A KNOWLEDGE BASED FRAMEWORK FOR THE DETECTION OF MEASUREMENT UNCERTAINTIES IN DERIVED SEA SURFACE CURRENT FIELDS

Benjamin Seppke[1], Leonie Dreschler-Fischer[1], Michael Wessel[2] and Martin Gade[3]

[1] University of Hamburg, Department of Informatics, Cognitive Systems Laboratory, Hamburg, Germany

[2] Hamburg University of Technology, Institute for Software, Technology, and Systems, Hamburg, Germany

[3] University of Hamburg, Centre for Marine and Atmospheric Sciences, Institute of Oceanography, Hamburg, Germany

E-mail: seppke@informatik.uni-hamburg.de

INTRODUCTION

In this paper we present a conceptual framework that uses a knowledge based description logic approach to decide whether or not a computed motion vector corresponds to a sea surface current. For this decision we need to take different sources of knowledge into account. In [7] we have shown that low level image processing algorithms can be used to improve the derived motion fields by means of smoothness. However there are still cases where this correction may result in unwanted motion measurements that do not correspond to the sea surface currents. To solve problems like this, image interpretation systems have proven to be adequate [5].

DERIVATION OF HIGH-RESOLUTION SEA SURFACE CURRENTS FROM SATELLITE IMAGERY

Currently there exist two well-known families of algorithms for the computation of sea surface currents from at least to satellite images of one region: the feature-based and the Optical Flow-based approaches. Both have in common that there have to be current tracers (e.g. sea surface films) visible in the images to be analyzed. The local sea surface currents cause the motion of these tracers, which allows an indirect and high-resolution measurement of the currents [6].

Before the feature-based local approaches can be used, one has to find the features of interest (e.g. algae signatures) in at least one satellite image a priori. After the detection of features different feature matching methods can be used (e.g. fast cross-correlation or shape-context matching). These methods usually assign a confidence value to each matching.

The Optical-Flow-based approaches do not depend on knowledge about specific features. They result in a global motion field that represents the displacement of one satellite image relative to the other. The Optical Flow methods do not explicitly assign a confidence value to each motion vector.

In this paper, we concentrate on results of feature-matching approaches based on surface film tracking on synthetic aperture radar (SAR) satellite imagery. The results of these approaches provide explicit information about the uncertainties of each motion measurement. Although the feature-matching methods are well known, and are highly optimized, there is always some chance of a measurement error. We can divide the error into two cases:

1. A high correlation value of a motion that does not correspond to a real sea surface current and
2. A low correlation value of motion that corresponds to a real current element.

Due to the matching strategy, the first case occurs more often than the second one and cannot be compensated by low-level post-processing steps [7]. The origin of these erroneous motion measurements is the tracking of unsuitable features. These features can be ship wakes or wind induced surface anomalies that result in similar signatures in the SAR images. To solve this ambiguity, we need to classify the motion target with scene specific knowledge.

SOURCES OF KNOWLEDGE

We distinguish between two sorts of knowledge: dynamic factual knowledge about the scene depicted in the current image that is analyzed and static interpretation knowledge that allows the automated reasoning over the facts. Examples of factual knowledge about the scene comprise wind information, tidal information, position of ships and waterways, chlorophyll ratio, and sea surface temperature information for each selected feature at each spatiotemporal point.

Moreover there are static interpretation rules given by domain experts

- Biogenic surface films are often associated with a locally enhanced chlorophyll-a concentration
- SAR signatures of wakes and surface films may look similar.
- Wind and tidal forces mainly affect surface currents.

The factual knowledge changes dynamically for each measurement. This may cause a huge amount of data, which cannot be represented inside the description logic efficiently. That is why there is a strong need of a multi-layer architecture with bottom-up reasoning for the task of detecting unreliable measurements.

The interpretation knowledge is often compact enough to be represented inside such a system. The representation of this “higher knowledge” is independent from the factual knowledge. Knowledge engineers can revise it according to domain experts and their knowledge [3].

THE DESCRIPTION LOGIC SYSTEM RACERPRO

Description logics (DLs) are a family of knowledge representation languages, which originated from early attempts in the 1970s to model knowledge with class- or concept-based knowledge structures, i.e., Minsky’s Frames, and the so-called Semantic Networks. Nowadays, DLs provide the semantic basis for the Semantic Web (e.g., OWL DL is basically a description logic). Most contemporary DLs can be considered as subsets of first-order logic, and hence, the inference services offered by the corresponding systems are well-defined. Knowledge in DL systems comes in two disguises: class- or concept-based knowledge, and individual-specific knowledge. Whereas the TBox models conceptual knowledge in terms of concept specialization and concept definition axioms, the assertional box, or ABox, contains a set of instance (or class) assertions and relationships between these instances. Standard DLs only support binary relationships, which are called roles. Together, TBox and ABox is called a knowledge base.

In this work, the DL system Racer is used [1]. RacerPro implements the expressive description logic SHIQ(Dn), which offers transitive, functional and inverse roles, role specialization hierarchies, reasoning with datatypes (e.g., strings, reals, integers, booleans), and some additional concept constructors (e.g., the qualified number restrictions of OWL2). Racer was the first system of a new generation of highly optimized

DL systems [1] that also supported ABoxes. Racer offers many advanced proprietary features, such as first-order (grounded or) epistemic queries, rules, programmatic "server-sided" scripting, extensibility, and some innovative inference services (such as abductive query answering). After more than 10 years of continues improvements, RacerPro is one of the fastest ABox reasoning system nowadays whose scalability for certain standard ABox benchmarks has been shown recently [2]. As such, it is an ideal basis for knowledge-intensive applications which require ABox reasoning and ABox query answering and was thus selected for this research. In addition, Racer has proven to fit well for reasoning by means of computer vision scene interpretation [4].

