

A Qualitative Model for Reasoning about 3D Objects using Depth and Different Perspectives

Zoe Falomir

Cognitive Systems (CoSy) Department
Spatial Cognition Centre
Universität Bremen

Enrique-Schmidt-Str. 5, 28359 Bremen, Germany
Email: zfalomir@informatik.uni-bremen.de

Abstract—A qualitative model for describing 3D objects (*Q3D*) using depth and different perspectives is presented in this paper. The front, right and up perspectives are considered as canonical. The *Q3D* model allow reasoning through logics defined to test the consistency of descriptions. The maximal volume of the object is also obtained logically using its *Q3D* description. Moreover, this model infers some features of the unknown perspectives of the object by defining logics based on the continuity of holes and the relative depth presented by opposite perspectives. The *Q3D* logics are implemented in Prolog and promising results are obtained, which can inspire approaches to solve 3D spatial problems computationally.

I. INTRODUCTION

QUALITATIVE Spatial and Temporal Representations and Reasoning (QSTR) [1]–[3] models and reasons about *time* (i.e. coincidence, order, concurrency, overlap, granularity) and also about properties of *space* (i.e. topology, location, direction, proximity, geometry, intersection, etc.) and their evolution between continuous neighbouring situations. Maintaining the consistency in space and time are the basics in qualitative reasoning when solving spatial and temporal problems. Spatio-temporal reasoning models deal with imprecise and incomplete knowledge on a symbolic level and have been successful in many areas and applications such as robotics [4], [5], computer vision [6], [7], ambient intelligence [8], [9], 2D shape description and recognition [10], colour naming and similarity [11], architecture and design [12], spatial query solving in geographic information systems [13], [14], etc. Furthermore, qualitative representations are thought to be closer to the cognitive domain, as shown in cognitive models of sketch recognition [15], spatial problem solving tasks (i.e. visual oddity tasks) [16]. However, further research is still needed to combine more aspects of QSTR with cognitive spatial thinking.

In the fields of computer vision, robotics and ambient intelligence, 3D object description and recognition are challenging tasks nowadays. Dealing with three dimensional data is a challenge because they usually suffer from distortions due to noisy sensors, viewpoint changes and point density variations. In the computer vision literature, approaches for object recognition usually use 3D descriptors to encode their shapes from different perspectives [17], [18]: feature-based approaches describe the local or global properties of the surface of the object (i.e., colour, curvature, texture, etc.); graph-based approaches describe the structure or skeleton of the object, that is, the relations between the object parts; and other approaches use other techniques like extended gaussian images, 3D moments, volumetric errors, etc.

Research in the field of 3D object recognition has been fostered by the availability of low-cost depth cameras based on structured infrared light (also called RGB-Depth cameras) such as the Microsoft Kinect and the Asus Xtion¹. Since the development of these sensors, diverse techniques have appeared to recognise real objects which learn their shape from the thousands of points which describe their surface from different perspectives [19]–[21]. Although these techniques are successful and applied in robotics and ambient intelligent systems, they are quite computational expensive, and they are not exploiting constraints in space to reduce this cost.

In the field of psychology, spatial cognition studies have demonstrated that there is a strong link between success in Science, Technology, Engineering and Math (STEM) disciplines and spatial abilities [22], [23]. Thus, it is important to maintain and train these abilities from the early stages. For example, children at 4 years old have already informal awareness of spatial relations such as parallel relations for two dimensional shape identification before they are properly taught about parallelism [24]. For this reason, researchers in US and Canada study the actualities and possibilities of training/including spatial reasoning in contemporary school mathematics [25], also because spatial learning and reasoning can be taught easily using visual and kinetic interactions offered by new digital technologies [26]. For example, touchscreen digital devices can facilitate geometrical expression for young children [27]. High spatial skills are also required in space teleoperation [28] (mental rotation and perspective-taking strategies are proved to be used by the operator-astronaut to move a robot arm around the workspace) and they are also decisive in Medicine [29].

Moreover, in cognitive psychology, games like *Upside Down World* are used to evaluate students' spatial skills when they are challenged to recreate buildings composed of multilink cubes and to use spatial language to describe the composition of these buildings so that their colleagues can build accordingly [25]. A test of the German Academic Foundation to find children with gifted brains among candidates for scholarships consists in finding out the consistent view/projection for a 3D object usually corresponding to a technological drawing².

This paper explores the challenge of describing 3D objects qualitatively and it is based on the levels of depth each object

¹Trade and company names are included for benefit of the reader and imply no endorsement or preferential treatment of the product by the author.

²Test der Studienstiftung: Gehirnjogging für Hochbegabte, see Spiegel Online: <http://www.spiegel.de/quiztool/quiztool.249771.html>

has at each perspective. This approach is inspired in designs of pieces which abstract the main features of the object from all their properties in the real world and describe them using 3 canonical views (top, lateral and front). Moreover, this approach is cognitively based, since in experimental psychology there is support for the general idea that human object recognition involves view-dependent representations, that is, people prefer to imagine, view, or photograph objects from certain “canonical” views [30]. Also this approach has been motivated by the fact that the German Academic Foundation uses consistent view/projection of a 3D object corresponding to a technological drawing to measure intelligence in humans². An example of a question in this test and the instructions given may be that provided in Fig. 1. Note that this example is made up for this paper to avoid copyright issues, and that real examples can be obtained online².

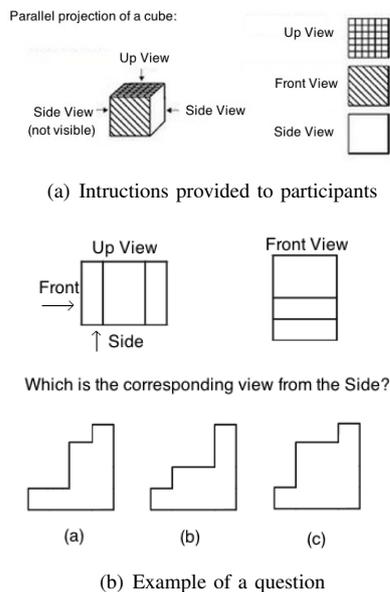


Fig. 1. (a) Instructions of the test translated to English; (b) Plausible example of a question regarding 3D projections in the German Academic Foundation test.

The rest of the paper is organised as follows. In Section II, properties of spatial substrates are explained. Section III presents a model for Qualitative Description of 3D objects. Section IV explains the logics encountered, and the consistency conditions for the $Q3D$ model are described in Section V. Section VI presents a logic approach to obtain the maximal volume of an object described by the $Q3D$ model. Section VII explains how to infer some features of the occluded views (back, down and left) from their opposite views (front, up and right). In Section VIII, the implementation of the model is described. Section IX discusses the closer related work. And conclusions and future work are presented in Section X.

