Same world, different words: augmenting sensor output through semantics

Anne-Laure Jousselme(1), Valentina Dragos(2), Anne-Claire Boury-Brisset(1), Patrick Maupin(1)

(1) Defence R & D Canada - Valcartier
Quebec (Qc), Canada
anne-laure.jousselme@drdc-rddc.gc.ca
anne-claire.boury-brisset@drdc-rddc.gc.ca
patrick.maupin@drdc-rddc.gc.ca

(2) ONERA-The French Aerospace Lab
F- 91761 Palaiseau, France
valentina.dragos@onera.fr

Abstract - This paper addresses the general problem of bridging the gap between sensor output and the structured reports needed and handled by end users in several practical applications. Sensors perform measurements used to characterize world entities according to their physical features, such as weight or speed or even more abstract features extracted from signal statistics such as spectral higher-order moments. The sensor outputs are usually provided to a human operator, allowing him to build, communicate and share with other humans a picture of a given situation. However in practice, humans often handle entities by considering their functions as it is the case in military intelligence reports rather than limiting their attention to physical or statistical properties. A possible solution to such a mismatch is to augment appropriately the vocabulary used to deliver the sensor outputs, allowing the operators to manipulate it efficiently in practical applications such as military intelligence analysis and dissemination. This paper presents a semantic approach developed for relating two complementary descriptions of data, one built upon experimental data collected using the SASNet system, the other being a military intelligence ontology called ONTO-CIF. The proposed method is based on ontology mapping and allows us to augment sensor data in order to cope with user needs.

Keywords: sensor network, military intelligence, ontology alignment, target classification.

1 Introduction

Information provided by sensor networks is being increasingly used especially in relation to security, defence and military application fields. Sensors produce a large amount of data possessing natural heterogeneity, as different features are considered. Those features represent physical properties whose values are measured by the sensors and returned to the end user for achieving his task. However in several applications, human operators handle and describe entities according to their functions or capabilities rather than just considering their physical or statistical properties. A semantic gap thus appears between sensor output and data as required from an operator point of view. One of the objectives of the work is to provide mechanisms in order to augment sensor output and classification results, so that these data can be presented to human operators in their operational language to facilitate interpretation, make further inference and ensure interoperability. We propose a semantic approach, using ontologies to model knowledge according to each point of view, and an ontology mapping process to create a bridge between two different representations of the same world.

The paper is structured as follows. In section 2, we describe the application context, as well as the sensor system used (in section 2.1) together with the overall sensor network architecture (in section 2.2). We also introduce the domain ontology used in support of this work in section 2.3. In section 3, we then present the classification problem based on sensor data, and show the proposed approach to attach semantics to the sensor output and provide meaning to facilitate interpretation by human operators. In section 4, we describe the mapping mechanism performed between the sensor output and the ontology entities in order to relate the two. Related approaches are presented in section 5 and finally, conclusions and future work perspectives end this paper.

2 Application context

2.1 The SASnet node

The Self-healing Autonomous Sensor Network (SASNet) project [1] is a 5 years (2006-2011) technological demonstration project conducted by a team of researchers from the Communications Research Center, Canada and Defence Research and Development Canada research center, Canada. This research project aims at testing in realistic operational situations an unattended sensor network together with its electronic and software components. With an emphasis on self-healing capacities and advanced communication protocols based in particular on logical neighbourhoods, the project focuses on the
design of efficient detection and classification techniques, in terms of precision but also of energy consumption. A more specific aim of this project is to develop a data classification approach adapted to surveillance-related queries made by operators of a network architecture composed of management nodes, communication relays as well as sensors used to monitor and protect a zone of interest. These sensors send contacts acquired locally to the communication relays, which before sending information to the human operators will fuse detection and identification reports. The users of this information are either men in the field, vehicles involved in operations or a commander in a tactical or operational headquarter (see Figure 2 for details). One of the principal problems resides in the management of surveillance queries involving a sequential detector/classifier system composed of one-class classifiers specialised in the detection of events of interest as well as in the classification of humans and vehicles.

Another problem, tackled by the present paper, is the translation of the sensor-level detection or identification reports, into messages formulated using some standard ontology for intelligence dissemination purposes. The latter, called ONTO-CIF, will be described in section 2.3.

Figure 1: The SASNet sensor node.

The SASNet sensor node is briefly presented in Figure 1 while more details about its 4 sensing modalities will be given in section 3.1.

