International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-10 Issue-1, November 2020

A Bit on the Right, A Bit on the Left: Towards Logical Bundle Adjustment

Mohamed BEN ELLEFI, Pierre DRAP

Abstract*: This article proposes an innovative approach fully based on logic to determine the relative positions and orientations of objects in a scene photographed from different points of view as well as those of the cameras used to take the pictures. The proposal is absolutely not based on 2D feature extraction, projective geometry or least squares adjustment but on a logical approach based on an enumeration of simple relationships between the objects visible in the photos. It is an approach imitating a natural and unconscious reasoning that each of us makes by observing a scene: is this object more to the right than this one? And is this other one further away from me than the one who's partially hiding it from me? It is therefore a question of approaching the problem by identifying and recognizing objects in photographs and not by measuring millions of points in space without having any idea of the object to which they belong. This article presents a "proof of concept" based on virtual experimentation: in a discrete 3D space, a simple scene, composed of spheres of different colors and cameras, is modelled in a 3D format. In this work the positioning of the spheres and cameras is limited to a plane. Cameras are placed in the scene in order to see the spheres and then for each camera an image is generated. The application reads each image and deducts relationships between object and camera. These relationships based on the visible occlusions between the projections of the objects onto the photographs, are formalized according to Allen's relationships. A knowledge base is implemented to allow an iterative process of SPARQL queries for qualitative spatial reasoning leading to a set of possible solutions. Finally, the system deduces the relative positions between objects and cameras and the result is imported and can be used within several photogrammetry software suites.

Keywords: Photography, Photogrammetry, Perspective, 3D, interval algebra, Knowledge Base, Ontology, SPARQL.

I. INTRODUCTION

The transition to the Digital Age has changed many technologies in recent decades. Photography, photogrammetry and computer vision have also been enormously impacted. Towards the end of the 20th century, photogrammetry software suites, reserved for professionals, cost several tens of thousands of euros. Now there are free online services, for example Arc3D[1] [1], [2], very good open-source software (MicMac[2], COLMAP[3] [3], [4]) as well as very efficient commercial software for a tenth of the price. Since the first years of this century, the confluence of photogrammetry and computer vision has given rise to an almost mature discipline. In a way, the construction of a 3D facsimile of a 3D scene by photogrammetry is no longer a matter of research Emerging research topics currently focus either on real-time surveys or mainly on the control of a survey's semantic component. The semantisation of surveys is now a crucial problem for all aspects of surveying, robotics, and satellite imagery, for all disciplines including biology, geology, heritage, archeology, tourism, oil industry, and civil engineering. There is thus a significant intrusion of symbolic and connexional Artificial Intelligence in the survey process, both upstream and downstream of the photogrammetry process. It is either a question of producing new knowledge and reasoning techniques with or identifying and segmenting the tens of millions of 3D points that photogrammetry software now provides.

A. Connectionism and Neural Networks. The classification and subsequent use of 2D and 3D recognition methods is developing intensely. The use of deep learning appears in optical measuring systems in the field of research (see [5], [6], [7], [8]) as well as in the industrial field (i.e., the automatic visual inspection system for automotive vehicles proposed by the company Tchek https:// www.tchek.fr/). Nevertheless the deep learning approach, mainly used for 2D images and very rarely directly on 3D models, only helps in segmenting 3D models (by pattern recognition on oriented images for example) and not in managing knowledge. A very interesting application is developed by Mrs. Gaia Pavoni in her thesis for the recognition and segmentation of coral colonies. It is a very promising approach where deep learning is combined with underwater photogrammetry [9]. A good overview of the massive use of neural networks in photogrammetry can be read in the work of Lei Ma and Brian Alan Johnson [10].

B. Symbolic Artificial Intelligence, the use of ontology. The need for semantics is more and more pressing in the production of surveys. For example, in the context of cultural heritage, the need to effectively share information in a restoration process requires the use of a common language among key players and stakeholders. The same is also the case in many areas where survey production is a prerequisite for the work of a specialist.

Revised Manuscript Received on November 30, 2020.
***** Correspondence Author
Mohamed BEN ELLEFI *, Aix Marseille University, University of Toulon, CNRS, UMR 7020, 13397 Marseille, France name.surname@univ-amu.fr
Pierre Drap, Aix Marseille University, University of Toulon, CNRS, UMR 7020, 13397 Marseille, France. Pierre.Drap@univ-amu.fr

[1] https://www.arc3d.be/

[2] MicMac from IGN, France https://micmac.ensg.eu/index.php/Accueil
[3] https://colmap.github.io/index.html

The representation of 3D models should follow certain existing standard data models and be able to communicate their geometric characteristics usefully to improve the knowledge of the survey's end users. Corpuses using the formalism of ontologies have already been developed and are now followed by many teams. [11], [12]. (INSPIRE[13], [14], CityGML [15], ImageProcessingOntology [16], CIDOC-CRM [17],[18], its extension MONDIS [19] and the vocabularies of the Getty Institute). If the graphical / semantic link is obvious in cartography for small-scale surveys, it is not the same for large-scale 3D surveys where 3D models are complex, visualization formats are not standardized and the tools to acquire them are emerging. This need for semantics is increasingly pressing, also linked to the fact that the production of digital images and 3D surveys has been growing exponentially for more than ten years. (see [20], [21], [22], [23], [24]). It is no longer possible to be satisfied with meaningless 3D models. The approach proposed here is fully based on knowledge, on the recognition of the objects photographed.

In another area, the use of description logics (DLs) as a knowledge representation and reasoning system addressed the problem of scene description which represents the relative positioning of objects and the points of view observing these scenes (see [25], [26]). Finally, the semantic relationships between objects in the same scene are more and more studied with the formalism of ontologies in order to use inference capabilities to validate or detect inconsistencies in the observed scenes as stated in [27], [28], [29]. This paper presents the outline of a new surveying approach based entirely on logic and the description of the relative positions of objects in photographs representing them. The proposed system produces an approximate solution to determining the position and orientation of the photographed objects and the cameras that took the pictures without any 2D point extraction or 3D point computation, and without using the well-known collinearity equations [30]. To synthesize, the problem addressed is the following: given a series of photographs of a scene, how to deduce the relative positions of the objects and cameras in 3D space by observing only the relative positions of the objects in the photographs without making matrix calculations?

This article describes a "proof of concept" based on virtual experimentation: In a discrete 3D space, a simple scene, composed of spheres of different colors and cameras, is modeled in a 3D format. The positioning of the spheres and cameras is restricted to a single plane. Eight cardinal relations are defined to link the objects with each other as well as the objects with the cameras (North, South, East, West, North-East, etc.). A Cardinal Direction Relations (CDR) model as defined in [31], [32], [33] The cameras are placed in the scene in order to see the spheres, then an image is produced by each camera. The application reads each image and infers the cardinal relationships between the visible spheres and expresses them in the camera reference system. These relationships are detectable in photographs because they do not consider the distance between the objects but only the angle (the heading) between the two objects they connect. The relations, based on the visible occlusions between the projections of the spheres on the photographs, are formalized according to Allen's relations.

From a logical point of view the key contribution of this paper is the conceptualization of a knowledge base for qualitative spatial reasoning over a 3D scene. Qualitative Spatial Reasoning (QSR) is concerned with providing cardinal relations allowing to represent and reason with spatial objects without resort to traditional quantitative techniques [34]. Building such knowledge base starts by populating the ontology where objects and cameras are instances of photogrammetry concepts extracted from the Arpenteur ontology. The next step is determining the orientation of the cameras and objects in the reference system. Their orientation is determined thanks to an iterative consistency checking process using SPARQL [35] queries to answer all the CDR constraints in the populated ontology. Furthermore, during this process, a positioning SPARQL query is formed in order to perform pattern matching between the CDR graph pattern and the plane grid pattern (CDR graph of tiles). An interactive demonstration of the resulted 3D scene is visible online via http://www.arpenteur.org/lba/3D/. This article is structured as follows: Section II traces historically the links between space measurement, optics, geometry and algebraic computation, with photogrammetry inheriting all this. Section III presents our vision of new logical bundle adjustment, combining interval algebra and cardinal relations. Section III.Bformalizes the cardinal direction relations for qualitative spatial reasoning. Section IV discusses CDR Knowledge Base, highlighting the uses of the photogrammetry ontology in our system. Section V describes precisely the constraints that limit this work as well as the *modus operandis*. Lastly, Section VI analyses our results before presenting certain conclusions and perspectives for further development.

II. HISTORICAL LINK BETWEEN PHOTOGRAMMETRY AND INTENSIVE COMPUTATION

Space measurement and optics are closely related. Space measurement must be designed, modelled, and represented in a systematic and reproducible process. It is also essential to have sophisticated mathematical tools: algebra, trigonometry and statistics.

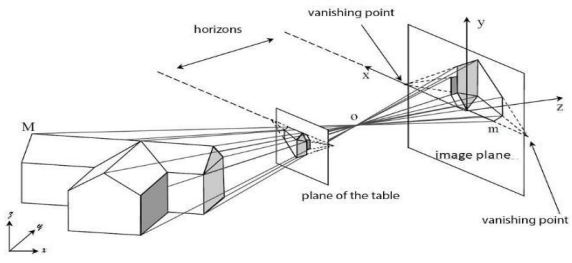


Fig. 1Perspective image formation.

The conceptual upheavals of the Quattrocento and the following century introduced a new vision of the world and are a key step in the construction of the knowledge necessary for photogrammetry. The models in place at that time for more than fourteen centuries have collapsed and have gradually been replaced by a set of homogeneous models that have paved the way for a rationalization of space and made optics a physics.

A. Photogrammetry: Perspective, geometry and algebra

Brunelleschi seems to be the first, in around 1410, to have realized a representation of a three-dimensional object on a flat (2D) surface using a central projection. Since then, the term perspective has been used to describe the type of representation that gives an object an image similar to the vision we would have of that object. "The perspective image and the natural vision of an object coincide, giving rise to the same retinal image provided that there is monocular vision, with the eye immobile, and at the correct distance." [36]. Perspective is an interdisciplinary technique at the crossroads of architecture, science and art. This type of representation of space became widespread in Italy in the 15th century. The problem solved by Brunelleschi is a problem of rigor in representation; he finds a systematic process to represent the apparent decrease of an object's size according to its relative position to the observer. The constructive systematism that he proposes, even if it is based on a global understanding of the phenomenon (intersection of the visual pyramid with a plane), is only in the plane of representation (plane of the table in Fig. 1). It is therefore able to correctly represent the object on the plane of the table without using the projective system illustrated in Fig. 1. It was only in the 17th century that the use of the camera obscura by painters became widespread, then in the 19th century and mathematicians Monge and Poncelet to see the relationship between the point M observed and expressed in the ground reference system and the coordinates of its image m expressed in the image reference system.

This formalism is used in the photogrammetric approach. The perspective image obtained on the plane of the image by central projection is linked to the observed object by the relationship:

$$X - X0 = \lambda R (X' - Xh') \quad (1)$$

Where we have:

X coordinates of the object point, X0 coordinates of the projection center O, λ Scale factor (specific to the observed points), R orthogonal rotation matrix, X' coordinates of the image point (in the image reference system), Xh' coordinates of the principal point (in the image reference system). (The principal point is the projection of the optical center on the image plane.) Projective geometry then becomes, with Monge and Poncelet, an algebraization of geometry making the models readable by algebraic equations. Finally, photogrammetry is part of the family of means of representing nature and inherits centuries of trial and error in this direction. It was in 1852 that Captain Laussedat, already a user of the 'camera obscura' for his surveys (Château de Vincennes, 1850), replaced the hand-drawn perspectives by photographs and named the process "Metrophotography". In Germany, the architect Albrecht Meydenbauer used photography for renovation work on Wetzlar Cathedral as early as 1858. In 1893 he introduced the term "Photogrammetry", which has since been used in all languages.