II. SPATIAL SUBSTRATES AND THEIR PROPERTIES

As Freksa [31] mentions, *properties of spatial objects and configurations are intrinsically highly interdependent. If we modify one spatial aspect (e.g. distance, orientation, topological relation) in a spatial structure, other spatial aspects will be*

changed automatically, as well. We call such a structure a spatial substrate. If we move an object in space, the spatial locations of all its parts as well as their relations to other objects will change. If we change a single spatial aspect in a spatial substrate, all these changes take place (for free); no computing (or otherwise) effort is required.

As far as we are concerned, there is no related literature about which are the properties spatial substrates may have. Here, we appeal to the intuition of the reader to formulate some properties which we envision they help in solving spatial problems:

- **Abstraction:** people abstract dimensions in space (i.e., by assuming one dimension as constant) and re-represent data in a way that helps visualising a problem. For example, a map represents 3D space in a 2D paper, sometimes assuming relief or altitude as constant.
- **Continuity:** dimensions in space are continuous. They can be abstracted or considered as constant in a representation, but this representation must be coherent with the space and transmit changes in the dimension abstracted, if produced. For example, if a change in relief is produced (i.e., a road is cut) this change should be transmitted to the dimensions not abstracted (i.e., an interactive or up-to-date map should represent this discontinuity in the road).
- **Interrelation:** most dimensions in space are relative or interrelated to each other. For example, when comparing roads in a map, people usually look for the shorter-path (wrt another) or the quicker path (wrt another). If the roads are represented by abstracting the same dimension, then they can be compared directly. If one road considers relief while the other does not, then they are not comparable.

In 3D engineering object design (see Fig. 2), objects are usually **abstracted or re-represented** using 3 canonical views. In each view, the object is abstracted by considering a dimension as constant. For example, in the *front view/perspective*, the dimensions involved are the width and height of the object, while the depth dimension is assumed as constant; in the *right view*, the dimensions depth and height are represented, whereas the width dimension is assumed as constant; and in the *up view*, the dimensions represented are width and depth, while height is assumed as constant.

Note that, in contrast to the 3D projection test by the German Academic Foundation where views are provided disconnected from each other, in 3D technological drawings, engineers assume **continuity** in their abstractions or re-representations of the object. When assuming a constant value for a dimension, it is assumed also that this dimension is continuous. If a change is produced in the dimension abstracted, this change has to be reflected in the other representations. In Fig. 2 this continuity is represented as grey lines.

Moreover, as Fig. 2 shows, in 3D technological drawings, perspectives are **relative** to each other. For example, the dimension height is involved in the views front and right; similarly, the dimension depth is involved in the views right and up; and the dimension width is involved in the views up and front. Therefore, following the continuity principle, a change in each common dimension must be reflected in the other two views involved.

After observing these properties in the spatial substrates, the following model for qualitative 3D object description was defined.

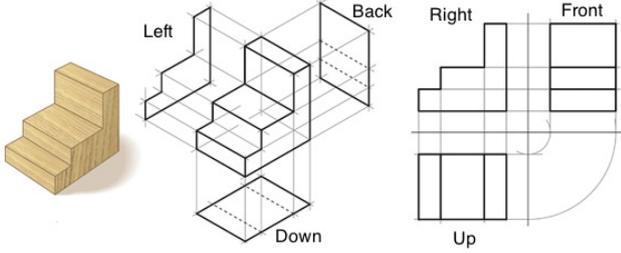


Fig. 2. Example of a 3D object in a technological drawing which shows the corresponding relationships among perspectives.

III. A QUALITATIVE DESCRIPTOR FOR 3D OBJECTS

When thinking qualitatively about real objects in space, humans usually think about volumes. For example, in pictures and paintings, observers assume sometimes depth in objects/scenes –differentiating foreground from background [11]– which is not easy to see in the absence of shadows. As a consequence, the minimal unit for the qualitative description presented here is considered a volume, specifically a cube of side $x \in \mathbb{R}$, which may be used to build an object similarly to how pixels are used to build digital images.

Therefore, a reference system for qualitative 3D object description is defined as follows:

$$Q3D_{RS} = \{F, R, U \in P \mid P \subseteq N_{depths}\}$$

$$N_{depths} = \{a, b, c, d, \dots, *\}$$

where F, R and U are the Front, Right and Up perspectives (P) or views of the object, and N is the total number of cubes which compose each edge of the object. That is, the edges of the object in each perspective are described by the volume of cubes of equal size, being the basic unit of measure considered a cube of side $x \in \mathbb{R}$ (i.e., $x = 1\text{cm}$, $x = 0.75\text{cm}$, $x = 5\text{m}$, etc.).

Thus, each perspective has N levels of depth, which can be named differently and sequentially as $\{a, b, c, d, \dots, *\}$ where a is the surface of the cube, b is the first level of depth (a previous cube in the row has been removed), c is the second level of depth (two previous cubes in the row have been removed) and so on, until $*$ is reached, which indicates that all the cubes in a row have been removed. The description is started from the upper-left part at each perspective.

As a first example, let us consider the object in Fig. 3 and its corresponding description according to the views: Front (F) in red, Right (R) in blue, and Up (U) in yellow. Starting from the

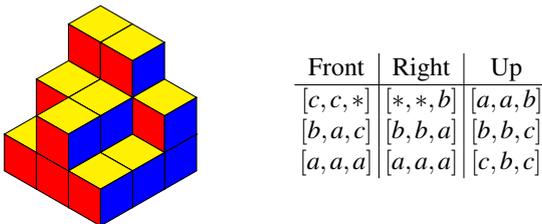
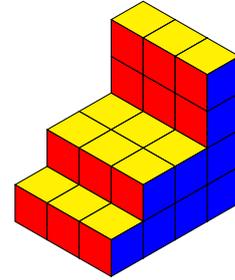


Fig. 3. Example of 3D object divided by a $3 \times 3 \times 3$ grid of cubes showing the front (red), right (blue) and up (yellow) views, and its corresponding Q3D description.

upper-left part of the *front* perspective, it can be observed that

2 cubes were removed in the first row, and also in the second row, so this is represented by the parameters c, c in the Q3D description. Then, all the cubes have been removed in the third row, so this is represented by the parameter $*$. Going down a level, it can be observed that only one cube is left in the first row (represented by b), then all the cubes are filling the second row (represented as a) and, in the third row, two cubes are missing (represented as c). Finally, in the basis of the object, all the rows are complete, which is represented as a, a, a . The perspectives *right* and *up* are explained similarly.

As a second example, let us consider the object in Fig. 4 extracted from the technological drawing in Fig. 2. The proportions of the object show that it can be modelled by a grid of $4 \times 4 \times 3$ cubes to be described qualitatively according to the different levels of depth at each perspective. Fig. 4 shows its corresponding Q3D description according to all the possible views.