2.2 The SASNet system architecture

As shown in Figure 2 the SASNet network is composed of 3 levels, each composed of nodes having different properties and roles. Management nodes can be either an headquarter facility or a command post in an operational theatre from which information queries originate. These management nodes receive in return detection and identification reports from either relay nodes or from the sensing nodes directly.

Relay nodes, having a more important energy budget, can be equipped with more evolved sensors such a infrared cameras or other actuators that can be controlled manually by operators of the management nodes or triggered by automated commands emitted by the sensing nodes. The sensing nodes, like the relay nodes, are equipped with a set of 3 pyroelectric motion sensors, as well as seismic, magnetic and acoustic sensors, but their task is limited to the detection and identification of targets of interest.

Practical constraints include the fact that the users of this system wish to maximise, on the one hand, the number of true positives detections as well as, on the other hand the number of good identifications, while minimising the energy consumption.

Figure 2: The SASNet sensor node architecture. Vehicles, personnel in the field and HQ can interact with the sensor system over large areas of operations.

Not surprisingly this requirement relates not only to the goodness of the detection/identification algorithms design, but also to the ability of this surveillance system to interface efficiently with the human operators in order to minimize the overall communication traffic between the sensor network and its users. In other words, that means letting the users task precisely the sensor network and allow the latter to transmit only useful information while allowing the users to understand as precisely as possible the meaning of the messages received.

For us, a good deal of this interaction quality is directly related to the quality of the mapping between the world of the sensor and the world of the sensor network users. This ontology alignment problem is the focus of the present paper.

In the surveillance problem targeted in this paper just like in other problems in which objects classification is involved, parenthood relationships exist. These relationships can be expressed using a tree representation for which branches express physical or functional relationships while the leaves represent types of objects of interest. This type of structure ressembles to formal ontologies used in the military intelligence domain. The work presented here is intended to complement a convoy protection scenario published earlier in [2]. In this scenario the objects of interest belong to the classes Human (civilian, military or security personnel) and Vehicle (diverse civilian and military vehicles).

The energy budget gain that can obtained by an efficient communication protocol managing intelligence messaging can be very important since radio communication involved at the sensor of relay level can cost as much as five to 10 times the energy consumed for local calculations. At the sensor network level, radio communication has thus a very important effect of life expectancy of the components and the logistics costs associated for the maintenance of such an infrastructure.
2.3 Describing a world of functions: the ONTO-CIF ontology

According to [7], an ontology is defined as a formal and explicit specification of a shared conceptualization. The ONTO-CIF ontology (see Figure 3) was built in order to have a general description of the military intelligence application field. It provides a standard model of various entities of this domain, along with the set of relations holding between them. The METHONTOLOGY methodology was used to build this ontology by exploiting several knowledge supports.

According to METHONTOLOGY, five steps are needed to build ONTO-CIF. The specification step allows us to identify the purpose of the ontology construction and to explain its intended use. ONTO-CIF was created in order to describe the military in telligence field according to a functional point of view. Therefore, we used as knowledge support several documents created by domain actors with emphasis on functional descriptions of entities. The goal of the overall process was to obtain a shared standard conceptual model of our field, providing a basis for the development of further automatic reasoning mechanisms, such as evaluating information credibility. The conceptualization step identifies field entities along with their relations. During this step domain concepts, relations between them and axioms are modeled. By exploiting knowledge sources, we have identified 58 concepts, clustered in 6 main categories, corresponding to: entities (i.e. vehicles or persons), locations (i.e. geographical area), structure (whether plain or hierarchical, for instance), status of entities (natural or man-made), goals (or function of entities) and events (significant actions involving several entities). We also identified 55 relations between concepts, expressing specializations (Organization, Social Grouping), compositions (Organization, Person), equivalences (Installation, Place) and field particular relations such as has-goal (Organization, Function).

Axioms of ONTO-CIF highlight 4 pairs of equivalent classes, 6 pairs of disjoint classes and 4 pairs of inversed roles. The formalization and implementation of ontology concern the choice of formalism and language to represent the ontology. We chose description logics and OWL-DL, a sub-language of OWL [22], as it offers a good compromise between expressiveness and computability. The last step of the construction process is ontology maintenance and it will be done all ontology life cycle long.

3 Describing a world of features: sensor data

3.1 SASNet data

The sensor network considered is the SASNet sensor network, designed at DRDC-Valcartier [1]. Each sensor node is composed of 4 sensing modalities: one PIR (motion detector) one acoustic sensor (noise detector), one seismic sensor (vibration detector) and one magnetic sensor (ferric metals detector). On each of the 4 signals, a series of features is computed as described in Table 1.

