A Visual Interface for Querying Ontologically and Socially Annotated 3D Worlds for the Web

Fabio Pittarello†
Università Ca’ Foscari Venezia - DAIS

Ivano Gatto†
Università Ca’ Foscari Venezia - DAIS

ABSTRACT
This work proposes a visual interface for querying sets of 3D worlds for the web, designed by different authors and annotated with both ontological top-down and social bottom-up styles. The goal is to overcome the current situation where most of these worlds are like independent islands that can be accessed only using the modalities that have been designed by their authors. Usually these worlds can’t be searched, because there is no high level information associated to the 3D entities that compose them. Our proposal is based on a specification, a software architecture and a set of interfaces that permit a full exploitation of the geometric and textual information contained inside these 3D environments. The applications can be different, from the retrieval of single entities for localization purposes by end users to the retrieval of complex spatial patterns for functional or typological analyses by domain experts. A case study related to cultural heritage is presented.

Categories and Subject Descriptors
H.3.3 [Information Search and Retrieval]: Query formulation; H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities; H.5.3 [Group and Organization Interfaces]: Web-based interaction;

General Terms
Design, Human Factors

Keywords
annotation, cultural heritage, folksonomy, information finding, ontology, search, social web, visual languages, Web3D

1. INTRODUCTION
In spite of the increased computing power of personal devices, the full exploitation on the web of the potential of the 3D representation is constrained by the unavailability of a high level description of the geometric components that make the worlds themselves. That is the reason why in recent years a number of researchers have focused their work on the proposal of methodologies and specifications that permit to associate a semantics to the geometric description of the components of the 3D worlds. This work takes advantage of a specification, a software architecture and a set of interfaces for annotating 3D worlds that have been described in previous works [9] [10] and shows how such enhanced semantic description can be used for building a system that allows to search and browse sets of Web3D worlds. In particular this work extends the architecture defined in previous works, introducing a search module and a visual interface for cross-querying sets of annotated 3D worlds, giving the users the possibility to define both keyword-based and topological-based searches of objects and spaces. One of the innovative features of this work is that the visual interface allows to query worlds labeled with both ontology-based and social-based annotations. The aim is take advantage of the entire annotation effort for exploiting the potential of the information contained inside the 3D worlds, even for these worlds lacking official ontology-based annotations. The retrievable entities include not only solid objects but also immaterial spaces, that are a fundamental component of the 3D environment definition. The system permits to retrieve single entities or objects embedded in specific spatial contexts. For geolocalized 3D environments representing real scenes such feature may be used for a variety of purposes, from tourism to shopping (e.g., find all the objects labelled as bench embedded in spaces labelled as park or find all the spaces labelled as furniture shop embedded in spaces labelled as mall). Besides, the system supports the definition of the relation of adjacency between spaces. Such relation can be combined with that one expressing the containment and used for the retrieval of complex spatial patterns by experts interested in functional analyses [11] or in architectural and urban study research [1]. For example the system may be used for analyzing the points of strength and weakness of the spatial solutions of complex buildings such as hospitals or for analyzing the recurring typological patterns in the works of a given architect. Another significant feature of the system is that it allows to couple in the same session the searching and the browsing mechanisms, permitting users to shift between the two interfaces, how it happens for the hypertextual web [13] but not still for the 3D web.

2. RELATED WORK
Searching the components of a 3D world is a complex task that researchers have faced through different techniques, that we can roughly categorize as shape-based, keyword-based and spatial-based search. The shape-based search is focused on the retrieval of objects whose shape is similar to that one specified by the users in the query composition (e.g., [5]). The text-based approach is based
on the use of one or more keywords for retrieving annotated 3D objects. This approach requires a preliminary annotation of the components of the 3D database that must be searched. Most of the available text-based approaches are based on a top-down annotation style, characterized by a predefined set of labels assigned by a single author and referred to a taxonomy or to an ontology [6] [12]. Even though the top-down approach has significant points of strength, such as the precision of classification, it may suffer from lack of expressivity and from the existence of significant differences with the language of the end users. That is reason why in the last decade in the hypertextual web there have been a number of proposals for extending the possibility to classify information to end users. Folksonomies [17] are often characterized by informational noise, due to problems such as homonynies, synonyms and others. In spite of that, they permit to classify the resources with the language of end users and, for what concerns the so-called broad folksonomies, they permit to emphasize let the most used keywords for annotating the entities. Because each annotation style has specific points of strength and weaknesses, in recent years a number of authors have proposed, for the hypertextual web, integrated uses of top-down and bottom-up classification styles [16] [4].