Right	Front	Up	Left	Back	Down
$[*, *, *, a]$	$[d, d, d]$	$[d, c, c, a]$	$[a, *, *, *]$	$[a, a, a]$	$[a, a, a, a]$
$[*, *, *, a]$	$[d, d, d]$	$[d, c, c, a]$	$[a, *, *, *]$	$[a, a, a]$	$[a, a, a, a]$
$[*, a, a, a]$	$[b, b, b]$	$[d, c, c, a]$	$[a, a, a, *]$	$[a, a, a]$	$[a, a, a, a]$
$[a, a, a, a]$	$[a, a, a]$		$[a, a, a, a]$	$[a, a, a]$	

Fig. 4. Three dimensional object representation extracted from the technological drawing in Fig. 2 which can be divided into a $4 \times 4 \times 3$ grid of cubes to be described qualitatively by the Q3D approach.

It is important to notice that a change in a parameter or letter in the Q3D description in Fig. 4 corresponds to a line of the sketch drawing in Fig. 2. This is easily seen in the *up* perspective descriptor, where a line may be drawn vertically separating all d letters from c letters and another line may be drawn vertically separating c letters from a letters (since each different letter correspond to a different depth), so that the technological drawing related to up perspective shown in Fig. 2 would be obtained. Therefore, some hints about the **shape** of the object are obtained. However, note that the complete shape of the object is not described at this stage, and also that circular or squared holes in an object would be represented equally by $*$, described by the change in depth they produce.

IV. Q3D LOGICS FOR DESCRIBING OBJECTS

The Q3D description of an object can be also described logically, as follows:

$$\forall X \ Q3DObject(X) \rightarrow view(front, X, N, N', Q3D) \wedge view(right, X, N, N', Q3D) \wedge view(up, X, N, N', Q3D) \quad (1)$$

where X is a particular object; N is the dimension in cubes of the edge of the object; and $Q3D$ is the qualitative description corresponding to each of the perspectives *front*, *right* and *up*, which is built by N lists of N' elements of depth each.

The Q3D logic description for the object in Fig. 3 is provided as follows:

$$\begin{aligned} Q3DObject(object_1) \rightarrow & \\ & view(front, object_1, 3, 3, [[c, c, *], [b, a, c], [a, a, a]]) \wedge \\ & view(right, object_1, 3, 3, [[*, *, b], [b, b, a], [a, a, a]]) \wedge \\ & view(up, object_1, 3, 3, [[a, a, b], [b, b, c], [c, b, c]]) \end{aligned} \quad (2)$$

The Q3D logic description for the object in Fig. 4 is provided as follows:

$$\begin{aligned} Q3DObject(object_2) \rightarrow & \\ & view(front, object_2, 4, 3, \\ & \quad [[d, d, d], [d, d, d], [b, b, b], [a, a, a]]) \wedge \\ & view(right, object_2, 4, 4, \\ & \quad [[*, *, *, a], [*, *, *, a], [*, a, a, a], [a, a, a, a]]) \wedge \\ & view(up, object_2, 3, 4, \\ & \quad [[d, c, c, a], [d, c, c, a], [d, c, c, a]]) \end{aligned} \quad (3)$$

V. REASONING WITH THE Q3D: CONSISTENT AND INCONSISTENT PERSPECTIVES

According to spatial reasoning, from the perspectives Front (F), Right (R) and Up (U), an object can be built in a three-dimensional space. In mechanical engineering, it is assumed as a convention that this canonical views correspond to the more detailed views. So, which are the common sense facts in spatial reasoning which guide this building? What are the 3D spatial facts which can or cannot happen?

Let us consider the representation in Fig. 5 to exemplify the following cases:

- Case 1: a change in an edge affects 2 perspectives at least. For example, if the cube $\{F_{1,2}, U_{3,2}\}$ disappears, this must be reflected at both perspectives F and U.
- Case 2: a change in a vertex affects 3 perspectives. For example, if the cube $\{F_{1,3}, R_{1,1}, U_{3,3}\}$ disappears, this must be reflected at perspectives Front and Up, but also at Right.
- Case 3: each hole affects 2 perspectives at least, two of them corresponding to opposite views. For example, a hole in the middle of the object (i.e., cube $\{F_{2,2}\}$ and followers disappear) would affect Front and Back perspectives, whereas a hole involving cubes $\{F_{2,3}, R_{2,1}, R_{2,2}, R_{2,3}\}$ would affect 3 perspectives: Front, Right and Back.

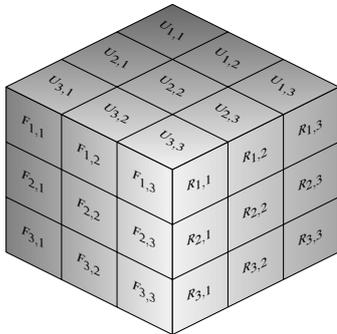


Fig. 5. Example of an object showing the constraints at the boundary of the canonical perspectives.

The spatial constraints appear along the boundary of the perspectives or the edges of the object, since a change in a perspective must be consistent with a change in another

perspective. In Fig. 5, each cube is named according to the perspectives Front (F), Right (R) and Up (U). Therefore, the descriptions must be consistent where the edges meet at F-R, F-U, and R-U perspectives:

$$\begin{aligned} consistent_Q3D(F, R, U) \rightarrow & \\ & consistent_perspective(F, R) \wedge \\ & consistent_perspective(F, U) \wedge \\ & consistent_perspective(R, U) \end{aligned} \quad (4)$$

Note that the problem is simplified by abstracting one dimension/view in each comparison, that is, the views meeting at each edge are those related and those that must be consistent.

Let us consider the edges meeting at cube $\{F_{1,3}, R_{1,1}, U_{3,3}\}$, then the consistent conditions for front (F) and right (R) perspectives can be defined as:

$$\begin{aligned} consistent_perspective([F_{11}, F_{12}, F_{13}], [F_{21}, F_{22}, F_{23}], [F_{31}, F_{32}, F_{33}], \\ [R_{11}, R_{12}, R_{13}], [R_{21}, R_{22}, R_{23}], [R_{31}, R_{32}, R_{33}]) \rightarrow \\ consistent_side(F_{13}, [R_{11}, R_{12}, R_{13}]) \wedge \\ consistent_side(F_{23}, [R_{21}, R_{22}, R_{23}]) \wedge \\ consistent_side(F_{33}, [R_{31}, R_{32}, R_{33}]) \wedge \\ consistent_side(R_{11}, [F_{11}, F_{12}, F_{13}]) \wedge \\ consistent_side(R_{21}, [F_{21}, F_{22}, F_{23}]) \wedge \\ consistent_side(R_{31}, [F_{31}, F_{32}, F_{33}]) \end{aligned} \quad (5)$$

