An Interactive Visual Exploration of Medical Data for Evaluating Health Centres

Tiziana Catarci, Giuseppe Santucci and Sonia Fernandes Silva

Dipartimento di Informatica e Sistemistica
Università di Roma "La Sapienza"
Via Salaria 113, zipcode 00198 – Roma, Italy
[catarci/santucci]@dis.uniroma1.it
Phone & Fax +39 06 49918331

It is well known that visualisation techniques are suitable for an effective exploration of large data sets. However, not much research has been carried out on applying visualisation techniques in the analysis of medical data apart from image data. Medical information systems collect a vast amount of monitored clinical data coming from specialised machines. The task of accessing and interpreting the portions of data that are relevant to the identification of a specific clinical problem can become a hard task. In order to address this problem, we propose to enhance medical information systems by providing an interactive visual exploration of large data sets. In this paper, we introduce our approach and show its use in order to enhance the quality and efficiency of health services in haemodialysis centres.

Categories and Subject Descriptors: H.5.2[Information Interfaces and Presentation]: User interfaces-graphical user interfaces (GUI); J.3[Life and Medical Sciences]: Medical information systems, health

Keywords: Medical information systems, medical data, user interfaces, information visualisation

1. INTRODUCTION

Information visualisation and data sharing constitute essential issues at development of medical information systems. In those systems there is the need of accurate storing, rapid access and sharing of a large amount of medical information, such as patient information, clinical symptoms, disease diagnosis and treatments.

Concerning the data sharing issue, in (Grimson, Grimson and Hasselring, 2000) it is described that a patient is managed by a team of health care professionals each specialising in one aspect of care. The quality of such shared care critically depends on the ability to share information easily among care providers. Recent mobile technologies (smart cards, telemedicine) are adopted in medical information systems in order to provide data sharing in real-time among different health care organisations. Easy mobility of smart cards (Shelfer and Procaccino, 2002) can be a potential solution for managing patient's medical records, whereas telemedicine connects geographically

Manuscript received: 28 September 2002
Communicating Editor: Associate Professor Jim Warren

dispersed health care facilities via video and telecommunications in order to perform remote clinical diagnoses and surgeries (Huston and Huston, 2000).

However, data sharing is not the only issue. The correct and complete interpretation of medical information for further investigation and analysis is vital for effective health-related decision making in order to improve the quality of care. In fact, some researchers (Plaisant, Mushlin, Snyder, Li, Hellen and Shneiderman, 1998; Brinck and York, 1998) have pointed out that much attention is put on developing standards and frameworks for gathering and sharing medical records, but little effort is devoted to design appropriate user interfaces and visualisation techniques for presenting and exploring medical records. Nevertheless, there are applications specialised in information retrieval of medical vocabulary databases, and there are many sophisticated visual techniques in the field of medical imaging, which require complex 3D visualisations and specialised knowledge for user-interaction.

Within this context, we have developed a system that deals with the problem of automatic visualisation and exploration of large sets of medical records in an appropriate way, where the visual representation is carefully chosen to effectively convey all and only patient information content. The system is mainly characterised by an homogeneous environment comprising different tools that fit the various users' tasks. Such tools exploit several visualisation paradigms, where relevant data are visualised and interactively explored.

In this paper, we see the application of this system to evaluating the quality of health services in haemodialysis centres. Haemodialysis centres of middle dimensions perform not less than 10,000 haemodialysis sessions per year and the number of patients that need haemodialysis is constantly increasing. A number of applications have recently been presented in the literature (Bellazzi, Larizza, Magni, Bellazzi and Cetta, 2001) as potential solutions for improvement of quality of care of haemodialysis treatments. Most of them use the telemedicine technology for automated remote monitoring of haemodialysis sessions.

However, as we previously said, an effective monitoring and sharing of data in haemodialysis sessions is not enough. It is fundamental that all data be appropriately visualised in order to help in extracting information about patterns, anomalies, and trends in such sessions so as to monitor and improve quality of care of haemodialysis centres. The visual exploration of such data helps, for example, to identify how many haemodialysis sessions reach the objectives prescribed in the therapeutic plan and to analyse what could have caused the objective of the therapeutic plan not to be reached. The foregoing would be instrumental in guaranteeing the patient safety during the treatment and in taking appropriate preventive measures, such as prevention of symptoms among sessions (e.g. hypertension), effective interventions in emergency situations, and so on.

