Data Analytics for Rapid Mapping: Case Study of a Flooding Event in Germany and the Tsunami in Japan Using Very High Resolution SAR Images

Corneliu Octavian Dumitru, Shiyong Cui, Daniela Faur, Mihai Datcu *IEEE Fellow*

In this paper we present data analytics for a quantitative analysis in a rapid mapping scenario applied for damage assessment of the 2013 floods in Germany and the 2011 tsunami in Japan. These scenarios are created using pre- and post-disaster TerraSAR-X images and a semi-automated processing chain. All our data set is tiled into patches and Gabor filters are applied as a primitive feature extraction method to each patch separately. A support vector machine with relevance feedback is implemented in order to group the features into categories. Once all categories are identified, these are semantically annotated using reference data as ground truth. In our investigation, non-damaged and damaged categories were retrieved with their specific taxonomies defined using our previous hierarchical annotation scheme. The classifier supports rapid mapping scenarios (e.g., floods in Germany, tsunami in Japan) and interactive mapping generation. The quantitative damages can be assessed by: (1) flooded agricultural areas (21.66% in the case of floods in Germany and 4.15% in case of tsunami in Japan) and destroyed aquaculture (2.33% in case of tsunami in Japan); (2) destroyed transportation infrastructures, such as airport (50% in the case tsunami in Japan), bridges, roads, etc.; and (3) debris that appears in post-disaster images (3.81% in the case of tsunami after the aquaculture was destroyed). The first analysis envisages the floods of Elbe river in June 2013: 30% of the investigated area, about 179 square km, including agricultural land, forest, river and some residential and industrial area close to the river, was covered by water. The second analysis, considering an area of 59 square km affected by the tsunami, led us to conclude that three months after the tsunami some of the categories returned to previous functionality -the airport, others return to partial functionality such as isolated residents, and some were totally destroyed - the aquaculture.
The flooded area was about 59 square km. The proposed approach goes beyond a simple annotation of the data and provides an intermediate product in order to produce a relevant visual analytics representation of the data. This makes it easier to compare data sets and different quantitative findings in a meaningful manner, accessible both to experts and regular users. Our paper presents an interactive and automatic, fast processing method applicable to large and complex data sets (such as image time series). In addition to enhancing the information content extraction (number of identified categories), this approach enables the discovery and analysis of these categories. The novelty of this paper resides in that this is the first time data analytics have been run on a large data set and for different scenarios using a semi-automated processing chain.