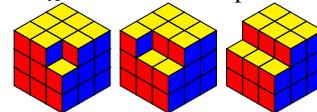
The conditions to obtain a consistent perspective in the sides R-U and F-U are defined similarly. It has been observed that the same constraints must be fulfilled for each edge, F-R, R-U and F-U, so they can be generalised:

$$\begin{aligned} consistent_side(F_{i3}, [R_{i1}, R_{i2}, R_{i3}]) \rightarrow \\ level_a(F_{i3}) \wedge level_a(R_{i1}) \end{aligned} \quad (6)$$

The logic rule (6) is explained as follows: if any cube exists on the right edge at *front* perspective, then it must exist also a cube on the left edge at *right* perspective, since a cube involves a volume which continues in both dimensions or perspectives. Note that *i* means *some row* and *1,2,3* means *column 1,2,3*, respectively. Note also that nothing is constrained on the R_{i2} and R_{i3} cubes.

$$\begin{aligned} consistent_side(F_{i3}, [R_{i1}, R_{i2}, R_{i3}]) \rightarrow \\ level_b(F_{i3}) \wedge level_a(R_{i2}) \wedge \\ (level_b(R_{i1}) \vee level_c(R_{i1}) \vee no_exist(R_{i1})) \end{aligned} \quad (7)$$

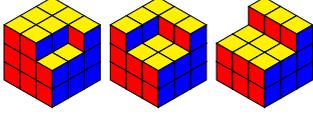
The logic rule (7) is explained as follows: if only a cube disappear on the right edge at *front* perspective (that is, $F_{i3} \equiv b$ in Q3D), then it must exist also a cube located on the second column at *right* perspective, since this is the cube seen from the *front* (that is, $R_{i2} \equiv a$). And the constraints on the cubes located on the left edge at *right* perspective (R_{i1}) are that: they cannot be level *a* of depth since that would mean that missing cubes, appeared again, which is inconsistent. For R_{i1} all the other possibilities can happen:



Note that nothing is constrained on the R_{i3} cubes.

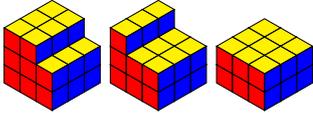
$$\begin{aligned}
 & \text{consistent_side}(F_{i3}, [R_{i1}, R_{i2}, R_{i3}]) \rightarrow \\
 & \text{level_c}(F_{i3}) \wedge \text{level_a}(R_{i3}) \wedge \\
 & (\text{level_b}(R_{i1}) \vee \text{level_c}(R_{i1}) \vee \text{no_exist}(R_{i1})) \wedge \\
 & (\text{level_b}(R_{i2}) \vee \text{level_c}(R_{i2}) \vee \text{no_exist}(R_{i2}))
 \end{aligned} \quad (8)$$

The logic rule (8) is explained as follows: if two cubes disappear on the right edge at *front* perspective (that is, $F_{i3} \equiv c$ in Q3D), then it must exist also a cube located at the third column at *right* perspective, since this is the cube seen from the *front* (that is, $R_{i3} \equiv a$). And the constraints on the cubes located on the left edge at *right* perspective (R_{i1} and R_{i2}) are that: they cannot be level *a* of depth since that would mean that the missing cubes appeared again, and that is inconsistent. For R_{i1} and R_{i2} all the other possibilities can happen:



$$\begin{aligned}
 & \text{consistent_side}(F_{i3}, [R_{i1}, R_{i2i}, R_{i3i}]) \rightarrow \\
 & \text{no_exist}(F_{i3}) \wedge \\
 & (\text{level_b}(R_{i1}) \vee \text{level_c}(R_{i1}) \vee \text{no_exist}(R_{i1})) \wedge \\
 & (\text{level_b}(R_{i2i}) \vee \text{level_c}(R_{i2i}) \vee \text{no_exist}(R_{i2i})) \wedge \\
 & (\text{level_b}(R_{i3i}) \vee \text{level_c}(R_{i3i}) \vee \text{no_exist}(R_{i3i}))
 \end{aligned} \quad (9)$$

The logic rule (9) can be explained as follows: if all the cubes disappear on the right edge at *front* perspective (that is, $F_{i3} \equiv *$ in Q3D), then no cube on the first row at right perspective must exist ($R_{i1} \neq a$, $R_{i2} \neq a$, $R_{i3} \neq a$) but all the rest of possibilities can happen for R_{i1} , R_{i2} and R_{i3} :

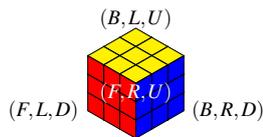


Note that F_{i3} denotes $F_{1,3}$, $F_{2,3}$ or $F_{3,3}$; that R_{i1} denotes $R_{1,1}$, $R_{1,2}$ or $R_{1,3}$; also R_{2i} denotes $R_{2,1}$, $R_{2,2}$ or $R_{2,3}$; and R_{3i} denotes $R_{3,1}$, $R_{3,2}$ or $R_{3,3}$, and also,

$$\begin{aligned}
 & \forall X \text{ level_a}(X) \rightarrow a \\
 & \forall X \text{ level_b}(X) \rightarrow b \\
 & \forall X \text{ level_c}(X) \rightarrow c \\
 & \forall X \text{ no_exist}(X) \rightarrow *
 \end{aligned} \quad (10)$$

If information is given about the rest of perspectives (Back -B-, Left -L-, Down -D-), the consistency conditions between the edges at each perspective are defined similarly.

Note that, as each vertex is proving consistency in 3 edges, only by proving the consistency conditions in 4 opposite vertices in the cube, all the edges of the cube are covered. Let us show an example:



Proving the consistency at the 4 vertices in the drawing above ($\text{consistent_Q3D}(F,R,U)$, $\text{consistent_Q3D}(B,L,U)$, $\text{consistent_Q3D}(F,L,D)$, $\text{consistent_Q3D}(B,R,D)$) is enough to cover the 12 edges of a complete consistent description of a 3D object.

The computational complexity of the consistency algorithm is calculated as follows. When choosing 4 opposite vertices in the cube where to apply the consistent_Q3D function, all the 12 edges of the cube are covered, since the consistent_Q3D function checks the consistency of the 3 edges meeting at a specific vertex. Then, the final computational complexity is 12 times the complexity of the $\text{consistent_perspective}$ function. And the complexity of this function is $2N$ where N is the edge size in volume-cubes. In summary, $12 \cdot 2 \cdot N = 24 \cdot N$, thus the computational cost is $\mathcal{O}(N)$.

VI. INFERRING THE MAXIMAL VOLUME OF THE OBJECT FROM THE Q3D

The maximal volume of an object described by a Q3D can be obtained as:

$$\text{Q3D}_{\text{volume}} = \min(\text{volume}_P(F), \text{volume}_P(R), \text{volume}_P(U)) \quad (11)$$

where $\text{Q3D}_{\text{volume}}$ refers to the volume of the object measured in cubes of side $x \in \mathbb{R}$; and \min refers to the minimum of the volumes corresponding to each perspective $F, R, U \in P$, that is (volume_P) which is defined as follows.

