Spatio-temporal framework for integrative analysis of zebrafish development studies

Belmamoune, M.

Citation

Belmamoune, M. (2009, November 17). *Spatio-temporal framework for integrative analysis of zebrafish development studies*. Retrieved from https://hdl.handle.net/1887/14433

CHAPTER 5

3D-VISQUS: A 3D VISUAL QUERY SYSTEM INTEGRATING SEMANTIC AND GEOMETRIC MODELS

Based on:

M. Belmamoune, E. Lindoorn and F. J. Verbeek.
3D-VisQuS: A 3D Visual Query System integrating semantic and geometric models.
InSCit2006 (Ed. Vicente P. Guerrero-Bote), Volume II "Current Research in Information Sciences and Technologies. Multidisciplinary approaches to global information systems", pp 401-405, 2006.

Abstract

For developmental studies, a wealth of anatomical data is being generated. This data is increasingly complex and is generally distributed across different systems. Our anatomical data is spatio-temporal having different levels of abstraction and distributed over different database systems. This kind of data impels to search for methods for its access and analysis in a usable way. Here we try to provide a solution to access this complex data by combining easy ways for data perception with information retrieval techniques. We introduce a visual query system for our zebrafish spatio-temporal resources. With this system, i.e. 3D Visual Query System (3D-VisQus) we strive to afford end-users with a simple and intuitive interface to visualize, interact and query complex anatomical data. The 3D-VisQus is a portal to simplify users' access to complex anatomical data and to facilitate data analysis and understanding.

5.1 Introduction

For developmental studies, a vast amount of experimental 3D spatio-temporal data is being produced and managed in different database systems. We expect that these data production will continue growing and from a scientific perspective these wealth of data creates unordinary opportunities for developmental studies. The challenge that we are facing is how to quickly and effectively turn this data into valuable knowledge for a wide range of users. In this chapter we describe our solution by developing communication processes that will bridge conceptual differences across our database systems.

We developed a spatio-temporal framework to study embryonic development of zebrafish model system. To this framework we introduced an additional component, i.e. the 3D Visual Query System (3D-VisQus). It is a prototype of a query system that provides a portal interface to the framework's main components, i.e. the 3D digital atlas of zebrafish (cf. chapter 3) and *in situ* gene expression patterns organized in the Gene Expression Management System (GEMS; cf. chapter 4). The 3D-VisQus uses ontological concepts of the Developmental Anatomy Ontology of Zebrafish (DAOZ; cf. chapter 2) to link phenotypic data of the 3D Atlas to bimolecular data in the GEMS.

The 3D-VisQus is an experimental approach that we developed to overcome problems related to accessing and searching 3D spatio-temporal anatomical data. Not all users have a direct knowledge of the anatomy, yet we want them to make use of our systems. Therefore, we investigated the possibility of designing a system to provide users with an environment for data visualization, integration and retrieval. Some of its required properties are: visual representation of 3D spatio-temporal models, different level-of-detail definitions when browsing in a graphical model and the possibility of representing a textual query in a visual way. We designed a prototype of such system, i.e. the 3D-VisQus to access and query our 3D spatio-temporal anatomical information. 3D-VisQus will be considered as a mental model that translates users' thoughts processes for how anatomical structures are related to each other into real 3D models. Moreover, it will assists users to formulate readily their search queries using visualized graphical data while underlying systems and the query language are transparent to users.

The concept of using visual perception to query data has already been introduced, e.g. geographical information systems (GIS). In these systems users are able to express their queries in a visual environment. Users formulate their search queries in a visual way by selecting appropriate image icons and putting them into the proper places. To facilitate information understanding the results of a search query are also presented in a graphical manner (Goncalves et al, 2000). In our framework we have different kind of data, i.e. 3D spatio-temporal models. The system should first visualize 3D models and second translates users' actions to formulate search queries.

