. Furthermore, the restriction of using only the most nadir viewing pixels causes 2.5D rendering to be "smeared" over vertical surfaces (perpendicular to the orthogonal projection plane), such as building facades. We propose instead to exploit the remaining data, and instead of using only (near) nadir viewing imagery as a data input, to choose those pixels with an optimized geometry for the rendering process. In our approach, the optimized pixel is defined as the pixel corresponding to the view angle most orthogonal to the DSM facet that models the true object surface. Figure 1 presents a simplified schema for choosing the optimal pixel from a set of three images.

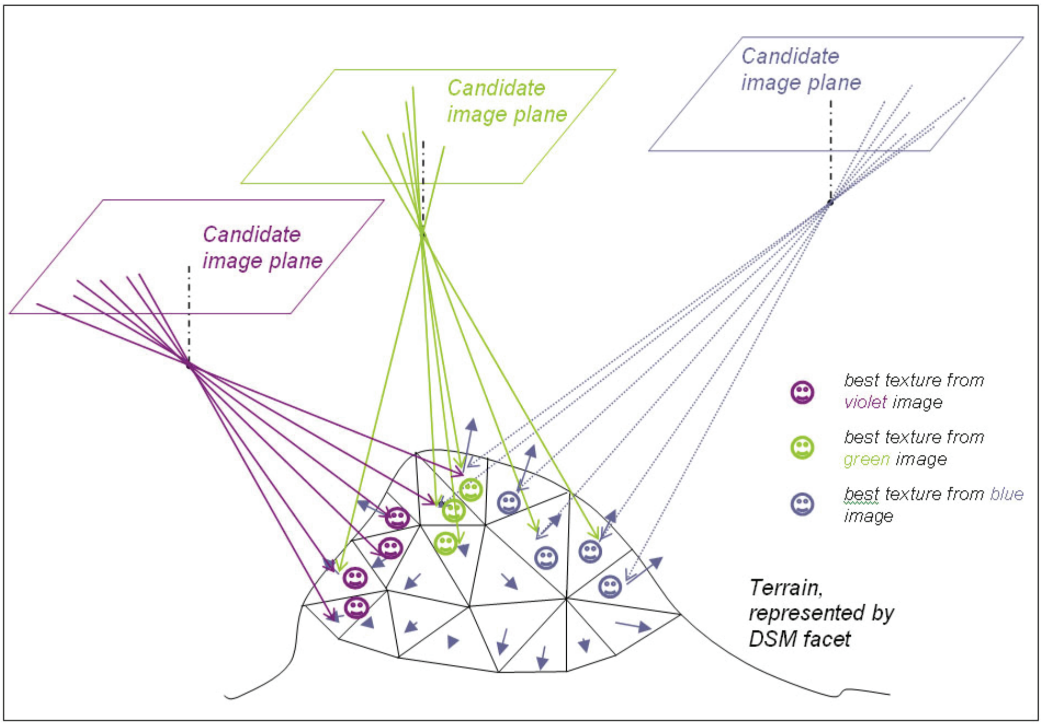

Figure 1: Pixel selection optimization schema, where several candidate images are assessed to determine which image plane was best positioned to permit rendering of a triangular facet terrain surface, by comparing the direction of the vector normal to the triangular facet with the view vector for each corresponding pixel in the candidate images.

Current approaches for rendering vertical surfaces

Scene reconstruction is a computer vision domain that has been growing rapidly in recent years. For example, Sequeira (3) reported on an integrated approach to the construction of textured 3D scene models of building interiors, using laser range data and image scanners. The prototype system was mounted on a mobile platform, and embedded software performs several functions, including triangulation of the range data, registration of video texture, as well as registration and integration of data acquired from different capture points. More recently, Boström (4) and Früh (5) present more sophisticated results of such an approach, with impressive demonstrations of highly detailed interactive rendering of city scenes, using data collected from mobile vehicle-platforms. The vehicle speed used was 7 - 17 km·h^{-1}, and facades were successfully imaged up to approximately two storeys, although these parameters are somewhat dependent upon the system configuration.

