The project ***Semantic and Cognitive Descriptions of Scenes for Reasoning and Learning in Ambient Intelligence***[1] **(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.

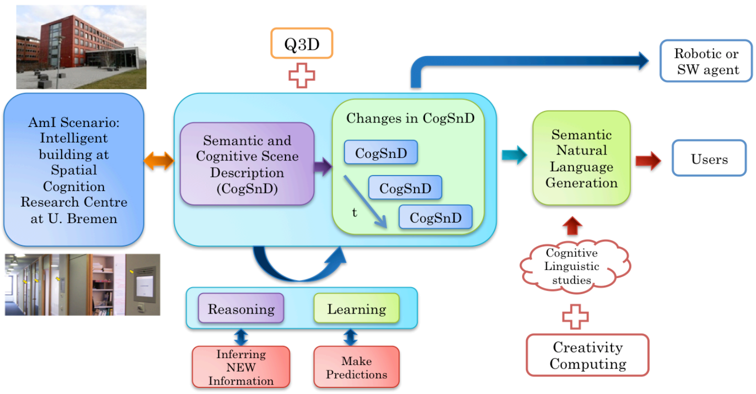


Figure 1. Graphical abstract of the Cognitive-AmI project.

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].

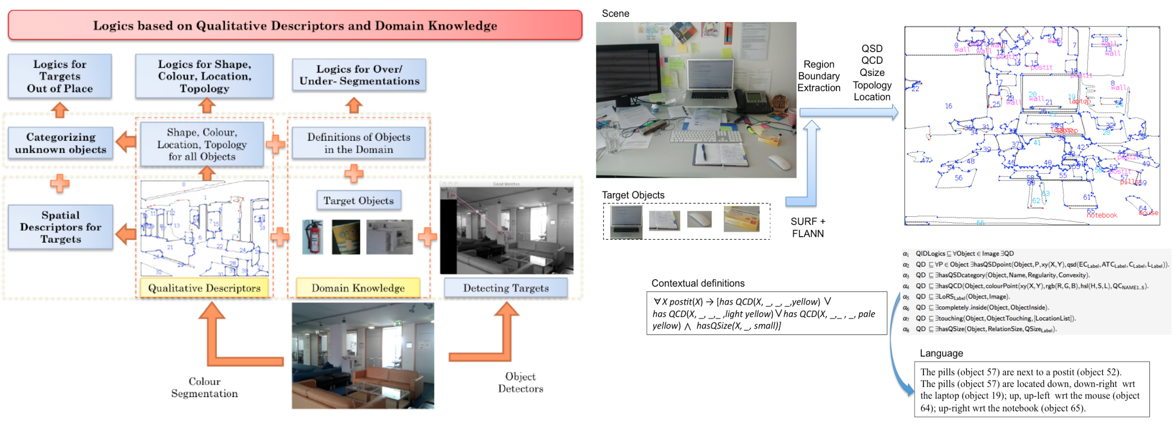


Figure 2. Graphical summary of the process followed to extract qualitative logic descriptions of indoor spaces.

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, Velàzquez 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].

[1] More details on Cognitive-AmI project website: **https://sites.google.com/site/cognitiveami/**

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, an so on)[Falomir and Rahman, 2015].

