Guiding Lens-based Exploration using Annotation Graphs

M. Alhsan, F. Marton, R. Pintus, and E. Gobbetti
CRS4, Italy

Abstract
We introduce a novel approach for guiding users in the exploration of annotated 2D models using interactive visualization lenses. Information on the interesting areas of the model is encoded in an annotation graph generated at authoring time. Each graph node contains an annotation, in the form of a visual markup of the area of interest, as well as the optimal lens parameters that should be used to explore the annotated area and a scalar representing the annotation importance. Graph edges are used, instead, to represent preferred ordering relations in the presentation of annotations. A scalar associated to each edge determines the strength of this prescription. At run-time, the graph is exploited to assist users in their navigation by determining the next best annotation in the database and moving the lens towards it when the user releases interactive control. The selection is based on the current view and lens parameters, the graph content and structure, and the navigation history. This approach supports the seamless blending of an automatic tour of the data with interactive lens-based exploration. The approach is tested and discussed in the context of the exploration of multi-layer relightable models.

CCS Concepts
- Information systems → Web applications; • Human-centered computing → Graphical user interfaces; • Computing methodologies → Graphics systems and interfaces; • Applied computing → Arts and humanities;

1. Introduction
With the advent of digitization, we see a significant increase in the sophistication of digital models that represent physical world objects. Effective exploration tools should enable users to understand, appreciate and interact with these virtual models. To achieve this goal, the exploration process must support the in-depth analysis of the models under inspection beyond plain visual replication. These virtual objects must, in particular, be capable of representing useful information at various levels of detail and display semantic attributes (textual, visual drawings, or abstractions) without crowding the view.

Annotations, i.e., blocks/additional layers of useful information embedded into the digital model or linked to various parts of it, are often used to provide better insights to the user [PCDS20], as well as to direct users towards interesting areas during exploration [BAMG21]. This latter approach falls within the realm of guidance techniques, which provide assistance in response to information on user needs [CGM’16]. The interactive experience can be further improved by coupling annotations with an exploration planning perspective, in which authors can describe, in a flexible way, not only single annotations, but possible meaningful tours through the data [SH10, HD12]. This aspect is of particular interest in the Cultural Heritage (CH) domain, as such planning can be used to support pre-visit (documentation and planning), visit (immersion and enhancement), as well as post-visit (emotional linking) phases [KWLS14] using simplified walk-up-and-use interfaces.

Interactive visualization lenses, i.e., movable tools that provide alternative visual representations for interactively controlled selected regions of interest of a display, have proved to offer flexible solutions to dealing with a variety of complex visualization problems using simplified interfaces [TGK’17]. They have been very recently extended to exploring annotated models [BAMG21], but without taking into account relations among annotations. The challenge is not only to display individual annotations without cluttering the display [CCJ14], but also to communicate the auxiliary (geometric, conceptual and semantic) information that is attached to 2D/2.5D/3D models in an order that supports a coherent, but context- and user-dependent narrative.

In this work, we focus on generating exploration paths for the navigation of annotated models with interactive lenses. In particular, we assume that the data under inspection is 2D and has been enriched with visual annotations, which mark interesting areas in the dataset. Annotations in form of visual cues can come from automated analysis or manual mark-up, and help viewers understand data [VETL18]. While current available methods organize those elements in partially-structured or completely unstructured schemes [PCDS20, BAMG21], the goal of our work is to harness the potential of highly structured relation-dependent graphs, with nodes and edges containing not only an encoded representations of visual markup of areas of interest, but also relations among the individual annotations. The introduction of edges allow us to introduce storytelling features by constraining the annotation presenta-
tion order. At run-time, the graph is exploited to assist users in their navigation by determining the next best annotation to be presented in the database, and to move the lens to explore it. This choice depends on the current view and lens parameters, the graph content and structure, and the navigation history. This automatic advice can be used to automatically generate intuitive tours through the data that dynamically respond to user actions. In this context, our main contributions are in the definition of an annotation graph with precedence relation to describe interesting paths for the exploration of a model, in the definition of a scoring system for determining the next best annotation taking into account these relations, as well as in the exploitation of this graph to create a context-dependent tour through the data that can be activated and restarted at any time.