In the domain of web 3D information there is a substantial lack of proposals for bottom-up classification styles and this prevents a full exploitation of the geometric information stored inside the 3D environment. That is the reason why in previous works [10] we proposed a unified specification, architecture and set of interfaces permitting to annotate sets of 3D worlds using both bottom-up and top-down classification styles. This paper completes the work done so far, introducing a search engine and an interface working on both types of annotation that is still missing for the 3D web.

Users are often interested in retrieving entities that are characterized by specific spatial relations. That is the reason why since a long time researchers have been defining languages and tools that enable to query databases specifying these properties. Most of the research on spatial query languages has been focused on the domain of geographical information system (GIS). Several approaches, such as SQL3 [15] and OpenGIS SQL [7], have extended SQL, the most widely known textual query language, for managing also spatial data. OpenGIS SQL [7], an extension of the SQL92 specification, is currently considered as the standard for manipulating spatial data. Textual languages for spatial queries are affected by the same problems that characterize traditional query languages: the syntax may be difficult to learn and to use, leading users to make errors while composing the query. Besides, textual languages translate spatial relationships into a non-spatial representation that can prevent an intuitive composition of the query. Alternative textual approaches based on the use of the natural language are appealing, but are characterized by verbosity and imprecision [2]. Visual languages represent an opportunity for bridging the gap between the query composition and the domain of interest, coping with the usability problems that characterize textual query languages. A variety of visual styles, including iconic, diagrammatic, graph-based and multimodal approaches, have been experimented in different contexts (e.g., [3]).

3. DESIGN CHOICES AND SOFTWARE ARCHITECTURE

In relation to the three searching techniques described in Section 2, our proposal is an hybrid approach that takes advantage of the keyword-based and the spatial-based search. Coupling both approaches, as it happens for the geometorphs described in [14], can give substantial benefits in terms of expressivity and precision of the search process. Concerning the keyword-based search, one of the novelties of our system is the possibility to query the database in relation to top-down and bottom-up annotation styles. Concerning the spatial-based search, in the design of the architecture we have considered different types of relations, related to the application domains described in Section 1 (e.g., tourism, shopping, typological and functional analyses). Others may be considered in future work.

Table 1: Grid representing the available topological relations for each combination of objects and spaces

<table>
<thead>
<tr>
<th>Object</th>
<th>Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object</td>
<td>Disjoint, Contains</td>
</tr>
<tr>
<td>Space</td>
<td>Disjoint, Within, Touches, Within, Contains</td>
</tr>
</tbody>
</table>

The OpenGIS specification [7] include the following topological relations: Equals, Disjoint, Touches, Within, Overlaps, Contains, Crosses and Intersects. While in other works all these relations have been considered, in our research we focused on the relations expressing concepts of adjacency, nidification and separation. Therefore we considered the Disjoint, Touches, Within and Contains relations, associating them to objects and/or spaces. Table 1 resumes the relations available for each combination of entities. The bold labels on the horizontal axis represent the first operand, while the bold labels on the vertical axis represent the second operand. The availability of a given relation as operator is displayed at the grid intersections. In spite of the reduced number of relations taken from the OpenGIS specification, we elaborate on them, introducing the relation that expresses the concept of indirect adjacency, useful to find the connections between spaces mediated by a user definable number of other spaces. Another feature that distinguishes our proposal from other ones, is the possibility to put in relation more than two entities in a single query, originating complex search patterns. We considered also metric relations, introducing the possibility to specify the Euclidean distance between entities, defined as the maximum distance that can separate them. Directional relations may be considered in future development of the system and targeted to geolocalized 3D worlds.