The paper is structured as follows. Section 2 presents an overview of the main visualisation techniques for interactive exploration of medical records. Section 3 describes the haemodialysis scenario, comprising the issues involved in evaluation of haemodialysis sessions. Section 4 describes the basic functionalities and the architecture of our system. Section 5 recalls the key issues of the system regarding 3D data visualisation. Section 6 describes an example of the application of the system. Section 7 reports some usability experiments. Finally, future work and conclusions are discussed in Section 8.

2. VISUALISATION TECHNIQUES IN MEDICAL DATA

The availability of rapid display of large data sets and *direct manipulation* strategies (Shneiderman, 1983) have in the last years given rise to a large diffusion of visualisation techniques for interactive visual exploration of such data sets. In a visual data exploration (Keim, 2001), the user is directly involved in the discovery process in order to detect patterns in data (e.g., trends, clusters, and exceptions). Visual data exploration is more intuitive and often more effective than any automatic statistical technique and machine learning since the central focus of data exploration is the human visual perception and interpretation.

In Keim (2001), the visualisation techniques are categorised according to the following criteria: the data to be visualised, the technique itself and the interaction and distortion methods. Concerning the first criteria, data are categorised as linear, 2D/3D and multidimensional data, network, hypertext, hierarchies, etc. There are geometric and pixel-oriented techniques which are suitable for multidimensional data, and hierarchical and graph-based techniques are suitable for representing hierarchies and networks respectively. Moreover, there are distortion methods which make better use of the available screen space, by integrating detailed and contextual views of information and are well suited for linear and hierarchic data.

The medical domain is appropriate for information visualisation. There are many sophisticated applications in the field of medical imaging (X-rays, anatomical structures, brain scans, etc) and in the field of information retrieval that are used in medical vocabulary databases. However, we are specifically concerned with the visualisation of patient records. A patient record is defined in Chan, Jiannong, Chan and Young (2001) as a representation of a key repository for information concerning health care. In medical information systems, the so-called electronic medical record is considered an evolution of the traditional paper-based patient record. Since the electronic medical record is always available in several places simultaneously, the information can be transferred. It can also support different record views for different health professionals (Grimson *et al*, 2000).

In this section we recall some interesting visual techniques which may be particularly suitable for visualising patient records. We specifically focus on visual techniques for displaying linear (e.g., medical history of patient), hierarchic (e.g., classification of medical terms in a medical vocabulary database) and multidimensional data (e.g., health data summary of patients).

Several visual techniques for linear data have been proposed in the literature. The distortion techniques, Perspective Wall (Mackinlay, Robertson and Card, 1991), Table Lens (Rao and Card, 1994), and Fisheye View (Furnas, 1986), for example, make use of focus+context, where portions of data of interest are shown with a high level of detail and other portions with a lower level of detail. In Perspective Wall, a 2D layout is distorted into a 3D visualisation, in which the user views the information on a wall in a 3D perspective. Besides distortion methods, other systems make use of interactive timelines (Kumar, Furuta and Allen, 1998). Such systems allow the user to explore relationships among historical events specifically for visualisation of time-oriented linear data (Silva and Catarci, 2000).

Considering the visualisation of hierarchies, notable proposals are Tree Maps (Johnson and Shneiderman, 1991), Cone Trees (Robertson, Mackinlay and Card, 1991), the tool Hyperbolic Browser (Lamping, Rao and Pirolli, 1995) and SinVis Magic Eye View (Kreuseler, Lopez and Schuhmann, 2000), where the hierarchy is first laid out in a 2D space and then mapped onto a hemispherical surface. Other systems making use of visual metaphors for a better visual perception of hierarchies are Information Pyramids (Andrews, Wolte and Pichler, 1997), and Botanical Visualisation (Kleiberg, van de Wetering and van Wijk, 2001).

Some visual techniques are used for the effective exploration of multidimensional data, such as geometric e.g., landscapes (Wright, 1995), parallel coordinates (Inselberg and Dimsdale, 1990) and pixel-oriented e.g., scatterplot, spiral technique (Keim and Kriegel, 1994), circle segments (Keim, 1997) techniques. For example, the approach presented in Keim (1997) makes use of a circle-segment technique which visualises k-dimensional data (resulting from the execution of a specific query) as a circle divided into k segments. The work of Yang (2000) presents a tool using 3D grand tour (Asimov, 1985; Cook, Buja and Cabrera, 1995), and volume rendering (Becker, 1997) for visualising very large data sets (with high dimensionality and high number of records).

The visualisation techniques described above are used in conjunction with statistics and data mining algorithms[1] for a faster data exploration process (Keim, 2001; Shneiderman, 2001). This approach is generally called visual data mining in the literature. Notable proposals can be seen in Hinneburg, Keim and Wawryniuk (1999); Ankerst, Ester and Kriege (2000) and Inselberg and Avidan (2000).

