

Using f-*SHIN* to represent objects: an aid to visual grasping

Nicola Vitucci, Mario Arrigoni Neri, and Giuseppina Gini

Politecnico di Milano - Dipartimento di Elettronica e Informazione
Via Ponzio 34/5, 20133 Milano, Italy
{vitucci,arrigoni,gini}@elet.polimi.it

Abstract. Description Logics (DLs) are nowadays used to face a variety of problems. When dealing with numerical data coming from the real world, however, the use of traditional logics results in a loss of useful information that can be otherwise exploited using more expressive logics. Fuzzy extensions of traditional DLs, being able to represent vague concepts, are well suited to reason on such objects.

In this paper we present an architecture for the automatic building and querying of a fuzzy ontology related to the representation of objects in terms of their composing parts. Our approach mainly aims to face the problem of visual grasping, which is of wide interest in the robotics field.

1 Introduction

The decomposition of an object in parts has been recognized as an important problem in artificial intelligence: it is considered both as a human-like way of reasoning on objects [1] and as a good way to reduce complexity in tasks like object recognition [14]. Apart of the actual image decomposition phase, a major issue is constituted by the semantic description of the extracted features and their mutual relationships. Due to the vagueness affecting real world data, some tolerance should be taken into account when formally representing the structure of an object; this is a reason to take advantage of novel tools as fuzzy DLs [11].

Fuzzy DLs extend crisp DLs by adding imprecision and vagueness in the reasoning process, thus giving some degrees of truth in place of binary answers as *yes* or *no*. Although the available fuzzy reasoners are not yet as powerful as their crisp counterparts, some interesting applications can be found. One of them lies in the robotics field, in which a symbolic representation of objects can improve the grasping capabilities of a robot by the use of some semantic information, regarding both the type of grasp itself and the structure of the object to be grasped.

To the best of our knowledge, the problem of semantic part decomposition is still an open problem and there are no tools available to automatically create a fuzzy ontology from raw concepts. The use of ontologies for object recognition has been investigated in some works as [4,5,6], but none of them makes explicitly use of fuzzy reasoning except for the creation of (crisp) descriptors as *Very_high* to be

used in the classical way; furthermore, they rely on a previous phase of semantic annotation by domain experts, while we focus on the automatic generation of simple concepts, which are sufficient for our purposes.

There are some recent works in which fuzzy DLs are thoroughly used to reason on multimedia information (see [7,8,10,12]) but little advantage is taken from the expressiveness given by cardinality restrictions (when available). Generally speaking, this is due to the fact that, for scene understanding purposes, it is sufficient to know whether a kind of object is present or not (see [15,16]). On the other hand, for object recognition purposes, it is often necessary to be able to count the instances of each kind of recognized component.

In this paper, we show why we use f-*SHIN* [9] as the underlying DL for addressing this problem, then we describe an architecture for the automatic building of a (crisp) ontology and its use for object recognition via fuzzy ABox reasoning services; eventually, in the last section, we make some considerations and propose some future work. The architecture we propose here is still far from being considered complete, yet we were able to obtain some interesting results.

2 The f-*SHIN* logic

The f-*SHIN* logic is the fuzzy extension of the *SHIN* logic [9]. The main improvement of this extension with respect to its crisp version is the possibility to use assertions like $\text{Concept}(p)[\geq 0.7]$, meaning that the individual p has a *minimum degree of participation* of 0.7 to the concept Concept , or $\text{role}(p,q)[\leq 0.3]$, meaning that the individuals p and q participate in the role role with a *maximum degree* of 0.3. The *greatest lower bound* (GLB) [11] is used to know “how much” an individual can be considered to belong to a certain class. A complete description of the f-*SHIN* logic can be found in [9].

For the f-*SHIN* logic there exists a reasoner called *FiRE*¹, while there exist other reasoners like *fuzzyDL*² which is based on the fuzzy extension of the *SHIF* logic. The reason why we chose to use *FiRE* as reasoner is, independently from the supported reasoning services, the high expressivity of the underlying f-*SHIN* logic as it supports cardinality restrictions; on the other hand, such a choice needs some functional blocks to be added to carry out operations like the definition of concepts in terms of membership functions.

