EARTH OBSERVATION SYMPOSIUM (B1)
Big Data, Data Cubes and new platforms to exploit large-scale, multi-temporal EO Data (6)

Author: Mr. Wolfgang Lueck
PCI Geomatics, Canada

CLOUD BASED PROCESSING OF FREE AND COMMERCIAL EARTH OBSERVATION DATA WITH PCI GXL, POPULATING AND ANALYZING DATA WITH THE AUSTRALIAN GEOSCIENCE DATA CUBE SOFTWARE.

Abstract

Over recent years, corporate and public satellite operators have provided API's to their image archives and distributed image processing has moved into the cloud. With this in place, the development of multi-sensor cloud enabled automated image preprocessing and analysis tools, feeding data cubes such as released by Geoscience Australia, have become necessary. Such a system is PCI GXL. The system can run on an Amazon cloud and extract time series of data from the archive of Planet Labs or ESA through their respective API's. A range of image preprocessing techniques can be undertaken with a set of software libraries and executables. GXL powered automated workflows would typically include image calibration, atmospheric corrections, image to image registration to 1/10th of the image pixel, image compositing and image fusion. Preprocessing is optimized in a distributed way to feed into a data cube released as open source by Geoscience Australia. Once imagery has been deposited in the data cube, quantitative time series analysis can be performed by GXL to enable rapid identification of changes. A new approach to the analysis of multi resolution data supports the combined use of high spatial resolution commercial data with lower spatial resolution imagery collected at a high temporal frequency. Image segmentation and feature calculation based on pure pixels only allow the abstraction of data to units of homogeneous surfaces reducing computational efforts and complexity for a multi resolution spatial and temporal image analysis. A set of software libraries dedicated to SAR image processing for polarimetry and interferometry enable the use of SAR data with all its attributes in the data stack. Near real time change detection and mapping applications for disaster management can be addressed in this way. Machine learning classifiers are able to convert data in the data cube to thematic geo-spatial information efficiently and accurately. The system can be integrated with cloud enabled multi-mission satellite ground segments such as FarEarth to complete the loop from satellite tasking to geo-spatial information delivery to end users.