ADAPTIVE LIDAR SCANNING OF HISTORIC BUILDINGS SUPPORTED BY AUGMENTED REALITY USER INTERFACES

V. Paelke [a], S. Filin [b], D. Eggert [c], S. Barnea [d]
[a] IG – Institut de Geomàtica, Casteldefells (Barcelona) – Spain - volker.paelke@ideg.es
[b] Technion – Israel Institute of Technology – Haifa – Israel - filin@technion.ac.il
[c] IKG - Leibniz Universitaet Hannover – Germany – daniel.eggert@ikg.uni-hannover.de
[b] Technion – Israel Institute of Technology – Haifa – Israel - barneas@technion.ac.il

KEY WORDS: 3D Modelling, Augmented Reality, LIDAR, Visualization, User Interaction

ABSTRACT:

3D models of buildings and spatial environments are rapidly becoming a standard tool in the documentation, management and presentation of architectural, archaeological and landscape cultural heritage. Despite advances in acquisition technologies like photogrammetry and LIDAR scanning, the costs for data collection are still a significant and often limiting factor, especially for large scale models. In the research project reported here we develop and evaluate techniques for the acquisition of large models driven by relevance, using an adaptive scanning approach that adjusts the resolution in which a model is acquired to its relevance. New processing techniques and a user-interface that integrates a fast augmented reality visualization of the current model state with quality and resolution metrics computed on the fly enable operators to interactively control the acquisition process and conduct quality control on-site.

1. INTRODUCTION

Despite technological advances the cost of creating and maintaining the 3D models required for the documentation, management and presentation of architectural, archaeological and landscape cultural heritage remains a challenging issue.
One central problem in applications using 3D models is the trade-off between detail and acquisition cost (during acquisition) as well as processing speed (during use). Much detail (even on a small scale) requires the complete scan to be conducted at high resolution and leads to long acquisition time, large amounts of data, and complex processing. Rapid scanning - in contrast - will be faster but only provides lower resolution and an overall coarse model.

In our approach we explore an alternative way to provide 3D information on a large scale, applying the concept of generalization - in the cartographic sense of meaningful abstraction - to 3D data derived from LIDAR scans, starting at the acquisition phase. In many use-cases a high-amount of detail is only required for those objects that are of high relevance to the user while others are only relevant as context. E.g. in a model of historic architecture a highly precise model of the building surface may be desirable but the same amount of detail is neither required nor desired for the leaves of plants in the environment.
An effective user interface is required to control the acquisition process to enable this selective acquisition of large-scale geospatial models where the amount of detail varies and is driven by relevance. An augmented reality (AR) user interface that overlays the current 3D model on a view of the real world environment provides an intuitive way to check the current state of the acquisition process. By simple touch interaction in the AR presentation operators can select objects of special interest and adjust the acquisition resolution on the fly. In addition quality metrics can be calculated and displayed on the model to enable quantitative checking of the acquired data on site.
The implementation of such a system in turn requires fast and reliable matching and processing techniques that operate at high speeds and can be used in the field.
As part of the ongoing research project it is planned to add facilities for semantic modelling and to address relevance driven presentation / visualization aspects at a later stage.

2. RELATED WORK

A growing number of applications and domains make use of laser-scanning technology as a means for modeling 3D cultural heritage (Barber et al., 2005; Stenberg, 2006; Visintini et al., 2006), and architectural modeling (Levoy et al., 2000; Akca et al., 2006). Laser scanning based documentation benefits from a very detailed depiction of complex objects that could not have been documented otherwise, and from a millimeter level of accuracy, which enables an accurate reconstruction. Nonetheless, the surveying process and later on the intensive geometric modelling incurs high costs that make such documentation expensive, and thus impractical in many cases.
In order to improve terrestrial laser scanning processing, research into autonomous and computationally efficient procedures has seen growing interest in recent year. The focus was mainly on autonomous registration of scans and on geometric primitive extraction. Registration of individual laser scans, aerial and terrestrial, into a common reference frame is a fundamental step in aligning all data into a common reference frame, thereby forming a seamless dataset. For terrestrial laser scans, commercial registration software does not provide a complete autonomous solution unless some specially constructed (and usually expensive) targets are deployed in the scanned area. Focusing on pairwise registration of terrestrial scans, a keypoint registration scheme has been developed and supplemented later on by the RGB image content (Barnea and Filin, 2007; 2008). Later on, inclusion of intensity data has show significant improvement in accuracy and efficiency. In reference to object modelling and identification, the challenges