The American Society of Photogrammetry defines photogrammetry as including remote sensing and image recording modes ranging from video to radar images to ultraviolet or infrared photographs; photogrammetry is "The art, science and technology of obtaining reliable information about physical objects and the environment through processes of recording, measuring and interpreting photographic images and pattern of electro-magnetic radiant energy and other phenomena" [37].In addition to being based on the perspective and ability to obtain a 2D image of the 3D world, photogrammetry uses the principles of projective geometry and the equations that govern it. Immersed in a context of metrology, and therefore of measurement, uncertainty and accuracy, photogrammetry uses the least squares approach intensively. It can therefore be said that since its beginnings photogrammetry has developed on the basis of intensive computing, for example Kruppa published in 1913 a 10-page article dedicated to the calculation of the relative orientation between two photographs [38].

And since then, photogrammetry, and thus the calculation of the positions and orientations of objects in space, has developed and intensified with the technological development of computing tools. In parallel with the development of photogrammetry, very quickly, i.e. around 1930, logicians became interested in the modelling of spatial relations, notably Tarski published two articles (1931 and 1941) on this subject. Tarski quotes Sanders Peirce and some of his articles published between 1870 and 1890 and defines him as the precursor of this approach [39], [40]. The work presented here is modestly located at the confluence of these works and proposes a method capable of elaborating, by logic, a solution approximated to a scene which could then be calculated finely by photogrammetry.

B. Two paradigm shifts in the 20th century

The 20th century has seen two fundamental upheavals in photogrammetry; two new ideas that have made possible the current level of maturation of photogrammetry: closer to us the appearance of 2D point descriptors in the search for homologous points and, more fundamentally, the introduction of beam adjustment allowing a simultaneous and global calculation of all the unknowns involved in the calculation of the scene.

1) Homologous point

The introduction of automatic detection of homologous points on a set of photographs was a revolution.

The appearance of algorithms such as SIFT [41] then SURF 42 and now many others have made possible the development of photogrammetry as we know it today (see [43], [44]).

Indeed, it is an important conceptual break, a paradigm shift of the problem that was blocking: how to find two homologous points on two photographs that are not at the same scale and not oriented in the same way?

The original idea was to 'recognize' a point on one image after having identified it on another, the paradigm shift that allowed us to free ourselves from this constraint was the introduction of a descriptor for each point. Thus, we no longer compare 'points' on an image, i.e. pixel matrices, but descriptors of remarkable points. In one decade, it has been possible to obtain automatically and with few false positives, a very important number of homologous points on two sufficiently textured images. This was a real technological breakthrough that could not have been predicted ten years earlier.

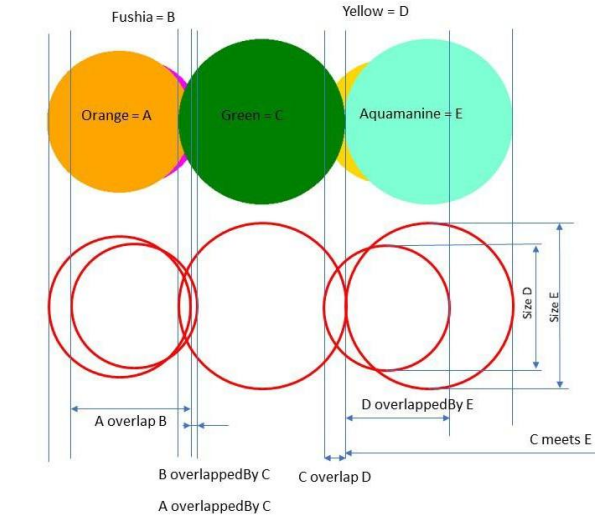
2) Bundle adjustment

Around 1957, Duane Brown (August 20, 1929 - July 30, 1994) developed new methods of camera calibration and a mathematical formulation of bundle adjustment [45]. (quoted in [46]). This work marked a radical change in the methods of photogrammetric computation. Bundle adjustment makes possible to estimate at the same time the internal parameters of the camera, its external parameters as well as the coordinates of the 3D points of the scene. Previously, the external orientation parameters were processed independently, and then the ground points were determined by triangulation. However, the implementation of simultaneous calculation of all these parameters involved the resolution of large matrix systems and the implementation and dissemination to potential users of these algorithms had to wait until the advent of the computer, i.e. the 1980s. Moreover, the confluence of the disciplines of photogrammetry and computer vision in the early 2000s also contributed greatly to the development of this method. For a comprehensive and accurate overview of bundle adjustment, see [47], [48], [49]. Bundle adjustment has developed along with technology: there are implementations for GPUs 50, intensive parallel bundle adjustments [51],[52], and of course the joint use of inertial fusion power plants, GNSS and SLAM applications [53],[54]. Bundle adjustment, as we have just seen, is a decidedly computational technique based on matrix calculation. This approach is now the keystone of photogrammetric orientation processes, and all software offers a declination of it. While the approach proposed in this paper is resolutely diverted from matrix computation and can in no way claim any relationship with a bundle adjustment, we have nevertheless opted for a Logical Bundle Adjustment in the sense that we propose a logical approach, without optimization, in a discrete space but offering, under some conditions, an approximate solution to the following problem: "Given an unordered set of images with known calibration data, we want to estimate the structure of the scene as well as the camera positions and orientations".

III. From Interval Algebra to Cardinal Direction Relations

A. Using interval algebra

This paper presents an approach for approximate modeling of a scene, composed of known objects, by analyzing a series of images of this scene. The objects are spheres of identical radius, identifiable by their unique color. This analysis is based on the study of occlusions and relative positions of the images of the objects visible in the pictures. The analysis of these occlusions is based on the Allen interval algebra dedicated to spatiotemporal reasoning created by James F. Allen in 1983 [55].Fig. 3 describes the 13 Allen relationships between objects.



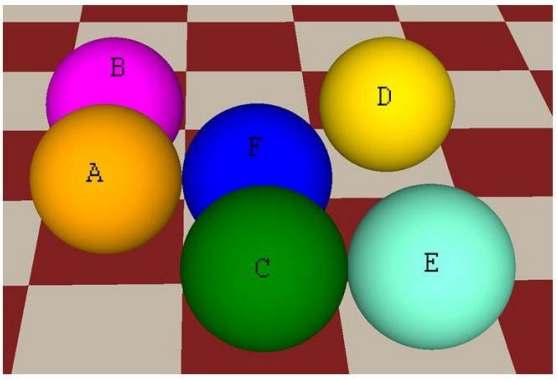


Fig. 2. **Photograph from camera 7 with computed overlaps and corresponding 3D scene.**

This approach imitates the natural and unconscious reasoning that we all do when observing a scene: an object hiding another is closer to us, the partially hidden object is further away. Otherwise, if there are two identical objects and the apparent diameter of one of them is smaller, then it is further away from me than the other. Our perception of the environment, beyond the perception of depth that we owe to our binocular vision, is due to a real-time analysis of the relative positions of known objects in the scene we are observing. To understand the scene, we use the relative positions of the objects, the partial occlusions of the objects in relation to each other and finally the relative apparent sizes of the objects we know. In fact, we understand scenes when they are composed of objects that we recognize and of which we know approximately their size. The a priori knowledge that we have of the objects composing a scene is a key element of its understanding and therefore of its measurability.

This aspect is an open problem in the field of robotics, with some emerging solutions (see the recent work of Peter R. Florence [56]).

However, we use the formalism proposed by Allen in a rather particular way: A relationship is established between the images of the spheres projected on the camera's image plane. The visible occlusions are treated as Allen intervals taking into account the respective sizes of the projected spheres. In the current version, the spheres and the cameras are distributed on the same plane. The studied intervals are therefore distributed on a straight line parallel to the OX axis of the cameras (see Fig. 2 and 9). In a future version, the intervals will be studied in 2D on the photograph and the objects will be distributed in a 3D space.

The key point of this work is the passage between the observations made on the photos and formalized with Allen's algebra doubled with metric considerations and the position of objects and cameras in the plane. These positions are formalized by their cardinal relationships. This approach has already been used successfully for a long time, for example by G. Libozat [57].

The calculation of cardinal points is a spatial formalism where the relationships express the relative positions of points in the plane with respect to north, east, west, and south [57]. These relationships are described in Section B.

We have already applied these cardinal relationships to represent visibility and a notion of neighborhood in photogrammetric surveys of the Xlendi wreck in Malta [58]. In this paper we show how a purely qualitative and formalized approach with cardinal relations can allow us to deduce neighborhood relations that are traditionally determined using an Euclidean approach.

From an operational point of view, the first step in the approach proposed here is an image processing applied to the generated photographs. Indeed, it is a question of reading a photograph generated from one of the cameras of the scene and to deduce the CDR between the objects and the camera. Since the camera is located in the same plane as the objects, the first analysis will provide Allen's relations between the spheres visible in the photograph (see Fig. 2). It should be noted that all the spheres have the same diameter, which is also the size of the tiles of the checkerboard grid plane. We can see very clearly in Fig. 2, showing the view from camera 7, that the overlap relationship between the discs, images of the spheres, present in the scene actually translates an in-depth relationship between these same spheres. For example if we consider spheres E and D which are represented in this figure by discs E and D. Disc D is covered by disc E, which induces that sphere D is located north of sphere E, in the reference system of camera 7.

Another example is that of spheres C and E linked by the Allen relation "meet" and by the analysis of the diameters of the two discs C and E. These two discs being of almost equal size the deduced CDR are of the East/West type. Sphere E is east of sphere C (from the point of view of camera 7). The image processing part is first of all identifying the perimeters, even partial, of each disk. Then calculate the least-squares circle corresponding to each of them. The Allen relations is evaluated by calculating the area of intersection of the discs, the relative positions in the picture of the centers of the spheres and the ratio of proportion between the different radii of the optimized circles. Indeed the relations used to deduce the CDR in the plane between the spheres are based on the algebra of the Allen intervals but also on metric considerations, for example the "meets" relation between two discs can be translated by an East/West relation for the spheres but also by a North-East/South-West relation according to the relative sizes of the two diameters of the discs (one sphere being further away from the camera than the other). This will formally result in the introduction of a threshold Γ in the formalism presented in Section B. This threshold is used when comparing disk sizes, images of the spheres by the cameras. It considers the differences in size due to possible differences in camera distance.

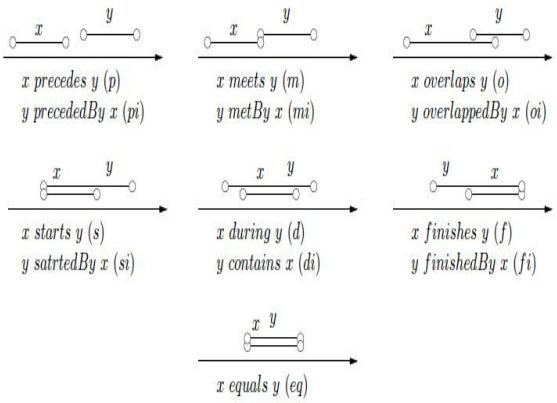


Fig. 3. Allen's Relations.

B. Cardinal Direction Relations for Qualitative Spatial Reasoning

Qualitative Spatial Representation and Reasoning (QSR) is a subfield of knowledge representation in Artificial Intelligence whose aim is to develop qualitative representations (and reasoning methods) involving spatial aspects of reasoning, such as spatial regions, directions, topology, location, proximity, geometry and intersection among others [59], [60], [61], [62], [63]. QSR models avoid the high computational cost of managing all quantitative information that can be collected from space; instead, they identify qualitative spatial relations/properties that are important for a particular problem. These relationships are generally modelled as disjointed but continuous, so that they can identify significant changes occurring in space. QSR models can deal with imprecise and incomplete data on a symbolic level since qualitative labels (i.e. close, far, in, touching) include already a margin for uncertainty and can be defined even if part of the numerical data is not known.

