FACADE DETAIL FROM INCOMPLETE RANGE DATA

Jan Böhm

Institute for Photogrammetry, Universität Stuttgart, Germany
jan.boehm@ifp.uni-stuttgart.de

Commission V, WG V/4

KEY WORDS: Range Data, LIDAR, Building, Façade, Modelling, Matching

ABSTRACT:

In recent years, the demand for highly detailed building models has clearly risen. Most of the details, which characterize a building, come from its façade. It is obvious that such façade detail is ideally acquired from ground-based sensors. For any street-level data acquisition system, a common problem arises: that of occlusions and hence incomplete data. In the past we have shown how multi-image coverage can overcome occlusions in image based façade recording. In this paper, we demonstrate new approaches on façade detail from range data in the case of incomplete data. The approach builds on our development of LASERMAPs, a simple and efficient way to use street-level LiDAR data to enhance existing prismatic building models.

1. INTRODUCTION

The impressive advances in the automatic generation of virtual city models achieved within the photogrammetric community in recent years, have contributed to the widespread dissemination of such models to various end-user platforms. So called 'digital globes' have become an especially popular platform for LOD1 (level-of-detail 1) models, which feature simple block models without roof structures. Lately LOD2 models have been released, which feature detailed roof structures, while the body remains a simple prism enriched by colour texture. See (Kolbe & Gröger, 2003) for a detailed definition of the LOD hierarchy.

However, the use of highly sophisticated computer equipment and enormous manpower in the modern media and entertainment industry, which can be experienced in stunning computer graphics special effects in movies or highly detailed three-dimensional video games, has further raised the expectations of today's audiences to the quality of computer-generated content and visualizations techniques. These expectations can not be fully met by today's city models. Furthermore the dominant business concept behind digital globes, the advertisement market, requires the spectacular and recognizable presentation of local businesses within the models. It is therefore foreseeable, that the level of detail has to be further increased and the time for widespread dissemination of LOD3 models is near. Increased detail can only come from three-dimensional façade detail, which in turn can only be achieved by ground-based sensors. An overview of how to create detailed 3D models for buildings in cities from terrestrial data is given in (Mayer et al., 2008).

We have developed an approach which efficiently combines the coarse geometry of an existing LOD2 building model with the detailed features form ground-based LiDAR data which is based on our concept of LASERMAPs. In the case of a LOD2 building model, the approach is straight forward, since the coarse geometry is described by a polyhedral boundary representation and facades are typically planar polygons. This property is also exploited when intensity images are used for texture mapping. In order to use a similar 2D-2D mapping for the integration of terrestrial LiDAR data, we have to derive a two-dimensional representation of the point cloud. Obviously, this is not possible for the whole point cloud. We rather have to split the point cloud into groups with respect to the facades of the building using a simple buffer operation for each facade polygon. The portion of the point cloud that belongs to a particular facade can then be interpolated into a regular raster, in a fashion very similar to digital elevation models derived from aerial LiDAR.

In ground-based façade scanning, as in any ground-based data acquisition, incomplete data acquisition is a major problem and a primary reason for insufficient data quality. Incomplete data acquisition can occur for multiple reasons. One cause is the partial occlusion of the façade by other objects, such as cars, trees, pedestrians, street signs and so on. Figure 1 shows an extreme case for an occlusion by a tree and the incomplete data it causes. Another cause is the self occlusion of the façade due to an oblique viewing angle. Protruded balconies or indented windows will cast shadows along the direction of measurement and cause incomplete data acquisitions. While such effect can be minimized by proper station planning, we have to keep in

