Applicability of KOMPSAT Image Data in Open Data Cube

Kwangseob Kim, Kiwon Lee *

Department of Electronics and Information Engineering, Hansung University, Seoul, Korea, engintruder@hansung.ac.kr, kilee@hansung.ac.kr

* Corresponding author

Keywords: Earth Observation Satellite Image, GEOSS, Open Data Cube, Open Source

Abstract:

Data cube terms a multi-dimensional stack of gridded datasets aligned for analysis. Open Data Cube (ODC) is an open-source based information processing and managing platform on the viewpoint from web-based infrastructure. This open platform is for a large volume of geo-spatial information with geo-rectified coordinates, and it has been applied by non-profit international organizations such as the Committee on Earth Observation Satellites (CEOS) for an international coordination and management of space-borne missions, Global Earth Observation System of Systems (GEOSS) in the Group on Earth Observations (GEO), as an intergovernmental organization to improve the applicability, accessibility and usability of Earth observations for benefit of human society.

The building of Analysis Ready Data (ARD), which means the preparation of radiometric calibration and geo-rectification, is for the data cube utilization. The platform converts large-scale satellite image data into analytic information, providing functions for time series analysis. Internationally, there has been an ever-increasing number of country-based data cube deployments with freely available satellite images, including Australia Data Cube, Vietnam Data Cube, Swiss Data Cube, and Colombia Data Cube, as a computing environment for information distribution, sharing, and analysis.

However, there is no program yet to register Korea Multi-purpose satellite (KOMPSAT) optical and radar images on this platform, so this study developed the registering and ingestion script codes for KOMPSAT optical and radar image sets into ODC. Data ingestion is the process of obtaining and importing data for immediate use or storage in a database. Thus, an ingestion process is required to add satellite data to the ODC platform and the process can be divided into three main stages in Figure 1. First, it is to define the data type in the YAML format. Then, the indexing process of datasets for metadata registration is necessary. The next step is a data ingestion process that users can be used directly data sets collected in ODC.

Figure 2 shows some of the Python module results for index datasets and the process of metadata generation. The metadata YAML, which is required for indexing, has many advantages in many respects in the creation of metadata through Python modules. This is why we added Python modules to create metadata YAML. In particular, the KOMPSAT data ingestion process was designed so that all of them were possible through one module.

Using script modules for these steps, the functional accuracy was tested with actual satellite data. Color composite images using RGB bands of KOMPSAT optical images were generated in the ODC environment in Figure 3. In this process, image data formats of GeoTiff and netCDF are also supported.

In this study, consideration points for implementation of ODC applications are also discussed. KOMPSAT data is basically commercial-based products, unlike other freely accessible satellite images in the ODC applications. For the practical contribution for ODC-GEOSS, careful considerations for data policy are needed, because it can be applied as a reference model for other commercial satellite data for GEOSS.

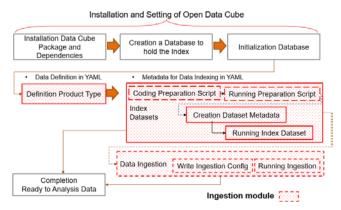


Figure 1. KOMPSAT ingestion modules and their processes in Open Data Cube.

```
Data
KOMPAST-2
                            KOMPAST-3
                                                        KOMPAST-3A
                                                                                       KOMPAST-5
      This preparse script supports only for Aux.xml metadata file
For example: python script.py [Path for dataset]""")
@click.argument('datasets', type=click.Path(exists=True, readable=True, writable=True), nargs=-1)
       def main(datasets):
        logging.basicConfig(format='%(asctime)s %(levelname)s %(message)s', level=logging.INFO) for dataset in datasets:
          path = Path(dataset)
          documents = prepare_datasets(path)
          dataset, folder = documents
           yaml path = str(folder.joinpath('datacube-metadata.yaml'))
           logging.info("Writing %s", yaml_path)
with open(yaml_path, 'w') as stream:
yaml.dump(dataset, stream)
            name__ == "__main__":
```

Figure 2. Metadata handling using the Python modules for indexing.

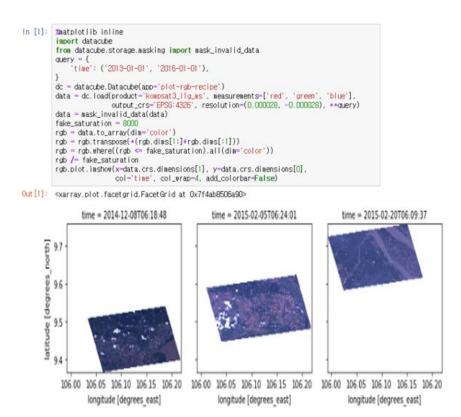


Figure 2. Testing processes of KOMPSAT data into ODC.

Acknowledgments

This research was supported by the National Land Space Information Research Program from the Ministry of Land, Infrastructure and Transport, Korea (No. 14NSIP-B080144-01).