to be addressed include object shape complexity as well as variations in depth within the scene (typical scans will cover hundreds of meter) and consequently in scale and resolution, even within a single scan. Zeibak and Filin (2007; 9), Barnea et al. (2007) and Gorte (2007) propose processing of range panoramas, reflection of the acquisition process, instead of processing of the actual 3-D point cloud. As the authors show, effective means to isolate objects, segment the data, compare scans, and recognize objects by non-parametric models are made possible this way.

The software currently used in the 3D data acquisition and modelling process is designed for operation in in-door office environments. A promising approach to improve the usability on-site is the use of the user interface concepts of mixed and augmented reality (Milgram et al. 1994, Azuma, 1997) that integrate the real environment into the user interface. Early work in this direction includes the AR outdoor modelling application by Piekarski and Thomas (2001). New hardware like smartphones makes mobile AR applications for outdoor use increasingly viable and the development of specialized devices like the GeoScope AR input/output device (Paelke and Brenner, 2007) allows to address problems in positioning and displaying for outdoor use-cases like data acquisition where high mobility is not required.

3. WORKFLOW AND SYSTEM STRUCTURE

The objective of our approach is the effective creation of 3D geospatial models required for the documentation, management and presentation of areas and buildings based on integrating global data (airborne laser or alternative sources) with local detail acquired using terrestrial laser scans.
The central idea is to control and limit the amount of detail in all processing stages to what is actually required while providing immediate feedback and on-site interaction capabilities. This reduces acquisition time. Modelling time can also be reduced in many cases because modelling problems can be resolved immediately with the original surfaces as the reference. An additional benefit is the reduction in the storage and computation requirements when using the resulting models.

Figure 1 shows the demand-driven workflow into which the acquisition, analysis, integration and presentation activities are embedded.

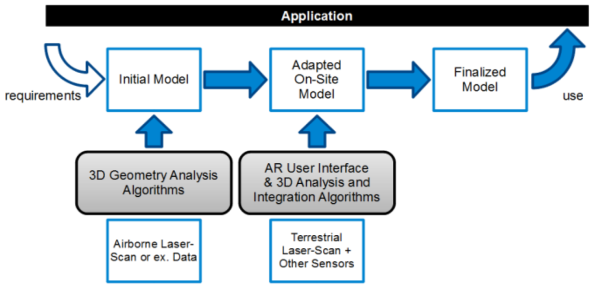

Figure 1: Acquisition and analysis workflow

In many cultural heritage applications initial information on the spatial context of the object/area of interest is already available in different forms, e.g. 2D map data or digital terrain models.
For the purpose of modelling we start with this (integrated) information as an initial model. It can be used to support user orientation during the acquisition process and serves as the spatial context in later use. Within this initial model areas of interest can be indentified in which more detailed models are required. Alternatively, an initial 3D model can be acquired by a cost effective method, e.g. through airborne laser scanning and photogrammetry.
To refine the areas of interest we employ on-site modelling using terrestrial laser scanning: The user interface is based on the paradigm of augmented reality (AR). In the visualization of the modelling interface information on the model (geometry, resolution and quality) is overlaid on a view of the real environment to be acquired. This spatially registered overlay of the 3D computer graphics rendering of the model allows controlling the acquisition and modelling process in a very intuitive way. The operator can directly see the current state of the model, has a straight-forward reference to the real environment and can therefore easily decide and select features for further detailing.
To make this information available on-site a rapid processing of the data is necessary to ensure that the information presented to the user incorporates all information acquired so far. We therefore develop a number of 3D geometry analysis and integration algorithms that match and integrate data from different scans and data sources and provide measures of model quality.
The resulting model can be further refined or extended with additional (e.g. non geometric) information either on-site or back in the office using conventional modelling tools.