3 Architecture

As anticipated in the previous section, due to the limitations of the reasoner, the whole architecture is complex and requires some functional elements to be split among different modules (e.g. the reasoner used on the definitions ontology is different from the one used on the objects ontology). The whole architecture of the system is depicted in Fig. 1.

¹ <http://www.image.ece.ntua.gr/~nsimou/FiRE/>

² <http://gaia.isti.cnr.it/~straccia/software/fuzzyDL/fuzzyDL.html>

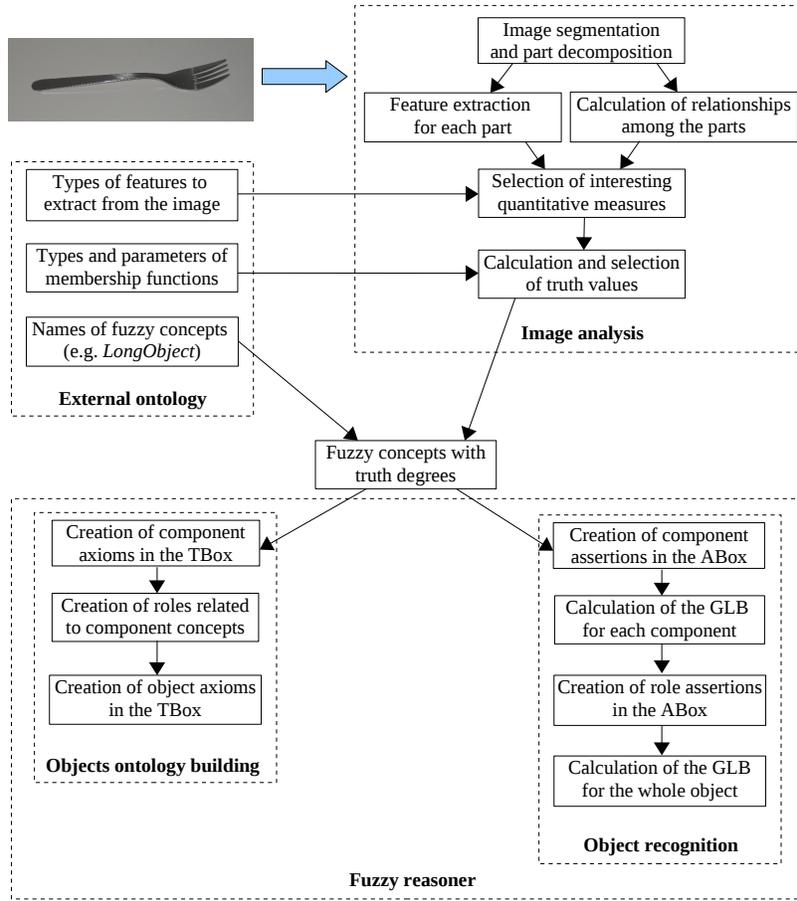


Fig. 1: The general architecture of the system

The “high level” information, which reflects the kind of knowledge that is to be extracted from the image, is encoded in the external ontology; the image analysis and the numerical calculations are performed with MATLABTM, while the intermediate steps are performed either in MATLABTM or in JavaTM. The FiRE reasoner is standalone, thus some steps are still to be carried out by hand.

As an example, we will model a fork in terms of its parts; thus, we will use the images shown in Fig. 2.

3.1 External ontology

The external ontology, also called the “definitions ontology”, is used to specify the kinds of membership functions to be used as well as the kinds of features to

be extracted from the objects found in the images (e.g. elongation, eccentricity, parallelism with respect to other objects and so on) and the meanings of concepts like `LongObj` and `SmallObj` in terms of membership functions.

Taking the ontology described in [2] as an example, we built a meta-ontology (based on the crisp logic $SHOIN(\mathbf{D})$ with datatypes) in which the features to be extracted from the image are subclasses of the meta-class `GeometricConcept` and the kinds of membership functions to use are subclasses of the meta-class `MembFunc`. The ontology presented in [2] makes use of some “concrete” concepts like `TrapezoidalConcreteFuzzyConcept` and `TriangularConcreteFuzzyConcept`, each one having several properties defined as `hasParameterX` (where X stands for A, B, K1 etc.) depending on the parameters needed by the considered membership function; an individual `tra1` represents a trapezoidal membership function with given parameters.