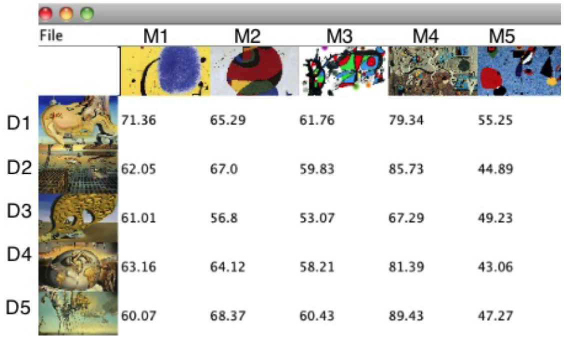
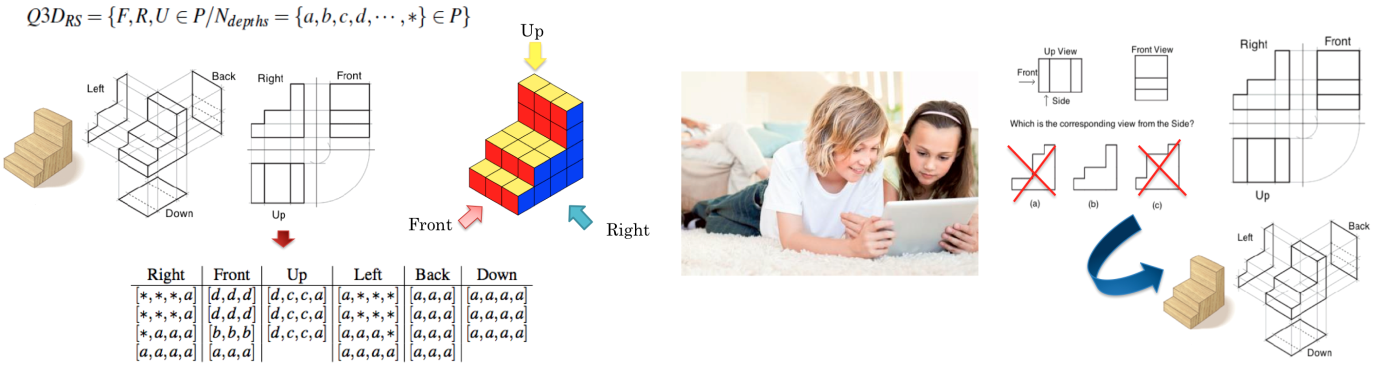


Figure 3. Similarity values obtained between some paintings by Dalí (D) and Miró (M).

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 a 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.

Moreover, 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].



(a) (b)

Figure 4. (a) Example of a description produced by the Q3D model for describing objects depending on volumes and depth; (b) The Q3D model may be useful to help students to understand the technical drawing descriptions of 3D objects.

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

[Ali et al., 2015] Ali A.L., Schmid F., Falomir Z., Freksa C., Towards Guided Classification for Volunteered Geographic Information, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., Vol. II-3/W5, pp. 211-217, DOI:10.5194/isprsannals-II-3-W5-211-2015, 2015.

[Bay et al., 2008] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (SURF). *Comput. Vis. Image Underst*., 110(3):346–359, June 2008.

[Bhatt and Freksa, 2015] M. Bhatt and C. Freksa, Spatial computing for design an artificial intelligence perspective, in S*tudying Visual and Spatial Reasoning for Design Creativity*, J. S. Gero, Ed., 2015, pp. 109–127.

[Cohn et al. 1994] Cohn, A., Randell, D., Cui, Z., Bennett, O., and Gooday, J. (1994). Taxonomies of logically defined qualitative spatial relations. In in N. Guarino and R. Poli (eds), Formal Ontology in Conceptual Analysis and Knowledge Representation, pages 831-846. Kluwer.

[Cohn and Renz, 2007] A. G. Cohn and J. Renz, *Qualitative Spatial Reasoning, Handbook of Knowledge Representation*, V. L. F. Harmelen and B. Porter, Eds. Wiley-ISTE, London: Elsevier, 2007.

[Egenhofer and Al-Taha, 1992] Egenhofer, M. J. and Al-Taha, K. K. (1992). Reasoning about gradual changes of topological relationships. In Frank, A. U., Campari, I., and Formentini, U., editors, Theories and Methods of Spatio-Temporal Reasoning in Geographic Space. Intl. Conf. GIS|From Space to Territory, volume 639 of Lecture Notes in Computer Science, pages 196-219, Berlin. Springer.

[Falomir et al. 2013] Falomir Z., Gonzalez-Abril L., Museros L., Ortega J. (2013), Measures of Similarity between Objects from a Qualitative Shape Description, *Spatial Cognition and Computation*, 13 (3): 181–218.