All these elements aim at enhancing the user experience in the exploration of annotated models. In particular, we test the presented approach within the Cultural Heritage (CH) domain, though it can be extended to a wider range of fields.

2. Related Work

Exploration of annotated models and using interactive lenses are broadly studied topics within the visualization community. In the following, we will analyze the approaches similar to ours. For a wide coverage of the topic we refer the reader to the established surveys on annotations [PCDS20], visualization lenses [TGK*17], and spatial interfaces [BYK*21].

Selecting and presenting relevant annotations in a comprehensible manner without clutter is one of the major challenges in effective visualization displays [CCJ14]. Displaying all of them at the same time is infeasible as it generates cluttering and cognitive overload [TGK*17]. Several attempts had been made to address the challenges of overcrowded displays and improving intuitive interaction. Some authors proposed to guide users towards interesting areas by controlling the camera, creating animation paths, or defining fixed video tours [FS97]. While others suggested enabling and disabling the visible interface categories [PCD*20], modifying the appearance, or distorting and zooming the images.

Serial temporal presentation to enhance the view [ED07] is one of the approaches that has been used to deal with overcrowded display, and, in conjunction with authoring or automatic determination of temporal precedence, it provides a way to deliver a narrative meaningful tours through the data [BRAMG15]. Manually writing or defining fixed key-frames and forcing a single path is one of the most adopted solutions [FS97], which has been used also by touring through annotations [WSA*18]. This approach, however, leads to the generation of static videos rather than interactive experiences.

Lenses are used widely in scientific and information visualization [TGK*17]. They support overview + details, focus+context and cue-based techniques [CKB09], and are often used to reduce clutter by sub-sampling [EBD05] or by selecting an annotation at a time [JVAP*21]. Bettio et al. [BAMG21] recently proposed a novel approach for assisting users in navigating with lenses, also exploiting data annotations. Their approach introduces a controller that guarantees maintenance of focus-and-context constraint by jointly adapting view- and lens-parameters, as well as a scheme to determine the next best annotation in the database based on the current view and lens parameters and the navigation history. We build on these prior approaches, significantly extending the annotation representation, moving from a simple flat list of annotations to an annotation graph, in which edges express semantic relationships among nodes, exploiting these relations for automatic data touring and guided suggestions. Annotation selection is based on a score that extends to annotated lens graphs the Degree-of-Interest (DOI) concept introduced by Furnas [Fur86] for trees and extended by van Ham and Perer [VHP09] to graphs. Similarly to Gladisch et al. [GST13], DOI computation also takes into account past behavior. The camera-control work of Balsa et al. [BRAMG15] is the most similar to ours, as it selects only a single item at a time from a viewpoint graph. In a different context, we also use a scoring system with a history term to help navigation. Our graph and scoring system is, however, targeted to support lens-based navigation of an annotated model.

3. Method

Our work targets the inspection of 2D datasets that have been enriched with visual annotations that mark and describe the areas of interest [PCDS20]. We assume that these annotations have been prepared at authoring time, and that, in addition the author has linked together annotations to define a preferred ordering of presentation, as described in Sec. 3.1. This annotation graph is used to support users in their navigation using an interactive lens, in particular through context-dependent automated tours that can be activated or stopped at any time, as well as intermixed with interactive exploration.

The automatic navigation starts upon request or when the user lens is steady for a while. The automatic tour finds the next best annotation, using a scoring system described in Sec. 3.2. The lens and rendering parameters are then smoothly changed to the ones defined by the annotation. The user can then possibly explore the annotated area by interactively controlling the lens or other parameters (e.g., lighting), or just wait for the next annotations to be selected. As each annotation selection takes into account the current lens configuration and changed rendering parameters, the selected layer to be visualized inside the lens, the history of visited annotations, and the loose precedence constraints described in the graph, this approach results in a highly dynamic storytelling experience where user actions are effectively integrated in the generated path at any time.

