reviewed paper

Infrastructure Acquisition and 3D Virtual Integration

Gerd Hesina, Bernd Leitner, Stephan Mantler, Friedrich Brimmer

(DI Dr. Gerd Hesina, M.Sc. Bernd Leitner, DI Dr. Stephan Mantler, VRVis Research Center for Virtual Reality and Visualization, Ltd., Donau-City-Strasse 1, A-1210 Vienna, Austria, hesina@vrvis.at, step@vrvis.at, bernd@vrvis.at)
(Ing. Friedrich Brimmer, ÖBB Infrastruktur Bau AG, Zieglergasse 6, A-1070 Vienna, Austria, friedrich.brimmer@bau.oebb.at)

1 ABSTRACT

Large and widespread infrastructures today are mostly monitored from only a few centralized locations. In events of crisis or catastrophe, command and control staff members need a comprehensive and precise overview of a site, its surroundings, and any other information that may be available. This is especially important if complex situations arise and measures have to be taken on short notice in order to save lives or to avert further damage. Numerous critical decisions have to be made within the first few minutes after a catastrophic event, including resource planning, selection of access routes, and the evaluation of additional risks and hazards. Otherwise, these vitally important first minutes will be lost and never be made up for.

A virtual model of a location, which is comparable to historic military sandbox games, but enriched with additional visualizations of all visible and invisible surrounding realities (ducts, conduits, roads, population density, and so on), including interaction with the model (e.g. taking measurements), allows command and control to immediately take the right measures in order to efficiently support and prepare all units even before reaching the site of accident or catastrophe. The virtual site model therefore helps to use the decisive first minutes following an incident and take the right measures, to save lives in the best possible ways, and to avert harm from human beings as well as goods, or to at least minimize them.

In order to create such a virtual site model, it is necessary to acquire the necessary data, process it for fast access and rendering, and to provide interactive rendering methods for using the model in an emergency. This paper will showcase the creation of such a model based on very high resolution laser scan data from railroad infrastructure, on very high resolution aerial images, and on GIS data. Finally, the interactive use and rendering of such a model will be demonstrated, which serves as a basis for the visualization part of an emergency command system.

2 RELATED WORK

With recent advances in both scanning technology and processing capabilities, the data sets resulting from aerial laser scanning have been growing rapidly. Therefore, the fusion of such data with other sources for organization and handling, as well as for visual analysis, has seen increasing research (McGaughey and Carson 2003). One of the main challenges in handling laser scan data is the sheer volume of the data sets. Typical airborne LIDAR (LIght Detection and Ranging) scanners produce 103 to 104 samples per second. The final data sets often contain 108 data points or more, which is well above the capabilities of current graphics hardware.

For closed surfaces, special point based algorithms such as QSplat exist that hierarchically reduce the data for interactive frame rates while preserving the full detail (Rusinkiewicz and Levoy 2000, Pauly et al. 2006), however the generally unstructured nature of laser scans obtained in landscape surveys requires different approaches such as adaptive sampling and out-of-core strategies (Wand 2004).

Rendering polygonal terrain is a well researched area, and many approaches exist for hierarchical level-of-detail rendering of terrain data up to a global level (for example, Lindstrom et al. 1997, Lindstrom and Pascucci 2002, Cignoni et al. 2003 or Wartell et al. 2003). Many of these algorithms assume that the management of associated color or texture information is straightforward, for example using a direct mapping between geometric quadtree subdivision and hierarchical texture tiling, as used in the Google Earth and NASA World Wind products. However systems such as GoLD are capable of taking irregular texture partitioning into account during adaptive subdivision, which optimizes resource usage and thus rendering performance (Borgeat et al. 2005).

A combination of different modes of representation, such as polygons and points or line data, has been typically studied only as alternative representations within a single model. For example, levels of detail of vegetation are often based on fully polygonal data in the near field, transitioning to line and point samples for less detailed far field representations, a method first proposed by Weber and Penn in 1995. However, combining several models of varying representation can be either straightforward, since ultimately the graphics hardware rasterizes all primitives to individual pixels, or nontrivial. For example, depicting linear information such as property boundaries from GIS data sets directly on a polygonal terrain model often leads to visual artifacts and interpenetrations of the models, caused by differences in the tessellation of the objects. Therefore, line data must be adapted to precisely match the terrain representation (Wartell 2003).

3 DATA ACQUISITION AND PROCESSING

For this project, several data sources were combined to create a highly detailed representation. The basic components were high-resolution laser scans and digital elevation models and aerial orthophotography; these were then merged with existing GIS information such as land classification boundaries and infrastructure metadata. Figure 1 illustrates the processing and merging of the individual sources.