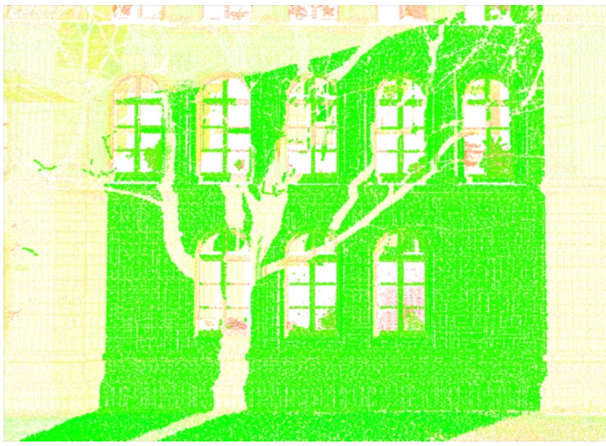

Figure 1. Incomplete ground-based LiDAR data due to an occlusion by a tree. The example is taken from a dataset shown figure 2Figure 2.

mind, that optimal stationing of the sensor is not always possible, especially in inner city areas, where traffic, property boundaries and other circumstances limit the choice for sensor stationing.

In colour texture acquisition it has been shown, that automatic fusion of images acquired at multiple stations can minimize occlusions (Böhm, 2004). However, the key to this approach, i.e. redundant data acquisition, is only feasible since image acquisition with close-range cameras is fairly inexpensive. This approach is not recommended in terrestrial laser scanning, since each station is associated with considerable costs. The same is true for vehicle-based scanning systems, as either the number of scanners mounted on the vehicle or the number of drive-bys need to be increased.

We therefore investigate the exploitation of self-redundancy for modelling facades from incomplete range data. This approach is based on the observation that typically façades of building are composed of repetitive patterns. This repetition of patterns can be used to substitute incomplete areas of the range data. One key question is how the structure of these repetitions can be detected and how it can be efficiently encoded. In this paper we propose a graph-based approach for the encoding of repetitions, which is efficiently derived from key point matching. Before we detail this procedure in section 5, we first give an overview of related work in section 2. In section 3 we give the fundamentals of our approach to façade modelling my LASERMAPs and in section 4 we show how defective areas can be substituted for.

2. RELATED WORK

In the computer vision literature the situation of replacing effective or occlude image parts has been dealt with extensively. Prominent solution include inpainting, a technique which is used to fill small gaps, typically by propagating linear structures from the border of the defective area into the area. A second approach is texture synthesis, which tries to copy repetitive texture to fill an occluded image area. There are many variations and also combinations of these methods, see for example (Criminisi, 2004). Our approach is most similar to exemplar-based inpainting methods. We are not aware of approaches specific to depth maps or range images.

As mentioned earlier we have dealt with the case of occluded colour texture images (Böhm, 2004). As the acquisition of LiDAR data is inherently more time-consuming than image acquisition, it is unrealistic to assume a highly redundant multi-station configuration to overcome occlusions, as it is possible in colour texture image synthesis. However, the idea to have a redundant description for one and the same area of an image still holds. And thus the proposed techniques for robust fusion can be transferred to LASERMAPS.

Self-similarity, repetitive structures and symmetry of buildings and facades in particular have recently attracted great attention in the research field (Müller, Wonka et al. 2006; Ripperda and Brenner 2006). A number of successful applications were developed which exploit these properties (Müller, Zeng et al. 2007). The aforementioned approaches use grammars to store the repetitive pattern of elements. This approach has shown to be successful for synthesizing complete façades and for creating variations of facades. In our work we chose a different representation scheme which is based on graphs. Our idea of using self-similarity for substituting incomplete data is also motivated by the work of (Pauly, Mitra et al. 2005). It differs in that we select the examples used to fill the gaps from the dataset itself rather than using a database of models.

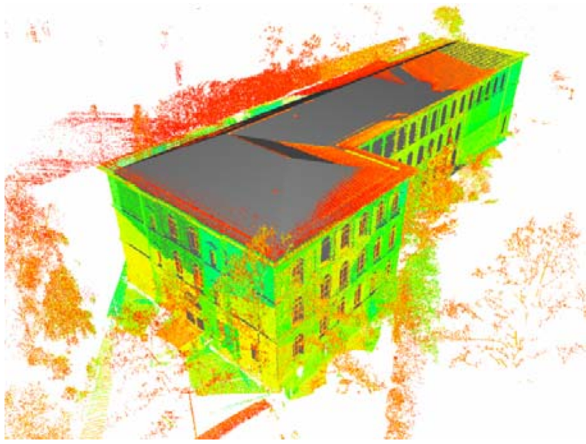

Figure 2. A prismatic building model with detailed roof structure and a registered point cloud acquired by ground-based LIDAR.