3.1. The annotation graph

We use a graph to represent the various relations between annotations and their spatial position with respect to the inspected 2D model. Each node in the graph is represented using the approach of Bettio et al. [BAMG21], in which each annotation stores, in addition to the visual overlay and the external annotation description, also the parameters that should be used for an effective lens-based exploration of the annotated area. Edges between these nodes are used to represent preferred ordering relations in the presentation of annotations.

In particular, each node points to a set of enabling nodes that should be seen before visiting a candidate node. The presence of
edges allow authors to define a predefined global order. This precedence can be used to define a story, or, e.g., to go from coarser to finer details as prescribed by the visualization mantra [Shn96].

Moreover, the nodes also contain a weight (ranging in the interval from zero to one), which defines the strength of the dependency. If the weight is 1 it means that there is a strict dependency and the node cannot be visited if its enabling node has not been already traversed. A weight of 0 implies that there is no dependency from that node. Intermediate values produce a less strict ordering, as will be detailed later in sec. 3.2.

Authoring details are orthogonal to the subject of this paper. For the sake of completeness, we mention here that we annotate the models during the lens navigation as described by Bettio et al. [BAMG21], and define the dependency relation by editing a node table in which annotation names and thumbnails of the annotation areas are made visible.

3.2. Scoring System

During the navigation the system selects the next best annotation for the automatic tour using a scoring system. Following Bettio et al. [BAMG21], we assign to each recorded annotation node \( i \) a score \( N_i = \gamma_i \sigma H_i \), where \( \gamma_i \) is the author-defined annotation importance, \( \sigma \) is the similarity score depending on spatial and semantic distance, and \( H_i \) is the history score depending upon the activity log of the active user. The details of each individual score are presented in the original publication [BAMG21]. In order to take into account dependencies, we extend this formulation by multiplying the node score \( N_i \) by a dependency score \( d_i \), which takes into account node precedence relations and their weights, to obtain a final annotation score \( S_i = d_i N_i \).

The dependency score must express the fact that the author would prefer a given node be presented after its enabling nodes, with a strength depending on the edge strength. We express this by taking the fuzzy logic AND (i.e., \( \min \)) of a per-edge quantity that expresses if the node has already been presented and which strength should have this information. We also boost the value of the score of nodes that have been enabled to be higher than the base score, so as to favor the selection of just enabled nodes over other ones, favoring an ordered visit of the graph. The dependency score of node \( i \) is thus given by

\[
d_i = (\min (1 - e_{ij}(1 - v_j m_j))) (1 + v_j e_j) \tag{1}
\]

where \( j \) loops over all enabling nodes, \( e_{ij} \) is the author-selected edge weight linking node \( i \) to node \( j \), and \( v_j \) is 1 if node \( j \) has been visited and 0 otherwise. The factor \( (1 + v_j e_j) \) is the boosting term that enhances the probability of selecting the dependents of a node which has been enabled. Its value ranges from 1, for not visited nodes, to 2, for visited enabling nodes with maximum edge weight. The index \( j^* \) is the one that produces the minimum value in the other term of the equation. The term \( m_j \) models the fact that users tend to forget nodes presented in the past. We compute it as:

\[
m_j = 1 - \min (v_{nj}, Mvn - 1)/Mvn \tag{2}
\]

where \( v_{nj} \) is the number of visited nodes after the last visit of the enabling node \( j \) and \( Mvn \) is a maximum number of potentially remembered nodes. Thus \( m \) is 1 as soon as a node is visited and decreases to a minimum value \( 1/Mvn \) after \( Mvn \) other nodes have been visited.

The next best annotation is then selected by taking into account these scores \( S_i \). However, rather than just selecting the node with the highest score, we perform a stochastic selection among a small set of nodes that have a similar high score. In particular, we select a cutoff score equal to 60% of the maximum achieved score, and extract the subset of \( K \) nodes which have a score higher than this threshold. We then assign to each node in this subset a picking probability \( p_i = \frac{k}{\sum K} \), and select the next best annotation according to this probability. In such a way, the exploration is open to a wide range of possible paths, while maintaining the author dependency requirements. In particular, the approach makes it possible to chose among several paths in case of similar scores for several nodes, increasing the variability of the exploration experience, while avoiding the selection of incoherent solutions. This variability is important for casual visitors, as it makes the visit more engaging and less repetitive.