It is worth noting that for an effective exploration of data, visual techniques must be integrated with suitable interaction techniques, following the principle of *visualisation information mantra* defined in Shneiderman (1996): *Overview first, zoom and filter, then details-on-demand*. This means that, starting from an overview of a large data set, one may zoom and filter this overview to extract a data subset. Then, more details can be obtained from the selected data subset. In this context, notable proposals include Spotfire (Ahlberg and Wistrand, 1995) and tools that explore the dynamic query and starfield display approaches (Ahlberg and Shneiderman, 1994).

Additionally, we recall that a data set can be large in terms of a high number of records and/or a high number of data attributes. Concentrating on applications that are specific to the visualisation of medical data, we identify two ways of visualising the patient records in the literature:

- *vertical visualisation*, that corresponds to the medical overview visualisation of one patient at a time. This kind of visualisation involves a high number of medical data attributes and small data sets;
- *horizontal visualisation*, that corresponds to the visualisation of some medical attributes of several patients. This kind of visualisation involves a low number of attributes and large data sets.

Each kind of visualisation is suited for specific medical purposes. The vertical visualisation is mainly used for diagnosis and treatment purposes, since it visualises the medical data attributes (biographical and clinical data) and events (consultations, medications, laboratory results, etc) related to one patient at a time. The respective summary view of some medical aspects related to several patients is mainly used for statistical evaluations of health communities and organisations in order to support clinical decision-making.

The medical systems that explore vertical visualisation are more common in literature. Since medical records are generally characterised by the temporal dimension (patient's medical history), most vertical visualisations use the interactive timeline technique (Powsner and Tufte, 1994; Plaisant *et al*, 1998; Kilman and Forslund, 1997). New approaches that explore vertical visualisation are arising such as the work described in Chan *et al* (2001), which integrates smart card technology and web-based applications. All medical information of a patient on a smart card can be visualised from a web browser.

The horizontal visualisation is exploited in medical systems that use data mining techniques. A proposal for exploratory mining of specific medical data sets is described in Hsu, Lee, Liu and Ling (2000), where the information is presented through the use of histograms for statistical purposes only. Scatterplot is a visualisation technique suited for horizontal visualisation. In this technique, a record is a geometric object and its position in the space is determined by two (2D space) or three (3D space) attribute values, while its size, shape, colour or texture are determined by other attribute values.

[1] Data mining is well defined in Witten and Frank (2000) as the extraction of implicit, previously unknown, and potentially useful information from data.

Within this context, one great benefit of our system is that it integrates the vertical and horizontal visualisations in the same environment. Regarding the evaluation of haemodialysis sessions, two types of information can be visualised in our system: summary data, which represent a summary view of many haemodialysis sessions, and detailed data, which represent a detailed view of a single session or related sessions (e.g., same patient, same machine, etc). Once the professionals have accessed the summary and detailed haemodialysis data, they are ready to investigate all significant data correlations that can be valuable for an effective evaluation of haemodialysis sessions. The next sections go into more details about the aforementioned integration.

3. HAEMODIALYSIS SCENARIO

As we pointed out in the introduction, we mainly concentrate on the following issues for the evaluation of haemodialysis sessions: how many haemodialysis sessions reach the objectives prescribed in the therapeutic plan, or what are the causes of failure. In the following, we describe the actual scenario in which the evaluation takes place. First of all, an automatic procedure (Intelligent Data Analysis Technique) is adopted, through which clinical data coming from a haemodialysis session is monitored, collected and automatically analysed. According to Bellazzi *et al* (2001), the quality of haemodialysis sessions is evaluated based on the following features:

- Dialytic dose
- Efficiency of the extra-corporeal circuit
- Body water reduction and hypotension episodes
- Nurse intervention

The dialytic dose is a fundamental measure for evaluating the quality of haemodialysis. It is calculated through a parameter, namely urea clearance fractional rate parameter (KT/V). This parameter describes the efficiency of the removal of protein catabolism products (urea, creatinine), where K is the clearance of diffuse elimination of urea and is dependent of the hematic flux in volume/time (QB ml/min), V is the urea distribution volume of the patient without excess of water, which can be estimated from the body weight, and T is the haemodialysis duration. Hence KT/V is evaluated by monitoring hematic flux (QB), weight loss (WL) at the end of haemodialysis and haemodialysis duration (T).

The efficiency of the extra-corporeal circuit is evaluated by measuring the arterial pressure (AP, before the dialyzer), and the venous pressure (VP, after the dialyzer). Such pressures are monitored to control the overall efficiency of the dialysis process, given that an elevated value of such pressures may cause a deterioration of vascular accesses.