4. COMPONENTS

4.1 Sensors and Data

Laser scanning technology makes direct acquisition of 3D data feasible, enabling coverage of wide regions constantly generating in recent years decimeter level point densities and sub-centimeter level of accuracy. Thus it provides a detailed object description in its actual 3D sense. Accompanying the range data are cameras that acquire high-resolution and texture rich content of objects, enabling to provide nearly photorealistic depiction of the studied scene with little processing. For the tests we use a Riegl LMS Z360I Scanner in combination with a Nikon SLR for texture capturing (Figure 2).


Figure 2: Nikon SLR camera and Riegl LMS Z360I scanner assembly

4.2 Data Processing:

The scanner and mounted camera, feature two reference frames which are co-aligned by a boresight transformation. The camera-scanner boresight relation can be encapsulated by a 3×4 projection matrix **P** which represents the relation between an object space point and an image point:

$$x = \mathbf{KR}\left(\mathbf{I} \mid -t\right)X \qquad (1)$$

where $X=[x\ y\ z\ 1]^T$ and $x=[u\ v\ 1]^T$, are object- and image-space points, respectively, in homogeneous representation; **K** the calibration matrix**,** **I** the identity matrix, and **R** and t, the rotation matrix and translation vector, respectively. Radial and decentering lens distortions are calibrated and corrected for.

For each scan, n images are acquired at predefined "stops" (every $360/n$ degrees). Assuming that, *i*) the camera is rigidly mounted to the scanner, *ii*) the intrinsic camera parameter are fixed and calibrated in advance, and *iii*) the acquisition position (of the "stop") is fixed across all scanning positions, enable using the same projection matrices for all images of the same "stop" within different scans.

The scanned data (ranging and intensity), is represented in polar coordinates

$$(\rho\cos\varphi\cos\theta, \rho\cos\varphi\sin\theta, \rho\sin\varphi)^T = (x, y, z)^T \quad (2)$$

with φ and θ latitudinal and longitudinal coordinates of the firing direction and ρ the measured range (Figure 3). Polar coordinates offer lossless raster data representation as the angular spacing is fixed. Range and intensity values set pixels content.

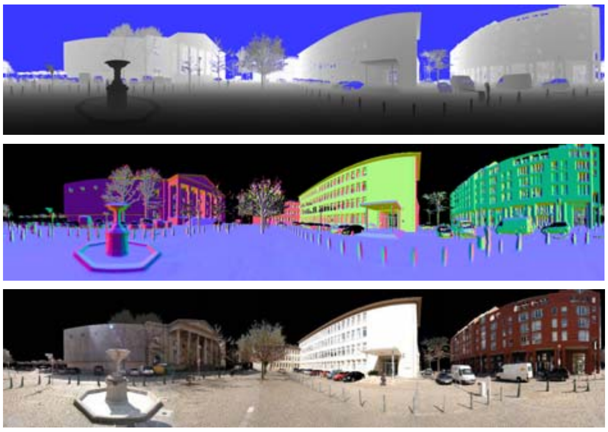

Figure 3: Polar representation of the segmentation channels. The horizontal and vertical axes of the images represent the values of φ, θ respectively. (top) distances as intensity values ρ (bright=far), with "no-return" pixels in blue, (middle) surface normals, (bottom) color (see text).

For data segmentation, we use the mean-shift segmentation (Comaniciu and Meer, 2002), a scheme that was chosen due to its successful results with complex and cluttered images.

As a non-parametric segmentation model, it requires neither model parameters nor domain knowledge as inputs. The algorithm is controlled by only two dominant parameters: the sizes of spatial and the range dimensions of the kernel. The first affects the spatial neighborhood while the latter affects the permissible variability within the neighborhood. These two parameters are physical in a sense. The mean shift segmentation is applied on the: range data in its panoramic forms, surface normals data which are computed from the range panoramas, and color content, which is derived from the acquired set of images. The range channel enables highlighting vertically dominant objects, like tree stems or poles, while the normal based segmentation reveals the ground, façades and other surface objects that appear as complete segments. Color content enables isolating objects which are consistent in their hue.