Table 1: List of extracted features for each of the 4 sensing modalities (PIR, acoustic, seismic, magnetic).

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>Event duration</td>
</tr>
<tr>
<td>maxP</td>
<td>Power at CPA</td>
</tr>
<tr>
<td>maxA</td>
<td>Maximum amplitude</td>
</tr>
<tr>
<td>totE</td>
<td>Total energy</td>
</tr>
<tr>
<td>avgF</td>
<td>Average frequency</td>
</tr>
<tr>
<td>1/3oct1...15</td>
<td>Thirds of octave</td>
</tr>
<tr>
<td>Samp1...4</td>
<td>Spectral amplitude statistics</td>
</tr>
<tr>
<td>Sshap1...4</td>
<td>Spectral shape statistics</td>
</tr>
</tbody>
</table>

The acoustic sensor helps mainly in distinguishing between motorized and non-motorized objects. Among the class of non-motorized objects, the seismic sensor captures the intensity of soil vibrations and thus helps to distinguish between a jogging and walking pedestrian. The seismic sensor can also help classifying between either tracked or wheeled vehicles. The magnetic sensor itself gives an indication of the mass of ferric metals contained in a passing by object and thus helps to distinguish bikes from pedestrians or within a pedestrian group, militaries from civilians. Generally a single sensing modality will not be sufficient to design an efficient surveillance system aimed at the detection and identification of passing by objects, and a combination of them will rather be required. For instance, a high value for the three modalities
generally corresponds to a truck (high noise, high vibrations, and high magnetic response). Also, a motorbike will generate a high level of noise but low vibrations and a low magnetic response. And so on for the other objects.

Note that the above discussion is based on the computed signal power of the three modalities although other less intuitive features have been extracted to train the classifier such as frequency-based features (see Table 1).

An experiment has been conducted (described in [4]) in which 17 types of vehicles transited through the sensor network for a total of 278 recorded events. Table 2 lists the 17 types of objects ranging from pedestrians to trucks, together with their class labels and the corresponding number of events of each type. The resulting dataset \( X \) will be used in the following as a training dataset for the classification purposes.

Table 2: 17 classes of the training dataset [4].

<table>
<thead>
<tr>
<th>Class labels</th>
<th>Title</th>
<th>Nb of events</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Pedestrian - Jogging</td>
<td>49</td>
</tr>
<tr>
<td>11</td>
<td>1 Pedestrian - Walking</td>
<td>20</td>
</tr>
<tr>
<td>12</td>
<td>2 Pedestrians - Walking</td>
<td>7</td>
</tr>
<tr>
<td>13</td>
<td>3 Pedestrians - Walking</td>
<td>6</td>
</tr>
<tr>
<td>14</td>
<td>4 Pedestrians - Walking</td>
<td>6</td>
</tr>
<tr>
<td>20</td>
<td>Bicycle</td>
<td>38</td>
</tr>
<tr>
<td>40</td>
<td>Gator John Deere SATUV</td>
<td>26</td>
</tr>
<tr>
<td>50</td>
<td>Subaru Legacy 4x4</td>
<td>17</td>
</tr>
<tr>
<td>51</td>
<td>Chevrolet 2500 HD 4x4</td>
<td>24</td>
</tr>
<tr>
<td>52</td>
<td>Chevrolet Express Van</td>
<td>13</td>
</tr>
<tr>
<td>53</td>
<td>Jeep Cherokee</td>
<td>12</td>
</tr>
<tr>
<td>54</td>
<td>Chrysler Sebring</td>
<td>12</td>
</tr>
<tr>
<td>55</td>
<td>Chevrolet Silverado</td>
<td>13</td>
</tr>
<tr>
<td>60</td>
<td>Cargo/ Deliver Truck 15'</td>
<td>13</td>
</tr>
<tr>
<td>61</td>
<td>Low Boy Van Trailer 50'</td>
<td>6</td>
</tr>
<tr>
<td>62</td>
<td>Minibus 20'</td>
<td>6</td>
</tr>
<tr>
<td>63</td>
<td>Dump truck (empty)</td>
<td>10</td>
</tr>
</tbody>
</table>

3.2 Ascending hierarchical classification

We adopt an unsupervised approach for the classification. On the training dataset \( X \) described in Table 2, an ascending hierarchical clustering algorithm is applied resulting in a tree whose leaves are the 278 individual objects of the dataset and whose root represents the whole set \( X \). Pairs of objects are successively built based on distances in the feature space of the objects. So, each node in the tree constitutes a cluster of training objects, the higher in the tree, the more populated the cluster will be. Figure 4 shows the resulting hierarchical clustering, we will denote by \( C \), built upon a standard Euclidean distance with complete linkage. The tree has been cut so that (1) the number of clusters (leaves in this tree, represented as circles) is less than 10, and (2) the number of objects in the clusters (reported in the circles) is less than 50. The number of individuals assigned to each cluster is displayed in circles.