The simple representation in the form of a LASERMAP enables us to use likewise simple image processing operators to extract structures. In order to detect repetitive patterns we use a similar procedure as has been proposed in (Wenzel, Drauschke et al. 2007). It is based on feature point extractors typically used for registration of separate data sets. Matching these feature points within the same dataset detects self-similarity.

3. FAÇADE MODELLING USING LASERMAPS

In order to able to combine ground-based laser data with pre-existing building models, the data has to be registered. There many possibilities to compute the registration, ranging from direct georeferencing, to manual and automatic alignment. Our approach to registration and georeferencing of terrestrial laser data and virtual city models is given in full detail in (Schuhmacher and Boehm, 2005).

For the rest of this paper we assume that the registration has been computed and the range data is given with respect to the same coordinate frame as the building model. This initial situation of a point cloud registered to a building model is shown in Figure 1. The dataset depicts the president's office at the Universität Stuttgart. The point cloud was acquired with a terrestrial laser scanner, a Leica HDS 3000, from more than 15 stations. The data covers the facades of the building at a point density better than 20 mm. The large number of stations was necessary to minimize shadowing of occluding objects. By removing selected stations, we can now control the completeness or incompleteness of the data.

Our method to modeling facades is motivated by concepts for modeling developed in computer graphics. In computer graphics the duality of coarse over-all geometry and fine detail has long been noted. The separation of the two is a fundamental modeling principle. Starting with the observations of Blinn (1978), that the effect of fine surface details on the perceived intensity is "primarily due to their effect on the direction of the surface normal … rather than their effect on the position of the surface", modeling concepts were developed, which keep fine surface detail separate as a perturbation of the normal direction or a displacement to the underlying coarser geometry.

As we have noted above, for a prismatic building model the situation is rather simple, since the coarse geometry is described facades are typically planar polygons. We split the point cloud into groups with respect to the facades of the building using a simple buffer operation for each facade polygon. Each subset of the point cloud that is assigned to a particular facade is then interpolated into a regular raster. We can either us nearest neighborhood computation for the interpolation or triangulation. In general the methods used are similar to digital elevation models derived from aerial LIDAR.

We refer to such a re-interpolated point cloud as a LASERMAP (Böhm, 2005). The term is composed from two terms describing the source of the data, a laser scanner, and the use of the data as a source for 2D mapping. Figure 3 shows a LASERMAP of the front facade of the aforementioned building. The gray values correspond to offsets relative to the plane of the facade. The map was computed at a resolution of 10 mm to preserve details, which gives an image of 2878 x 1778 pixels. Each pixel stores the offset in 16 bits.

3.1 Generation of Normal Map and Displacement Map

A LASERMAP is simply a grey value image and can be processed as such. An example for such a simple processing step is the generation of a normal map. The normal map stores the perturbations of the normal vector at each pixel to model the variation of the surface. The unit normal vectors for each pixel can be computed from the partial derivatives of the surface functions as represented by the LASERMAP. This is easily done applying derivative filters to the LASERMAP. Figure 3Figure 3 shows the rendering of a single facade polygon using a normal map derived from the above LASERMAP. While a normal map gives the impression of fine surface detail, this is only achieved by varying the shading of each output pixel; the actual geometry is still a flat polygon. This is advantageous as it does not increase the polygon count of the fully detailed model, when compared to the original model. However, the 'flatness' of the surface is revealed to the observer under very oblique viewing angles.

