Automatic Image Orientation for Accurate Texture Mapping of 3D City Models

L. Grammatikopoulos[1*], I. Kalisperakis[1]and E. Petsa[1]

[1] Department of Civil Engineering and Surveying & Geoinformatics Engineering, Technological Educational Institute of Athens, Greece, *lazaros@teiath.gr

Keywords: Texture Mapping, Lidar, 3D city models**,** Image orientation

Abstract

In the last decade, 3D city model generation is proved to be an indispensable task for a variety of applications related to tourism, real-estate management, 3D GIS, cadastre, extending common geospatial products such as maps and orthophotos to more realistic representations. Besides the acquisition of high fidelity 3D data, which is usually performed by airborne Lidar scanners or automated photogrammetric procedures, a crucial step for the generation of visually detailed 3D models is texture mapping from aerial imagery. In this contribution we propose a framework for accurate texture mapping of 3D models with the integration of vertical and oblique aerial images, which consists of three main steps: In the beginning, the image set is oriented through a Hierarchical Structure from Motion scheme. Afterwards, a registration is performed with respect to the existing 3D city model and finally, texture mapping is applied using a multi-image blending technique. The proposed approach has been evaluated on a dataset covering a densely build neighbourhood of Athens in Greece which includes a DSM generated from Lidar an orthomosaic and a set of unoriented overlapping vertical and oblique aerial images.

Introduction

The generation of photorealistic 3D city models is a task of great interest for many researchers in the fields of Photogrammetry and Computer Vision. The two main issues that are involved and studied are image to model registration and texture mapping. Concerning the registration stage, Delara *et al*, 2004 and Habib *et al*, 2005 extract manually Ground Control Points required for the registration phase and apply a bundle adjustment solution. Armenakis *et al*, 2012 use a semi-automatic approach for the alignment of photogrammetric and Lidar data based on the identification of corresponding buildings in the two datasets. Gneeniss *et al*, 2014 optimize their registration by extracting automatically Lidar control points and integrating them into a final self-calibrating aerotriangulation. On the other hand, Yang and Chen, 2015 combine mini-UAV images with Lidar data. In their approach images are first oriented via a Structure from Motion (SfM) algorithm and the obtained point cloud is georeferenced based on a building matching scheme and a subsequent ICP procedure. Concerning texture mapping, it is advisable to blend colour from more images in a weighted average of corresponding values (Baumberg, 2002, Orzan & Hasenfratz, 2005). In this contribution, we extend our previously presented approach for automatic orthophotos generation (Karras *et al*, 2007) to produce photorealistic 3D models from Lidar and image data.

Automatic orientation of aerial imagery

The first step of our approach is to automatically compute the relative orientation of all available aerial images. This is performed at different image scales in order to handle high resolution imagery. We start at a low resolution where 2D SIFT (Lowe, 2004) and SURF (Bay *et al*, 2008) features with descriptors are extracted on all images. Afterwards, sparse matching is performed and outliers are identified through RANSAC on the fundamental matrix computation. Combining the inlier matches across different stereopairs we obtain multi image point correspondences. A bucketing algorithm is also applied to reduce the number of tie points without affecting their distribution. Then, by means of closed-form algorithms and successive small bundle adjustment solutions we estimate initial image orientations which are refined through a typical self-calibrating bundle adjustment solution (Fig. 1). This procedure is repeated at higher image resolutions but this time the matching is restricted by the estimated epipolar geometry and a rough 3D reconstruction from the tie points of the previous resolution. The final interior and exterior orientation of all images is calculated again by a bundle adjustment solution.