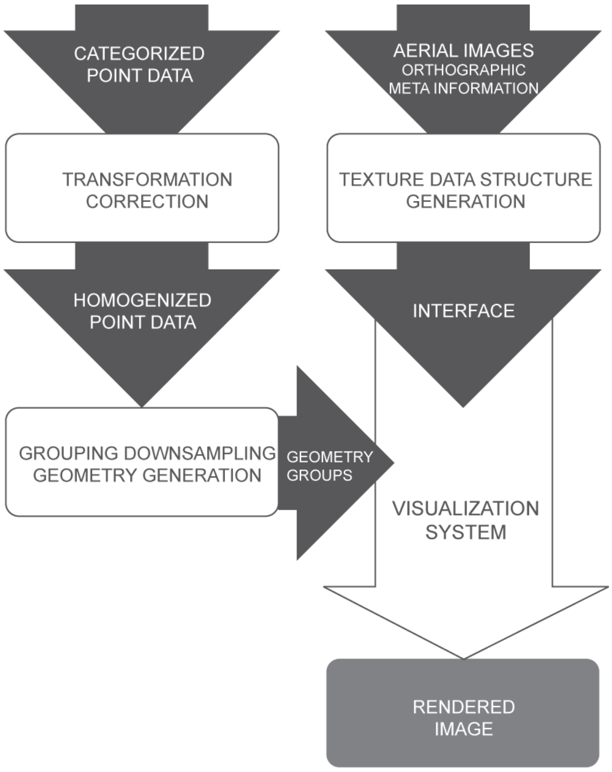

Figure 1: System Architecture Overview

3.1 Data Acquisition

Laser scan data was acquired from a low flying helicopter (200m altitude) at 45kt speed. The measurement equipment consisted of a RIEGL Full-waveform-processing Airborne Laserscanner LMS-Q560, an Inertial Measurement Unit IMAR INAV_FJR-001, a GPS/GLONASS receiver Topcon Legacy-E and a calibrated CCD camera to synchronously acquire images with a resolution of 22 Mpixels. The resulting density was approximately 20 points per square meter.

In addition to these images, further photographic data was produced using aerial orthophotography. Using a Vexcel UltraCamX, georeferenced aerial images with a ground level resolution of few centimeters were obtained, processed and merged with the laser scan images.

The digital elevation and object surface models used for further processing had a minimal resolution of 50cm, and additional orthophotography at 5cm resolution was also available. In total, the acquired laser scan data consisted of roughly 600 million point samples, and 40GB of raster images were produced with image dimensions between 1900x5000 and 12000x14000 pixels.

3.2 Processing

After acquisition, some processing was required to combine the various data sources with high quality. This step also included optimizations to allow rendering of the final model in real time.

Although all acquired data was georeferenced, the very high resolution of the individual parts exacerbated minor registration errors and variations caused by using varying geospatial reference systems. To compensate for these effects, the individual components were therefore manually fine tuned to achieve optimal matching. Alongside these manual adjustments, an automatic preprocessing pipeline was implemented to streamline the preparation of data for the rendering system, and for the extraction of laser scan attributes.

One important detail was to shift the acquired laser scan points from a global reference frame to the coordinate system origin used in the rendering system. In very large scenes, the extents of the terrain to be rendered exceeds the available precision, and direct rendering would lead to visualization errors in the rendered scene. The reason for this lies in the architecture of current graphics hardware. Most graphics adapters available today internally process data with 32-bit float precision, which is sufficient in most cases. However, when covering very large areas or handling data with geospatial coordinates, the values cannot be handled with the precision required for artifact free rendering; using a local coordinate frame avoids these issues at the expense of a static translation and some overhead when mapping between graphics and geospatial coordinate values.

Laser scan data inherently consists only of location information, but does not have intrinsic color. However, the data points were classified by object type, for example ground / vegetation, buildings, and overhead contact lines (see Figure 2). However, for visualization purposes color information from the associated location of the orthophotography was mapped to each laser scan point (see Figure 3). In the final rendering, ground data is displayed directly from orthophotographic and digital elevation data, so points classified as ground were also removed in this preprocessing step.

The mesh used for rendering the ground level terrain was created from several data sources: the laser scan data classified as ground level, and a digital elevation model represented as a regular grid of height samples. Due to the low variance of ground level data, the data set was intensively decimated, and a low polygon mesh was created using 2D Delaunay triangulation. Even though using only a fraction of the original point set for creating the ground mesh introduces minor changes to the scanned representation, the decimation algorithm guaranteed that the differences remained negligible.