Moreover, QSR models help in human-machine interaction because they align human cognitive concepts with the numerical perception of computational systems. Another advantage of a description based on qualitative relations is that semantics can be assigned to them by means of logics and ontologies [64]. Most previous works assume only the representation and inferences related to the spatial perception of a single agent, leaving aside issues related to reasoning about multiple viewpoints in a scene.

In order to cope with this issue, [65] reports the development of a multiple-view spatial reasoning system called Interval Occlusion Calculus (IOC), based on Allen's Interval Algebra [66].

This formalism explicitly represents various distinct viewpoints whereby the notions of object occlusion. In other words, IOC extends Allen by defining the notion of "layered interval", and the ordering of the intervals with respect to represents their proximity to the observer (the closer an object is to the observer, the greater is its associated layer).

Hence, what we adopt from previous works [59], [65] is how to represent objects in a picture with Allen's intervals. Next, we will extend Allen's relations to the Euclidean plane.

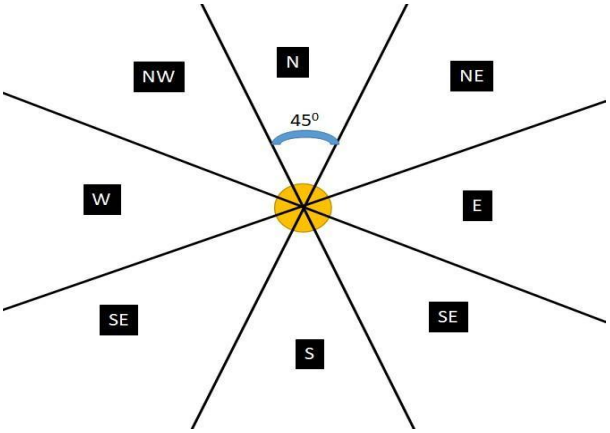


Fig. 4. The 8 cardinal areas for the orange sphere. Each area is represented by an angle of 45◦.

In this paper we formalize the CDR for Qualitative Spatial Reasoning in term of the cardinal and intercardinal directions of an 8-wind compass rose, where the plane is divided into 8 cone-shaped directions such that each direction refers to a region of 45 degrees (see Fig. 4). The set of directional symbols for this system is V 8 = {N,NE,E,SE,S,SW,W,NW} where N for North, NE for North East, E for East, SE for South East, S for South, SW for South West, W for West and NW for North West. Note here that, [67] identified two examples of systems of cardinal directions that can be implemented, both using the same set of eight directional symbols: the first one uses cone-shaped directions, which we adopt in our CDR system. The second one is based on projections, where the four main directions (N, E, S, W) are interpreted as half-lines, while the intermediate regions (NE, SE, SW, NW) refer to quadrants.

In the scene we assume that the grid is plane. This plane represents the 8 cardinal relations between objects (cameras and spheres). The reference system of each camera is oriented by its optical axis (defining the North) and the 'up' axis normal to the plane of the grid: i.e, all objects are in the north of the camera in this reference system. For instance, assuming that in a camera ci with i ∈ [0,n] where n is the number of the camera, we have objects presents in a picture pi oriented with a referential r_i. Let A, B be two distinct objects having the same size. A and B are represented respectively by their images Ai, Bi in the picture pi. Now we formally define when r_i satisfies the above spatial relationships. If rel is any of the above relationship symbols, we write rel_i(A,B) to denote the relationship B rel A is satisfied with the referential r_i. CDR are deduced from the Allen relations to which metric conditions are added to take into account the differences in size of objects according to their distance from the camera. Two metric thresholds are introduced: the first, Γ, on the difference between the apparent sizes of the objects and the second, β, on the percentage of overlap when there is occlusion. Respectively two functions have been defined delta size(A_i,B_i) and delta overlaps(A_i,B_i). The first one to evaluate the difference in radius of two visible discs on the same image and the second one the ratio between the hidden and visible areas in case of partial occlusion. A computation of the area of the intersection of two disks is performed during the image analysis in order to experimentally evaluate the value of these thresholds as a function of the image resolution, the shooting distance and the intrinsic parameters of the cameras (see Section A). In the scenario of Fig. 2, the thresholds that perform best are β = 24 and Γ = 14. Here are the rules used to deduce the cardinal relationships from Allen's relationships:

- if A_i precedes B_i :
  - if $delta\_size(A_i, B_i) \simeq 0.0$ then $E_i(A, B)$ can be read as "B is in the **east** of A in r_i".
  - if $delta\_size(A_i, B_i) > \Gamma$ then $SE_i(A, B)$ can be read as "B is in the **south east** of A in r_i".
  - if $delta\_size(A_i, B_i) < \Gamma$ then $NE_i(A, B)$ can be read as "B is in the **north east** of A in r_i".
- if A_i is preceded by B_i :
  - if $delta\_size(B_i, A_i) \simeq 0.0$ then $E_i(B, A)$ can be read as "B is in the **east** of A in r_i".
  - if $delta\_size(B_i, A_i) > \Gamma$ then $SE_i(B, A)$ can be read as "A is in the **south east** of B in r_i".
  - if $delta\_size(B_i, A_i) < \Gamma$ then $NE_i(B, A)$ can be read as "A is in the **north east** of B in r_i".
- if A_i meets B_i :
  - if $delta\_size(A_i, B_i) \simeq 0.0$ then $E_i(A, B)$ can be read as "B is in the **east** of A in r_i".
  - if $delta\_size(A_i, B_i) > \Gamma$ then $SE_i(A, B)$ can be read as "B is in the **south east** of A in r_i".
  - if $delta\_size(A_i, B_i) < \Gamma$ then $NE_i(A, B)$ can be read as "B is in the **north east** of A in r_i".
- if A_i overlaps B_i :

  - if $delta\_overlaps(A_i, B_i) \leq \beta$ then $NE_i(A, B)$ can be read as "B is in the **north east** of A in r_i"
  - if $delta\_overlaps(A_i, B_i) > \beta$ then $N_i(A, B)$ can be read as "B is in the **north** of A in r_i"

- if A_i is overlapped by B_i :
  - if $delta\_overlaps(B_i, A_i) \leq \beta$ then $SE_i(A, B)$ can be read as "B is in the **south east** of A in r_i"
  - if $delta\_overlaps(B_i, A_i) > \beta$ then $S_i(A, B)$ can be read as "B is in the **south** of A in r_i"

Here, in this step we extend Allen's relations to the CDR spatial relations in the Euclidean plane, corresponding to Fig. 2. Note that CDR contains inverse relations following Table 1 and are defined as follows:

- if $N_i(B, A)$, then $S_i(A, B)$
- if $NE_i(B, A)$, then $SW_i(A, B)$
- if $E_i(B, A)$, then $W_i(A, B)$
- if $SE_i(B, A)$, then $NW_i(A, B)$
- if $S_i(B, A)$, then $N_i(A, B)$
- if $SW_i(B, A)$, then $NE_i(A, B)$
- if $W_i(B, A)$, then $E_i(A, B)$
- if $NW_i(B, A)$, then $SE_i(A, B)$

For a better understanding, let's look at the scenario of Fig. 2 that represents Allen's relations corresponding to the picture p_7 observed by camera c_7 with a referential r_7. We have the following rules in the referential r_7 :

- A_7 (Orange) precedes E_7 (Aquamarine) and $delta\_size(A_7, E_7) < \Gamma$ then $SE_7(A, E)$.
- C_7 (Green) meets E_7 (Aquamarine) and $delta\_size(A_7, E_7) \simeq 0.0$ then $E_7(C, E)$.
- D_7 (Yellow) overlapped by E_7 (Aquamarine) and $delta\_overlaps(E_7, D_7) > \beta$ then $S_7(D, E)$.
- A_7 (Orange) overlapped by C_7 (Green) and $delta\_overlaps(C_7, A7) < \beta$ then $SE_7(A, C)$.

This paper falls within a small subset of QSR formalism that considers the camera's position and orientation.

Table- I: Inversion table of basic cardinal relations.

R*i*	inverse R*i*
N_i	S_i
W_i	E_i
NW_i	SE_i
NE_i	SW_i
S_i	N_i
E_i	W_i
SW_i	NE_i
SE_i	NW_i

IV. ONTOLOGY FOR SPATIAL REASONING

The challenge here is the building of a knowledge system that can extract a QSP model (positions and orientations of objects and cameras) that satisfies all constraints identified in Section V.A without any calculations. Such knowledge is based on ontologies as a formal representation to bridge the conceptual framework. Ontologies can be used to cover different terminologies and to represent a clear specification of the different meanings. As a knowledge base, ontologies have two components: a TBox (intensional knowledge in the form of a terminology) and an ABox (extensional knowledge that is specific to the individuals of the domain of discourse). Furthermore, the use of ontologies will help in maintaining a strict distinction between data and an interpretation based on the data. A particular conceptual framework along with the associated ontology is the optimal way to create a formal representation fit for different abstraction level. An ontology is a set of data elements within a domain that are linked together to denote the types, properties, and relationships between them. In other words, as depicted in [68], the main features of ontologies are: *(i)* Entities and relations, i.e. referred to by a standard identifier such as an Internationalized Resource Identifier (IRI), a Uniform Resource Identifier (URI), that enables data integration since the same identifiers can be used across multiple databases, files, or websites; *(ii)* Domain vocabulary, i.e. a list of terms associated with the ontology's concepts and relations that can be exploited for applications ranging from natural language processing, creation of user interfaces, etc. (iii) Metadata description, i.e. textual definitions and descriptions that provide additional information about what kind of things a concept (/properties) refers to, which enable domain experts to understand the precise meaning of the concept (/properties) in the ontology. (iiii) Domain axioms, i.e. statements that are considered to be true within that domain and which provide background knowledge about a domain.

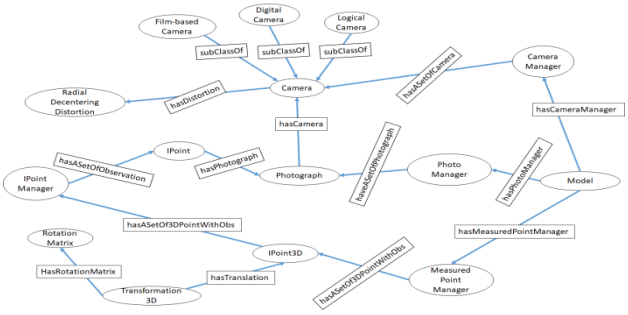


Fig. 5. A partial view of the photogrammetry ontology Arpenteur.

Ontologies provide a common way of representing knowledge about a specific domain and a way to share a common understanding of its information structure.

Once we have common understanding, we can try to reason/query over this information, i.e. inference, consistency checking, etc.

We serialized our ontology with the Web Ontology Language OWL2, which is a World Wide Web Consortium (W3C) recommendation[4]. OWL2 ontologies enable the representation of concepts (classes), instances (individuals), attributes (data properties for which the value is a data) and relationships (object properties for which the value is an individual).

A. Arpenteur: an ontology for photogrammetry

The Arpenteur ontology[5] covers all photogrammetry items for objects measurement. This generic upper ontology can be used to conceptualize any photogrammetry system. A partial view of the Arpenteur ontology is shown in Fig. 5.