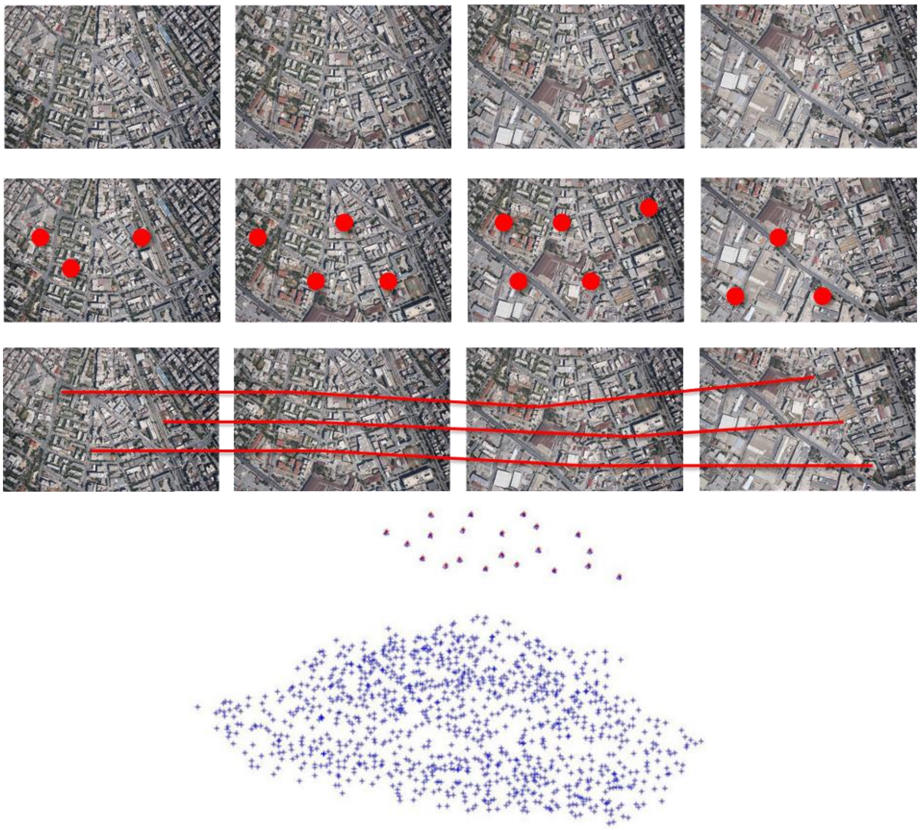

Figure 1. Automatic Image Orientation Scheme using SIFT and SURF features

From the evaluation with our dataset, the standard error of the relative orientation using hierarchical SFM was about 0.5 pixel.

3D model to image registration

In order to apply texture mapping to the existing 3d city model, image orientations acquired from previous step, need to be transformed to the reference projection system of the DSM. For this purpose different approaches can be followed. Conventionally, this is performed by the use of Ground Control Points measured in the field with typical surveying techniques such as GPS which are manually identified on the available images. However, several approaches have been proposed for the automation of this step. Here we apply a framework suggested by our team (Grammatikopoulos et al., 2015) that uses an available orthophoto as a reference. Based on this, a DSM is generated from the relative oriented image set using dense stereo algorithms (Stentoumis et al., 2015) and a new orthomosaic is computed (Karras et al., 2007). Then, Ground Control Points are extracted by matching 2D features between the newly created and the reference orthophotos and 3D coordinates are assigned to them via interpolation on the

reference DSM. These Control Points are automatically identified on the aerial images and a final bundle adjustment solution allows the optimal registration among 3D model and aerial images.
Applying this approach to our dataset, we had a final RMS error for the GCPs of about 40 cm, which is about two times the size of the reference orthomosaic pixel and half the 3D model resolution.

Texture mapping
Once all available data (3d surface model and aerial images) have been registered together, an accurately textured 3d city model can be acquired by employing texture mapping algorithms. The simplest approach is to attribute RGB colour values to every polygon vertex of the 3D model through back-projection on the aerial images. However, this leads to a less detailed representation, where the texture of each surface triangle is simply generated from the interpolation of the corresponding vertices' colours. Hence, details smaller than the size of a triangle are not visible.