The integration scheme originates from the realization that the different channels exhibit different properties of the data. Consequently, they provide "good" segments in some parts of the data and "noisy" ones in other parts. Therefore each channel is segmented independently and then a segmentation that integrates them is constructed by selecting the better segments from each channel. In this scheme the addition of other channels can be accommodated without many modifications. The objective is to obtain segments that are uniform in their measured property, where optimally, all data units belonging to the segment will have similar attributes. Additionally, we aim for segments that are spatially significant and meaningful. As such, we wish to assemble large group of data units, preferably of significant size in object space. These segments should not lead however to under-segmentation. In order to meet the need for significant grouping in object space, we set the score of a segment with respect to its 3D coverage. The proposed model is applied as follows. First, the largest segment is selected from all channels, if the segment quality is satisfactory it is inserted into the integrated segmentation. All pixels relating to this segment are then subtracted from all channels and the isolated regions in the other channels are then regrouped and their attribute value is computed. Following, is the extraction of the next largest segment and the repetition of the process until reaching a segment whose size is smaller than a prescribed value and/or preset number of iterations. We note that due to the non-parametric nature of the mean-shift segmentation, re-segmenting the data between iterations has little effect. Figure 4 shows the segmentation results for the scene in Figure 3.

Figure 4: Results of the segmentation

4.3 AR UI

For the implementation of the user interface for on-site interaction we address two device types.

The first category is smartphones using the android operating systems. They provide coarse localization and orientation through the integrated GPS and digital compass. The user interface of our application is realized as video-see-through augmented reality, using the video-stream provided by the camera on the back of the phone and augmenting it with computer graphics of the model and additional information.

Smartphones are highly mobile and relatively inexpensive – within the acquisition system they are best used to interactively mark areas of specific interest and for initial quality control (coverage checking) of acquired data. E.g. (multiple) domain experts on a cultural heritage site can use smartphones to mark areas in the initial coarse model that must be acquired in higher resolution and later check that the acquired data covers all intended areas. We currently use HTC Hero and HTC Desire smartphones with the Android operating system (HTC, 2010).

Figure 5: AR UI setup using HTC Desire

Unfortunately, the resolution and accuracy of the GPS position and compass orientation in smartphones is limited and can be insufficient for detailed modelling. We therefore employ a second device when a higher resolution is required. The GeoScope (Paelke and Brenner, 2007) is an augmented reality device that can be installed on a standard geodetic tripod at arbitrary locations. Its main components are a high resolution LCD display with a touch-screen that faces the user and a high resolution camera that is mounted on the back, looking into the environment. Similar to a telescope the GeoScope can be turned in two degrees of freedom (pitch and yaw). The rotation angles are captured with high resolution and precision by mechanical sensors. In combination with the position of the tripod (which can be determined precisely by geodetic means) all position parameters of the GeoScope can be determined with high precision, allowing for spatially correct augmentation of the video images in a similar setup to the smartphones, but with higher resolution and precision. When acquiring data with a laser scanner the same tripod can be used both for the scanner and the GeoScope, simplifying logistics. The precise visual overlay of the acquired model on the real environment can be used for detailed quality checking, refined modelling and semantic annotation.

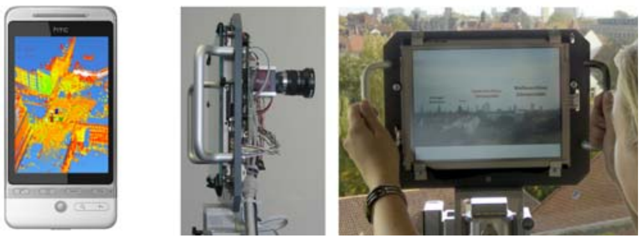

Figure 6: AR devices used: HTC Hero Smartphone (left), GeoScope (right)

4.3.1 Visualization

For the visualization of the acquired data we use a point based rendering (PBR) approach that can operate on the point data with minimal preprocessing. While many tools that operate on 3D models require a polygonal reconstruction of the surface, a visual depiction can be rendered directly from point data (Levoy et al., 2000).

For the acquisition application this has two benefits:

First, it is not necessary to construction (good) polygonal models before the information can be displayed – this is very desirable for models where (during the acquisition process) some parts maybe highly detailed and others only coarsely represented. It also results in a much simpler and faster way to present the data after acquisition. Second, the direct depiction of the data provides an intuitive presentation of the data density – to exploit this we have adapted existing point-based rendering techniques to be able to display either continuous surfaces (as is usually the goal in PBR techniques) or distinguishable point-sets. As the rendering of the point based data is very simple (each point is rendered as a splat) it can be easily implemented on a Smartphones and similar devices, especially if the data handling and preprocessing is handled on a more capable PC that is networked to the Smartphone. The effect of direct geometry rendering and the possibility to integrate secondary information like quality data can be seen in the Welfenschloss Example in Section 5.