The proposed ontology models most of the concepts and relationships needed to describe a photogrammetric model and the associated computational processes. The proposed modeling reflects as closely as possible the physical reality of the phenomenon. Thus, the ontology contains all the concepts necessary for photogrammetric computation but also the objects potentially measured by photogrammetry and the interface necessary for this measurement. The 3D points have a set of 2D observations on photographs that allow their coordinates to be calculated by optimized triangulation. The 2D points are observed on photographs that are produced by cameras whose intrinsic parameters are faithfully described. The current model contains the focal length, the eccentricity represented by the projection of the optical center on the image plane as well as different models of distortion (radial or tangential). The management of the sense of distortion (correction or application) is also taken into account because it varies according to the existing software. (See, for example, a conversion solution proposed in [69]). A *Photograph* is thus the image produced by a camera (digital or logical in our case) and the Camera is the object that produces the Photograph. Furthermore, a Photograph is related to the concept Transformation3D that locates the position of its optical center in the 3D world space, as well as its orientation. A direct connection between Photograph and Camera(s) through the relation hasCamera allows to use a calibrated camera for a set of photographs. This modeling approach was initially thought to solve interoperability problems between traditional photogrammetry software (Bingo [70], Photomodeler [71], Agisoft [72], OpenCV [73]). Indeed, the use of an ontology allows an easy adaptation between various formalisms. However, in recent years, the evolution of information and communication technologies has stimulated the recovery of spatial, temporal, and spatio-temporal data from moving objects. We are now considering an alignment of this ontology with the ontologies managing semantic trajectories [74]. The concept of Camera considers digital cameras but also cameras using film with fiduciary marks. (This type of camera is practically no longer used nowadays, but this model can be useful for the analysis of archive photos, as for example in this work on distortion models [75]). Another kind of Camera has been introduced for the purposes of this work, the concept of LogicalCamera. (which could have been called VirtualCamera). It is a virtual camera positioned in the 3D models used in this work and producing images according to the collinearity equation (see Equation 2). The singularity of this camera also lies in the fact that it evolves in a discrete world and that, in its current version, the rotation matrix is limited to an angle around the OZ axis.

In the following section, we describe the population of the Arpenteur ontology within an example of the *image 7* taken with *camera 7*.

B. CDR Knowledge Base

An ontological knowledge base can be seen as a structured system of fundamental concepts and relationships and of an agreed epistemology [76]. In other words, it can be seen as a graph of concepts and data connected with typed relations. Ontology can be used to guide the design of knowledge-based applications and to store the various experimental data as well as the reasoning process in a knowledge manner. Accordingly, the key contribution of this paper is the population of the Arpenteur ontology by the CDR information in order to produce a knowledge base (KB) where the reasoning is performed by an iterative SPARQL query over the CDR KB. The KB contains a schema part (TBox) and a data part (ABox). The schema part is represented by the photogrammetry ontology Arp (stand for Arpenteur) which is serialized in an OWL2 ontology representation. The data part consists of the graph of instances corresponding to the concepts in the ontology.

In Fig. 6, we have a partial view of the knowledge base representation in our CDR model corresponding to camera 7 and the objects seen by camera 7, a green sphere and an aquamarine sphere. In the schema part, all the objects are defined by an upper concept Arp:IdentifiedObject. The concept Arp:Camera in Arp is a super concept of three types of phtogrammetric cameras: Arp:DigitalCamera, Arp:FilmBasedCamera and Arp:LogicalCamera. In the presented scenario, cameras are instances of the concept

Arp:LogicalCamera and spheres are instances of the concept Arp:FitBall as follows:

- Arp:Green and Arp:Aquamarine are instances of the concept Arp:FitBall and represent respectively the spheres green and aquamarine.
- Arp:LogicalCameraView_7 is an instance of Arp:LogicalCamera and represents the camera 7. This camera has an orientation of value 5 which corresponds to the South direction.
- Arp:LogicalOrientation220038608 represents the observation of the logical camera 7. This observation is represented by a domain, target and an orientation that corresponds to a QSR rule as defined in Section III.B. The observation rule is an oriented relation where we have Orientation (Source, Target), i.e, E(Green, Aquamarine) read as ”Aquamarine is east of Green”.

The ontological KB is populated during the orientation process where the consistency is verified throughout the CDR knowledge base building, as described in Section V.

[4]https://www.w3.org/TR/owl2-overview/
[5]http://www.arpenteur.org/ontology/Arpenteur.owl

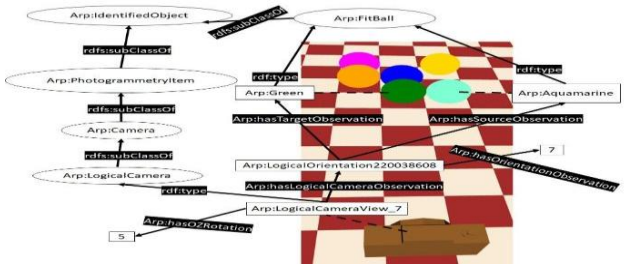


Fig. 6. **Disposition of spheres as seen by camera 7 with Arpenteur ontology descriptions.**

V. MODUS OPERANDI

This paper presents a reduced but functional implementation of a “Logical Bundle Adjustment” as a proof of concept. Of course, this implementation contains a set of constraints and limitations that are there to make the implementation easier without however affecting the general presentation of the concept. This section details the implementation and the limitations needed to make these early developments easier.

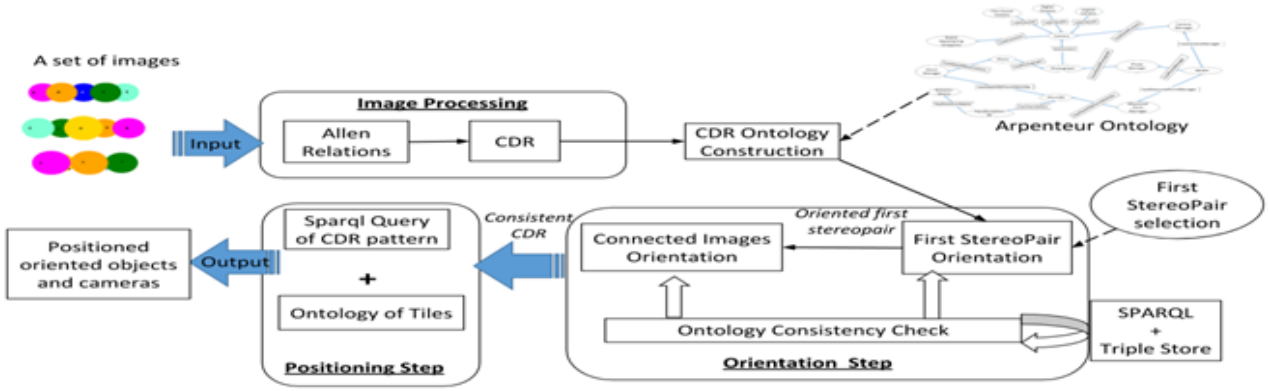


Fig. 7 Modus operandi of the logical bundle adjustment system.

First, a quick reminder of the objectives: for a 3D scene in a given space, composed of a set of objects, observed from several points of view by photographs taken from different positions. The problem we want to solve is the determination of the relative positions of the objects present in the scene and visible on the photographs as well as the positions and orientations of the cameras by simply interpreting the relative positions of objects in the photographs.

A. Limitations and image processing

We therefore generated a virtual scene, positioned virtual cameras in this scene and calculated the images obtained by each of the cameras. These images were then automatically analyzed to extract the cardinal relationships between the visible objects. A set of limits and constraints have been set for the generation of the scene (see Fig. 8):

1) Objects are spheres. This is to simplify the calculation of the shape projected on the images.
2) All objects are the same size. This simplifies the establishment of cardinal relationships during the analysis of photographs.
3) Each sphere has a unique color. This allows us to recognize the object in the images. We therefore have a set of homologous objects seen on the generated images.
4) All the objects of the scene are relatively close to each other, the scene is homogeneous. This allows the cameras to be distributed over a circle approximately centered on the scene.
5) The objects, spheres, and cameras are distributed on a plane as it is shown in Fig. 8 and in the web interactive 3D interface.
6) The space in which the objects are distributed is discrete. The objects, here the spheres and cameras, are distributed on a uniform checkerboard grid plane where each tile can only receive one object.
7) The number of tiles is approximately related to the size of the scene, spheres and cameras included, too many tile would lead to a significant increase of equivalent solutions.
8) The local reference frame of a camera is such that its optical axis is oriented towards the North. Furthermore, according to condition 4, the objects are not scattered on the plane but relatively grouped. These two conditions result in the fact that all objects visible by a camera are in its North. In a future version, the focal length and sensor size of the camera will be considered so that the perspective and scale of the projected objects can be considered when analyzing the image.The virtual scene that was used for this experiment is accessible on the net via an interactive 3D interface. It is available at the following addresshttp://www.arpenteur.org/lba/3D/. The objects, photographs and each iteration of the inference mechanism leading to the result is visible.

The scene is built using the 3D online tool, 3DHOP (3D Heritage Online Presenter) which is an open-source software package for the creation of interactive Web presentations of high-resolution 3D models (see [77]) developed by the Visual Computing Lab of ISTI-CNR.

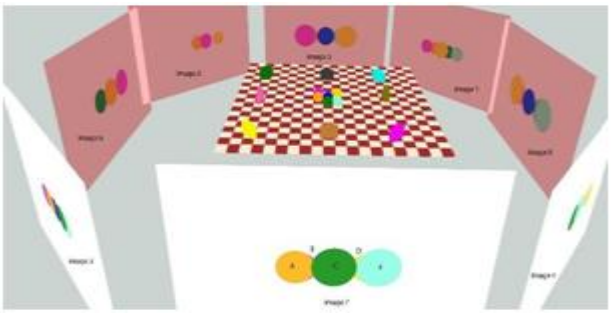


Fig. 8. Virtual scene: spheres and cameras on tiled plane with generated images.
See *http://www.arpenteur.org/lba/3D/*

This interactive application shows the virtual scene used for the demonstration and provides access to the ontology used in order to see, for example, on which image a sphere is seen or what cardinal relation can be deducted from a given camera.

A scene is built with 6 spheres grouped more or less on the central area of a 20x20 tile paved plane. The eight cameras are distributed in a circle around the scene and are pointing to its center.

Each camera produces an image of the scene with only the spheres, neither the checkerboard plane nor the other cameras are projected on the image.

The first step is to obtain the images from each camera. An internal orientation has been chosen for the generation of the images but does not count for the rest of the processing. Each camera has a 1800 x 1200 pixel sensor with 0.02 mm sides. The focal length is 30 mm and, of course, there is no distortion nor eccentricity.

Images are produced simply by using collinearity relations (from Kraus 30):

$$u = u_0 - c\frac{r_{11}(X - X_{0)} + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_{0)} + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} \quad (2)$$

$$v = v_0 - c\frac{r_{12}(X - X_{0)} + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_{0)} + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)}$$

where

- (u,v) unknown image coordinates.
- (u_0, v_0) coordinate of the principal point.
- c principal distance.
- r_{ij} elements of a rotation matrix R, which then defines the position of the photo in space relative to the object coordinate system (X, Y, Z). The r_{ij} elements can be expanded according to the three rotations.
- (X_0, Y_0, Z_0) perspective center coordinate in object space (i.e. camera location).

