

CENTRO UNIVERSITÁRIO FEI

ADELINE CECILIA SÉCOLO

**COLLABORATIVE SPATIAL REASONING FOR ENVIRONMENT MAPPING
USING UNMANNED AERIAL VEHICLES**

São Bernardo do Campo

2022

ADELINE CECILIA SÉCOLO

**COLLABORATIVE SPATIAL REASONING FOR ENVIRONMENT MAPPING
USING UNMANNED AERIAL VEHICLES**

Doctoral thesis presented to FEI University Center to obtain the title of Doctor in Electrical Engineering, supervised by Prof. Dr. Paulo Eduardo Santos.

São Bernardo do Campo

2022

Sécolo, Adeline Cecilia.

Collaborative Spatial Reasoning for Environment Mapping Using Unmanned Aerial Vehicles / Adeline Cecilia Sécolo. São Bernardo do Campo, 2021.

103 p. : il.

Tese - Centro Universitário FEI.

Orientadora: Prof.^a Dra. Paulo Eduardo Santos.

Coorientadora: Prof.^a Dra. Reinaldo Augusto Bianchi.

1. Collaborative Spatial Reasoning. 2. Qualitative Spatial Reasoning. 3. Allen's Interval Algebra. 4. Environment Mapping. 5. Unmanned Aerial Vehicles. I. Santos, Paulo Eduardo, orient. II. Título.

Aluno: Adeline Cecília Secolo

Matrícula: 516301-9

Título do Trabalho: COLLABORATIVE SPATIAL REASONING FOR ENVIRONMENT MAPPING USING UNMANNED AERIAL VEHICLES.

Área de Concentração: Inteligência Artificial Aplicada à Automação e Robótica

Orientador: Prof. Dr. Paulo Eduardo Santos

Data da realização da defesa: 21/10/2021

ORIGINAL ASSINADA

Avaliação da Banca Examinadora

A defesa de tese de doutorado da candidata Adeline C. Sécolo iniciou-se às 9:00am (horário de Brasília) do dia 21/10/2021 e contou com a presença de todos os membros da banca avaliadora: Prof.^a Esther Colombini, Prof.^a Leliane Nunes, Prof. Flavio Tonidandel, Prof. Reinaldo Bianchi e eu (orientador) Prof. Paulo Santos, convidados e familiares também estavam presentes. A banca ocorreu por meio de vídeo conferência, sem ter havido alguma ocorrência que mereça destaque.

A aluna apresentou o seu trabalho em 40 min, respeitando o prazo previsto, e respondeu aos questionamentos dos membros da banca de maneira satisfatória.

A banca considerou o trabalho com nível suficiente para justificar sua aprovação por unanimidade.

São Bernardo do Campo, 06 / 02 / 2022 .

Adeline C. Sécolo

MEMBROS DA BANCA EXAMINADORA

Prof. Dr. Paulo Eduardo Santos	Ass.: _____
Prof. Dr. Reinaldo Augusto da Costa Bianchi	Ass.: _____
Prof. Dr. Flávio Tonidandel	Ass.: _____
Prof. ^a Dr. ^a Leliane Nunes de Barros	Ass.: _____
Prof. ^a Dr. ^a Esther Luna Colombini	Ass.: _____

A Banca Examinadora acima-assinada atribuiu ao aluno o seguinte:

APROVADO ☒

REPROVADO ☐

VERSÃO FINAL DA TESE

**ENDOSSO DO ORIENTADOR APÓS A INCLUSÃO DAS
RECOMENDAÇÕES DA BANCA EXAMINADORA**

Aprovação do Coordenador do Programa de Pós-graduação

Prof. Dr. Carlos Eduardo Thomaz

Adeline Cecilia Sécolo

COLLABORATIVE SPATIAL REASONING FOR ENVIRONMENT MAPPING
USING UNMANNED AERIAL VEHICLES

Doctoral thesis presented to FEI University Center to obtain the title of Doctor in Electrical
Engineering

Judging Committee

Prof. Dr. Paulo Eduardo Santos

Prof. Dr. Reinaldo Bianchi - FEI

Prof. Dr. Flávio Tonidandel - FEI

Prof. Dra. Leliane Nunes - USP

Prof. Dra. Esther Colombini - UNICAMP

São Bernardo do Campo, SP

September, 2021

To my parents Estanislau and Lindaura.

ACKNOWLEDGMENT

I thank my parents Estanislau and Lindaura for giving me all support and love I needed to not give up, my siblings Andreza and André for staying by my side, as well as my sister-in-law Aline.

My deep gratitude to Prof. Paulo Eduardo Santos, for his orientation, advising, patience and for sharing his huge knowledge to me. I also thank Prof. Reinaldo Bianchi for his advice, friendship and coffees during the afternoons at FEI, the RoboCup laboratory team Danilo Perico, Isaac Silva and Douglas Rizzo, and also Prof. Carlos Eduardo Thomaz. I thank you all for your understanding and helping to overcome the adversities that have appeared in recent years.

Patrick Doherty and Zoran Sjanic for the receptivity, guidance, teachings, for letting me be part of the team during my stay at Linköping University and to enable the experiments with modern unmanned aerial vehicles. Pietro, Mariuz, Tom and Cirille for the friendship, happy hours and help with the experiments and data management at Artificial Intelligence and Integrated Computer Systems laboratory. Karin, Ana, Daniel, Mathias, Vanessa, Tim, Sam, Evelina and Robin for the *fikas*, for your friendship, and for making my stay at Linköping an unforgettable experience.

To Natalia Sécolo, Nivea Ivo and Sidnei Silva for being the best friends a person can have.

I thank CAPES and FEI for the scholarship, essential so I could develop this work. And the partnership CISB-SAAB for providing the enriching experience of exchanging knowledge abroad. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

What sort of people live about here?"

"In that direction," the Cat said, waving its right paw round, "lives a Hatter: and in that direction," waving the other paw, "lives a March Hare. Visit either you like: they're both mad."

"But I don't want to go among mad people," Alice remarked.

"Oh, you can't help that," said the Cat: "we're all mad here. I'm mad. You're mad."

"How do you know I'm mad?" said Alice.

"You must be," said the Cat, "or you wouldn't have come here."

Lewis Carroll/ Charles Lutwidge Dodgson (1832–1898)

RESUMO

O objetivo deste projeto é a investigação dos formalismos de raciocínio espacial existentes para sistemas colaborativos, a fim de interpretar uma cena a partir de múltiplos pontos de vista em tarefas de mapeamento de ambientes. Motivadas pela crescente necessidade de interação entre humanos e robôs, as teorias do Raciocínio Espacial Qualitativo (QSR) são integradas em um único formalismo para modelar as percepções de veículos aéreos não tripulados (VANTs) operados remotamente. As teorias qualitativas possibilitam a troca de informações entre humanos e agentes robóticos, para que possam realizar tarefas em missões colaborativas envolvendo objetivos de busca e monitoramento na agricultura, desastres naturais, tarefas de busca e resgate, entre outros. A combinação das teorias espaciais estudadas levou ao desenvolvimento de dois formalismos: o Cálculo de Intervalo LH e o Raciocínio Espacial Colaborativo. O Cálculo de Intervalo LH consiste na combinação do Cálculo de Conexões de Regiões com a Álgebra de Intervalos de Allen para descrever as relações entre dois objetos de um ponto de vista aéreo. O Raciocínio Espacial Colaborativo combina o Cálculo de Direção Cardinal com o Cálculo do Intervalo LH para a tarefa de mapeamento do ambiente onde os agentes têm uma visão parcial da cena. VANTs equipados com câmeras são a plataforma utilizada para testar o formalismo deste projeto, captando imagens com uma visão parcial do ambiente, de diferentes direções de voo. Os resultados obtidos mostraram que os dois formalismos propostos tiveram sucesso na tarefa de mapeamento do ambiente.

Palavras-chave: Raciocínio Espacial Colaborativo. Raciocínio Espacial Qualitativo. Mapeamento de Ambientes. Veículos Aéreos Não Tripulados.

ABSTRACT

The goal of this project is the investigation of existing spatial reasoning formalism for collaborative systems, in order to interpret a scene from multiple viewpoints in the task of environment mapping. Motivated by the increasing need of interaction between humans and robots, Qualitative Spatial Reasoning (QSR) theories are integrated into a single formalism for modeling the perceptions of remotely operated Unmanned Aircraft Vehicles (UAV). Qualitative theories enables the exchange of information between humans and robotic agents, so that they can perform tasks in collaborative missions involving searching and monitoring objectives in agriculture, natural disasters, searching and rescue tasks, among others. The combination of the studied spatial theories led to the development of two formalism: the LH Interval Calculus and the Collaborative Spatial reasoning. LH Interval Calculus consists in the combination of Region Connection Calculus and Allen's Interval Algebra to describe the relations of two objects from an aerial point of view. Collaborative Spatial Reasoning combines the Cardinal Direction Calculus with LH Interval Calculus to the task of environment mapping where agents have a partial view of the scene. UAVs equipped with cameras are the platform used to test the formalism of this project, capturing images with a partial view of the environment, from different directions of flight. The results obtained showed that the two formalism proposed were successful in the task of mapping the environment.

Keywords: Collaborative Spatial Reasoning. Qualitative Spatial Reasoning. Allen's Interval Algebra. Environment Mapping. Unmanned Aerial Vehicles.

LIST OF FIGURES

Figure 1	– UAVs executing monitoring of civil application functions	16
Figure 2	– Operating system <i>Safe2Ditch</i>	17
Figure 3	– UAV Northrop-Grumman Global Hawk fixed wing	18
Figure 4	– UAV Northrop-Grumman NASA Global Hawk	18
Figure 5	– UAV Saab AB Skeldar V-200	19
Figure 6	– Semantic Object Map representation.	26
Figure 7	– Relations of length among Tangram pieces.	29
Figure 8	– Composition table for length in Qualitative Shape Descriptors.	29
Figure 9	– RCC-8 Conceptual Neighborhood Diagram.	31
Figure 10	– Cardinal Direction Calculus representation, quadrant projection.	33
Figure 11	– Allen’s Relations	34
Figure 12	– Qualitative Angles.	36
Figure 13	– Spatial relations between angles.	37
Figure 14	– Composition Table for angle feature in Qualitative Shape Descriptors when composing convex (cx) angles.	38
Figure 15	– Composition Table for angle feature in Qualitative Shape Descriptors when composing convex (cx) and concave (cv) angles.	38
Figure 16	– Two different examples where relation <i>meets</i> applies.	38
Figure 17	– Basic Relations of Interval Occlusion Calculus	40
Figure 18	– Qualitative map	44
Figure 19	– Projective Preposition Regions	46
Figure 20	– Simulated Experiment	47
Figure 21	– Real Data Experiment	48
Figure 22	– Example of camera resection and landmark triangulation.	51
Figure 23	– Icon Example: an arrow shape and a C-like shape.	53
Figure 24	– Location Reference System.	53
Figure 25	– Process of understanding the visual representation of the icon <i>Exit</i>	54
Figure 26	– Grey intervals representing object’s projections from an aerial point of view.	56
Figure 27	– Set of Allen’s relations for Disconnected objects.	60

Figure 28 – Set of Allen’s relations for Externally Connected objects.	60
Figure 29 – Set of Allen’s relations for Partially Overlapped objects.	61
Figure 30 – Set of Allen’s relations for Tangential Proper Part objects.	61
Figure 31 – Set of Allen’s relations for the inverse of Tangential Proper Part objects. . .	61
Figure 32 – Set of Allen’s relations for Non Tangential Proper Part objects and its inverse configuration.	62
Figure 33 – Patterns of scan trajectories.	63
Figure 34 – Cardinal directions.	63
Figure 35 – Axis of reference’s positions according to agents orientation of flight. . . .	65
Figure 36 – Allen’s Composition Table	68
Figure 37 – DJI Matrice 600 Pro experimental platform.	69
Figure 38 – DJI Matrice 100 experimental platform.	70
Figure 39 – Motala Flygklubb scenario	71
Figure 40 – VGG object identification and labels	73
Figure 41 – Yolo V5 object identification and labels	74

LIST OF TABLES

Table 1	– RCC relations and interpretations.	31
Table 2	– RCC-8 Composition Table.	32
Table 3	– Cardinal Direction Calculus Composition Table.	33
Table 4	– RCC-8 Composition Table.	67
Table 5	– Accuracy for LH Interval Calculus relations between the objects at Motala Flygklubb.	75
Table 6	– LH Interval Calculus Expected Results x Obtained Results	76
Table 7	– Collaborative Spatial Reasoning inference for two images with one object in common	82

CONTENTS

1	INTRODUCTION	15
1.1	Main Objective	21
1.2	Contributions	22
1.2.1	First Contribution	23
1.2.2	Second Contribution	23
1.3	Thesis Structure	24
2	THEORY	25
2.1	Cognitive Robotics	25
2.2	Qualitative Spatial Reasoning	27
2.2.1	Region Connection Calculus	30
2.2.2	Cardinal Direction Calculus	32
2.2.3	Allen's Interval Algebra	34
2.2.4	Rectangle Calculus	35
2.2.5	Interval Occlusion Calculus	39
2.3	Simultaneous Localization and Mapping	40
3	LITERATURE REVIEW	43
3.1	Qualitative Spatial Reasoning	43
3.2	Multi-agent Systems	45
3.3	Robotic Localization	49
3.4	Image Interpretation	51
4	METHOD	55
4.1	Problem Definition	55
4.2	LH Interval Calculus	55
4.3	Collaborative Spatial Reasoning for Environment Mapping	62
4.3.1	Partial View Combination in Collaborative Spatial Reasoning	65
5	EXPERIMENTS	69
5.1	Domain	69
5.1.1	Environment Description	70
5.1.2	Data Collection Flight	70
5.2	LH Interval Calculus Evaluation	71
5.2.1	Evaluation Method	72

5.2.2	LH Interval Calculus Results	74
5.3	Collaborative Spatial Reasoning Evaluation	81
5.3.1	Collaborative Spatial Reasoning Results	81
5.3.2	Discussion	85
6	CONCLUSION	86
	BIBLIOGRAPHY	88
	INDEX	94
	ANNEX A – Publications	94

1 INTRODUCTION

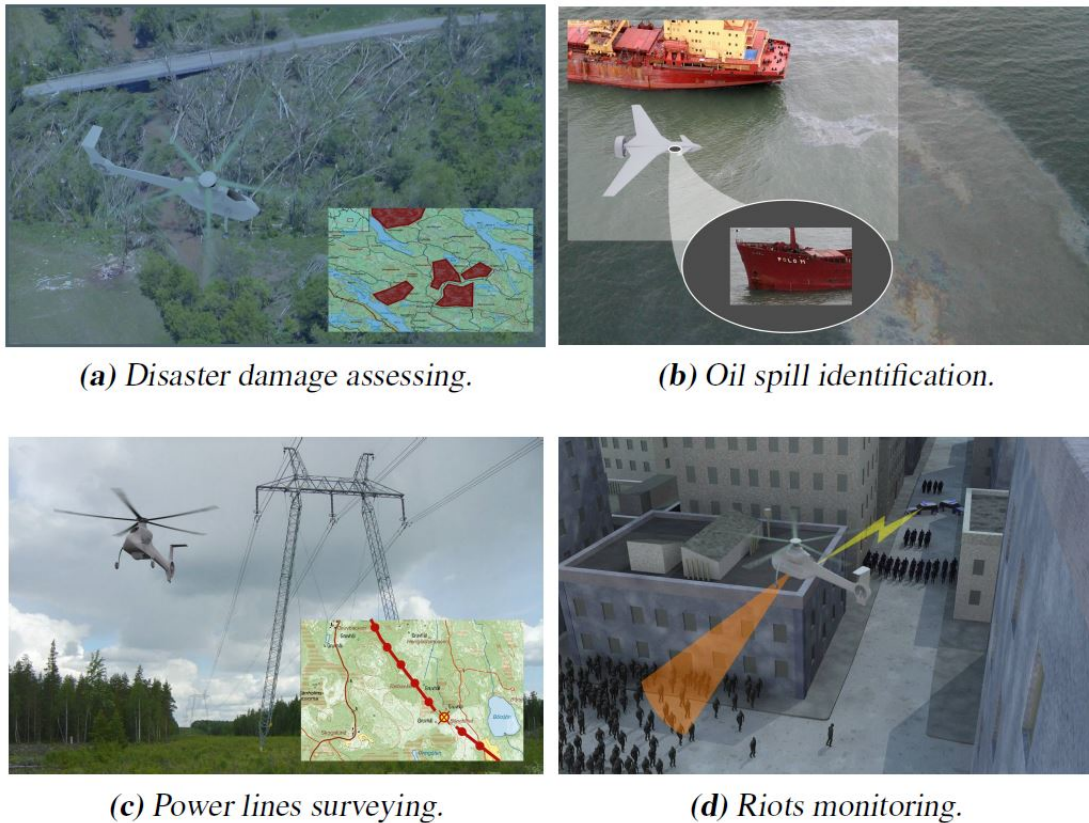
The natural human characteristic of exploring the environment around himself led to the continuous development and improvement in the art of navigation, to know where we are in relation to a previously known reference point, whether at sea, land or air. Navigation systems can serve as auxiliaries in operations performed by humans in remotely controlled vehicles, or even act as the main operator, in the case of robots, cars, submarines or unmanned aircraft equipped with autonomous systems. In this way, it is possible to carry out missions where the risk of vehicle loss is great, such as in wars, during the monitoring of natural disasters, or even when the mission can last a long time, for example, the recognition of certain environment for more than 24 hours, making it impossible for a person on board to perform such a task with the expected quality.

The increasing interaction between people and intelligent systems, leads to a high demand of reliable representations of human knowledge. Remote-controlled robots are becoming a reality in the world, and its applications vary from delivery systems to complex surgeries. Specific aspects of interaction as cooperation and collaboration raises the ability of heterogeneous teams composed by humans and different kinds of robots, or with different resources, to solve difficult problems that rely on real time data analysis and quick responses.

To contribute to the evolution of existing navigation systems, this project is focused on the improvement of a system applied to Unmanned Aerial Vehicles (UAVs). Currently, UAVs are being intensively used for military protection purposes, recognizing battlefields or the location of the opposing army (SJANIC, 2013), but there is great potential for UAV applications for civil use, by the police, firefighters, environmental activities such as fire tracking, damage survey after catastrophes such as earthquakes, tsunamis, among others, also considering the automatic pilots of commercial aircraft, already widely used and which can also benefit from the evolution of the system. Figure 1 shows some examples of using UAVs in the civil sector. Item *a* shows the mapping and assessment of damage caused by disasters, in *b* it is possible to identify the oil spill in the ocean, the image of item *c* reveals the survey of the conditions of electric power lines and the item *d* reveals riot monitoring.

Unmanned Aerial Vehicles (UAVs) are considered tools for emergency informatics (MURPHY, 2016), a scientific field that approaches the use of different data-sets to save lives in natural or man-made disasters, through acquisition, organization and visualization of data and, consequently, which actions can be taken into account. Besides the use of UAVs for

Figure 1 – UAVs executing monitoring of civil application functions



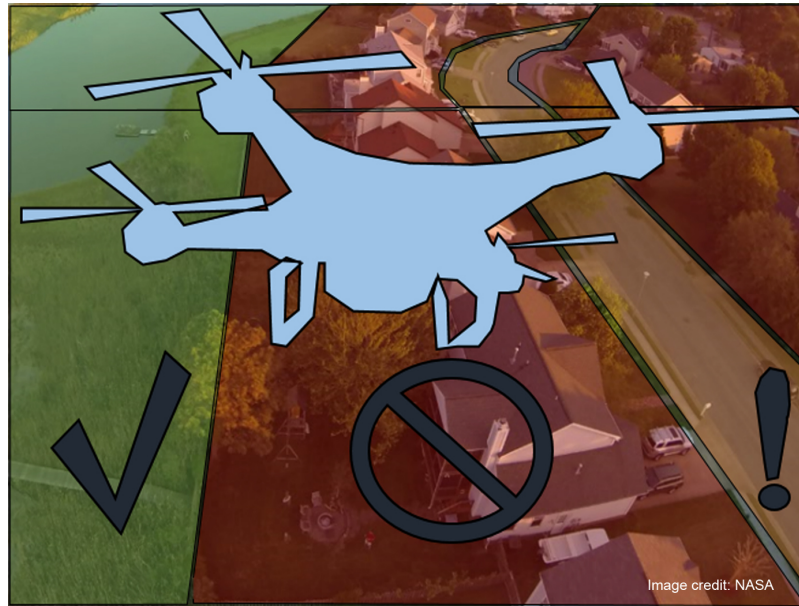
Source: (SJANIC, 2013). Images ©Saab AB.

military proposes, there is potential for civil applications, as forest fire tracking, damage survey after earthquakes or tsunamis, monitoring of riots, search and rescue of missing people, and so on. All the information contained in images, scan data, laser data, GPS position or any other knowledge captured by UAVs might not be readily interpreted by the final user, instead, a qualitative description that includes cognitive comprehensibility ensuring efficient understanding by humans without wide computational complexity can yield better results.

When we think about Unmanned Aerial Vehicles, it is likely that images come to our minds as small UAVs, known as drones, which have served mainly for recreational purposes. Drones can also be used as a platform for testing systems to be applied to large UAVs, such as the *Safe2Ditch System*¹ of Figure 2, designed to manage emergencies and safely landing the vehicle during events of an unexpected critical problem. The *Safe2Ditch System* uses the control of the UAV and the rest of the vehicle's battery in order to reach the safest possible landing location,

¹Nasa Website - <https://technology.nasa.gov/patent/LAR-TOPS-243> - accessed on 11/06/2017.

Figure 2 – Operating system *Safe2Ditch*.



Source: Nasa website - <https://technology.nasa.gov/patent/LAR-TOPS-243> - accessed 06/11/2017.

bypassing houses or people, and executing the mission autonomously, without any assistance from a pilot, security or ground station.

Figure 3 shows an example of a large-sized UAV from the company *Northrop-Grumman*, *Global Hawk* fixed wing model, used for surveillance and reconnaissance of large geographic areas, providing information about high accuracy in real time, and being able to stay in missions for more than 32 hours. This UAV is equipped with radar, high resolution optical camera and infrared sensor. The captured data can be fused in a processor or individually selected to transmit information to the operator, distinguishing different types of vehicles, aircraft, people, military devices, even under adverse weather conditions, during day or night. As it is a large equipment, measuring 40 meters, it can fly over large open areas ².

The company *Northrop-Grumman* has a partnership with NASA using the *NASA Global Hawk* model, shown in Figure 4 to help scientists on missions during hurricanes and severe storms off the US coast. The UAV control can be done from two distinct NASA bases, with the

²*Northrop-Grumman* website - 2016 - http://www.northropgrumman.com/Capabilities/GlobalHawk/Documents/GH_Brochure_B30.pdf - accessed on 11/02/2017.

Figure 3 – UAV Northrop-Grumman Global Hawk fixed wing



Source: *Northrop-Grumman website*².

Figure 4 – UAV Northrop-Grumman NASA Global Hawk



Source: *Northrop-Grumman website*³.

support of meteorologists, engineers, pilots and scientists. The UAV flew over the area of the Atlantic Ocean where hurricanes begin to form, tracking the storms to Caribbean ³.

Another UAV model already on the market is the *Skeldar V-200* from Saab AB, which measures approximately 4 meters and weighs 235 kilos, as shown in Figure 5. This model performs missions capturing information, flying over land or sea, due to his vertical takeoff

³*Northrop-Grumman website* - 2013 - http://www.northropgrumman.com/Capabilities/NASAGlobalHawk/Documents/pageDocuments/NASA-GH_Data-Sheet.pdf - accessed 11/02/2017.

Figure 5 – UAV Saab AB Skeldar V-200



Source: Saab AB website - <http://saab.com/air/airborne-solutions/unmanned-aerial-systems/skeldar-v-200-land/> - accessed 11/02/2017. Image ©Saab AB.

and landing system. It is also equipped with a high resolution optical camera, infrared sensor and synthetic and electronic aperture radar. The UAV counts on surveillance and 3D mapping capabilities which can be used during the day or at night, with a flight capacity of approximately 5 hours. Due to its small size, even being larger than recreational drones, this equipment can access smaller places, being used to monitor industrial infrastructure, pipelines in petrochemical areas, hydroelectric plants, nuclear plants and mines, reducing the risks of human inspection and sending real-time data to remote control stations ⁴.

Individually, UAVs have been able to capture and process the acquired information, according to the applications mentioned. The next challenge is to develop a system that allows a team of UAVs to map the environment using information from several points of view. Multi-agent systems are often supposed to have several advantages over single agent systems, as the capability to accomplish a single task faster, or efficiently exchange information about their position to precisely localize themselves whenever they sense each other (BURGARD et al., 2000). When multiple agents are observing a scene and occlusion prevents sensors from assessing parts of the objects present in the scene, distinct observers can provide multiple viewpoint descriptions

⁴Saab AB website - 2017 - <http://saab.com/air/airborne-solutions/unmanned-aerial-systems/skeldar-v-200-land/> - accessed 11/02/2017.

about the objects they can detect, improving the completeness of the information observed (P. E. SANTOS; LIGOZAT; SAFI-SAMGHABAD, 2015).

Furthermore, there is a problem for using external systems to aid navigation systems. The GNSS (*Global Navigation Satellite System*) can contribute significantly to improve the accuracy of UAVs auto-locating and assistance in mapping the environment, but signal availability can be affected by occlusions or by the intentional blocking of the signal, in case of more hostile environments (SJANIC, 2013). The intelligent formalism developed in this work for environment mapping from partial aerial point of view using sensors that can capture signals to scene interpretation, could turn the agents less dependent of external systems.

For intelligent systems to be able to positively interact in human environments, some characteristics are essential, such as accurate representations, efficient inference methods and proximity to human knowledge. That is why this work proposes the studying of knowledge representation and reasoning processes related to space, the development of new formalism of Qualitative Spatial Reasoning (QSR) (VAN HARMELEN; LIFSCHITZ; PORTER, 2008), an area of Artificial Intelligence that seeks to formalize spatial knowledge using elementary entities , such as spatial regions, directions, line segments, and the application of these formalism in groups of UAVs that aim to jointly solve a common spatial problem or activity, as mapping a domain.

Collaborative systems are required to integrate not only robots, but also human resources. For complex emergency scenarios such as search and rescue missions, providing assistance and guiding people to a safe destination from inhabitants lost in wilderness regions, at sea scenarios, places devastated by earthquakes, flooding or forest fires (DOHERTY et al., 2010) it is an essential role to provide an effective communication and continuous interaction between people and robotic agents in order to achieve mission goals, specially the ones related to environment exploration, so that each individual can explore different areas simultaneously (BURGARD et al., 2000).

This project aims to develop a new Qualitative Spatial Reasoning formalism for collaborative systems, used in the interpretation of scenes from multiple points of view. The intention is to develop a system based on Qualitative Spatial Reasoning (QSR) theories, integrating such theories into a single formalism to model the perceptions captured by Unmanned Aerial Vehicles (UAVs).

Unmanned Aerial Vehicles (UAVs) were chosen in this project to obtain an aerial view of the environment to be mapped, as well as to use the sensors available in these equipment as a

source of information, seeking to improve the quality of the scene to be interpreted. In addition, UAVs are safe, operated remotely, with the smaller ones operating outside regular airspace and, with the evolution in the development of sensors, communication methods and electronics, they can be easily controlled through tablets or laptops (HOLMBERG, 2015).

In the development of this work a set of formalism from Qualitative Spatial Reasoning are joined, providing a qualitative description of a scene, that can be observed by humans and robots, from different aerial points of view, in a way that all agents involved in the goal of environment mapping have the possibility to understand the position of objects and the relations between them. The reasoning presented in Region Connection Calculus (D. RANDELL; CUI; A. COHN, 1992), Allen's Interval Algebra (ALLEN, 1990) and Rectangle Calculus (BALBIANI; CONDOTTA; CERRO, 1998) are used to provide the possible relations between a pair of objects seen from an aerial point of view, while the Cardinal Direction Calculus (FRANK, 1996) provides the basis for the orientation definition of those points of view. Humans and robots will be able to exchange information and infer data about their partial view of the domain through the Spatial Collaborative Reasoning, an extension of the concepts previously developed in the Interval Occlusion Calculus (P. E. SANTOS; LIGOZAT; SAFI-SAMGHABAD, 2015).