The absence of data for taller buildings, as well as roof textures, has led researchers in this area to examine integration with airborne or satellite imagery. In particular, Früh and Zakhor (6) used aerial imagery and a DSM to improve the truck-based laser scanning of their ground observations; however the airborne imagery was only integrated in their products in further experiments (7,8). In this latter case, oblique imagery was acquired from a helicopter using a standard (i.e., non-photogrammetric system) digital camera. The camera positions for each oblique image were then estimated using image matching with line segments collected on the ground. Ground based data were then fused with the airborne imagery, which replaced image elements not sufficiently well imaged from the ground based system.

METHODS

The study we present here is more aligned to a standard photogrammetric approach, by contrast with other work in the computer vision field. We attempt to use nominally vertical imagery, collected under typical photogrammetric conditions, to improve the rendering of vertical surfaces and facades. Image orientation is achieved only through the use of a precise camera model (including metric calibration information) and camera positions are provided by the aerotriangulation of a photo block. The DSM is furthermore also created using a standard photogrammetric approach and is of a quality commensurate with that used for 2D orthoimage production.

Study dataset

The images used in this study were collected in Maussane les Alpilles, France, on the 24th of May 2005 by Aerodata International Surveys (Deurne, Belgium). A Vexcel UltraCamD digital frame camera was used to collect 80 images in 5 strips, with 80% forward overlap and 60% side overlap. The flight plan was essentially of classical photogrammetric form (2), with nominally vertical images and a flying height of around 5000 m above ground level. The data used for prototype development were four pansharpened multispectral image frames (henceforth termed *candidate images*), with a ground sampling distance of the GSD of approximately 0.5 m. Exterior orientation ancillary data derived from a standard aerotriangulation process were also made available, comprising the centres of projection (in UTM projection zone 31N, WGS-84 ellipsoid), camera rotation angles, as well as the calibration report of UltraCamD (Vexcel serial no.18) camera.

The terrain model used in this study was a digital surface model prepared by ISTAR S.A. (Sophia Antipolis, France) with a 2 m raster cell size. It was extracted (using spatial autocorrelation techniques) from a Leica Geosystems ADS40 airborne sensor image set, acquired 14th of May 2003. The flight altitude was also approximately 5000 metres above ground level, corresponding to an image ground sampling distance of 0.5 m. The vertical accuracy of this DSM was assessed using 54 independently surveyed check points (9). The vertical quality measure expressed as $RMSE_z$, on well defined points, was computed to be 0.34 m.

Algorithm optimization

The methodology can be divided into four parts:

- Absolute orientation of candidate image arrays: The main goal of this part is to calculate coordinates of each pixel of each candidate image frame in the project global coordinate system (UTM, zone 31N, ellipsoid WGS-84). The process is similar to a solution of the interior and absolute orientation parameters of each pixel in every image;
- DSM processing: The DSM is reprojected back into the candidate image space; each triangular facet of the DSM is represented as an irregular network in the candidate image space;
- The information concerning the DSM is used to determine which candidate image frame is best suited for rendering into the 2.5D scene; virtual image files are produced;

- Finally, the image rendering is undertaken in a specialized visualization software package (Reconstructor) using the virtual image files.

Figure 2 presents a schematic overview of the processing flow, described in more detail below.