4. Results

We implemented the proposed approach on a web-based platform, where relightable images and all corresponding metadata are made available by a standard web server to a web client running in a browser on top of WebGL2, a JavaScript API that closely conforms to OpenGL ES 3.0 and can be used in HTML5 <canvas> elements without requiring plugins. The preparation of the relightable images and their layers, all the annotations, and the authored definition of each node and edge attributes in the relation-dependent graph are done off-line. They are stored in a repository that contains the set of image layers, a configuration file that manages the arrangement of those layers, and a file that includes both the text annotations with all the graph structure. At run-time, the viewer loads a scene description that includes the annotation database, and starts navigation by placing a lens in the middle of the screen. A simple user interface makes it possible to jointly control the lens and camera with the recently-introduced approach of Bettio...
et al. [BAMG21], that always ensures a good focus-and-context placement of the lens within the view. Automated tours can be activated at any time, either upon user request, or after the lens remains idle for 10 seconds with the auto-tour option active.

In this paper, we provide a preliminary demonstration and evaluation of the approach for annotated data exploration in the context of cultural heritage (see accompanying video).

The test dataset is a relightable multi-layered rendered image of three representative models from the Mont’e Prama collection of prehistoric stone sculptures [BJM'15, BRAB’16]: Archer n.5, Boxer n.15, and Warrior n.3 (see Fig. 1). The annotation database concerns reconstruction hypotheses, artistic details and part descriptions. It contains 44 annotations at multiple scales connected by 196 edges. Fig. 2 depicts a subset of the graph. In the image, each node is represented by the image of the context area of the optimal lens associated to its annotation. Edge width is proportional to dependency weight. The arrows indicate that, after the pointing node has been already viewed/visited, the pointed node will have a probability to be visualized proportional to the dependency weight.

We first tested the automatic navigation without the free movements introduced by the user interaction. In this setup, since the graph was authored with a single root node (the statue overview) required for all further inspection (dependency weight=1), we expect that the navigation starts from the root and, from there, a relevance-based order would be followed by navigation, taking into account node and edge priorities. In addition, since many of the subsequent weights are less than one, we expect our navigation to enable a good level of variety.

We performed 20 automatic tours, each of them visiting 10 nodes, always starting from an initial position at the center of the screen and viewing all the dataset. The 20 tours visited a total of 32 nodes with respect to the 44 contained in the graph. This fact shows how stochastic component in path selection avoids full repetitiveness, providing different exploration experiences to the users, also in fully automated mode. Fig. 3 (bottom three rows) shows three runs of the auto-navigation, with time going from left to right. It is clear how the first view is always the same, i.e., the graph root presenting the annotation related to the whole set of statues. Even if the lens is centered, providing a higher node priority to the center statue, there are several situation in which one of the other statues is selected first, due to our selection strategy with picking probability proportional to weight. From there, the navigation continues with a spatial consistency (e.g., if a high-level node of a statue has been visited, the navigation continues with higher probability in the leaf nodes of that statue), since the visualization of a statue’s annotation enables the visiting of the detail nodes of that statue, while the visiting of the details of nearby ones are blocked, until their root nodes have been enabled. The introduction of authored edges enables the introduction of semantic constraints.

When edges are removed, as in previous work on lens navigation [BAMG21], such constraints are not possible, and the next best annotation in a navigation path may be selected from a nearby statue based on pure proximity consideration. We repeated the 20 tests, but with edge dependencies disabled. In such a situation we explored in total 42 nodes, which almost covers all the nodes on the database, since more freedom is allowed by the lack of node dependencies. However, the paths are less structured, as they jump more frequently, for instance, from a statue to another. The first row of Fig. 3 shows an example of that kind of navigation, where edges are removed, and navigation proceeds purely by selecting the most similar annotations. It is clear that, without taking into account a hierarchy of nodes, the storytelling aspect of the automatic navigation gets lost.