1.1 MAIN OBJECTIVE

The goal of this work is to develop a formalism based on Qualitative Spatial Reasoning theories, for the task of environment mapping from partial aerial viewpoint, performed by multi-agent teams including humans.

To reach this goal it was studied scene interpretation using appropriate spatial constraints, in order to enable agents to assimilate knowledge about an environment in situations of partial view of the scene from multiple points of view, captured by UAVs.

The objective of applying such restrictions to multi-agent systems with multiple points of view, including humans in the team, is to collectively solve problems considered too difficult or even impossible to be solved by a single agent, enabling the interaction between humans and robots. Qualitative theories enables the exchange of information between humans and robotic agents, so that they can perform tasks in collaborative missions involving searching and monitoring objectives in agriculture, natural disasters, searching and rescue tasks, among others.

This work also develops the method to mediate the process of obtaining joint solutions, studying and implementing a system in which solutions generated by multiple agents are integrated in a collaborative process, where qualitative spatial constraints are satisfied.

1.2 CONTRIBUTIONS

To achieve the objective of this work through purely qualitative techniques, the starting point was the study of Qualitative Spatial Reasoning (QSR), a sub-field of knowledge representation research that deals with knowledge about an infinite spatial domain using a finite set of qualitative relations, modeling human common sense to understand space (WOLTER; WALLGRÜN, 2013). Even though quantitative approaches can provide more accurate information, specific numerical data can be unavailable to humans (J. CHEN et al., 2015). One of the conditions to reach the objective of this work is to enable the active interaction between humans and robotic agents, so, qualitative constraints was chosen to be the language of communication.

Starting from the most embracing formalism about object (that can be considered a spatial region) relations at a scene, it was studied the Region Connection Calculus, that describes the basic relations between spatial regions through a constraint language containing 8 JEPD (Joint Exclusive and Pairwise Disjoint) base relations, which allows reasoning about topological distinctions, with possibility to infer new spatial relations and transitions from incomplete spatial knowledge (D. RANDELL; CUI; A. COHN, 1992; RENZ, 2002), a key point for the development of this work.

Considering that a spatial region can be outlined by a bounding box, the edges of the bounding box can be considered intervals, from where it is possible to get information from Allen's Interval Algebra. This theory describes a temporal representation and reasoning where temporal interval is considered as a primitive. This method represents the relations between pairs of reference intervals taking into account their upper and lower limits in a hierarchical manner, resulting in a set of 13 JEPD base relations (ALLEN, 1990).

As a bounding box has edges parallel to axis x and y , the Rectangle Calculus (BALBIANI; CONDOTTA; CERRO, 1998) defined as a model for reasoning about bi-dimensional temporal relations, whose objects are rectangles which sides are parallel to the axes of some orthogonal basis in a 2-dimensional Euclidean space can also contribute with qualitative information about the objects present in an environment.

To represent the qualitative information from each agent's point of view, this work is based on the notation developed in the Interval Occlusion Calculus, presenting a qualitative description of a set of basic relations between pairs of objects observed from a point of view, given the object's lines of sight (P. E. SANTOS; LIGOZAT; SAFI-SAMGHABAD, 2015).

At least, to get information about the direction of flight of each UAV used to capture images of the environment to be mapped, it was used the Cardinal Direction Calculus (FRANK, 1996) information. The Cardinal Direction Calculus is a formalism to reason about cardinal directions between objects and its relations, being composed by 9 basic relations including a neutral region.

The research and combination of information resulting from the formalism above described, resulted in two contributions, described in the following subsections.

1.2.1 First Contribution

The first contribution of this work is the development of a formalism to identify objects in an environment, consider these objects as regions and enable the acquisition of information related to the Region Connection Calculus, Allen's Interval Algebra, Rectangle Calculus and Interval Occlusion Calculus into a single constraint, which can be acquired from different points of view. This formalism called LH Interval Calculus uses the bounding boxes around the objects present in the environment to be mapped, as well as the projection of its outlines on axis s and y to obtain the relations between the objects from an aerial perspective. The complete development of LH Interval Calculus is presented in Section 4.2.

1.2.2 Second Contribution

The second contribution is a formalism that allows the combination of the information contained into two images with partial and complementary views, captured in the same environment, by different agents. The information acquired through the LH Interval Calculus is associated with the Cardinal Direction Calculus, resulting in sentences that contains the complete information about the environment. This formalism is called Collaborative Spatial Reasoning and is presented in Section 4.3.

1.3 THESIS STRUCTURE

This work is divided into chapters, according to the structure:

- a) Chapter 2 - Literature Review: concepts necessary for the development of this work and research related to qualitative reasoning theorems;
- b) Chapter 3 - Theoretical Review: theories that form the foundation of this work;
- c) Chapter 4 - Method: LH Interval Calculus and Collaborative Spatial Reasoning description;
- d) Chapter 5 - Experiments: description of the experiments to verification of the formalisms developed;
- e) Chapter 6 - Conclusion and future work.

2 THEORY

This chapter describes the theories on which this research proposal is based. The main point of these investigations involve the concepts of Cognitive Robotics (Section 2.1), Qualitative Spatial Reasoning (QSR) (Section 2.2), Region Connection Calculus (RCC) (Section 2.2.1) and Cardinal Direction Calculus (CDC) (Section 2.2.2), Allen's Interval Algebra (Section 2.2.3), Rectangle Calculus (Section 2.2.4), Interval Occlusion Calculus (Section 2.2.5) and Simultaneous Localization and Mapping (Section 2.3).

2.1 COGNITIVE ROBOTICS

The decision-making for performing tasks in autonomous agents continuously works with perception of the agents themselves and the environment in which they are immersed, and action (NEHMZOW, 2012), determining which task to perform and how to perform it. It is expected to create robots which can learn, understand, and execute tasks as human beings would do (PAULIUS; SUN, 2019). The integration of different levels of control architecture, involving the choice and execution of an integrated formalization up to the cognitive aspects, is one of the challenges of cognitive robotics (PIRRI; MENTUCCIA; STORRI, 2003).

The search for the best training techniques for a robot to perform tasks autonomously aims at making machines to perform tasks in an acceptable and reliable way by human evaluation, many of them performed in collaboration between humans and robots (KANDA et al., 2007).

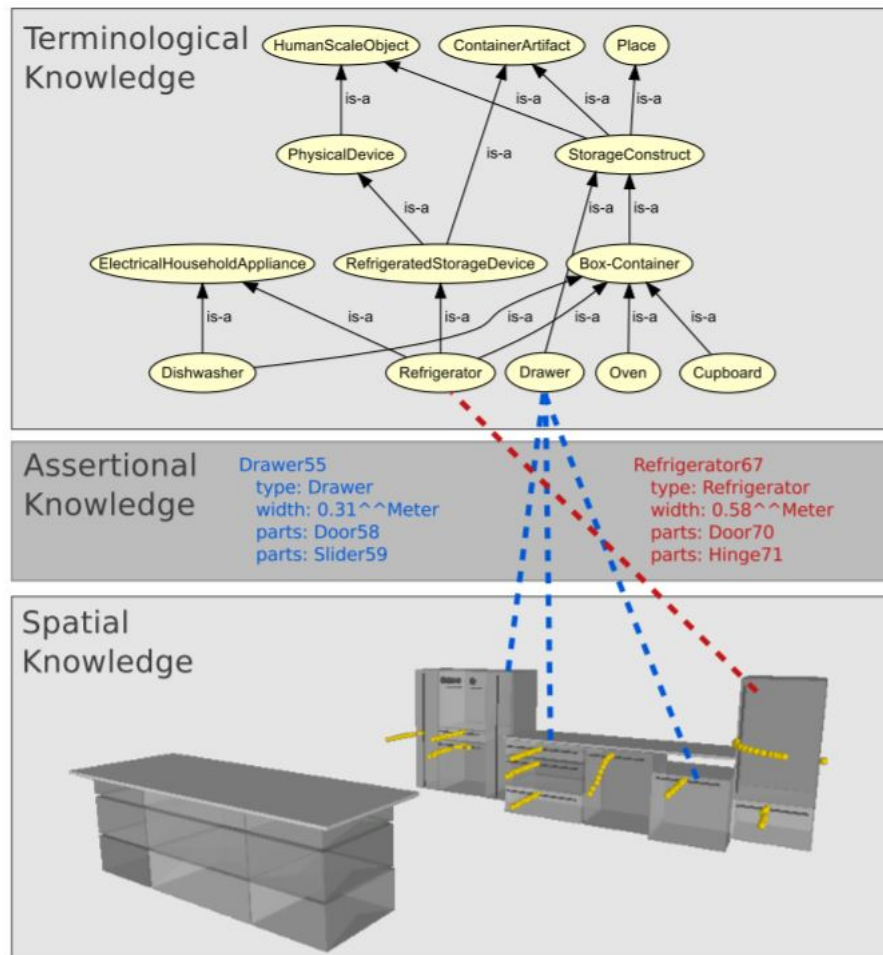
To perform meaningful tasks, the robot should be able to understand its environment and actions, integrating learning methods and high-level knowledge. In artificial intelligence, knowledge representation is grounded in formal logical expressions, symbolically and semantically represented and manipulated by reasoning programs (PAULIUS; SUN, 2019). It is common to write relations (facts or rules) as clauses that can be used in queries to reasoning and infer new concepts initially unknown to the agent as implicit knowledge.

High-level knowledge denotes a semantic, structural representation of relations between different components, giving meaning to low-level representations inputs as perception modules as audio, vision, or touch, sensor inputs, motion primitives and actuation systems, motion and task planning or inference systems. The ambivalent description of objects/tools and manipulations in high and low level representations enables the robot to communicate with humans (PAULIUS; SUN, 2019).

A constrained set of terms or language defines an ontology, used to label and describe the working space of a robotic agent, combining several types of information, in a format understandable by humans. Cognitive architectures emphasize on how knowledge is learned and retained, and are activated once their arguments have been fulfilled by identifying them through the robot's perception system (PAULIUS; SUN, 2019).

The combination of geometric data with semantic data is proposed in Semantic Object Maps (SOM) (PANGERCIC et al., 2012), with the objective of answering queries to determine whether certain action can be executed given present circumstances in its environment. Figure 6 shows an instance that can be inferred to be a kitchen because there are items within the environment that are typical of a kitchen, such as a stove or a fridge (PANGERCIC et al., 2012).

Figure 6 – Semantic Object Map representation.



Source: (PANGERCIC et al., 2012).

Researches focused on using robotic perception are usually carried out in controlled environments and performing specific tasks, given the degree of uncertainty existing in dynamic

environments, such as (SCHERL; LEVESQUE, 1993), which used the information perceived by a robot to produce knowledge from its actions, or in (P. E. SANTOS; SHANAHAN, 2002) where the data collected by the sensors of a mobile robot formed the basis for the construction of a spatial reasoning system.

However, researchers have realized that it is possible to elevate robotic cognitive capacity, adding different algorithms developed to perform tasks (MCCARTHY, 2007), as proposed in this research, using Qualitative Spatial Reasoning in the formalization of Collaborative Spatial Reasoning for scene interpretation.

2.2 QUALITATIVE SPATIAL REASONING

Qualitative Spatial Reasoning (QSR) is an area of Artificial Intelligence that seeks the formalization of spatial knowledge and the inference of facts about the structure of a domain from elementary entities such as spatial regions (D. RANDELL; CUI; A. COHN, 1992), line segments (MORATZ; RENZ; WOLTER, 2000; SCHLIEDER, 1996), directions (FREKSA, 1992; LIGOZAT, 1998), topology, distance between objects (A. G. COHN; HAZARIKA, 2001), among others (STOCK, 1998; A. G. COHN; HAZARIKA, 2001; J. CHEN et al., 2015), in addition to making explicit common sense knowledge, coupled with abstractions to create mathematical models (A. G. COHN et al., 1997). From the mentioned entities and their transitions, axioms are created to represent the knowledge about space and adequate methods to represent continuous properties in the world, using a system based on discrete symbols.

In a QSR system for a robotic agent, the agent is equipped with computer vision to interpret the three-dimensional world from two-dimensional information images, where bodies can be defined as physical objects, and regions can be divided into two sets: three-dimensional, which denote the volume occupied by the body, and two-dimensional, which denote the image of the projected body, seen from a point of view (D. RANDELL; WITKOWSKI; SHANAHAN, 2001).

Uncertainty is an intrinsic feature on capturing information or representing knowledge in the real world, especially in dynamic environments due to human error or contradictory data of different sources (SIOUTIS; WOLTER, 2021). Even a qualitative approach dealing with a certain degree of uncertainty, in some situations it is not possible to represent the qualitative relations properly (J. CHEN et al., 2015). A more comprehensive synthesis of this problem can be found in (A. G. COHN; HAZARIKA, 2001; J. CHEN et al., 2015).

Even so, the works (YAMAN; NAU; SUBRAHMANIAN, 2004; MULLER, 1998) were developed in dynamic environments and address the use of QSR in the formalization of a Logic of Motion (LOM). The applications of this logic ranges from real-time monitoring of several aircrafts till information about the movement of people with cell phones, for better choice of positioning of transmission towers, for example.

Maintaining the consistency in space and time are the basics in qualitative reasoning when solving spatial and temporal problems (FALOMIR; PICH; COSTA, 2020). Qualitative Shape Descriptors (QSD) include qualitative angles and lenght, and is an area of Qualitative Spatial Reasoning applied to compute shape similarity, being effective to mosaic building, for example (FALOMIR; PICH; COSTA, 2020).

Qualitative Shape Descriptors can describe straight lines, geometrically defined by a segment with a starting and an ending point, and also curves. In object composition operators, the length comes from relating an edge from an object with another edge from a different object. As all the objects involved in the composition are known, the largest object is taken as a reference and the rest of the edges are compared in relation to it (FALOMIR; PICH; COSTA, 2020).

The Length Reference System is a set of qualitative symbols and structure relations for interval values, obtained by trigonometry. The set of relations of the Length Reference System is composed by the following relations: smaller-short (ss), short (s), larger-short (ls), quarter-longest (ql), smaller-medium (sm), medium (m), larger-medium (lm), half-longest (hl), smallerlong (sl), long (l), larger-long (ll), longest (lst), largertan-longest (llst) (FALOMIR; PICH; COSTA, 2020).

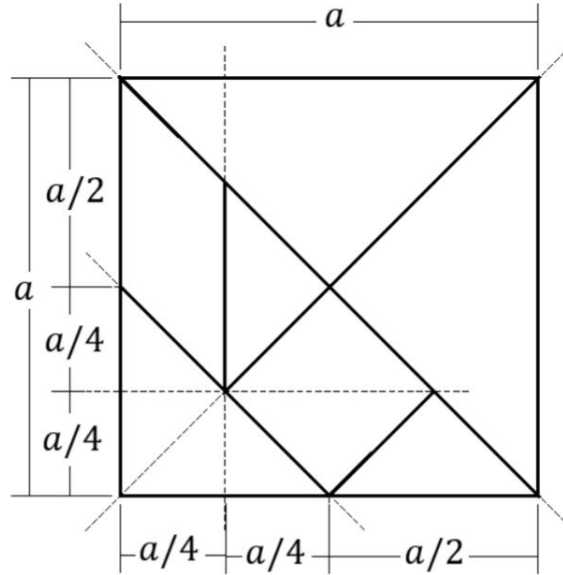
The baseline for the Length Reference System is showed in Figure 7.

The composition table of Figure 8 relates all the different combinations of qualitative lengths.

It is found the use of QSR in geographic applications, such as in the works of (SCHULTZ; GUESGEN; AMOR, 2006; TORRES et al., 2016; WALLGRÜN, 2010), aiming at learning topological maps in unknown environments, methods to deal with ambiguity in topological information, in addition to human-computer and systems integration of geographic information. These themes are important for this research proposal, for the aspects of environment mapping and the human-computer interaction coping with the ongoing efforts to implement tools for practical tasks such as data verification, repair and visualization.

Among the QSR attributes, the characteristic "region" was used for the development of the Region Connection Calculus, discussed in Section 2.2.1, one of the best known theories to

Figure 7 – Relations of length among Tangram pieces.



Source: (FALOMIR; PICH; COSTA, 2020).

Figure 8 – Composition table for length in Qualitative Shape Descriptors.

	ss	s	ls	ql	sm	m	lm	hl	sl	l	ll	lst	llst
ss	{ss,s,ls,ql,sm}	{ls,ql,sm}	{ls,ql,sm,m,lm}	{sm,m,lm}	{sm,m,lm,hl,sl}	{lm,hl,sl}	{lm,hl,sl}	sl	{sl,l,ll}	ll	{ll,lst,llst}	llst	llst
s	{ls,ql,sm}	m	lm	lm	{lm,hl,sl}	sl	sl	sl	{sl,l,ll}	ll	{ll,lst,llst}	llst	llst
ls	{ls,ql,sm,m,lm}	lm	lm	lm	{lm,hl,sl}	sl	{sl,l,ll}	{sl,l,ll}	{sl,l,ll}	ll	{ll,lst,llst}	llst	llst
ql	{sm,m,lm}	lm	lm	hl	sl	sl	{sl,l,ll}	ll	ll	ll	{ll,lst,llst}	llst	llst
sm	{sm,m,lm,hl,sl}	{lm,hl,sl}	{lm,hl,sl}	sl	sl	sl	{sl,l,ll}	ll	{ll,lst}	{ll,lst}	{ll,lst,llst}	llst	llst
m	{lm,hl,sl}	sl	sl	sl	sl	l	ll	ll	{ll,lst,llst}	llst	llst	llst	llst
lm	{lm,hl,sl}	sl	{sl,l,ll}	{sl,l,ll}	{sl,l,ll}	ll	ll	ll	{ll,lst,llst}	llst	llst	llst	llst
hl	sl	sl	{sl,l,ll}	ll	ll	ll	ll	lst	{ll,lst,llst}	llst	llst	llst	llst
sl	{sl,l,ll}	{sl,l,ll}	{sl,l,ll}	ll	{ll,lst,llst}	{ll,lst,llst}	{ll,lst,llst}	{ll,lst,llst}	llst	llst	llst	llst	llst
l	ll	ll	ll	ll	{ll,lst,llst}	llst	llst	llst	llst	llst	llst	llst	llst
ll	{ll,lst,llst}	{ll,lst,llst}	{ll,lst,llst}	{ll,lst,llst}	{ll,lst,llst}	llst	llst	llst	llst	llst	llst	llst	llst
lst	llst	llst	llst	llst	llst	llst	llst	llst	llst	llst	llst	llst	llst
llst	llst	llst	llst	llst	llst	llst	llst	llst	llst	llst	llst	llst	llst

Source: (FALOMIR; PICH; COSTA, 2020).

represent and reason about topological relations, being the basis for others Qualitative Spatial Reasoning approaches (J. CHEN et al., 2015).

2.2.1 Region Connection Calculus

The Region Connection Calculus (RCC) (D. RANDELL; CUI; A. COHN, 1992) consists of a model of representation and axiomatic reasoning about relations concerning spatial regions in first order logic, which is based on a primitive binary relation of symmetric and reflexive connectivity. The basis of these relations is the binary primitive “connected”, given two regions x and y , $C(x,y)$, read as “the region x is connected to the region y ”, being true if and only if the regions x and y have at least one point in common.

The Region Connection Calculus represents the mereotopological relations between spatial regions that are independent from the observer’s point of view, unlike the *Lines-of-Sight Calculus* (GALTON, 1994) which, from one point of view, represents the relative positions between pairs of non-overlapping convex bodies.

The connectivity relations defined by the Region Connection Calculus obey the following axioms (D. RANDELL; CUI; A. COHN, 1992; GOTTS, 1994):

$$\begin{aligned} &\forall x C(x,x); \\ &\forall x y C(x,y) \rightarrow C(y,x). \end{aligned}$$

Assuming the binary relation $C/2$ and the variables x , y and z for spatial regions, different relations between pairs are defined. The relations established by the RCC and its interpretations are described in Table 1 (D. RANDELL; CUI; A. COHN, 1992; RENZ, 2002). The relations DC , EQ , PO , EC , TPP , $NTPP$, $TPPI$ and $NTPPI$ form a set of mutually exclusive and jointly exhaustive and pairwise disjoint, that is, two regions cannot be represented by more than one RCC relation. This set is referred as RCC-8 (D. RANDELL; CUI; A. COHN, 1992; STOCK, 1998; RENZ, 2002).

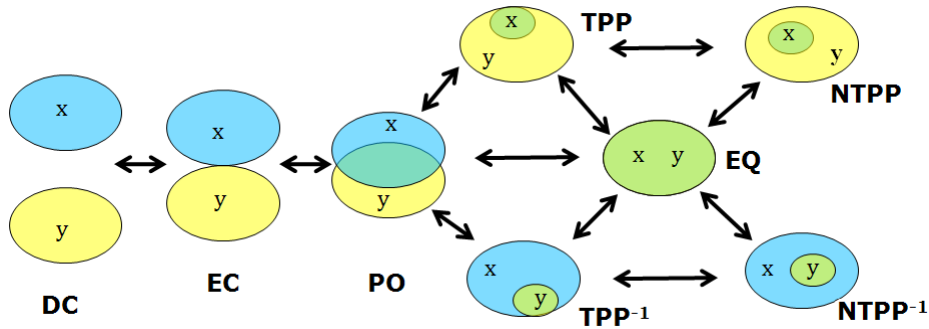
The change in the relations between regions defined by the RCC occurs continuously, that is, when a relation changes, it results in a close or neighboring relation. Figure 9 graphically represents a subset of the RCC, containing the eight basic relations called RCC-8. In Figure 9 the transitions are represented by arrows determining that, between two relations connected by an arrow, there is no other possible relation. This representation is known as Conceptual Neighborhood Diagram (CND) (CUI; A.G. COHN; D.A. RANDELL, 1992; D. RANDELL; CUI; A. COHN, 1992).

Table 1 – RCC relations and interpretations.

RCC Relations Definition	RCC Relations Interpretation
$DC(x,y) \equiv_{def} \neg C(x,y)$	' x is disconnected from y '
$P(x,y) \equiv_{def} \forall z[C(z,x) \rightarrow C(z,y)]$	' x is part of y '
$PP(x,y) \equiv_{def} P(x,y) \wedge \neg P(y,x)$	' x is proper part of y '
$EQ(x,y) \equiv_{def} P(x,y) \wedge P(y,x)$	' $x = y$ ': ' x is equal to y '
$O(x,y) \equiv_{def} \exists z[P(z,x) \wedge P(z,y)]$	' x overlaps y '
$PO(x,y) \equiv_{def} O(x,y) \wedge \neg P(x,y) \wedge \neg P(y,x)$	' x partially overlaps y '
$DR(x,y) \equiv_{def} \neg O(x,y)$	' x is discrete from y '
$EC(x,y) \equiv_{def} C(x,y) \wedge \neg O(x,y)$	' x is externally connected to y '
$TPP(x,y) \equiv_{def} PP(x,y) \wedge \exists z[EC(z,x) \wedge EC(z,y)]$	' x is tangential proper part of y '
$NTPP(x,y) \equiv_{def} PP(x,y) \wedge \neg \exists z[EC(z,x) \wedge EC(z,y)]$	' x is non-tangential proper part of y '
$PI(x,y) \equiv_{def} P(y,x)$	' y is part of x '
$PPI(x,y) \equiv_{def} PP(y,x)$	' y is proper part of x '
$TPPI(x,y) \equiv_{def} TPP(y,x)$	' y is tangential proper part of x '
$NTPPI(x,y) \equiv_{def} NTPP(y,x)$	' y is non-tangential proper part of x '

Source: (D. RANDELL; CUI; A. COHN, 1992; RENZ, 2002).

Figure 9 – RCC-8 Conceptual Neighborhood Diagram.



Source: (FENELON, 2014), adapted from (CUI; A.G. COHN; D.A. RANDELL, 1992).

Extending the RCC concept to three objects a , b and c , Table 4 represents the RCC-8 Composition Table, where the first column represents the relations between a and b ($R_1(a,b)$) and the first line the relations between b and c ($R_2(b,c)$). The set of possible relations between a and c ($R_3(a,c)$) are represented in the cells defined by the intersection between the column of $R_1(a,b)$ and the line from $R_2(b,c)$.

The Composition Table is a tool used to represent and reason about the disjunctions and conjunctions of pairs of the set of binary relations of the RCC, both spatial and temporal (D. RANDELL; WITKOWSKI, 2002). This table contains the transitivities between three distinct regions a, b and c , where, given the information about the relations between a and b and between b and c , it is extracted a set of possible relations between the regions a and c . (D. RANDELL; WITKOWSKI, 2002).

The information shown in Table 4, where it is possible to infer information about objects that did not have their data collected, is interesting to this research work because, as different points of view are present in a scene, it is possible to infer environmental characteristics or

Table 2 – RCC-8 Composition Table.

$R_1(a,b)$	$R_2(b,c)$							
	DC	EC	PO	TPP	TPPI	EQ	NTPP	NTPPI
DC	DC, EC, PO, TPP, TPPI, EQ, NTPP, NTPPI	DC, EC, PO, TPP, NTPP	DC, EC, PO, TPP, NTPP	DC, EC, PO, TPP, NTPP	DC	DC	DC, EC, PO, TPP, NTPP	DC
EC	DC, EC, PO, TPPI, NTPPI	DC, EC, PO, TPPI, TPPI, EQ	DC, EC, PO, TPP, NTPP	EC, PO, TPP, NTPP	DC, EC	EC	PO, TPP, NTPP	DC
PO	DC, EC, PO, TPPI, NTPPI	DC, EC, PO, TPPI, NTPPI	DC, EC, PO, TPP, TPPI, EQ, NTPP, NTPPI	PO, TPP, NTPP	DC, EC, PO, TPPI	PO	PO, TPP, NTPP	DC, EC, PO, TPPI, NTPPI
TPP	DC	DC, EC	DC, EC, PO, TPP, NTPP	TPP, NTPP	DC, EC, PO, TPP, TPPI, EQ	TPP	NTPP	DC, EC, PO, TPPI, NTPPI
TPPI	DC, EC, PO, TPPI, NTPPI	EC, PO, TPPI, NTPPI	PO, TPPI, NTPPI	PO, TPP, TPPI, EQ	TPPI, NTPPI	TPPI	PO, TPP, NTPP	NTPPI
EQ	DC	EC	PO	TPP	TPPI	EQ	NTPP	NTTPI
NTPP	DC	DC	DC, EC, PO, TPP, NTPP	NTPP	DC, EC, PO, TPP, NTPP	NTPP	NTPP	DC, EC, PO, TPP, TPPI, EQ, NTPP, NTPPI
NTPPI	DC, EC, PO, TPPI, NTPPI	PO, TPPI, NTPPI	PO, TPPI, NTPPI	PO, TPPI, NTPPI	NTPPI	NTPPI	PO, TPP, TPPI, EQ, NTPP, NTPPI	NTPPI

Source: (D. RANDELL; WITKOWSKI, 2002).

relations through the different information captured by each agent, following a logical formalism such as the method shown in the RCC-8 Composition Table.

A qualitative reasoning to deal with different perspectives perceived by multiple agents observing a scene called Cardinal Direction Calculus is described in the next section.

2.2.2 Cardinal Direction Calculus

The Cardinal Direction Calculus (CDC) (FRANK, 1996) it is another formalism from Qualitative Spatial Reasoning to reasoning about directions between spatial objects and their relations. The formalism is composed by nine basic relations, considering a neutral region: north

Table 3 – Cardinal Direction Calculus Composition Table.

