SEMANTIC AND COGNITIVE DESCRIPTIONS OF SCENES FOR REASONING AND LEARNING IN AMBIENT INTELLIGENCE

Reporting

Project Information

COGNITIVE-AMI
Grant agreement ID: 328763

Funded under
Specific programme “People” implementing the Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007 to 2013)

Start date End date
1 November 2013 31 October 2015

Total cost
€ 161 968,80

EU contribution
€ 161 968,80

Coordinated by
UNIVERSITAET BREMEN
Germany

This project is featured in...
Final Report Summary - COGNITIVE-AMI (SEMANTIC AND COGNITIVE DESCRIPTIONS OF SCENES FOR REASONING AND LEARNING IN AMBIENT INTELLIGENCE)

[Please note that there is a version including Figures attached. Moreover, a longer version of the report for all audiences is also available on-line in English, Spanish and Catalan languages: https://sites.google.com/site/cognitiveami/-in-few-words ]

The project Semantic and Cognitive Descriptions of Scenes for Reasoning and Learning in Ambient Intelligence (Cognitive-AmI), funded by the Marie Curie Intra-European actions under the 7th European Framework, deals with the extraction of qualitative information from images/videos taken indoors. Why qualitative information? Because qualitative representations abstract unnecessary details and can deal with uncertain data (i.e. noise). Qualitative representations also align cognitive linguistic concepts -easily understood by people- with machine numerical perception, thus enhancing human-machine communication. Qualitative Spatio-Temporal Reasoning (QSTR) [Cohn and Renz, 2007; Ligozat, 2011] is a field of research connecting Artificial Intelligence (computer science) with cognitive spatial perception (psychology) and communication about space (linguistics). Moreover, QSTR has defined useful models to reason about location/orientation [Hernández, 1991], topology [Egenhofer and Al-Taha, 1992; Cohn et al. 1994], direction [Freksa, 1992], visibility [Tarquini et al., 2007], shape [Falomir et al. 2013], etc. and QSTR models have been also applied to different fields such as robotics [Kunze et al., 2014; Falomir et al., 2013b], architecture and design [Bhatt and Freksa, 2015], geographic information systems [Fogliarioni, 2013; Ali et al., 2015], sketch recognition [Lovett et al. 2006], etc.

In Cognitive-AmI project (Figure 1) images and videos were captured by cameras located on a robot or inside a building, such as Cartesium building at Universität Bremen where users can interact with the building using several displays.

Digital images/videos discretize space and represent it as a matrix of colour points or pixels (i.e. Red Green and Blue) which are not connected to each other, that is, those points do not preserve the
properties of space (i.e. continuity, interrelations, etc.). They only preserve their location and their colour, so a lot of effort in computer vision is carried out to find out which pixels belong together and identify an object for instance by studying pixel colour/texture similarity (i.e. segmentation methods and feature detectors). A simple spatial cognitive problem as to know where a cup finishes and the table starts is not so easy to solve using digital images. RGB-depth sensors can obtain also the depth of the pixels in the space, converting a digital pixel matrix into a point cloud. Then the problem to recognize where the cup finishes and where the table starts can be solved calculating where the points belong together in a vertical or horizontal plane. However, a cognitive approach for a human would be to interact with the cup and the table, for example, taking the cup and trying to separate it from the table. If possible, the human/cognitive agent would learn that those objects are not attached to one another, so they can be disjoint.

The objective of this project is to use the methods available in computer vision field for recognizing objects [Bay et al., 2008; Muja and Lowe, 2009], regions [Felzenszwalb and Huttenlocher, 2004] or movements [Zivkovic, 2004], and from the data obtained, try to abstract concepts that preserve the properties of space and that can try to describe scenes in a more cognitive manner. Indoor scenes at Cartesium building were captured, where the Bremen Spatial Cognition Centre is located at the Universität Bremen, in order to obtain a dataset to apply the model developed for the qualitative and logic descriptions of scenes (QIDL). This model obtains a logic and narrative description of spaces using qualitative features of shape, colour, topology, location and size [Falomir, 2015a]. The main aim is to describe the location of the objects which are needed for a task or known a priori, but also to describe unknown objects from which the system only knows its colour, shape or location, so that it can provide these features to a user which can categorize the object with a name. The logic description provided by QIDL uses Horn clauses implemented in Prolog which can reason about spatial locations (Figure 2). The experiments carried out at Cartesium building in common areas and in offices have shown the utility of the developed model [Falomir and Olteteanu, 2015]. So that artificial software agents can understand indoor environments those descriptions have also been obtained as ontological description logics too [Falomir, 2014; 2013b].

