3D Content-Based Visualization of Databases

Anestis KOUTSOUDIS
Dept. of Electrical and Computer Engineering, Democritus University of Thrace
12 Vas. Sofias Str., Xanthi, 67100, Greece

Fotis ARNAOUTOGLOU
Cultural and Educational Technology Institute, Research Centre "Athena"
58 Tsimiski Str., Xanthi, 67100, Greece

George PAVLIDIS
Cultural and Educational Technology Institute, Research Centre "Athena"
58 Tsimiski Str., Xanthi, 67100, Greece

and

Christodoulos CHAMZAS
Dept. of Electrical and Computer Engineering, Democritus University of Thrace
12 Vas. Sofias Str., Xanthi, 67100, Greece

ABSTRACT

The scope of this paper is to propose an innovative interactive system for 3D visualization of multimedia databases. In such databases, the digital content is accompanied by complementary information and metadata that describe the content. The proposed 3D visualization environment is able to provide multiscale representations driven by appropriate clustering processes that exploit content descriptions. It is also able to provide direct access to all the digital content, both in terms of physical access to the multimedia files as well as in terms of searching and retrieving information.

Keywords: 3D graphical user interface, 3D visualization, content-based description, content-based visualization.

1. INTRODUCTION

As the 3D technologies become more and more common to the wide public and gain even more attention, acceptance and appreciation, there is an increasing demand for more sophisticated user interfaces that can use metaphors of real life and keep the user far apart from low level information and data structures. Nowadays, users prefer to access data in a more human-oriented and comprehensive way instead of using typical text-based interfaces. Additionally, there is an increasing demand for interfaces that reflect the content of the data and are able to provide content-based retrieval mechanisms and representations. These facts are complemented by a continuous development of 3D display hardware, based on both active and passive technologies, that contribute on the generation of an open field for research and development of new human-computer interaction interfaces on both algorithmic and software levels.
The system described in this work is an innovative interactive 3D visualization system, and it is actually a proposal for future user interfaces and front-end virtual environments for multimedia and multi-dimensional databases, such as image databases, video archives, musical archives and collections. The digital content is usually accompanied by complementary information and metadata that describe its content. The procedures of extraction and insertion of the accompanied information in the database is not a subject of this work. We focus mainly on the 3D visualization environment that is able to perform multiscale representations and direct access to the whole digital dataset. The system provides a virtual 3D space based on taxonomies either on high or low level features retrieved from basic categories and sub-categories. More specifically, the proposed approach performs 3D data clustering and visualization according to metadata and content-based descriptions that have already been extracted and stored within the archive. The system is composed by two main sub-systems:

- the clustering sub-system, which uses Self-Organizing Maps (SOM) [1]-[3] to cluster the data and construct the 3D virtual space in terms of 3D coordinates derived from the data, and
- the visualization and user interaction sub-system, which uses features of the OpenGL graphics libraries to render in real-time the virtual space and provide the user with the required interactivity

For the evaluation of the proposed system we considered as a music archive of Greek traditional songs. The clustering subsystem was trained and evaluated using a collection of Greek traditional music database that was formed in CETI as a part of a national R&D project named "Polymnia", containing about 1000 files. The collection follows an XML native database format and includes the whole dataset that includes information such as:

- links to digital musical files and audio samples
- accompanied information (such as scores, references, analysis work) and
- metadata (such as the genre, the beat and the meter, or other automatically extracted low level characteristics).