Figure 4: Tree resulting from the ascending hierarchical classification with complete linkage based on a Euclidean distance.

This unsupervised classification can be seen as the level 0 of semantics where only distinctiveness between objects is considered. The objects are classified according to a proximity criterion based on their measured features so that two objects are assigned to the same cluster if the distances between their respective features vectors are small. The only information available at this step is the number of objects in each cluster. During the classification phase, each new observed object is then compared (through its features vector) to the training clusters (either through a barycentre or through the individual themselves) and assigned to one particular cluster.

3.3 Augmenting semantics

To the clustering of Figure 4, some semantics can be added by labelling the objects in the clusters with the available ground truth information, i.e the labels of the training dataset (see Table 2). The labelled cluster is shown in Figure 5 where the 17 kinds of objects have been considered. Each cluster can then be qualified according to the proportion of objects of pre-defined classes. For instance, it can be established that Cluster #1 (first from left in Figure 5) is composed of 44% of Jogging-Pedestrians, 35% of Bikes, 13% of Walking-Pedestrians, 6% of Subaru Legacy 4x4 and 2% of Chrysler Sebring.

This approach can be qualified as “semi-supervised” as we use the labels to characterise each node of the hierarchical clustering.

Another level of semantics to be possibly added to this unsupervised hierarchical classification is the structuring of information among class labels into a taxonomy. More precisely, the 17 classes considered can be naturally structured according to the discriminative properties of the three modalities of the sensor node (the PIR serves as a detector of any moving object), resulting in what we call the “experimental taxonomy” \((ET)\) of Figure 6.
The enriched clustering is shown in Figure 7 where 6 classes have been considered. The repartition of the objects among the 6 classes within each cluster is displayed within the circles and detailed in text, and the class with the highest likelihood is displayed in bold. For instance, Cluster #1 contains 56% of Pedestrians, 36% of Bikes and 8% of small cars. These values can further be used to quantify the uncertainty relatively to a new incoming target.

3.4 Measuring uncertainty and distances of clusterings

Let \( C = \{ c_1, \ldots, c_K \} \) be the set of clusters in \( C \) and let \( L = \{ l_1, \ldots, l_J \} \) be the set of class labels; \( c_k \) denotes then the set of objects belonging to cluster \( k \), and \( l_j \) denotes the set of objects of class \( j \). Depending on the set of labels considered (i.e., the more or less fine grained partition of the set of labels \( L \)), we obtain several labelled versions of the same clustering \( C \). Let us denote by \( L_{17} \) the partition of \( L \) with 17 class labels (the most fine grained partition), and by \( L_6 \) the partition with the six labels selected for Figure 7. We use an entropy measure to quantify how uniform the repartition of objects among the partition of \( X \) is:

\[
H(C) = - \sum_{k=1}^{K} \frac{|c_k|}{|X|} \log \left( \frac{|c_k|}{|X|} \right) \tag{1}
\]

\( H \) is independent on the labelling used, and we obtain \( H(C)=2.18 \). To quantify the impact of the labelling of the cluster, we compute the mutual information between \( C \) and \( L_j \):

\[
I(C, L_j) = \sum_{k=1}^{K} \sum_{j=1}^{J} \frac{|c_k \cap l_j|}{|X|} \log \left( \frac{|X|}{|c_k \cap l_j|} \right) \tag{2}
\]

Depending on the number of labels considered, the mutual information varies. For instance, we have \( I(C, L_{17})=0.65 \), \( I(C, L_6)=0.5 \) and \( I(C, L_5)=0.43 \). The experimental taxonomy (ET) can thus be seen as second clustering from which a distance to \( C \) can be computed. We use here as a distance, the variation information proposed in [9]:

\[
d_{VI}(C, ET) = H(C) + H(ET) - 2I(C, ET) \tag{3}
\]

