Automated Classification of Road-Surface Types Based on Crowd-Sourced Data

1. Problem Statement
For cyclists, it is crucial to know which kind of road-surfaces they have to expect along a cycling route. Traditional land-cover classification by analyzing aerial imagery is inapplicable for various reasons:
- It is expensive
- Image resolutions are too low
- Images are not up to date
- Image quality is inconsistent

2. Research Question
Can road-surface information be deduced from crowd-sourced geographic data?

3. Methodology
We combine heterogeneous data from various sources and create a decision tree for classifying road surfaces. OpenStreetMap (OSM) is our main data source, because it is the biggest open-source project offering crowd-sourced geographic data.

4. Data Analysis
OSM tags are textual annotations which provide various information about streets. Thorough data analysis was necessary to identify useful tags. Some streets already have a surface description among other tags which describe properties like the street type or the speed limit.

5. Additional Data
External Datasets are used to complement OSM data. CORINE Landcover Data, which contains information about the type of an area (e.g., if it is an urban or agricultural area, a forest, etc.) and height information are mapped to each OSM way. Our system is designed so that further data can be added in the future.

6. Taxonomy
Due to the crowd-sourced nature of OSM, the data has many missing values. A two-level taxonomy is used to adapt classification granularity dependent on the available data.

7. Classification
Streets which already have surface information are used to build decision trees for coarse and fine-grained classification. Streets which have no surface tag can now be classified.

8. Results and Conclusion
Our method was tested on different areas in Austria.
- ~90% of roads are correctly classified on Level 1
- ~70% of roads are correctly classified on Level 2
The best results are achieved when using areas with similar geographic properties for creating the tree and classifying the streets.
Future work could improve our method by adding further external data sources and tags.

References

Kontakt: silvana.podaras@gmail.com