Similarity methods to compare scenes [Falomir et al., 2014], shapes [Museros et al., 2015] and paintings [Falomir et al., 2015a] have been defined based on conceptual neighbourhood relations between qualitative concepts. For example, similarities between colours of painters such as Dalí, Miró, el Greco, Velázquez and Hundertwasser were obtained automatically (Figure 3) and compared to those provided by the participants of a survey. The results obtained were correlated. Moreover, as a cognitive system must have the capacity to learn, some learning techniques (i.e. support vector machines) have been applied to the categorization of different painting styles (i.e. Barroque, Impresionism and Post-Impresionism) [Falomir et al., 2015b] which seems to follow some logic on colour palettes. Besides, the adaptability and usability of the qualitative colour model defined has been showed when customized by its users [Sanz et al., 2015].

Qualitative models have also shown their applicability to describe movements in videos (i.e. location and direction of an object at a moment or for a period of time). This movement can be described using Horn clauses in Prolog, which can be used for reasoning about the information obtained in order to categorize movements (i.e. parabolic, straight, and so on) [Falomir and Rahman, 2015].

Moreover, in order to improve human-machine communication, a grammar has been developed to generate sentences in natural language from the qualitative descriptions obtained from the scenes, so that
A narrative description can be obtained [Falomir, 2013a]. Moreover, some studies on cognitive linguistics have been taken into account, specifically, how people refer to objects when we need to describe them to another person. Results show that people try to discriminate between the more characteristic features of the objects, so that other people know what they are referring to. Accordingly, a model has been created [Mast et al., 2015] which obtain features of shape, colour and location of objects, in an absolute and in a vague manner and these features have been used to describe objects in a context. For example, the colour of an object can be perceived as red, pink, brown, etc. depending on the person observing, but also depending on the context, people could refer to the same object as dark/light/pale red if there is another object which is also red. The descriptions produced by absolute and vague modes have been compared to those produced by participants in a study and results showed that vague models are more adaptable to the context/dialogues with people.

Furthermore, a qualitative model for describing 3D objects (Q3D) based on depth and different perspectives has been developed (Figure 4a). If we consider 3 perspectives of an object as canonical (i.e. front, right and up perspectives) we can take into account the continuity relations between those perspectives to define the conditions to hold in each perspective. If those conditions are not fulfilled, then the description is not consistent. Moreover, Q3D descriptions can be used to infer the rest of perspectives of the object which are occluded. For example, if an object has a transversal open hole, it must be described in all the perspectives where the hole is seen (i.e. front-back, up-down, etc.). Accordingly, logic descriptions have been defined, implemented and tested in Prolog [Falomir, 2015b, Falomir 2015c]. Results obtained are promising and they can be useful to help students solve the intelligence test by the German Studienstiftung (Figura 4b). Furthermore, an approach has been also developed to cognitively describe in natural language real 3D scenes which contain oriented objects (i.e. chairs) which have a front side different than the one used by the speaker [Kluth and Falomir, 2013].

Finally, cognitive tests have been carried out about creativity and its relation with the common or uncommon associations we people do between linguistic and visual concepts. A computational method (comRAT-C) [Olteteanu and Falomir, 2015] was developed which provides a concept related with three other concepts presented and produces similar results to the test carried out by Mednick and Mednick [1971] on humans to measure their level of creativity. For example, what would be a remote associate (RAT) to the following 3 concepts: Cottage-Swiss-Cake? The studies by Mednick and Mednick [1971] provided Cheese as the convergent concept, since there exist cottage cheese, Swiss cheese and cheesecake. The computational method comRAT-C can provide other possibilities like Chocolate, since there exist also Swiss chocolate, chocolate cake and the chocolate cottage in Hansel and Gretel fairy tale.

Acknowledgements: The collaboration in this project of the following researchers is very acknowledged: Ana-Maria Olteteanu (U. Bremen), Thomas Kluth (U. Bielefeld), Vivien Mast (U. Postdam), Diedrich Wolter (U. Bamberg), Lledo Museros (U. Jaume I), Ismael Sanz (U. Jaume I) and Luis Gonzalez-Abril (U. Sevilla).

REFERENCES


Related documents

[final1-cognitive-ami-finalreport.pdf]

Last update: 19 February 2016

Permalink: https://cordis.europa.eu/project/id/328763/reporting

European Union, 2023