Equation 2 is therefore used to fill the images from the 3D points calculated on the spheres. The mesh of the spheres is obtained from successive subdivisions of an original icosaedron. More than 5 million 3D points are calculated on each icosphere and projected on the images. This allows a fine smoothing of the contours. Each vertex is projected on the image while keeping its color. No notion of shading is used, so the spheres are uniformly represented on the image. A simple version of the painter's algorithm, projecting the furthest spheres first, makes it possible to obtain an image where the occlusions of the discs represent a difference in depth along the camera's optical axis. The cameras represented in the scene (see Fig. 8) are arbitrarily positioned and it is from these positions and orientations that the images were generated. The same is true for the spheres. This is the field truth against which we will compare the results. The images are automatically analyzed which allows the identification of the spheres by their color. The identification of the colored areas allows the calculation of the perimeter segments for each sphere and thus the calculation of the circles using the least-squares method (see Fig. 2). Once the projection of the center of each sphere on the image is obtained, the calculation of the relative positions of the spheres in the image as well as the determination of the overlap and relative sizes is easy. When the relative positions, overlap and ratios between the diameters of the discs in the photo have been evaluated, a set of rules transforms these data into CDR facts describing the relationships between the spheres on the grid (see Fig. 2). As can be seen in this figure and in Section III.A, Allen's relationships describing the intervals or occlusions between the images of the spheres in the photographs are translated into the expression CDR. A sphere a bit on the right or a bit on the left in respect to another is translated in depth positioning on the 2D plane. This is an immediate and intuitive translation of the effect of parallax in photogrammetry, however the rules have been empirically refined using context-specific thresholds. For example, the fact of not using focal length in the analysis of photographs led us to use these thresholds. In a future version, the focal length may be used or even be an unknown to the system. These CDR expressions are then used to populate an ontology that models the scene. Of course, these expressions are evaluated in the reference system of the camera that have generated the photo. The North is therefore parallel to the camera's optical axis and all spheres are set to the North of the camera.

The output of the image analyzing consists of a set of CDR facts describing the relationships between spheres on the grid in the reference system of the camera. Let's consider an example taken from the 3D scene presented in Fig. 8 and composed of images 1, 3 and 6.

In this example illustrated in Fig. 9 and Table II, we consider an LBA scenario applied on spheres and cameras. The 6 spheres are identified by their color and a letter: (orange, A), (fushia, B), (green, C), (gold, D), (aquamarine, E) and (blue, F). The three cameras used in this example are cameras 1, 3 and 6 which have generated images 1, 3 and 6 visible in Fig. 9. The CDR, as Cardinal Direction Relation, presented in Section III.B is computed for each image in this example and the results are presented in Table II. Each of cardinal direction relation is in the referential of the corresponding camera. For example the relations NE1(aquamarine, green), S6(aquamarine, green) and SW3(aquamarine, green) reflect the CDR between the spheres aquamarine and green in the referentials r_1, r_2 and r_3 for respectively image 1, image 2 and image 3.

From this stage onwards, most of the process will be done by inference on ontologies using a typically photogrammetric approach. In fact, there is a clear parallel between the results obtained by inference and those that we could imagine, under similar geometric circumstances, in photogrammetry.

B. Inference and photogrammetric process

The first step in the process of orienting a set of unordered photographs is to generate all possible pairs and then to order the pairs according to a quality criterion.

Here, considering that the poses of each camera are unknown and that all the intrinsic parameters of the cameras are equal, the quality criterion used to order the stereopairs is the number of spheres that the two cameras have in common. For example, the couple (view 1, view 6) have 5 common objects in their visions while the couple (view 1, view 3) see only 4.

This quality criterion is reflected in a maximization of the CDRs linking the spheres and the cameras, which leads to a minimization of the number of possible solutions and therefore to more robust solutions on which the other stereopairs will be implanted.

To do this, our pipeline consists of two main steps: (1) the orientation of the first stereo pair (2) the orientation of the rest of the connected images.

This looks like an iterative bundle adjustment, but we are in a discrete space where there is no optimization, the solutions found at iteration i are preserved and the solutions at iteration i+1 keep intact the previously existing solutions.

Algorithm 1. First Stereo Pair Orientation.

```
Input: Arpenteur Ontology, Selected Stereopair (camera_1, camera_2);
Output: Oriented Stereopair;
fix camera_1 orientation to the north;
fix camera_2 orientation to the north;
while (checking ontology consistency) is inconsistent do
 increment orientation camera_2;
end
```

Table- II: CDR facts between spheres and also cameras, inferred from the images presented in Fig. 9. We note that for each camera, all the visible spheres are considered to be north of the camera.

Image 1 - $r1$	**Image 6** - $r6$
NE_1(aquamarine, green)	E_6(fuchsia, orange)
E_1(aquamarine, fuchsia)	NE_6(fuchsia, green)
E_1(aquamarine, orange)	N_6(fuchsia, gold)
SE_1(aquamarine, gold)	NE_6(fuchsia, aquamarine)
W_1(fuchsia, aquamarine)	W_6(orange, fuchsia)
W_1(fuchsia, green)	NE_6(orange, green)
NW_1(fuchsia, orange)	NW_6(orange, gold)
SW_1(fuchsia, gold)	NE_6(orange, aquamarine)
NE_1(gold, fuchsia)	SW_6(green, fuchsia)
N_1(gold, green)	SW_6(green, orange)
N_1(gold, orange)	NW_6(green, gold)
NW_1(gold, aquamarine)	N_6(green, aquamarine)
E_1(green, fuchsia)	S_6(gold, fuchsia)
E_1(green, orange)	SE_6(gold, orange)
S_1(green, gold)	SE_6(gold, green)
SW_1(green, aquamarine)	E_6(gold, aquamarine)
W_1(orange, aquamarine)	SW_6(aquamarine, fuchsia)
W_1(orange, green)	SW_6(aquamarine, orange)
S_1(orange, gold)	S_6(aquamarine, green)
SE_1(orange, fuchsia)	W_6(aquamarine, gold)
N_1(camera 1, orange)	N_6(camera 6, orange)
N_1(camera 1, green)	N_6(camera 6, green)
N_1(camera 1, fuchsia)	N_6(camera 6, fuchsia)
N_1(camera 1, aquamarine)	N_6(camera 6, aquamarine)
N_1(camera 1, gold)	N_6(camera 6, gold)

- **First Stereopair Orientation**

This step is represented by Algorithm 1. The first step is to load the Arpenteur ontology. This ontology does not yet contain any instance of CDR.

Then comes the choice of the starting stereopair. This is done with the quality criterion where a weight is given to each pair of photographs according to the number of spheres visible in the stereopair. In the example shown in Fig. 9 it is a question of processing 3 views. Three stereopairs are possible (view 1, view 6), (view 1, view 3) and (view 3, view 6). The stereopair (view 1, view 6) will be chosen because it has 5 objects in common against only 4 for the two others. The next step is to find the orientation of the selected stereopair. The process begins with fixing the stereopair orientation to the initial direction, the north. This orientation results in the storage of the oriented CDR facts of viewed objects in the ontology as shown in the Table- II where a consistency checking is performed. The consistency checking consists of a SPARQL query answering procedure over the ontology. This procedure is to verify the existence of any conflict within the oriented CDR facts stored in the ontology. The performed SPARQL query is presented in form of SPARQL ASK[6] query.

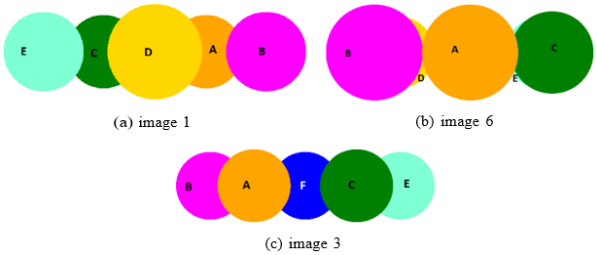


Fig. 9. **Images obtained from cameras 1, 3 and 6. The CDR translation is shown in Table- II.**

Our applications use the ASK query to test whether the query pattern has a solution. No information is returned about the possible query solutions, just whether a solution exists. If the result of the SPARQL query is empty, it means that the ontology is consistent, and the system orientation is correct. Otherwise, if the result contains at least one answer, it means that the ontology is inconsistent, and the system orientation is incorrect. In the case of inconsistency, the

[6] https://www.w3.org/TR/2013/REC-sparql11-query-20130321/##ask

camera orientation is incremented following the orientations shown in Fig. 4. The system stops if the ontology is consistent.The performed SPARQL query is presented in the List. 1.

List. 1. **SPARQL Orientation consistency Checking.**

```
prefix Arpenteur: <http://www.arpenteur.org/ontology/Arpenteur.owl#>

ASK WHERE{
?cam1 Arpenteur:hasLogicalCameraObservation ?lo1;
Arpenteur:hasName ?nameCam1.

?lo1 Arpenteur:hasSourceObservation ?source;
Arpenteur:hasTargetObservation ?target;
Arpenteur:hasOrientationObservation?o1.
?cam2 Arpenteur:hasLogicalCameraObservation ?lo2;
Arpenteur:hasName ?nameCam2.
?lo2 Arpenteur:hasSourceObservation ?source;
Arpenteur:hasTargetObservation ?target;
Arpenteur:hasOrientationObservation ?o2.
?source Arpenteur:hasName ?nameSource.
?target Arpenteur:hasName  ?nameTarget.
FILTER(?cam1!= ?cam2)
FILTER(?o1!= ?o2)
}
Orderby ?nameCam1 ?nameCam2
```

Let's take the stereopair scenario (view 1, view 6) of Fig. 9 representing respectively the source view and the target view. Orienting view 1 to the NE means orienting all the CDR facts of the Table II of view 1 by following the composition table of the basic cardinal relations (see Table III). For example NE1(aquamarine, green) oriented to NE becomes E(aquamarine, green). Similarly, orienting view 6 to NE will result in transforming S6(aquamarine, green) to SW(aquamarine, green). The system then stores these facts in the ontology and performs the SPARQL query. In this scenario, the SPARQL query will detect conflicts between the different CDR orientations, i.e. the conflict consists of E(aquamarine, green) and SW(aquamarine, green). Hence the result of the consistency checking is "inconsistent". The ontology is then reset to remove the newly introduced facts and check the consistency for the next orientation. The system will stop when the SPARQL is empty. In this scenario, it stops when view 1 is oriented towards NE and view 6 is oriented towards W, giving the same orientation for all viewed object, i.e. the same E orientation (aquamarine, green) for view 1 and view 6.

Table- III: Composition table of basic cardinal relations. (R_c is the composition of relations R_i and R_j)

R_i	R_j	R_c
N_i	N_j	N_c
N_i	NE_j	NE_c
N_i	E_j	E_c
N_i	SE_j	SE_c
N_i	S_j	S_c
N_i	SW_j	SW_c
N_i	W_j	W_c
N_i	NW_j	NW_c
S_i	N_j	S_c
S_i	NE_j	SW_c
S_i	E_j	W_c
S_i	SE_j	NW_c
S_i	S_j	N_c
S_i	SW_j	NE_c
S_i	W_j	E_c
S_i	NW_j	SE_c
NE_i	N_j	NE_c
NE_i	NE_j	E_c
NE_i	E_j	SE_c
NE_i	SE_j	S_c
NE_i	S_j	SW_c
NE_i	SW_j	W_c
NE_i	W_j	NW_c
NE_i	NW_j	N_c
SW_i	N_j	SW_c
SW_i	NE_j	W_c
SW_i	E_j	NW_c
SW_i	SE_j	N_c
SW_i	S_j	NE_c
SW_i	SW_j	E_c
SW_i	W_j	SE_c
SW_i	NW_j	S_c
E_i	N_j	E_c
E_i	NE_j	SE_c
E_i	E_j	S_c
E_i	SE_j	SW_c
E_i	S_j	W_c
E_i	W_j	NW_c
E_i	W_j	N_c
E_i	NW_j	NE_c
W_i	N_j	W_c
W_i	NE_j	NW_c
W_i	E_j	N_c
W_i	SE_j	NE_c
W_i	S_j	E_c
W_i	SW_j	SE_c
W_i	W_j	S_c
W_i	NW_j	SW_c
SE_i	N_j	SE_c
SE_i	NE_j	S_c
SE_i	E_j	SW_c
SE_i	SE_j	W_c
SE_i	S_j	NW_c
SE_i	SW_j	N_c
SE_i	W_j	NE_c
SE_i	NW_j	E_c
NW_i	N_j	NW_c
NW_i	NE_j	N_c
NW_i	E_j	NE_c
NW_i	SE_j	E_c
NW_i	S_j	SE_c
NW_i	SW_j	S_c
NW_i	W_j	SW_c
NW_i	NW_j	W_c

- **Connected Images Orientation**

Once we are able to orient two images, the first one in respect to the second, the problem is to orient a large set of images which are strongly interconnected. This problem is well known in photogrammetry and is addressed by all approaches using bundle adjustment. Usually a connectivity graph is generated ([47], [78], [79], [80]) and the algorithm used runs through the graph by choosing pairs of photos, one of which is already oriented, calculating the orientation of the new pair introduces the new photo into the set of oriented photos. This provides the approximate values needed to calculate the BA.