In contrast to the ground mesh visualization, points classified as vegetation or buildings are rendered as point clouds. In order to ensure high rendering performance, users can select a detail level for a given data set which controls the point cloud density, i.e. the fidelity of the representation of vegetation or buildings.

After the input data has passed all stages of the preprocessing pipeline, the results are written to cache files on disk. This way a given data set has to pass through the pipeline just once and can then be streamed from the filesystem without any preprocessing. This ensures very fast application startup and model switching, which is highly desirable in emergency response scenarios.

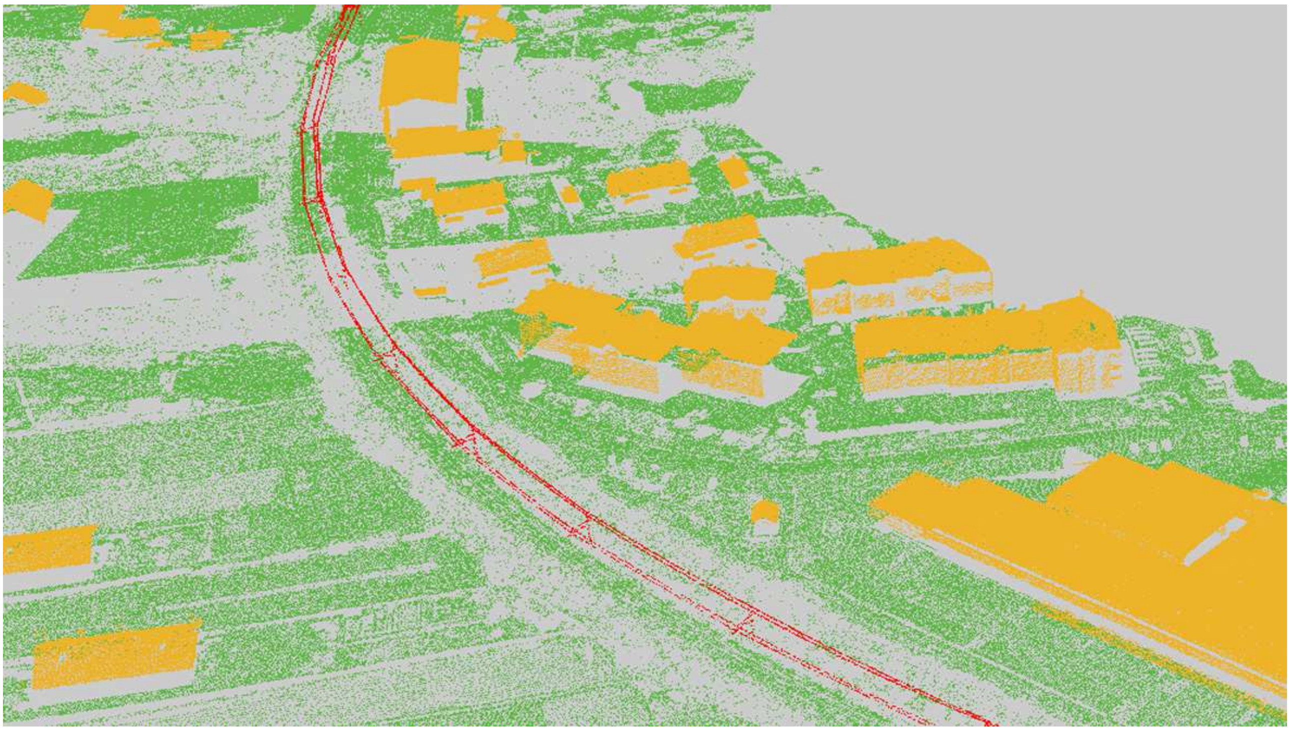

Figure 2: Laser scan data classified by object type green: ground and vegetation; yellow: buildings; red: overhead contact lines

Figure 3: Laser scan data colored by orthophotography.

4 RENDERING

The rendering application was developed as a prototype to support the visualization of these combined data sets at very high performance. Photographic data is mapped onto the terrain model and rendered as a closed surface; laser scan data is directly rendered as point data and the GIS meta-information is displayed as colored lines (actually rendered as thin polygonal slabs due to the limitations of DirectX). The point size of the laser scan data is automatically changed as the viewer navigates through the scene, but can also be interactively adjusted to achieve either a dense, more realistic impression or a sparse rendering that allows the viewer to see the covered metadata.