We chose X3D as the target language for describing 3D entities and annotations. X3D [19], ISO standard, is the most widely diffused and long-standing language for describing 3D worlds for the web. Its openness, expressivity and capability to host metadata were the reasons why we chose it. As extensively described in [9] [10], the specification for annotating 3D worlds was designed on the top of it, permitting to annotate preexisting worlds and preserving the possibility to browse annotated 3D worlds with a standard X3D browser. In [10] we described a software architecture built on the top of the specification. The software components support all the phases of the management of 3D worlds, from the initial upload and pre-processing, to the annotation and the navigation of the annotated 3D worlds. In this work we propose an enhanced version of this architecture, featuring a set of components and interfaces for managing the search of annotated 3D worlds. All the components of the software architecture have been implemented using web technologies. Because of the limits of the space, in this paper we will focus mainly on the description of the components related to the search functionality (see Fig. 1). As discussed in [10], the design choice of managing bottom-up annotations required a centralized management. 3D worlds are pre-processed at the upload time, for univocally identifying each entity that characterizes them, and stored into a repository. The metadata describing the
annotations are stored in a separate MYSQL database. Each time the user enters a 3D world and selects an entity belonging to it, the server components take care of retrieving from the MYSQL database all the associated annotations and pass them to the client interface for visualization. The design choice of using a separate database structure for the dynamic part of the X3D worlds grants a higher scalability of the system, enabling faster accesses to data and an easier management of concurrent queries. Registered users are enabled to annotate the entities of the 3D worlds. The top-down annotations are made in relation to a predefined domain ontology, specified in OWL [8], that may change according to the specific nature of the managed set of 3D worlds. The user that uploads the 3D world is entitled to use both top-down and bottom-up annotation styles. The other registered users can take advantage only of the bottom-up annotation style. Both categories of users can define new groups of entities and annotate them with the authorized top-down and/or bottom-up annotation styles, for improving the semantic specification of the 3D world. Fig. 1 displays the client and the server components involved in the query process. The queries are composed through a visual interface, described in Section 4, based on Javascript and HTML5. The queries are passed to the search component, residing on the server side, through an asynchronous mechanism based on AJAX. The search component on the server side searches the annotated 3D worlds for retrieving total or partial matchings of the keyword-based and spatial-based query. Besides, as it will be detailed in Section 4, for a better exploitation of the different user annotation styles, the search is extended also to terms similar to the keywords used in the query, taking advantage of WordNet [18], the biggest free lexical database for the English language developed by the Princeton University.

4. VISUAL ENVIRONMENT FOR QUERIES

Fig. 2 shows a snapshot of the visual environment for querying the set of 3D worlds. The interface is divided in four parts: (1) the iconic toolbar, representing the entities and the spatial relations; (2) the area for composing the query, dragging or selecting the icons of the toolbar; (3) the area for managing the properties of the entities and relations, after having selected their representations on the central area; (4) the footer area, where the results of the query are displayed. For the sake of comprehension, the visual environment is described in relation to the case study chosen for testing our system: a set of annotated buildings and villas of the Renaissance, designed by the famous Italian architect Andrea Palladio and modeled by the students of the Fine Arts Academy of Venice. These worlds represent a good testbed for our system in terms of quantity and quality, because they are complex in terms of entities and are characterized by spatial schemes suitable of interesting cross-world spatial analyses.

The toolbar is composed by a set of visual icons, representing the different types of entities and spatial relations. The first two icons are defined for representing single objects and spaces. The user can drag them from the toolbox to the central area of the interface for adding single entities to the query. The other icons represent the spatial relations. The third and the fourth icons represent the topological Touches and Indirectly Touches relations for specifying direct and indirect connections between spaces. In the latter case the user can specify, through the property area, the number of intermediate spaces. The fifth icon represents the metric relation expressing the maximum distance between entities. The relations Contains and Within are not associated to icons of the tool-
bar, but they can be specified nesting the representations of objects and spaces in the central area. Applying iteratively these visual operations the user can compose complex hierarchies. Overall, the composition of the different types of entities and spatial relations can originate very complex spatial queries. Moreover, the representations of objects and spaces can be associated, formulating the query, to top-down and bottom-up annotations. For example the user may specify an object belonging to the class chair and associated to the tag seat or a space belonging to the class room and associated to the tag my room. For improving the user awareness of the available classes and tags, a textual cloud of annotations is displayed (see Fig. 2, next to the toolbar) each time the mouse pointer clicks the toolbar icons representing the objects or the spaces. The textual cloud displays the classes of the ontology (in red) and the most used tags (in black) related the set of worlds belonging to the database. The user may specify annotated entities in the query just dragging these labels on the central area of the interface. For example, if the user drags the label courtyard belonging to the textual cloud of the spaces, the result will be the visualization on the central area of a squared shape representing a space, labelled as courtyard. Similar textual representations can be visualized on the properties area, for helping the users to manage the properties of entities that have already been positioned on the central area. The users, clicking on the yellow cross icons positioned after the available fields (e.g., class and tag for the objects and the spaces), can visualize textual clouds displaying the classes or the most used tags for that type of entity. The properties area can be used also for specifying additional parameters, such as the name of the authors that have inserted the textual annotations. The search will be restricted to the annotations of the specified author. In the case of the author held the yellow cross icon will allow to visualize the textual cloud of authors that have created the highest number of annotations. There are several applications where the specification of the author may be useful. For example students may restrict the search to the entities annotated by their teacher or by a given art historian.