In our ontology, a concept like “a long object” is modeled as an individual `longObject` of meta-class `Length` which has, as its membership function, another individual `longMF` of a subclass of `MembFunc` with the function parameters given as datatype properties. By means of the Jena Ontology API³ and the Pellet reasoner⁴, information like the kind and the parameters of a membership function representing a concept related to the image is extracted to feed the image analysis module; thus, a SPARQL query like:

```
SELECT *
WHERE {
  ?x rdfs:subClassOf :GeometryConcept .
  ?y rdf:type ?x .
  ?y :hasMembershipFunction ?z .
  ?z rdf:type ?w .
  ?w rdfs:subClassOf :MembershipFunction .
  FILTER (?w != :MembershipFunction) .
  ?z :hasParameter1 ?k1 .
  ?z :hasParameter2 ?k2 .
  OPTIONAL {?z :hasParameter3 ?k3} .
  OPTIONAL {?z :hasParameter4 ?k4}
}
```

is used to extract the individuals representing the actual fuzzy geometry concepts (e.g. `LongObject`) used in the objects ontology and their related membership functions data (e.g. a sigmoidal function with two parameters).

The ontology is built by a domain expert to reflect the physical characteristics of the robot, so that for example an object can be considered “long” with respect to the maximum aperture of the robot hand. Although a system of measurement has to be established, we now use only pixel measures.

³ <http://jena.sourceforge.net/>

⁴ <http://clarkparsia.com/pellet/>

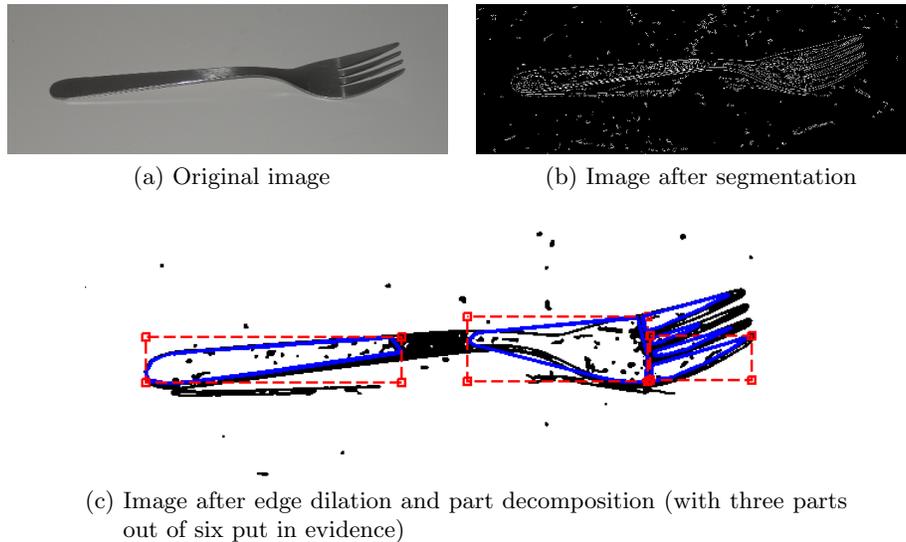


Fig. 2: Steps of the image analysis phase

3.2 Image analysis

In this phase, the original image is converted in a binary image after thresholding and edge recognition performed by Canny method [17] (Fig. 2b); the resulting edges are dilated, then the parts having an area over a threshold are selected (Fig. 2c). This segmentation and decomposition phase is actually non-robust, so that the use of fuzzy relationships can be better shown.

After the first phase, some features like the area, the length of the major axis of the ellipse having the same normalized second central moments as the selected region, and so on, are extracted from each found part (see Tab. 1 for some examples of extracted values); then, some quantitative characteristics are computed: for example, the measure of parallelness π , given α and β as the angles between the major axes of the two objects and the x axis of the image, is defined as $\pi = |\cos(\alpha - \beta)|$, while the distance between two parts, instead, is defined as the minimum distance between their convex hulls.