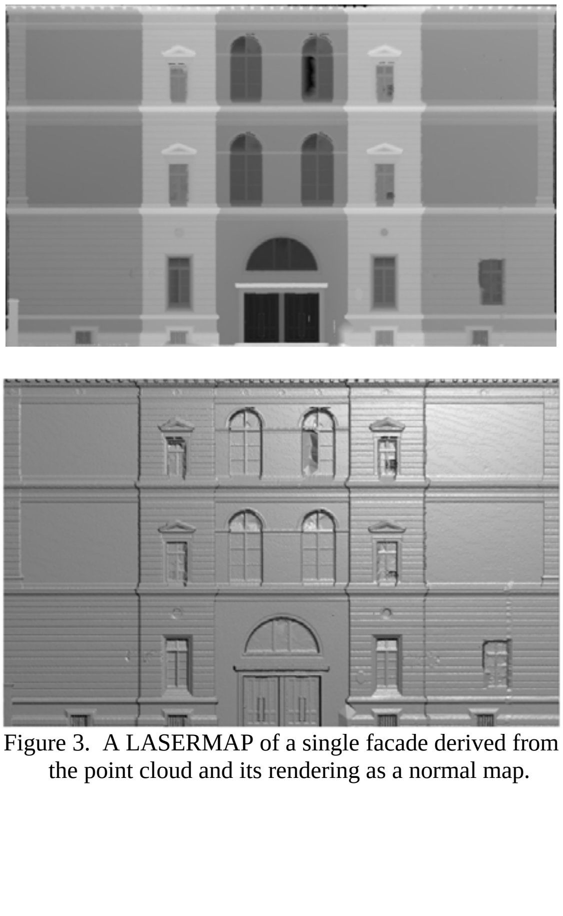

Figure 3. A LASERMAP of a single facade derived from the point cloud and its rendering as a normal map.

To overcome this 'flatness' true three-dimensional geometry has to be generated. This can only be achieved by adding vertices to the model and thereby generating new polygons. This is easily done by subdividing the façade polygon, a standard procedure in modelling. This procedure iteratively subdivides a polygon into smaller polygons, until the desired resolution is achieved. This procedure was initially suggested by (Catmull, 1974). The advantage of subdivision surfaces is that instead of generating vertices explicitly, they are generated implicitly, by storing the subdivision scheme and level. After the subdivision is defined the offset values stored in the LASERMAP are used to displace the generated vertices. The method is therefore referred to as displacement mapping. Figure 4 shows the model mentioned before using displacement mapping.

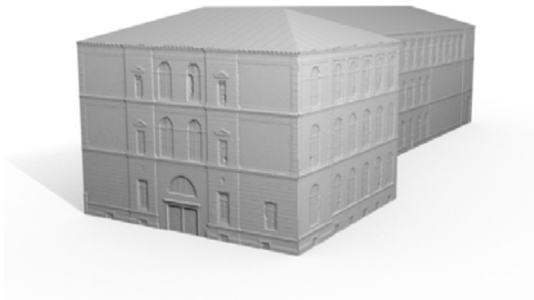

Figure 4. Rendering of a building model with displacement maps.

4. SUBSTITUTION OF DEFECT AREAS

Within most terrestrial laser scanning projects we frequently encounter scanning artifacts, which impair the quality of the point cloud. Such artifacts are often created by occlusions, as mentioned above, but can also be caused by varying surface reflectivity, beam deflections and other problems. It is rather difficult to correct these defects directly in the point cloud. However, it can be rather simple to treat these situations in a LASERMAP.

Since facade architecture does not consist of purely random geometry, but is composed of repetitive elements, we can simply replace defective areas with a copy of an intact element. Since the representation of the LASERMAP is essentially the same as an image, image processing operations can be employed to automate this task. In figure 5 we show an example of a semi-automated repair process. The example is a detail from the faced already shown in figure 3. The right one of the two windows clearly has a defect, due to the window being half-opened at the time of scanning. The repair process starts by interactively marking the defective area in the LASERMAP, shown as a white box in the image. Then we automatically search for a similar area in the LASERMAP. This is implemented using simple template matching. We perform a global template matching across the full LASERMAP and the best match (depicted by a black box) is copied over the defect area. The result of the repair operation is shown in the bottom rows of figure 5.

