EXPLORING HIGH RESOLUTION SATELLITE IMAGE COLLECTIONS USING THEIR HIGH-LEVEL FEATURES

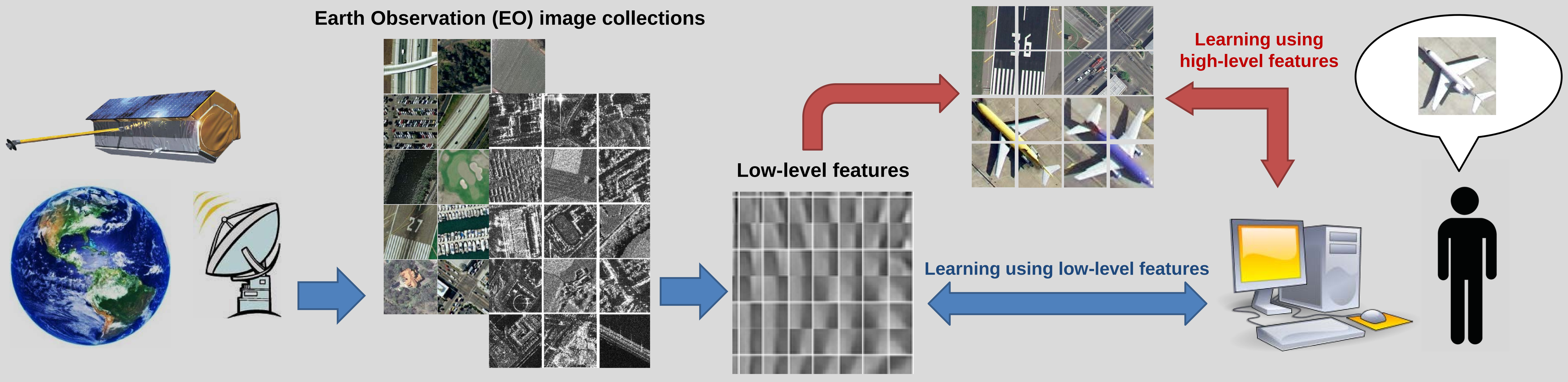
Reza Bahmanyar[1,*], Wei Yao[2], Shiyong Cui[1], Otmar Loffeld[2], Mihai Datcu[1,*]

[1] Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany

[2] Center for Sensor Systems (ZESS), University of Siegen, Siegen, Germany

* The authors are also affiliated by Munich Aerospace Faculty, Munich, Germany

1) Introduction :



The collected EO data volumes are increasing immensely with a rate of several Terabytes of data a day. With the current EO technologies these figures will be soon amplified, the horizons are beyond Zettabytes of data. The challenge is the exploration of these data and the timely delivery of focused information and knowledge in a simple and understandable format.

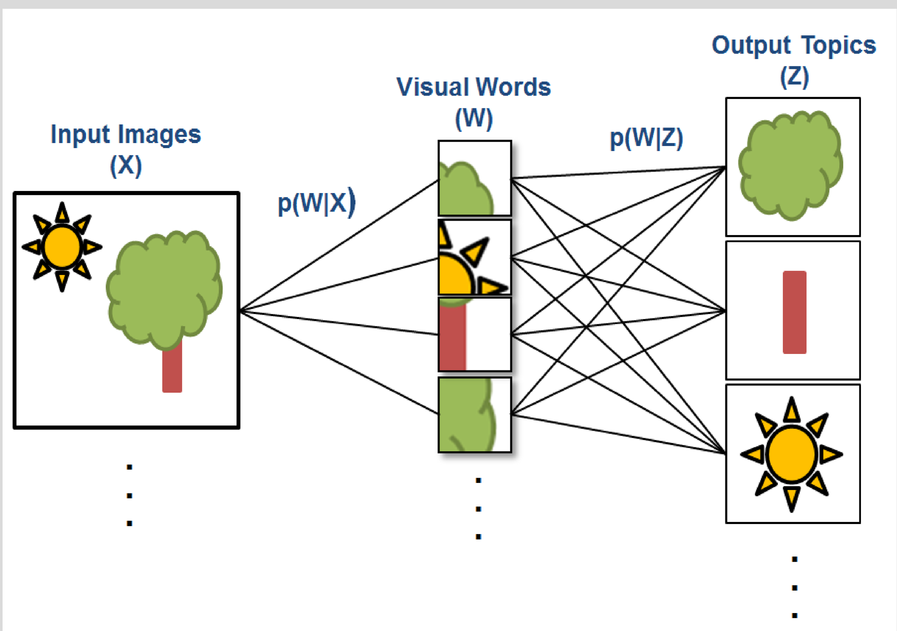
Although the main intention of content-based image retrieval and classification systems is to provide results which satisfy user's semantic queries, the provided results are still not user satisfactory. The fundamental reason is that users understand images based on their high-level contents (e.g., tree, water, building); however, most of the current CBIR systems perform on the primitive features of the images (e.g., shape, texture, color).