Using the definitions from the external ontology, for every part we calculate the degree of membership of each feature to its related membership functions. For example, for the feature “length” (i.e. the length of its major axis), the truth values for the functions “LongObj”, “MediumLengthObj” and “ShortObj” are calculated; if a `MediumLengthObj` is associated to a generalized bell curve membership function with parameters $a = 240$, $b = 2.5$, $c = 600$ and the length of the major axis of the considered object is 456.61 pixels, the object will belong to the class `MediumLengthObj` with a truth degree $\mu = 0.93$.

Table 1: Examples of features extracted from the image

(a) Measures of parallelness between every pair of parts

	p1	p2	p3	p4	p5	p6
p1	1.00	0.97	0.97	0.97	0.97	0.97
p2	0.97	1.00	0.89	0.90	0.88	0.89
p3	0.97	0.89	1.00	0.99	0.99	1.00
p4	0.97	0.90	0.99	1.00	0.99	0.99
p5	0.97	0.88	0.99	0.99	1.00	0.99
p6	0.97	0.89	1.00	0.99	0.99	1.00

(b) Other features (area and lengths are in pixels)

	Area	Eccentricity	Major axis	Minor axis	Axes ratio
p1	14860	0.99	456.61	47.33	0.10
p2	12351	0.95	288.23	88.41	0.30
p3	2194	0.98	151.78	23.11	0.15
p4	500	0.99	93.09	8.81	0.09
p5	2617	0.98	181.07	25.79	0.14
p6	771	0.99	141.47	9.39	0.06

3.3 Objects ontology building

Using the results from the previous phase, and taking as a working hypothesis that all the found parts belong to the same object (i.e. there is just one object in the scene), for each part only the membership functions which give the highest truth value for each feature are selected; for example, if a part has a truth degree over a threshold for the membership function “MediumLengthObj”, the concept MediumLengthObj is added to the concept representing that part in the fuzzy ontology. At the end, we obtain a concept like (for the sake of simplicity we list only some concepts and roles):

$\text{ObjClass1} \equiv \text{MediumLengthObj} \sqcap \text{SmallObj} \sqcap \geq 5 \text{ parall} \sqcap \geq 1 \text{ near} \sqcap \dots$

where ObjClass1 is the newly created concept related to the part which has been considered. A new fuzzy concept is created only if the current analyzed part does not belong to any existing concept, i.e. there is no concept that fully describes the part (it can be verified via the fuzzy reasoner). Since FiRE does not let us write fuzzy TBox axioms, the degrees of truth are discarded in this phase.

When there are no parts left, a role for each concept is created. For example, from the class ObjClass1 the role hasObjClass1 is created, so that the class Fork can be created using the previously found number of objects per class:

$\text{Fork} \equiv \geq 1 \text{ hasObjClass1} \sqcap \geq 4 \text{ hasObjClass2} \sqcap \geq 1 \text{ hasObjClass3}$

This is due to the fact that the f-SHIN logic lacks of the qualified cardinality restrictions, so a general hasPart role cannot be used. We use a “typographical” operation, yet the problem of role creation has been faced in [3]. For the sake of completeness, domain and range role axioms should be added to qualify the new roles introduced, but the used reasoner does not fully support them yet.

3.4 Object recognition

Once the objects ontology TBox has been built, it is possible to find whether an object, after it has been decomposed in parts, belongs to a class or not (i.e.

how much it can be considered to belong to the considered class with respect to a certain threshold); the image analysis steps are the same for the ontology building phase.

When for every part all the pertaining concepts and roles can be written in the ABox, the fuzzy reasoning is performed to find the GLB of that part belonging to a certain class; then, roles like `hasObjClass1` are created with the same value of the found GLBs and, at the end, the GLB of the main object is calculated.

This procedure can be applied to determine whether a specific kind of grasp can be performed or not on the selected object. For example, given the concept defined as (for the sake of simplicity using no roles):

`GraspableByPinch` \equiv `MediumLengthObj` \sqcap `HighlyEccentricObj`

representing objects that are graspable by a pinch grip, we can find which part of the object (if any) can be grasped this way via a subsumption check.

