Bayesian Enhancement of SVM based Search Engine

Mihai Costache, Mihai Datcu
mihai.costache@enst.fr, mihai.datcu@dlr.de
GET - Télécom Paris, CNRS UMR 5141

Introduction:
Context: High number of satellites acquired and acquire large number of Earth Observation (EO) images thus collecting huge volume of data.
Satellite image resolution per pixel is increasing (0.8 – 2.5 meters/pixel) thus the satellites collect a high volume of information.

Needs: Develop automatic indexing and annotation of satellite image archives and solve the problem of fast and accurate detection of structures within satellite images (classification).

Category based Search Engine
- Multiple levels of information representation needed for connecting the user’s interpretations with image content.
- At each level models and/or user's interest are inferred.
- Novelty: category Bayesian learning supported by SVM classification.
- Human Machine Communication with incremental / decremental learning.

System diagram

1. User level
2. Classes level
3. Categories level
4. Image level

Adaptation to the user
Learning / unlearning
Learning data models
Features extraction

Support Vector Machine – SVM
- Kernel type: Gaussian
- Distance function g(x)
- Training finds vector w and real b so that the class members be on the same side: y(w \cdot \Phi(x)+b)+1 for all i, i \in \{1, \ldots, N\}

Bayesian method for categories generation:
- Given the observed data D = \{(x_1, y_1), \ldots, (x_N, y_N)\}, data model M, model's parameter Θ
- Using Bayes rule for M and parameter Θ, models are inferred:

\[ p(\Theta, M | D) = \frac{p(D | M, \Theta)p(M)}{p(D)} \quad \text{and} \quad p(M | D) = \frac{p(D | M)p(M)}{p(D)} \]

- Model data selection based on Maximum A Posteriori (MAP)

Bayesian framework for SVM classes generation:
- Objective - Subjective

Results
1. Classification (structure detection / recognition)
- Detection and classification of structures within EO images. Example of building detection in SPOT5 image above Paris region.

Data simulation and evaluation measures
- Database: 8 categories; 100 examples/category (clouds, sea, forest, desert, city, fields, airports, village houses).
- Textural features: Quadrature Mirror Filters (QMF).
- Evaluation measures: Precision-Recall and Mean Precision

Main Precision
- A relevant images; B retrieved images

\[ P = \frac{A \cap B}{|B|}, R = \frac{A \cap B}{|A|} \quad \text{and} \quad p = \frac{1}{N} \sum_{i=1}^{N} p(i) \]

Search and indexation of image areas in EO SPOT5 images
1. Query and system’s suggestion
2. User’s indication: relevant / irrelevant
3. System’s suggestion
4. User’s indication: relevant / irrelevant
5. Results

Search example in SPOT5 database.
- Airports and
- Village houses.

Performances evaluation

<table>
<thead>
<tr>
<th>Search engine</th>
<th>Speed of learning</th>
<th>Mean precision p</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM based</td>
<td>8 RF steps</td>
<td>0.97</td>
</tr>
<tr>
<td>Category based</td>
<td>1 RF step</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Percent of good category selection: 98%

• Category Bayesian based learning enhances the search capabilities by speeding the learning process.
• SVM supports classification of EO scenes.