4.3.2 Rapid Calculation of Quality Metrics

In order to determine the quality of the acquired data geometry analysis algorithms must be applied. Given the large data-sets this processing usually does not work at interactive speeds. To provide this information rapidly in the on-site acquisition context we have conducted experiments with the use of modern graphics processors (GPUs) that provide highly parallel data processing capabilities. As a simple example that provides a useful quality metric we discuss the calculation of k nearest neighbors (kNN): KNN is useful in many aspects, e.g. algorithms like triangle-mesh reconstruction use it for surface reconstruction. However, in the acquisition context our main use is the use of the density value in each point as a quality indicator.

The density d_P of point P is the inverted sum of the distances between P and the k nearest neighbors P_i of P.

$$d_P = \frac{k}{\sum_{i=0}^{k} |PP_i|}$$

The association of colors to the minimum and maximum density value (e.g. max density = green, min density = red) allows to visualize the calculated density values (see Figure 12).

We have implemented several algorithms for point set analysis on the GPU, using OpenCL to achieve platform independence. In a test setup on a PC with Intel Core 2 Duo E8400 with 3.0 GHz, 8 GB of dual channel DDR2 RAM and a Nvidia Geforce GTX 285 with 240 stream processors we compared performance between the GPU implementation using OpenCL and a C++ implementation on the CPU. As shown in Figure 7 a simple brute-force implementation of kNN the GPU calculation was 45 times faster, making the use of this quality measure on-site viable. More complex algorithms (e.g. partioned kNN) are possible and result in further speedup, but require more detailed tailoring to the parallel structure GPU to realize the acceleration potential. For a detailed discussion see (Eggert and Paelke, 2010).

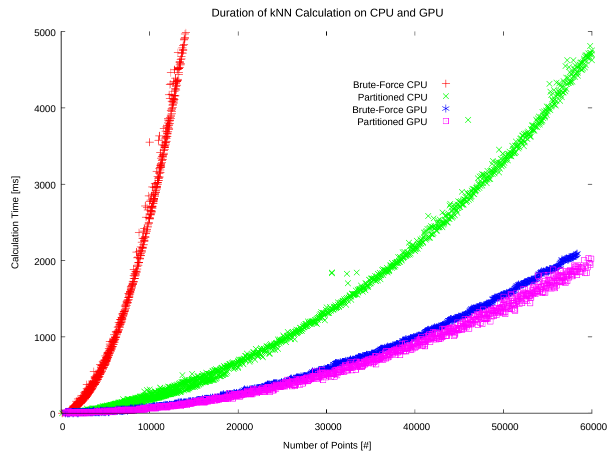

Figure 7: Runtime comparison between CPU- and GPGPU-based kNN implementations

5. EXAMPLE

The Welfenschloss (Figure 8) is a former castle in the northern part of the city of Hannover. It was planned by the architect Christian Heinrich Tramm and built between 1857 and 1866. Since the kingdom of Hannover ceased to exist as an independent entity when it was annexed by Prussia in 1866, the Welfenschloss was never used as a castle. In 1879 it became the main building of the University of Hannover (Pietsch, 2003).

Figure 8: the Welfenschloss (Photo: Andree Stephan)

In the following example we use the Welfenschloss as a test-case for demonstrating our methods.

Figure 9 shows an initial model of the Welfenschloss – in this case data acquired from airborne laserscanning. The data is sufficient to enable a rough orientation and thus initial mark-up of areas of interest, but lacks all the detail required for detailed documentation or presentation. Information of this kind is often available from airborne laserscanning, airborne photogrammetry or extruded cadastral data.

