

Visual Query Tools for Uncertain Spatio-Temporal Data

Katherine Malan
Department of Computer Science
University of Cape Town
7701 Rondebosch
+27 21 6502665
kath @ cs.uct.ac.za

Gary Marsden
Department of Computer Science
University of Cape Town
7701 Fiondebosch
+27 21 6502666
gaz@cs.uct.ac.za

Edwin Blake
Department of Computer Science
University of Cape Town
7701 Rondebosch
+27 21 6503661
edwin@cs.uct.ac.za

ABSTRACT

Some multimedia archives contain data which have vague locations in time and space. By this we mean that, although there is some idea of when and where the entity is located, the precise information is unknown. In this paper, we present a novel approach to displaying and querying such uncertain data. We use the concepts of dynamic queries, add to this a 2D query tool for performing spatial queries and enable Boolean combinations of queries. We have implemented these ideas in a pilot system for querying African artwork. In this way, we show how it is possible for novice users to easily query large multimedia archives with complex uncertain attributes.

Keywords

Uncertain Data, Dynamic Queries, Graphical User Interfaces, Query Preview, Cultural Archive, Visual Languages.

1. INTRODUCTION

Uncertainty in data is a common problem. The literature on this topic refers most frequently to fuzzy uncertainty, where areas have *vague boundaries*. The study described in this paper is not about areas with vague boundaries, but rather about uncertain *point* data, where the point could occur anywhere within a given area. This kind of uncertainty is known as probabilistic uncertainty

This research was initiated in collaboration with CAMA, the Contemporary African Music & Arts Archive [2]. The ultimate aim of CAMA is to develop an Africa-wide distributed database of African culture. In the CAMA archive, some data items have precise locations (e.g., an artifact which comes from a particular town). However, there are many items which do not have precise locations. These items originate from somewhere within an area of uncertainty, which can be anything from a province within a country to the whole of Africa (e.g., a sculpture in a gallery in Paris, which comes from *somewhere* in Africa). In the same way, while the year of origin of some data items is known precisely,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM'01, Sept. 30-Oct. 5, 2001, Ottawa, Canada.

Copyright 2001 ACM 1-581 13-394-4/01/0009...\$5.00

other items have uncertain origins in time. For example, one artifact could be dated 1940, another dated "late 19th Century" and yet another "1390 - 1352 BP".

Our aim was to develop a visual querying interface to the CAMA archive. There were a number of requirements for this application:

1. The application should provide an overview of the data and visual feedback on the results of queries.
2. The application should not treat uncertain items as any less important or available than items which have definite origins.
3. The user interface should be accessible to novice users.

2. RELATED WORK

The challenge in developing a user interface for our application, lies in not only in presenting the user with a meaningful representation of the data, but also in designing an easy-to-use query mechanism. The dynamic query approach lends itself well to this problem.

Ahlberg *et al.* in [1] were the first to define the concept of "dynamic queries". These allow users to formulate queries with graphical widgets, such as sliders. The goal of dynamic queries is that users see the results of query refinements as they make them. There have been many applications that employ dynamic queries and the benefits of using this approach have been well established [1,4,5,6]. Several experiments have demonstrated that dynamic queries help users find trends and exception conditions [4,5,6]. Dynamic queries are clearly effective, particularly for novice users. However, on their own, they are not expressive enough to cover all the combinations of queries needed for our application.

Mirel [3] has investigated the kinds of problems that users encounter in expressing complex combination queries for business problems. She proposed a **toolbar** of options for selecting data, which included replace (=), add to (+), subtract from (-) and intersect (&). She claims that with this **toolbar**, users' performance improved dramatically.

We have investigated if combining dynamic queries with such a **toolbar** for performing combination queries, would present a powerful solution to our problem.

3. A VISUAL QUERY SYSTEM

The overall user interface of the pilot Visual Query System is shown in Figure 1. All items in the database are shown in the central data display area. Any item can be queried (by right

clicking on the blue marker), and a thumbnail image of the relevant artifact will be displayed on the bottom right corner.

A **timeline** is shown above the data display. The horizontal axis denotes time, whereas the vertical axis represents uncertainty. Each item in the database is represented as a blue bar on the display. The more uncertain items are plotted further away from the timeline, and are shown as time ranges (longer bars). At a glance the user can gain insight into the data stored in the archive. For example, vertical clusters show a concentration of artifacts in a particular time period. By manipulating query tools, items on the data display are selected. The final aim is to generate a virtual gallery of the **artworks** automatically, which will be populated by the selected items.

3.1 Combinations of Queries

We have implemented a full set of Boolean operators based on Mirel's toolbar [3]. Through manipulation of the dynamic query tools, items on the data display are "tagged". The Boolean operators are then used for combining the current selection with the tagged items in different ways.