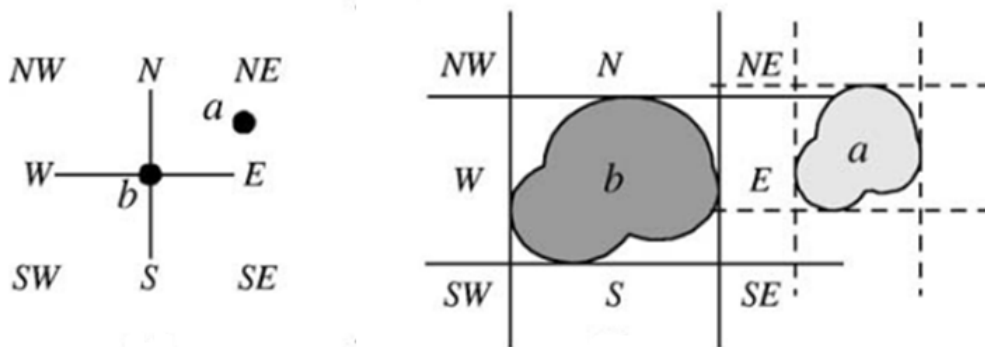
	N	NE	E	SE	S	SW	W	NW	0
N	N	N	NE	e	o	w	NW	NW	N
NE	NE	NE	NE	e	e	o	n	n	NE
E	NE	NE	E	SE	SE	s	o	n	E
SE	e	e	SE	SE	SE	s	s	o	SE
S	o	e	SE	SE	S	SW	SW	w	S
SW	w	o	s	s	SW	SW	SW	w	SW
W	NW	n	o	s	SW	SW	W	NW	W
NW	NW	n	n	o	w	w	NW	NW	NW
0	N	NE	E	SE	S	SW	W	NW	0

Source: (LIGOZAT, 1998).

(*n*), east (*e*), west (*w*), south (*s*), northwest (*nw*), northeast (*ne*), southeast (*se*), southwest (*sw*) and equal (*eq*) corresponding to the central neutral region. Each relation can be seen in Figure 10.

At first, (FRANK, 1996) studied the division of spaces into quadrants and into cone-shaped, concluding that the quadrant shape contains a greater number of exact deductions in a composition table. The quadrant approach is shown in Figure 10 being the related objects represented by points (Figure 10 left) or by regions (Figure 10 right).

Figure 10 – Cardinal Direction Calculus representation, quadrant projection.



(LIGOZAT, 1998).

Source:

As well as the Region Connection Calculus, the Cardinal Direction Calculus also has a Composition Table, showed in Table 3.

Through the CDC relations it is possible to infer the direction between two objects *A* and *C*, from the knowledge about the direction between *A* and (another object) *B* and between *B* and *C*. This inference derives from the Composition Table between the nine basic relations, shown in Table 3, being the result a set of exclusively disjoint possibilities (LIGOZAT, 1998). The concept of neighborhood is also applicable due to the topology of the relations defined by the proximity between regions, for example, the transition that can occur from the southeast relation *se* is the

subset composed of the relations s (south), e (east), se (southeast) and eq (equal) (LIGOZAT, 1998).

In (FENELON, 2014) the concepts related to direction relations were used in the task of self-localization of a mobile robot in relation to multiple objects, where the transition relations and distance calculus were used with topological calculus based on the relations of occlusion to formalize the robot's location through a qualitative map.

The Qualitative Spatial Reasoning theories described here are used to infer information from images obtained by cameras installed in UAVs. In a real environment, the system has to deal with uncertainties caused by sensor noise or any kind of information loss.

2.2.3 Allen's Interval Algebra

From the perspective of artificial intelligence, Allen's Interval Algebra describes a temporal representation and reasoning where temporal interval is considered as a primitive. This method represents the relationships between pairs of reference intervals taking into account their upper and lower limits in a hierarchical manner, resulting in a set of 13 jointly-exhaustive and pairwise-disjoint base relations (ALLEN, 1990). Figure 11 shows Allen's relations for two intervals, x and y .

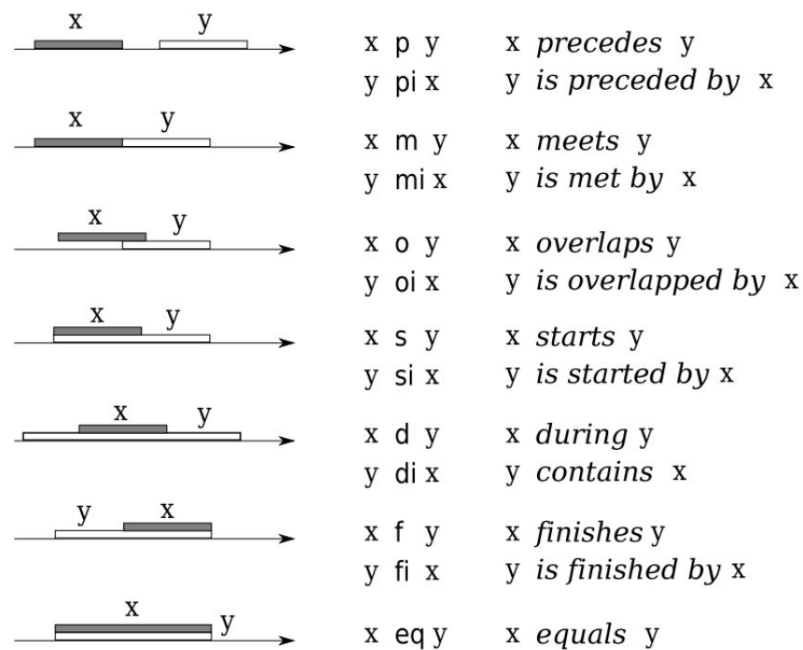


Figure 11 – Allen's Relations

(P. E. SANTOS; LIGOZAT; SAFI-SAMGHABAD, 2015), adapted from (ALLEN, 1990).

Considering a constraint network pair (N, C) , N being the set of vertices where each domain element is represented by a vertex, and C representing the set of constraints defined by the basic Allen's relations, it is possible to verify the existence of a consistent scenario by imposing algebraic closure on the network of constraints, to confirm whether the configurations provided by the information from the agents are feasible in at least one scenario of a domain described by the set of Allen's relations (P. E. SANTOS; LIGOZAT; SAFI-SAMGHABAD, 2015).

A network (N, C) is algebraically closed if its three vertices $(i, j, k) \in N^3$, and is consistent with the composition $C(i, j) \subseteq C(i, k) \circ C(k, j)$. This concept supports the Spatial Collaborative Reasoning proposed in this work and the definition of the relations of the objects identified by agents with a partial view of the environment.

In the work of (FALOMIR; PICH; COSTA, 2020), a qualitative model for angles is proposed, where an angular relation α is defined by three points: a, b, c . The Angle Reference System is composed by a set of spatial relations [very-acute (va), half-right (hr), acute (a), right (r), obtuse (o), 3q-right (3qr), very-obtuse (vo), plane (pl), full], which contains a set of qualitative symbols and a set of structure relations defining the acceptance areas or interval values for each qualitative symbol.

Figure 12 shows the Qualitative Angles and their Convexities, where a convex angle (cx) refers to the angle located in the interior of the object, while a concave angle (cv) refers to the angle on the exterior of the object.

The spatial relations between angles can be defined similarly to Allen's relations between time intervals, according to the information of Figure 13 (FALOMIR; PICH; COSTA, 2020).

The spatial relations between angles resulted in two composition tables, one for convex (cx) angles (Figure 14) and another for convex (cx) and concave (cv) angles (Figure 15).

An extension of Allen's Interval Algebra called *Rectangle Calculus* was developed proposing to reason about bi-dimensional temporal relations, and is described in the next section.

2.2.4 Rectangle Calculus

The Rectangle Calculus can be defined as a model for reasoning about bi-dimensional temporal relations, whose objects are rectangles which sides are parallel to the axes of some orthogonal basis in a 2-dimensional Euclidean space (BALBIANI; CONDOTTA; CERRO, 1998). The rectangle model is defined by the relational structure $M = (R, m_1, m_2)$, where R is a nonempty set of rectangles and m_i a binary relation on R .

Figure 12 – Qualitative Angles.

A_{LAB_2}	A_{RS} A_{INT_2}	Convexity	Geometry depiction
va	$(0, 45)$	cx	
hr	≈ 45	cx	
a	$(45, 90)$	cx	
r	≈ 90	cx	
o	$(90, 135)$	cx	
$3qr$	≈ 135	cx	
vo	$(135, 180)$	cx	
pl	≈ 180	cx/cv	
vo	$(180, 225)$	cv	
$3qr$	≈ 225	cv	
o	$(225, 270)$	cv	
r	≈ 270	cv	
a	$(270, 315)$	cv	
hr	≈ 315	cv	
va	$(315, 360)$	cv	
$full$	$\approx 360 \vee \approx 0$	cv/cx	

Source: (FALOMIR; PICH; COSTA, 2020).

Figure 13 – Spatial relations between angles.

	$(\alpha, \alpha') \in d$ $a \neq a'$ $b \neq b'$ $c \neq c'$	α is disjoint (d) α'
	$(\alpha, \alpha') \in e$ $a = a'$ $b = b'$ $c = c'$	α same (e) α'
	$(\alpha, \alpha') \in i$ $a = a'$ $b = b'$ $c = c'$	α inverse (i) α'
	$(\alpha, \alpha') \in m$ $(\alpha', \alpha) \in m^{-1}$ $b = a' \vee b' = a$ $c = c'$ α, α' convex	α meets (m) α' α' is met by α
	$(\alpha, \alpha') \in o$ $(\alpha', \alpha) \in o^{-1}$ $b \neq a' \wedge b' \neq a$ $c = c'$ α, α' convex	α overlaps (o) α' α' is overlapped by α
	$(\alpha, \alpha') \in s$ $(\alpha', \alpha) \in s^{-1}$ $a = a'$ $c = c'$ α is concave α' is convex	α starts (s) α' α' is started by α
	$(\alpha, \alpha') \in f$ $(\alpha', \alpha) \in f^{-1}$ $b = b'$ $c = c'$ α is concave α' is convex	α finishes (f) α' α' is finished by α
	$(\alpha', \alpha) \in d$ $(\alpha, \alpha') \in d^{-1}$ $c = c'$ α is concave α' is convex α is divided into $\alpha_1(aca')$ $\alpha_2(b'cb)$	α' during (d) α α includes α'

Source: (FALOMIR; PICH; COSTA, 2020).

Figure 14 – Composition Table for angle feature in Qualitative Shape Descriptors when composing convex (cx) angles.

meet _{op} : cx/cx	va	hr	a	r	o	3qr	vo	pl
va	{va, hr, a}	a	{a,o,r}	o	{o,3qr,vo}	vo	{vo,pl}	vo
hr	a	r	o	3qr	vo	pl	vo	3qr
a	{a,r,o}	o	{o,3qr,vo}	vo	{vo,pl}	vo	{vo,3qr,o}	o
r	o	3qr	vo	pl	vo	3qr	o	r
o	{o,3qr,vo}	vo	{vo,pl}	vo	{vo,3qr,o}	o	{o,r,a}	a
3qr	vo	pl	vo	3qr	o	r	a	hr
vo	{vo,pl}	vo	{vo,3qr,o}	o	{o,r,a}	a	{a,hr,va}	va
pl	vo	3qr	o	r	a	hr	va	full

Source: (FALOMIR; PICH; COSTA, 2020).

Figure 15 – Composition Table for angle feature in Qualitative Shape Descriptors when composing convex (cx) and concave (cv) angles.

starts _{op} , finish _{op} : cx/cv	pl	vo	3qr	o	r	a	hr	va
va	vo	{vo,3qr,o}	o	{o,r,a}	a	{a,hr,va}	va	{va,full,#}
hr	3qr	o	r	a	hr	va	full	#
a	o	{o,r,a}	a	{a,hr,va}	va	{va,full,#}	#	#
r	r	a	hr	va	full	#	#	#
o	a	{a,hr,va}	va	{va,full,#}	#	#	#	#
3qr	hr	va	full	#	#	#	#	#
vo	va	{va,full,#}	#	#	#	#	#	#
pl	full	#	#	#	#	#	#	#

Source: (FALOMIR; PICH; COSTA, 2020).

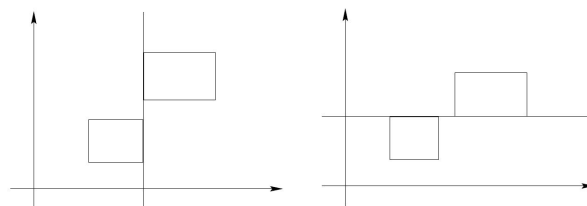


Figure 16 – Two different examples where relation *meets* applies.

(BALBIANI; CONDOTTA; CERRO, 1998).

In this work, the bounding box that enfolds the objects present in an image are considered rectangles, following the same reasoning of the Rectangle Calculus, but extended to two axis projections, as explained in Section 4.2. The left configuration of Figure 16 represents a situation that can be read as $x m_1 y$ or the same vertical line reaches the right side of x and the left side of y . Consequently, the right configuration can be read as $x m_2 y$ or the same horizontal line reaches the top side of x and the bottom side of y (BALBIANI; CONDOTTA; CERRO, 1998).

According to (ALLEN, 1990), Interval Algebra has 13 atomic relations: $A_{int} = \{p, m, o, s, d, f, pi, mi, oi, si, di, fi, eq\}$. Extending this reasoning to Rectangle Algebra, it is possible to say that it has 13×13 atomic relations, which constitute the exhaustive list of 169 pairs of relations which can hold between two rational rectangles. Let “ \times ” denotes the cartesian product and “ \circ ” denotes the composition of a pair of relations, the consistency of the composition $C(i, j) \subseteq C(i, k) \circ C(k, j)$ for the rectangle algebra includes:

The cartesian product of two pairs of relations, defined by $(A, B) \times (C, D) = (A, C), (A, D), (B, C), (B, D)$.

The rectangle composition, defined by $(A, B) \circ (C, D) = (A \circ C) \times (B \circ D)$ (BALBIANI; CONDOTTA; CERRO, 1998).

Allen’s Interval Algebra is also extended to the *Interval Occlusion Calculus*, developed to reason about occlusion from multiple points of view, and described in the next section.

2.2.5 Interval Occlusion Calculus

Interval Occlusion Calculus (IOC) is a qualitative description of a set of basic relations between pairs of objects observed from a point of view, given the object’s lines of sight (P. E. SANTOS; LIGOZAT; SAFI-SAMGHABAD, 2015). Considering that two objects A and B can be observed from a point of view Σ , the function *image* defined by $a = image(A, \Sigma)$ and $b = image(B, \Sigma)$, maps the image of a physical body seen from a viewpoint Σ . Figure 17 illustrates an example of two bodies A and B , as well as a map based on the object’s lines of sight, enabling the observers to locate themselves with respect to the qualitative relations observed between the images of the objects. If Σ is located in the region of the map marked by p , then the observer will notice that a precedes b , or if Σ is located in the region o^+ , it could be seen that a overlaps and is in front of b . The same reasoning applies to all other positions of the map.

Reasoning about different points of view, it is possible to interpret that if Σ_1 sees $\{a p b\}$, $\{\sigma_2 p_i a\}$, and $\{\sigma_2 p b\}$; then Σ_2 is in Region (2) of the map and its possible observations is the

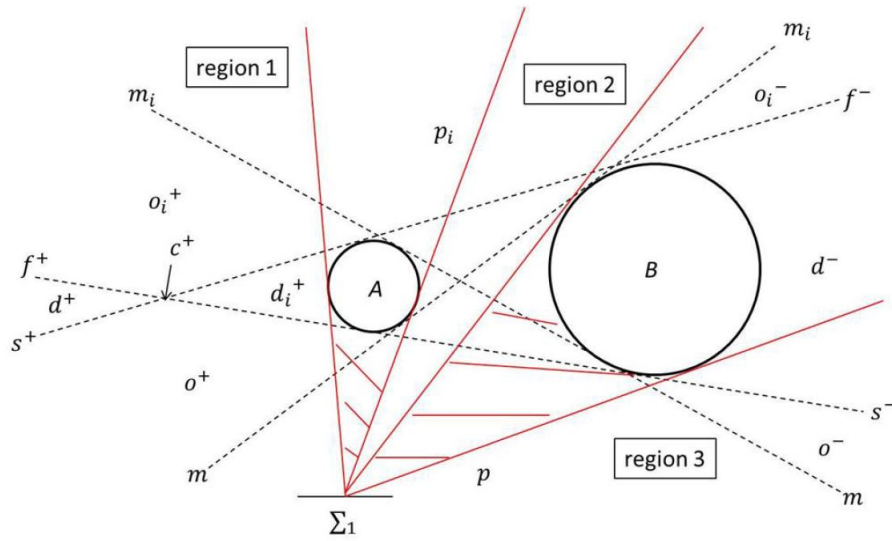


Figure 17 – Basic Relations of Interval Occlusion Calculus

(P. E. SANTOS; LIGOZAT; SAFI-SAMGHABAD, 2015).

set of relations $\{a' \{p_i; p\} b'\}$. As another example, if Σ_2 is in the red-dashed region between A and Σ_1 , then Σ_1 observes that $\{\sigma_2 \{s^+, f^+, d^+\} a\}$, and the set of relations from Σ_2 would be $\{a' \{p, m, o^+\} b'\}$ (P. E. SANTOS; LIGOZAT; SAFI-SAMGHABAD, 2015).

Following the same idea of IOC, the Spatial Collaborative Reasoning first introduced in this work aims to identify the objects positions and its relations from multiple points of view, but the formalism is not based on occlusion, taking into account that the objects rarely occlude each other from an aerial point of view.

2.3 SIMULTANEOUS LOCALIZATION AND MAPPING

When a mobile robot has no access to the map of the environment, nor its own location, the concept of Simultaneous Localization and Mapping (SLAM) is used so that the agent can estimate a map, or the environment spatial model, and simultaneously locate itself in relation to this map (THRUN; BURGARD; FOX, 2005). Mobile robots must be able to perceive the environment through sensors.

The oldest and most influent SLAM algorithm is based on the Extended Kalman Filter. Among its features, we have the reference points map. The algorithm processes only positive observations of reference points, being unable to detect the absence of any of them in the sensor measurement. As less ambiguous are the reference points, better the algorithm works. In addition, the Extended Kalman Filter assumes noise in the robot's perception and movement, but the

amount of uncertainty caused by such noise must be relatively small so that the error considered in the creation of the map is tolerable (THRUN; BURGARD; FOX, 2005).

One of the techniques used to deal with noise, as false markings identified by the agent, is to create a provisional reference list. Each new reference point found is added to a list and not straight to the map, decreasing the influence in the inference in the position of the robot. Once the reference point has been consistently observed, it is transferred to the map created by the agent (THRUN; BURGARD; FOX, 2005). This process reduces the number of markings added to the map, making them more likely, and facilitating the removal of inconsistent landmarks. The greater the number of markings, greater the chance of erroneous reference points to be present, however, few markings turns difficult the automatic location of the robot. So, it is necessary to define the ideal amount of markings for the agent to achieve a certain objective (THRUN; BURGARD; FOX, 2005).

A wide range of algorithms were derived from the concepts mentioned and applied to the SLAM problem. One of the alternatives to algorithms based on the Kalman Filter is the Expectation Maximization algorithm, capable of generating consistent maps of the environment, executing the process cyclically. Expectation Maximization algorithms do not generate maps incrementally, but search for the most likely map in the set of all generated maps at each cycle (THRUN, 2002). The Incremental Maximum Likelihood Method is an hybrid model for the incremental creation of a single map as the sensor data is observed, but without the uncertainty caused by noises (THRUN, 2002).

Considering that the robot's position is previously known, the Occupancy Grid Mapping algorithm aims at generating a consistent metric map, starting from noisy or incomplete sensor data, being used mainly in agents with long-range sensors, such as sonar or laser, both characterized by the presence of noises (THRUN, 2002). The Object Mapping algorithm addresses the problem of constructing maps composed by basic geometric shapes or objects, such as lines or walls, to represent 2D maps using lines instead of grids. The advantage of this method is a compact map, with greater accuracy, representations able to describe dynamic environments and closer to human perception of environment (THRUN, 2002).

For dynamic environment mapping, the Dynamic Occupancy Grid Mapping Algorithm (DOGMA) learns dynamic object models, represented by occupation grid maps of objects in specific locations, for example, the shape of a chair, which can be used to model a new object. As long as the environment is being observed, the algorithm assumes that the object may have been moved, resulting in a different global occupancy map (THRUN, 2002).

SLAM concepts were studied for the development of a system to allow environment mapping, considering qualitative spatial characteristics, observed from different points of view, and using the software that best adapts to these characteristics.

3 LITERATURE REVIEW

This chapter describes the necessary concepts for the development of this project and the works where such concepts were applied. Much of the investigations involve the theory about Qualitative Spatial Reasoning (QSR), Multi-agent Systems, Robotic Localization and Image Interpretation.

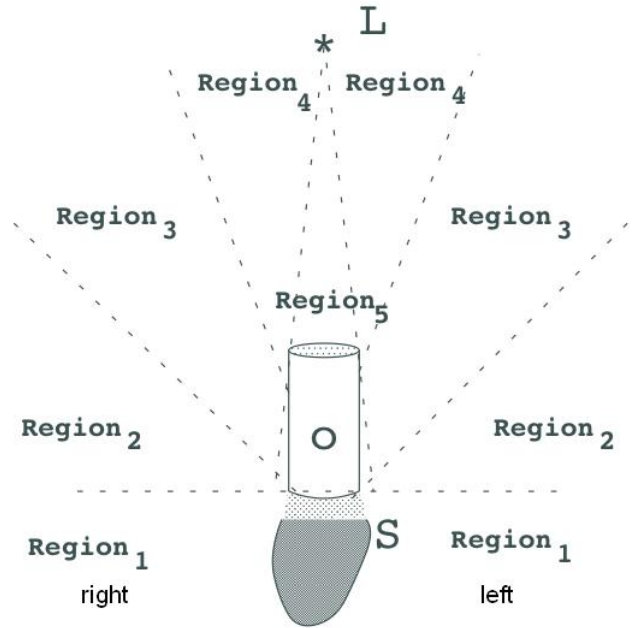
3.1 QUALITATIVE SPATIAL REASONING

The development of this project is based on the theories of Qualitative Spatial Reasoning (QSR), area of Artificial Intelligence that purposes to formalize spatial knowledge using elementary entities, such as spatial regions, directions, line segments, among others, without the use of traditional quantitative techniques (VAN HARMELEN; LIFSCHITZ; PORTER, 2008), aiming at representing the spatial perception of an agent with high-level representation.

It is common the usage of QSR concepts in conjunction with other methods of information interpretation, enabling the interaction between quantitative and qualitative representations in an agent's learning (A. G. COHN et al., 2006). An example of this application is the work presented by (P. E. SANTOS et al., 2016), showing a probabilistic localization algorithm based on QSR representation formalism. This algorithm was used in a mobile robot self-localization system that performs probabilistic localization without the acquisition of metric information, defined through Qualitative Spatial Reasoning. The concepts of the Region Occlusion Calculus (KÖHLER, 2002) and Region Connection Calculus (D. RANDELL; CUI; A. COHN, 1992) were used in the creation of a map of relative positions, divided into regions, using colored and static objects, observed through a camera installed on a mobile robot.

Another work that combines qualitative data with probabilistic techniques were presented in (FENELON, 2014). The qualitative model Perceptual Qualitative Relations about Shadows (PQRS) (P. E. SANTOS; H. M. DEE; FENELON, 2009) were combined with the probabilistic technique Bayes Filter resulting in the work called Image Interpretation with Probabilistic Qualitative Spatial Reasoning. The qualitative-probabilistic approach is applied to robot localization, where a robot can locate itself through observations of objects and their shadows. The agent reasons about the relations between shadows and objects, based on the relations from the Region Occlusion Calculus (KÖHLER, 2002) and from the Region Connection Calculus (D. RANDELL; CUI; A. COHN, 1992).

Figure 18 – Qualitative map



Source: (FENELON, 2014).

In the experiment, the robot collected images in an office environment, and should locate itself with respect to a black bucket (the object) and its shadow. The agent inferred its qualitative position, according to the qualitative map showed in Figure 18, where the space is divided into five distinct regions. In each region of the map the robot perceives a different object-shadow relation (FENELON, 2014).

Figure 18 shows the representation of the qualitative map containing five distinct regions, showing the left and right sides and regions ($Region_1$ to $Region_5$). The lines of sight between the L light source, the O object and its shadow define the boundaries between the regions. In this diagram, only the top of the shadow is used to defining the regions (FENELON, 2014).

As the PQRS is based on the Region Occlusion Calculus and Region Connection Calculus, first order logics, the method cannot handle uncertainty. To deal with the noisy of the vision system, the combination of PQRS with a Bayes filter, was used to describe the robot's belief in the state s_t , given all the evidences $e_0 : t$ up to time t . The state s_t represents a region of the qualitative map and the evidences are the terms denoted by the PQRS (FENELON, 2014). The results showed that the accuracy of the robot self localization using only the PQRS is 58% while the probabilistic qualitative self localization resulted in 80% of accuracy.

The present work investigated the use of Qualitative Spatial Reasoning for a multi-agent system, enabling the exchange of information between two or more robots, taking into account aerial points of view, captured by the UAVs. From an aerial point of view, the occlusion and connection relations of regions are identified from a different perspective compared to those observed by a terrestrial agent.

The proposed multi-agent system can be applied to dynamic formalisms, as in (SOUCHANSKI; P. E. SANTOS, 2008), where a reasoning about actions and changes is shown, to facilitate autonomous inferences about the behavior of other agents (observed from a point of view) and (P. E. SANTOS; SHANAHAN, 2003); (P. E. SANTOS, 2007), where the building blocks of a system show the qualitative changes observed by a mobile robot in a qualitative spatial reasoning formalism.

3.2 MULTI-AGENT SYSTEMS

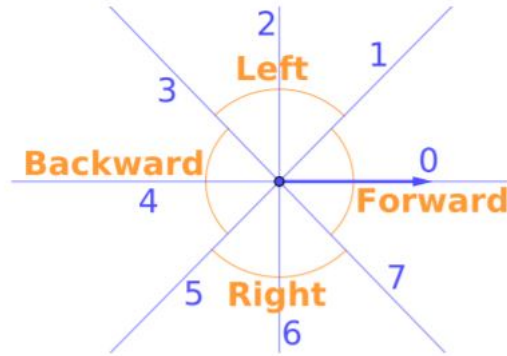
In the works mentioned in the previous section, it is assumed a continuous input of data received from a single agent, which performs the representation and inferences related to spatial perceptions. The spatial knowledge cannot be maintained by the system, if the robot or sensor suffers a fault causing an interruption in the signal transmission, or when there is more than one robot exploring the environment.

In this project, we propose the use of the Collaborative Spatial Reasoning System, seeking to develop a spatial reasoning system with multiple points of view, capable of representing the different points of view of a multi-robot system and providing communication between them, in a way that each agent can efficiently reestablish the flow of information received or transmitted to other agents, or even acquired through its own sensors.

Detailed theoretical studies regarding multi-agent systems and empirical tests were carried out in a system capable of describing the relations between visible objects in the field of view of a group of agents, with the information acquired in a point of view shared between the agents (P. E. SANTOS; LIGOZAT; SAFI-SAMGHABAD, 2015).

In the context of global location, collaborative systems primarily aim at exchanging information between two or more agents because when one robot determines the location of another robot relative to its own location, both can refine their internal beliefs based on where the other robot is, improving the accuracy of its location. In a set of different robots, the point

Figure 19 – Projective Preposition Regions



Source: (PERICO; P. E. SANTOS; BIANCHI, 2021).

of view of each robot can reduce the uncertainties in their respective estimated locations (FOX et al., 2000).