The approach proposed in this article could be similar to this step of determining the approximate values of the objects and cameras present in the scene.

These are approximate values because we are in a discrete universe both in terms of positions in space and angular determination. Nevertheless, after analysis of the images, a connectivity graph is generated and ordered. A weight is given to each pair of photographs according to the number of spheres visible in the stereopair.

Algorithm 2.**Connected Images Orientation.**

Input: Arpenteur Ontology, First stereopair, Sorted list of unoriented cameras;
Output: List of oriented cameras;
add first stereopair to the sorted list of oriented cameras;

For(target camera **in** List of unoriented cameras)
fix target camera orientation to north;
  While (checking ontology consistency) is inconsistent **do**
increment orientation target camera;
end While
add oriented target camera to the list of oriented cameras
end For

This process is represented in Algorithm 2. The process begins with the acquisition of the ontology resulting from the previous step. Then the algorithm selects an unoriented target camera from the list.

The order here is based on the weight given to the unoriented list of cameras according to the number of common visible spheres in the images from the cameras whose orientation has been previously determined.

After the selection of a target camera, the orientation is set to the north and we append its viewed cardinal direction relations to the ontology. As for the single stereopair algorithm, its consistency is verified using the SPARQL ASK query as depicted in Listening 1. This SPARQL query answering checks the consistency of all CDR facts in the ontology, including all previous orientations of the analyzed views. If the ontology is inconsistent, a reset is applied to its state before the inconsistency. Also, the orientation of the target camera is incremented. If the ontology is consistent, the ontology is saved, and the loop continues with the next target view in the sorted list of stereopairs. Once all views are analyzed, all CDR facts are used as constraints to position the different objects in their corresponding tiles in the plane grid.

- **Positioning the objects on a checkerboard plane grid**

In this version of the LBA all objects, spheres and cameras are distributed on a checkerboard plane grid. After having generated the CDRs of all these objects, these objects must be positioned in the plane grid. The set of positioning constraints from CDRs, is described as a pattern linking all the objects. To solve the positioning of the objects in the plane, we look for a similar pattern among the elements on the plane grid. To do this, the tiles are also organized by CDRs, under the 8 cardinal directions (see Fig. 4) where all possible relationships between tiles are expressed in a graph pattern. Hence, the system will look for a match between the graph pattern of cameras and the objects and, on the other side, all possible graph patterns of tiles in the grid plane.

Thus, the proposed approach is based on the fact that tiles, as elements of the paving of space, are also bound by cardinal direction relationships (CDR). For example, in the 20X20 plane, thetile (0,1) of the grid is east of thetile (0,0), i.e. E($tile_{00}$, $tile_{01}$).

The pattern matching is detected by SPARQL reasoning over the CDR populated ontology. In fact, the SPARQL query for the pattern matching is constructed during the orientation steps via the method.

This SPARQL query reflects the consistent CDR graph linking objects and cameras from the produced consistent ontology. An example of consistency from our scenario where view 1 is oriented towards NE and view 6 is oriented towards W resulting to a pattern matching relation E(aquamarine, green). A binding between the CDR consistent graph and its corresponding tiles graph is added to the query, i.e. E(tileAquamarine, tileGreen) is the binding of E(aquamarine, green). In such a manner, the result of the query is merely the tile in the grid plane corresponding to the objects, i.e. the value of tileAquamarine corresponds to the position of the tile where the aquamarine object is positioned. The final SPARQL query must contains all CDR tiles binding from the CDR relations of objects and cameras. Finally, the complete modus operandi of the presented LBA system is depicted in Fig. 7 The results and performance of this SPARQL query answering solution are discussed in Section VI.

List. 2. **SPARQL Orientation consistency Checking.**

```
prefix Arpenteur: <http://www.arpenteur.org/ontology/Arpenteur.owl#>

SELECT   ?tileAquamarine ?tileGreene ...
WHERE {
?tileAquamarine Arpenteur:E ?tileGreene.
...
Filter( ?tileAquamarine !=  ?tileGreene)
}
```

In fact, the SPARQL query for the pattern matching is constructed during the orientation steps via the method. This SPARQL query reflect the consistent CDR graph linking objects and cameras from the produced consistent ontology. An example of consistency from our scenario where view 1 is oriented towards NE and view 6 is oriented towards W resulting to a pattern matching relation E(aquamarine, green). A binding between the CDR consistent graph and its corresponding tiles graph is added to the query as reflected in List. 2 where E (tileAquamarine, tileGreen) is the binding of E (aquamarine, green). In such a manner, the result of the query is merely the tile in the grid plane corresponding to the objects, i.e. the value of tileAquamarine corresponds to the position of the tile where the aquamarine object is positioned. The final SPARQL query must contains all CDR tiles binding from the CDR relations of objects and cameras. Finally, the complete modus operandi of the presented LBA system is depicted in Fig. 7 The results and performance of this SPARQL query answering solution are discussed in Section VI.

VI. CONCLUSIONS AND PERSPECTIVES

A. Implementation and performance

The proposed system software is developed in the Java programming language.

The generation of images, 3D models and ontology management are based on the ARPENTEUR platform (See the Arpenteur website [81]).

The orientation steps described in Section V.Aconsist of ensuring that images are correctly oriented by a coherence check performed via querying the ontology model by SPARQL. These orientation steps generate an ontology of all CDRs orientations relations (graph pattern) of the cameras and the viewed objects. The positioning step described in Section V.Bconsists of querying the produced ontology to find a matching pattern in the grid plane. Since SPARQL querying is the key concept of our system, we implemented an RDF triplestore – Apache Jena Fuseki server [82] offering a persistent storage of the ontology and a W3C SPARQL 1.1[7] quering the stored data. Essentially, SPARQL is based on matching graph patterns against RDF graphs. Graph patterns are based on triple patterns. RDF graphs represent the ontologies. The SPARQL querying process works by transforming the query into a subgraph matching query over the stored RDF graph (ontologies). In other words, answering a SPARQL query means finding all

[7] https://www.w3.org/TR/sparql11-query

instances (subgraph) from a given data graph matching the pattern specified by a query graph. This particular subgraph matching feature fits perfectly into our development strategy of matching the graph pattern extracted from images (graph pattern of cameras and the viewed objects) into the grid plane in order to find the corresponding match of tiles graph patterns.

Table- IV: SPARQL query answering duration in Fuseki triplestore of the pattern matching step by dataset size and grid plane dimensions.

Grid dimensions	Number of triples	Query answering duration
10X10	17370	80 ms
15X15	64630	158 ms
20X20	183270	2770 ms
25X25	816446	5577 ms
30X30	859770	835 ms
35X35	1567620	312 ms

The system's performance was evaluated using a PC with a CPU of i7-8700k where the process of images orientation (including ontology development, consistency checking via SPARQL, image processing and CDR fact generation) takes around 14 seconds. Table IV shows our experiments where the ontology size (number of triples) is proportional to the grid plane dimensions (number of tiles). On the other hand, measuring the performance of the triplestore response for one result is not linear since several features (detailed in [83]) could impact the overall query execution time, such as the number of projection variables, join vertices, triple patterns, the result sizes, the join vertex degree, etc. Evaluating a triplestores in such a setup can only provide very partial evidence on performance. Hence, we opted for a live demonstration where we made available a demo[8] showing a live performance of the SPARQL answering by Apache Fuseki corresponding to different plane grid dimensions depicted in Table IV. Note here, that the grid dimensions could be set as highlighted in Section V.A where the number of tiles must be approximately related to the size of the cameras included. Since there is no universal winner amongst triplestores [84], [85], a customizable SPARQL benchmark generation framework like [86] can be our future direction to compare the performance of different triplestores (i.e. Virtuoso[9], GraphDB[10], etc.) within our application. Also, the scalability of our system can be handled by different existing research work on distributed/parallel SPARQL queries over large-scale RDF graphs such as HadoopRDF [87] and the distributed gStore system [88].

B. Results

It is interesting to note that the approach presented here behaves like a real photogrammetric process because the constraints are of course similar: observation of the photographs, interpreting the parallax in depth, possible resolution as soon as an object is observed on two photographs, greater imprecision in the direction of the optical axis, etc. As can be seen on the online simulation (http://www.arpenteur.org/lba/3D/), in the first iteration, four spheres are positioned by two cameras. In the second iteration, a fifth sphere is added as well as a third camera. The third, fourth and fifth iterations do not add any spheres, only one more camera per iteration. The sixth iteration adds a camera and the last sphere which is in the center of the model and which was strongly hidden by the others. The last two iterations add only one camera each. Each time an iteration adds only one camera, it means that the new observations on the spheres corroborate the previous ones and do not contradict them. The working universe being discrete there is no optimization possible only agreement or inconsistency with the previous observations. Nevertheless, a notion of precision can be evaluated. Indeed, if an optimization is not directly possible because we are in a discrete universe, the system produces a set of possible solutions. We can see in Fig. 10 the trace of the various possible solutions given by the system: more than 500 solutions are consistent with the observations; all give the same position for the spheres, only the cameras can move. The Fig. 10 shows the accuracy obtained for the camera positions. The trace of the possible positions for each camera is colored with the color of the camera. This figure shows all possible positions that satisfy the CDR observations resulting from the image analysis. In fact the focal length not being used in the constraints the cameras can be positioned anywhere on the grid as long as they are north of the spheres because there are no constraints between the cameras but only sphere/camera relations. The precision related to the position of the cameras therefore only affects the position and not their orientation. In all the solutions obtained, the orientation of the cameras is always the same, only their position changes.

C. Photogrammetric results

The application presented here infers the positions of the spheres observed in the photos as well as the positions and orientations of the cameras in a discrete space. In this way, this application is very close to a photogrammetry application but does not use the homologous points extracted from the images.

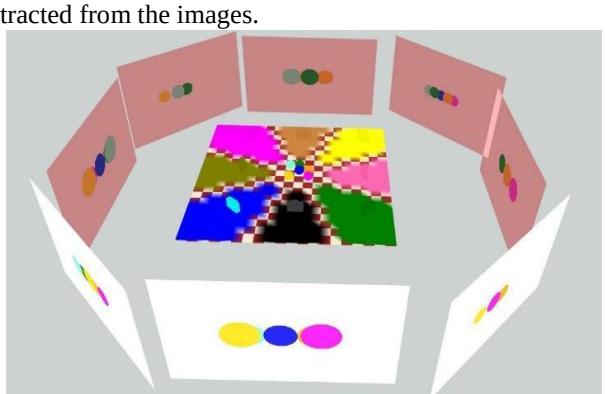


Fig. 10. Camera precision: colored traces represent the possible positions of the camera with the same position of spheres.

[8] http://www.arpenteur.org/lba/relax.html
[9] https://virtuoso.openlinksw.com
[10] http://graphdb.ontotext.com

Apart from the non-existence of the points, we can easily translate the results to standard photogrammetry software. We have opted for two softwares suites: Photomodeler from EOS and Metashape from Agisoft. These software suites were chosen due to their ease in introducing external data.

For Photomodeler, the interface foresees it and for Metashape, in the latest version (1.6), a JAVA binding is proposed, which makes it easy to communicate with the LBA developments also in JAVA.