Body water is eliminated by means of a different pressure (ultrafiltration) between the blood and the dialysis solution. A fast removal of extra-cellular water may conduct to the hypertension event, where the patient may collapse. The prevention of hypertension is controlled by measuring the cardiac frequency, and the systolic and diastolic pressures. As indicated in Bellazzi *et al* (2001), if during a dialysis session the systolic and diastolic pressures fall below a certain threshold, the hematic flux may be manually decreased in order to avoid collapse. In this case, a reduction of the KT/V parameter is induced by a blood pressure problem.

Finally, it must be noted if and how many times a nurse intervention is needed in a haemodialysis session so that the therapeutic plan can be respected.

According to the features described above, six relevant parameters for evaluating haemodialysis sessions have been chosen: hematic flux, arterial and venous pressures, haemodialysis duration, weight loss, and nurse intervention. The chosen parameter values are compared with the reference values established in the prescribed therapeutic plan (PTP), in order to obtain a quality index of the overall treatment and discover the causes of unsuccessful treatments. The parameters and their corresponding reference values can be seen in Table 1.

Relevant Parameter	Reference Value
Hematic Flux (Qb ml/min)	≥ to that of PTP
Arterial Pressure (mmHg)	≥ –250
Venous Pressure (mmHg)	≤ 350
Haemodialysis Duration (min)	≥ to that of PTP less 10 min
Weight Loss	≤ to that of the PTP + 1, 1%
Nurse Intervention	yes/no

Table 1: The relevant parameters for evaluating haemodialysis sessions

The result of a haemodialysis session is categorised as follows:

- Successful Session: all parameter values agree with those of the PTP without nurse intervention.
- Successful Session with intervention: all parameter values agree with those of the PTP with nurse intervention.
- Unsuccessful Session: one or more parameter values do not agree with those prescribed.

Data returned from the automatic procedure are confronted with the manual analysis performed by the nurses at the end of each haemodialysis session. In unsuccessful sessions, it has to be possible to carry out a data analysis in order to emphasise the causes of the failures, some patterns or anomalies (e.g. increased or decreased trend of some intra-session parameters). The evaluation of the significant difference between the two analyses, automatic and manual, reveals the subjective interpretation of a part of the PTP by the nursing personnel.

Finally, the necessity of a data mining activity to summarise haemodialysis outcomes and to automatically highlight the intra- and inter-session variability of the clinical parameters was indicated in Bellazzi *et al* (2001).

4. SYSTEM FUNCTIONALITIES AND ARCHITECTURE

In this section, we present the basic functionalities of our system and relate them to the main users' tasks. We consider that a haemodialysis treatment is evaluated by different professionals at different levels, from a global evaluation of a haemodialysis centre to a detailed investigation of an individual treatment. It follows that the typical users of the system are hospital directors, doctors, specialists, and nurses.

The main tasks of the aforementioned users are:

- Analysis of the overall quality of health services in a haemodialysis centre;
- Analysis of an individual treatment (haemodialysis sessions related to a single patient);
- Analysis of a single haemodialysis session.

The medical procedures for corrective therapeutic actions are derived from the results of these analyses.

The current system determines what information to present and how the relevant information should be presented and explored depending on the task being performed, as indicated in the following.

Let us consider that a director of an haemodialysis centre is interested in having a global view of health services in her/his centre. In this case, the overall summary data about the results of haemodialysis sessions are presented, such as percentage of the successful and unsuccessful sessions and the percentage of the causes of unsuccessful sessions. The goal here is to allow the director to have some basic ideas about the quality of the overall haemodialysis treatment. The system presents the overall summary data using standard statistical graphs such as histograms, pie charts, etc.

Once the director has obtained the overall summary data described above, s/he (or another doctor indicated by her/him) is ready to investigate an individual treatment in detail. In this case, an overall detailed data view about the patient is presented. Such data represents the average values of the parameters indicated in Table I. The goal here is to allow the professionals to have a global idea about the impact of these parameters on the haemodialysis sessions related to a single patient, in order to identify trends and anomalies in some intra-session parameters.

The system adopts two visual techniques for visualising the foregoing data: a 2D timeline and a 3D scatterplot. In the 2D timeline, the average values of a specific parameter (e.g., arterial pressure) are visualised in a chronological sequence of haemodialysis sessions. This kind of visualisation is limited since it presents only one parameter at a time. To overcome this limitation, the system may visualise more parameters at a glance in an advanced 3D scatterplot. In this kind of visualisation, the data are visualised as points or little cubes (depending on the number of records) and their position in the 3D space, size and colour are determined by patient identification, session numbers and average values of the parameters. The user explores this multidimensional data set by zooming, filtering and selecting a specific point. In turn, more details can be obtained for the selected point, which corresponds to a single haemodialysis session.