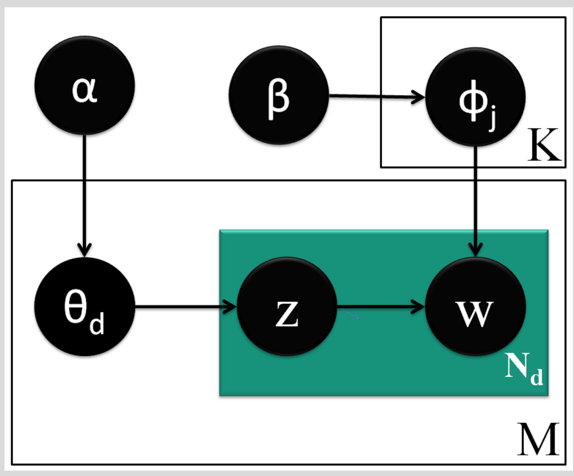
In this work, the high-level features of the EO images are discovered using a statistical topic model, so-called Latent Dirichlet Allocation (LDA). Then learning algorithms are applied to the high-level representation of the images. Experimental results demonstrate that high-level features can provide comparable results to BoW model; while the dimensionality of data is much lower in BoT model. Consequently, high-level features increase the scalability of learning systems as well as providing results relevant to the users' queries.

2) High-level feature representation:

Latent Dirichlet Allocation (LDA) :

LDA is a probabilistic topic model. This model assumes the images are distribution over set of visual words drawn from a certain vocabulary. Then it discovers the latent structure behind the given collection of images. The discovered latency is called **topic** which suppose to represent different **concepts** of the images. The generative property of the LDA allows to estimate the image collection using the discovered topics.



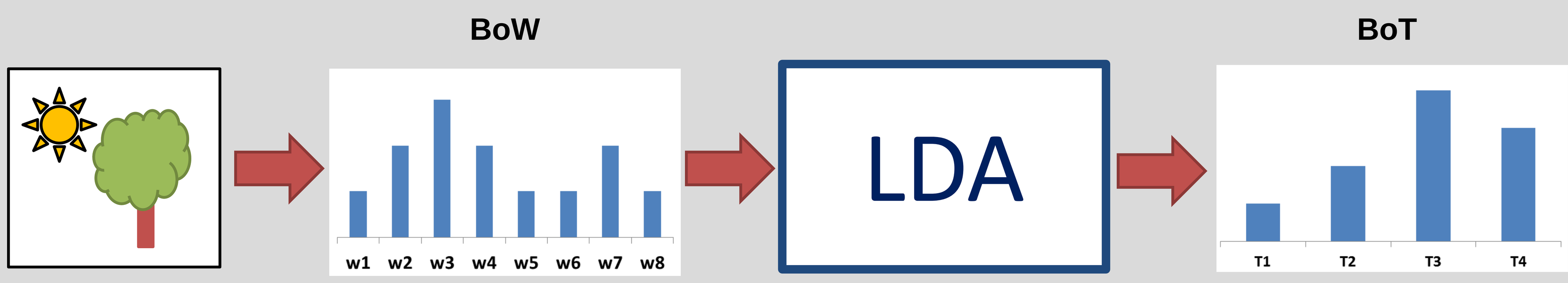


$$p(w_{nd}|\alpha,\beta) = \int p(\theta_d|\alpha)\left(\sum_{j=1}^{K} p(z_j|\theta_d)p(w_n|z_j,\beta)\right)d\theta_d$$

$$p(\theta_d|\alpha) = \frac{\Gamma(K\alpha)}{\Gamma^K(\alpha)}\prod_{j=1}^{K}\theta_{dj}^{\alpha-1}$$

Representing images by Bag-of-Topics :