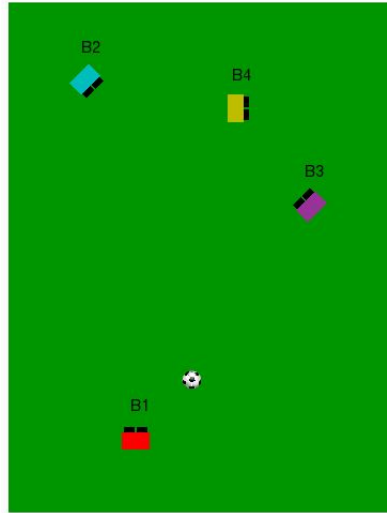
The proposal of the work (PERICO; P. E. SANTOS; BIANCHI, 2021), consists in the guided navigation of a single agent that can be either a human, as a robot (terrestrial or aerial), which does not perceive the world. The guided navigation uses linguistic spatial instructions based on projective relations, which are obtained by combining the perspective of multiple non-human agents. An agent should be guided by a set of other non-human autonomous agents with computer vision, to a goal. The guided agent receive more abstract and high-level instructions that may be understandable by any kind of agent. The navigation is carried out through instructions of relative directions based on four projective linguistic spatial prepositions: right, left, front and back.

In the movement model proposed, the agent can act in five different ways: “stop”, “move forward”, “turn right and move forward”, “turn left and move forward”, “turn back and move forward”. Figure 19 shows the projective prepositions regions.

Regardless of the number of observer agents, instructions related to movements are transmitted to the agent guided by only one observer agent, who will be defined as coordinator. The coordinator makes the necessary inferences, given the perspectives and guidance of all other observer agents involved, except for the stop instruction, that can be given by any observer agent (PERICO; P. E. SANTOS; BIANCHI, 2021).

Two hybrid models, qualitative-probabilistics, were proposed. The first model presented, $QPF\mathcal{GN}_m$, uses qualitative data to feed a particle filter that is responsible for tracking the guided

Figure 20 – Simulated Experiment



Source: (PERICO; P. E. SANTOS; BIANCHI, 2021).

agent along the entire route until the objective is reached. The prediction of the translation movement and the update are performed based on the qualitative regions generated by joining the angular sectors defined in $StarVars_m$ (PERICO; P. E. SANTOS; BIANCHI, 2021).

The second model, $NPFQI_m$, uses direction and orientation numerical data to feed a conventional particle filter and, through these particles, track and infer numerical instructions that are transformed into instructions based on projective prepositions for the guided agent, in order to lead him to the goal. The prediction of the translation movement is done with random data and the update of the particle weights is done based on a Gaussian distribution with defined deviation (PERICO; P. E. SANTOS; BIANCHI, 2021).

The experiments to verify and evaluate the behavior of the proposed models were carried out with humanoid robots of different colors, that follow the rules and restrictions imposed by the domain found in RoboCup *Soccer Humanoid League*, with five space entities: three humanoid robots observers, a guided humanoid robot and a goal, represented by a ball (PERICO; P. E. SANTOS; BIANCHI, 2021).

Simulated experiments demonstrated that the $NPFQI_m$ model is more efficient than the $QPFQI_m$ model because it had the smallest path sizes and managed to get the robot closer to the objective entity, even considering that the sizes of the objective regions were the same for all the models (PERICO; P. E. SANTOS; BIANCHI, 2021).

Figure 21 – Real Data Experiment



Source: (PERICO; P. E. SANTOS; BIANCHI, 2021).

For real situation experiments, the both models tested proved to be effective, however, several problems were found in the application of the models with real robots, as delays in the network, detection of the agents and the objective entity through computer vision using only color segmentation (PERICO; P. E. SANTOS; BIANCHI, 2021).

Applying the multi agent system concepts to Unmanned Aerial Vehicles (UAVs) and operators or pilots in highly complex missions, a conceptual framework and an architecture to deal with the specification, generation and execution of collaborative missions was developed (DOHERTY; HEINTZ; KVARNSTRÖM, 2013). In order to clarify, verify and validate the different types of interaction between operators or pilots and UAVs, a triad was created with fundamental and independent concepts: delegation, mixed initiative interaction and adjustable autonomy. In this way, the interaction and negotiation between pilots or operators and UAVs can take advantage of the skills, capabilities and knowledge of each component involved in the development of a mission (DOHERTY et al., 2010).

Still in the field of multi-agent systems, the main objective of the logic probabilistic structure developed in the work of (POTYKA et al., 2016) was to obtain the preference of a set of agents, based on the value of their beliefs and utility of such preferences, instead of using the individual preference of each agent in a decision-making situation. For this purpose, an individual knowledge base was developed for each agent, containing their individual beliefs about the available alternatives, represented by conditional probabilistics that express their subjective

beliefs, in addition to the criteria considered important for the decision. In this way, it was possible to obtain an interval of expected utilities regarding the agents' group of beliefs, which gave rise to three group preference relations: an optimistic, a pessimistic and a cautious one.

In other works, initially the individual preference of each agent is calculated, then voting rules are applied or weights are given to the alternatives and finally the group decision process is applied. In the work of (POTYKA et al., 2016), the first step is to gather the individual beliefs of each agent, forming groups of preferences, and then the group decision processes start to be executed, with the proper weights or voting system. This decision-making method can be studied for application in other axioms and properties in different systems, enabling a comparison between the performance, accuracy and robustness of information fusion methods.

3.3 ROBOTIC LOCALIZATION

The use of robotic technology can currently be observed not only in industries, but also in domestic environments, offices and public places. In these environments, there is an interaction with people, requiring robotic agents to move safely to fulfill the tasks to be performed (KRUSE et al., 2013). Seeking to meet this necessities, robotic localization consists of estimating the coordinates of an agent relative to the external environment, providing the agent with autonomous navigation skills as locating, planning, building or interpreting maps and motion control (SIEGWART; NOURBAKHSH; SCARAMUZZA, 2011; NEHMZOW, 2012) .

In this work, qualitative localization is addressed, following the ideas originated in (LEVITT; LAWTON, 1990), where a topological map was constructed, having the edges of the regions defined by a set of lines connecting pairs of landmarks; or even in the spatial representation called panorama, which restricts the location of a point relative to the visual ordering of objects observed around the robot (SCHLIEDER, 1996). The notion of qualitative navigation is defined in (SCHLIEDER, 1996), by changes in the ordering of information, which occurs when the observer crosses each one of the virtual lines defined by the reference objects, an idea applied to a domain of the *RoboCup* reported in the work of (WAGNER; VISSER; HERZOG, 2004).

Publications that address qualitative mapping of environments are also the basis for this research, such as (P. E. SANTOS; H. DEE; FENELON, 2008), that presented the theory of Perceivable Qualitative Relations on Shadows (PQRS) for the construction of qualitative maps, where the observed space was divided into five regions , based on a logical formalization of

occlusion and shadow observation. (FOGLIARONI et al., 2009) proposed a model of object visibility and occlusion to generate the regions in navigable space, while (MCCLELLAND; CAMPBELL; ESTLIN, 2013) introduced a method for the construction of maps using the relative positions of reference points. These methods are generally based on human concepts of space observation, and therefore, are incapable of dealing with sensor noise, limiting their applicability to controlled and deterministic domains. A comprehensive overview of qualitative representations for robotic localization is presented by (WOLTER; FREKSA; LATECKI, 2007).

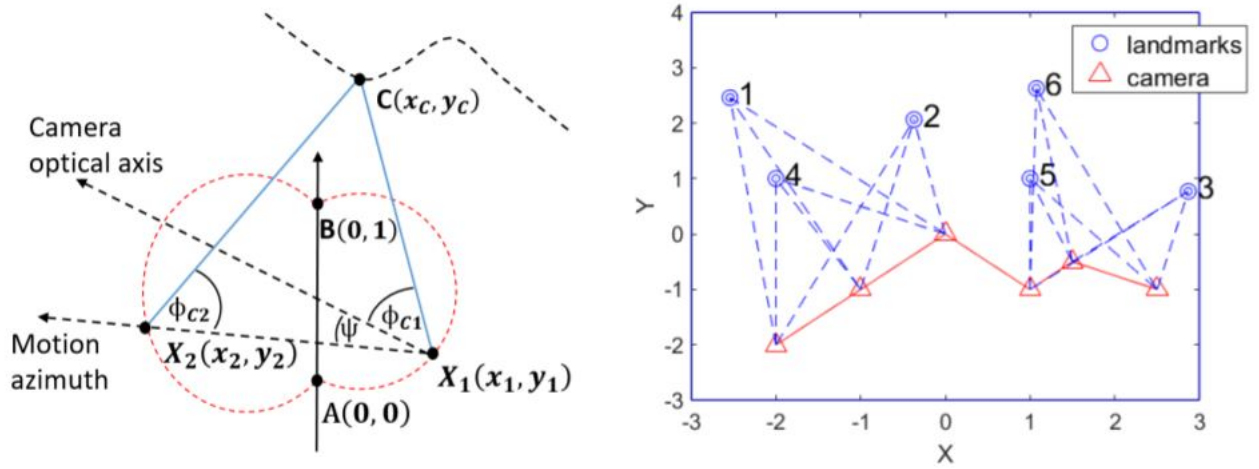
Initially, robotic localization methods based on probabilistic algorithms, defined as Bayesian filters, were presented in (THRUN, 2002; THRUN; BURGARD; FOX, 2005). This theme was based on research on Markov localization (FOX; BURGARD; THRUN, 1999) and localization procedures using Kalman filters (S. Y. CHEN, 2011) with high acceptance in the field of robotics (DELLAERT et al., 1999). Methods known as Monte Carlo algorithms (DELLAERT et al., 1999) extend the basic Markov localization by applying a particle filter to represent the distribution of possible states. Given the robot's movement, the particles are updated according to the prediction of the next state and are resized according to the agent's perception.

(MOR; INDELMAN, 2020) developed a probabilistic QSR for concurrent qualitative localization and mapping framework designed to be used for large scale navigation with simple sensors and with low complexity, using a small number of high-quality landmarks, unlike metric SLAM which is noise sensitive and typically exploits many tracked features to average noise probabilistic qualitative approach for localization and mapping as well as a probabilistic composition for propagating information between different landmark triplets. The goal is to qualitatively describe the environment and the agent's camera trajectory.

This work formulated the probabilistic inference of camera and landmark triplet qualitative states with multiple views, incorporating a motion model. Given a motion model for each camera transition, it is possible to infer the posterior probabilities of a landmark qualitative state, and camera qualitative trajectory, both in a frame. The system is able to work with a small number of measurements, and no prior knowledge. Under these conditions, a sampling based approach to solve the global non-linear small SLAM problem was chosen. This approach avoids linearization, and gets a global robust solution. Also, integration over coarse resolution qualitative states compensates some sampling errors (MOR; INDELMAN, 2020).

Considering that two landmark triplets share two common landmarks, information can be propagated from one to the other. This "composition" operation enables to enhance estimation

Figure 22 – Example of camera resection and landmark triangulation.



Source: (MOR; INDELMAN, 2020).

for existing overlapping triplets, and to infer triplets that were not viewed together at all (MOR; INDELMAN, 2020).

The tests demonstrated that incorporating a motion model in qualitative estimation improves results. This method can easily be adapted to various types qualitative space partitions, and different underlying SLAM based methods. Qualitative inference has inherent ability to accommodate small errors and generate a low compute approximated algorithm, being a practical alternative for low cost robotic systems or for active planning, in cases where exact metric location is not important (MOR; INDELMAN, 2020).

3.4 IMAGE INTERPRETATION

The concept of logical approach to image interpretation was initially proposed by (REITER; MACKWORTH, 1989) through a constraint satisfaction procedure, composed of a set of three axioms. This system was adapted in (MATSUYAMA; HWANG, 1990) to a reasoning procedure based on abduction of hypotheses (POOLEV; GOEBEL; ALELIUNAS, 1987) to generate descriptions of aerial images.

Subsequently, a logical formalism was proposed to rigorously define the transition of information from the sensors of a mobile robot to symbols representing the existence, location and shape of the observed objects (SHANAHAN, 1996). In this proposal, scenes were interpreted from a global point of view, in which objects were always visible. In contrast, the works of

(P. E. SANTOS; SHANAHAN, 2002; 2003; P. E. SANTOS, 2007) consider changes in a dynamic environment, observed from an egocentric point of view.

Adding the variable *time*, the work (HAZARIKA; A. G. COHN, 2002) describes a system for the assimilation of scenes using the concept of spatiotemporal histories, with the creation of spatiotemporal regions to represent the temporal development of relations topological data between physical objects, model based on the proposal of (FERNYHOUGH; A. G. COHN; HOGG, 2000) for the automatic construction of event models from visual information. In (BENNETT; A. COHN; MAGEE, 2005), a voting criterion elects the most consistent story to explain a given sequence of images out of multiple possible stories.

Considering the logic of processes of (BONNER; KIFER, 1994), whose semantics mirror the process of state transitions occurred in the observations of objects in dynamic scenes, the work proposed in (M. SANTOS et al., 2009) was based on applying it in a formalism of actions, as well as the work of (SECOLO; P. E. SANTOS, 2011), with a system capable of interpreting the depth profile data generated by a simulator of vehicular traffic situations, and suggesting actions to an autonomous vehicle so that it can interact with safely with other vehicles in the same environment.

The theory of understanding novel concepts by an agent developed by (FALOMIR; PLAZA, 2020), uses the Concept Blending Theory (CBT) and the CoInvent project computational model of blending, arguing that understanding a novelty requires not only disintegration and decompression, given the blend the task to find the input spaces and an adequate generic space, but also a creative process by the receiver agent that creates a blend in the process, using its own cognitive resources.

The agent understanding process involves re-creation rather than reconstruction, not disintegrating a mental space which does not exist in the receiving agent, if the concept is a novelty. The mental space of the blend exists in the computational model only after re-constructing the blending network that yields the blend, when the receiver agent understands that blend (FALOMIR; PLAZA, 2020).

The top level of communication of internal creative process based on conceptual blending is necessary to deal with unknown or novel concepts in a blend network created by an agent, but new or unknown to another agent. This communication level requires the transmission of signs or descriptions of concepts to enable a creative agent to generate a novelty by means of a blend network where a new concept is constructed, and then the communication level is used to transmit a sign about this novelty to a receiver agent (FALOMIR; PLAZA, 2020).

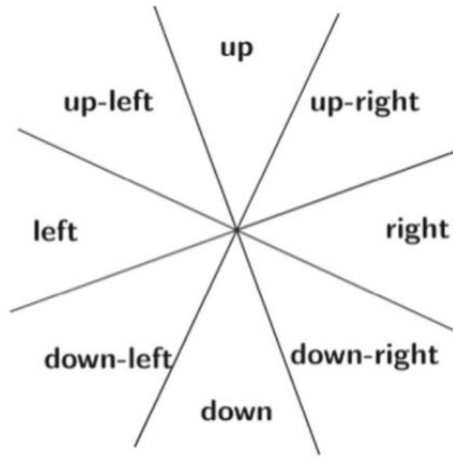
The experiments used the visual Exit sign of Figure 23 to evaluate the bending theory.

Figure 23 – Icon Example: an arrow shape and a C-like shape.



Source: (MOR; INDELMAN, 2020).

Figure 24 – Location Reference System.



Source: (MOR; INDELMAN, 2020).

The icon was deconstructed into two elements: an arrow shape and a C-like shape, to be qualitatively described. The arrow shape was qualitative described as:

$$tail(a,x) \wedge head(a,y) \wedge body(a,x,y) \Rightarrow arrow(a,x,y).$$

The C-like shape was qualitative described as:

$$has - border(a,b_a) \wedge has - point(b_a,p_a) \wedge has - border(c,b_c) \wedge has - point(b_c,p_c) \wedge point - in - border(p_a,b_c) \vee point - in - border(p_c,b_a) \Rightarrow touching(a,c)$$

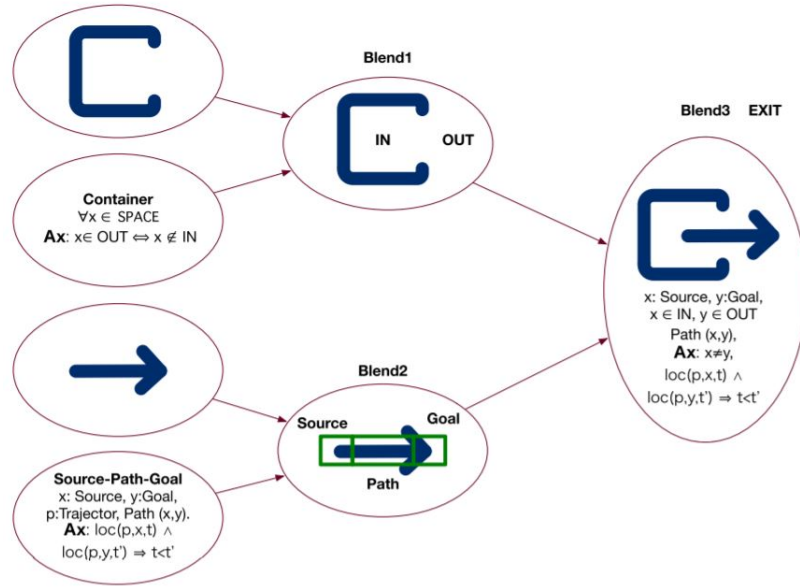
$$\neg touching(c,a) \Rightarrow disjoint(a,c)$$

Figure 24 shows the Location Referent System to obtain the location an object (arrow shape, for example) with respect to another object (C-like shape).

The complete blend re-creation process is shown in Figure 25.

This work shows that blend generation and understanding requires the same components and processes. Are they the whole gamut of mental space creation for the two input spaces,

Figure 25 – Process of understanding the visual representation of the icon *Exit*.



Source: (MOR; INDELMAN, 2020).

generic space determination, identifying the commonalities in both input spaces and blend generation of the two input spaces. To understand a novel or unknown concept, a creative artefact, is also a creative process in the visual language scenario of icons, pictograms and signage where meaning can be grounded whenever pre-existing mental spaces can be used to reconstruct such blending network.

Considering that for the evaluation of image sequences (NAGEL, 2000) some definitions of probabilistic reasoning are used, without however using logical formalisms, the propose of using strict definitions of semantics and inference methods in a formalism of Collaborative Spatial Reasoning to extract spatial information from the images captured by UAVs that enable the inference of information for environment mapping.

The works referenced in this chapter form the context for the development of this research project, where the main additional contribution compared to the reference works is the combination of different qualitative calculus instead of the combination of a qualitative calculus with probabilistic techniques.

4 METHOD

This chapter describes the original contributions of this work.

4.1 PROBLEM DEFINITION

The task of integrating multi-agent systems where robots and humans interact, exchanging information which can be fast understood by all members of the team is a challenge.

The usage of qualitative formalism enables this exchange of information, specially in an environment where the transmission of signal can be jeopardized for many reasons. Using purely quantitative data can impair the assimilation and fast response of humans in a multi-agent team. On the other hand, qualitative information contain uncertainty.

The theories explained in this chapter relates different qualitative theories, in order to combine information into a single sentence, so that it is possible to exchange complete information about objects perceived in an environment, and make it possible to humans and robots to share knowledge from partial and different points of view.

4.2 LH INTERVAL CALCULUS

Assuming that a system is composed by two or more agents and that one or more of those agents are humans, the mission to describe the position of objects present at a previously unknown environment might be impaired without the use of sentences that take into account the way people are used to describe spatial information. To deal with this issue, LH Interval Calculus is a formalism extended from Interval Occlusion Calculus (P. E. SANTOS; LIGOZAT; SAFI-SAMGHABAD, 2015). To qualitatively describe the position of two or more perceived objects in a scene and how they are related to each other, from various distinct global aerial points of view. The qualitative description is based on Region Connection Calculus formalism (D. RANDELL; CUI; A. COHN, 1992) and Allen's Interval Algebra (ALLEN, 1990).

In this work, the aerial observers are autonomous unmanned aerial vehicles (UAVs) equipped with a camera, a GPS and a compass. The viewpoint of each UAV is represented as pairs $\Sigma_i = (g_i, v_i)$, $i \in \mathbb{N}$, where g_i is the global position of the observer indicated by the GPS, and v_i is a unit vector representing the observer's direction of flight indicated by the compass.

The observed scenario is read by the cameras installed in the UAVs from right to left and from top to bottom, and the origin is located in the upper left corner of the scene. The function

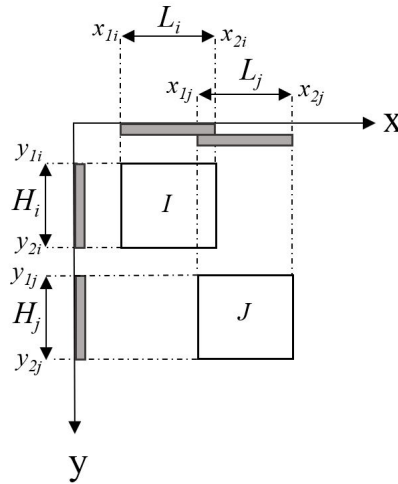


Figure 26 – Grey intervals representing object's projections from an aerial point of view.

$j = image(J, \Sigma)$ maps the objects perceived in the world, where j is the image of an object J , seeing from an aerial point of view Σ .

Every observer is able to describe the relations between pairs of objects in its field of view, and exchange information with each other by means of a message-passing procedure. From a global aerial viewpoint, the perceived spatial objects can be considered as regions, enabling the system to extract information about object's positions applying RCC-8 (D. RANDELL; CUI; A. COHN, 1992) constraints. However, it is possible to obtain more complete information about object's disposition in the environment, taking into account that the system count on more than one observer. Introducing orientation on a linear axis of reference, the Allen's relations can be applied to agents' viewpoints.

The reference system is positioned in a 2D space, its origin is located in the upper left corner of the scene, the horizontal line of reference is oriented from right to left (axis x), while the vertical line of reference is oriented from top to bottom (axis y), in conformity with the process of reading the scene by the cameras installed in the UAVs. The projections of the shape of each object in the line axis of reference is considered a layered interval. The layered interval L is the projection of the shape of the object in the axis of reference x , being $L = (x_1, x_2)$; $x_1 < x_2$, x_1 is the lower limit of the object's projection on axis x and x_2 is the upper limit of the object's projection on axis x . The layered interval H is the projection of the shape of the object in the axis of reference y , being $H = (y_1, y_2)$; $y_1 < y_2$, y_1 is the lower limit of the object's projection on axis y and y_2 is the upper limit of the object's projection on axis y . Figure 26 shows an example of objects projections.

Inspired by the work defined in the Block Algebra (BALBIANI; CONDOTTA; CERRO, 2002), the projections of a pair of objects into the axis x hold one of the 13 basic Allen's relations, as well as the projections of the objects into axis y . The results of Allen's relations constraints in a 2D scene for a pair of objects I and J are represented by the sentence $I(r_x, r_y)J$, where r_x is the Allen relation for the projections of the objects I and J in the axis of reference x and r_y is the Allen relation for the projections of the objects I and J in the axis of reference y .

In this work, the objects are not necessarily blocks, they can be perceived in different formats, once the relations are given by the projections of the existing interval between the extreme points of the objects.

Given two distinct objects I and J and their intervals $L(I)$, $L(J)$, $H(I)$ and $H(J)$, the notation $I(r_x)J$ applies iff $L(I) < L(J)$ and $I(r_y)J$ applies iff $H(I) < H(J)$. The inverse situation follows the notation $I(-r_x)J$ applies iff $L(I) > L(J)$ and $I(-r_y)J$ applies iff $H(I) > H(J)$.

According to the distance between the observer and the objects, the image projected in a point of view Σ can be larger or smaller, but their projections on the x and y axis will result in the same pair of relations (r_x, r_y) .

To illustrate how RCC-8 and Allen's interval algebra information interact (LIGOZAT, 2013), let's consider the example of Figure 26, where a *DC* (*disconnected*) RCC-8 relation is identified between objects I and J . For disconnected objects, a *precedes* or *is preceded* Allen's relation will be detected in the projection of those objects in the axis x or y , or in both axis. Figure 26 shows the Allen's relation "*overlaps*" in axis x and the Allen's relation "*precedes*" in axis y . Therefore, the information obtained is $I(o,p)J$.

The same reasoning is extended to the *EC* (*externally connected*) RCC-8 relation, that leads to the *meets* or *is met* Allen's relations, detected in the projection of the objects in the axis x or y . In case two objects are externally connected by one of its corners, the *meets* or *is met* relations can be detected in the projections of both axis.

For *PO* (*partially overlapped*) objects, it is expected the Allen's relations *overlaps* or *is overlapped* projection of the objects in the axis x or y , or in both axis.

The *TPP* (*tangential proper part*) objects needs special attention as they stand for *starts* and *finishes* objects projections on axis x or y , while *TPPI* (*tangential proper part inverse*) objects stand for *is started* and *is finished* Allen's relations.

When the RCC-8 relation *NTPP* (*non tangential proper part*) is identified between two objects, the Allen's relation *during* is expected to be projected on axis x or y , the relation

NTPPI (*non tangential proper part inverse*) leads to the *contains* Allen's relations, and *EQ* (*equal*) RCC-8 relation will result in the *equal* Allen's relation projected on axis x or y .

To formally define the relations of the LH Interval Calculus in terms of RCC-8 and Allen's relations, note that $L = (x_1, x_2); x_1 < x_2$, and $H = (y_1, y_2); y_1 < y_2$. The functions $ext(L)$ and $ext(H)$ are assumed to map the layered intervals L and H to its *extent*, delimited by its upper and lower projected limits: $ext(L) = (l_1, l_2); l_1 < l_2$ and $ext(H) = (h_1, h_2); h_1 < h_2$.

RCC-8 relations are described in capital letters R and Allen's relations are represented by r . The definitions below apply to the axis of reference x . The construction of the sentence $I(r_x, r_y)J$ results from the application of the constraints below to the object's projections $ext(L_i)$ and $ext(L_j)$ on axis x to obtain the r_x relation.

- a) $I p_x J : \Sigma$, read as "I precedes J from Σ " and defined by $(I) DC(J) \cap ext(L_i) p_x ext(L_j)$;
- b) $I -p_x J : \Sigma$, read as "I is preceded by J from Σ " and defined by $(I) DC(J) \cap ext(L_j) p_x ext(L_i)$;
- c) $I m_x J : \Sigma$, read as "I meets J from Σ " and defined by $(I) EC(J) \cap ext(L_i) m_x ext(L_j)$;
- d) $I -m_x J : \Sigma$, read as "I is met by J from Σ " and defined by $(I) EC(J) \cap ext(L_j) m_x ext(L_i)$;
- e) $I o_x J : \Sigma$, read as "I overlaps J from Σ " and defined by $(I) O(J) \cap ext(L_i) o_x ext(L_j)$;
- f) $I -o_x J : \Sigma$, read as "I is overlapped by J from Σ " and defined by $(I) O(J) \cap ext(L_j) o_x ext(L_i)$;
- g) $I s_x J : \Sigma$, read as "I starts J from Σ " and defined by $(I) TPP(J) \cap ext(L_i) s_x ext(L_j)$;
- h) $I -s_x J : \Sigma$, read as "I is started by J from Σ " and defined by $(I) TPPI(J) \cap ext(L_j) s_x ext(L_i)$;
- i) $I f_x J : \Sigma$, read as "I finishes J from Σ " and defined by $(I) TPP(J) \cap ext(L_i) f_x ext(L_j)$;
- j) $I -f_x J : \Sigma$, read as "I is finished by J from Σ " and defined by $(I) TPPI(J) \cap ext(L_j) f_x ext(L_i)$;
- k) $I d_x J : \Sigma$, read as "I is during J from Σ " and defined by $(I) NTPP(J) \cap ext(L_i) d_x ext(L_j)$;
- l) $I -d_x J : \Sigma$, read as "I contains J from Σ " and defined by $(I) NTPPI(J) \cap ext(L_j) d_x ext(L_i)$;
- m) $I eq_x J : \Sigma$, read as "I is equal to J from Σ " and defined by $(I) EQ(J) \cap ext(L_i) eq_x ext(L_j)$.

The application of the constraints below to object's projections $ext(H_i)$ and $ext(H_j)$ on axis y results in the r_y relation of the sentence $I(r_x, r_y) J$.