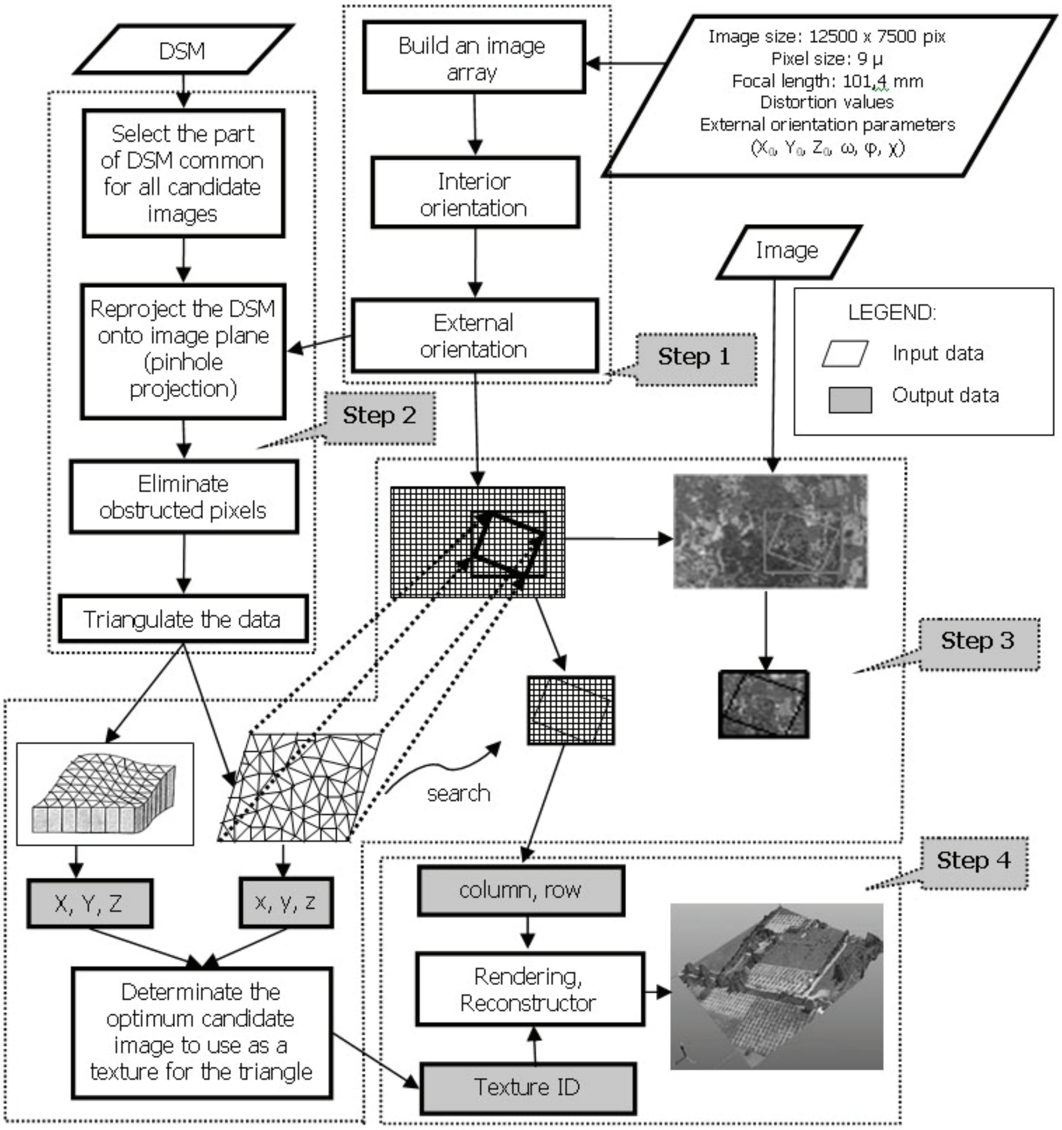

Figure 2: Pixel selection optimization is a step-wise set of computations or transformations made on a particular data type (DSM, candidate image, etc.) Step 1 is mainly concerned with candidate image absolute orientation, Step 2 with DSM preprocessing, Step 3 with assigning indexes of view geometry to candidate image pixels in preparation for rendering, and Step 4 the rendering itself in a specialist visualization package.

Image array absolute orientation

The first step in the absolute orientation is the transformation, for each candidate image, of the camera pixel coordinate system to the metric coordinate system. The algorithm proceeds by correcting for radial distortion around the origin of the image coordinate system; in case the PPS (principal point of best distortion's symmetry) is given in the camera calibration information, there is a need to move the origin to that point, correct the distortion, and return to the original point.

Each pixel on the candidate image frame is then assigned geodetic coordinates in the world coordinate system, using the classical collinearity equation.

In a typical exterior orientation model (built using at least two images), the scale coefficient λ describes the difference between the model and terrain scaling. In our (single) image exterior orientation, λ is always equal to 1; there is no change of scale, because the goal is to know the geodetic coordinates of each image pixel on the image plane and then be able to rebuild the image geometry at the moment of acquisition.

DSM processing

Re-projection back into candidate image space

When the image position in the global coordinate system is known, it is possible to form an equation that describes the image plane in the geodetic coordinate system.