The proposed framework, in addition, allows one to mix purely automatic navigation with interaction, since the user may take control of the lens during any path, and auto navigation restarts from the new user-updated lens and view configuration. We show this behaviour in Fig. 4. The red frames depict lens positions decided by the user (not by the automatic generated path); the next annotation selected by the algorithm takes into account the dependency graph and the history, while being consistent with the lens positioning provided by the user. In Fig. 4, after the first four automatic frames, the user interrupts the automatic navigation three times, in order to move and inspect all the three statues. The accompanying video shows additional examples of this behavior, in which we seamlessly move from automatic touring to interactive exploration, and, each time, the tour restarts taking into account the possibly largely modified local context.

5. Conclusions and Future Work

We have proposed a framework that aims at presenting annotations in a structured way. The approach is meant to support casual users to explore, at their own pace, spatially annotated 2D models using an interactive lens that moves from an interesting area to the next, while also responding to user input for interactive exploration. The presentation order is defined depending on lens position, navigation history and authoring information encoded in an annotation graph. The integration of a stochastic recommendation system that interprets context-dependent scores as transition probabilities makes it possible to increase the variability of exploration paths. Moreover,
Figure 3: **Automatic navigation.** Top row (yellow outline): an example of automatic navigation without using the dependency graph. The path proceeds by going from an annotation to the most similar one, without taking into account semantic aspects. Other rows (blue outline): several examples of automatic navigation with the dependency graph. All exploration paths start from the same annotation, and all tours share the same flow, dictated by authored graph dependencies, but introduce variations due to our stochastic next-best annotation selection process. The dependencies introduce semantic aspects, in this example favoring the presentation of a statue’s detail after presenting its overview.

Figure 4: **Mixing automatic and free exploration.** Our framework enables both automatic and free navigation. As soon as the user moves the lens (images outlined in red), the automatic navigation stops. When it restarts, the next frame is selected by taking into account both the dependency graph, the navigation history, and the user-updated lens and view configuration.

The user can freely mix personal/free exploration with automatic touring.

Our very preliminary evaluation has shown the potential interest of the approach, but also highlighted areas for future research. First of all, the current approach is targeted towards the exploration of areas that fit well on a circular lens, but should be refined when pointing at areas where linear or extended features should be explored. We plan to address this problem by storing at each node not only a single lens position, but a lens path for the exploration of the annotated area. Second, the dependencies presented here currently target the definition of simple precedence relations expressed by taking the fuzzy AND of values coming from enabling nodes. It is worth exploring whether fully supporting other logical operators (i.e., at least OR, XOR, and NOT) would be beneficial for improving the authoring expressiveness. Moreover, authoring, orthogonal to this work, also deserves attention, in particular in case of extension of the dependency logic. Finally, our current evaluation was very preliminary, and focused mostly on testing the feasibility of the approach. More work is required to objectively assess the effectiveness of our user interface. It will be also interesting to evaluate whether the proposed approach, currently tuned to museum applications, can be extended to more complex situations requiring specific visualization tasks to be solved.

**Acknowledgments** The authors thank CRBC Sassari and Soprintendenza Archeologia, belle arti e paesaggio per la città metropolitana di Cagliari e le province di Oristano e Sud Sardegna for the access to the artworks for the purpose of digitization and for annotation information. We also acknowledge the contribution of Raffaella Chierici for content creation. The project received funding from the European Union’s H2020 research and innovation programme under grant 813170 (EVOCATION), and from Sardinian Regional Authorities under project VIGECLAB (POR FESR 2014-2020).

**References**


[BJM+15] Bettio F., Jaspe Villanueva A., Merella E., Marton F., Gobbetti E., Pintus R.: Mont’e Scano: Effective shape and color digitization of cluttered 3D artworks. ACM Journal on Com-
puting and Cultural Heritage (JOCCH) 8, 1 (2015), 4:1–4:23. doi:10.1145/2644823.4