- a) $I p_y J : \Sigma$, read as “I precedes J from Σ ” and defined by $(I) DC(J) \cap ext(H_i) p_y ext(H_j)$;
- b) $I -p_y J : \Sigma$, read as “I is preceded by J from Σ ” and defined by $(I) DC(J) \cap ext(H_j) p_y ext(H_i)$;
- c) $I m_y J : \Sigma$, read as “I meets J from Σ ” and defined by $(I) EC(J) \cap ext(H_i) m_y ext(H_j)$;
- d) $I -m_y J : \Sigma$, read as “I is met by J from Σ ” and defined by $(I) EC(J) \cap ext(H_j) m_y ext(H_i)$;
- e) $I o_y J : \Sigma$, read as “I overlaps J from Σ ” and defined by $(I) O(J) \cap ext(H_i) o_y ext(H_j)$;
- f) $I -o_y J : \Sigma$, read as “I is overlapped by J from Σ ” and defined by $(I) O(J) \cap ext(H_j) o_y ext(H_i)$;
- g) $I s_y J : \Sigma$, read as “I starts J from Σ ” and defined by $(I) TPP(J) \cap ext(H_i) s_y ext(H_j)$;
- h) $I -s_y J : \Sigma$, read as “I is started by J from Σ ” and defined by $(I) TPPI(J) \cap ext(H_j) s_y ext(H_i)$;
- i) $I f_y J : \Sigma$, read as “I finishes J from Σ ” and defined by $(I) TPP(J) \cap ext(H_i) f_y ext(H_j)$;
- j) $I -f_y J : \Sigma$, read as “I is finished by J from Σ ” and defined by $(I) TPPI(J) \cap ext(H_j) f_y ext(H_i)$;
- k) $I d_y J : \Sigma$, read as “I is during J from Σ ” and defined by $(I) NTPP(J) \cap ext(H_i) d_y ext(H_j)$;
- l) $I -d_y J : \Sigma$, read as “I contains J from Σ ” and defined by $(I) NTPPI(J) \cap ext(H_j) d_y ext(H_i)$;
- m) $I eq_y J : \Sigma$, read as “I is equal to J from Σ ” and defined by $(I) EQ(J) \cap ext(H_i) eq_y ext(H_j)$.

The parameters formalized in the LH Interval Calculus express the qualitative localization of two or more objects from an aerial point of view in terms of regions, through the RCC-8 relations and in terms of object detection sequence through the Allen's Algebra.

As already mentioned, each RCC-8 relation has an expected Allen's relation projection that can be identified on axis x , y or both. In case the expected relation is identified in only one axis, any other relation can be identified on the other axis, resulting in a set of possible pairs

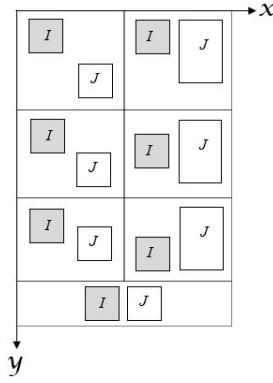


Figure 27 – Set of Allen's relations for Disconnected objects.

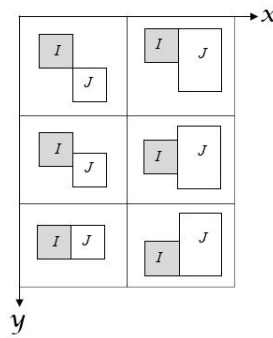


Figure 28 – Set of Allen's relations for Externally Connected objects.

of Allen's relations. Figure 27 shows the set of possible Allen's relations combinations for two objects identified as *Disconnected*. The projections on x axis represent the relation *precedes*, while the projections on y axis represent the basic Allen's relations $[p, m, o, s, d, f, eq]$. The inverse positions of objects are also valid, for example, the top left square of Figure 27 shows the (p, p) configuration, but it is possible to have also $(p, -p)$, $(-p, p)$ and $(-p, -p)$ configurations in the set of *Disconnected* objects. The same rule is valid for all configurations of the figure.

In case the relation between two objects is defined as *Externally Connected*, Figure 28 shows the projections of *meets* configuration on axis x and the set of possible Allen's relations projections on y axis, excluding the relation *precedes*, already showed in Figure 27 to be possible only for *Disconnected* objects. Again, the inverse and transverse configurations are possible.

The set of possible Allen's relations for *Partially Overlapped* objects is shown in Figure 29, for *Tangential Proper Part* in Figure 30, the inverse of *Tangential Proper Part* in Figure 31 and for *Non Tangential Proper Part* and its inverse in Figure 32. The *Equal* relation identified in RCC-8 constraints will result in the pair (eq, eq) in the LH Interval Calculus, or, the image will be considered a single object.

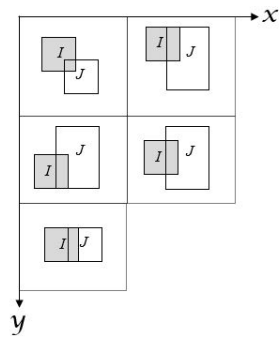


Figure 29 – Set of Allen's relations for Partially Overlapped objects.

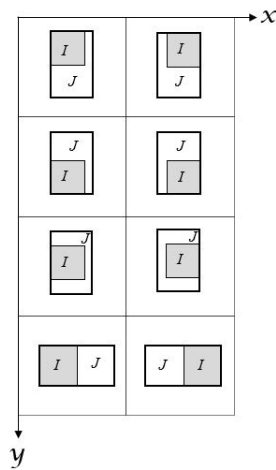


Figure 30 – Set of Allen's relations for Tangential Proper Part objects.

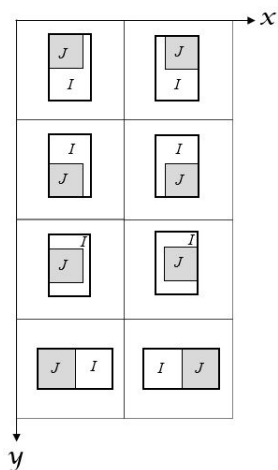


Figure 31 – Set of Allen's relations for the inverse of Tangential Proper Part objects.

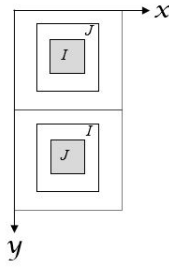


Figure 32 – Set of Allen's relations for Non Tangential Proper Part objects and its inverse configuration.

Finally, in a scenario composed of two or more objects, the relations between them, perceived by one or more agents, will be described in the sentence $I(r_x, r_y) J: \Sigma_i$ for each pair of objects I and J .

The LH Interval Calculus identify the qualitative relations between two objects, associating two different calculus that take into account the interaction of humans with robots, so that the information can be delivered and/or read by any member of the team, considering the abstraction present in the high level information, but maintaining the effectiveness of the messages that need to be exchanged in a mission of environment recognition.

In real situations of an efficient environment mapping task, the agents fly over different portions of the region to acquire information, from different directions of flight. It is necessary the development of a formalism to deal with different viewpoints of the agents as well as a consistency check for the information exchanged by the UAVs.

The *Collaborative Spatial Reasoning for Environment Mapping*, first proposed in this work, is a formalism to reason about objects positions and their qualitative relations, combining partial fields of view, from different aerial viewpoints. This formalism is introduced in the next section.

4.3 COLLABORATIVE SPATIAL REASONING FOR ENVIRONMENT MAPPING

The mission of mapping an environment or perform coordinated missions using two or more agents can be done more efficiently when a team of agents collaborate to scan a single domain. The partitioning of a region take into account the particular capabilities of the team's participants, such as the speed of flight or how wide can the area be covered by sensors, to define which sub area of the terrain, as well as the width of the area that can be scanned in a given period of time (DOHERTY et al., 2016).

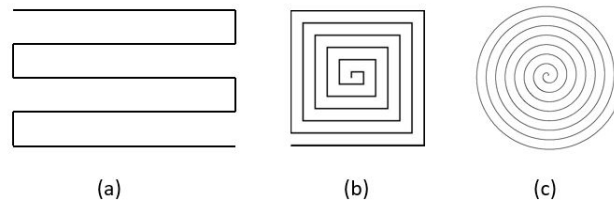


Figure 33 – Patterns of scan trajectories.

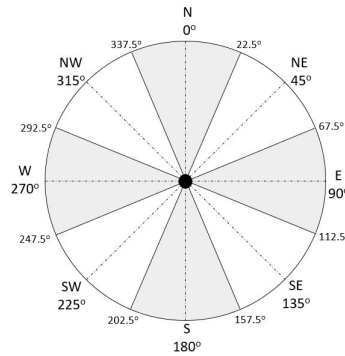


Figure 34 – Cardinal directions.

Some common patterns of scan trajectories are spirals (Figure 33), expanding squares and lawnmower, and all of them result in partial views of the region, seen from diverse directions of flight. To combine the spatial information collected by the agents, the knowledge about the cardinal direction of flight is necessary. Since in this work it is considered the participation of humans in the mission of environment mapping, the numerical information delivered by the drones is translated to a qualitative approach through the Cardinal Direction Calculus (FRANK, 1996).

The cardinal direction indicated by a compass in the case of robotic agents and by human sense of direction in case of human agents, determines the vector v to represent the flight orientation of a viewpoint Σ observing the objects in a scene.

To qualitatively classify the set of symbolic directions, the NED (North-East-Down) geographic frame used in the drones to inform latitude, longitude and height to the operator is shared in 8 equal cone shaped areas, covering a range of $\pm 22,5^\circ$ for each cardinal direction, as shown on Figure 34. The set of directions is: *North* (N) = 0° , *North East* (NE) = 45° , *East* (E) = 90° , *South East* (SE) = 135° , *South* (S) = 180° , *South West* (SW) = 225° , *West* (W) = 270° and *North West* (NW) = 315° .

From the 3 common patterns of scan trajectories, in this work it is possible to process the information delivered according to expanding squares or lawnmower methods, where the vector v changes during the scanning of the area, rotating approximately 90° , 180° or 270° for each part

of the scanned route. The spiral trajectory is not considered because the vector v changes all the time, making it impossible to set an equivalence between the information generated by different agents. The correspondence process enables the system to identify objects positions and their projections relations on the axis x and y , relative to the first direction of the vector v indicated in Σ .

Given two agents flying to different directions over the same domain, the information generated by LH Interval Calculus for two observed objects I and J by each drone is $I(r_x, r_y) J: \Sigma_i$, being $\Sigma_i = (v_i, g_i)$. The direction of flight vector v_i is added as an index in the relations r_x and r_y below just to denote a reference vector, becoming $I(v_i r_x, v_i r_y) J: \Sigma_i$.

Supposing a first agent flying to a v_i direction, and a second agent flying to a $v_i + 90^\circ$ direction, observing the same objects I and J , the correspondence between $I(v_i r_x, v_i r_y) J: \Sigma_1$ and $I(v_i + 90^\circ r_x, v_i + 90^\circ r_y) J: \Sigma_2$ is given by

$$v_i + 90^\circ r_x = v_i r_y$$

$$v_i + 90^\circ r_y = -v_i r_x$$

The notation “ $-$ ” indicates the reverse Allen’s relation. Following the same reasoning, an observer flying to a $v_i + 180^\circ$ direction, observing the same objects I and J as the first agent flying to a v_i direction, the correspondence between $I(v_i r_x, v_i r_y) J: \Sigma_1$ and $I(v_i + 180^\circ r_x, v_i + 180^\circ r_y) J: \Sigma_2$ is given by

$$v_i + 180^\circ r_x = -v_i r_x$$

$$v_i + 180^\circ r_y = -v_i r_y$$

For an observer flying to a $v_i + 270^\circ$ direction, observing the same objects I and J as the first agent flying to a v_i direction, the correspondence between $I(v_i r_x, v_i r_y) J: \Sigma_1$ and $I(v_i + 270^\circ r_x, v_i + 270^\circ r_y) J: \Sigma_2$ is given by

$$v_i + 270^\circ r_x = -v_i r_y$$

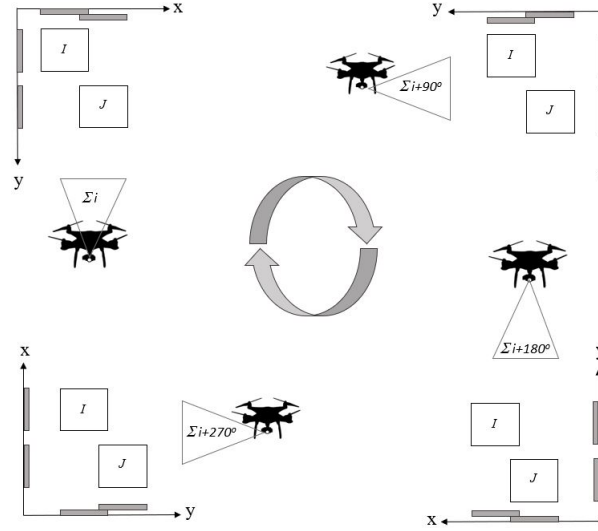


Figure 35 – Axis of reference's positions according to agents orientation of flight.

$$v_i + 270^\circ r_y = v_i r_x$$

This correspondence analysis works not only for changes in the direction of flight of a single agent, but also for consistency check, to correctly represent relations between objects from any point of view. Figure 35 shows that, whether object $I(-o_x, p_y) J: \Sigma_i$ is perceived by an agent flying to v_i direction, the same scene seen from a Σ_2 flying to $v_i + 180^\circ$ direction will be described as $I(o_x, -p_y) J: \Sigma_2$.

It is important to point out that Allen's relations *during*, *contain* and *equal* don't change when saw from different directions of flight, this way, the inverse relations of the correspondence rules don't apply to *during*, *contain* and *equal*. One object *during*, *containing* or *equal* another one will keep its relation independent of the viewpoint, the change will happen only in the axis x or y where these relations will be identified. The relation "start" is the inverse relation of "finish", "is started" is the inverse relation of "is finished", and vice versa.

Up to now it is being considered objects in the viewpoint of the agents. To combine 2 or more fields of view, the Collaborative Spatial Reasoning applies the concept of RCC-8 Composition Table developed by (D. RANDELL; CUI; A. COHN, 1992) and Allen's Composition Table developed by (ALLEN, 1990). The combination process is described in the next subsection.

4.3.1 Partial View Combination in Collaborative Spatial Reasoning

In this work, a terrain contains the objects to be identified, as buildings of different sizes, cars of different colors and an airplane. The complete domain can be scanned by two or more

agents, being each agent responsible to fly over a portion of the terrain. The size of this portion is defined according to the capabilities of each agent, as speed of flight and the resolution of the cameras installed in the drones. The portion scanned by each agent provides a partial view of the environment.

To reason about objects that are not in the agent's field of view, the composition tables express the implications obtained by the composition of each pair of basic relation, in this work, RCC-8 and Allen's Interval Algebra domains. In case 2 or more different agents detect at least one object in common, it is possible to infer information about objects that are not in their field of view, through the concept of composition.

To identify if one object was already detected or not, a labeling system is used to tag every object detected by each agent, based on the GPS position showed in the moment when the object is perceived by the agent. Object detection is part of this work, but it is not its focus. Therefore, an off-the-shelf software for object detection and tagging is used to identify the objects perceived in a scene and label them so that different agents can share objects' information and reason about them. When every object is identified and tagged, this information is shared between all members of the team, making it possible to combine the partial view of the agents and guarantee the consistency of the information that each agent delivers to the system.

The consistency check is made through a constraint network pair (N, C) , being N the set of vertices where each domain element is represented by a vertex, and C representing a set of constraints, to verify the existence of a consistent scenario by imposing algebraic closure on the network of constraints, to confirm whether the configurations provided by the information from the agents are feasible in at least one scenario of a domain (P. E. SANTOS; LIGOZAT; SAFI-SAMGHABAD, 2015). In this work it is considered two domains: one defined by the set of RCC-8 constraints and a second one defined by the set of Allen's relations constraints.

A network (N, C) is algebraically closed if its three vertices $(i, j, k) \in N^3$, and is consistent with the composition

$C(i, j) \subseteq C(i, k) \circ C(k, j)$ (LIGOZAT, 2013). This concept is the basis for the combination of the partial view of each agent to describe a single domain, where one agent scans part of the terrain and detects the objects I and K , and a second agent scans the remaining part of the terrain and detects the objects K and J , the Collaborative Spatial Reasoning combine the information about objects to get the relation of the objects I and J , regarding to the RCC-8 and Allen's relations, delivering the information under the same frame of LH Interval Calculus, i.e., $I(r_x, r_y) J: \Sigma_i$, and complying the same constraints.

Table 4 – RCC-8 Composition Table.

$R_1(a,b)$	$R_2(b,c)$							
	DC	EC	PO	TPP	TPPI	EQ	NTPP	NTPPI
DC	DC, EC, PO, TPP, TPPI, EQ, NTPP, NTPPI	DC, EC, PO, TPP, NTPP	DC, EC, PO, TPP, NTPP	DC, EC, PO, TPP, NTPP	DC	DC	DC, EC, PO, TPP, NTPP	DC
EC	DC, EC, PO, TPPI, NTPPI	DC, EC, PO, TPPI, TPPI, EQ	DC, EC, PO, TPP, NTPP	EC, PO, TPP, NTPP	DC, EC	EC	PO, TPP, NTPP	DC
PO	DC, EC, PO, TPPI, NTPPI	DC, EC, PO, TPPI, NTPPI	DC, EC, PO, TPP, TPPI, EQ, NTPP, NTPPI	PO, TPP, NTPP	DC, EC, PO, TPPI	PO	PO, TPP, NTPP	DC, EC, PO, TPPI, NTPPI
TPP	DC	DC, EC	DC, EC, PO, TPP, NTPP	TPP, NTPP	DC, EC, PO, TPP, TPPI, EQ	TPP	NTPP	DC, EC, PO, TPPI, NTPPI
TPPI	DC, EC, PO, TPPI, NTPPI	EC, PO, TPPI, NTPPI	PO, TPPI, NTPPI	PO, TPP, TPPI, EQ	TPPI, NTPPI	TPPI	PO, TPP, NTPP	NTPPI
EQ	DC	EC	PO	TPP	TPPI	EQ	NTPP	NTTPI
NTPP	DC	DC	DC, EC, PO, TPP, NTPP	NTPP	DC, EC, PO, TPP, NTPP	NTPP	NTPP	DC, EC, PO, TPP, TPPI, EQ, NTPP, NTPPI
NTPPI	DC, EC, PO, TPPI, NTPPI	PO, TPPI, NTPPI	PO, TPPI, NTPPI	PO, TPPI, NTPPI	NTPPI	NTPPI	PO, TPP, TPPI, EQ, NTPP, NTPPI	NTPPI

Source: (D. RANDELL; CUI; A. COHN, 1992).

The composition of the relations follows the RCC-8 Composition Table developed by (D. RANDELL; CUI; A. COHN, 1992) and Allen's Composition Table developed by (ALLEN, 1990). Table 4 represents the disjunction of RCC-8 relations resulting from the composition of each pair of basic RCC-8 relation. Table 36 shows the composition table of Allen's relations.

Keeping in mind that each RCC-8 relation has an expected Allen's relation projection identified on axe x , y , or both, the resulting composition of the relations between two pairs of regions I and K , and K and J , to infer the relation of a third pair I and J , will lead to an Allen's relation resulting from the composition of the projections of this third pair.

In this work, each pair of objects has a pair of Allen's relations (r_x , r_y) representing the relation of the projections of objects' intervals into axis x and y . To combine partial views, the Collaborative Spatial Reason applies the concept developed by (BALBIANI; CONDOTTA;

o	p	m	o	s	d	fi	eq	f	di	si	oi	mi	pi
p	p	p	p	p	[p, d]	p	p	[p, d]	p	p	[p, d]	[p, d]	[p, pi]
m	p	p	p	m	[o, d]	p	m	[o, d]	p	m	[o, d]	[fi, f]	[di, pi]
o	p	p	[p, o]	o	[o, d]	[p, o]	o	[o, d]	[p, di]	[o, di]	[o, oi]	[di, oi]	[di, pi]
s	p	p	[p, o]	s	d	[p, o]	s	d	[p, di]	[s, si]	[d, oi]	mi	pi
d	p	p	[p, d]	d	d	[p, d]	d	d	[p, pi]	[d, pi]	[d, pi]	pi	pi
fi	p	m	o	o	[o, d]	fi	fi	[fi, f]	di	di	[di, oi]	[di, oi]	[di, pi]
eq	p	m	o	s	d	fi	eq	f	di	si	oi	mi	pi
f	p	m	[o, d]	d	d	[fi, f]	f	f	[di, pi]	[oi, pi]	[oi, pi]	pi	pi
di	[p, di]	[o, di]	[o, di]	[o, di]	[o, oi]	di	di	[di, oi]	di	di	[di, oi]	[di, oi]	[di, pi]
si	[p, di]	[o, di]	[o, di]	[s, si]	[d, oi]	di	si	oi	di	si	oi	mi	pi
oi	[p, di]	[o, di]	[o, oi]	[d, oi]	[d, oi]	[di, oi]	oi	oi	[di, pi]	[oi, pi]	[oi, pi]	pi	pi
mi	[p, di]	[s, si]	[d, oi]	[d, oi]	[d, oi]	mi	mi	mi	pi	pi	pi	pi	pi
pi	[p, pi]	[d, pi]	[d, pi]	[d, pi]	[d, pi]	pi	pi	pi	pi	pi	pi	pi	pi

[p, o] = {p, m, o}	[p, d] = {p, m, o, s, d}	[p, di] = {p, m, o, fi, di}	[p, pi] = all relations
[d, oi] = {d, f, oi}	[d, pi] = {d, f, oi, mi, pi}	[s, si] = {s, eq, si}	[o, di] = {o, fi, di}
[o, oi] = {o, s, d, fi, eq, f, di, si, oi}	[o, d] = {o, s, d}	[di, pi] = {di, si, oi, mi, pi}	[di, oi] = {di, si, oi}
[oi, pi] = {oi, mi, pi}	[fi, f] = {fi, eq, f}		

Figure 36 – Allen's Composition Table

(ALLEN, 1990)

CERRO, 1998), defining that the composition of pairs of Allen's relations is given by

$$(A, B) \circ (C, D) = (A \circ C) \times (B \circ D),$$

being A and C a set of relations r_x and B and D a set of relations r_y . The reasoning for Allen's composition table is analogous to that used in RCC-8 composition table.

Supposing that the information captured by Σ_1 shows two objects A and B and their relation is PO . Σ_2 shows two objects B and C and their relation is DC . Σ_1 and Σ_2 are flying to the same direction v_i . The relations of the projections of the objects A and B on axis x and y is the pair (d, o) . The relations of the projections of the objects B and C is the pair $(p, -m)$. According to the composition table of Figure 4, $PO \circ DC = (DC, EC, PO, TPPI, NTPPI)$, that are the possible relations for the objects A and C . The composition of the pairs

$$(d, o) \circ (p, -m) = (p) \times (-o, -s, -d) = ((p, -o), (p, -s), (p, -d)),$$

according to the composition table of Figure 36, that shows the set of possible pairs of relations of the projections of the objects A and C . As the results show that the projections on axis x is *precedes* for the three possible pairs of relations, it is possible to deduce that the objects A and C are *Disconnected*, and its projections on axis x and y can be the relations $((p, -o), (p, -s), (p, -d))$.

5 EXPERIMENTS

The experiments described in this chapter were performed aiming at evaluating the LH Interval Calculus and the Collaborative Spatial Reasoning theories. Three aerial navigation were performed by two UAVs to collect real environment information using cameras.

5.1 DOMAIN

Two different UAV platforms were used to execute the experiments. The first one is a DJI Matrice 600 Pro¹ research platform presented in Figure 9. It has the maximum takeoff weight of 15.1 kg with the maximum payload of 5.5 kg. Its largest dimension is approximately 1.6 meters. The second platform is a DJI Matrice 100² presented in Figure10. It has the maximum takeoff weight of 3.6 kg, with maximum payload capacity of approximately 1 kg. It's tip-to-tip length is 1 meter.



Figure 37 – DJI Matrice 600 Pro experimental platform.

Both platforms are equipped with Intel NUC computers using Core i7-7567U processors, 16 GB of memory, and 500 GB SSD storage.

Three real data collection flights were performed at Motala Flygklubb, Sweden. The DJI Zenmuse Z3³ cameras were used to collect video and images during the experimental flights. The collected video was of 1080p resolution, and the images were of size 4000×2250 pixels. The UAV's positions, altitudes and cardinal direction of flight were logged as well as ROS bag files.

¹www.dji.com/matrice600/info

²www.dji.com/matrice100

³www.dji.com/zenmuse-z3

5.1.1 Environment Description

The data collection took place at Motala Flygklubb (Figure 39), Sweden, during the author's visit at Linköping University in 2018. The Flygklubb is a non-profit association for motor and gliders airplanes, founded in 1967. The club is used in the works for the development and popularization of non-commercial aviation.

It was considered 14 objects present in the Flygklubb area at the moment of the data collection flights. They are: 7 buildings of different sizes and colors, 5 cars, 1 van and 1 airplane. The background contains areas covered by grass, asphalt, trees and a road.

5.1.2 Data Collection Flight

The scanning patterns flight to collect the real experimental data were automatically generated to cover a designated area, defined by the perimeter allowed to drone flights at Motala Flygklubb. The platforms were positioned at the grass, in the right side of the picture showed in Figure 39, in a region proper for departure. The UAVs took off manually and after reaching a safe altitude flown autonomously over the region and performed the lawnmower scan according to the pre-determinate altitude. The flights were performed by each UAV at three different altitudes: 30, 50 and 80 meters above the ground level (AGL).

The videos and photos collected register a partial top view of the environment and were taken while the UAVs flew to different directions, to show partial perspectives of each scene. Each drone flew over a particular portion of the area, capturing images of some objects in common in order to use those objects as references to infer the position of other objects that were not in its field of view.



Figure 38 – DJI Matrice 100 experimental platform.



Figure 39 – Motala Flygklubb scenario

89 pictures were obtained by the UAVs flying at 30 meters, 65 pictures at 50 meters and 69 pictures at 80 meters, adding up 223 pictures. The pictures were taken at every 5 seconds from takeoff to landing.

It was also recorded 6:51 minutes of video movie of the 30 meters flight, 5:48 minutes of the 50 meters flight and 5:56 minutes of the 80 meters flight. The videos and pictures were recorded in the memory cards of the drones.

5.2 LH INTERVAL CALCULUS EVALUATION

The objective of the experiments described in this section is to verify if the system is able to deliver the correct RCC and Allen's relations that composes the LH Interval Calculus theory using different methods of object identification.

Figure 39 shows the environment and the 14 objects considered in the LH Interval Calculus evaluation. Based in Figure 39, a gold standard table was created by the author, observing and annotating the LH Interval Calculus relations, to serve as a comparison to the answers of the LH Interval Calculus algorithm.

The first experiment employing the pictures collected were done using the VGG Image Annotator (VIA) (DUTTA; ZISSERMAN, 2019), a simple and standalone manual annotation

software for image, audio and video. VIA is an open source project based solely on HTML, Javascript and CSS (no dependency on external libraries). VIA is developed at the Visual Geometry Group (VGG) and released under the BSD-2 clause license which allows it to be useful for both academic projects and commercial applications.

The identification of the objects detected in the scene was done manually, outlining the object within a rectangle. From the total amount of 223 pictures, a set of 102 images was chosen because they contain two or more objects in the scene. After the annotation of the objects present in the 102 pictures, the QSRLib⁴ library was used to allow computation of the Qualitative Spatial Relations. LH Interval Calculus algorithm was developed in Python which makes calls to QSRLib in order to abstract the input data and form a qualitative representation of the perceived world scene.

A second method of object detection was tested. The Yolo V5 ⁵ is Deep Neural Network focused on detecting objects in scenes, with architectures and models pre-trained on the COCO data set. YOLO is an acronym for 'You only look once', due to its algorithm that divides images into a grid system. Each cell in the grid is responsible for detecting objects within itself. The training data set of Yolo V5 has 80 classes, more than 1.500 images per class and about 10.000 labeled objects per class. The object detection is a tool for this work and not its focus, so the Yolo V5 trained data set was enough to detect the bounding boxes of the 14 objects of the experiment.