We obtain \( d_{VI}(C, ET)=2.92 \) for \( ET=L_6=\{\text{Pedestrian; Bike; Small-Medium-Car; Large-Car; Small-Truck; Large-Truck}\} \), while \( H(ET)=1.61 \). The maximum value possibly reached by \( d_{VI} \) is \( \log(278)=5.63 \). The distance \( d_{VI} \) can still be reduced by increasing the classification quality and thus can be used as an optimization criterion for the classification process: On the training dataset, we will thus built the clustering as closest as possible to the experimental taxonomy built for a given class partition.
4 Mastering the gap through ontology mapping

The goal of this work is to augment sensor data so they can be presented in a meaningful way to end users, facilitating therefore their further processing. From a practical point of view, the classification process creates a link between features, as highlighted by sensors, and labels, as assigned by the classification expert (Link (C)-(ET) in Figure 8).

Over the 28 concepts of the taxonomy, 19 are labelled by 7 distinct ontology concepts while 9 remain unassigned, as they correspond to visions unaddressed by the ontology. For instance walking pedestrian is not assigned, as the ontology declines persons as civilian, military or insurgent.

By using the outcome provided by ontology mapping, it becomes possible to enrich sensor data by assigning to each class individual a semantic label.

4.1 Measuring the quality of semantic enrichment

Two measures are defined in order to estimate the semantic gap between resources. They attempt to quantify differences between two distinct representations of the world, and potentially can express the quality of the semantic enrichment.

Therefore, the semantic mismatch \( SM(ET,O) \) is defined as the ratio of non-assigned taxonomy concepts, over the entire set of taxonomy concepts. This measure tries to expresses the dissimilarities between points of view highlighted by our resources.

The semantic overlap \( SO(ET,O) \) is defined as a ratio of the number of distinct ontology concepts assigned to taxonomy concepts, over the entire number of labelled taxonomy concepts. With this definition our goal is to capture the richness of assignations provided by the mapping process. The higher the number of different concepts used, the better the “translation” done.

Those measures range from 0 to 1, and are complementary. While using various data sets and different resources, the quality of translation from one point of view to another improves as \( SM \) decreases and the \( SO \) augments. In our case, \( SM(ET,O)=0.32 \) and \( SO(ET,O)=0.36 \). Although several definitions may be considered, we define an aggregate measure of the two above as a distance between \( ET \) and \( O \):

\[
d_{SO}(ET,O) = SM(ET,O)\cdot(1-SO(ET,O)) \quad (4)
\]

And we obtain \( d_{SO}(ET,O)=0.20 \). This distance is 0 if and only if either \( SM=0 \) (perfect match) or \( SO=1 \) (maximum overlap). Moreover, it is maximal if both \( SM=1 \) and \( SO=0 \) (null match and no overlap).

Table 3 : Excerpt of mapping outcome

<table>
<thead>
<tr>
<th>Taxonomy concept</th>
<th>Ontology concept</th>
<th>Type of relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delivery Truck</td>
<td>Truck</td>
<td>Inclusion</td>
</tr>
<tr>
<td>Car</td>
<td>Automobile</td>
<td>Equality</td>
</tr>
<tr>
<td>Motorbike</td>
<td>Motorcycle</td>
<td>Equality</td>
</tr>
<tr>
<td>Pedestrian</td>
<td>Person</td>
<td>Inclusion</td>
</tr>
<tr>
<td>Jeep Cherokee</td>
<td>Automobile</td>
<td>Inclusion</td>
</tr>
</tbody>
</table>
4.2 From sensors output to end user

During the classification phase, objects will be detected by the sensor network. As a first step, a feature vector is extracted from the recorded signal, which is then compared to the clustering of Figure 7 through a distance measure. The object is then assigned to the closest cluster and the classes’ information of this cluster is then sent to the user. For instance, if Cluster #1 is found to be the closest to the observed object a vector representing the likelihood of the classes is sent, in this case, \([0.56 \ 0.36 \ 0.8 \ 0 \ 0]\) meaning that the object is most likely to be a pedestrian to a level of 56%. Based on this piece of information and with the alignment of Table 3 which provides the mapping between the labels and the operator language, the sensor’s report will be assigned to one concept in the functional ontology (the Person concept in this case). The series of uncertainty measures associated to (1) the classifier’s output, (2) the clustering and (3) the ontological mapping help the operator to assess the quality of the piece of information received (in terms of credibility and reliability) and produce the corresponding intelligence report. Furthermore, semantic enrichment of sensor output facilitates their exploitation by various inference mechanisms and ensures interoperability with other entities.