The volume of a perspective is the opposite to its levels of depth. If an object is described by N cubes, then the volume is calculated as:

$$\text{volume}_P(P) = \sum_{i=1}^{N-N} \text{volume}(\sigma) \quad (12)$$

where, the volume of an element σ is defined as:

$$\text{volume}(\sigma) = \sum_{i=1}^{\sigma} N - i \quad (13)$$

that is, for example, for $N=3$, $\text{volume}(a) \equiv N$, $\text{volume}(b) \equiv N - 1$, $\text{volume}(c) \equiv N - 2$ and $\text{volume}(*) \equiv 0$.

As an example, the volume of the object in Fig. 3 is calculated as follows:

Object	Front view	Right view	Up view
if $N=3$,	$[c, c, *] = 1 + 1 + 0$	$[*, *, b] = 0 + 0 + 2$	$[a, a, b] = 3 + 3 + 2$
$a=3, b=2$	$[b, a, c] = 2 + 3 + 1$	$[b, b, a] = 2 + 2 + 3$	$[b, b, c] = 2 + 2 + 1$
$c=1, *=0$	$[a, a, a] = 3 + 3 + 3$	$[a, a, a] = 3 + 3 + 3$	$[c, b, c] = 1 + 2 + 1$
	= 17	= 18	= 17

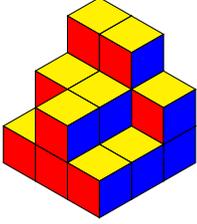
Note that the minimal result obtained is 17, which is the correct volume of the object, as it can be checked in Fig. 3.

The minimum volume of all canonical perspectives is calculated, since, for example, some holes may only be seen from a specific perspective. Then, it is important to notice that the volume obtained is the maximal that the object can have when being observed from the FRU perspective. Note that it is important to select the canonical perspectives as FRU, otherwise the correct volume might not be obtained since an object might have a hole at a side which would not be appreciated.

The computational cost of the maximal volume algorithm is obtained as follows. The complexity of calculating the volume of a Q3D view is the complexity of getting the value of the N depths at each row and the N depths at each column, being N the size of the edge, thus the cost is $N \cdot N$. This volume must be computed using the 3 views at a vertex (i.e., FRU), that is $3 \cdot N \cdot N$, so the computational complexity of the algorithm is $\mathcal{O}(N^2)$.

VII. INFERRING SOME DEPTHS IN UNKNOWN PERSPECTIVES FROM OPPOSITE VIEWS

Taking into account the spatial relations showed by the drawing in Fig. 6, from the views Front (F), Right (R) and Up (Up), how much can we deduce logically from the rest of the object? Can the rest of the views be computed?



Front	Right	Up
$[c, c, *]$	$[*, *, b]$	$[a, a, b]$
$[b, a, c]$	$[b, b, a]$	$[b, b, c]$
$[a, a, a]$	$[a, a, a]$	$[c, b, c]$

Fig. 6. Q3D description of object 1 used to explain inferences.

According to the properties of continuity and relativity of spatial substrates, some features of the unknown views can be inferred:

Hole Continuity: all the holes observed in a view affect the opposite views of the object, that is, front-back, right-left, up-down. Taking into account this property, the following features could be inferred regarding the *back*, *left* and *down* views of object 1:

Back	Left	Down
$[*, _, _]$	$[_, *, *]$	$[_, _, _]$
$[_, _, _]$	$[_, _, _]$	$[_, _, _]$
$[_, _, _]$	$[_, _, _]$	$[_, _, _]$

Note that features that remain unknown are represented by ‘ $_$ ’.

Depth Relativity: features indicating the last level of depth in a view (i.e., level c in a $3 \times 3 \times 3$ description), involve the existence of the first level of depth in the opposite view (always level a). Taking into account this property, the following features could be inferred regarding the *back*, *left* and *down* views of object 1:

Back	Left	Down
$[*, a, a]$	$[_, *, *]$	$[_, _, _]$
$[a, _, _]$	$[_, _, _]$	$[a, _, _]$
$[_, _, _]$	$[_, _, _]$	$[a, _, _]$

The consistency properties mentioned in Section V must be fulfilled also by the unknown perspectives and they can be used to infer the depths in them. For example, the neighbouring perspectives of *back* are *right*, *up*, and *left*. If *Right* and *Up* perspectives are known (R, U) or given by a *Q3D*, some features regarding the perspective *Left* can be inferred. These inference inter-relationships between neighbouring perspectives to discover more features in unknown views are currently under study.

VIII. IMPLEMENTATION

First-order logic knowledge bases are usually built using Horn clauses [32], which contains at most one positive literal. Prolog programming language [33] is based on Horn clause logic and it was selected as the logic programming language for implementing the logics of the *Q3D* description. SwiProlog³ was the testing platform [34], and the Prolog Contest book [35] was a guide.

³SWI-Prolog: <http://www.swi.2prolog.org/>

The *Q3D* description of objects was written using Prolog facts as: `view(View, Object, N, Q3D)`.

For example, the *Q3D* description of the object in Fig. 3 is described as:

```
view(front,obj1,3,[[c,c,*],[*,*,b],[a,a,b]]).
view(right,obj1,3,[[b,a,c],[b,b,a],[b,b,c]]).
view(up,obj1,3,[[a,a,a],[a,a,a],[c,b,c]]).
```

The correctness of the input *Q3D* was checked. The consistency logics were also programmed and tested. The maximal volume of the objects regarding the *Q3D* was also programmed and tested. And the inference of some features of the unknown perspectives from their opposite perspectives were also programmed and checked.

As an example, the results of the Prolog implementation for the *Q3D* description of the object in Fig. 3 are given:

```
?- qualitative_3D(object1).
Front:[[c,c,*],[b,a,c],[a,a,a]] Correct description.
Right:[[*,*,b],[b,b,a],[a,a,a]] Correct description.
Up:[[a,a,b],[b,b,c],[c,b,c]] Correct description.
Consistent Q3D F,R,U views.
Maximal volume:17
Back constrained wrt Front:[[*,a,a],[a,?,?],[?,?,?]]
Left constrained wrt Right:[[?,*,*],[?,?,?],[?,?,?]]
Down constrained wrt Up:[[?,?,?],[a,?,?],[a,?,a]]
true.
```

More examples of the testings are provided in the Appendix. All the Prolog code corresponding to the *Q3D* is available for downloading⁴. For easily testing, the on-line platform Pengines⁵ can be used.

IX. DISCUSSION ABOUT RELATED WORK

In the literature, objects are also described using 3D shape grammars [36]. As linguistic grammars build sentences and paragraphs, shape grammars follow also rules (i.e., recursively subdivision) to build 3D objects. In these grammars, although the rules applied are logical, the obtained description of the object is not qualitative, in contrast to the one proposed in this paper.