Aerial images were used as textures for the ground mesh and for coloring the point samples. Besides the pure image data, each aerial image also contains additional registration data, such as geospatial image coordinates, actual pixel size, skew and rotation. A common problem when using orthographic photos for rendering were varying pixel dimensions and of course memory - high resolution images typically amount to hundreds of megabytes, making them difficult to handle in a realtime rendering application.

To address these problems, we exploit the meta information contained within the aerial images. An intermediate data structure maps the set of aerial images to a single coherent texture, also considering potential rotations that may have occurred during acquisition. The data structure serves as a facade between the rendering system and an arbitrary number of textures in any resolution and size. It provides a convenient interface for resampling and extracting arbitrary sub-images, which makes it easy to generate homogenous tiles of the input images. Furthermore, the actual image data is loaded from disk dynamically as required for the extraction process, reducing memory consumption and improved performance.

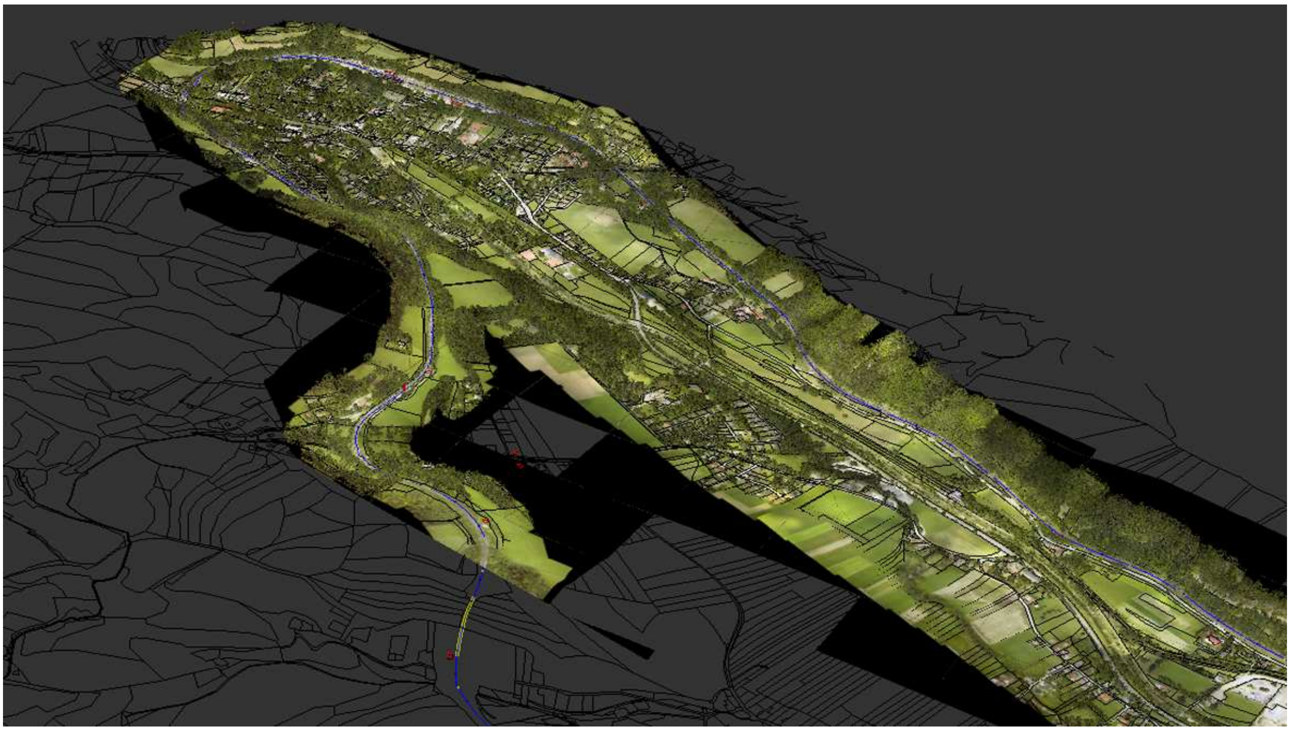

Figure 4: Combined display of orthophotography mapped onto a digital elevation model, laser scan data and vector based GIS metadata.

Figure 5: Close-up view. In addition to the terrain and vegetation data, GIS meta-information depicts property boundaries and railway chainage labels (in kilometers).