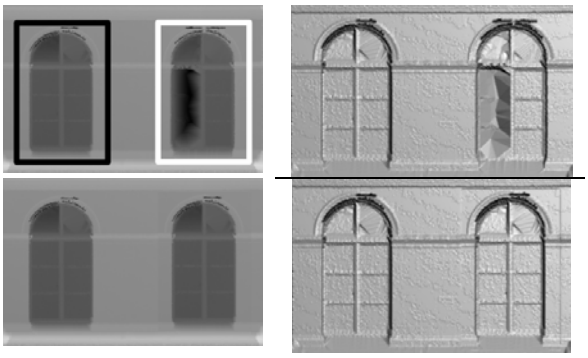

Figure 5. Semi-automatic repair of a defect shown in the top row. The bottom row shows the result of the substitution.

The process is an easy to use tool to remove defects caused by scanning from the facade model. While it is limited to repetitive structures, it can cope with any kind of shape and is not limited to certain geometries. One disadvantage is the need for interactive marking of defect areas, which prevents the application of the approach to mass production of building models. The second disadvantage is the need for a global template matching, since this is a very costly operation. However the repetitions of the elements are spread across the façade, so a simple localized search in the region around the defect area is not sufficient. We will address this issue later.

4.1 Automated identification of defect areas

Since most defects are caused by incomplete data, we devise a simple method to detect areas which have insufficient sampling. In other words we define a defect area as an area of the LASERMAP which has a point density which is considerably lower than the average point density. The average point density can be determined from the pre-planning stage of the scanning campaign, when the density was specified as a mission parameter. Another method is to estimate the average point spacing by computing the nearest neighbor to each scanned point. Obviously doing so for all points is computationally too expensive. But it is sufficient to sample the point density over a subset of points, randomly chosen from the point cloud. Figure 6Figure 6 shows a plot of the number of samples over the distances to the nearest neighbor for a subset of 100000 points. The plot exhibits a clear peak at just under 10 mm, which gives the average point density.

Figure 7 shows the result of such computation. It shows a LASERMAP and the corresponding error map, which marks areas of insufficient point density. We can see that most areas of incomplete data occur at windows. This is due to the laser beam penetrating through the glass and thus not giving a return. But we can also see how shadowing caused by the dominating oblique view from the left creates areas of incompleteness. Even small occlusions caused by trees are visible at the right border. The sensitivity of the error map can be controlled by simple thresholding. The regions of the error map can be used to replace the interactive marking of defect areas.

The method for computing the error map integrates well with the computation of the LASERMAP itself, since both are based on the computation of planar nearest neighbors. We use an accelerated search structure for the computation, which has to be initialized only once for both computations.

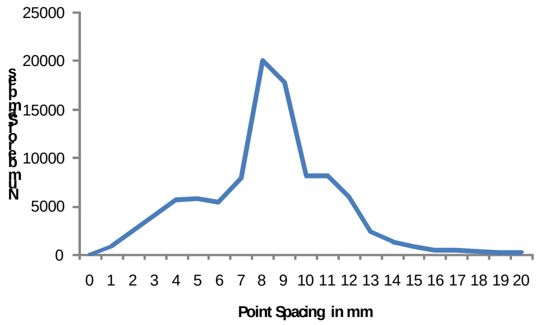

Figure 6. Distribution of the point spacing to estimate average point density.

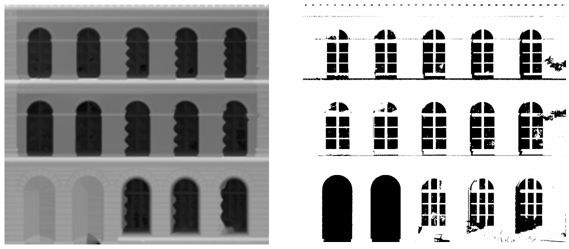

Figure 7. A LASERMAP computed by nearest neighborhood interpolation and the corresponding error map, which marks areas of insufficient scan density.