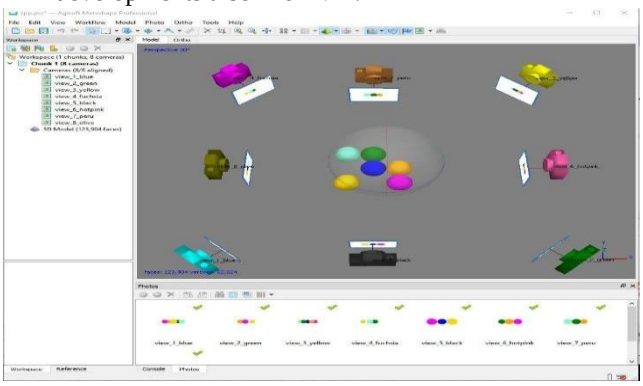


Fig. 11. Result imported in Metashape with JAVA binding approach.

We have developed a JAVA bridge between the results proposed by LBA and Metashape version 1.6. The results are shown in Fig. 11. The cameras and photos have been imported and Metashape is able to display them as well as the 3D model of the calculated spheres. We also added the 3D cameras in the final model imported in Metashape.

The result is fully satisfying even if Metashape itself cannot orient such a model at all: there are no 2D or 3D points and no homologous points can be extracted from the photographs because the colors of the discs are completely homogeneous.

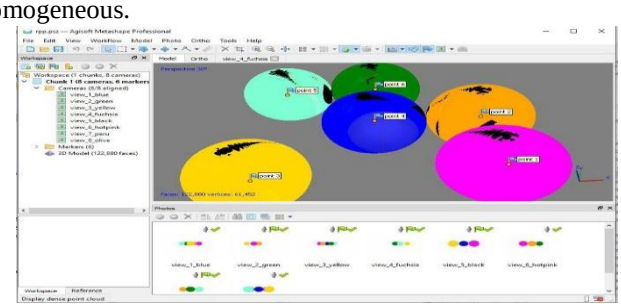


Fig. 12. Managing model computed wilt LBA in Metashape: Sphere center added using markers on photographs and texturing the model using oriented photographs.

Fig. 12 shows the use of the orientation results produced by the LBA in Agisoft's Metashape software. This figure shows new 3D points, markers identified manually on the photographs as the centers of the color discs. These points were then calculated in 3D and represented in the model reconstructed in Metashape. On the same figure, the spheres are colored by projecting the texture from the photographs. It is perfectly visible that the texturing process, which is based on a correct orientation of the photographs in relation to the 3D model, is of good quality. Some imperfections are visible on the tangent parts (3D model / projective ray) where the texture is less defined.

D. Future work

The work presented here is a proof of concept and not an operational application. We have shown that it is possible to obtain the positions and orientations of the spheres visible on the photos as well as the camera poses by simply inferring on the relative positions of the spheres projected on the photos. However, the presented application has a number of constraints that we propose to overcome in the future.

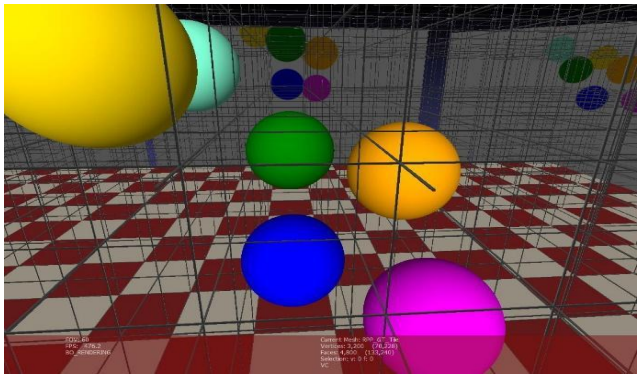


Fig. 13. Future work: Use of a discrete 3D space by voxels.

First, all the objects are currently on the same plane. The discrete space is in fact the surface of the grid plane. We have started the developments necessary for the management of objects in space. In this case, the objects are distributed in a space structured by voxels (see Fig. 13). It is then necessary to widen the CDRs by adding up, down, and their relative declinations (upNorth, upNorthWest, etc.). The most interesting part will be the extension of Allen relations to handle two-dimensional occlusions on images. The second point we are going to work on is the use of camera focal length in determining CDR relationships. Indeed, this will have the effect of constraining more strongly the position of the camera with respect to the spheres, which is where the current system generates uncertainty (see Fig. 10). Once these two points have been resolved we will introduce objects more complex than spheres, objects whose orientation in space will need to be found. The problem will then be to recognize them in the images even if they are occluded by others. We are thinking about proposing a deep learning approach for the 2D recognition aspects which will be in fact the main problem brought by this point of development.

APPENDIX

The following abbreviations are used in this manuscript:
3DHOP: 3D Heritage Online Presenter
BA: Bundle Adjustment
CDR: Cardinal Direction Relations
KB: Knowledge Base
LBA: Logical Bundle Adjustment
OWL: Web Ontology Language
QSR: Qualitative Spatial Reasoning
SPARQL: Protocol and RDF Query Language
RDF: Resource Description Framework

KB: Knowledge Base
A: Orange sphere
B: Fuchsia sphere
C: Green sphere
D: Yellow sphere
E: Aquamarine sphere
F: Blue sphere

ACKNOWLEDGMENT

The authors would like to warmly thank Mrs. Odile PAPINI for her pertinent proofreading and her unfailing friendship.

REFERENCES

1. D. Tingdahl and L. V. Gool, "A public system for image based 3D model generation," in Lecture Notes in Computer Science, vol. 6930/2011 of Computer Vision/- Computer Graphics Collaboration Techniques 5th International Conference, MIRAGE 2011, pp. 262–273, SpringerLink, 2011.
2. M. Vergauwen and L. V. Gool, "Web-based 3d reconstruction service," Mach. Vision Appl., vol. 17, no. 6, pp. 411–426, 2006.
3. J. L. Sch¨onberger and J.-M. Frahm, "Structure-from-motion revisited," in Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
4. J. L. Sch¨onberger, E. Zheng, M. Pollefeys, and J.-M. Frahm, "Pixelwise view selection for unstructured multi-view stereo," in European Conference on Computer Vision (ECCV), 2016.
5. A. Rahimian Boogar, H. Salehi, H. R. Pourghasemi, and T. Blaschke, "Predicting habitat suitability and conserving juniperus spp. habitat using svm and maximum entropy machine learning techniques," Water, vol. 11, 09 2019.
6. J. Eastwood, D. Sims-Waterhouse, S. Piano, R. Weir, and R. Leach, "Autonomous close-range photogrammetry using machine learning," ISMTII, vol. 1, 09 2019.
7. P. Borin and F. Cavazzini, "Condition assessment of rc bridges. integrating machine learning, photogrammetry and bim," ISPRS - International Archives of the Pho- togrammetry, Remote Sensing and Spatial Information Sciences, vol. XLII-2 W15, pp. 201–208, 08 2019.
8. J. Gong and S. Ji, "Photogrammetry and deep learning," Cehui Xuebao/Acta Geo- daetica et Cartographica Sinica, vol. 47, pp. 693–704, 06 2018.
9. P. Gaia, Approche ontologique pour la mod´elisation et le raisonnement sur les tra- jectoires : pris en compte des aspects th´ematiques, temporels et spatiaux. PhD thesis, University of La Rochelle, France, 2020.
10. L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 152, pp. 166 – 177, 2019.
11. E. Colucci, F. Noardo, F. Matrone, A. Spano, and A. LINGUA, "High-level-of-detail semantic 3d gis for risk and damage representation of architectural heritage," IS- PRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLII-4, pp. 107–114, 09 2018.
12. M. Belgiu and T. Lampoltshammer, "Ontology based interpretation of very high resolution imageries – grounding ontologies on visual interpretation keys," in 16th AGILE conference on Geographic Information Science, 05 2013.
13. F. Noardo, A Spatial Ontology for Architectural Heritage Information, pp. 143–163. Cham: Springer International Publishing, 2017.
14. C. Heipke, M. Madden, Z. Li, and I. Dowman, "Theme issue "state-of-the-art in photogrammetry, remote sensing and spatial information science"," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 115, pp. 1–2, 05 2016.
15. C. Blut and J. Blankenbach, "Three-dimensional citygml building models in mobile augmented reality: a smartphone-based pose tracking system," International Journal of Digital Earth, vol. 0, no. 0, pp. 1–20, 2020.
16. R. Clouard, A. Renouf, and M. Revenu, "An ontology-based model for representing image processing application objectives," International Journal of Pattern Recogni- tion and Artificial Intelligence, vol. 24, no. 08, pp. 1181–1208, 2010.
17. L. Albers, P. Große, and S. Wagner, Semantic Data-Modeling and Long-Term Interpretability of Cultural Heritage Data—Three Case Studies, pp. 239–253. Cham: Springer International Publishing, 2020.
18. K. Vavliakis, G. Karagiannis, and P. Mitkas, "Semantic web in cultural heritage after 2020, "what will the semantic web look like 10 years from now?"," in Workshop held in conjunction with the 11th International Semantic Web Conference 2012 (ISWC 2012), Boston, USA., At Boston, USA, 01 2012.
19. M. Bla˘sko, R. Cacciotti, P. K˘remen, and Z. Kouba, "Monument damage ontology," in Progress in Cultural Heritage Preservation (M. Ioannides, D. Fritsch, J. Leissner, R. Davies, F. Remondino, and R. Caffo, eds.), (Berlin, Heidelberg), pp. 221–230, Springer Berlin Heidelberg, 2012.
20. S. Manzoor, S.-H. Joo, Y. G. Rocha, H. uk Lee, and T.-Y. Kuc, "A novel semantic slam framework for humanlike high-level interaction and planning in global environment," The 1st International Workshop on the Semantic Descripto, vol. 1, 2019.
21. C.-L. Kuo, Y.-M. Cheng, Y.-C. Lu, Y.-C. Lin, W.-B. Yang, and Y.-N. Yen, "A frame- work for semantic interoperability in 3d tangible cultural heritage in taiwan," in Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection (M. Ioannides, E. Fink, R. Brumana, P. Patias, A. Doulamis, J. Martins, and M. Wallace, eds.), (Cham), pp. 21–29, Springer International Publishing, 2018.
22. M. Weinmann, B. Jutzi, S. Hinz, and C. Mallet, "Semantic point cloudinterpretation based on optimal neighborhoods, relevant features and efficient classifiers," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 105, pp. 286 – 304, 2015.
23. N. Maillot and M. Thonnat, "Ontology Based Complex Object Recognition," Image and Vision Computing, vol. 26, no. 1, pp. pp 102–113, 2008.
24. P.-Y. Rabattu, B. Mass´e, F. Ulliana, M.-C. Rousset, D. Rohmer, J.-C. L´eon, and O. Palombi, "My Corporis Fabrica Embryo: An ontology-based 3D spatio-temporal modeling of human embryo development," Journal of Biomedical Semantics, vol. 6, pp. 36:1–15, 2015.
25. B. Neumann and R. M¨oller, On Scene Interpretation with Description Logics, pp. 247–275. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006.
26. E. Kalogerakis, S. Christodoulakis, and N. Moumoutzis, "Coupling ontologies with graphics content for knowledge driven visualization," in IEEE Virtual Reality Con- ference (VR 2006), pp. 43–50, March 2006.
27. N. Vitucci and G. Gini, "Reasoning on objects and grasping using description logics," Advanced Robotics, vol. 33, no. 13, pp. 616–635, 2019.
28. L. F. Sikos, "Vidont: a core reference ontology for reasoning over video scenes," Journal of Information and Telecommunication, vol. 2, no. 2, pp. 192–204, 2018.
29. B. Neumann and R. M¨oller, "On scene interpretation with description logics," Image Vision Comput., vol. 26, p. 82–101, november 2008.
30. K. Kraus, Photogrammetry vol 1 and 2, vol. 2 of Advanced method and applications. Bonn, Germany: Dummlers, dummlerbush ed., 1997.
31. S. Skiadopoulos and M. Koubarakis, "Composing cardinal direction relations," Arti- ficial Intelligence, vol. 152, no. 2, pp. 143–171, 2004.
32. J. Chen, D. Liu, C. Zhang, and Q. Xie, "Combinative reasoning with rcc5 and cardinal direction relations," in International Conference on Knowledge Science, Engineering and Management, pp. 92–102, Springer, 2007.
33. R. Goyal and M. J. Egenhofer, "Cardinal directions between extended spatial objects," IEEE Transactions on Knowledge and Data Engineering, vol. 21, pp. 22–31, 2000.
34. F. Christian, "Qualitative spatial reasoning," in Cognitive and linguistic aspects of geographic space, pp. 361–372, Springer, 1991.
35. S. Harris, A. Seaborne, and E. Prud'hommeaux, "Sparql 1.1 query language," W3C recommendation, vol. 21, no. 10, p. 778, 2013.
36. M. D. Emiliani, La question de la perspective. In La perspective comme forme sym- bolique, Erwin Panofsky. Les E´ ditions de Minuit, 1961.
37. ASP, Manual of photogrammetry, Fourth Edition. American Society of Photogram- metry(Asprs Pubns), 1980.
38. E. Kruppa, Zur Ermittlung eines Objektes aus zwei Perspektiven mit innerer Orien- tierung. H¨older, 1913.
39. C. Kuratowski and A. Tarski, "Les op´erations logiques et les ensembles projectifs," Fundamenta Mathematicae, vol. 17, no. 1, pp. 240–248, 1931.
40. A. Tarski, "On the calculus of relations," The journal of symbolic logic, vol. 6, no. 3, pp. 73–90, 1941.
41. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Com- put. Vision, vol. 60, pp. 91–110, november 2004.
42. H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in Computer Vision – ECCV 2006 (A. Leonardis, H. Bischof, and A. Pinz, eds.), (Berlin, Heidelberg), pp. 404–417, Springer Berlin Heidelberg, 2006.
43. M. Donoser and H. Bischof, "Efficient maximally stable extremal region (mser) track- ing," in 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), vol. 1, pp. 553–560, June 2006.
44. A. S´luzek, "Multi-distinctive mser features and their descriptors: A low-complexity tool for image matching," in Advanced Concepts for Intelligent Vision Systems (S. Battiato and P. Scheunders, eds.), (Cham), pp. 672–680, Springer International Publishing, 2015.
45. D. C. Brown, "A solution to the general problem of multiple station analytical stereo- triangulation," RCA MTP Data Reduction Technical Report AFMTC TR 59-25, vol. 54, 1958.
46. D. C. Brown, "The bundle adjustment - progress and prospects," International archive of photogrammetry, vol. 21, no. 3, pp. 0–33, 1976.
47. B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, "Bundle adjust- ment — a modern synthesis," in Vision Algorithms: Theory and Practice (B. Triggs, A. Zisserman, and R. Szeliski, eds.), (Berlin, Heidelberg), pp. 298–372, Springer Berlin Heidelberg, 2000.
48. R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision. Cam- bridge University Press, 2 ed., 2004.
49. M. I. A. Lourakis and A. A. Argyros, "Sba: a software package for generic sparse bundle adjustment," ACM Transactions on Mathematical Software, vol. 1, pp. 1–30, 2009.
50. R. H¨ansch, I. Drude, and O. Hellwich, "Modern Methods of Bundle Adjustment on the Gpu," ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. III3, pp. 43–50, jun 2016.
51. H. Mayer, "Rpba – robust parallel bundle adjustment based on covariance informa- tion," 2019.
52. M. Cao, L. Zheng, W. Jia, and X. Liu, "Fast incremental structure from motion based on parallel bundle adjustment," Journal of Real-Time Image Processing, vol. ., 04 2020.
53. D. A. Cucci, M. Rehak, and J. Skaloud, "Bundle adjustment with raw inertial obser- vations in uav applications," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 130, pp. 1 – 12, 2017.
54. H. Liu, M. Chen, G. Zhang, H. Bao, and Y. Bao, "Ice-ba: Incremental, consistent and efficient bundle adjustment for visual-inertial slam," in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1974–1982, 2018.
55. J. F. Allen, "Maintaining knowledge about temporal intervals," Commun. ACM, vol. 26, pp. 832–843, november 1983.