2. DATA VISUALIZATION IN THREE DIMENSIONS

It is a fact that during the last decade, an increase in the volume and complexity of the digital information has occurred. As a result, more advanced techniques are now required in order to visualize this complex information. The mature era of computer graphics can be considered as a domain where digital information can be moved from concrete representation to abstraction. The use of computer graphics has eventually enabled significant improvements in the everyday accessibility of large data sets [4]. 3D computer graphics can improve the process of database exploration by providing a virtual environment which enhances the processes of both data searching and analysis. Many visual query languages have been already been implemented under WIMP (window, icon, menu, pointer) environments. The motivation behind database visualization is the development of a novel system that will be derived from experimentation in the areas of graphics rendering and viewing parameterization. The main development of the system should be focused on the capability to bring out the most relevant information from the vast data set. Thus, a useful data retrieval and visualization system can be characterized by its ability to allow interactive explorations of a parameterized 3D space. The more efficient, a system is, the fewer the number of iterations needed for parameter selection [5]. A three dimensional depiction of a logically organized node-based network can be considered as a representation that exploits the human visual processing in order to reduce the cognitive load of many tasks that require understanding of a global or a local text-based structure [6]. Recent research work has been focused on the efficient real-time visualization and interaction of data by using volume renderings, isosurfaces, 3D contours, slices, scatter and vector plots.

3. THE VISUAL UNIVERSE AS AN INTERFACE

The driving force of this work was to capture the beauty, diversity and scale of the universe in all its colorful formations and structures and combine it with today's 3D computer graphics technologies in order to bring it to a common computer system as an enhanced, friendly and appealing user interface environment for multi-dimensional data visualization. Contemporary digital archives can be though of as multi-dimensional databases consisting of both multimedia content and its description. The descriptions of digital data can be multiple and multi-dimensional at the time, varying from low to high levels. High level descriptions often refer to properties easily perceived by the human observer while low level descriptions refer to mathematically expressed properties. Low level descriptions, usually, take the form of numerical vectors of high dimensions. Considering the fact that each one of these descriptions can only capture a single or a small segment of features from the original content, one can imagine the amount of different dimensions that can be used in order to provide with an overall signature-like description. Nevertheless, provided that such descriptions already exist within our database structure, we describe a way of exploiting high-dimensional descriptions to provide an interactive three dimensional data access interface base on the visual metaphor of the universe.

The visualization and interaction scenario in the proposed virtual space is as follows:

- a "planet" is considered as the data unit of the system and represents a digital music file from the collection. It is placed within a "planetary system" that consists of similar and directly related files in terms of a selected low-level characteristic (like dominant color for images or the beat in case of music). Each planet carries apart from a link to the digital file, additional information and metadata. Furthermore,
- a "planetary system", that groups similar planets, belongs to a "galaxy" that consists of planetary systems with a common higher-level characteristic (like the musical genre for music), but are distinguishable through a lower-level characteristic. Finally,
- the galaxies form the "universe". The universe visually describes the complete dataset in an organized manner.

In the following paragraphs a brief review of relative works is given and a detailed description of the system is provided.

Data Organization and Visualization

Data access interfaces of digital archives are mainly based on access modes that allow searching via full text and metadata (the most common interface) or via high level characteristics (i.e. content type, titles, etc.) and, more recently, via low level features based on signal processing techniques.

Most of these approaches in today's systems are highly dependent on the user knowledge of what he or she is looking for. Most of the archives provide means of interacting with their content in textual form. This king of organization (using mainly textual descriptions) is based on manual content description by the holder, which is rather subjective and, possibly, inconsistent with a user's point of view. As a result, it is difficult for such a user to browse and search within the content. A way to overcome the subjective nature of the categorizations is to move to a lower level and incorporate signal processing techniques to describe the content and use appropriate models for the final classification.

Implementation of a Universe

The proposed visualization system, in its experimental form, consists of two main sub-systems:

- Clustering sub-system for the Universe formation: the sub-system is responsible for the database content clustering and the formation of the Universe by using a three-step process
- Enhanced user interface sub-system: the sub-system is responsible for the 3D representation and interaction with the database

Clustering sub-system for the Universe formation: The first sub-system is employing a strategy for the formation of the Universe. It is critical for the system as it is responsible for preparing the data and metadata for visualization. The sub-system relies on neural networks, and especially on SOM. It employs a three-dimensional neural formation which is fed with the appropriate data, is trained to fit the data and labeled accordingly and, finally, is used to represent the multi-dimensional data in a virtual 3D space. Specifically, it includes the following processing steps:

1. *Data structure formation*: the information that the system uses in order to produce the appropriate clustering is a multi-dimensional combination of selected metadata from the database. In some cases, several such combinations have to be tested, each of which produces different clustering results.
   The data undergo a Principal Component Analysis (PCA) [7], [8] that reduces the dimensions to three. It is significant to note here that the dataset can consist of several thousands of entries, as a digital collection can grow enormously.
2. *Neural network formation and training*: the three-dimensional training dataset is fed to a three-dimensional SOM of a specific number of neurons. The number of neurons can either be chosen according to the dataset or can be fixed to an appropriate, relatively small, number. The small number of neurons is chosen so that the final visualization data are also kept to a small amount, in order to be able to provide high quality 3D visualization even on low processing power sys-

tems (with old 3D acceleration graphics cards). Next, the SOM is trained and the neurons are moving to positions that reflect the data positions. This way a reduced model of the dataset is obtained. The final step of this process involves the labeling of neurons accordingly to the data that they represent and their position coordinates are normalized. These neuron positions are stored and are used to render the Universe. Such a clustering result is shown in FIG. 1, where the final neuron positions are drawn and different gray-levels represent different galaxies. The stars represent the centroids of the various classes.

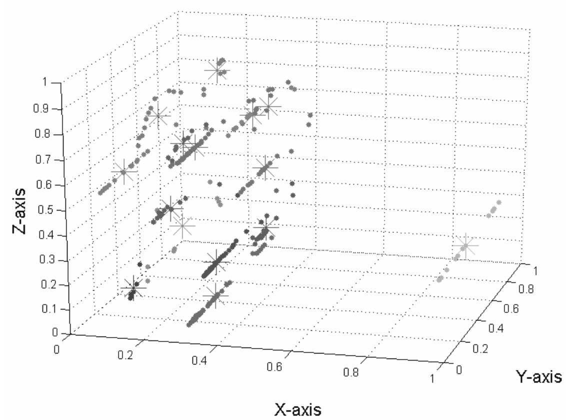

FIG. 1. Positions of neurons after data dimension reduction and clustering

3. *Unsupervised clustering to sub-categories*: after the training of the SOM and the labeling of the neurons, an additional clustering is imposed on the data for each of the classes (galaxies). Each set of data that belong to a specific galaxy undergo an unsupervised clustering based on the same or other features from the content descriptions. The classifier employed is a k-means classifier [9], which can be though of as an equivalent to a SOM where the affected neighborhood of the neurons is set to zero. K-means leads to the classification of each galaxy dataset to a variable number of clusters according to the spatial distribution of the data and can be explicitly set not to exceed a predefined number of clusters. This way the neurons that represent each of the database entries that belong to a specific galaxy are clustered according to their content-based descriptions into sub-classes. At the final step, the actual data of the database are mapped to the three-dimensional virtual space according to a best matching neuron topology representation: the data are labeled according to the neurons that best represent them and, thus, the planetary systems are formed. Each galaxy's data positions are stored and are used to render the galaxies in the visualization environment.

Enhanced user interface sub-system - 3D visualization and interactivity: The second main sub-system consists of the 3D visualization and user interaction environment. As already stated, in order to ensure a high level of compatibility to various hardware configurations and versions, as well as to provide high quality visualization, OpenGL was chosen as the basis for the implementation of the 3D user interface. FIG. 2 depicts a screenshot of a 3D Universe.