In order to determine the viewing geometry of each candidate image with respect to the DSM elements, we must determine a vector normal to the image plane. For any 3 non-collinear points on the plane, it is possible to define two vectors within that plane. The cross product of two vectors is always a vector orthogonal to both of these vectors, thus orthogonal to the plane. It is therefore convenient to use the cross product result as a normal vector.

For each cell of the DSM it is possible to write an equation of a line connecting ground centre of the pixel with the image centre of projection. Each of these lines will intersect the image plane in a particular point and it is possible to calculate coordinates of intersection points for all cells. These points act as an index of the terrain in the same coordinate geometry as the candidate images.

Elimination of obstructed pixels

The reprojection of the DSM array to the plane has some disadvantages. The reprojection is calculated for all the DSM cells from array, even those which are not visible from the centre of projection (for example, on facets facing away from the camera position or occluded from view by tall objects). These obstructed pixels need to be identified and eliminated. We implemented a simple algorithm that evaluates the reprojected position of each DSM cell on the image plane, and determines if adjacent pixels are obstructed (Figure 3).

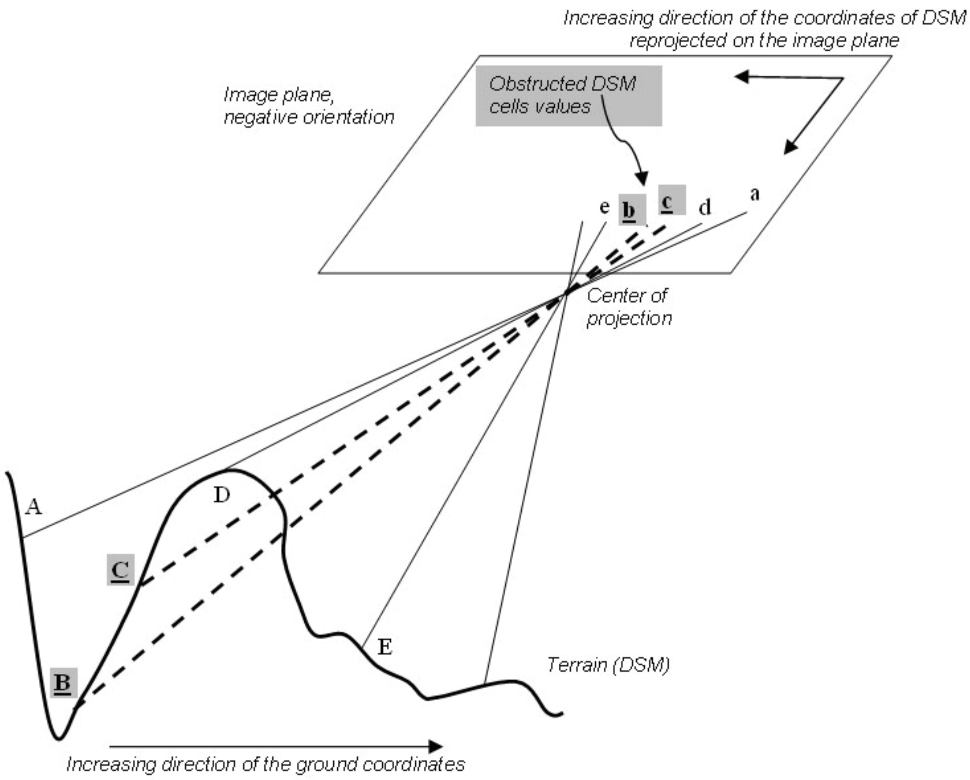

Figure 3: When the analysis of the DSM cell coordinates on the candidate image plane shows a reversal of direction, the cells are marked as obstructed and eliminated from subsequent processing.

Finally, the now irregular grid of reprojected DSM data must be networked as triangular facets. Only those pixels visible from the centre of projection of each exposure are selected for this process.

When it is determined which pixel on the candidate image corresponds to the cell in the DSM array, it becomes possible to find coordinates on the image plane corresponding to vertices of triangles on the ground.