Moreover, another approach related to shape grammars is constructive solid geometry [37] (or computational binary solid geometry), that is, a technique used in solid modelling which can define the steps of building/synthesising complex objects by combination of other objects and Boolean operations (i.e., intersection, union, difference). It has a broad application in computer graphics for generating objects in computer games [38].

Both shape grammars and constructive solid geometry methods are useful for object building/synthesis, but challenging to use for object description/analysis because they sometimes use not-reversible operations. Moreover, there is not a specific set of grammar rules or constructive geometry methods to obtain a specific object, since different methods can produce the same result. Therefore, the descriptions obtained might be not unique and then difficult to use for object identification. The *Q3D* model defined here can be useful for designing the plan to

⁴Data download: <https://sites.google.com/site/zfalomir/projects/cognitive.2ami>

⁵Pengines by SWI-Prolog: <http://pengines.swi.2prolog.org/apps/swish/index.html>

synthesise/build the object, but also to uniquely describe that object when it is created.

In the literature, the main theoretical approaches in qualitative 3D representation which are studied in psychology of object perception are:

- Marr and Nishihara’s approach [39] which uses a 3D composition of generalised cylinders to describe a sketch or skeleton of the objects and their parts. This model is hierarchical, that is, component parts can also be decomposed into parts and recognition is achieved when matching a description derived from an image to a previous stored 3D object type. The *Q3D* approach presented in this paper is similar to Marr and Nishihara’s approach [39] in the sense that it uses a generalised cube to describe the structure of the object, similarly to Marr and Nishihara’s cylindre. However the *Q3D* approach represented the whole object, not only its skeleton.
- Biederman’s approach [40] describes 3D objects using more geometric shapes or geons, not only cylinders or cubes. However, they are obtained from a 2D image representation rather than from a 3D representation as in Marr and Nishihara’s approach [39]. According to Biederman, geons are detected on the basis of certain properties of contours in the image (i.e., linearity, parallelism, curvilinearity, symmetry) or at regions of concavity. Therefore, Biederman’s approach tries also to represent the different shapes of the components of the object, not only the volume parts.
- Guesgen’s approach [1] approximate objects to polygons with parallel sides which are projected to a coordinate axis. The relations between the objects (or intervals in the axis) are represented qualitatively (i.e., left of, attached to, overlapping, inside) similarly to the relations between temporal intervals defined by Allen’s model [41]. This model is similar to the *Q3D* in the sense that both approximate objects, Guesgen’s approach to polygons and the *Q3D* to arrangements of cubes. However, the *Q3D* uses the cube as a unit which allows to calculate the volume of the object, whereas Guesgen’s method do not obtain it, but it is independent of it.

The approaches above describe objects based on their 3D structural skeletons or sides and produce object centred descriptions. That is, view-independent descriptions are obtained which are not designed to detect inconsistencies in object perception from different views. The *Q3D* approach presented here describes the depth of 3D objects in its canonical views. The representation obtained is object-centred, but allows comparison between perspectives in order to detect inconsistencies and also in order to infer unknown perspectives, which is a novel aspect in the literature, as far as we are concerned. Moreover, there is support for the general idea that human object recognition involves view-dependent representations, that is, people prefer to imagine, view, or photograph objects from certain “canonical” views [30]. Therefore, the *Q3D* has a cognitive basis.

X. CONCLUSION AND FUTURE WORK

This paper presents the definition of a qualitative 3D object descriptor based on a qualitative concept of depth which considers a cube as the minimal unit of volume. This representation

can be considered a spatial substrate [31], since if a single cube is added or removed at an edge, this change is produced for free to 3 perspectives (i.e., *front*, *right* and *up*) without no computing effort needed to readjust the views.

Abstraction, continuity and interrelation properties are proposed in this paper as the basis to define spatial substrates. The logics to test the consistency of the Front, Right and Up *Q3D* descriptions corresponding to real 3D objects are presented. Then, the maximal volume of an object is calculated logically from the *Q3D* obtained. Moreover, logics to infer some features of the unknown back, down and left views are proposed.

All the *Q3D* logics described above have been implemented in Prolog and tested using the SwiProlog platform. Results are promising and they inspire future work towards an approach which could give humans a hint about which projections of a 3D object are impossible when solving spatial problems, so that they could understand and reason about 3D object representations such as those in the test by the German Academic Foundation for scholarships².

As future work, it is intended to: (i) implement further reasoning methods to infer the rest of features of depth of the unknown *back*, *left* and *down* perspectives, from the data known regarding its neighbouring perspectives; (ii) extend the *Q3D* description to include hidden concavities; and (iii) define an approach to describe the boundary shape of each of the perspectives of the object taking into account the *Q3D* as a basis.

Applications in education are envisioned when helping students in engineering to understand conventions in technical drawing. Other applications in computer vision would be interesting, for example when computing 3D attention saliency in proto-objects [42], a *Q3D* description could help to store a short memory narrative of the evolution of the proto-objects in these attention artificial systems.

ACKNOWLEDGMENTS

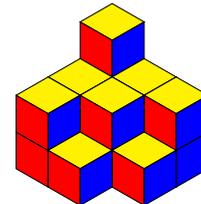
Dr.-Ing. Zoe Falomir acknowledges funding by the project COGNITIVE-AMI (GA 328763) by the European Commission through FP7 Marie Curie IEF actions and the support by the Universität Bremen and the Spatial Cognition Centre.

The author also thanks the anonymous reviewers’ comments which helped to improve this paper and the plane partition latex style fonts by Jang Soo Kim.

APPENDIX

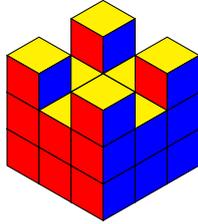
Other examples of *Q3Ds* implemented in Prolog and used in the testings.