5. ENCODING OF REPETITIONS

To avoid global template matching, a problem mentioned above, we restrict the matching to key points. We use the non-maximum suppression on the difference-of-Gaussian to detect robust key points. This is in effect the same operation used to localize SIFT key points. However, we do not compute SIFT features. To detect corresponding key points within the same dataset, we compute the normalized correlation coefficient of the patches surrounding the key points. Matching key points are assigned to the same equivalence class. In Figure 8 the image on top shows the key points detected in a LASERMAP.

We use a graph to store the relations among the key points. A Graph *G* consists of a set of vertices *V* and a set of edges *E*. The edges connect the vertices and thus are represented by pairs. If two key points *k1* and *k2* have matching patches, we add the two vertices *v1* and *v2* to the graph and add an undirected edge *e(v1, v2)*. The Euclidian distance of *k1* and *k2* is set as the weight of the edge *e*.

A problem in the matching of key points is the selection of a proper threshold for the correlation coefficient. Because of this problem we often encounter the following situation: while key point *k1* is matched to *k2* and *k2* is matched to *k3*, *k1* is not matched to *k3*. Using the arrow symbol for the matching we can write $k1 \rightarrow k2, k2 \rightarrow k3, k1 \nrightarrow k3$. However, since we think of matching key points as equivalence classes, we have to generate the missing relation. In graph theory this is known as transitive closure. So if vertex v2 can be reached from vertex v1 ($v1 \Rightarrow v2$), we add the edge *e(v1, v2)* to the graph. This creates fully connected sub-graphs, representing equivalence classes of matching key points. Figure 9 shows in the upper image a portion of a graph after transitive closure.

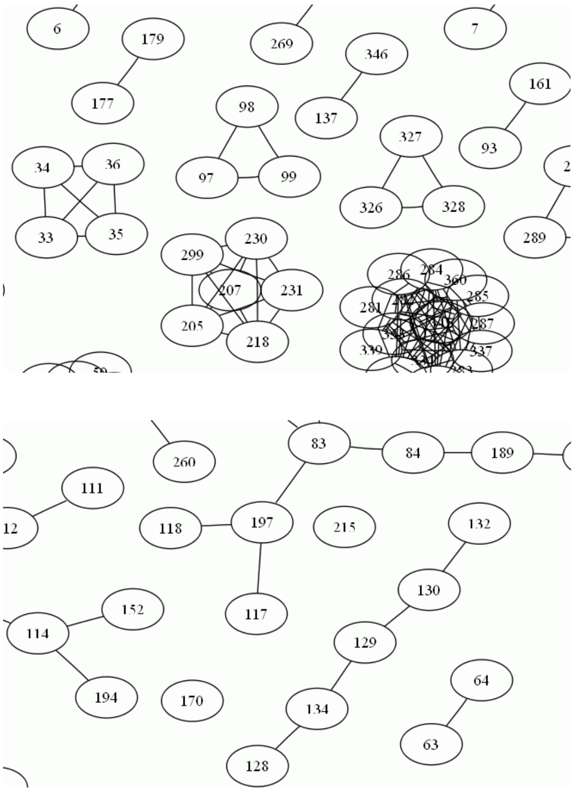

Figure 9. The representation of matching key points as vertices in a graph. The top image shows a detail of the fully connected sub-graphs. The bottom image shows examples for the minimal spanning tree(s).

While this graph fully represents the matching relation of the key points, when we plot the vertices and edges onto the corresponding image features, it does not exhibit the expected row and column structure we typically expect from a façade (see Figure 8Figure 8 second row).