Integrating DSM and candidate images

Extracting image "textures" for rendering

To extract the texture from candidate image we have to find geodetic coordinates of all vertices in each triangle on the ground in the pixel coordinate system (column, row) of the image. In the previous step the coordinates (irregular, geodetic) on the image plane of all vertexes on the ground were computed. Thus the next step is to find the location of these points (irregular) in the image geodetic coordinates array (regular). After this task is completed it is possible to establish the relation between any triangle vertex on the ground and corresponding triangle on each candidate image.

Optimization of rendering source

To determine which candidate image is the best choice for texturing any individual DSM triangular facet, it is necessary to compare geometry of acquisition (view angle from the centre of projection to the centre of the triangular facet) of each candidate image with the normal to this triangle. The best texture is the candidate image pixel with a view angle most orthogonal to the DSM facet, so the angle between the normal vector and the 'view on triangle' vector should be as close to 180° as possible.

Algorithm implementation and visualization of rendered imagery

The algorithm was implemented in Interactive Data Language (10), an array-oriented language with numerous mathematical analysis image processing and graphical display techniques. Use was made of a Linux operating system on a PC platform due to better possibilities of memory management, a critical factor with the processing of the large datasets (the whole image matrix consists of 11500x7500 pixels, double-precision data). Processing times on a 2.4GHz machine were quite demanding, in the order of 6 hours for image sets of only 800x800 pixels.

The resulting output files from our prototype algorithm are then processed using a script written in MatLab programming language, which prepares files readable for the visualization software Reconstructor (11). The examples presented in the paper all use this software for rendering the 2.5D visualization.

RESULTS

The prototype algorithm was able to correctly select candidate images from the test dataset, and make virtual images available for rendering in the visualization software environment. Examples of the test results can be viewed in Figures 4 and 5 below.

Figure 4 shows a typical rural landscape from the Mausanne dataset. Flat fields are bounded by hedges and trees, which provide vertical barriers in the landscape. The raised objects in the DSM have been rendered very accurately with the appropriate image data. The textures appear to be very compatible with adjacent pixels derived from different candidate images, and no artifacts are apparent in the 2.5D virtual image.

Figure 5 shows a pseudo-colored image, with each of four colors representing one of the input candidate image frames. It can be seen easily that raised objects (hedges, copses) have acquired textures from up to four images, which is rendered in accordance with the aspect and slope of the DSM facet being processed.

Although these images show the successful execution of the optimization algorithm, the generalization due to the DSM cell size (2m) still causes a large degree of smoothing in the rendered image. Furthermore, candidate images were not chosen optimally with respect to overlap, resulting in a set of candidate images that in fact were near-vertical viewing.