Object 2:



Front	Right	Up	Back	Left	Down
$[c, *, *]$	$[*, *, c]$	$[a, b, b]$	$[*, *, a]$	$[a, *, *]$	$[a, a, a]$
$[a, b, c]$	$[c, b, a]$	$[b, b, c]$	$[a, a, a]$	$[a, a, a]$	$[a, a, a]$
$[a, a, b]$	$[b, a, a]$	$[b, c, *]$	$[a, a, a]$	$[a, a, a]$	$[*, a, a]$

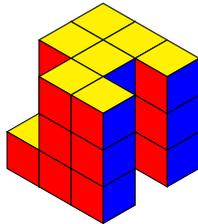
```
?- qualitative_3D(object2).
Front: [[c,*,*],[a,b,c],[a,a,b]] Correct description.
Right: [[*,*,c],[c,b,a],[b,a,a]] Correct description.
Up: [[a,b,b],[b,b,c],[b,c,*]] Correct description.
Consistent Q3D F,R,U views.
Maximal volume:15
true .
```



Object 3:

Front	Right	Up	Back	Left	Down
$[a,*,a]$	$[a,*,a]$	$[a,b,a]$	$[a,*,a]$	$[a,*,a]$	$[a,a,a]$
$[a,a,a]$	$[a,a,a]$	$[b,b,b]$	$[a,a,a]$	$[a,a,a]$	$[a,a,a]$
$[a,a,a]$	$[a,a,a]$	$[a,b,a]$	$[a,a,a]$	$[a,a,a]$	$[a,a,a]$

```
?- qualitative_3D(object3).
Front: [[a,*,a],[a,a,a],[a,a,a]] Correct description.
Right: [[a,*,a],[a,a,a],[a,a,a]] Correct description.
Up: [[a,b,a],[b,b,b],[a,b,a]] Correct description.
Consistent Q3D F,R,U views.
Maximal volume:22
true .
```



Object 4:

Front	Right	Up	Back	Left	Down
$[b,a,a]$	$[a,b,a]$	$[a,a,a]$	$[a,a,a]$	$[a,a,b]$	$[a,a,a]$
$[b,a,a]$	$[a,b,a]$	$[a,a,*]$	$[a,a,a]$	$[a,a,b]$	$[*,a,a]$
$[a,a,a]$	$[a,b,a]$	$[c,a,a]$	$[a,a,a]$	$[a,a,a]$	$[a,a,a]$

```
?- qualitative_3D(object4).
Front: [[b,a,a],[b,a,a],[a,a,a]] Correct description.
Right: [[a,b,a],[a,b,a],[a,b,a]] Correct description.
Up: [[a,a,a],[a,a,*],[c,a,a]] Correct description.
Consistent Q3D F,R,U views.
Maximal volume:22
true .
```

REFERENCES

- [1] H. W. Guesgen, "Spatial Reasoning Based on Allen's Temporal Logic," International Computer Science Institute, Tech. Rep., 1989.
- [2] A. G. Cohn and J. Renz, *Qualitative Spatial Reasoning, Handbook of Knowledge Representation*, V. L. F. Harmelen and B. Porter, Eds. Wiley-ISTE, London: Elsevier, 2007.
- [3] G. Ligozat, *Qualitative Spatial and Temporal Reasoning*. Wiley-ISTE, London: MIT Press, 2011.
- [4] L. Kunze, C. Burbridge, and N. Hawes, "Bootstrapping probabilistic models of qualitative spatial relations for active visual object search," in *Qualitative Representations for Robots, Proc. AAAI Spring Symposium, Technical Report SS-14-06*, 2014, pp. 81–80, ISBN 978-1-57735-646-2.
- [5] Z. Falomir, L. Museros, V. Castelló, and L. Gonzalez-Abril, "Qualitative distances and qualitative image descriptions for representing indoor scenes in robotics," *Pattern Recognition Letters*, vol. 38, pp. 731–743, 2013. [Online]. Available: <http://dx.doi.org/10.1016/j.patrec.2012.08.012>
- [6] Z. Falomir, E. Jiménez-Ruiz, M. T. Escrig, and L. Museros, "Describing images using qualitative models and description logics," *Spatial Cognition and Computation*, vol. 11, no. 1, pp. 45–74, 2011. [Online]. Available: <http://dx.doi.org/10.1080/13875868.2010.545611>
- [7] A. Cohn, D. Hogg, B. Bennett, V. Devin, A. Galata, D. Magee, C. Needham, and P. Santos, "Cognitive vision: Integrating symbolic qualitative representations with computer vision," in *Cognitive Vision Systems*, ser. Lecture Notes in Computer Science, H. Christensen and H.-H. Nagel, Eds. Springer Berlin/Heidelberg, 2006, vol. 3948, pp. 221–246.
- [8] M. Bhatt and F. Dylla, "A qualitative model of dynamic scene analysis and interpretation in ambient intelligence systems," *I. J. Robotics and Automation*, vol. 24, no. 3, 2009. [Online]. Available: <http://dx.doi.org/10.2316/Journal.206.2009.3.206.23274>
- [9] Z. Falomir and A.-M. Olteşanu, "Logics based on qualitative descriptors for scene understanding," *Neurocomputing*, vol. 161, pp. 3–16, 2015. [Online]. Available: <http://dx.doi.org/10.1016/j.neucom.2015.01.074>
- [10] Z. Falomir, L. Gonzalez-Abril, L. Museros, and J. Ortega, "Measures of similarity between objects from a qualitative shape description," *Spatial Cognition and Computation*, vol. 13, pp. 181–218, 2013. [Online]. Available: <http://dx.doi.org/10.1080/13875868.2012.700463>
- [11] Z. Falomir, L. Museros, and L. Gonzalez-Abril, "A model for colour naming and comparing based on conceptual neighbourhood. An application for comparing art compositions," *Knowledge-Based Systems*, vol. 81, pp. 1–21, 2015. [Online]. Available: <http://dx.doi.org/10.1016/j.knsys.2014.12.013>
- [12] M. Bhatt and C. Freksa, "Spatial computing for design an artificial intelligence perspective," in *Studying Visual and Spatial Reasoning for Design Creativity*, J. S. Gero, Ed., 2015, pp. 109–127.
- [13] P. Fogliaroni, *Qualitative Spatial Configuration Queries. Towards Next Generation Access Methods for GIS*, ser. Dissertations in Geographic Information Science. IOS Press, 2013, ISBN 978-1614992486.
- [14] R. Al-Salman, "Qualitative spatial query processing: Towards cognitive geographic information systems," Ph.D. dissertation, University of Bremen, 2014, supervised by Prof. Christian Freksa (University of Bremen) and Prof. Christian Jensen (Aalborg University).
- [15] A. Lovett, M. Dehghani, and K. Forbus, "Learning of qualitative descriptions for sketch recognition," in *Proc. 20th Int. Workshop on Qualitative Reasoning (QR)*, Hanover, USA, July, 2006.
- [16] A. Lovett and K. Forbus, "Cultural commonalities and differences in spatial problem-solving: A computational analysis," *Cognition*, vol. 121, no. 2, pp. 281 – 287, 2011. [Online]. Available: <http://dx.doi.org/10.1016/j.cognition.2011.06.012>
- [17] I. Kazmi, L. You, and J. J. Zhang, "A Survey of 2D and 3D Shape Descriptors," in *Computer Graphics, Imaging and Visualization (CGIV), 2013 10th International Conference*, Aug 2013, pp. 1–10. [Online]. Available: <http://dx.doi.org/10.1109/CGIV.2013.11>
- [18] J. Tangelder and R. Veltkamp, "A survey of content based 3d shape retrieval methods," *Multimedia Tools and Applications*, vol. 39, no. 3, pp. 441–471, 2008. [Online]. Available: <http://dx.doi.org/10.1007/s11042.2007.20181.20>
- [19] T. Kluth and Z. Falomir, "Studying the role of location in 3D scene description using natural language," in *XV Workshop of the Association on Qualitative Reasoning and its Applications (JARCA13). Qualitative Systems and their applications to Diagnosis, Robotics and Ambient Intelligence*, I. Sanz, L. Museros, and J. A. Ortega, Eds. Murcia, Spain: Proceedings from the University of Seville, 2013, pp. 33–36.
- [20] K. Lai, L. Bo, X. Ren, and D. Fox, "Sparse Distance Learning for Object Recognition Combining RGB and Depth Information," in *IEEE International Conference on on Robotics and Automation*, 2011. [Online]. Available: <http://dx.doi.org/10.1109/ICRA.2011.5980377>
- [21] L. Bo, X. Ren, and D. Fox, "Depth kernel descriptors for object recognition," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2011, San Francisco, CA, USA, September 25-30*. IEEE, 2011, pp. 821–826. [Online]. Available: <http://dx.doi.org/10.1109/IROS.2011.6095119>
- [22] N. Newcombe, "Picture this: Increasing math and science learning by improving spatial thinking," *American Educator*, vol. 34, no. 2, pp. 29–35, 2010.
- [23] J. Wai, D. Lubinski, and C. P. Benbow, "Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance," *Journal of Educational Psychology*, vol. 101, no. 4, pp. 817–835, 2009. [Online]. Available: <http://dx.doi.org/10.1037/a0016127>
- [24] N. Sinclair, E. de Freitas, and F. Ferrara, "Virtual encounters: the murky and furtive world of mathematical inventiveness," *ZDM. The International Journal on Mathematics Education*, vol. 45, no. 2, pp. 239–252, 2013. [Online]. Available: <http://dx.doi.org/10.1007/s11858.2012.20465.23>
- [25] N. Sinclair and C. D. Bruce, "Spatial reasoning for young learners, Research Forum," in *Proc. of the 38th Conference of the International Group for the Psychology of Mathematics Education and the 36th Conference of the North American Chapter of the Psychology of Mathematics Education (PME 38 / PME-NA 36)*, Vancouver, Canada, July 2014, pp. 173–205.