To query the archive, the user can select items in the data display area. This is done by selecting individual items (using the left mouse button), or by selecting ranges of items using one of the query tools provided (in this case time and space query tools). The idea is that the user "tags" items using the query tools and then adds these to (or subtracts from, replaces, intersects these with) the current selection using the Boolean operator buttons (see the vertical button bar to the right of the data display). In this way the user is able to change the current selection. Each tagged item is surrounded by a red rectangle for user feedback, while selected items are shown in yellow (note that an item can be both tagged and selected). The number of items tagged and the number of items selected are fed back to the user below the data display area.

3.2 Querying 1-D Data

The time query tool is just below the **timeline** and above the data display area: a slider with two drag boxes. This allows the user to choose items within a given time period. Vertical dashed lines provide further user feedback. As the user drags one of the **timeline** boxes, the items which fall in the given range, are tagged. Once again, these tagged items can be used to change the current selection.

In order to use a dynamic query tool with data, the data must be ordered in some way. Although time is clearly ordered, uncertain time is not so straightforward. Does an item which is very uncertain in time (i.e., with a long uncertainty range), occur before an item which is very certain within the same range? For example, say one item (say item 1) has a time value of 1650 and another item (say item 2) has a time value of "17th Century" (i.e. it occurs somewhere in the range 1600 – 1700). Which of these items should come first in an ordering? Since we would like a query tool to return both certain and uncertain elements, we assume that item 2 is both before and *after* item 1, with respect to time.

The application will therefore tag items when either of the vertical dashed lines of the time slider touches an uncertainty range. The user is able to exclude the uncertain items to limit the number of elements selected. There is an additional control on the left of the data display. This slider is used for queries based on the level of

uncertainty of the time attribute. By dragging the slider upwards, the more uncertain elements are excluded from the query (i.e. not tagged). In this way, the user can specify that only elements with relatively precise time origins should be included in the query. Figure 1 shows these two tools in operation. Some of the elements have been selected (those shaded yellow, or grey in black and white print). The **timeline** controls are used to tag elements in the given range (tagged items are surrounded by rectangles), while the uncertainty control excludes the two very uncertain elements below the horizontal dashed line.

3.3 Querying 2-D Data

To the right of the Boolean operator buttons is the space query tool. This tool differs from the usual sliders used in dynamic queries. When querying large geographic areas (such as the whole of Africa), users will usually think in terms of countries, rather than a continuous space of latitude and longitude. It therefore makes more sense for the tool to use countries as the minimal query unit, rather than some distance measurement (such as kilometers).

So, instead of providing the user with horizontal and vertical range sliders, we have provided the user with tools for selecting countries or sub-regions (recognized groups of countries). There are three tools above the map display. The first button (i) is used for querying the map. If this tool is selected and the user clicks on the map, information about that country will be displayed: the name of the country and the number of items which occur in that country. The other two buttons are used for selection.

The country selector button (C) is used for selecting countries, whereas the sub-region selector button (S) is used for selecting sub-regions. The idea is that the user first chooses a tool for selecting areas. Depending on the tool selected, they will be in a particular selection mode. If they are in a country selection mode, and they click on the map, the relevant country will be highlighted. If they are in a sub-region selection mode, and they click on the map, the relevant sub-region will be highlighted. In this way, the user can select multiple areas on the map. As countries are selected, the items which fall into that region, are tagged on the data display area.

The user is also given the option of either including items with uncertain geographic attributes in the query, or not. The user does this by using the check box below the map tool. If this option is selected, additional items, which could possibly occur in the selected area, are also tagged. It is important to note that a query on a region will include all items, at different levels of precision, which fall into that region. For example, if the user selects Zimbabwe and Mozambique, then the items which will be tagged will include: items which originate from either country, as well as items which originate from a more precise point or province within either country. In addition, if the user has chosen to include uncertain elements, then items which have an area which is a **superset** of either country as their origin (Southern Africa, for example), will also be included.

In this way, using the query tools, the user is able to continually refine their query until they have the desired selection and number of artifacts. The visual representations of tagged items and selected items ensures that the user knows which items will be involved in any operation before the operation is done.