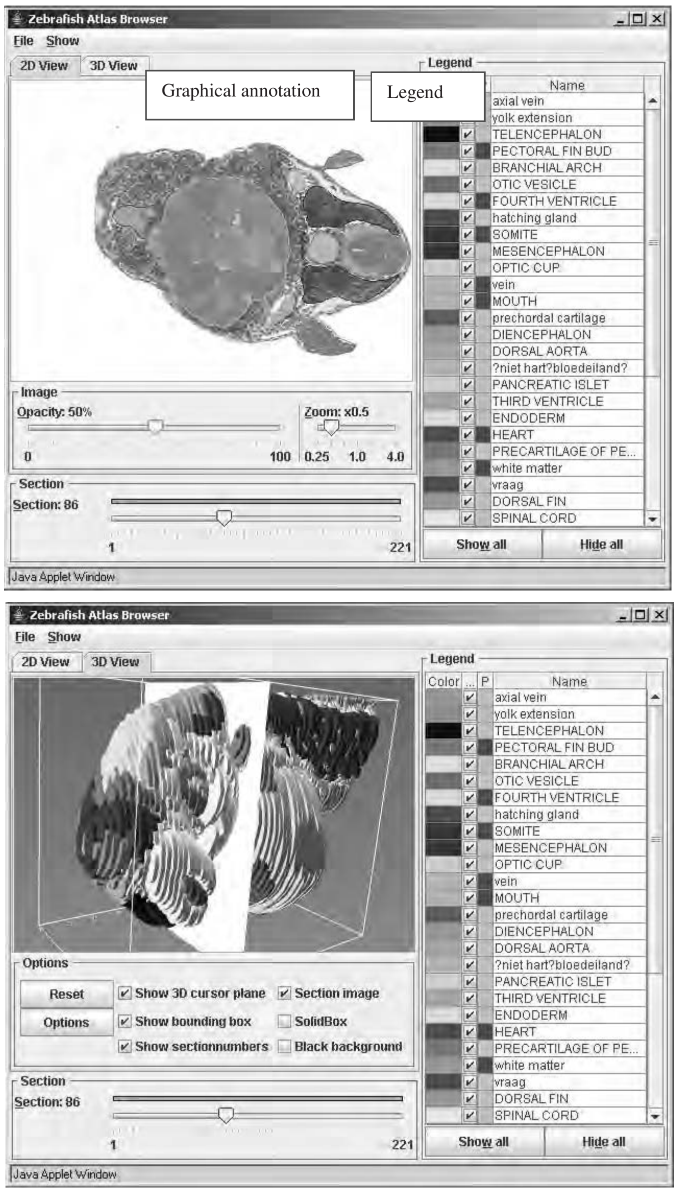
The 3D-VisQus is based on the 3D Atlas Browser (cf. chapter 2) which was designed to visualise and explore anatomical data of zebrafish model system in 3D. Such 3D data gives users the opportunity to explore anatomical structures in the correct manner. The 3D-VisQus extended the 3D Atlas Browser to use a visualized 3D model as a query interface to search for patterns of gene expression. We are applying the concept of visual queries by using each anatomical domain in a visualized image as an abstract object to generate a query. Anatomical domains are annotated graphically and are labeled with a textual annotation. Therefore from choosing a graphical domain, the system extracts the spatial and temporal information of that domain and generates a search query clause. The query is then executed against the DAOZ and GEMS database. The 3D-VisQus prototype is implemented on top of our framework databases. The 3D-VisQus is, therefore, a portal to our database resources to which a typical retrieval environment is attached. Searching with a system such as the 3D-VisQus will make querying complex data more intuitive and users do not have to worry about query languages nor the underlying databases.

The 3D digital atlas of zebrafish (cf. chapter 2) is used as a spatial reference framework for mapping data. It is our representation of zebrafish anatomy that provides a visual standardized coordinate system for the search and analysis of expression data. 3D patterns of markers genes expression are obtained from whole mount in situ hybridization (Welten et al, 2006) and are organized in the GEMS database (cf. chapter 4). The gene expression database must describe the time and space of gene expression in a standardized way. To achieve this goal, we used DAOZ (cf. chapter 2) for common nomenclature for data annotation in GEMS and 3D atlas models. This common annotation enables mapping GEMS data onto 3D atlas of zebrafish development.