We can compute one more transformation on the graph, known as the minimum spanning tree. The minimum spanning tree problem is described by the following condition: find a subset of edges $T \subset E$ which connects all vertices of the Graph and whose total weight is minimal. The total weight is the sum of all weights of the edges contained in *T*. One popular algorithm to compute the minimum spanning tree is given by Kruskal (1956). In our case, since we have unconnected sub-graphs, we do not compute a single spanning tree, but several unconnected spanning trees. The Euclidian distance of the key points, which was chosen as the weights of the edges, should favour the horizontal and vertical edges of the graph and partially eliminate edges running traverse. We can further filter the graph, eliminating edges which violate some boundary conditions, such as minimum or maximum length or direction. The result of the computation is shown in the bottom image of Figure 8. The graphic shows the main structure of the façade, exhibiting the strong horizontal similarity on both upper levels. The two doors on the lower level have absolutely no vertical connectivity.

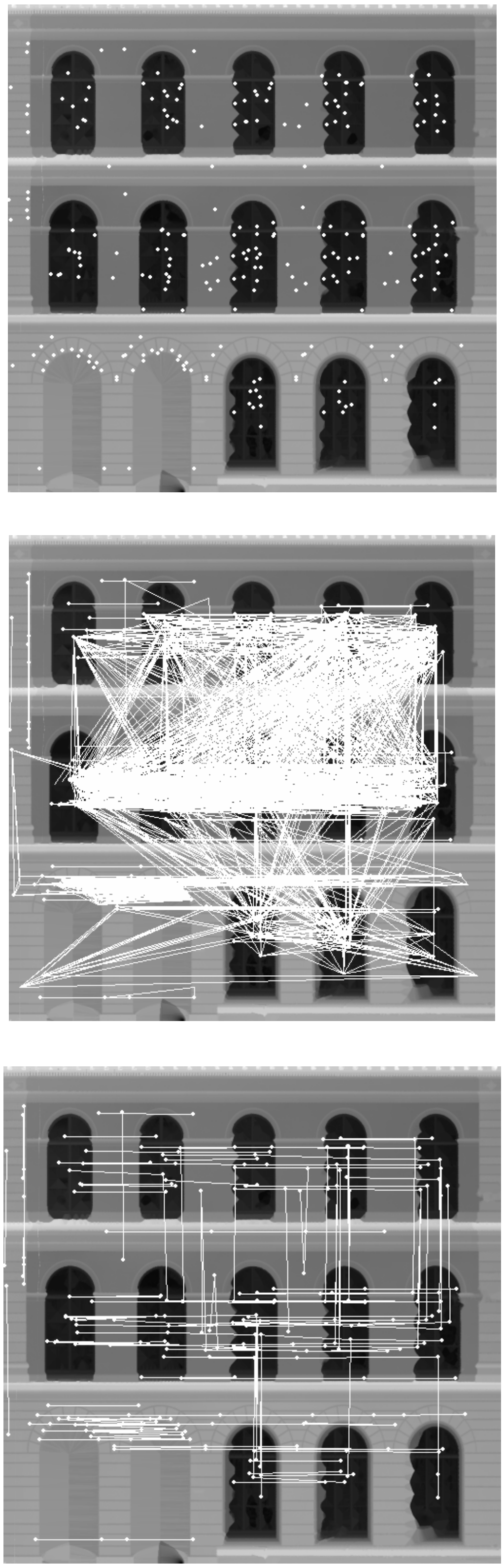

Figure 8. (Top) The key points detected from the LASER-MAP. (Middle) The fully connected sub-graphs of matching key points. (Bottom) The minimum spanning tree(s).

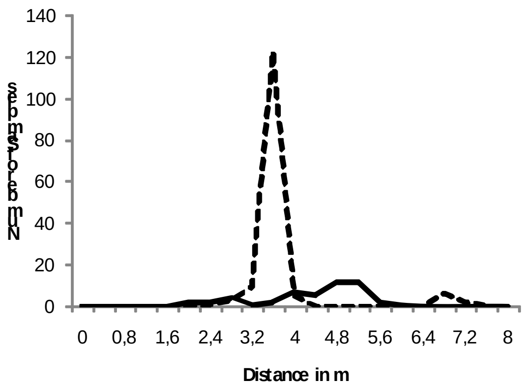

Figure 10. Distribution of the horizontal (dashed line) and vertical (solid line) distances of the graph edges.

