Do we need a voxel-based approach for LiDAR data in geomorphology?

Balázs Székely (1,2), Peter Dorninger (3), Robert Faber (4), and Clemens Nothegger (1)

(1) Vienna University of Technology, Institute of Photogrammetry and Remote Sensing, Vienna, Austria
(balazs.szekely@ipf.tuwien.ac.at), (2) Department of Geophysics and Space Science, Eötvös University, Budapest, Hungary,
(3) Vienna University of Technology, Christian Doppler Laboratory, Institute of Photogrammetry and Remote Sensing,
Vienna, Austria, (4) TerraMath, Vienna, Austria

Generations of geomorphologists have developed a multi-faceted approach to model the Earth's (and planetary) surface and the corresponding processes. This set of models is based on data, more specifically on conspicuously increasing amount of data. Obviously, all geomorphologists wish themselves more accurate and increasingly high resolution data on, or related to the Earth surface. This evolution also means that the studied boundary is not anymore a single surface; instead it is considered mostly a 2.5D object, sometimes a real 3D object.

LiDAR technology can cope with this challenge: the data accuracy and resolution requirements can be fulfilled by applying this method. Although it is yet somewhat still expensive, more and more areas will be scanned, and in some regions the topographic point clouds are already multitemporal (causing of course other types of processing and evaluation problems). It is rather obvious that for certain, geomorphologically very interesting areas very dense and severalfold multitemporal LiDAR data will be available in the near future.

These data sets will have various differences concerning the data density, accuracy, data acquisition technique (conventional or full-waveform), and perhaps most importantly, concerning the actual state of the surface. Similar to the satellite imagery integration problems, soon we all have to face with the LiDAR data integration problem. What type of surface or surfaces can be derived from this multitude of data sources with acceptable ambiguity? What conclusions can be drawn from these data that were originally acquired for various other purposes using various acquisition concepts? Will it be advantageous for geomorphic use to have a coverage of the surface with 100-200 points/m^2 density?

Clearly, these data are, if they are once collected, still too expensive not to be integrated for further analyses. Consequently, we need a data reduction concept that effectively decreases the computer capacity needed to store, process and visualize the results. To reduce the amount of originally collected data for further applications, in general, continuous model surfaces are derived from the point clouds using interpolation approaches. Commonly, grid-based or triangulation models are used for that purpose. Typical models are Digital Surface Models (DSM) representing the whole topography including all natural (e.g. vegetation) and artificial (e.g. buildings) objects and Digital Terrain Models (DTM) representing the topography only.

In the visual computing industry the voxel-based approach is quite common for various purposes. Although this technology is quite straightforward concerning the data reduction, it is hardly applied in the geomorphic context. An argument can be against its application that mostly we are interested in a surface, not a volume. Of course in the strict sense it is true, however, if we consider the technology itself how the data, especially the ground data are derived, it turns out that actually it is a volume with a certain accuracy that we are sampling of. The position of this "relatively thin" volume also varies, especially in mountainous areas. Here, depending on the slope angles, the accuracy also varies, especially for the integrated data set consisting of a multitude of sources, e.g. mixing conventional (first echo/last echo) and full-waveform data. These point clouds also contain attributes that otherwise could be very valuable, but during the integration, their meaning may be lost, or it cannot be integrated to the data set. A large scale application of such approaches is mainly prevented by the problems introduced by the high amount of data, making on-the-fly processing a challenging task.

To overcome these restrictions and to enable taking advantage of the new possibilities provided by the waveform analysis, we propose a voxel-based data representation approach. The multichannel/multilayer design with an a priori unlimited number of layers enables storing an unlimited number of additional parameter per point. We expect that such a voxel structure enables to represent and analyze huge datasets of large areas (e.g., connected regions which are geologically relevant to be analyzed at once) in applicable processing times in order to bridge the gap between the original point cloud and the user and interpretation level. The challenging task to be solved will be to reduce the amount of data significantly by means of the proposed structure while preserving the content of the original data.