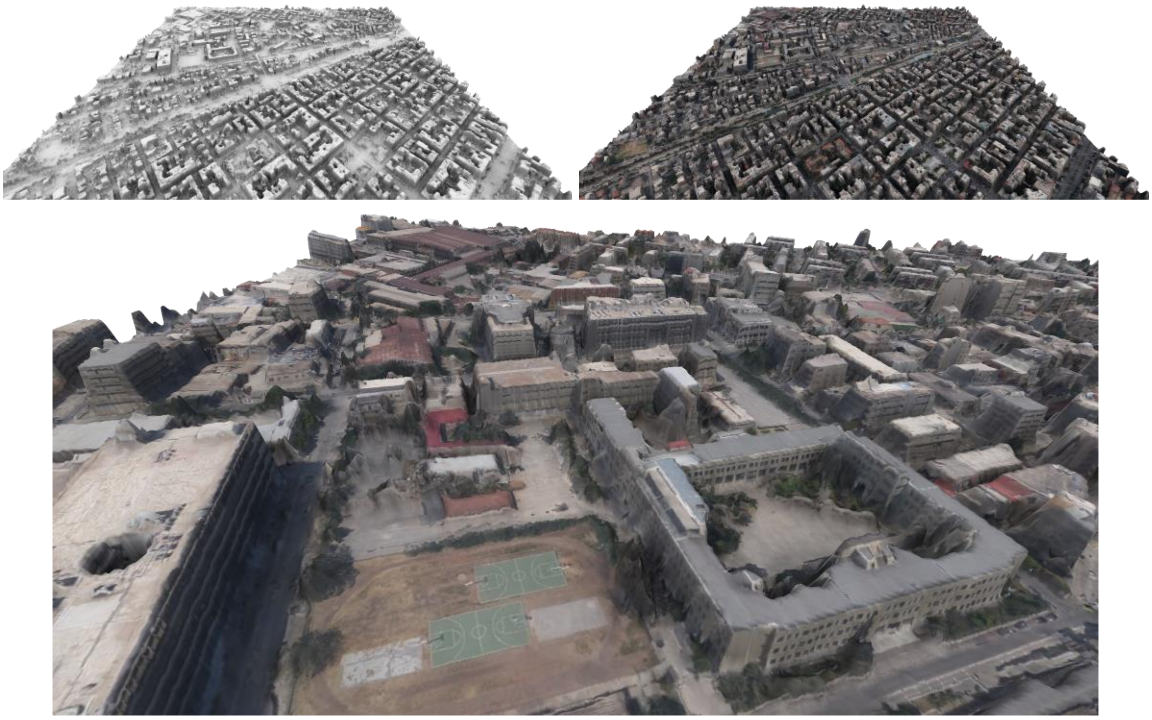

Figure 2. Overall and detailed view of 3D City Model with Multi-image Texture Mapping

This drawback can be overcome by applying texture instead of colour to each triangle. A straightforward approach is to compute for all vertices of the model *u,v* coordinates on the original images. In this way, texture from a suitably chosen image (most nadir, closest distance etc.) is applied to every model triangle. The main disadvantages of this approach is memory inefficiency since all original images are required to display the model and colour discontinuities for neighbouring triangles textured from different images.
A more accurate approach is ensured by blending texture from multiple images. A new image *(texture atlas*) is generated that stores texture for the whole model. In a first step *u, v* coordinates on the texture atlas are estimated for each triangle via a surface parameterization algorithm and the final texture is generated by combining texture from corresponding image regions using a weighted blending scheme that can compensate for different orientations, scale

and resolution of the images involved (Fig. 2). Occlusion handling is also incorporated along with means for automatically eliminating colour outliers from individual images, especially near image occlusion borders.

Concluding Remarks

The work presented here is part of an ongoing research project on 3D city models that focuses on multimodal representations of urban areas. A framework was proposed for the automatic orientation of aerial imagery with respect to an existing 3D model obtained from Lidar scanning. First results have shown the effectiveness of the suggested approach.

Acknowledgements

This research has been co-funded by the European Union (European Social Fund) and Greek national resources under the framework of the Archimedes III: Funding of Research Groups in T.E.I. of Athens project of the Education & Lifelong Learning Operational Programme.

We would also like to thank GeoIntelligence S.A. for providing the DSM elevation data as well as the National Cadastre & Mapping Agency of Greece for supplying us with the high resolution aerial photographs.

References

Armenakis, C., Gao, Y. and Sohn, G., 2012. Semi-Automatic Co-Registration of Photogrammetric and Lidar Data Using Buildings. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences I-3, 13–18.

Baumberg A., 2002. Blending images for texturing 3D models. British Machine Vision Conference, pp. 404-413.

Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L., 2008. SURF: Speeded Up Robust Features. Computer Vision and Image Understanding, Vol. 110, No. 3, pp. 346–359.

Delara, Jr. R., Mitishita, E. A. and Habib, A., 2004. Bundle Adjustment of Images From Non-Metric CCD Camera Using Lidar Data As Control Points. International Archives of XXth ISPRS Congress, Commission III, 13–19.

Gneeniss, A. S., Mills J. P. and Miller, P. E., 2014. In-Flight Photogrammetric Camera Calibration and Validation via Complementary Lidar. ISPRS Journal of Photogrammetry and Remote Sensing 100, 3–13.

Grammatikopoulos L., Kalisperakis I., Petsa E., Stentoumis C., 2015. 3D city models completion by fusing lidar and image data. Videometrics, Range Imaging, and Applications XIII .Proc. SPIE, vol. 9528.

Habib, A., Ghanma, M., Morgan, M. and Al-Ruzouq, R., 2005. Photogrammetric and Lidar data registration using linear features. Photogramm. Eng. Remote Sensing 71 (6), 699–707.

Karras, G., Grammatikopoulos, L., Kalisperakis, I. and Petsa, E., 2007. Generation of orthoimages and perspective views with automatic visibility checking and texture blending. Photogrammetric Engi¬neering and Remote Sensing, 73(4), 403-411.

Lowe, D., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60 (2), 91–110.

Orzan, A., Hasenfratz, J.-M., 2005. Omnidirectional texturing of human actors from multiple view video sequences. IEEE Int. Workshop on Human Computer Interaction, Beijing.

Stentoumis, C., Grammatikopoulos, L., Kalisperakis, I. and Karras, G., 2014. On accurate dense stereo-matching using a local adaptive multi-cost approach. ISPRS Journal of Photogrammetry and Remote Sensing, Volume 91, May 2014, Pages 29-49.

Yang, B. and Chen, C., 2015. Automatic registration of UAV-borne sequent images and Lidar data. ISPRS J. Photogramm. Remote Sens., 101, 262–274.