The final graph can be used in different ways to aid in the retrieval of replacement patches. For one it can be used to determine standard values for floor height and horizontal separation of façade elements. To do so we detect the peaks in the histogram of vertical and horizontal edge lengths. Figure 10 shows such a plot. We can read a clear peak of about 3.3 m in horizontal distance. A peak of 5.1 m denotes the vertical separation of the floor levels. Both values were verified in the original point cloud. The graph is further used to reduce the template search for substituting defect areas. Instead of a global search, we only have to search at the vertices of the sub-graph, which has a vertex closest to the defect area.

6. SUMMARY

We have presented our strategy on adding façade detail to existing building models using ground-based LiDAR. A typical problem which occurs in street-level data acquisition is incomplete data. We have shown approaches on how to handle such cases. We have shown that the representation of façade detail in a LASERMAP is advantageous, since it allows the use of sophisticated image processing algorithms. Such algorithms can be used to substitute defective areas of the LASERMAP, caused by missing data. Our graph-based representation scheme reduces the number of comparisons needed to find suitable matches.

ACKNOWLEDGEMENT

We acknowledge the work of Angelos Tsakiroglou, Thanasis Georgantas and Christos Stamatopoulos of the Aristotle University of Thessaloniki, who have acquired the laser data set during their stay in Stuttgart in the ERASMUS exchange program.

REFERENCES

Blinn, J. F., 1978. Simulation of wrinkled surfaces. *ACM SIGGRAPH Computer Graphics* 12(3), pp. 286-292.

Böhm, J., 2004. Multi-Image Fusion for Occlusion-Free Façade Texturing. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Istanbul, Turkey, Vol. XXXV, Part B5, 867-872.

Böhm, J., 2005. Terrestrial Laser Scanning - A Supplementary Approach for 3D Documentation and Animation, In: Photogrammetric Week '05, Fritsch (ed.), Wichmann, 263-271.

Catmull, Ed. A., 1974. Subdivision Algorithm for Computer Display of Curved Surfaces. PhD thesis, Computer Science Department, University of Utah, Salt Lake City, UT, Report UTEC-CSc-74-133.

Criminisi, A.; Perez, P. & Toyama, K., 2004. Region filling and object removal by exemplar-based image inpainting. In: Image Processing, IEEE Transactions on, 13, 1200-1212.

Kolbe, T. H. & Gröger, G., 2003. Towards unified 3D city models. In: Proceedings of the ISPRS Comm. IV Joint Workshop on Challenges in Geospatial Analysis, Integration and Visualization II.

Kruskal, J., 1956. On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem. In: Proceedings of the American Mathematical Society, American Mathematical Society, 7, 48-50.

Mayer, H.; Hinz, S. & Stilla, U., 2008. Automated extraction of roads, buildings, and vegetation from multi-source data. In: Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences: 2008 ISPRS Congress Book, Li, Z. (ed.).

Müller, P., P. Wonka, et al. (2006). Procedural modeling of buildings. ACM Transactions on Graphics.

Müller, P., G. Zeng, et al. (2007). Image-based procedural modeling of facades. SIGGRAPH '07: ACM SIGGRAPH 2007 papers, New York, NY, USA, ACM Press.

Pauly, M., N. J. Mitra, et al. (2005). Example-Based 3D Scan Completion. Proceedings of the 3rd Symposium on Geometry Processing (SGP).

Ripperda, N. and C. Brenner (2006). Reconstruction of Facade Structures Using a Formal Grammar and RjMCMC. DAGM-Symposium.

Wenzel, S., M. Drauschke, et al. (2007). Detection of repeated structures in facade images. 7th Open German / Russian Workshop on Pattern Recognition and Image Understanding, Ettlingen.

Schuhmacher, S. & Boehm, J., 2005. Georeferencing of terrestrial laserscanner data for applications in architectural modeling 3D-ARCH 2005: Virtual Reconstruction and Visualization of Complex Architectures, XXXVI, PART 5/W17.