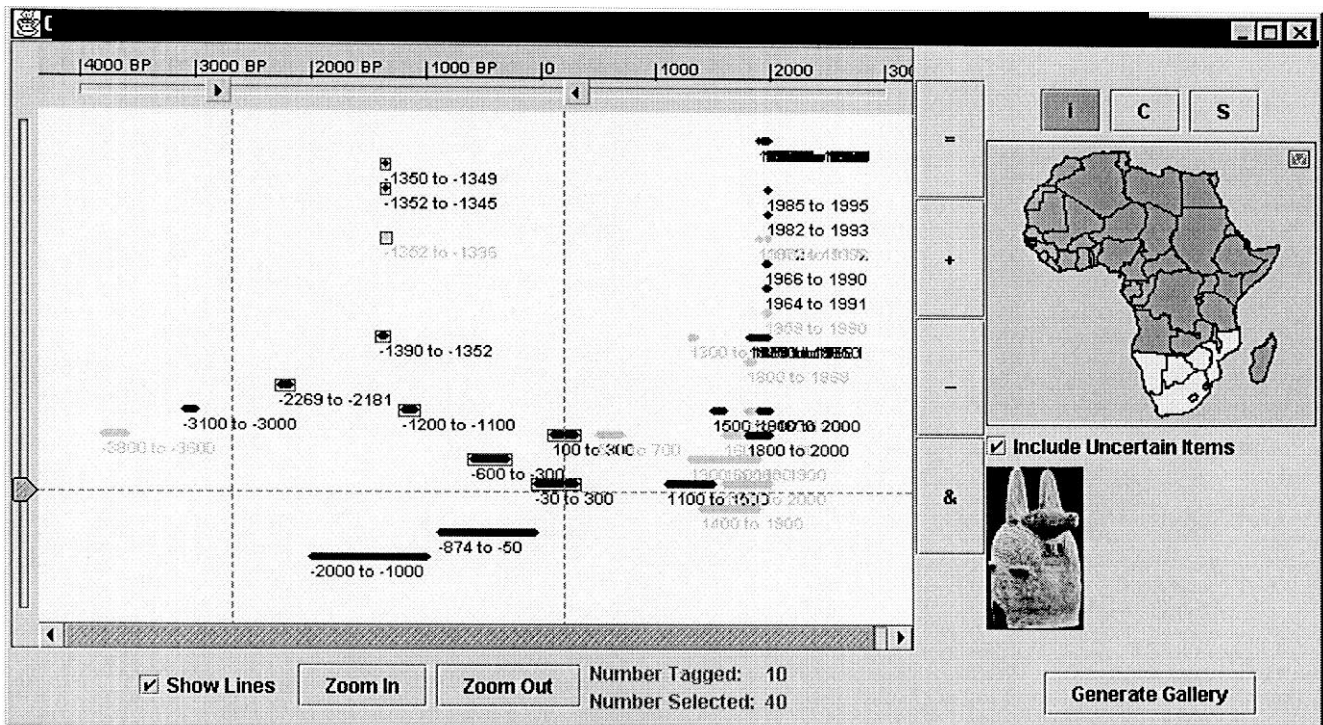


Figure 1. Pilot Visual Query System. Tagged elements are enclosed in rectangles, while selected items are highlighted. The time query tools (arrowed buttons above the data display), map query tool (on the right) and Boolean operator buttons, are used to tag elements.

4. DISCUSSION

By using the concept of dynamic queries, we have achieved the aim of providing an overview of the data with visual feedback on the results of queries. In addition, we have shown how both uncertain and certain items can be displayed on the same map. We have devised query mechanisms which catch both certain and uncertain data in the same query. In this way, the uncertain items are given the same level of importance as certain data elements.

The interface of the application used metaphors familiar to most users (a time line and a geographic map). The use of dynamic queries, with the possibility of fully combining queries, means that users can do all their querying visually, making the application accessible to novice users.

5. CONCLUSION

In this paper, we have presented approaches to visualizing and querying uncertain data in archives, which ensure that uncertain data are given the same level of importance as certain data. A further aim was to incorporate these approaches with techniques for querying multimedia archives, that would be appropriate for novice users.

We implemented our ideas in a pilot application for African artifacts. In this application, we show how dynamic queries can be used with Boolean operators, to produce an expressive, yet easy-to-use user interface. We have shown how data with uncertain I-D attributes (such as time) can be visualized and queried, within this framework. We have also shown how data

with uncertain 2-D attributes (such as geographic location) can be queried using a 2-D query tool.

6. ACKNOWLEDGMENTS

Our thanks to the NRF for funding this research. We are also grateful to CAMA for their assistance.

7. REFERENCES

- [1] Ahlberg, C., Williamson, C. and Shneiderman, B. Dynamic queries for information exploration: an implementation and evaluation. In Proceedings of CHI'92 (1992), 619-626.
- [2] Contemporary African Music and Arts Archive, <http://www.cama.org.za>.
- [3] Mirel, B. Complex queries in information visualizations: Distributing instruction across documentation and interfaces. In Proceedings on the 17th annual international conference on Computer Documentation (1999), 1-8.
- [4] Plaisant, C and Jain, V. Dynamaps: Dynamic Queries on a Health Statistics Atlas. In Proceedings of CHI'94, ACM Conference on Human Factors in Computing Systems (April 1994), 439.440.
- [5] Shneiderman, B. Dynamic Queries: for visual information seeking. In IEEE Software, volume 11, Nov 1994, 70-77.
- [6] Williamson, C. and Shneiderman B. The Dynamic HomeFinder: Evaluating Dynamic Queries in a Real-Estate Information Exploration System. In hoc of ACM SIGIR 92 (June 1992). 338 - 346.