In the DAOZ anatomical structures are modeled hierarchically from functional system, body region to substructures. Each anatomical concept could have more than one hierarchical parent using more than one relationship specifying in which way these concepts are related to each other. This hierarchy provides different level of granularity that facilitates data organization, analysis and retrieval. The ontology concepts represent the pivot of the 3D-VisQus to integrate the 3D atlas and GEMS components and to compose and execute search queries against the 3D atlas and GEMS database systems using the different levels of data abstraction. Through the visual query interface of the 3D-VisQus users are able to compose spatio-temporal queries when interacting with the visualized graphical entities. The system extracts the spatio-temporal information of the region of interest and generates metadata. These represent the basic components of the queries that the system creates and executes against our databases models, e.g. “select all patterns of gene expression where the expressed patterns are located in anatomical domain X and present at time Y”. By this system a mapping between gene expression data, the 3D atlas and DAOZ is realized and moreover a novel paradigm for searching based on visual understanding of data is introduced and explored.

5.2 3D-VisQus Usability

Our anatomical data could be considered as complex since it is 3D, spatio-temporal and distributed over different database systems. The learnability concept is the key of a successful system such as the 3D-VisQus to access and search these data. The system should therefore be easy to lean with an intuitive interface. This interface should present anatomical information with the focus being to make such information easy to access in order to sufficiently engage the user. Anatomical data should be available in different views and scales. Multiple view approach is wished for users to deal with related data at different levels of abstraction. The points of interest should be visible in both global and detailed views; related aspects only visible, but not detailed, until the user chooses them. In this context of these usability requirements, the 3D-VisQus has been designed.



Graphical annotation

Legend

Opacity: 50%

0 100

Zoom: x0.5

0.25 1.0 4.0

Section: 86

1 221

	Name
	axial vein
	yolk extension
✓	TELENCEPHALON
✓	PECTORAL FIN BUD
✓	BRANCHIAL ARCH
✓	OTIC VESICLE
✓	FOURTH VENTRICLE
✓	hatching gland
✓	SOMITE
✓	MESENCEPHALON
✓	OPTIC CUP
✓	vein
✓	MOUTH
✓	prechordal cartilage
✓	DIENCEPHALON
✓	DORSAL AORTA
✓	?niet hart?bloedeiland?
✓	PANCREATIC ISLET
✓	THIRD VENTRICLE
✓	ENDODERM
✓	HEART
✓	PRECARTILAGE OF PE...
✓	white matter
✓	vraag
✓	DORSAL FIN
✓	SPINAL CORD

☑ Show 3D cursor plane ☑ Section image

☑ Show bounding box ☐ SolidBox

☑ Show sectionnumbers ☐ Black background

Section: 86

1 221

Color	...	P	Name
	✓		axial vein
	✓		yolk extension
	✓		TELENCEPHALON
	✓		PECTORAL FIN BUD
	✓		BRANCHIAL ARCH
	✓		OTIC VESICLE
	✓		FOURTH VENTRICLE
	✓		hatching gland
	✓		SOMITE
	✓		MESENCEPHALON
	✓		OPTIC CUP
	✓		vein
	✓		MOUTH
	✓		prechordal cartilage
	✓		DIENCEPHALON
	✓		DORSAL AORTA
	✓		?niet hart?bloedeiland?
	✓		PANCREATIC ISLET
	✓		THIRD VENTRICLE
	✓		ENDODERM
	✓		HEART
	✓		PRECARTILAGE OF PE...
	✓		white matter
	✓		vraag
	✓		DORSAL FIN
	✓		SPINAL CORD

Figure 1: (Above) 3D atlas: 2D and (Below) 3D views of a zebrafish embryo at 36 hours post fertilization (hpf).

The 3D-VisQus is an extension of the 3D Atlas browser (cf. Figure 1). The extension resides in 3D-VisQus capability to compose queries based on user's perception of the data. We have explored the possibility to use section images of a 3D model as a query interface. The system uses a low level of data visualization in the form of images, i.e. the Graphical User Interface (GUI) to retrieve high level elements in other database systems.