Each noted rectangle (manually or automatic) present in the scene is considered an object. The horizontal lines of the rectangle represent the dimension and localization of the object in the axis x and the vertical lines of the rectangle represent the dimension and localization of the object in the axis y .

The output of the algorithm is a qualitative relation for the objects in the axis x and a qualitative relation in the axis y .

5.2.1 Evaluation Method

The parameter to evaluate the LH Interval Calculus is the accuracy of the algorithm response for the whole process: get information of the objects identified in VGG and Yolo V5 from an agent point of view, identify RCC relations and identify Allen's Interval Algebra Relation projected in the axis x and y . The output is compared to the gold standard table, defined as the

⁴<https://qsrlib.readthedocs.io/en/latest/index.html>

⁵<https://github.com/ultralytics/yolov5>



Figure 40 – VGG object identification and labels

true base of knowledge. The gold standard table comprises all possible qualitative relations between the 14 objects to each other, and was created manually by an expert.

The accuracy is the proportion of correct results, among all the results and is given by the formula:

$$accuracy = \frac{\text{correct results}}{\text{all results}} \quad (1)$$

This method was applied to the LH Interval Calculus as well as to the Collaborative Spatial Reasoning evaluation. The result is considered correct when the output of the algorithm shows the same relations of the gold standard table and considered incorrect when the output shows any other relation.

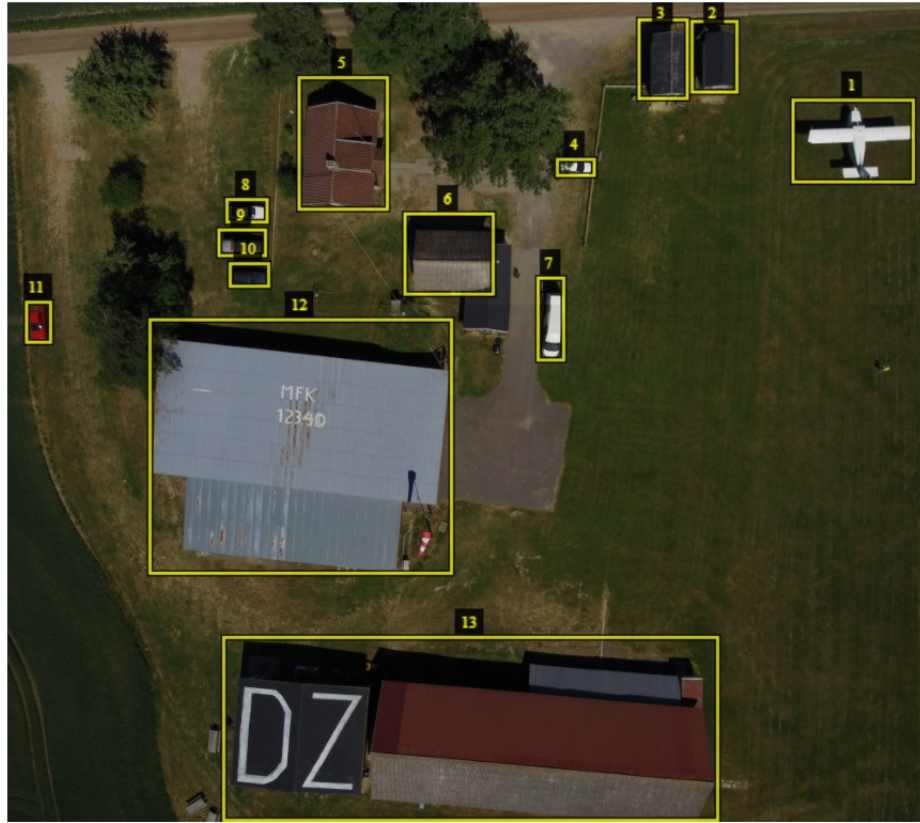


Figure 41 – Yolo V5 object identification and labels

5.2.2 LH Interval Calculus Results

The gold standard table considers all RCC and Allen's relations possible to the 14 objects each other ($Object_1$ and its relations to $Object_2$, $Object_3$, $Object_4$ and so on, $Object_2$ and its relations to $Object_1$, $Object_3$, $Object_4$ and so on, till $Object_{14}$ and its relations to $Object_1$, $Object_2$, $Object_3$ and so on), adding up 364 pairs of relations, being 182 RCC relations and 182 Allen's relations, with the UAV flying to North cardinal direction. Figure 40 shows the 14 objects and its labels.

The values used in the algorithm that makes calls to the QSRLib are the medium position of each rectangle in the axis x and y , and the rectangle extent in the axis x and y . These coordinates give the position of the rectangle and hence the position of the object that the rectangle represents. The QSRLib algorithm gives two outputs for each pair of objects. The first one is the RCC relation and the second one is the Allen's relation for the object in the axis x and y . The combination of these two outputs is tested according to the constraints that define the LH Interval Calculus and were presented in Section 4.2.

The 102 images have from 2 to 11 different objects in the environment. The total amount of LH Interval Calculus sentences extracted from the 102 images is 996. Table 6 shows 5

Table 5 – Accuracy for LH Interval Calculus relations between the objects at Motala Flygklubb.

Objects	<i>Object</i> ₁	<i>Object</i> ₂	<i>Object</i> ₃	<i>Object</i> ₄	<i>Object</i> ₅	<i>Object</i> ₆	<i>Object</i> ₇
Accuracy	92%	77%	77%	85%	69%	69%	77%
Objects	<i>Object</i> ₈	<i>Object</i> ₉	<i>Object</i> ₁₀	<i>Object</i> ₁₁	<i>Object</i> ₁₂	<i>Object</i> ₁₃	<i>Object</i> ₁₄
Accuracy	77%	92%	85%	85%	85%	23%	0%

sentences of example for each object, demonstrating the expected results x obtained results, because to show the 996 sentences was considered unnecessary.

The VGG identification of the objects resulted in 100% of correct responses comparing to the gold standard table, because the rectangle outlining each object present in the scene was done manually and contains only the object itself, discarding shadows or any background.

The automatic object identification using Yolo V5 offered a different result. The shadow of all objects were identified as part of the object by Yolo V5 algorithm (Figure 41). Besides this, the black roof of the *Object*₆ is a balcony, part of the building, but was excluded from the rectangle that surrounds the object. The *Object*₁₃ is composed by two buildings. One is covered by the black roof with the letters "DZ" and the other is covered by the red and gray roof, but they were identified as a single object. Consequently, *Object*₁₄ was not identified anyway.

The RCC relations obtained using Yolo V5 object identification are correct because the rectangles remain "*disconnected*" from each other, despite the problems already mentioned. The Allen's relations presented some incorrect responses, to be analyzed below.

Table 5 shows the accuracy of the responses of the LH Interval Calculus to the 14 objects present at Motala Flugklubb (Figure 41), regarding RCC relations between the objects present in the images captured by the UAVs and identification of Allen's Interval Algebra Relation projected in the axis x and y .

LH Interval Calculus presented 92% of accuracy for the relations between *Object*₁, an airplane, and the others 13 objects of the scene. This means that for all images where *Object*₁ was identified, 92% of the relations between *Object*₁ and any other objects identified in the scene is correct. The 8% of wrong answers refers to the relations between *Object*₁ and *Object*₁₄, not identified by Yolo V5 algorithm.

*Object*₂ and *Object*₃, are small buildings, and presented 77% of accuracy for the LH Interval Calculus sentences. The 23% of incorrect responses showed up in the relations with

Table 6 – LH Interval Calculus Expected Results x Obtained Results

<i>Object₁</i>	
<i>Expected</i>	<i>Obtained</i>
Obj1,Obj13:(DC: 'p,p')	Obj1,Obj13:(DC: 'p,p')
Obj1,Obj10:(DC: 'p,p')	Obj1,Obj10:(DC: 'p,p')
Obj1,Obj5:(DC: 'p,d')	Obj1,Obj5:(DC: 'p,d')
Obj1,Obj4:(DC: 'p,-d')	Obj1,Obj4:(DC: 'p,-d')
Obj1,Obj14:(DC: 'p,p')	\emptyset
<i>Object₂</i>	
<i>Expected</i>	<i>Obtained</i>
Obj2,Obj1:(DC: 'p,p')	Obj2,Obj1:(DC: 'p,p')
Obj2,Obj14:(DC: 'o,p')	\emptyset
Obj2,Obj10:(DC: 'p,p')	Obj2,Obj10:(DC: 'p,p')
Obj2,Obj5:(DC: 'p,p')	Obj2,Obj5:(DC: 'p,o')
Obj2,Obj13:(DC: 'p,p')	Obj2,Obj13:(DC: 'o,p')
<i>Object₃</i>	
<i>Expected</i>	<i>Obtained</i>
Obj3,Obj1:(DC: 'p,p')	Obj3,Obj1:(DC: 'p,p')
Obj3,Obj14:(DC: 'd,p')	\emptyset
Obj3,Obj10:(DC: 'p,p')	Obj3,Obj10:(DC: 'p,p')
Obj3,Obj5:(DC: 'p,p')	Obj3,Obj5:(DC: 'p,o')
Obj3,Obj13:(DC: 'p,p')	Obj3,Obj13:(DC: 'd,p')
<i>Object₄</i>	
<i>Expected</i>	<i>Obtained</i>
Obj4,Obj13:(DC: 'p,p')	Obj4,Obj13:(DC: 'd,p')
Obj4,Obj9:(DC: 'p,p')	Obj4,Obj9:(DC: 'p,p')
Obj4,Obj3:(DC: 'p,-p')	Obj4,Obj3:(DC: 'p,-p')
Obj4,Obj2:(DC: 'p,-p')	Obj4,Obj2:(DC: 'p,-p')
Obj4,Obj12:(DC: 'p,p')	Obj4,Obj12:(DC: 'p,p')

Object₁₄, *Object₁₃* and *Object₅* as shown in Table 6. As *Object₁₄* was not identified, *Object₁₃*

<i>Object₅</i>	
<i>Expected</i>	<i>Obtained</i>
Obj5,Obj13:(DC: 'o,p')	Obj5,Obj13:(DC: 'd,p')
Obj5,Obj3:(DC: 'p,p')	Obj5,Obj3:(DC: 'p,o')
Obj5,Obj2:(DC: 'p,p')	Obj5,Obj2:(DC: 'p,o')
Obj5,Obj12:(DC: 'd,p')	Obj5,Obj12:(DC: 'd,p')
Obj5,Obj6:(DC: 'p,p')	Obj5,Obj6:(DC: 'p,p')
<i>Object₆</i>	
<i>Expected</i>	<i>Obtained</i>
Obj6,Obj2:(DC: 'p,p')	Obj6,Obj2:(DC: 'p,p')
Obj6,Obj12:(DC: 'o,p')	Obj6,Obj12:(DC: 'o,p')
Obj6,Obj11:(DC: 'p,p')	Obj6,Obj11:(DC: 'p,o')
Obj6,Obj7:(DC: 'p,o')	Obj6,Obj7:(DC: 'p,o')
Obj6,Obj8:(DC: 'p,o')	Obj6,Obj8:(DC: 'p,p')
<i>Object₇</i>	
<i>Expected</i>	<i>Obtained</i>
Obj7,Obj10:(DC: 'p,p')	Obj7,Obj10:(DC: 'p,o')
Obj7,Obj5:(DC: 'p,p')	Obj7,Obj5:(DC: 'p,p')
Obj7,Obj4:(DC: 'o,p')	Obj7,Obj4:(DC: 'o,p')
Obj7,Obj13:(DC: 'p,p')	Obj7,Obj13:(DC: 'd,p')
Obj7,Obj9:(DC: 'p,p')	Obj7,Obj9:(DC: 'p,p')
<i>Object₈</i>	
<i>Expected</i>	<i>Obtained</i>
Obj8,Obj13:(DC: 'o,p')	Obj8,Obj13:(DC: 'd,p')
Obj8,Obj3:(DC: 'p,p')	Obj8,Obj3:(DC: 'p,p')
Obj8,Obj2:(DC: 'p,p')	Obj8,Obj2:(DC: 'p,p')
Obj8,Obj12:(DC: 'd,p')	Obj8,Obj12:(DC: 'd,p')
Obj8,Obj6:(DC: 'p,p')	Obj8,Obj6:(DC: 'p,o')

is considered bigger than it really is, so, the *is overlapped* relation for the object projections on axis x of the output sentence $Object_2, Object_{13}(DC: -o_x, p_y)$: Σ_I is inconsistent.

A correct bounding box around $Object_{13}$ should satisfy the following sentence: $Object_2 DC Object_{13} \cap ext(L_{13}) p_x ext(L_2)$.

<i>Object₉</i>	
<i>Expected</i>	<i>Obtained</i>
Obj9,Obj6:(DC: 'p,d')	Obj9,Obj6:(DC: 'p,d')
Obj9,Obj11:(DC: '-p,p')	Obj9,Obj11:(DC: '-p,p')
Obj9,Obj7:(DC: 'p,p')	Obj9,Obj7:(DC: 'p,p')
Obj9,Obj8:(DC: 'o,-p')	Obj9,Obj8:(DC: 'o,-p')
Obj9,Obj4:(DC: 'p,-p')	∅
<i>Object₁₀</i>	
<i>Expected</i>	<i>Obtained</i>
Obj10,Obj12:(DC: 'd,p')	Obj10,Obj12:(DC: 'd,p')
Obj10,Obj6:(DC: 'p,d')	Obj10,Obj6:(DC: 'p,d')
Obj10,Obj11:(DC: '-p,p')	Obj10,Obj11:(DC: '-p,p')
Obj10,Obj7:(DC: 'p,o')	Obj10,Obj7:(DC: 'p,p')
Obj10,Obj8:(DC: '-o,-p')	Obj10,Obj8:(DC: '-o,-p')
<i>Object₁₁</i>	
<i>Expected</i>	<i>Obtained</i>
Obj11,Obj9:(DC: 'p,-p')	Obj11,Obj9:(DC: 'p,-p')
Obj11,Obj3:(DC: 'p,-p')	Obj11,Obj3:(DC: 'p,-p')
Obj11,Obj2:(DC: 'p,-p')	Obj11,Obj2:(DC: 'p,-p')
Obj11,Obj12:(DC: 'p,o')	Obj11,Obj12:(DC: 'p,o')
Obj11,Obj6:(DC: 'p,-p')	Obj11,Obj6:(DC: 'p,-o')
<i>Object₁₂</i>	
<i>Expected</i>	<i>Obtained</i>
Obj12,Obj5:(DC: '-d,-p')	Obj12,Obj5:(DC: '-d,-p')
Obj12,Obj4:(DC: 'p,-p')	Obj12,Obj4:(DC: 'p,-p')
Obj12,Obj13:(DC: 'o,p')	Obj12,Obj13:(DC: '-d,p')
Obj12,Obj9:(DC: '-d,-p')	Obj12,Obj9:(DC: '-d,-p')
Obj12,Obj3:(DC: 'p,-p')	Obj12,Obj3:(DC: 'p,-p')

As the bounding box for *Object₁₃* identification is bigger in axis *x* than it should be, the sentence satisfied is the following: *Object₂* *O* *Object₁₃* \cap *ext(L₁₃)* *o_x* *ext(L₂)*.

<i>Object₁₃</i>	
<i>Expected</i>	<i>Obtained</i>
Obj13,Obj1:(DC: 'p,-p')	Obj13,Obj1:(DC: 'p,-p')
Obj13,Obj10:(DC: '-d,-p')	Obj13,Obj10:(DC: '-d,-p')
Obj13,Obj5:(DC: '-d,-p')	Obj13,Obj5:(DC: 'o,-p')
Obj13,Obj4:(DC: '-d,-p')	Obj13,Obj4:(DC: 'p,-p')
Obj13,Obj9:(DC: '-o,-p')	Obj13,Obj9:(DC: '-o,-p')
<i>Object₁₄</i>	
<i>Expected</i>	<i>Obtained</i>
Obj14,Obj3:(DC: '-d,-p')	∅
Obj14,Obj2:(DC: 'o,-p')	∅
Obj14,Obj12:(DC: '-o,-p')	∅
Obj14,Obj6:(DC: '-d,-p')	∅
Obj14,Obj11:(DC: '-p,-p')	∅

The same reasoning applies to *Object₅*, identified as bigger than it really is on axis y due to its shadow. The sentence that should be satisfied is *Object₂ DC Object₅ ∩ ext(H₂) p_y ext(H₅)* instead of *Object₂ O Object₅ ∩ ext(H₂) o_y ext(H₅)*.

Object₄, a car, has 85% of accuracy in the LH Interval Calculus responses. The 15% of inconsistencies refers to the relations between *Object₄*, *Object₁₃* and *Object₁₄* for the same reasons explained in the paragraph above. *Object₄ DC Object₁₃ ∩ ext(L₁₃) p_x ext(L₄)* is the sentence that should be satisfied.

Object₅, a medium sized building, presented an accuracy of 69% due to its rectangle of identification which takes into account the building's shadow. 31% of the LH Interval Calculus inconsistent sentences are related to *Object₂*, *Object₃*, *Object₁₃* and *Object₁₄*, for the same reasons exposed above.

Object₆ is also a medium sized building, but the rectangle that outlines the object is smaller than the object. The balcony of the building was left outside the annotation while the shadow of the building was included. These facts leads to an accuracy of 69% for the LH Interval Calculus. The 31% of the relations that presented problems refers to *Object₁₃*, *Object₁₄*, *Object₈* and *Object₁₁*. The inconsistency of the relation with *Object₈* appears due to the inclusion of the building's shadow in the annotation and the problem with *Object₁₁* due to the balcony exclusion.

The correct sentence that should be satisfied for $Object_6$ and $Object_8$ position is $Object_6 \text{ DC } Object_8 \cap ext(H_8) p_y ext(H_6)$. For $Object_6$ and $Object_{11}$, the sentence $Object_6 \text{ O } Object_{11} \cap ext(H_6) o_y ext(H_{11})$ is the sentence that should be satisfied.

$Object_7$ is a van, and presented an accuracy of 77%. The 23% of inconsistent sentences refers to $Object_7$ relations to $Object_{13}$, $Object_{14}$ and $Object_{10}$. Again, the inclusion of the shadow of the van in the rectangle of object identification caused the inconsistency. The target sentence that should be satisfied is $Object_7 \text{ DC } Object_{10} \cap ext(H_{10}) p_y ext(H_7)$.

$Object_8$ is a car and its accuracy in the LH Interval Calculus is 77%. Over again $Object_{13}$ and $Object_{14}$ didn't contribute to consistent outputs as well as $Object_6$, due to its shadow included in the annotation rectangle. To be succinct, the sentence for $Object_8$ regarding to $Object_6$ is complementary to the sentence demonstrated in the paragraph describing $Object_6$, keeping in mind that if $Object_6$ is preceded by $Object_8$, $Object_8$ precedes $Object_6$.

$Object_9$ is also a car. Its accuracy is 92%. The only problem is the relation with $Object_{14}$, not identified by Yolo V5.

$Object_{10}$, another car, presented an accuracy of 85% in its relations with the objects considered in this experiment. The 15% of inconsistencies are related to $Object_{14}$ and $Object_7$, already explained. The sentence for $Object_{10}$ and $Object_7$ is the complementary of $Object_7$ and $Object_{10}$.

$Object_{11}$ represents a car and presented an accuracy of 85% for LH Interval Calculus. 15% of incorrect sentences refer to the relation with $Object_{14}$ and $Object_6$, as explained above.

$Object_{12}$ is a big building, and its accuracy is 85%. The $Object_{13}$ and $Object_{14}$, are responsible for the 15% of inconsistencies.

$Object_{13}$ is another big building, with 23% of accuracy in the LH Interval Calculus, followed by $Object_{14}$ with 0%.

The global accuracy of the LH Interval Calculus using automatic annotation of objects is 77%. This result shows that qualitative spatial reasoning can assimilate some imprecision in the object identification. On the other hand, when the object outline identification is largely imprecise, the inconsistencies are unavoidable.

The next section will describe the discussion about the imprecision at LH Interval Calculus and its consequences for the Collaborative Spatial Reasoning.

5.3 COLLABORATIVE SPATIAL REASONING EVALUATION

The input data for the Collaborative Spatial Reasoning system are the results of the LH Interval Calculus. To combine two partial fields of view that complement each other, the information about the cardinal direction of flight becomes necessary. As demonstrated in Section 4.3, Figure 35, there is a correspondence position according to agents orientation of flight. As the scan pattern performed in the experiment was the lawnmower (Figure 33), the agents flew following the North, East, South and West directions, according to the information registered by the compass of the drones and available for conference in the bag file.

First of all, the LH Interval Calculus qualitative relations responses were normalized to North direction of flight, as explained in Section 4.3, Figure 35. The images that had at least one object in common were used to test the Collaborative Spatial Reasoning algorithm, aiming at a partial view combination of two images, to infer information about objects not present in both images. Each object has a RCC relation and a pair of Allen's relations, projected in the axis x and y . The RCC relation of each pair of objects in common in a picture results in a set of one or more possible relations for a third object, given by the composition table rules for RCC (Figure 4). The same method of composition is applied to the axis x and axis y relations of the pair of objects, submitted to the composition table of Allen's relations (Figure 36).

As explained in Section 4.2, each RCC relation interacts with 2 Allen's relations. So, a cross check is done between the results of the RCC composition table and Allen's composition table to compose the location of the third object related to the two first ones. Only the set of relations that agrees with the RCC and Allen's relations interactions are considered in the final result.

5.3.1 Collaborative Spatial Reasoning Results

From 102 images showing two or more objects in the scene, 30 were chosen to test the Collaborative Spatial Reasoning theory. All relations between the objects is normalized to the North direction view according to the correspondence analyses explained in Section 4.3. In these 30 images, all 14 objects are present in the scene at least once.

The result of a partial view combination of 2 images with one object in common, in the most cases, is a set of possible relations between the objects that don't appear in both images. Table 7 shows examples of partial view images combination with one object present in both images and the inference information about a third object.

Table 7 – Collaborative Spatial Reasoning inference for two images with one object in common

Obj Pair Image 1	Obj Pair Image 2	Expected Result	Obtained Result
Obj2,Obj1: (DC: 'p,p')	Obj1,Obj4: (DC: '-p,-d')	Obj2,Obj4: (DC: '-p,p')	Obj2,Obj4: DC:[('p', 'p'), ('m', 'p'), ('o', 'p'), ('s', 'p'), ('d', 'p'), ('-f', 'p'), ('eq', 'p'), ('f', 'p'), ('-d', 'p'), ('-s', 'p'), ('-o', 'p'), ('-m', 'p'), ('-p', 'p')]
Obj14,Obj13: (DC: 'm,d')	Obj13,Obj7: (DC: '-d,-p')	Obj14,Obj7: (DC: 'p,p')	Obj14,Obj7: DC:[('p', '-p')]
Obj5,Obj10: (DC: '-p,p')	Obj10,Obj11: (DC: '-p,p')	Obj5,Obj11: (DC: '-p,p')	Obj5,Obj11: DC:[('p', 'p')]
Obj4,Obj5: (DC: '-p,d')	Obj5,Obj8: (DC: '-p,o')	Obj4,Obj8: (DC: '-p,p')	Obj4,Obj8: DC:[('p', 'p'), ('-p', 'm'), ('p', 'o'), ('-p', 's'), ('p', 'd')]
Obj6,Obj4: (DC: 'p,-p')	Obj4,Obj1: (DC: 'p,d')	Obj6,Obj1: (DC: 'p,p')	Obj6,Obj1: DC:[('p', 'd'), ('p', 'f'), ('p', '-o'), ('p', '-m'), ('p', '-p')]
Obj12,Obj14: (DC: '-d,p')	Obj14,Obj13: (DC: 'm,d')	Obj12,Obj13: (DC: 'o,p')	Obj12,Obj13: DC:[('o', 'p'), ('o', 'm'), ('o', 'o'), ('o', 's'), ('o', 'd'), ('f', 'p'), ('-f', 'm'), ('f', 'o'), ('-f', 's'), ('f', 'd'), ('-d', 'p'), ('d', 'm'), ('-d', 'o'), ('d', 's'), ('-d', 'd')]
Obj11,Obj9: (DC: 'p,-p')	Obj9,Obj5: (DC: 'p,-p')	Obj11,Obj5: (DC: 'p,-p')	Obj11,Obj5: DC:[('p', '-p')]
Obj2,Obj3: (DC: '-p,d')	Obj3,Obj4: (DC: '-p,p')	Obj2,Obj4: (DC: '-p,p')	Obj2,Obj4: DC:[('p', 'p')]
Obj1,Obj2: (DC: '-p,-p')	Obj2,Obj5: (DC: '-p,p')	Obj1,Obj5: (DC: '-p,d')	Obj1,Obj5: DC:[('p', 'p'), ('-p', 'm'), ('p', 'o'), ('-p', 's'), ('p', 'd'), ('-p', '-f'), ('p', 'eq'), ('-p', 'f'), ('p', '-d'), ('-p', '-s'), ('p', '-o'), ('-p', '-m'), ('p', '-p')]

Obj Pair Image 1	Obj Pair Image 2	Expected Result	Obtained Result
Obj14,Obj12: (DC: 'd,-p')	Obj12,Obj6: (DC: 'o,-p')	Obj14,Obj6: (DC: 'p,-p')	Obj14,Obj6: DC:[('p', '-p'), ('m', '-p'), ('o', '-p'), ('s', '-p'), ('d', '-p')]
Obj7,Obj6: (DC: '-p,-o')	Obj6,Obj1: (DC: 'p,-p')	Obj7,Obj1: (DC: 'p,-p')	Obj7,Obj1: DC:[('p', '-p'), ('m', '-p'), ('o', '-p'), ('s', '-p'), ('d', '-p'), ('-f', '-p'), ('eq', '-p'), ('f', '-p'), ('-d', '-p'), ('-s', '-p'), ('-o', '-p'), ('-m', '-p'), ('-p', '-p')]
Obj8,Obj11: (DC: '-p,p')	Obj11,Obj13: (DC: 'p,p')	Obj8,Obj13: (DC: 'o,p')	Obj8,Obj13: DC:[('p', 'p'), ('m', 'p'), ('o', 'p'), ('s', 'p'), ('d', 'p'), ('-f', 'p'), ('eq', 'p'), ('f', 'p'), ('-d', 'p'), ('-s', 'p'), ('-o', 'p'), ('-m', 'p'), ('-p', 'p')]
Obj12,Obj7: (DC: 'p,-o')	Obj7,Obj4: (DC: 'o,-p')	Obj12,Obj4: DC: 'p,-p')	Obj12,Obj4: DC:[('p', '-p')]
Obj10,Obj8: (DC: '-o,-p')	Obj8,Obj9: (DC: '-o,p')	Obj10,Obj9: (DC: '-o,-p')	Obj10,Obj9: DC:[('o', 'p'), ('-o', 'm'), ('-o', 'o'), ('-o', 's'), ('-o', 'd'), ('-o', '-f'), ('-o', 'eq'), ('-o', 'f'), ('-o', '-d'), ('-o', '-s'), ('-o', '-o'), ('-o', '-m'), ('-o', '-p'), ('-m', 'p'), ('-m', 'm'), ('-m', 'o'), ('-m', 's'), ('-m', 'd'), ('-m', '-f'), ('-m', 'eq'), ('-m', 'f'), ('-m', '-d'), ('-m', '-s'), ('-m', '-o'), ('-m', '-m'), ('-m', '-p'), ('-p', 'p'), ('-p', 'm'), ('-p', 'o'), ('-p', 's'), ('-p', 'd'), ('-p', '-f'), ('-p', 'eq'), ('-p', 'f'), ('-p', '-d'), ('-p', '-s'), ('-p', '-o'), ('-p', '-m'), ('-p', '-p')]

In the tests considering VGG identification, the reference relation of the gold standard was present in 100% of the resulting sets of possible relations showing in Table 7, Column “Obtained Result”, given by RCC composition table and Allen’s relations composition table results.