The central area of the interface displayed in Fig. 2 shows how annotated entities and relations can be combined for composing a complex query. The system is asked to retrieve all the occurrences in the available sets of 3D where: there is a sequence of three directly connected spaces annotated with labels belonging to top-down and bottom-up styles: atrium, great staircase and courtyard; the space atrium contains an object of class column; the space courtyard contains an object tagged as fountain.

4.1 Query Results

The results can be progressively displayed as the user visually composes the query and changes the parameters that characterize the different entities. The bottom of Fig. 2 shows the list of results following the visual query. The first result string states that a complete matching of the query pattern has been found in the world Villa Badoer. The user can choose, by clicking over the link following the world name, to enter the annotated world embedding the results. As displayed in Fig. 3-1, a 3D camera will be dynamically generated for displaying all the entities that have been matched by the search engine. The user, clicking on the icon preceding each result, may also visualize details related to each retrieved space and object. The user can then click the link associated to each entity for being teleported in front of it. For example, the snapshot displayed in Fig. 3-2 shows the view dynamically created in the world Villa Badoer when the user clicks over the link associated to the entity fountain. The system has been designed for returning not only results corresponding to all the entities and relations specified in the query, but also partial results, ordering them according to the completeness of the answer. For example, the second string of results of the query (see Fig. 2) is related to a partial matching of the query, in the world Villa Emo, where not all the objects have been retrieved (i.e., the courtyard of Villa Emo doesn’t include a fountain). Fig. 3-3 displays the retrieved entities. The third string of results displayed in Fig. 2 is related to another partial matching, in the world Villa Pisani, where three spaces have been retrieved, but only two have the same labels specified in the query and the third one is labeled with court, synonym of the class courtyard requested in the query. Fig. 3-4 displays the retrieved entities. As stated in Section 3, this feature is obtained through the integration in the system of the WordNet lexical database. Such integration permits a better exploitation of the user annotations, enlarging the search to objects and spaces annotated with synonyms of the keywords specified by the user composing the query.

5. IMPLICIT COMPOSITION OF QUERIES

The user that needs to find objects and spaces stored in the 3D database may decide to begin a session composing a query through the visual interface or browsing the list of available worlds. As a result of the searching or browsing activity, the user will be teleported to the world and to the 3D entities of interest. The interface displaying the 3D worlds (Fig. 3-1) permits the user to perform different activities, from navigation to tagging. Most interestingly, the navigation interface doesn’t represent a dead end for the searching activity, but stimulates, as suggested in [13], the associative mechanisms typical of the serendipitous information finding. The user can select any object or space (i.e., spaces can be visualized - on request - as clickable semitransparent rectangles. A sample of this representation can be seen in the 3D area of Fig. 3-1. The selected object, a column, is highlighted through a bounding box that helps the user to distinguish it from the other entities of the 3D world. The selection of any annotation triggers the visualization of all the associated information on the right side of the interface. In Fig. 3-1 the selection of the label ionic column causes the visualization of the associated author Anna. The selection of any annotation causes also the execution of an implicit query, for retrieving all the occurrences of the annotations in the set of the 3D worlds belonging to the system. On the right side of the interface displayed in Fig. 3-1 are visualized all the occurrences of the objects labelled as ionic column. Besides, the system analyzes the context of the entity associated to the selected annotation and offers the additional possibility to trigger a query for finding all the occurrences with the same spatial context (e.g., in Fig. 3-1 all the occurrences of the objects labelled as ionic column in the spatial context of class atrium).

6. CONCLUSION

The work done so far has led to the design and implementation of an environment for searching and navigating a set of 3D worlds that can be annotated with different styles. Current work is focused on improving the level of integration between the query and the navigation interfaces that currently have been designed as related but separate interfaces. While in a future implementation we’ll experiment the integration in a single interface, the selection of 3D entities from the list of results represents anyway a situation of disruption in the user experience because of the time needed to download and to initialize the rendering of the associated 3D world. Of course, the use of an interface organized in a single web page would minimize the lag, at least for all those situations where the users se-
select different results associated to the same 3D world. Solutions for obtaining an immediate visualization of different worlds, such as multiple-view interfaces, are not easy to implement because they usually augment the computational load associated to the management of the 3D environment on the client, that in many cases is already heavy even for a single world. Previewing the 3D worlds through a set of snapshots can be a viable alternative that we are going to explore in our future work.

7. REFERENCES