[Falomir, 2013a] Z. Falomir. Towards cognitive image interpretation qualitative descriptors, domain knowledge and narrative generation. In V. Botti K. Gibert and R. Reig-Bolao, editors, *Artificial Intelligence Research and Development, Frontiers in Artificial Intelligence and Applications*, vol. 256, pages 77-86, IOS Press, Amsterdam, 2013.

[Falomir, 2013b] Z. Falomir. Towards scene understanding using contextual knowledge and spatial logics. In J. Dias, F. Escolano, and R. Marfil, editors, *Proc. of the 2nd Workshop on Recognition and Action for Scene Understanding (REACTS)*, pages 85–100, 2013. ISBN 978-84-616-7092-5.

[Falomir et al. 2013b] Z. Falomir, L. Museros, V. Castelló, and L. Gonzalez-Abril, Qualitative distances and qualitative image descriptions for representing indoor scenes in robotics, *Pattern Recognition Letters*, 38: 731–743, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.patrec.2012.08.012

[Falomir et al., 2014] Z. Falomir, L. Museros, and L. Gonzalez-Abril. Towards a similarity between qualitative image descriptions for comparing real scenes. In *Qualitative Representations for Robots, Proc. AAAI Spring Symposium*, Technical Report SS-14-06, pages 42–49, 2014. ISBN 978-1-57735-646-2, Palo Alto, California, USA, 2014.

[Falomir, 2014] Z. Falomir. An approach for scene interpretation using qualitative descriptors, semantics and domain knowledge. In *Knowledge Representation and Reasoning in Robotics, AAAI Spring Symposium Series*, pages 95–98, 2014. ISBN 978-1-57735-646-5. Palo Alto, California, USA, 2014.

[Falomir, 2015a] Zoe Falomir. A qualitative image descriptor QIDL+ applied to ambient intelligent systems. In *Proceedings of the 10th International Workshop on Artificial Intelligence Techniques for Ambient Intelligence (AITAmI15), co-located at IJCAI-2015*, Accepted. Buenos Aires, Argentina, 2015.

[Falomir, 2015b] Falomir Z., A Qualitative Model for Reasoning about 3D Objects using Depth and Different Perspectives, *1st Workshop on Logics for Qualitative Modelling and Reasoning (LQMR).* Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 1--9, Lodz, Poland, September 2015.

[Falomir, 2015c] Falomir Z., A Qualitative Model for Describing 3D Objects using Depth. In *Spatio Temporal Dynamics (STeDy) Workshop at International Joint Conference on Artificial Intelligence (IJCAI)*, Buenos Aires, Argentina, July 2015.

[Falomir and Olteteanu, 2015] Z. Falomir and A-M. Olteteanu. Logics based on qualitative descriptors for scene understanding. Neurocomputing, 161:3–16, 2015. Available: http://dx.doi.org/10.1016/j.neucom.2015.01.074

[Falomir and Rahman, 2015] Z. Falomir and S. Rahman. From qualitative descriptors of movement towards spatial logics for videos. In J. Dias, F. Escolano, and R. Marfil, editors, *Proc. of the 3rd Workshop on Recognition and Action for Scene Understanding (REACTS)*, accepted, 2015.

[Falomir et al., 2015a] Z. Falomir, L. Museros, and L. Gonzalez-Abril. A model for colour naming and comparing based on conceptual neighbourhood. An application for comparing art compositions. *Knowledge-Based Systems*, 81:1–21, 2015. Available: http://dx.doi.org/10.1016/j.knosys.2014.12.013

[Falomir et al., 2015b] Z. Falomir, L. Museros, I. Sanz, and L. Gonzalez-Abril. Guessing art styles using qualitative colour descriptors, SVMs and logics. In Artificial Intelligence Research and Development, *Frontiers in Artificial Intelligence and Applications*. IOS Press, accepted, Amsterdam, 2015.

[Felzenszwalb and Huttenlocher, 2004] P. F. Felzenszwalb, D. P. Huttenlocher, Efficient graph-based image segmentation, *Int. J. Comput. Vis*. 59 (2) (2004) 753 167–181.