An example result of the preprocessing pipeline can be seen in Figures 4 and 5. The input data consisted of 307.170.000 categorized point samples, and 23GB of georeferenced aerial images with individual resolutions of 5000x5000 pixels. The preprocessing, data reduction and terrain mesh generation steps took 15 minutes (filtering and grouping 8 minutes, terrain triangulation and texture generation 7 minutes), resulting in 1.1 million point samples and a terrain mesh of 78662 triangles, using 29387 vertices; this data is rendered with approximately 30FPS on a 2.8GHz Intel QuadCore CPU with 8GB RAM and NVidia GeForce GTX 280 graphics.

5 CONCLUSION AND FUTURE WORK

We have demonstrated a preprocessing and rendering framework for the interactive visualization of combined geospatial polygonal, point and line data. The current rendering application can be seen as a proof of concept implementation that demonstrates the feasibility of a semi-automated process to preprocess and render a multitude of highly detailed geospatial data sources.

Currently, the system depends on the user to make certain decisions for optimal rendering performance (for example, level of detail or point cloud density). This mode of operation provides a straightforward way to tweak the application for delivering the best possible performance on particular hardware very quickly. However, at the same time this approach also offers the most important aspect for future improvement. An automated, view dependent approach to level of detail and point cloud thinning would deliver near optimal performance in a wider range of usage scenarios, perhaps also dynamically adapting from a lower level of detail for interactive navigation to a highly detailed still image.

Also, further research will focus on the integration of interaction and visual query techniques. It is envisioned that the framework will ultimately support direct communication to geospatial databases for real-time queries and updates, a feature that would be highly useful for emergency response and incident management applications.

6 ACKNOWLEDGEMENTS

This work was part of the VEI-3D research project, which is supported by the Austrian Research Promotion Agency FFG as part of the KIRAS suport programme for security research under contract nr. 813769.

7 REFERENCES

BORGEAT, L., GODIN, G., BLAIS, F., MASSICOTTE, P. and LAHANIER, C.: GoLD: Interactive Display of Huge Colored and Textured Models. In ACM Transactions on Graphics (Proceedings of SIGGRAPH 2005), Vol. 24, Issue 3, pp. 869-877. New York, 2005.

BOUBEKEUR, T., DUGUET, F. and SCHLICK, C.: Rapid Visualization of Large Point-Based Surfaces. In: Proceedings of the 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST), pp. 75-82. Pisa, 2005.

CIGNONI, P., GANOVELLI, F., GOBBETTI, E., MARTON, F., PONCHIO, F. and SCOPIGNO, R.: BDAM - Batched Dynamic Adaptive Meshes for High Performance Terrain Visualization. In: Computer Graphics Forum, Vol. 22, Issue 3, pp. 505-514. Granada, 2003.

LINDSTROM, P., KOLLER, D., RIBARSKY, W., HODGES, L.F.: An Integrated Global GIS and Visual Simulation System. GVU Technical Report GIT-GVU-97-07, Graphics, Visualization and Usability Center, Georgia Institute of Technology, 1997.

LINDSTROM, P. and PASCUCCI, V.: Terrain Simplification Simplified: A General Framework for View-Dependant Out-of-Core Visualization. In: IEEE Transactions on Visualization and Computer Graphics, Vol. 8, Issue 3, pp. 239-254. 2002.

MCGAUGHEY, R.K. and CARSON, W.W.: Fusing LIDAR Data, Photographs, and other Data Using 2d and 3d Visualization Techniques. In: Proceedings of Terrain Data: Applications and Visualization - Making the Connection, pp. 16-24. Charleston, SC, 2003.

PAULY, M., KOBBELT, L.P. and GROSS, M.: Point-Based Multiscale Surface Representation. In: ACM Transactions on Graphics, Vol. 25, Issue 2, pp. 177-193. New York, 2006.

RUSINKIEWICZ, S. and LEVOY, M.: QSplat: A Multiresolution Point Rendering System for Large Meshes. In: Proceedings of the 27th annual conference on computer graphics and interactive techniques (SIGGRAPH), pp. 343-342. New York, 2000.

WAND, M.: Point-Based Multi-Resolution Rendering (Dissertation). Tübingen, 2004.

WARTELL, Z., KANG, E., WASILEWSKI, T., RIBARSKY, W. and FAUST, N.: Rendering Vector Data over Global, Multi-Resolution 3D Terrain. In: Proceedings of the Joint EUROGRAPHICS-IEEE TVCG Symposium on Visualization, pp. 213-222. Grenoble, 2003.

WEBER, J. and PENN, J.: Creation and rendering of realistic trees. In Computer Graphics (SIGGRAPH '95 Proceedings), Pages 119-128. New York, 1995.