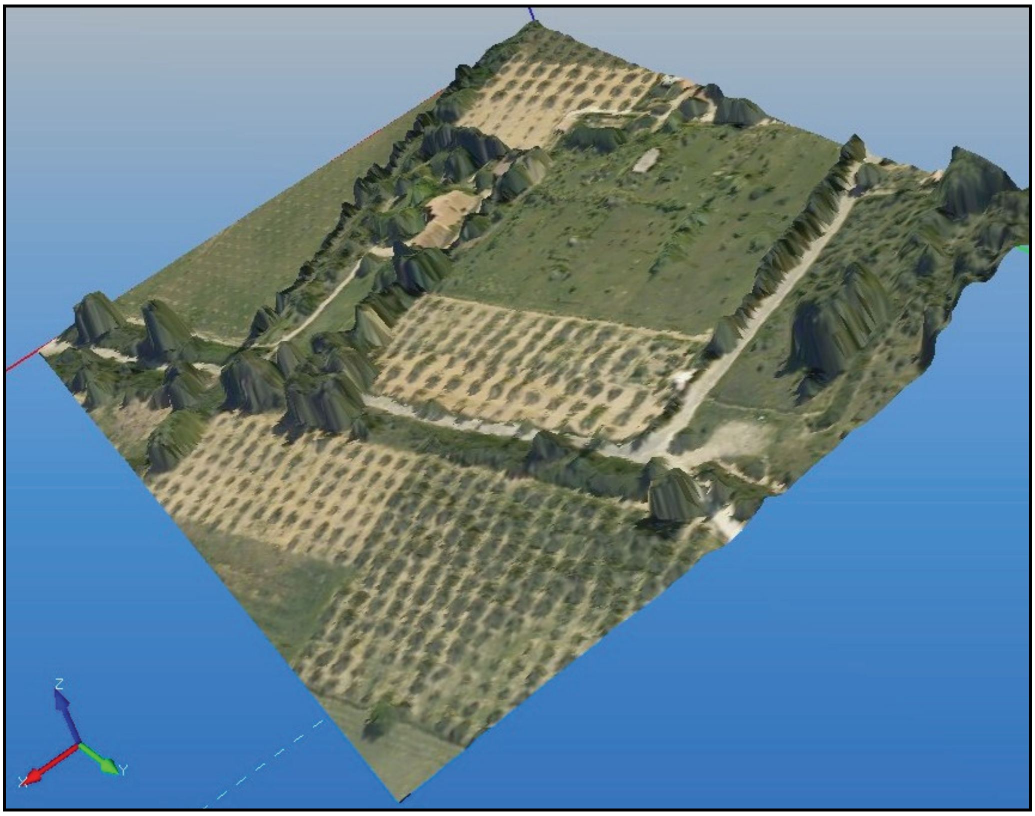

Figure 4: The optimized 2.5D view sample: texturing with real images. It can generally be noted that raised objects (trees, hedges, buildings) are rendered in the correct geometric position.

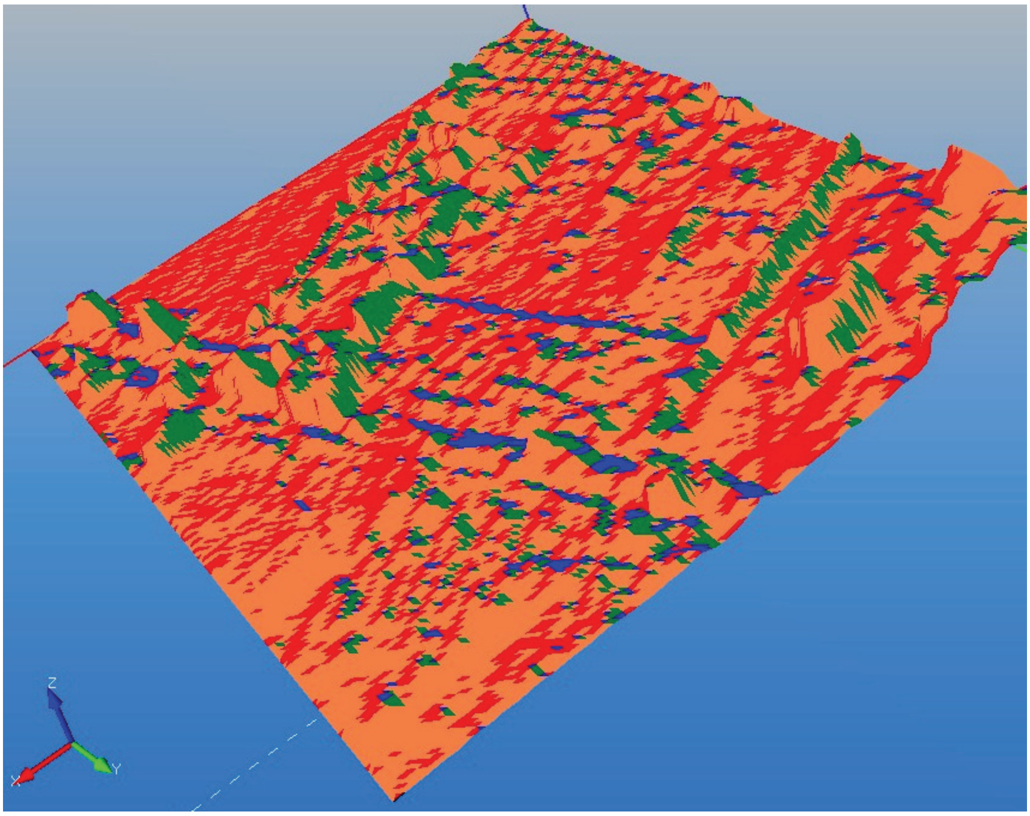

Figure 5: Texturing of optimized 2.5D view sample. The image and view point is identical to Figure 4 above. In this graphic, each color is a substitute of the texture from one of the four candidate images. It can thus be observed how the selection is determined by the orientation of each facet, not just in terms of aspect, but also of slope.