4 Conclusions

In this paper we have presented a possible architecture for the generation and the use of a fuzzy ontology for object recognition by means of objects decomposition in parts. We take advantage of the use of fuzzy cardinality restrictions which, to the best of our knowledge, have not been fully exploited in the current fuzzy DLs applications (e.g. multimedia retrieval). Our results are preliminar and prone to errors, partly due to limitations in the modules in use (e.g. the fuzzy reasoner is still experimental), partly due to the approximations induced by the use of a *SHIN* logic, while at least qualified cardinality restrictions would be needed.

As future work, we plan to take advantage of a more powerful fuzzy DL as it seems to be needed for object modeling purposes, so we will work on a more powerful reasoner and on a better integration between classical and fuzzy knowledge bases; furthermore, as we plan to use the system as an aid to the grasping task, we will add physical information (that can be obtained via different sensors, e.g. haptic devices) and further information on the grasping types along with their quality measurements.

References

1. Biederman, I.: Recognition-by-components: A theory of human image understanding. *Psychological Review* **94** 2 (1987) 115-117
2. Bobillo, F., Straccia, U.: An OWL Ontology for Fuzzy OWL 2. *Proceedings of the 18th International Symposium on Methodologies for Intelligent Systems (2009)*
3. Haarslev, V., Lutz, C., Möller, R.: Foundations of spatioterminological reasoning with description logics. *Proceedings of Sixth International Conference on Principles of Knowledge Representation and Reasoning (1998)* 112-123
4. Hudelot, C.: Towards a cognitive vision platform for semantic image interpretation; Application to the recognition of biological organisms. PhD thesis. University of Nice Sophia Antipolis (2005)

5. Maillot, N.: Ontology based object learning and recognition. PhD thesis. University of Nice Sophia Antipolis (2005)
6. Hudelot, C., Atif, J., Bloch, I.: Fuzzy spatial relation ontology for image interpretation. In: Fuzzy Sets and Systems , **159** 15 (2008) 1929–1951
7. Stoilos, G., Stamou, G., Pan, J.Z., Simou, N., Tzouvaras, V.: Reasoning with the fuzzy description logic *f-SHIN*: Theory, practice and applications. In P.C.G. da Costa et al. (eds): Uncertainty Reasoning for the Semantic Web I (2008) 262–281
8. Simou, N., Athanasiadis, T., Tzouvaras, V., Kollias, S.: Multimedia reasoning with *f-SHIN*. Second International Workshop on Semantic Media Adaptation and Personalization (2007) 44–49
9. Stoilos, G., Stamou, G., Tzouvaras, V., Pan, J.Z., Horrocks, I.: The fuzzy description logic *f-SHIN*. International Workshop on Uncertainty Reasoning For the Semantic Web (2005)
10. Mylonas, P., Simou, N., Tzouvaras, V., Avrithis, Y.: Towards semantic multimedia indexing by classification and reasoning on textual metadata. Knowledge Acquisition from Multimedia Content Workshop (2007)
11. Lukasiewicz, T., Straccia, U.: Managing uncertainty and vagueness in Description Logics for the Semantic Web. Journal of Web Semantics **6** 4 (2008) 291–308
12. Straccia, U.: Towards Spatial Reasoning in Fuzzy Description Logics. Proc. of the 2009 IEEE International Conference on Fuzzy Systems (2009)
13. Suh, I. H., Lim, G. H., Hwang, W., Suh, H., Choi, J.-H., Park, Y.-T.: Ontology-based multi-layered robot knowledge framework (OMRKF) for robot intelligence. IEEE Int. Conf. on Intelligent Robots and Systems (2007) 429–436
14. Wan, L.: Parts-based 2D shape decomposition by convex hull. IEEE International Conference on Shape Modeling and Applications (2009) 89–95
15. Dasiopoulou, S., Kompatsiaris, I., Strintzis, M.G.: Applying Fuzzy DLs in the extraction of image semantics. Journal of Data Semantics **14** (2009) 105–132
16. Meghini, C., Sebastiani, F., Straccia, U.: A model of multimedia information retrieval. Journal of ACM **48** 5 (2001) 909–970
17. Canny, J.: A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, **8** 6 (1986) 679-698