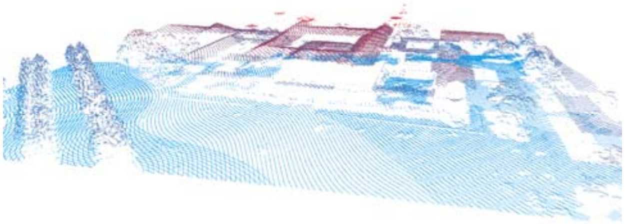

Figure 9: Initial model from airborne laserscan with color indicating height

To generate models that are suitable for documentation, analysis or presentation additional detail must be acquired and integrated into a coherent (geometric) model. We use a terrestrial laser scanner (Scanner Riegl LMS Z360I) to acquire this detail information on the geometry.

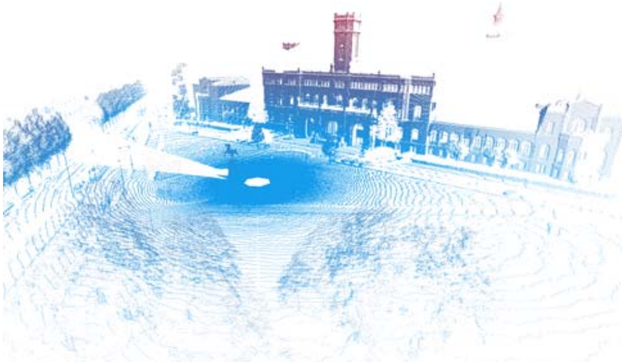

Figure 10: Terrestrial laserscan with color indicating height

Figure 11: Terrestrial laserscan with texture colors

Figure 12: Terrestrial laserscan using splat-based PBR with texture colors

To check the data acquired on site both for coverage (completeness) and quality a rapid visualization is required. As described in section 4 we employ a point based rendering (PBR) approach that generates images like Figure 9-13 for display on smartphones, the GeoScope or Laptops. In the PB rendering of the pure geometry operators can get an intuitive overview of the coverage of the data acquired at a certain stage. Combing the rendering with the mark-up of areas of interest previously assigned by a content expert in the initial model allows an intuitive check for completeness. However, it is difficult to check the quality of the acquired data from a pure geometry depiction. Therefore, quality information like density is derived from the data and this can be integrated into the rendering (Figure 13).

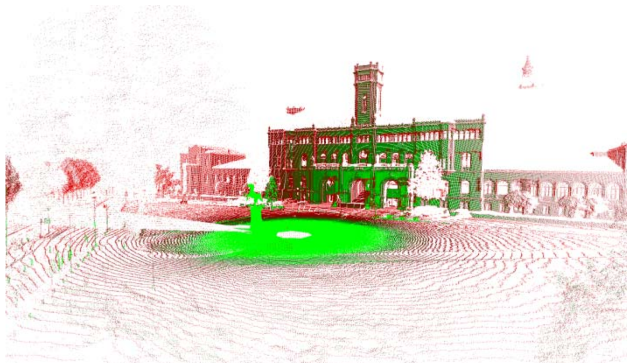

Figure 13: Visualization of scan density.

Using this presentation a rapid check for completeness and quality with regards to the area of interest becomes possible.

6. OUTLOOK AND FUTURE WORK

Most heritage monuments are complex in their shape and span over wide areas. Therefore they require complex acquisition setups and very elaborate processing. Both consequences translate into high cost and reduced accessibility to laser scanning technology driven modelling. In this regards, this paper proposed efficient acquisition of data, efficient processing of the laser point cloud, and easy do display and interact visualization tools. The main focus here was on the acquisition process and primitive extraction. In this regards, it offered utilizing the range panorama concept as a means for both data integration and efficient feature extraction. The paper has proven that the basic technologies required for implementing a system for geometry acquisition with terrestrial LIDAR that adapts the resolution to the requirements using an AR based user interface. Using an iterative approach the individual components are currently refined, and integrated to develop additional and more complex algorithms for data analysis, fusion based on segmentation and presentation. In order to face the demands of on-site quality evaluation we employed GPGPU-based analysis-algorithm implementations.

Extension of this research will explore three areas: first, studying the effect of the new acquisition tools, gathering user feedback on the functionality of an AR user interface and the suitability of different devices to support them, especially comparing the use of smartphones and the GeoScope. Secondly, extending the functionality of the acquisition process and on-site modelling, incorporating the results from the user study and additional requirement from real world users. Specifically, we want to integrate enhanced quality measure requiring more complex analysis algorithms, advanced matching algorithms for multiple laser-scans that are usable “on-the-fly” and additional modelling functionality to support semantic annotation. The third research thread will focus on the use of the data gathered, addressing the presentation of the acquired models and interaction with them from a user perspective.