In order to try and overcome these circumstances – as well as to test the implementation on an independent set of candidate images not used for the development – a second dataset (Brasschaat, Belgium), acquired by Aerodata International, was processed using the prototype algorithm implementation. The input data were similar to the original test site, acquired with the same digital camera; however, image scale was larger (20 cm ground sampling distance) and with a 1 m resolution DSM (also created through autocorrelation but this time from the original imagery). Moreover, the scene is composed of an urban environment, with numerous buildings as well as landscape objects such as trees and bushes. The result of the 2.5D rendered image can be seen in Figure 6.

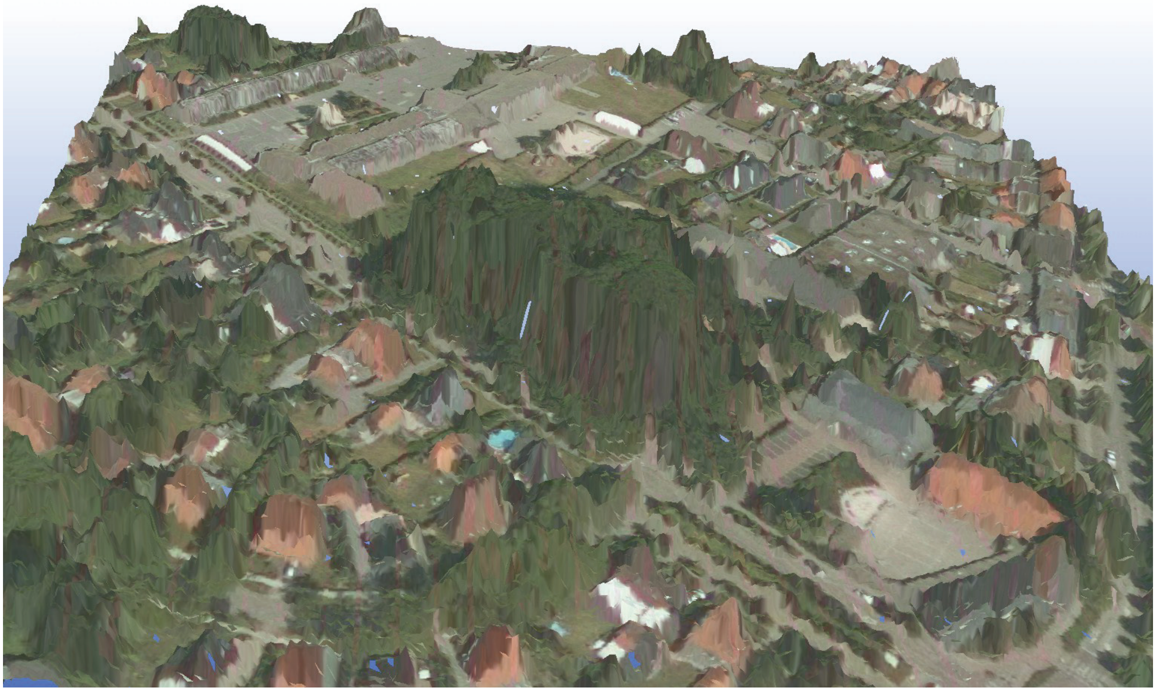

Figure 6: Supplementary trial with higher resolution dataset (Brasschaat, Belgium). Even though the DSM and image data are of higher resolution, visualization is not yet fully optimized. This is mainly due to a) the relatively vertical nature of the image acquisition, and b) the smoothing of the DSM, typical of terrain information originally destined for high quality 2D orthorectification.

Again, the algorithm performs satisfactorily with respect to the general selection and positioning of image texture. However, the DSM is again rather smooth, causing the detail of buildings to be disrupted. Although some facades appear to be rendered with texture colours differing from those used for building roofs, in fact there is still very little data on vertical surfaces available for the rendering algorithm from such a predominantly nadir viewing flight configuration.

CONCLUSIONS

We have demonstrated here a proof of concept algorithm for the optimization of imagery for 2.5D visualization. The algorithm has been implemented in a standard image processing tool and is able to operate on frame camera imagery, such as the Vexcel UltraCamD. Results show image quality at least equivalent to state of the art 2.5D rendering where standard “most nadir” vertical orthoimages are used.