It was observed in some results that 2 images tested with 3 objects, being 2 objects present in one image and 2 in another image (for example: *Object*₁₁ and *Object*₇ in one image and *Object*₇ and *Object*₁ in another image) the result for the relations between *Object*₁₁ and *Object*₁ is not a set, but the exact pair of relations. It happened for 57% of the combinations, which can be seen as a good result, taking into account that qualitative theories are not considered precise. This result shows that combinations of qualitative relations doesn’t increase the imprecision that already exists in the system.

From the remaining 43% combinations of objects, the result is a set of pairs of relations (Table 7, Column “Obtained Result”), which 78% resulted in the exact relation for one of the axis (relation for axis *x* or relation for axis *y*) and the accuracy of the set of possible relations for the remaining axis varies from 8% to 20%.

The accuracy of the 22% combinations of objects which none of the axis received as result an exact relation, the accuracy of the set of possible relations varies from 3% to 7%.

Using Yolo V5 for identification of objects, and keeping in mind that the input for the Collaborative Spatial Reasoning is the LH Interval Calculus, the accuracy showed in Section 5.2.2 guaranteed that only 7% of the combinations of objects tested from the 30 images chosen, resulted in a set of possible relations where the expected pair of relations is not present, this means 0% of accuracy.

Trying to get a better evaluation about the implications of the imprecision from LH Interval Calculus on Collaborative Spatial Reasoning results, a second test was done focused on the inconsistent pairs of relations imputed by LH Interval Calculus.

The failure in the automatic identification of *Object*₁₃ and *Object*₁₄ is the cause of 68% of wrong pairs of relations at Collaborative Spatial Reasoning results. The exclusion of the balcony’s roof and addition of the shadow of *Object*₆ caused 13% of wrong outputs. The shadow is also the reason for 13% of wrong pair of relations when *Object*₅ is present in the scenario. The remaining 6% of inconsistent results is caused by imprecision in the outlining of all other objects.

5.3.2 Discussion

This chapter presented the results of the real situation experiments using two unmanned aerial vehicles (UAVs). The images captured by the UAVs were used for the tasks of object identification, followed by the application of Spatial Qualitative Reasoning to extract the relations between the objects.

The first contribution of this work, the LH Interval Calculus, a formalism that merges the concepts of Region Connection Calculus (RCC) and Allen's Interval Algebra (AIA) demonstrate very satisfactory results. The relations between the pairs of objects in RCC and AIA for axis x and axis y showed successful results for manual identification of objects as well as for automatic identification using Yolo V5.

The imprecision in the object identification led to the inconsistent results already explained in section 5.2.2. This problem can be softened exploring the theories of Computer Vision, to get a better object identification and a more precise bounding box around each object. In this work a proper net was not trained at Yolo V5 due to the few amount of images and also because Computer Vision techniques are not the focus of the work. But it is known that a specific trained network could result in better accuracy for the system.

Even with an imperfect outlining of the objects, the relations precede, is preceded, overlaps, is overlapped, during and is contained are able to absorb some imprecision in case the projection of a pair of objects in the axis x and y have enough distance to deal with the imprecision. On the other hand, a minimum deviation in the rectangle that sketches the objects totally affects the relations meets, is met, starts, is started, finishes, is finished and equal.

The second contribution of this work, the Collaborative Spatial Reasoning formalism, showed good results for the objective of infer the relative position of a third object given the relation of two other objects. Inferring the relative position of objects present in two different images, captured by UAVs flying over different regions of an environment and capturing pictures with a partial view of the scene, it is possible to combine two complementary scenarios.

The precision of the object identification impacted the accuracy of the Collaborative Spatial Reasoning that uses the LH Interval Calculus as input. Besides the Computer Vision techniques, the system can be improved if more than one object in common is used to infer the position of a third object. A deeper study in this matter is suggested for future work.

6 CONCLUSION

The problem of interaction and communication between multi agent teams involving humans and robotic agents in a common mission or task is not completely solved, however, for tasks where precision is not the main objective, the use of qualitative data showed to be a good solution.

The combination of diverse qualitative calculus can be used in many tasks, as search and rescue of victims in disaster scenarios, environment mapping or any other domain where humans and multiple robots need to cooperate for a common objective.

To cope with the interaction and communication problem between multi agent systems, this work introduced two novel solutions. The first, LH Interval Calculus allows the exchange of information between agents, from different points of view, through Allen's Interval Algebra relations projected in axis x and axis y , associated with Region Connection Calculus relations, which can be easily seen and identified by robots as well as by specialists.

The second solution, the Collaborative Spatial Reasoning, allows different agents to feed the system with its own partial perceptions of the environment and associate those information to get knowledge about the complete scenario.

The two contributions of this work were tested using real data, captured by autonomous unmanned aerial vehicles, in the environment and methodology described in Chapter 5. The images captured showed 14 objects, identified in two different ways: manually and using the deep neural network Yolo V5.

The results presented in Chapter 5 showed that LH Interval Calculus and Collaborative Spatial Reasoning bring relevant contributions to environment mapping using qualitative data. While image transmission can be a big issue in certain regions, the transmission of simple sentences describing the environment disposition of the objects is much easier and, for the experiments executed in this work, can be easily understood by the specialist operator in ground, promoting a more intuitive and natural interaction between robots and humans.

For future work, it is very interesting to study the amount of objects in common in two complementary images of an environment are necessary to increase the accuracy of the Collaborative Spatial Reasoning. As the association of objects not present in both images take into account the composition table, the association of more than two pairs can improve the percentage of correct answers.

It is also relevant to include different kinds of agents, and not only autonomous unmanned aerial vehicles. To test the LH Interval Calculus and Collaborative Spatial Reasoning in agents with aerial and terrestrial points of view can bring interesting results, including the addition of the Occlusion Calculus in the terrestrial point of view.

To integrate the contributions of this work with probabilistic models of environment mapping and robots self localization is recommended. The works described in Chapter 3 show that associating qualitative formalism with probabilistic models increases the accuracy of systems of robots localization, motion, guidance and mapping.

BIBLIOGRAPHY

ALLEN, J. F. Maintaining knowledge about temporal intervals. **IEEE Computer Society Press**, Los Alamitos, CA, USA, pp. 248–259, 1990.

BALBIANI, P.; CONDOTTA, J. F.; CERRO, L. F. A model for reasoning about bidimensional temporal relations. **Principles of Knowledge Representation and Reasoning**, Morgan Kaufmann Publishers, pp. 124–130, 1998.

BALBIANI, P.; CONDOTTA, J. F.; CERRO, L. F. Tractability results in the block algebra. **Journal of Logic and Computation**, Oxford University Press, vol. 12, no. 5, pp. 885–909, 2002.

BENNETT, B.; COHN, A.; MAGEE, D. Enforcing global spatio-temporal consistency to enhance reliability of moving object tracking and classification. **Künstliche Intelligenz**, vol. 2, pp. 32–35, 2005.

BONNER, A. J.; KIFER, M. An Overview of Transaction Logic. **Theoretical Computer Science**, vol. 133, pp. 205–265, Oct. 1994.

BURGARD, W. et al. Collaborative multi-robot exploration. **IEEE Computer Society Press**, vol. 1, pp. 476–481, 2000.

CHEN, J. et al. A survey of qualitative spatial representations. **The Knowledge Engineering Review**, Cambridge University Press, vol. 30, no. 1, pp. 106–136, 2015.

CHEN, S. Y. Kalman filter for robot vision: a survey. **IEEE Transactions on Industrial Electronics**, IEEE, vol. 59, no. 11, pp. 4409–4420, 2011.

COHN, A. G.; HAZARIKA, S. M. Qualitative Spatial Representation and Reasoning: An Overview. **Fundamenta Informaticae**, vol. 46, no. 1-2, pp. 1–29, 2001.

COHN, A. G. et al. Cognitive Vision: Integrating Symbolic Qualitative Representations with Computer Vision. **Cognitive Vision Systems**, Springer, Berlin, Heidelberg, vol. 3948, pp. 221–246, 2006.

COHN, A. G. et al. Representing and Reasoning with Qualitative Spatial Relations about Regions. Ed. by Oliviero Stock. **Spatial and Temporal Reasoning**, Springer, pp. 97–134, 1997.

CUI, Z.; COHN, A.G.; RANDELL, D.A. Qualitative Simulation Based on A Logic of Space and Time. **AAAI**, California, U.S., pp. 679–684, 1992.

DELLAERT, F. et al. Monte carlo localization for mobile robots. **IEEE Robotics and Automation**, vol. 2, pp. 1322–1328, 1999.

DOHERTY, P.; HEINTZ, F.; KVARNSTRÖM, J. High-Level Mission Specification and Planning for Collaborative Unmanned Aircraft Systems Using Delegation. **Unmanned Systems**, vol. 01, no. 01, pp. 75–119, 2013. DOI: 10.1142/S2301385013500052. eprint: <http://www.worldscientific>.

com/doi/pdf/10.1142/S2301385013500052. Available from: <http://www.worldscientific.com/doi/abs/10.1142/S2301385013500052>.

DOHERTY, P. et al. A Collaborative Framework for 3D Mapping Using Unmanned Aerial Vehicles. **PRIMA 2016: Principles and Practice of Multi-Agent Systems**, Springer International Publishing, Cham, vol. 9862, pp. 110–130, 2016.

DOHERTY, P. et al. Research with Collaborative Unmanned Aircraft Systems. **Proceedings of the Dagstuhl Workshop on Cognitive Robotics**, Schloss Dagstuhl, Wadern, Germany, Feb. 2010.

DUTTA, A.; ZISSERMAN, A. The VIA Annotation Software for Images, Audio and Video. **Proceedings of the 27th ACM International Conference on Multimedia**, ACM, Nice, France, 2019. DOI: 10.1145/3343031.3350535. Available from: <https://doi.org/10.1145/3343031.3350535>.

FALOMIR, Z.; PICH, A.; COSTA, V. Spatial reasoning about qualitative shape compositions. **Annals of Mathematics and Artificial Intelligence**, Springer, vol. 88, no. 5, pp. 589–621, 2020.

FALOMIR, Z.; PLAZA, E. Towards a model of creative understanding: deconstructing and recreating conceptual blends using image schemas and qualitative spatial descriptors. **Annals of Mathematics and Artificial Intelligence**, Springer, vol. 88, no. 5, pp. 457–477, 2020.

FENELON, V. **Interpretação de imagens com raciocínio espacial qualitativo probabilístico**. 2014. PhD thesis – Universidade de São Paulo.

FERNYHOUGH, J.; COHN, A. G.; HOGG, D. C. Constructing Qualitative Event Models Automatically from Video Input. **Image and Vision Computing**, Elsevier, vol. 18, pp. 81–103, 2000.

FOGLIARONI, P. et al. A Qualitative Approach to Localization and Navigation Based on Visibility Information. **COSIT**, vol. 9, pp. 312–329, 2009.

FOX, D.; BURGARD, W.; THRUN, S. Markov localization for mobile robots in dynamic environments. **Journal of Artificial Intelligence Research**, vol. 11, pp. 391–427, 1999.

FOX, D. et al. A probabilistic approach to collaborative multi-robot localization. **Autonomous robots**, Springer, vol. 8, no. 3, pp. 325–344, 2000.

FRANK, A. U. Qualitative spatial reasoning: Cardinal directions as an example. **International Journal of Geographical Information Science**, Taylor & Francis, vol. 10, no. 3, pp. 269–290, 1996.

FREKSA, C. Using Orientation Information for Qualitative Spatial Reasoning. **Theories and Methods of Spatial-Temporal Reasoning in Geographic Space**, Springer-Verlag, vol. 629, 1992.

GALTON, A. Lines of Sight. **Proc. of the Seventh Annual Conference of AI and Cognitive Science**, Dublin, Ireland, pp. 103–113, 1994.

GOTTS, N. M. How far can we C? Defining a doughnut using connection alone. **Principles of Knowledge Representation and Reasoning: Proceedings of the 4th International Conference (KR94)**, Morgan Kaufmann, pp. 246–257, 1994.

HAZARIKA, S. M.; COHN, A. G. Abducing Qualitative Spatio-Temporal Histories from Partial Observations. **Proc. of KR**, Toulouse, France, pp. 14–25, 2002.

HOLMBERG, G. Unmanned Aviation – New Opportunities and Challenges. **Proceedings of Symposium The Digital Body and a Digital World – Communication and Visualization Technology Serving Humans**, Royal Swedish Academy of Sciences, Linköping and Norrköping, p. 8, 2015. Available from: <https://www.kva.se/globalassets/kalendarium/2015/150410-abstracts.pdf>.

KANDA, T. et al. A humanoid robot that pretends to listen to route guidance from a human. **Autonomous Robots**, Springer, vol. 22, no. 1, p. 87, 2007.

KÖHLER, C. The Occlusion Calculus. **Proc. of Cognitive Vision Workshop**, Zürich, Switzerland, 2002.

KRUSE, T. et al. Human-aware robot navigation: A survey. **Robotics and Autonomous Systems**, Elsevier, vol. 61, no. 12, pp. 1726–1743, 2013.

LEVITT, T. S.; LAWTON, D. T. Qualitative navigation for mobile robots. **Artificial intelligence**, Elsevier, vol. 44, no. 3, pp. 305–360, 1990.

LIGOZAT, G. **Qualitative Spatial and Temporal Reasoning**. London, England: John Wiley & Sons, 2013.

LIGOZAT, G. Reasoning about Cardinal Directions. **Journal of Visual Languages & Computing**, Elsevier, vol. 9, no. 1, pp. 23–44, 1998.

MATSUYAMA, T.; HWANG, V. S. **SIGMA: A Knowledge-Based Image Understanding System**. New York, U.S.: Plenum Press, 1990.

MCCARTHY, J. What is artificial intelligence. URL: <http://www-formal.stanford.edu/jmc/whatisai.html>, 2007.

MCCLELLAND, M.; CAMPBELL, M.; ESTLIN, T. Qualitative relational mapping for planetary rovers. **Proceedings of Workshop on Intelligent Robotic Systems AAI**. Washington, USA, pp. 110–113, 2013.

MOR, R.; INDELMAN, V. Probabilistic Qualitative Localization and Mapping. **2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)**, pp. 5009–5016, 2020.

MORATZ, R.; RENZ, J.; WOLTER, D. Qualitative Spatial Reasoning about Line Segments. **ECAI**, pp. 234–238, 2000.

MULLER, P. A qualitative theory of motion based on spatio-temporal primitives. **KR**, vol. 98, pp. 131–141, 1998.

MURPHY, R. R. Emergency informatics: using computing to improve disaster management. **Computer**, IEEE, vol. 49, no. 5, pp. 19–27, 2016.

NAGEL, H-H. Image Sequence Evaluation: 30 years and still going strong. **Proc. of ICPR**, Barcelona, Spain, pp. 1149–1158, 2000.

NEHMZOW, U. **Mobile robotics: a practical introduction**. London, England: Springer Science & Business Media, 2012.

PANGERCIC, D. et al. Semantic object maps for robotic housework-representation, acquisition and use. **2012 IEEE/RSJ International Conference on Intelligent Robots and Systems**, pp. 4644–4651, 2012.

PAULIUS, D.; SUN, Y. A survey of knowledge representation in service robotics. **Robotics and Autonomous Systems**, Elsevier, vol. 118, pp. 13–30, 2019.

PERICO, D. H.; SANTOS, P. E.; BIANCHI, R. A. C. Guided navigation from multiple viewpoints using qualitative spatial reasoning. **Spatial Cognition & Computation**, Taylor & Francis, vol. 21, no. 2, pp. 143–172, 2021.

PIRRI, F.; MENTUCCIA, I.; STORRI, S. The domestic robot-a friendly cognitive system takes care of your home. **Ambient Intelligence: Impact on Embedded System Design**. Kluwer Academic, Boston, Springer, pp. 131–159, 2003.

POOLEV, D.; GOEBEL, R.; ALELIUNAS, R. Theorist: A Logical Reasoning System for Defaults and Diagnosis. Ed. by Nick Cercone and Gordon McCalla. **The Knowledge Frontier – Essays in the Representation of Knowledge**, Springer-Verlag, pp. 331–352, 1987.

POTYKA, N. et al. Group Decision Making via Probabilistic Belief Merging. **KI 2016: Advances in Artificial Intelligence: 39th Annual German Conference on AI, Klagenfurt, Austria, September 26-30, 2016, Proceedings**, IJCAI/AAAI Press, vol. 9904, p. 291, 2016.

RANDELL, D.; CUI, Z.; COHN, A. A Spatial Logic Based on Regions and Connection. **Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning**, Cambridge, U.S., pp. 165–176, 1992.

RANDELL, D.; WITKOWSKI, M. Building Large Composition Tables via Axiomatic Theories. **Proc. of KR**, Toulouse, France, pp. 26–35, 2002.

RANDELL, D.; WITKOWSKI, M.; SHANAHAN, M. From Images to Bodies: Modeling and Exploiting Spatial Occlusion and Motion Parallax. **Proc. of IJCAI**, Seattle, U.S., pp. 57–63, 2001.

REITER, R.; MACKWORTH, A. A Logical Framework for Depiction and Image Interpretation. **Artificial Intelligence**, vol. 41, no. 2, pp. 125–155, 1989.

RENTZ, J. **Qualitative spatial reasoning with topological information**. Berlin, Germany: Springer-Verlag, 2002.

SANTOS, M. V. et al. Logic-Based Interpretation of Geometrically Observable Changes Occurring in Dynamic Scenes. **Applied Intelligence**, vol. 31, no. 2, pp. 161–179, 2009.

SANTOS, P. E. Reasoning about Depth and Motion from an Observer's Viewpoint. **Spatial Cognition and Computation**, vol. 7, no. 2, pp. 133–178, 2007.

SANTOS, P. E.; DEE, H. M.; FENELON, V. Qualitative robot localisation using information from cast shadows. **IEEE International Conference on Robotics and Automation**, pp. 220–225, 2009.

SANTOS, P. E.; DEE, H.; FENELON, V. Notes on a qualitative theory of shadows. **Proceedings of the Cognitive Robotics Workshop**, 2008.

SANTOS, P. E.; LIGOZAT, G.; SAFI-SAMGHABAD, M. An Occlusion Calculus Based on an Interval Algebra. **Brazilian Conference on Intelligent Systems (BRACIS)**, pp. 128–133, 2015.

SANTOS, P. E.; SHANAHAN, M. A logic-based algorithm for image sequence interpretation and anchoring. **IJCAI**, Acapulco, Mexico, pp. 1408–1410, 2003.

SANTOS, P. E.; SHANAHAN, M. Hypothesising object relations from image transitions. **Proceedings of the 15th European Conference on Artificial Intelligence**, pp. 292–296, 2002.

SANTOS, P. E. et al. Probabilistic self-localisation on a qualitative map based on occlusions. **Journal of Experimental & Theoretical Artificial Intelligence**, Taylor & Francis, vol. 28, no. 5, pp. 781–799, 2016.

SCHERL, R. B.; LEVESQUE, H. J. The frame problem and knowledge-producing actions. **AAAI**, vol. 93, pp. 689–695, 1993.

SCHLIEDER, C. Qualitative Shape Representation. Ed. by P. A. Burrough and A. U. Frank. **Geographic Objects with Indeterminate Boundaries**, Taylor & Francis Inc., pp. 123–140, 1996.

SCHULTZ, C. P. L.; GUESGEN, H. W.; AMOR, R. Computer-human interaction issues when integrating qualitative spatial reasoning into geographic information systems. **Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI**, pp. 43–51, 2006.

SECOLO, A.C.; SANTOS, P. E. Raciocínio Espacial Qualitativo para um Sistema de Segurança em Tráfego de Veículos. **10º Simpósio Brasileiro de Automação Inteligente (SBAI)**, 2011.

SHANAHAN, M. Robotics and the Common Sense Informatic Situation. **Proceedings of ECAI**, Budapest, Hungary, pp. 684–688, 1996.

SIEGWART, R.; NOURBAKHSI, I. R.; SCARAMUZZA, D. **Introduction to autonomous mobile robots**. Massachusetts, EUA: MIT press, 2011.

SIOUTIS, M.; WOLTER, D. Qualitative Spatial and Temporal Reasoning: Current Status and Future Challenges. **IJCAI**, 2021.

SJANIC, Z. **Navigation and mapping for aerial vehicles based on inertial and imaging sensors**. 2013. PhD thesis – Linköping University Electronic Press.

SOUCHANSKI, M.; SANTOS, P. E. Reasoning about Dynamic Depth Profiles. **Proceedings of the 18th European Conference on Artificial Intelligence (ECAI)**, IOS Press, Amsterdam, Netherlands, pp. 30–34, 2008.

STOCK, O. **Spatial and Temporal Reasoning**. Ed. by Oliviero Stock. Netherlands: Springer Science & Business Media, 1998.

THRUN, S. Robotic mapping: A survey. **Exploring artificial intelligence in the new millennium**, vol. 1, pp. 1–35, 2002.

THRUN, S.; BURGARD, W.; FOX, D. **Probabilistic robotics**. Massachusetts, EUA: MIT press, 2005.

TORRES, M. et al. Qualitative spatial reasoning methodology to determine the particular domain of a set of geographic objects. **Computers in Human Behavior**, Elsevier, vol. 59, pp. 115–133, 2016.

VAN HARMELEN, F.; LIFSCHITZ, V.; PORTER, B. **Handbook of Knowledge Representation**. Oxford, UK: Elsevier, 2008. Qualitative Spatial Representation and Reasoning, pp. 551–596.

WAGNER, T.; VISSER, U.; HERZOG, O. Egocentric qualitative spatial knowledge representation for physical robots. **Robotics and Autonomous Systems**, Elsevier, vol. 49, no. 1, pp. 25–42, 2004.

WALLGRÜN, J. O. Qualitative spatial reasoning for topological map learning. **Spatial Cognition & Computation**, Taylor & Francis, vol. 10, no. 4, pp. 207–246, 2010.

WOLTER, D.; FREKSA, C.; LATECKI, L. J. Towards a generalization of self-localization. **Robotics and cognitive approaches to spatial mapping**, Springer, pp. 105–134, 2007.

WOLTER, D.; WALLGRÜN, J. O. Qualitative spatial reasoning for applications: New challenges and the SparQ toolbox. **Geographic Information Systems: Concepts, Methodologies, Tools, and Applications**, IGI Global, pp. 1639–1664, 2013.

YAMAN, F.; NAU, D. S.; SUBRAHMANYAN, V.S. A Logic of Motion. **AAAI**, pp. 85–94, 2004.

ANNEX A – PUBLICATIONS

Qualitative Environment Mapping Based on Unmanned Aerial Vehicles 3D Images

Adeline Sécoto¹, Paulo Santos¹, Zoran Sjanic², Piotr Rudol², Tommy Persson²,
Mariusz Wzorek², Patrick Doherty²

¹ Centro Universitário FEI, São Paulo, Brazil

² Linköping University, Linköping, Sweden

secolo@fei.edu.br, psantos@fei.edu.br, zoran.sjanic@liu.se, piotr.rudol@liu.se, tommy.persson@liu.se,
mariusz.wzorek@liu.se, patrick.doherty@liu.se

Abstract

This paper describes a preliminary work towards a collaborative system involving humans and unmanned aerial vehicles to perform environment mapping using a qualitative description of 3D images. The images collected by a team of unmanned aerial vehicles flying over a region are pre-processed to reconstruct 3D images. The qualitative description of the 3D images aims to simplify the information delivered by the system to the users engaged in a team of search and rescue mission, enhancing their understanding and knowledge about the environment.

1 Introduction

The interaction between people and intelligent systems including robots, leads to a high demand of reliable representations of human knowledge. Remote-controlled robots are becoming a reality in the world, and its applications vary from delivery systems to complex surgeries. Specific aspects of interaction as cooperation and collaboration raises the ability of heterogeneous teams composed by humans and different kinds of robots, or with different resources, to solve difficult problems that rely on real time data analysis and quick responses.

Unmanned Aerial Vehicles (UAVs) are considered tools for emergency informatics [Murphy, 2016], which is a scientific field that approaches the use of different data-sets to save lives in natural or man-made disasters, through acquisition, organization and visualization of data and consequently which actions can be taken into account. Besides the use of UAVs for military proposes, there is potential for civil applications, as forest fire tracking, damage survey after earthquakes or tsunamis, monitoring of riots, search and rescue of missing people, and so on.

This paper presents an extension of the work developed in [Doherty *et al.*, 2016], where an integrated collaborative system (composed of humans and heterogeneous autonomous Unmanned Aerial Vehicles) pursues the goal of generating 3D models of an environment. The information contained in 3D images might not be readily interpreted by the final user, instead, a qualitative description that includes cognitive comprehensibility ensuring efficient understanding by humans

without wide computational complexity can yield better results for tasks such as search and rescue missions.

In this work a set of subfields of Qualitative Spatial Reasoning are joined to provide a qualitative description of a scene, observed by humans and robots, from different aerial points of view, in a way that all agents involved in the system have the possibility to understand where the objects are and the relations between them. The reasoning presented in Region Connection Calculus, Allen's Interval algebra and Interval Occlusion Calculus are used to provide the possible relations between a pair of objects seen from an aerial point of view, while the Cardinal Direction Calculus provides the basis for the orientation definition of those points of view.

Three new relations, *above*, *in the top of* and *in the bottom of* are presented to describe Allen's relations *precedes*, *starts* and *finishes* seen from a lateral point of view. Combinations of the relations are also developed to describe the positions of the objects when the projections of their shapes in axes x and y are not totally aligned, and cannot be described by a single relation from Allen's set of relations. Humans and robots will be able to exchange information and infer data about their partial view of the domain.

2 Qualitative Spatial Reasoning

Representing and reasoning about spatial knowledge can naturally be done through description of relations between two or more objects, specifying how spatial entities are related in space with others, particularly when numerical information is unavailable or even unnecessary for humans [Chen *et al.*, 2015].

In this research a 3D space is considered and the spatial entities are defined by a set of spatial points and lines. The set of relations is JEPD (Joint Exhaustive and Pairwise Disjoint), considering that in the set of spatial relations there is one and only one that can be satisfied. JEPD relations can be called basic relations for representing definite relationships between spatial entities. The union of basic information can express indefinite information. All base relations are possible in case no information is known [Chen *et al.*, 2015].

2.1 Region Connection Calculus

The basic relations between spatial regions are described in the theory of *Region Connection Calculus (RCC)* [Randell *et al.*, 1992], through the primitive *Connected* ($C(x,y)$), where

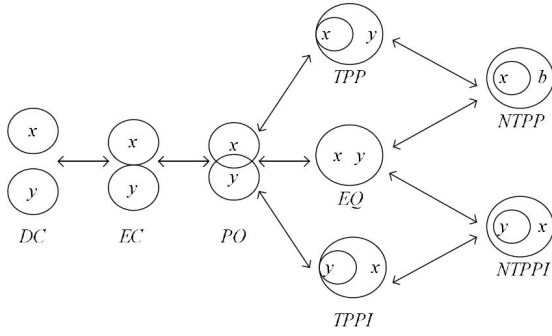


Figure 1: Conceptual neighborhood diagram [Randell *et al.*, 1992]; [Renz, 2002].

two non-empty regions of some topological space x and y are connected if and only if their topological closures share at least one common point.

Following the same theory, the constraint language RCC-8 contains eight JEPD base relations which allows reasoning about topological distinctions, with possibility to infer new spatial relations and transitions from incomplete spatial knowledge. Figure 1 shows the eight qualitative spatial relations covered by RCC-8 and its possible transitions illustrated by regions x and y on a conceptual neighborhood diagram [Randell *et al.*, 1992; Renz, 2002].

To define a reasoning for interpreting images collected by the UAVs, the RCC-8 theory is associated with Allen’s Interval Algebra, described in the next section.

2.2 Allen’s Interval Algebra

From the perspective of artificial intelligence, Allen’s Interval Algebra describes a temporal representation and reasoning where temporal interval is considered as a primitive. This method represents the relationships between pairs of reference intervals taking into account their upper and lower limits in a hierarchical manner, resulting in a set of 13 jointly-exhaustive and pairwise-disjoint base relations [Allen, 1990]. Figure 2 shows Allen’s relations considering two intervals, x and y .

