Information theoretical approach for domain ontology exploration in large Earth observation image archives

Mihai Datcu, Mariana Ciucu
DLR Oberpfaffenhofen

IGARSS 2004
Objectives: a prototype system of a next generation architecture to help the users to gather relevant information rapidly, manage and add value to the huge amounts of historical and newly acquired satellite data-sets.

Data: ERS, Landsat, Ikonos, MERIS, E-SAR, DEDALUS...

Evaluators and users: ESRIN, EUSC, NERSC, DLR

Output: prototype system

Train a Label

Select an image to define a label:

- Define a cover-type label on a Landsat TM image of Laren
- Define a cover-type label on an E-SAR image of Laren
- Define a cover-type label on a MERIS image of Laren
- Define a cover-type label on a DEDALUS image of Laren

Inventory of Cover-types

Currently there exist two users to recognize two default labels:

- The Atlantic Ocean
- The tropical rainforest

They can make two non-overlapping sets of all labels that have been defined with the annotation tool. Note that the two inventories offer two different types of overview. The complete inventory offers an overview of all labels.
We have to design systems to enable people to analyse TeraBytes of data!

People have normally trouble in caching more than 7 items at a time
HIERARCHY OF INFORMATION REPRESENTATION

- $D \in$ level 0: image data
- $f \in$ level 1,2: image and meta features
- $\omega_i \in$ level 3: image classification/segmentation
- $A_{\nu} \in$ level 4: user-specific semantics

Parameter estimation and unsupervised clustering are shown in the diagram. Inventory categories include mountains, drainage areas, snow, ice, and fields. The conditional probability $p(\omega_i | A_{\nu})$ is also depicted.
Clustering and coding

Morse alphabet

E .
I ..
A _.
O ___
Semantic coding

- **Spectral**
- **Texture**
Image semantics

Pixel
- Spectral
- Texture
- Line/Edge

Object
Components

Objects
and
Scene
Components

Region
- Classification
- Segmentation

Contextual
Structures

"Semantic" abstraction in signal models

"Semantic" abstraction in user ontology
Advanced communication

Image:
\[ \{ I_k \} \]

Cluster:
\[ \{ S_j \} \]

Semantic labels:
- CITY
- VILLAGE
- ROAD
- RIVER
- FOREST

\[ p(\text{L} | \ ?) \]

\[ p(\text{L} | I) \]

\[ p(\ ? | I) \]
- applicable to a number of fields
- knowledge acquisition and sharing within user communities
- multi-type data handling
- knowledge communication
Semantic grouping - method

- Supervised Bayesian classification:
  \[ p(L_0 \mid \omega_i) > p(L_\mu \mid \omega_i), \forall \nu \neq \mu \]

- Inference from data to aggregated labels:
  \[ p(\Lambda_\tau \mid D) = p(\Lambda_\tau) \sum_\nu \frac{p(L_\nu \mid \Lambda_\tau) p(L_\nu \mid D)}{p(L_\nu)} \]

- Interactive learning the probabilistic link \( p(L_\nu \mid \Lambda_\tau) \):
  \[ p(L_\nu \mid \Lambda_\tau, \psi) = \psi_\nu \quad \psi = \{\psi_1, \ldots, \psi_c\} \]
  \[ p(\psi) = \Gamma(c) = (c - 1)! \]
  \[ p(\psi \mid T) = \text{Dir}(\psi \mid \alpha) \]
  \[ p(L_\nu \mid \Lambda_\tau) = \frac{\alpha_\nu}{\sum \alpha_\nu} \]
In the system there may be *labels* with different names and *with the same information content* (the same meaning)
Controlling the Dynamic Indexing

Solution: similarity measure, Kullback-Leibler divergence

\[
KL(L_1 \mid L_2) = \sum_i \sum_j p(\omega_{ij} \mid L_1) \cdot \log \left( \frac{p(\omega_{ij} \mid L_1)}{q(\omega_{ij} \mid L_2)} \right)
\]

Example

The initial label for two classfiles

The labels from database for the same classfiles

After Kullback-Leibler procedure the sorted labels list is: 2, 1, 6, 4, 5, 9, 7, 8, 3
For the same label we can have many realisations, with the same features (like L2).

When a user is training a new label on an image, the system may try to identify the other labels closest to it (KL between 0 and 1) and give the user the possibility to choose an old label or to define a new one.
**Problem**

- User needs to learn from another user
- We can predict what the user wants.

**Solution**

Similarity measure -> Kullback-Leibler divergence between two ontologies

\[
KL'(r_1, r_2) = KL(r_1, r_2) + \sum_{i=1}^{n} p(k_i \mid r_1) * KL'(k_i, p_i)
\]

Where,

\[
\begin{align*}
    r_1 & \text{ – root node of sub-trees for one of users} \\
    r_2 & \text{ – root node of sub-trees for other user} \\
    k_i & \text{ child nodes of } r_1 \\
    p_i & \text{ child nodes of } r_2
\end{align*}
\]
Domain Ontology Search

Examples

the entropy from the root node = hd_cloud6 is 7.232032711267443
the entropy from the root node = md_cloud is 8.75741191905023
KL (hd_cloud6,md_cloud) 0.7136991789787241

the entropy from the root node = ac_acqua_veloce is 8.134289966218503
the entropy from the root node = hd_water2 is 6.1539319369981
KL (ac_acqua_veloce,hd_water2) 0.6401859742548874
Conclusions

The system helps users by

• cumulative learning
• incremental acquisition of knowledge
• building on knowledge acquired in earlier steps
• take advantage of what was learned before
• share knowledge