Limitations of the results so far are due to the rather vertical nature of the datasets tested, rather than any specific issues with the algorithm itself. The visualization results are still inferior to datasets where either terrestrial imagery has been used for facades (for example, (4)), or where specially acquired oblique airborne imagery is integrated in the rendering process (8). We believe that our results would be improved if imagery were acquired with a wider field of view and from a lower flying height; we conclude that the algorithm would nonetheless perform effectively with such data since there is no fundamental change in the geometry of image acquisition. Additionally, a more refined DSM (again, a likely benefit of such a modified data acquisition approach) would significantly improve visualization results.

Many areas for optimization of the algorithm implementation remain. The triangulation performed by IDL for grid data is not optimized, resulting in a large number of unnecessary triangles in areas where the terrain is smooth. Data processing is consequently slow, requiring around 6 hours to prepare the optimized 2.5D visualization of an area represented by approximately 800 × 800 pixel image sample; much potential exists to improve processing speeds. Re-implementing the algorithm in a dedicated program – and not a scripting environment – would also make significant improvements in terms of speed. Other areas that could be optimized are the procedure for identifying obstructed pixels, and of course incorporating the observer's viewing geometry into the rendering algorithm to ultimately render the best possible data in the virtual 2.5D image.

ACKNOWLEDGEMENTS

We would like to thank: Fred Hagman and Willem Philipse (Aerodata International Surveys, Deurne, Belgium), Gunnar Boström and Vitor Sequeira (Nuclear Safeguards Unit, JRC Ispra, Italy), Giovanni Banchini (CGR S.p.A., Parma, Italy), Arthur Rohrbach and Udo Tempelmann (Leica Geosystems AG, St. Gallen, Switzerland).

REFERENCES

1 Devos W & S Kay, 2005. Assessing geometric accuracy of ortho-imagery: Towards product specification of fully orthorectified imagery for urban areas. GIM International, 19(6): 14-17

2 Graham R & A Koh, 2002. Digital Aerial Survey: Theory and Practice (CRC Press, Boca Raton, Florida) 248pp.

3 Sequeira V, K Ng, E Wolfart, J G M Goncalves & D Hogg, 1999. Automated reconstruction of 3D models from real environments. ISPRS Journal of Photogrammetry and Remote Sensing, 54(1): 1-22

4 Boström G, M Fiocco, D Puig, A Rossini, J G M Gonçalves & V Sequeira, 2004. Acquisition, modelling and rendering of very large urban environments. In: Second International Symposium on 3D Data Processing, Visualization and Transmission, Sept 6-9, 2004 (Thessaloniki, Greece) 191-198

5 Früh C, S Jain, & A Zakhor, 2005. Data processing algorithms for generating textured 3D building facade meshes from laser scans and camera Images. International Journal of Computer Vision, 61(2): 159-184

6 Früh, C & A Zakhor, 2001. 3D model generation for cities using aerial photographs and ground level laser scans. In: Computer Vision and Pattern Recognition Conference, December 2001 (Kauai, Hawaii) 2: 31-38

7 Früh C & A Zakhor, 2003. Constructing 3D City Models by Merging Ground-Based and Airborne Views. Computer Graphics and Applications, 23(6): 52-61

8 Früh C, R Sammon & A Zakhor, 2004. Automated Texture Mapping of 3D City Models With Oblique Aerial Imagery. In: Second International Symposium on 3D Data Processing, Visualization and Transmission, September 6-9, 2004 (Thessaloniki, Greece) 191-198

9 Spruyt P & S Kay, 2004. Quality assessment of digital airborne orthoimagery: A test study using the Leica Geosystems ADS40. GIM International, 18(6): 35-37

10 ITT, 2007. IDL: The Data Visualization & Analysis Platform. ITT Corporate Headquarters, 4990 Pearl East Circle, Boulder, CO 80301, USA (last accessed: October 2007)

11 Inn.Tec., 2007. JRC 3D Reconstructor. Consortium For Innovation Technology, 38, Branze, I-25123 Brescia, Italy (last accessed October 2007)