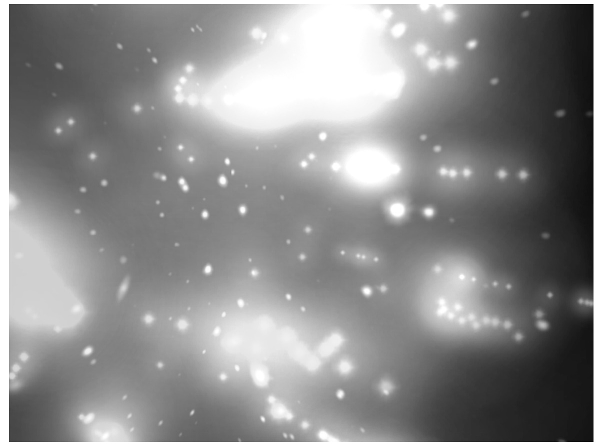

FIG. 2. 3D visualization of a Universe

In this phase, the user can navigate within this view (translate and rotate) and retrieve information regarding the total amount of the data in the collection. Obviously, since the positions of the galaxies are determined by a clustering process, the relative volume of each galaxy represents its relative amount of information, and the relative positions of the galaxies represent the relation between them according to the features selected for clustering. When the mouse moves over a galaxy the name of the galaxy appears as a pop-up text within the universe. The user can additionally get textual information regarding the universe, such as statistics. Coloring and special shading techniques (glowing and blending) are being used to make the different classes easily distinguishable.

If the user selects a galaxy, the user interface "zooms in" one level and displays the sub-categorization within the selected galaxy. That is, it presents the planetary systems within the galaxy grouped by similar characteristics and distinguishable by appropriate coloring and glowing techniques. Again the interaction follows the same principles: the user can navigate in order to meet a preferred viewpoint of the data, read the titles on top of the planets, get all the information regarding the content as well as a link to the actual digital file and get the total table of contents in the galaxy in textual form. This way the user reaches intuitively to the final step of this visualization scenario which is to get the information about an entry of the digital collection.

4. A CASE STUDY: A MUSICAL UNIVERSE – VISUALIZING A COLLECTION OF DIGITAL MUSIC

The idea of visualizing a database with music data has been explored by several researchers over the last years. It is a fact that a large segment of the previously mentioned volume increase of digital information is related with the rapid developments of the World Wide Web, portable music players, and multimedia compression schemes. It is more than common nowadays that users store hundreds of music tracks in their personal computer systems as it allows a better management of their personal music collection. Managing such a collection is by its nature an interesting and challenging research field. Most of the commercial music management tools are based on conventional directory-based organization tree structures with textual information presented as a list [10], [11]. We have implemented a music database management system that depicts the attributes of our universe and it establishes a fundamental platform for the future development of advanced software applications that will provide novel and more efficient methodologies in music organization, retrieval and playback. It is important when designing a visualization software to take into account the characteristics of the target user's task domain so that the system performs more effectively and scale to larger datasets [6].

Musical data organization and visualization
Human-computer interaction interfaces aimed for electronic music archives may basically be distinguished in four primary modes of access. These are:

- *database-oriented* access allowing search via metadata, which is the most common interface provided by archives
- *text-based* access searching the text of songs for particular phrases
- *searching* for a particular melody or tune
- *looking for* a particular type of music, i.e. for titles within a specific musical genre

All of these approaches address situations where a user knows, more or less, exactly what he or she is looking for. None of the approaches naturally supports browsing of a collection, searching for a certain type of music, rather than for a specific title. Still, this approach to exploring music archives is one of the most prominent in conventional record stores, where customers explore the total wealth of information in textual form. A similar form of organization is found with most people's private record collections [12]. A main drawback in such hierarchical organizations is that they are based on subjective criteria imposed by the holders of the collections. One way to overcome the subjective nature of the categorizations is to move to a lower level and incorporate signal processing techniques to describe the audio signal of the music and use psychoacoustic models for the final classification.
Among the many works in the field of audio data organization and visualization we are referencing here just two. These two works have played a very important role in the conception of the case study of the Musical Universe:

- In 2001, Pampalk in his diploma thesis "Islands of music" [13], based on the work of Freuhwirth and Rauber [14] as well as the work of Rauber and Freuhwirth [12], developed a system that extracts musical features and uses Self Organizing Maps to construct a 2D visualization environment using the visual metaphor of islands in the sea. His environment displays a pseudo-3D landscape where island formations represent groups of musical pieces that have similar common characteristics. The placement of the various islands and their relative distances indicate how the characteristics that have been chosen to represent the music can distinguish the data.
- In 2001 Tzanetakis et.al [15] proposed a set of features to describe the musical genre in musical collections and shown how their genre classification scheme can be accompanied by visualization either as a histogram, that they named GenreGram, or as a virtual space that they named the GenreSpace. GenreSpace is a 3D interactive space that represents the classes formed after genre classification and after the reduction of the dimensions of the data to 3.

Implementation of the Musical Universe
In order to test the proposed system to real audio data, we were given permission to work with a database of Greek traditional music from an R&D project named "Polymnia"[1]. The musical database is empowered by eXist[2], which is an open-source native-XML and Java-based platform-independent database technology. Each musical piece in the database is represented by an XML file that contains links to the digital file as well as a small sample of it, and content descriptions in a form that is compliant with the MPEG7 standard [16]. The schema of the MPEG7 standard has been extended in order to include more low-level audio features that have been used for the scopes of the project "Polymnia" in order to facilitate content-based search and retrieval. Of these features, four have been selected for the content clustering for the purposes of the Musical Universe:

- *the musical genre*: a text feature that was manually added in the database
- *the genre feature*: a vector of 32 numbers describing the genre of a musical piece based on [17]
- *the meter*: a number describing the meter of the musical piece [18]
- *the beat*: a vector of two numbers describing the beat of the musical piece and the certainty of the beat detection [18]

The selection of these features was driven primarily by the notion of testing the system using features that focus on high level descriptions (such as the genre) and easily perceivable acoustic features (such as the meter and beat).
The Musical Universe, in its experimental form, consists of two main sub-systems:

- Clustering sub-system for Universe formation
- Enhanced user interface sub-system

Clustering sub-system for Universe formation: As stated in the general system description, the clustering sub-system includes three processing steps:

1. Data structure formation: for the purposes of this case study, the final selected dataset was seven-dimensional and consisted of the following features:
   - the textual musical genre (represented by an index)
   - the mean value of the genre feature
   - the meter
   - the beat
   - the beat detection certainty
2. Neural network formation and training
3. Unsupervised clustering to sub-categories: K-means clustering to a maximum of three sub-categories. A clustering for a particular genre of Greek traditional music is shown in FIG. 3. In this figure, the dots depict music files while the stars represent centroids of the classes (represented by different gray levels). As seen in this figure, the unsupervised clustering performed by the k-means algorithm can efficiently recognize the different data characteristics.

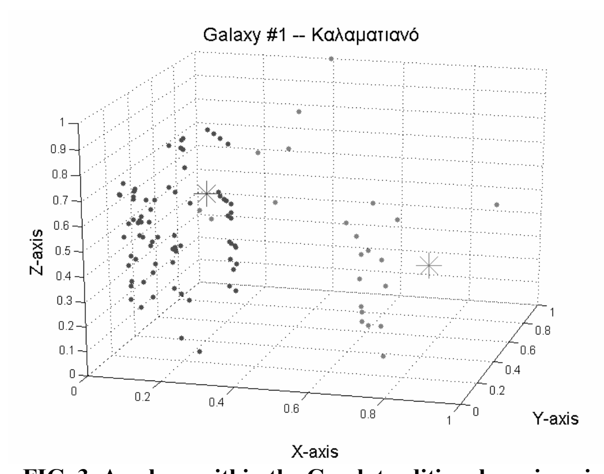

FIG. 3. A galaxy within the Greek traditional music universe

1 http://www.polymnia.gr
2 http://exist.sourceforge.net

Enhanced user interface sub-system - 3D visualization and interactivity: At this stage, the user is able to view and interact with the database content within the 3D visualization sub-system. The user interacts with the system as described before and additionally he/she can get textual information concerning the universe as shown in FIG. 4.