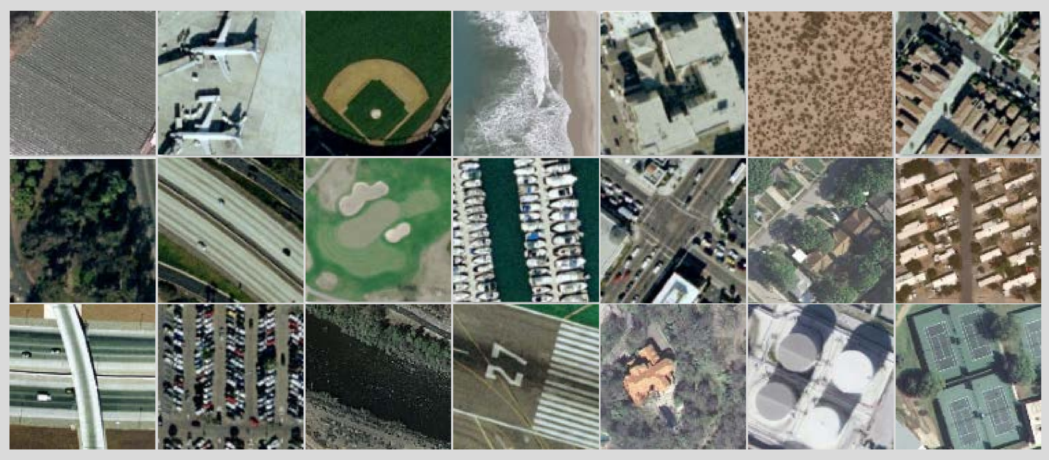
In order to discover the high-level features, first, LDA model is applied to the Bag-of-Word (BoW) model of images. BoW represents an image based on the occurrence of the primitive feature descriptors in the image. The discovered high-level features, the so-called topics, are then used to describe images, we call it Bag-of-Topics (BoT) model. BoT represents each image as a vector, so-called topic vector, where each dimension of this vector shows the frequency of a particular topic in the image. These topic vectors are then used in learning tasks.

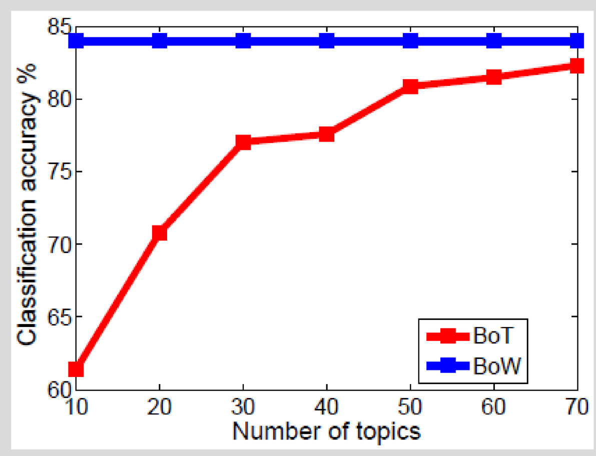


3) Experiments and results :

The level of the discovered high-level features highly depends on the number of topics in LDA model. In other words, more number of topics leads to lower-level features (i.e., detailed features) while less number of topics result in higher-level features (i.e., concept features). In order to evaluate the different levels of discovered features, LDA is run for different number of topics. Then images are represented by BoT model. Finally, the classification accuracy of SVM is used as the measure of informativeness of the high-level features compared to low-level representation of images (i.e., BoW). The evaluation is performed on a multi-spectral dataset, so-called UCMerced-LandUse, and a TerraSAR-X dataset.

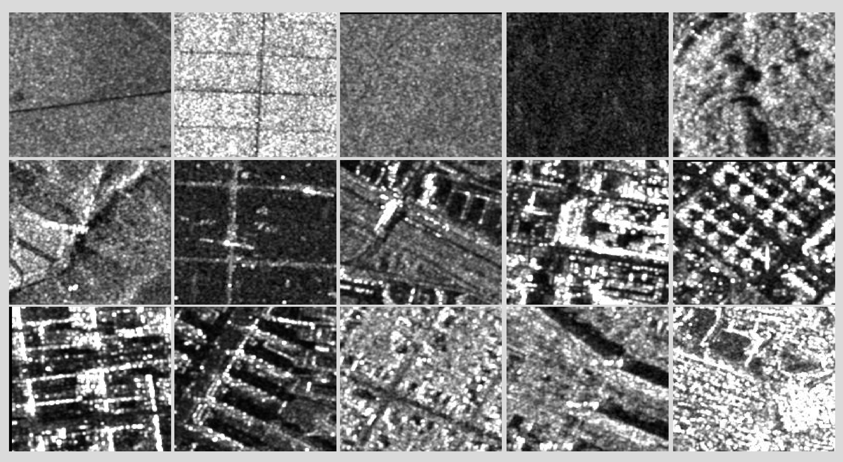
UCMerced_LandUse dataset:

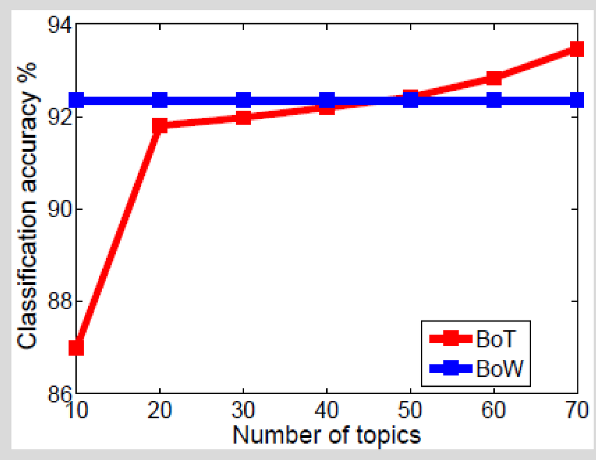




This dataset contains 2100 images grouped into 21 land-use scenes, where each class contains 100 images.

TerraSAR-X dataset:



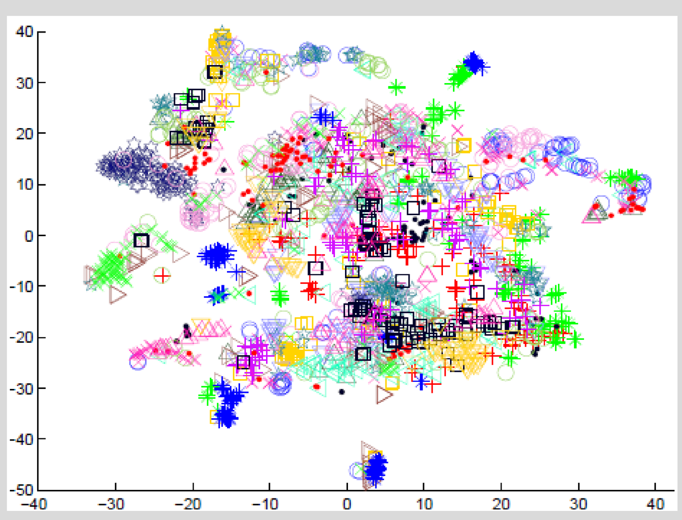


This dataset contains 3434 TerraSAR-X satellite images grouped into 15 classes. Top to bottom, left to right, the first two classes are Agricultural Fields. Then Grass Fields, Water Surfaces, Forests, Mountains, Flooded Fields, Highways, Industrial Areas, and the rest are different kinds of Urban Areas.

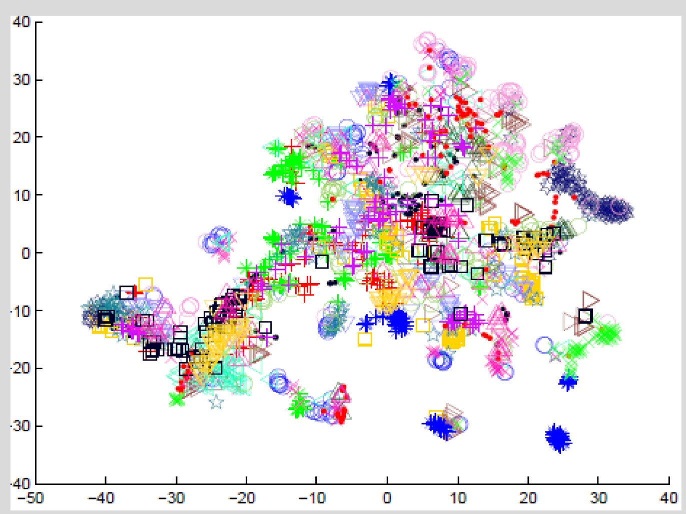
According to the plots, on TerraSAR-X data, for 50 topics, SVM obtains the same accuracy by BoT and BoW models; however, the dimensionality of feature vectors in BoT is one fourth of the one in BoW model. BoT can even outperform BoW model with sufficient number of topics. Thus, using high-level features of the data not only leads learning methods to provide relevant results to the users' queries, but also increase the scalability of the learning systems by reducing the dimensionality of the data.