Actually, the 2D view of a visualized 3D model represents a portal user interface with two major capabilities. On one hand it is used to explore a 3D model and the section images to get more insight of the 3D data. On the other hand each section image within the 2D view is used to generate a query, which is then executed against the underlying databases, i.e. DAOZ and GEMS databases. Within a section image, each anatomical structure is bounded by its contour representing the graphical annotation to which an anatomical name is associated that represents the semantic annotation. Each anatomical domain is associated with a specific label that facilitates distinguishing and highlighting information of interest (cf. chapter 2).

The 3D-VisQus offers two query forms. First, users are able to query into the visualized model itself. In a 2D view, visualizations are synchronized. In the legend at the left pane (cf. Figure 1), when an anatomical name is selected, the main pane is immediately updated with the first section image containing the anatomical domain related to the selected name (Verbeek et al, 2000 and 2002). Second, each anatomical domain is used to query the GEMS for patterns with expression in that domain. The query is an association between the spatial and temporal characteristics of the anatomical structure. The spatial information is provided by both the graphical and semantic annotations of the visualized anatomical domain while temporal information is obtained from the developmental stage of the visualized embryo (3D model). When an end-user wants to make a query, there are three main steps: what to query, where to query and how to query. In the following, we give an example to illustrate the visual query process. Assume an end user wants to know expression patterns within structure X at developmental stage or time Y. First, end user selects the proper 3D model (where to query) which is then displayed with the 3D-VisQus. Second, the user looks for the structure X in the displayed section images (what to query) and interacts with (clicks on) the anatomical domain of this structure to generate a system action. At this point, the system, i.e. 3D-VisQus translates this operation to system's internal query in the form of: "select all patterns of gene expression in structure X and at time Y". This query is then executed against the underlying database systems and the patterns of gene expression are displayed.

The system is based on three spatio-temporal database management systems (DBMS) (cf. Figure 2). The main spatio-temporal database is this of the 3D digital atlas. The visualized 3D models are stored in this database while queries are executed against the other two databases: the DAOZ database and the GEMS database.

Visualization of the query result is critical to the end user. The output is given as a combination of two forms: graphical and textual. The system is able to translate textual results to a graphical output. This is given in the form of thumbnails of patterns of gene expression that are generated in real time. Each thumbnail image is a 2D projection of the original multi-channel in situ hybridization images stored in the GEMS (cf. chapter 3). Additionally, each image carries expression information of a given gene and is indicated with this gene symbol representing the textual information. To explore data, end users can move from global overview to more detailed information. This principle is applied to explore the data and is also respected in the visualized output data. The output list provides a global visual data analysis. At this stage, the system also provides details on demand. When a user selects a thumbnail image a second query is submitted to the underlying system. A detailed view about the selected pattern is then generated in real time (cf. Figure 4).

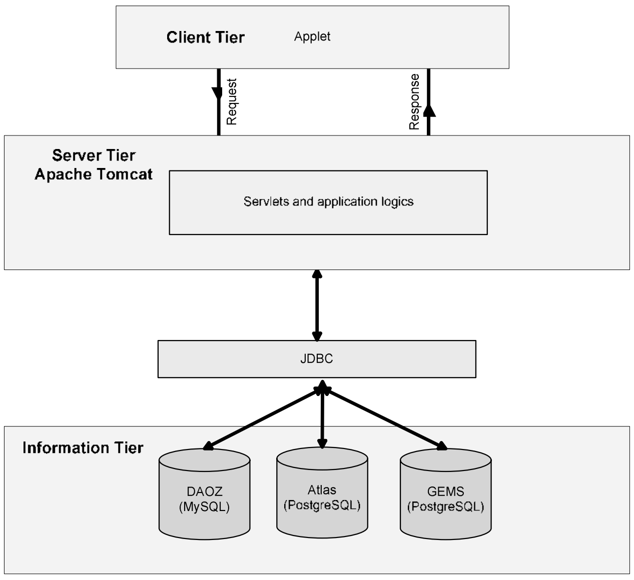
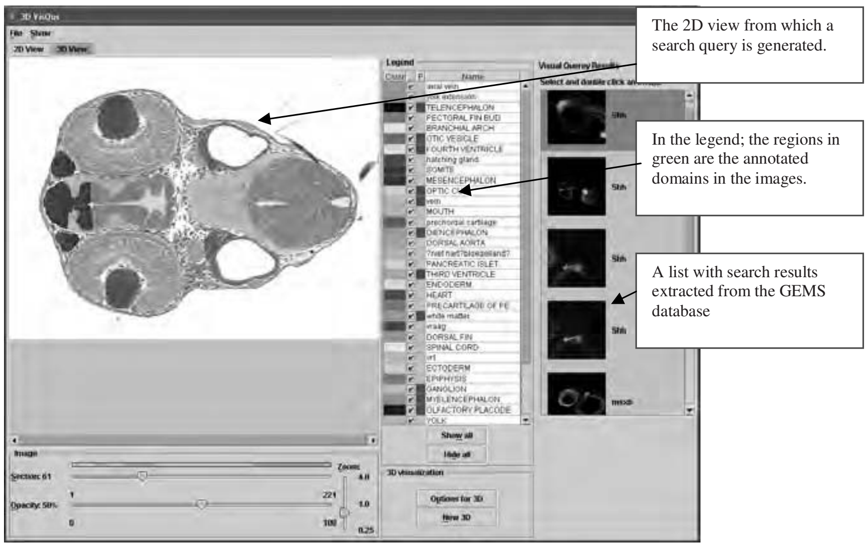