FIG. 4. Statistics of the Musical Universe in textual form

If the user selects a galaxy, the user interface displays the sub-categorization within a galaxy. The interaction once again follows the same principles. This way the user reaches intuitively to the final step of this visualization scenario which is to get the information about a song in a digital musical collection.

5. CONCLUSIONS

An enhanced 3D visualization interface has been developed in order to provide an intuitive data access environment to large digital collections. The system can exploit content-based descriptions and metadata that exist along with the actual data in order to provide a more user friendly representation of low-level descriptions. The imposed clustering is based on 3D SOM structures and k-means. The implementation uses the universe metaphor to represent the content-based clustering within a database and was developed based on OpenGL for high 3D computer graphics efficiency and compatibility. For the evaluation of the system, a case study was presented, which addresses the visualization of a musical archive.

6. ACKNOWLEDGEMENTS

This work was supported by the Greek GSRT grant PENED03#679 and it is in partial fulfillment of the requirements for the PhD degree.

7. REFERENCES

[1] Kohonen, T. (1982), *Self-organizing formation of topologically correct feature maps,* Biological Cybernetics, 43, pp. 59-62.

[2] Kohonen, T. (1995), *Self-organizing maps,* Springer-Verlag, Berlin, Germany.

[3] Kohonen, T. (2001), *Self-organizing maps,* Springer Series in Information Sciences, 30, 3rd edition, Berlin, Germany.

[4] David K. Modjeska (2000), *Hierarchical Data Visualization in Desktop Virtual Reality*, PhD Thesis, Department of Computer Science, University of Toronto.

[5] Kwan-Liu Ma (2000), *Visualizing Visualizations: User Interfaces for Management and Exploration of Scientific Visualization Data*, IEEE Computer Graphics and Applications.

[6] Tamara Munzner (2000), *Interactive Visualization of Large Graphs and Networks*, PhD. dissertation, Stanford University.

[7] Pearson, K. (1901), On Lines and Planes of Closest Fit to Systems of Points in Space, *Philosophical Magazine*, 2/6, pp. 559–572.

[8] Smith, L. (2002), *A tutorial on Principal Component Analysis*, Available on http://csnet.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf.

[9] McQueen, J. B. (1967), Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1, pp. 281-297.

[10] Jiajun Zhu, and Lie Lu (2005), *Perceptual Visualization of A Music Collection*, Proc. IEEE International Conference on Multimedia and Expo (ICME05), Amsterdam, The Netherlands.

[11] Sahori Imai, Shuichi Kurabayashi, Asako Ijichi, Yasushi Kiyoki (2006), *A Music Retrieval System Supporting Intuitive Visualization by the Color Sense of Tonality*, Databases and Applications 2006, pp. 153-159.

[12] Rauber, A. and Freuhwirth, M. (2001), Automatically analyzing and organizing music archives*, in Proceedings of the 5th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2001*).

[13] Pampalk, E (2001), *Islands of music, analysis, organization and visualization of music archives*, Diploma Thesis, Vienna, Austria.

[14] Freuhwirth and M. Rauber, A. (2001), Self organizing maps for content-based music clustering*, in Proceedings of the 12th Italian Workshop on Neural Nets (WIRN01), Vietri sul Mare, Italy.*

[15] Tzanetakis, G., Essl, G. and Cook, P. (2001), Automatic musical genre classification of audio signals, *International Symposium on Music Information Retrieval (ISMIR 2001)*, Indiana, USA.

[16] Martinez, J. eds. (2004), *MPEG7 Overview*, Available from http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm, accessed 29 April 2006.

[17] Tzanetakis, G. and Cook, P. (2002), Musical genre classification of audio signals*, IEEE Transactions on Speech and Audio Processing*, 10/5, pp. 293-302.

[18] Pikrakis, A., Antonopoulos J. and Theodoridis S. (2004), Music meter and tempo tracking from raw polyphonic audio*, Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR 2004)*, Barcelona, Spain.