KNOWLEDGE BASED DETECTION OF VALID SEA SURFACE CURRENT MEASUREMENTS

We developed a prototypical framework that models the domain expert's knowledge inside the TBox and uses the derived currents in conjunction with the other factual knowledge about the scene as ABox contents (Figure 1). To transfer the quantitative information into the ABox, we use an abstraction on the middle layer of the framework that maps the quantitative values to qualitative symbols. One example of a very simple TBox could be like this:

```
(all-disjoint water land)
(all-disjoint high-chlorophyll  medium-chlorophyll low-chlorophyll)
coastal <-> (and (some next-to water) (some next-to land))
valid-current -> (or (not (some next-to ship)) (not (some next-to waterway)))
valid-chlorophyll-amount -> (or high-chlorophyll medium-chlorophyll)
valid-current -> (all next-to valid-chlorophyll-amount)
```

The abstraction from quantitative to qualitative values can be observed e.g. in the case of the chlorophyll amount. Instead of model floating-point values we use three categories: high, medium and low. Another abstraction can be found at the spatial neighborhood that is modeled as role "next-to". Note that the system can derive knowledge about a "coastal" relationship using a TBox equivalence rule.

We now present an example for an ABox:

```
valid-current(motionA)
next-to(motionA, water)
next-to(motionA, land)
next-to(motionA, high-chlorophyll)
```

In this ABox, we assume that the measured motion is a valid current. We perform a so-called ABox consistency check, to finally get the answer to our question: Is the measured motion vector representing a valid sea surface current or not? Please note that the example above is a very simple one, just to demonstrate the basics needed for our approach. For this example, the measured motion "A" represents a valid current given the TBox above, because it is next to some higher chlorophyll amount than usual, and is not next to some waterway or ship. On the other hand the system derived that it belongs to a coastal region.

CONCLUSIONS

We have presented the main concepts of a flexible knowledge based framework and have given a first example of an application. A conceptual diagram is given in (Figure 1). Due to the multi-layer architecture of the framework and the RacerPro DL system, it is highly scalable and can therefore be used for reasoning in many areas of remote sensing. One of the key features is the separation of knowledge into static expert knowledge and highly dynamic knowledge. We implemented a running prototype, which results in promising but yet preliminary results [5]. The next steps are the integration of other knowledge sources and of more expert knowledge to improve the automatic reasoning.

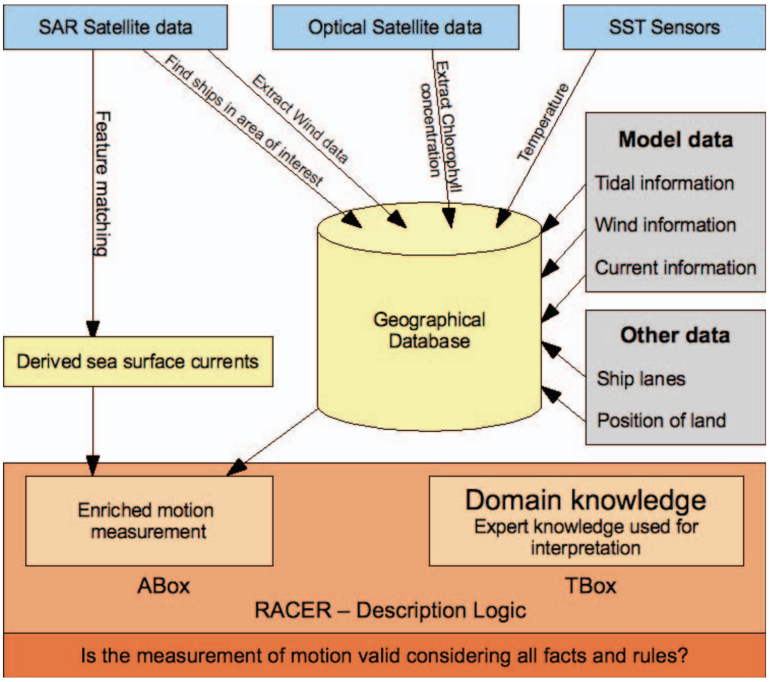

Figure 1: Diagram of the conceptual Framework showing the bottom-up reasoning approach. The geographical database represents the middle (integration) layer.

LITERATURE

[1] Haarslev and Möller, 2001. RACER System Description. In *International Joint Conference on Automated Reasoning, IJCAR'2001*, June 18-23, Siena, Italy, pages 701–705. Springer- Verlag.

[2] Haarslev and Möller, 2008. On the Scalability of Description Logic Instance Retrieval. Journal of Automated Reasoning, 41(2):99–142.

[3] Kudryavtsev, Akimov, Johannessen and Chapron, 2005. On radar imaging of current features: 1. model and comparison with observations, *J. of Geophys. Res. 110, No. C07016*

[4] Neumann and Möller, 2006. On Scene Interpretation with Description Logics. In: *Cognitive Vision Systems,* H.-H. Nagel and H. Christensen (Eds.), Springer, LNCS 3948, 2006, 247-275

[5] Panic, 2009. *Räumliche Constraints zur wissensbasierten Interpretation von Radarbildern der Meeresströmung*. Diploma Thesis, Dept. Informatics, University of Hamburg

[6] Seppke, Gade, Dreschler-Fischer, 2009. Sea Surface Current Fields in the Baltic Sea Derived From Multi-Sensor Satellite Data. In *Proceed. ISPRS Workshop 2009*

[7] Seppke, Gade, Dreschler-Fischer, 2010. *The Role of Spatial Constraints in the Derivation of Mesoscale Sea Surface Current Fields from Multi-Sensor Satellite Data, Proceed. IGARSS 2010.*