5 Related work

In order to standardize the semantics of sensors, observations and measurements, a number of research efforts have recently focused on better integration of the sensor world and semantic representations. Several initiatives have been conducted to provide formal representations of sensor models in the form of ontologies based on semantic web techniques. This aims at facilitating sensor data interpretation and exploitation for data fusion or other applications.

In this context, a major initiative of the Open Geospatial Consortium (OGC), the Sensor Web Enablement (SWE) [11], provides data models and encodings for describing sensors and their observations and specifications of web services interfaces to make sensors discoverable and accessible on the web, and enable the integration of heterogeneous sensor data. In particular, this includes the definition of the Sensor Model Language (SensorML). Several research efforts toward sensor ontology definitions leverage from this ongoing effort. OntoSensor [16] is a sensor ontology used to mark-up data from sensors for effective data fusion. It includes or extends concepts from SensorML and other standards (ISO 19115, IEEE SUMO upper-ontology). In order to discriminate the sensor taxonomy hierarchy, OntoSensor uses the attribute measurand to impose the order upon the sensor hierarchy that is thus decomposed into acoustic, metal-detector, seismic, etc. Barmaghi et al [12] also describe a sensor data ontology to provide a uniform model to describe the heterogeneous sensor data based on standards according to the SWE and SensorML.

To deal with the heterogeneity of sensor networks, most research efforts propose to attach semantics to sensor data for enhanced interpretation. In particular, Sheth et al [20] propose in their Semantic Sensor Web vision, to annotate sensor data with spatial, temporal, and thematic semantic metadata. Also in the context of the SWE, in order to register new sensors on the Sensor web and publish observations to a Sensor Observation Service, Bröring et al [15] address semantic matching challenges between observations, features and properties, and propose to use ontologies to add more semantics to the OGC services as well as alignment operations. Jung [14] exploits semantic annotation of sensor streams and ontology alignment. Ni and colleagues [18] propose the Semantic Sensor Net, an extensible framework for dynamic tagging of semantic information (metadata and context information) to sensory data to allow efficient handling of the environment dynamics.

In an acoustic vehicle classification problem, Guo et al [17] extract semantic attributes from acoustic training data, formalized in a domain ontology, and exploit the semantic annotation to enhance vehicle classification and fusion. Gomez et al [19] make use of ontologies as formal models for representing Intelligence Surveillance Reconnaissance (ISR) requirements, ISR capabilities, sensors, sources and platforms to support the effective allocation of ISR assets to multiple competing missions. A sufficiently rich representation of these elements (e.g., platform and sensor ontology) associated with deductive reasoning mechanisms facilitates the matchmaking process.

From another interesting perspective, Dietze and Domingue [13] consider the existing bridge between symbolic knowledge representations and measured data collected by sensors. They propose the use of conceptual spaces to facilitate interoperability, through mapping, between observations and measurements provided by sensors and symbolic semantic web representations. Conceptual spaces represent knowledge in geometrical vector spaces in order to enable computation of similarities between knowledge entities by means of distance metrics.

In our approach, we do not enrich the classification process by ontological information but rather provide a semantic mapping between the sensor network output and the functional standard ontology as shown in Figure 8. This way we ensure the reusability of the classification system.

6 Conclusions and future work

Sensor data are abstract and intrinsically heterogeneous, therefore their processing, sharing and reusability appear as difficult tasks without understanding the underlying meaning. In this paper, by establishing mappings between
elements of Figure 8 and associated measures quantifying the quality of these mappings, we have linked the basic building blocks of a decision support system which bridges the gap between the world of features (as understood by the sensors) and the world of functions (as understood by the operator). The interest of quantifying the mappings is twofold: (1) measures provide the user with a global reliability assessment of the whole data flow from sensor network to the end user, (2) these measures can be used in future works as optimization criteria to improve the chain of processing between the sensors and the experimental taxonomy on the one hand and between the taxonomy and the ontology on the other hand. Moreover, a more precise ontology mapping approach estimating the correctness of matches could allow us to refine the definitions of those measures.

Also, we intend to improve the pertinence of the experimental terminology by learning this resource from sensor data, based on objects properties, as proposed for instance by Sequeda et al [6]. We can also use this approach in order to enrich sensor data with multi-layer information, as mentioned by Huang and colleagues, [5] expressing for instance the level of threats of entities.

References