- [26] K. Highfield and J. Mulligan, "The role of dynamic interactive technological tools in preschoolers' mathematical patterning," in *Proc. of the 30th annual conference of the Mathematics Education Research Group of Australasia*, J. Watson and K. Beswick, Eds., vol. 1. MERGA, 2007, pp. 372–381. [Online]. Available: <http://hdl.handle.net/1959.14/150138>
- [27] L. Museros, Z. Falomir, I. Sanz, and L. Gonzalez-Abril, "Sketch retrieval based on qualitative shape similarity matching: Towards a tool for teaching geometry to children," *AI Communications*, vol. 28, no. 1, pp. 73–86, 2014. [Online]. Available: <http://dx.doi.org/10.3233/AIC.2140614>
- [28] M. A. Menchaca-Brandan, A. M. Liu, C. M. Oman, and A. Natapoff, "Influence of perspective-taking and mental rotation abilities in space teleoperation," in *Proceedings of the ACM/IEEE International Conference on Human-robot Interaction*, ser. HRI '07. New York, NY, USA: ACM, 2007, pp. 271–278. [Online]. Available: <http://doi.acm.org/10.1145/1228716.1228753>
- [29] M. Hegarty, M. Keehner, C. A. Cohen, D. R. Montello, and Y. Lippa, "The role of spatial cognition in medicine: Applications for selecting and training professionals," in *Applied spatial cognition: From research to cognitive technology*, G. L. Allen, Ed. Mahwah, NJ: Lawrence Erlbaum, 2007, pp. 285–315.
- [30] S. Palmer, E. Rosch, and P. Chase, "Canonical perspective and the perception of objects," *Attention and Performance IX*, pp. 135–151, 1981.
- [31] C. Freksa, "Spatial computing – how spatial structures replace computational effort," in *Cognitive and linguistic aspects of geographic space*, M. Raubal, D. Mark, and A. F. (Eds.), Eds. Heidelberg: Springer, 2013.
- [32] J. W. Lloyd, *Foundations of logic programming. Symbolic computation: Artificial intelligence*. Springer-Verlag, 2nd, extended edition, 1987.
- [33] L. Sterling and E. Shapiro, *The Art of Prolog (2nd Ed.): Advanced Programming Techniques*. Cambridge, MA, USA: MIT Press, 1994.
- [34] J. Wielemaker, T. Schrijvers, M. Triska, and T. Lager, "SWI-Prolog," *Theory and Practice of Logic Programming (TPLP)*, vol. 12, no. 1-2, pp. 67–96, 2012. [Online]. Available: <http://dx.doi.org/10.1017/S1471068411000494>
- [35] B. Demoen, P.-L. Nguyen, T. Schrijvers, and R. Tronçon, *The First Programming Contests*, Belgium, 2005. [Online]. Available: <http://www.cs.kuleuven.be/~dtai/ppcbook/>
- [36] A. I.-K. Li, L. Chen, Y. Wang, and H. Chau, "Editing Shapes in a Prototype Two-and Three-dimensional Shape Grammar Environment," in *Computation: The New Realm of Architectural Design: 27th eCAADe Conference Proceedings*, ser. eCAADe: Conferences. Istanbul, Turkey: Istanbul Technical University, Faculty of Architecture, 2009, pp. 243–250.
- [37] A. A. G. Requicha, "Representations for rigid solids: Theory, methods, and systems," *Computing Surveys*, vol. 12, no. 4, pp. 437–464, Dec. 1980.
- [38] S. van Rossen and M. Baranowski, *Real-Time Constructive Solid Geometry*. A K Peters/CRC Press, 2011, pp. 79–96. [Online]. Available: <http://dx.doi.org/10.1201/b10946.211>
- [39] D. Marr and H. K. Nishihara, "Representation and Recognition of the Spatial Organization of Three-Dimensional Shapes," *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 200, no. 1140, pp. 269–294, 1978.
- [40] I. Biederman, "Recognition-by-components: A theory of human image understanding," *Psychological Review*, vol. 94, pp. 115–147, 1987.
- [41] J. F. Allen and L. F. Allen, "Maintaining knowledge about temporal intervals," *Communication of ACM*, pp. 832–843, 1983.
- [42] P. Lanillos, J. F. Ferreira, and J. Dias, "Multisensory 3D saliency for artificial attention systems," in *Proc. of the 3rd Workshop on Recognition and Action for Scene Understanding (REACTS)*, Malta, G. Azzopardi, F. Escolano, and R. Marfil, Eds., 2015, pp. 1–14, ISBN 978-84-606-9592-9.