[Fogliarioni, 2013] P. Fogliaroni, Qualitative Spatial Configuration Queries. Towards Next Generation Access Methods for GIS, ser. *Dissertations in Geographic Information Science. IOS Press*, 2013, ISBN 978-1614992486.

[Freka, 1992] Freksa, C. (1992). Using orientation information for qualitative spatial reasoning. In Frank, A. U., Campari, I., and Formentini, U., editors, Theories and Methods of Spatio-Temporal Reasoning in Geographic Space. Intl. Conf. GIS|From Space to Territory, volume 639 of Lecture Notes in Computer Science, pages 162-178, Berlin. Springer.

[Hernandez, 1991] Hernandez, D. (1991). Relative representation of spatial knowledge: The 2-D case. In Mark, D. M. and Frank, A. U., editors, Cognitive and Linguistic Aspects of Geographic Space , NATO Advanced Studies Institute, pages 373-385. Kluwer, Dordrecht.

[Kluth and Falomir, 2013] T. Kluth and Z. Falomir. Studying the role of location in 3D scene description using natural language. In J. A. Ortega I. Sanz, L. Museros, editor, *XV Workshop of the Association on Qualitative Reasoning and its Applications (JARCA13). Qualitative Systems and their applications to Diagnosis, Robotics and Ambient Intelligence.* Proceedings from the University of Seville, pages 33–36, 2013. ISBN 978-84-616-7622-4.

[Kunze et al., 2014] L. Kunze, C. Burbridge, and N. Hawes, Bootstrapping probabilistic models of qualitative spatial relations for active visual object search, in *Qualitative Representations for Robots, Proc. AAAI Spring Symposium*, Technical Report SS-14-06, 2014, pp. 81–80, ISBN 978-1-57735-646-2.

[Ligozat, 2011] G. Ligozat, Qualitative Spatial and Temporal Reasoning. Wiley-ISTE, London: MIT Press, 2011.

[Lovett et al. 2006] A. Lovett, M. Dehghani, and K. Forbus, Learning of qualitative descriptions for sketch recognition, in Proc. *20th Int. Workshop on Qualitative Reasoning (QR)*, Hanover, USA, July, 2006.

[Mast et al., 2015] V. Mast, Z. Falomir, and D.Wolter. Probabilistic reference and grounding with PRAGR for dialogues with robots. *Journal of Experimental & Theoretical Artificial Intelligence*, under revision, 2015.

[Mednick and Mednick, 1971] Mednick, S.A., Mednick, M.: *Remote associates test: Examiner's manual*. Houghton Mifflin (1971).

[Muja and Lowe, 2009] M. Muja and D. G. Lowe. Fast approximate nearest neighbors with automatic algorithm configuration. In *VISAPP Int. Conf. on Computer Vision Theory and Applications*, pages 331–340, 2009.

[Museros et al., 2015] L. Museros, Z. Falomir, I. Sanz, and L. Gonzalez-Abril. Sketch retrieval based on qualitative shape similarity matching: Towards a tool for teaching geometry to children. *AI Communications*, 28(1):73–86, 2015.

[Sanz et al., 2015] I. Sanz, L. Museros, Z. Falomir, and L. Gonzalez-Abril. Customizing a qualitative colour description for adaptability and usability. *Pattern Recognition Letters, SI: Cognitive Systems for Knowledge Discovery*, http://dx.doi.org/10.1016/j.patrec.2015.06.014

[Olteteanu and Falomir, 2015] Olteteanu, A.M., Falomir, Z.: comRAT-C - A computational compound Remote Associates Test solver based on language data and its comparison to human performance. *Pattern Recognition Letters (2015)*, http://dx.doi.org/10.1016/j.patrec.2015.05.015

[Tarquini et al. 2007] Tarquini, F., De Felice F., Fogliaroni P., Clementini E., A qualitative model for visibility relations. KI, Advances in Artificial Intelligence 2007.

[Zivkovic, 2004] Zoran Zivkovic. Improved adaptive gaussian mixture model for background subtraction. *In Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on,* volume 2, pages 28 – 31. IEEE, 2004.