Figure 2: The 3 tier architecture of the 3D-VisQus.

In accordance to the architecture of the atlas, the DAOZ and GEMS, we also adopted a multi-tier architecture for the 3D-VisQus application (cf. Figure 2). 3D models are extracted from the atlas database and are visualized with a java applet that can be deployed on the client side. We are using the applet as the application front-end. The applet communicates with the server side through a servlet. We need to capture the user action and pass this information to the servlet. Since servlets support the HTTP interface, we communicate with the servlet over HTTP socket connections. The applet opens a connection to the specified servlet URL. Once this connection is made, then the applet can get an output stream or input stream on the servlet.

The applet translates the user's action to a request and sends this to the server tier. This tier processes client requests and sends results back to the client. The server tier formulates and executes the query based on the action type, the spatial feature of the region of interest and the temporal information, i.e. development stage of the visualized atlas model. Both the DAOZ and the GEMS databases are addressed. The first is queried to get control identifiers of the spatio-temporal concepts of the interested region while the second uses this information to retrieve patterns of gene expression responding to these search criteria (cf. Figure 3).



The 2D view from which a search query is generated.

In the legend; the regions in green are the annotated domains in the images.

A list with search results extracted from the GEMS database

Figure 3: A search example of 3D-VisQus. By clicking on a region of interest in the visualized section image the system creates and executes a query against the GEMS and the search results are displayed. If there is a void query result, the system looks in the proximity location of the indicated anatomical domain. In this case the following query has been generated: "select all patterns with gene expression domain located in head and not older than 48 hpf".

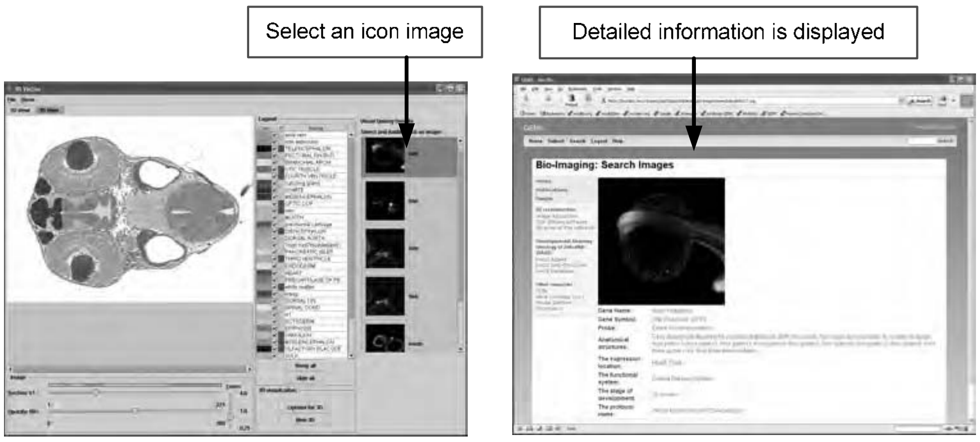


Figure 4: From the output list (left) a detailed view of a selected pattern can be generated (right).