4) Visualization:

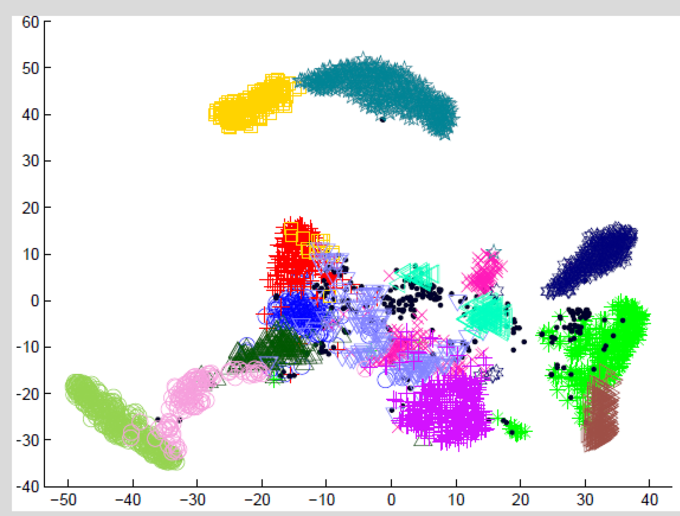
Since visualization is a massive aid to understand the structure of data, the feature spaces resulted by BoT and BoW representations of the images are visualized using t-Distributed Stochastic Neighbor Embedding technique. BoT and BoW provide similar distributions of classes in the feature space which leads to achieve similar classification accuracies. In these figures, colors represent various classes.