56. P. R. Florence, L. Manuelli, and R. Tedrake, "Dense object nets: Learning dense visual object descriptors by and for robotic manipulation," CoRR, vol. abs/1806.08756, 2018.
57. G. Ligozat, "Reasoning about cardinal directions," Journal of Visual Languages and Computing, vol. 9, pp. 23–44, 1998.
58. M. B. Ellefi, P. Drap, L. Garcia, F. Garreau, C. Lef`evre, O. Papini, I. St´ephan, and E. Wu¨rbel, "Query answering with non-monotonic rules: A case study of archaeology qualitative spatial reasoning," in GCAI 2019. Proceedings of the 5th Global Conference on Artificial Intelligence, vol. 65 of EPiC Series in Computing, pp. 94–107, 2019.
59. H. W. Guesgen, Spatial reasoning based on Allen's temporal logic. International Com- puter Science Institute Berkeley, 1989.
60. Trav´e-Massuy`es, "Conceptual neighborhood and its role in temporal and spatial rea- soning," in ..., 1991.
61. A. Cohn and J. Renz, "Qualitative spatial reasoning, handbook of knowledge repre- sentation," 2007.
62. F. Van Harmelen, V. Lifschitz, and B. Porter, Handbook of knowledge representation, vol. 1. Elsevier, 2008.
63. E. M. Mikhail, J. S. Bethel, and J. C. McGlone, "Introduction to modern photogram- metry," New York, vol. 1, 2001.
64. A. L. Ali, Z. Falomir, F. Schmid, and C. Freksa, "Rule-guided human classification of volunteered geographic information," ISPRS journal of photogrammetry and remote sensing, vol. 127, pp. 3–15, 2017.
65. P. E. Santos, G. Ligozat, and M. Safi-Samghabad, "An occlusion calculus based on an interval algebra," in 2015 Brazilian Conference on Intelligent Systems (BRACIS), pp. 128–133, IEEE, 2015.
66. J. F. Allen, "Maintaining knowledge about temporal intervals," in Readings in quali- tative reasoning about physical systems, pp. 361–372, Elsevier, 1990.
67. A. U. Frank, "Qualitative spatial reasoning: Cardinal directions as an example," In- ternational Journal of Geographical Information Science, vol. 10, no. 3, pp. 269–290, 1996.
68. T. Berners-Lee, J. Hendler, O. Lassila, et al., "The semantic web," Scientific american, vol. 284, no. 5, pp. 28–37, 2001.
69. M. Ben Ellefi and P. Drap, "Semantic export module for close range photogrammetry," in The Semantic Web: ESWC 2019 Satellite Events (P. Hitzler and R. Verborgh, eds.), (Cham), pp. 3–7, Springer International Publishing, 2019.
70. E. Kruck, "The world of bingo, web page," Available (accessed 10.06. 2020): http://bingo-atm. de/ , vol. 1, 2020.
71. Photomodeler, "Photomodeler, web page," Available (accessed 10.06. 2020): https:// www. photomodeler. com/ / , vol. 1, 2020.
72. A. cp., "Agisoft, web page," Available (accessed 10.06. 2020): https:// www.agisoft. com/ , vol. ., .
73. OpenCV, "Opencv, web page," Available (accessed 10.06. 2020): https: // opencv. org/ , vol. 1, 2020.
74. M. Wafa, Automatizing the large-scale analysis of underwater optical data. PhD thesis, University of Pisa, Italy, 2015.
75. P. Drap and J. Lef`evre, "An exact formula for calculating inverse radial lens distor- tions," Sensors, vol. 16, p. 807, 06 2016.
76. N. Guarino, D. Oberle, and S. Staab, "What is an ontology?," in Handbook on on- tologies, pp. 1–17, Springer, 2009.
77. M. Potenziani, M. Callieri, M. Dellepiane, M. Corsini, F. Ponchio, and R. Scopigno, "3dhop: 3d heritage online presenter," Computers and Graphics, vol. 52, pp. 129 – 141, 2015.
78. N. Snavely, S. M. Seitz, and R. Szeliski, "Skeletal graphs for efficient structure from motion," in Proc. Computer Vision and Pattern Recognition, 2008.
79. S. Agarwal, N. Snavely, S. M. Seitz, and R. Szeliski, "Bundle adjustment in the large," in Computer Vision – ECCV 2010 (K. Daniilidis, P. Maragos, and N. Paragios, eds.), (Berlin, Heidelberg), pp. 29–42, Springer Berlin Heidelberg, 2010.
80. C. Sweeney, T. Sattler, T. H¨ollerer, M. Turk, and M. Pollefeys, "Optimizing the viewing graph for structure-from-motion," in 2015 IEEE International Conference on Computer Vision (ICCV), pp. 801–809, Dec 2015.
81. P. Drap, "The arpenteur project web site: http://www.arpenteur.org," 2017.
82. A. J. Fuseki, "Apache jena org, web page," Available (accessed 06.04. 2018): https:// jena. apache. org/ documentation/ fuseki2 , vol. 1, 2020.
83. M. Saleem, G. Sza´rnyas, F. Conrads, S. A. C. Bukhari, Q. Mehmood, and A.-C. Ngonga Ngomo, "How representative is a sparql benchmark? an analysis of rdf triple- store benchmarks?," in The World Wide Web Conference, pp. 1623–1633, ACM, 2019.
84. 84. K. Rohloff, M. Dean, I. Emmons, D. Ryder, and J. Sumner, "An evaluation of triple- store technologies for large data stores," in OTM Confederated International Con- ferences" On the Move to Meaningful Internet Systems", pp. 1105–1114, Springer, 2007.
85. G. A. Atemezing and F. Amardeilh, "Benchmarking commercial rdf stores with publi- cations office dataset," in European Semantic Web Conference, pp. 379–394, Springer, 2018.
86. M. Saleem, Q. Mehmood, and A.-C. N. Ngomo, "Feasible: A feature-based sparql benchmark generation framework," in International Semantic Web Conference, 2015.
87. M. Husain, J. McGlothlin, M. M. Masud, L. Khan, and B. M. Thuraisingham, "Heuristics-based query processing for large rdf graphs using cloud computing," IEEE Transactions on Knowledge and Data Engineering, vol. 23, no. 9, pp. 1312–1327, 2011.
88. P. Peng, L. Zou, M. T. O¨ zsu, L. Chen, and D. Zhao, "Processing sparql queries over distributed rdf graphs," The VLDB Journal—The International Journal on Very Large Data Bases, vol. 25, no. 2, pp. 243–268, 2016.

AUTHORS PROFILE

Mohamed BEN ELLEFIis a doctor in computer science with a focus in modeling knowledge and automatic reasoning systems for artificial intelligence (AI). He worked in R&D projects in French research laboratories, in particular LIRMM Montpellier and LIS Marseille. Within the Khresterion R&D team, he is currently working on the implementation of artificial intelligence systems inkeeping a scientific watch on advanced technologies related in particular to knowledge graphs.

Pierre DrapA senior researcher at the CNRS, he works at the LIS laboratory (Laboratoire d'Informatique & Systèmes) UMR CNRS 7020 at the of Aix-Marseille University. His main research activities concern close range photogrammetry applied to underwater archaeology, marine biology and medieval archaeology. He develops a knowledge base approach to photogrammetry and works on the semantisation of photogrammetric survey. Head of the Image & Models team he has coordinated or participated during the last twenty years in more than 25 international or national projects funded by Europe, the French Agency for Scientific Research (ANR) and other ministries.