5.3 Users Analysis and System evaluation

There are two main groups of target users comprising creators of the 3D models and biologists interested on the 3D data. The first group is of experienced biologists. They create atlas 3D models using the TDR-3Dbase (cf. chapter 2). The second group represents our main audience; they could be experienced biologists as students interested on using our spatio-temporal data. For this group, it is important to access the information as quickly and easily as possible.

To investigate how the 3D-VisQus will be accepted, we focused on measuring its capacity to meet its requirements. Therefore a set of task scenarios were developed. These represent realistic situations wherein the tester performs a list of tasks. The evaluation was carried out with a small group of target users comprising collaborator biologists and novice users. The evaluation involved a walk through the developed scenarios to capture typical user tasks in a normal work flow and to allow users to explore offered functionalities. Based on uses feedback, we concluded that the system meets the most important requirements. Users also made suggestions for improvements and changes to the system. Such as enabling query storage for later use, and the possibility of representing a visual query in a textual way as a form of feedback to the user.

5.4 Conclusions and future work

Integration of spatio-temporal gene expression patterns with the anatomy from the 3D atlas is important for genetic and developmental studies. It is crucial for researchers to be able to access, search and combine such information for an effective understanding of the anatomical development. The 3D-VisQus is an experiment that we derived to investigate a mechanism to intuitively map genotype into phenotype data. With a system such as the 3D-VisQus we could realize a primary mapping between in situ data of different expressions patterns into the 3D digital atlas. Moreover, the 3D-VisQus offers many other advantages. It provides the possibility to explore 3D data and dynamically formulate correct and exact visual query clauses. When compared to a semantic query system, this form of data retrieval offers a great flexibility in formulating complex search queries. In our case, users are not required to have a deep knowledge of development in order to formulate a search query with exact anatomical terms. With a few maneuvers (mouse clicks), users formulate their search queries based on their visual perception and data recognition. The 3D-VisQus extracts and executes text-based SQL statements from the submitted visual query while the underlying databases models and the query language are completely transparent to the users.

The 3D-VisQus is our proposed prototype of an interface to visualize and query our complex anatomical data. The 3D-VisQus is a portal to our database resources. The system allows search based on visual understanding of the data, and as such bridges conceptual and linguistic differences between systems as well as users. Early usability evaluation involved expert and novice users. This primary evaluation revealed that users appreciate the interface and the possibilities it offers to query spatio-temporal databases in an intuitive manner. Therefore, we are encouraged to improve this system to be used as a valuable visual query system on the top of our database systems. We need to add advanced search methods by taking full advantage of the combinatorial relationships among anatomical structures and to mirror these in the search capabilities of the system. In this initial prototype the focus was on the spatial relation: the *located_at* operator of the selected anatomical domain. As a future improvement, the system needs to generate a query based on the selected anatomical entity; in case of no results the system should come up with a new query with a higher granularity level based on the spatial characteristics of the selected graphical entity. These characteristics can easily be extracted from the DAOZ. In order to describe the spatial relationships more clearly between anatomical entities, operations like *part_of*, *near_of*, *located_at* can be used. For example, if there are no patterns associated with the anatomical term ‘hindbrain’ then the system should be able to generate a new query using granularity and relationship covering the whole ´brain´ region by using the *part_of* relationship. This feature is made possible because of the advanced hierarchical structure of the anatomical ontology of the zebrafish (cf. chapter 4). In future embellishments the system should enable query storage for later use, and the possibility of representing a visual query in a textual way as a form of feedback to the user. The actual output of a search query is given as 2D images, i.e. patterns of gene expression, while, the ideal output should be a spatial mapping of the expression domains onto the standard atlas model. Spatial concepts from the DAOZ are used to annotate anatomical and expression domains in the 3D atlas models and patterns of gene expression in GEMS respectively. To improve the 3D-VisQus, these concepts could be reused to provide a spatial mapping between the atlas and in situ gene expression data. The 3D-VisQus is a promising on-going research project intended to contribute to our user’s satisfaction.