7. ACKNOWLEDGEMENT

This joint research project was financially supported by the State of Lower-Saxony and the Volkswagen Foundation, Hannover, Germany.

8. REFERENCES

Akca, D., Gruen, A., Alkis, Z., Demir, N., Breuckmann, B., Erduyan, I., Nadir, E., 2006. 3D modeling of the weary herakles statue with a coded structured light system. International Archives of Photogrammetry and Remote Sensing. XXXVI(5).

Azuma, R., 1997: “A Survey of Augmented Reality” in: Presence: Teleoperators and Virtual Environments, Vol. 6, No. 4, 1997.

Barber, D. M., Mills, J. P., Bryan, P. G., 2005. Maintaining momentum in terrestrial laser scanning: A UK case study. Archives of ISPRS WG III/3, III/4, V/3 Workshop "Laser scanning 2005". September 12-14, 2005, Enschede, the Netherlands.

Barnea, S., Filin, S., 2007. Registration of terrestrial laser scans via image based features. ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007. September 12-14, 2007, Espoo, Finland.

Barnea, S., Filin, S., 2008. Keypoint based autonomous registration of terrestrial laser point-clouds. ISPRS Journal of Photogrammetry & Remote Sensing. 63: 19-35.

Eggert, D., Paelke, V., 2010 : Relevance-driven Acquisition and Rapid On-site Analysis Of 3D Geospatial Data, Proc. Joint International Conference on Theory, Data Handling and Modelling in GeoSpatial Information Science , vol. 38 , no. 2 , p. 118-123 , 2010.

Gorte B. (2007). Planar feature extraction in terrestrial laser scans using gradient based range image segmentation, ISPRS Workshop on Laser Scanning, pp. 173-177.

HTC, 2010: HTC Hero Smartphone product webpage: http://www.htc.com/www/product/hero/overview.html; accessed 4.6.2010

Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S. E., Davis, J., Ginsberg, J., Shade, J., Fulk, D., 2000. The digital Michelangelo project: 3D scanning of large statues. In Proceedings of the 27th Annual Conference on Computer Graphics (SIGGRAPH), New Orleans, Louisiana. pp. 131-144.

Milgram, P. Takemura, H, Utsumi A. & Kishino, F., 1994: "Augmented Reality: A class of displays on the reality-virtuality continuum". in: SPIE Vol. 2351-34, Telemanipulator and Telepresence Technologies, 1994.

Paelke, V., Brenner, C., 2007: “Development of a Mixed Reality Device for Interactive On-Site Geo-visualization“ in: Proc. Simulation und Visualisierung 2007, Magdeburg, März 2007.

Piekarski, W. and Thomas, B. H., 2001: "Tinmith-Metro: New Outdoor Techniques for Creating City Models with an Augmented Reality Wearable Computer”, in: 5th Int'l Symposium on Wearable Computers, Zurich, Switzerland, Oct 2001.

Pietsch, W, 2003: „Vom Welfenschloss zum Campus Center“ in: Die Universität Hannover. Ihre Bauten, ihre Gärten, ihre Planungsgeschichte. Auffarth, S., Pietsch, W. (Eds.), Imhof 2003.

Sternberg, H. 2006. Deformation measurements at historical buildings with terrestrial laser scanners. International Archives of Photogrammetry and Remote Sensing. XXXVI(5).

Visintini, D., Crosilla, F., Sepic, F., 2006. Laser scanning survey of the Aquileia basilica (Italy) and automatic modeling of the volumetric primitives. International Archives of Photogrammetry and Remote Sensing. XXXVI(5).

Zeibak, R., Filin, S., 2007. Change detection via terrestrial laser scanning. International Archives of Photogrammetry and Remote Sensing. 36(3/W52): 430-435.

Zeibak, R., Filin, S., 2009. Object Extraction from Terrestrial Laser Scanning Data. in Proceedings of FIG working week 2009, Eilat, Israel, published on CD-ROM