Considering a constraint network pair (N, C) , N being the set of vertices where each domain element is represented by a vertex, and C representing the set of constraints defined by the basic Allen’s relations, it is possible to verify the existence of a consistent scenario by imposing algebraic closure on the network of constraints, to confirm whether the configurations provided by the information from the agents are feasible in at least one scenario of a domain described by the set of Allen’s relations [Santos *et al.*, 2015].

A network (N, C) is algebraically closed if its three vertices $(i, j, k) \in N^3$, and is consistent with the composition $C(i, j) \subseteq C(i, k) \circ C(k, j)$.

The Allen’s relations can be applied to agents’ viewpoints considering two rigid body convex entities, in an Euclidean plane. The *observers* (also referred as *viewpoint* Σ_i), are represented as pairs $\Sigma_i = (x_i, v_i)$, where x_i express the position of observer’s centroid and v_i is the unit vector representing observer’s orientation, and *objects*, identified by the 2D

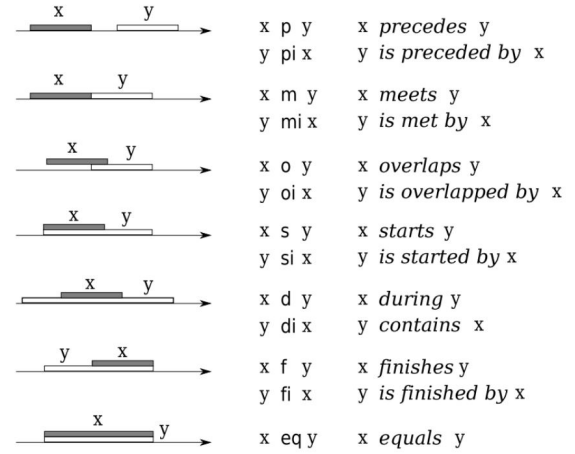


Figure 2: Allen’s Relations [Santos *et al.*, 2015], adapted from [Allen, 1990].

position of their centroids. The function $image(x, \Sigma)$ maps the projection of an object x perceived by an observer from a point of view Σ . In case more than one observer is present in the domain, each observer is perceived as an object when located within the field of view of another viewpoint [Santos *et al.*, 2015]. Furthermore, each observer is able to describe relations between pair of objects in its field of view.

An extension of Allen’s Interval Algebra called *Interval Occlusion Calculus* was developed proposing to reason about occlusion from multiple points of view, and is described in the next section.

2.3 Interval Occlusion Calculus

Interval Occlusion Calculus (IOC) is a qualitative description of a set of basic relations between pairs of objects observed from a point of view, given the object’s lines of sight [Santos *et al.*, 2015]. Considering that two objects A and B can be observed from a point of view Σ , the function $image$ defined by $a = image(A, \Sigma)$ and $b = image(B, \Sigma)$, maps the image of a physical body seen from a viewpoint Σ . Figure 3 illustrates an example of two bodies A and B , as well as a map based on the object’s lines of sight, enabling the observers to locate themselves with respect to the qualitative relations observed between the images of the objects. If Σ is located on the region of the map marked by p , then the observer will notice that a precedes b , or if Σ is located on the region o^+ , it could see that a overlaps and is in front of b . The same reasoning applies to all other positions of the map. Notations “region 1, 2, 3” and the red dashed region will be explained hereafter.

According to the distance between the observer and the objects (the depth), the image projected in a point of view Σ can be bigger or smaller. A layered interval is defined by $I = (I_a, l)$ where $I_a = (x_1, x_2)$, $x_1 < x_2$ are real numbers, x_1 is the lower limit of I , x_2 represents its upper limit and l is the layer of I . The function $ext(I)$ maps the extension (the upper and lower limits) of a layered interval I and the function $l(I)$ maps the proximity between the object and the observer (as closer observers and objects are, greater is the

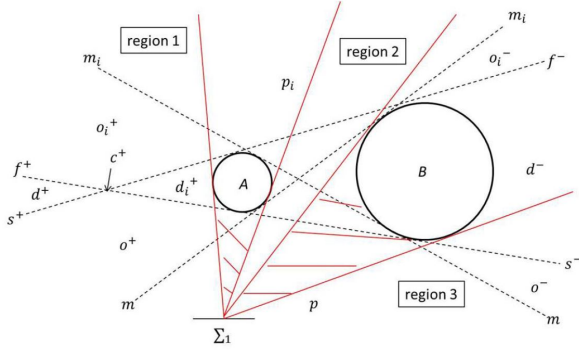


Figure 3: Basic Relations of Interval Occlusion Calculus [Santos *et al.*, 2015].

value of l). Finally, for any IOC relation r between two intervals I and J , the notation $I r^+ J$ applies iff $l(I) > l(J)$ and $I r^- J$ applies iff $l(I) < l(J)$ [Santos *et al.*, 2015].

It is possible to apply IOC relations for any pair of objects, including environments where multiple agents are able to share each other's spatial observations from different perspectives, interpreting and checking the consistency of the information received from other agents according to the translation table that can be found in [Santos *et al.*, 2015]. Let two distinct point of views Σ_1 and Σ_2 observe two objects A and B . The translation table gives the set of possible relations R considering exhaustively all the possible locations of Σ_2 from the point of view Σ_1 , taking into account that $a = \text{image}(A, \Sigma_1)$, $b = \text{image}(B, \Sigma_1)$, $a' = \text{image}(A, \Sigma_2)$, $b' = \text{image}(B, \Sigma_2)$, $\sigma_1 = \text{image}(\Sigma_1, \Sigma_2)$ and $\sigma_2 = \text{image}(\Sigma_2, \Sigma_1)$.

Reasoning about the environment showed in Figure 3, it is possible to interpret that if Σ_1 sees $\{a p b\}$, $\{\sigma_2 p_i a\}$, and $\{\sigma_2 p b\}$; then Σ_2 is in Region (2) of the map and its possible observations is the set of relations $\{a' \{p_i; p\} b'\}$. As another example, if Σ_2 is in the red-dashed region between A and Σ_1 , then Σ_1 observes that $\{\sigma_2 \{s^+, f^+, d^+\} a\}$, and the set of relations from Σ_2 would be $\{a' \{p, m, o^+\} b'\}$ [Santos *et al.*, 2015].

2.4 Cardinal Direction Calculus

A qualitative reasoning to deal with different perspectives perceived by multiple agents observing a scene is called Cardinal Direction Calculus.

The Cardinal Direction Calculus (CDC) is a formalism to reason about cardinal directions between objects and its relations. CDC is composed by 9 basic relations including a neutral region: *north* (n), *east* (e), *west* (w), *south* (s), *northwest* (nw), *northeast* (ne), *southeast* (se), *southwest* (sw) and *equal* (eq), that corresponds to the neutral region [Frank, 1996].

Through CDC relations it is possible to infer the direction of two objects I and J from the knowledge of the direction between I and K , and between K and J .

In [Frank, 1996] it is possible to find more information related to cardinal reasoning, as a composition table of projection-based directions, a study about cone-shaped area for which a symbolic direction is applicable and the quadrant-projection that includes the neutral area.

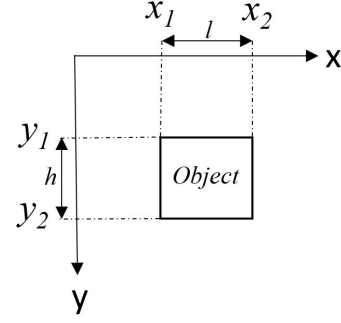


Figure 4: Object's projections from an aerial point of view.

The multiple point of views and relations mentioned in this chapter will be joined with the collaborative reasoning to be described in next sections in order to provide the appropriate constraints to interpret 3D images.

3 Collaborative Spatial Reasoning for Environment Mapping

Assuming that a system is composed by two or more agents and that one or more of those agents are humans, the mission to describe the position of objects present at a previously unknown environment might be impaired without the use of sentences that take into account the way people are used to describe this kind of information.

Considering an image composed by two objects I and J and observed from an aerial point of view Σ pointing to a direction v , the observed relations r are described in the sentence $I r J: \Sigma, v$ and based on Allen's Interval Algebra, Interval Occlusion Calculus (IOC) and Region Connection Calculus, described on Section 2.

The image of the objects seen from an aerial point of view is processed by projecting its shape on two axes, x and y , whose origin is located in the upper left corner of the image (Figure 4). Whilst in IOC a single layered interval is considered, an aerial view requires at least two intervals: l and h where $l = (x_1, x_2)$; $x_1 < x_2$ and $h = (y_1, y_2)$; $y_1 < y_2$. To reason about 3D images, an additional interval c where $c = (z_1, z_2)$; $z_1 < z_2$ is added with the projections of the object on axis z . The formalization of this theory and a deep study about its implementation can be seen in [Guesgen, 1989] and [Balbiani *et al.*, 2002].

The cardinal direction indicated by a compass in the case of robotic agents and by human sense of direction in case of human agents, determines the vector v to represent the orientation of a point of view Σ observing the objects in a scene. The aerial view can be made by a camera on board the drones or a person looking at an environment from the window of a building, a range of $\pm 22,5^\circ$ is considered to classify each cardinal direction, as showed on Figure 5. The logic to describe the observed relations for objects present in a scene requires the possibility to be transferred, to correctly describe relations from any point of view, for example, if object I precedes object J when Σ_1 is directed to *North*, the same scene seen from a Σ_2 directed to *South* will be described as object J precedes object I . The directions v considered are: *North*

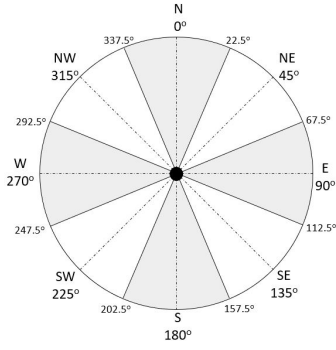


Figure 5: Cardinal directions.

(N) = 0° , North East (NE) = 45° , East (E) = 90° , South East (SE) = 135° , South (S) = 180° , South West (SW) = 225° , West (W) = 270° and North West (NW) = 315° .

The notation “Object₁ relation Object₂ : Σ , v ” is used to describe the possible relations between two objects observed from an aerial point of view Σ oriented to the direction v . Extending the single layered interval to a representation in two axes, more than one relation can apply to a pair of objects, resulting in the first set of reflexive relations described below, considering the arrangement of objects as shown in Figure 6. The relations written in capital letters refers to the RCC set of relations.

- $I p J : \Sigma, v$, read as “I precedes J from Σ oriented to v ” and defined by $l(I) p l(J) \cap l(I) DC l(J)$;
- $I \{p, m\} J : \Sigma, v$, read as “I precedes and meets J from Σ oriented to v ” and defined by $l(I) p l(J) \cap l(I) EC l(J)$;
- $I \{p, o\} J : \Sigma, v$, read as “I precedes and overlaps J from Σ oriented to v ” and defined by $l(I) p l(J) \cap l(I) O l(J)$;
- $I s J : \Sigma, v$, read as “I starts J from Σ oriented to v ” and defined by $l(I) s l(J) \cap h(I) TPP h(J)$;
- $I d J : \Sigma, v$, read as “I is during J from Σ oriented to v ” and defined by $l(I) d l(J) \cap h(I) NTPP h(J)$;
- $I f J : \Sigma, v$, read as “I finishes J from Σ oriented to v ” and defined by $l(I) f l(J) \cap h(I) TPP h(J)$;
- $I eq J : \Sigma, v$, read as “I is equal to J from Σ oriented to v ” and defined by $l(I) eq l(J) \cap h(I) eq h(J)$.

The same object’s positions showed in Figure 6 seen from a lateral point of view results in the second set of relations of Figure 7. The relations *precedes*, *starts* and *finishes* are represented by three new relations: *above*, *top* and *bottom*, defined below, to avoid misunderstandings regarding object’s positions when seen from different directions. The relation *during* is not described here because it keeps the same definition from the first set of relations.

- $I a J : \Sigma, v$, read as “I is above J from Σ oriented to v ” and defined by $h(I) a h(J) \cap h(I) DC h(J)$;
- $I \{a, m\} J : \Sigma, v$, read as “I is above and meets J from Σ oriented to v ” and defined by $h(I) a h(J) \cap h(I) EC h(J)$;

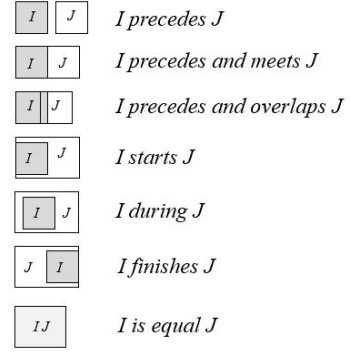


Figure 6: First set of relations and its definitions derived from Allen’s Relations Calculus and Region Connection Calculus.

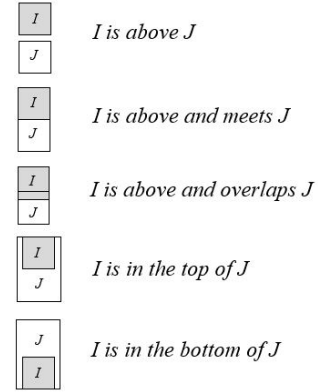


Figure 7: Second set of relations and its definitions derived from Allen’s Relations Calculus and Region Connection Calculus.

- $I \{a, o\} J : \Sigma, v$, read as “I is above and overlaps J from Σ oriented to v ” and defined by $h(I) a h(J) \cap h(I) O h(J)$;
- $I t J : \Sigma, v$, read as “I is in the top of J from Σ oriented to v ” and defined by $h(I) t h(J) \cap l(I) TPP l(J)$;
- $I b J : \Sigma, v$, read as “I is in the bottom of J from Σ oriented to v ” and defined by $h(I) b h(J) \cap l(I) TPP l(J)$.

A third set of relations was developed associating the first and second sets to express the relations for objects whose projections l and h are not totally aligned, and here the formalization is different from [Guesgen, 1989] because even working with intervals projected in two axes, it can express three possible relations, while in [Guesgen, 1989] it is possible to represent only two relations. Figure 8 shows the position of the objects described below.

- $I \{p, a\} J : \Sigma, v$, read as “I precedes and is above J from Σ oriented to v ” and defined by $l(I) p l(J) \cap h(I) a h(J) \cap l(I) DC l(J)$;
- $I \{p, m, a\} J : \Sigma, v$, read as “I precedes, meets and is above J from Σ oriented to v ” and defined by $l(I) p l(J) \cap l(I) EC l(J) \cap h(I) a h(J)$;
- $I \{p, o, a\} J : \Sigma, v$, read as “I precedes, overlaps and is

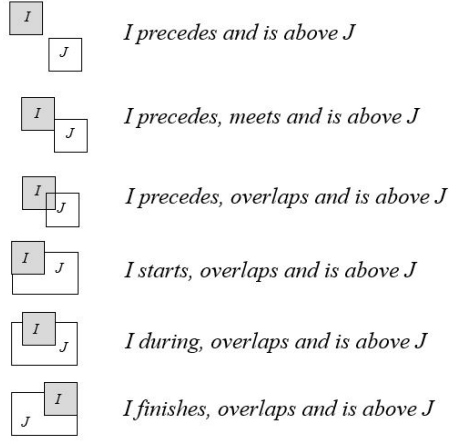


Figure 8: Third set of relations and its definitions derived from Allen's Relations Calculus and Region Connection Calculus.

- above J from Σ oriented to v " and defined by $l(I) p l(J) \cap l(I) PO l(J) \cap h(I) PO h(J) \cap h(I) a h(J)$;
- $I \{s, o, a\} J : \Sigma, v$, read as "I starts, overlaps and is above J from Σ oriented to v " and defined by $l(I) s l(J) \cap h(I) PO h(J) \cap l(I) PO l(J) \cap h(I) a h(J)$;
- $I \{d, o, a\} J : \Sigma, v$, read as "I is during, overlaps and is above J from Σ oriented to v " and defined by $l(I) d l(J) \cap h(I) PO h(J) \cap l(I) PO l(J) \cap h(I) a h(J)$;
- $I \{f, o, a\} J : \Sigma, v$, read as "I finishes, overlaps and is above J from Σ oriented to v " and defined by $l(I) f l(J) \cap h(I) PO h(J) \cup l(I) PO l(J) \cap h(I) a h(J)$.

To conclude, it is perceived that the basic relations *precedes*, *meets*, *overlaps*, *starts*, *during*, *finishes*, *equal* and the new ones *above*, *top* and *bottom* are enough to describe all possible configurations for a pair of objects.

Qualitative solutions for the robotic navigation problem was discussed in [Schlieder, 1993] and [Santos *et al.*, 2016], where reference points was encoded in the tessellation of a plane into regions to enable agents to navigate and localize themselves in an environment through the identification of landmarks. This approach takes the occlusion factor as an important role, making it difficult to implement the same reasoning for an aerial point of view, as the occlusion identified in most cases are in the vertical plane, for example, the leafs occluding a vehicle parked under a tree.

The effectiveness of the logic presented in this chapter will be tested using video image collected by drones in an experiment described in Section 5.

4 Collaborative Reasoning

Multi-agent systems are often supposed to have several advantages over single robot systems, as the capability to accomplish a single task faster, or efficiently exchange information about their position to precisely localize themselves whenever they sense each other [Burgard *et al.*, 2000]. When

multiple agents are observing a scene and occlusion prevents sensors from assessing parts of the objects present in the scene, distinct observers can provide multiple view-point descriptions about the objects they can detect, improving the completeness of the information observed [Santos *et al.*, 2015].

Collaborative systems are required to integrate not only robots, but also human resources. For complex emergency scenarios such as search and rescue missions, providing assistance and guiding people to a safe destination from inhabitants lost in wilderness regions, at sea scenarios, places devastated by earthquakes, flooding or forest fires [Doherty *et al.*, 2010] it is an essential role to provide an effective communication and continuous interaction between people and robotic agents in order to achieve mission goals, specially the ones related to environment exploration, so that each individual can explore different areas simultaneously [Burgard *et al.*, 2000].

5 Experiment

Three data collection flights were performed at Motala Flygklubb and two different UAV platforms were used.

The first one is a DJI Matrice 600 Pro¹ research platform and The second platform is a DJI Matrice 100².

Both platforms are equipped with Intel NUC computers using Core i7-7567U processors, 16 GB of memory, and 500 GB SSD storage. The DJI Zenmuse Z3³ cameras were used to collect video and images during the experimental flights.

The scanning patterns flown to collect the experimental data were automatically generated to cover a designated area. The platforms took off manually and after reaching a safe altitude flew autonomously over the region and performed the scan according to the pre-determinate altitude. The flights were performed at three different altitudes: 30, 50 and 80 m above the ground level (AGL).

The videos and photos collected register a partial view of the environment and were taken from different directions, to show different partial perspectives of each scene. The images contain many buildings with different sizes and colors, some cars, trees and a small airplane. The background contains areas covered by grass, asphalt and a road.

Each drone flew over a different portion of the area, but capturing images of some objects in common, in order to use those objects as references to infer the position of all other objects that are not in its field of view. The information can be easily exchanged by agents via wi-fi, including the ground operator. It is possible to transmit images, but a large broadband is necessary and it is not always available. The developed sentences in the format "*Object₁ {relations} Object₂ : Σ, v* " can be transmitted using only sparse data and thus cope with the bandwidth problem, independent of the number of axes used to generate the relations. The complexity will be related to the process of object identification, projections on axes and relations generation, and not to data transmission.

¹www.dji.com/matrice600/info

²www.dji.com/matrice100

³www.dji.com/zenmuse-z3



Figure 9: Image captured by drone DJI Matrice 100, flying in the East direction

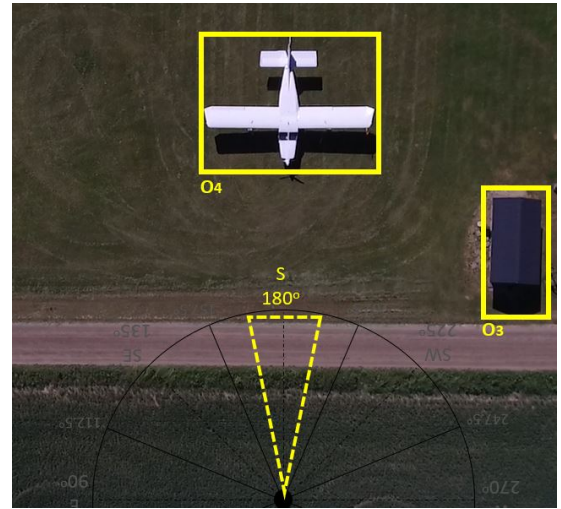


Figure 11: Image captured by drone DJI Matrice 600, flying in the South direction

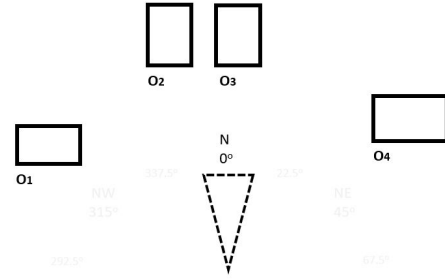


Figure 12: Sketch of a visual map representing object's positions

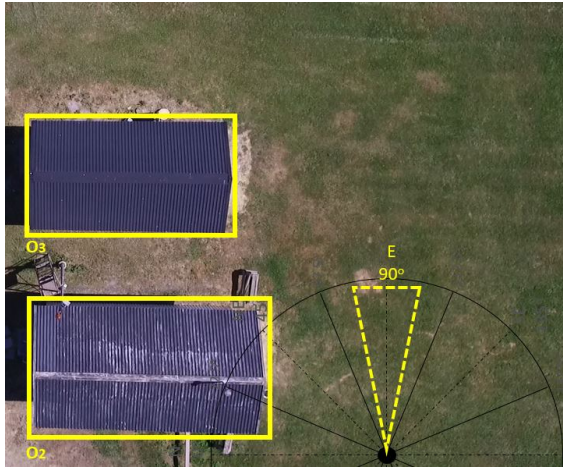


Figure 10: Image captured by drone DJI Matrice 100, flying in the East direction

Figures 9 and 10 show images captured by one drone flying in the East direction, at an altitude of 30 meters. As soon as the system recognize a pair of objects, each object is labeled and the relations between them are identified according to the set of restrictions of Section 3, and the information is transmitted to the other agents. The information contained in Figure 9 is transmitted in the format: $Object_2 \{p, a\} Object_1 : \Sigma_1, E$, while Figure 10 should transmit the information $Object_3 \{a\} Object_2 : \Sigma_1, E$.

The second agent produced the image seen in Figure 11 flying in the South direction, at an altitude of 50 meters. The information transmitted by this agent is: $Object_4 \{p, a\} Object_3 : \Sigma_2, S$. The airplane labeled as $Object_4$ is not identified as a square or rectangle in the image, but its extreme points can be projected in the axes enabling the system to fit it under a square format.

The information exchanged by the agents is supposed to produce a result similar to the Figure 12, where it is possible to see a sketch of a map with the objects translated to the North direction, to complement the information already received in the format of sentences. The Figure 13 shows the environment used in this example. The image was captured by DJI Matrice 600, flying to North direction, at 80 meters of altitude.



Figure 13: Real environment

Currently, those images and videos are under analyses to evaluate which computer vision techniques and filters will be applied to identify and label the objects present in each scene or frame for posterior implementation and test of the relations of Collaborative Spatial Reasoning.

6 Next Steps

After implementation, test and correction of any inconsistencies found in the Collaborative Spatial Reasoning for Environment Mapping using videos and 2D photographs captured by drones, the same reasoning will be applied to 3D images, using an additional axis projection. The implementation using 2D and 3D data will be compared to evaluate pros and cons of adding complexity to the system. Tuples of possible relations will be studied to avoid an explosion of number of relational combinations.

The laser data collected in the experiment will be used to improve the quality of the input data, as it contains the three-dimensional shape of the objects and its GPS position. Point clouds will also be used to create a 3D version of the scene. The delegation framework [Doherty *et al.*, 2013] is planned to be used in the future, in order to get collaboration between agents.

It is expected that the integration of qualitative and quantitative data, as well as robot's and human's abilities increase the robustness of the system, enabling all involved in a mission to describe and understand an environment to reach their goals.

Acknowledgments

This work was carried out with the support of CISB, Swedish-Brazilian Research and Innovation Center, and Saab AB. The author also would like to thank the scholarship provided by CAPES, and the staff of Autonomous Unmanned Aircraft Systems Technologies Lab from Linköping University for sharing the experiments data. Paulo Santos acknowledges support from IBM-FAPESP Proc. 2016/18792-9.

References

- [Allen, 1990] James F Allen. Maintaining knowledge about temporal intervals. In *Readings in qualitative reasoning about physical systems*, pages 361–372. Elsevier, 1990.
- [Balbiani *et al.*, 2002] Philippe Balbiani, Jean-François Condotta, and Luis Fariñas Del Cerro. Tractability results

in the block algebra. *Journal of Logic and Computation*, 12(5):885–909, 2002.

- [Burgard *et al.*, 2000] Wolfram Burgard, Mark Moors, Dieter Fox, Reid Simmons, and Sebastian Thrun. Collaborative multi-robot exploration. In *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on*, volume 1, pages 476–481. IEEE, 2000.
- [Chen *et al.*, 2015] Juan Chen, Anthony G Cohn, Dayou Liu, Shengsheng Wang, Jihong Ouyang, and Qiangyuan Yu. A survey of qualitative spatial representations. *The Knowledge Engineering Review*, 30(1):106–136, 2015.
- [Doherty *et al.*, 2010] Patrick Doherty, Jonas Kvarnström, Fredrik Heintz, David Landén, and P-M Olsson. Research with collaborative unmanned aircraft systems. In *Dagstuhl Seminar Proceedings*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2010.
- [Doherty *et al.*, 2013] Patrick Doherty, Fredrik Heintz, and Jonas Kvarnström. High-level mission specification and planning for collaborative unmanned aircraft systems using delegation. *Unmanned Systems*, 1(01):75–119, 2013.
- [Doherty *et al.*, 2016] Patrick Doherty, Jonas Kvarnström, Piotr Rudol, Marius Wzorek, Gianpaolo Conte, Cyrille Berger, Timo Hinzmann, and Thomas Stastny. A collaborative framework for 3d mapping using unmanned aerial vehicles. In *PRIMA 2016: Principles and Practice of Multi-Agent Systems*, volume 9862, pages 110–130, Cham, 2016. Springer International Publishing.
- [Frank, 1996] Andrew U Frank. Qualitative spatial reasoning: Cardinal directions as an example. *International Journal of Geographical Information Science*, 10(3):269–290, 1996.
- [Guesgen, 1989] Hans Werner Guesgen. *Spatial reasoning based on Allen's temporal logic*. International Computer Science Institute Berkeley, 1989.
- [Murphy, 2016] Robin R Murphy. Emergency informatics: using computing to improve disaster management. *Computer*, 49(5):19–27, 2016.
- [Randell *et al.*, 1992] David A Randell, Zhan Cui, and Anthony G Cohn. A spatial logic based on regions and connection. *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR 1992)*, 92:165–176, 1992.
- [Renz, 2002] Jochen Renz. *Qualitative spatial reasoning with topological information*. Springer-Verlag, 2002.
- [Santos *et al.*, 2015] Paulo E Santos, Gérard Ligozat, and Marjan Safi-Samghabad. An occlusion calculus based on an interval algebra. In *Intelligent Systems (BRACIS), 2015 Brazilian Conference on*, pages 128–133. IEEE, 2015.
- [Santos *et al.*, 2016] Paulo E Santos, Murilo F Martins, Valquiria Fenelon, Fabio G Cozman, and Hannah M Dee. Probabilistic self-localisation on a qualitative map based on occlusions. *Journal of Experimental & Theoretical Artificial Intelligence*, 28(5):781–799, 2016.

[Schlieder, 1993] Christoph Schlieder. Representing visible locations for qualitative navigation. *N. Piera Carrete & M. G. Singh (Eds.) Qualitative Reasoning and Decision Technologies*, 1993.