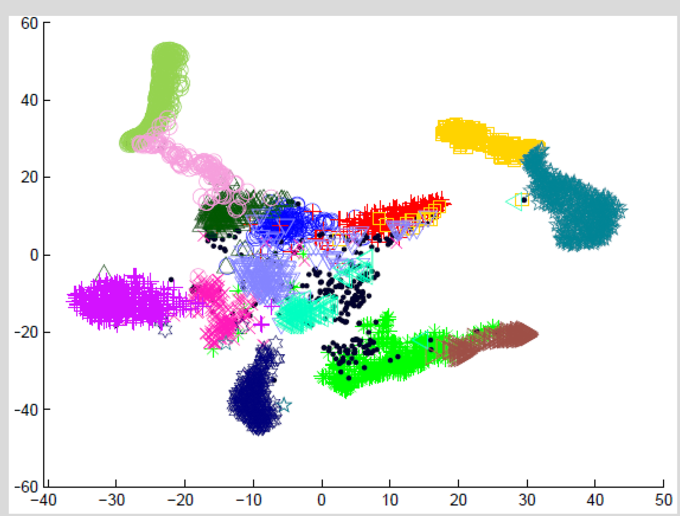
BoW representation of UCMerced data



BoT representation of UCMerced data



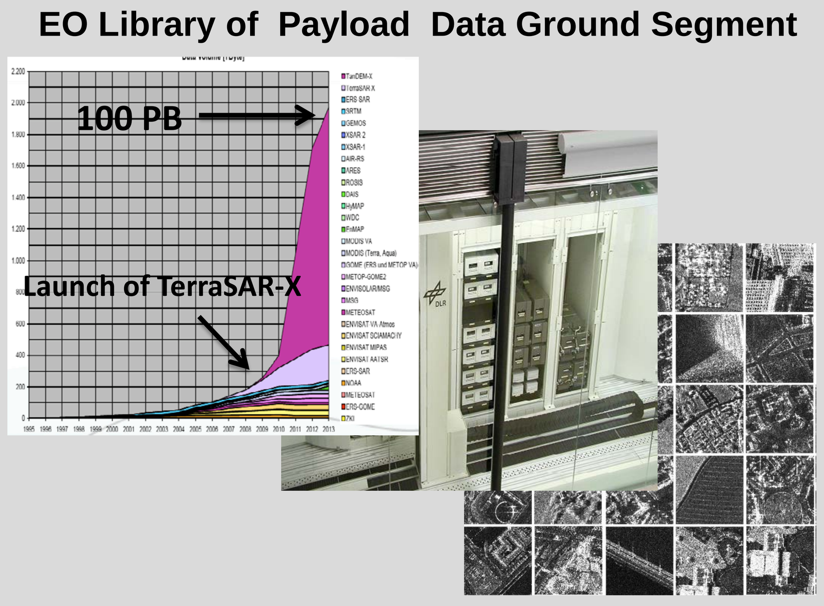
BoW representation of TerraSAR-X data



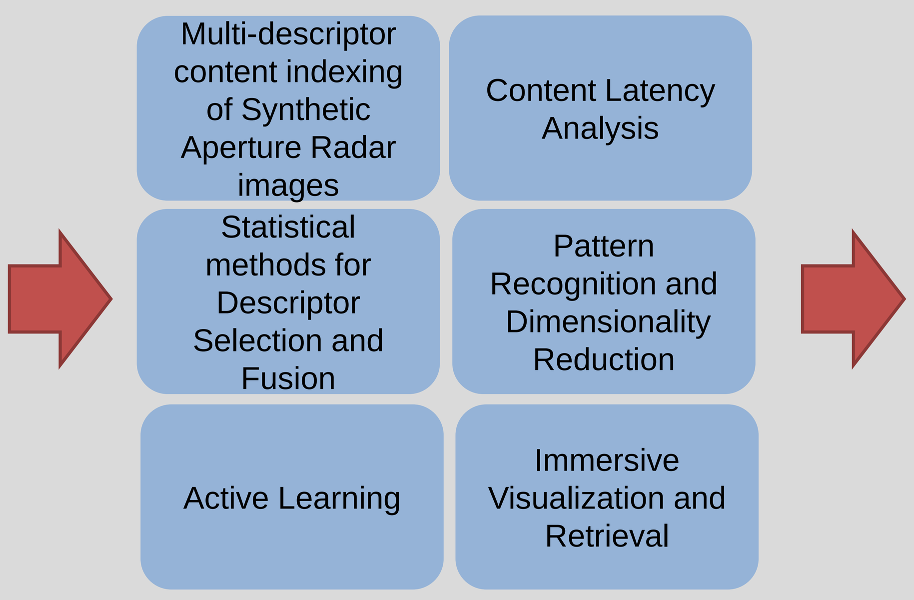
BoT representation of TerraSAR-X data

5) Future works :

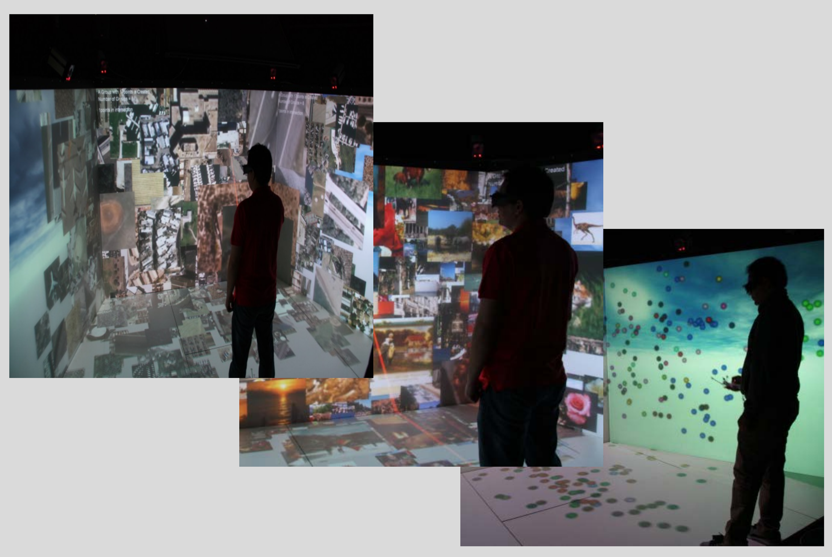
EO Library of Payload Data Ground Segment



Main Process module



Immersive information visualization (cave)



Immersive visual information mining: SAR data from the DLR EO Digital Library will be processed for **descriptor extraction**. Then descriptor space will be analyzed **and projected adaptively in 3D space**, **visualized in the CAVE**, jointly with **multi-modal rendering of the images** and their content. The analyst, immersed in the CAVE, will be enabled to **interact with the data content** using learning algorithms and navigate, explore and analyze the information in the archive.
The CAVE is located in Institute for Human-Machine Communication in Technical University of Munich, Germany.