Finally, the analysis of a single session is performed by presenting summary and detailed data about a particular session. The summary data represents the quality of a single session, indicating that a session is successful or not and presenting a quality index of such a session. On the one hand, the system adopts the 3D scatterplot for visualising the summary data. On the other hand, the system adopts the 2D timeline for visualising the detailed data of a single session, where a temporal sequence (time instants) of values of a specific measure (e.g., arterial pressure) taken during such session are visualised.

It is worth noting that our system uses a 3D scatterplot as the main visual technique given that there are few involved attributes, but it may adopt other visualisation techniques, such as Parallel Coordinates (Inselberg and Dimsdale, 1990), Spiral (Keim and Kriegel, 1994) and Circle Segments (Keim, 1997), which are suitable for an effective data exploration whenever the number of dimensions is high.

In the next subsection, we describe the overall system architecture with reference to the functionalities described above.

4.1 System Architecture

The system architecture is illustrated in Figure 1. Haemodialysis data are automatically collected through a specialised software for data acquisition (DIALMASTER© software). A program periodically (or on-demand) filters the data and creates a filtered database. Starting from this filtered database, a program periodically (or on-demand) creates a database of summary data through the specification of a temporal query (*Temporal Query* module). The parameters needed in a temporal query are defined in a knowledge base of a specific haemodialysis centre. The knowledge base management is performed by a knowledge acquisition tool (KA tool). In the rest of the paper we do not deal further with these components.

Finally, the *User Interface* module represents the front-end of the system, where the different users may visualise the database of filtered data and/or the summary data exploiting different kinds of visualisation (2D and 3D visualisation modules).

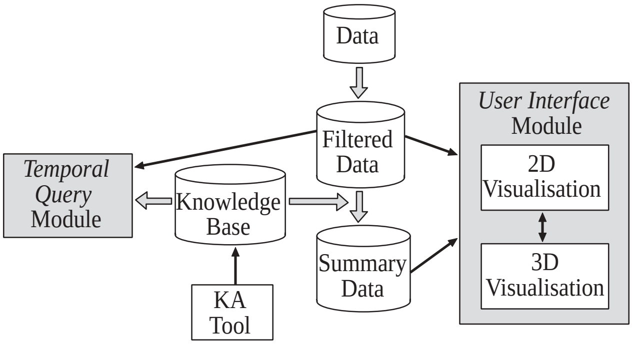

Figure 1: The System Architecture

In the remaining part of this paper we mainly concentrate on the 3D visualisation module.

5. 3D VISUALISATION MODULE

The 3D visualisation module is a specialisation of the Dare system (Catarci, Costabile and Santucci, 1999). In this section we recall the key issues of Dare and discuss its specialisation to our specific medical domain. The Dare system deals with the general problem of automatic visualising information contained in a database.

The aim of the system is to produce a correct, complete, and possibly highly effective visual representation. Such properties of the visual representation are made possible by the consistency of a set of logic rules contained in a knowledge base. These rules express facts and relationships about the visual domain, the data domain, the mapping between both domain, and the perceptual domain, namely:

1. *Visual rules* characterise the different kinds of visual symbols (e.g., they list the *visual attributes*, which are associated with the different kinds of visual symbols);
2. *Data rules* specify the characteristics of the data model, the database schema, and the database instances;
3. *Mapping rules* specify the link between data and visual elements;
4. *Perceptual rules* tell us how the user perceives a visual symbol (i.e., a line, a geometric figure, an icon, etc.).

5.1 Visual Rules

Visual rules define the types of fundamental visual symbols and their associated visual attributes. Examples of main predicates used in such rules are:

- *VisSymbType(x)* means that *x* is a visual type. We consider four predefined types of visual symbols only, i.e. *point* (a triple $\langle x, y, z \rangle$ in the space), *line* (a continuous sequence of points), *plane* (a surface passing through line) and *3D figure* (the portion of space enclosed by a set of planes).
- *VisSymb(x)* means that *x* is a visual symbol.
- *VisCateg(x)* means that *x* is a *visual category*. A visual category is a specialisation of a visual type, characterised by fixing the value of some visual attributes.
- *VisAttr(x)* means that *x* is a visual attribute. Visual attributes characterise the external properties of visual symbols, i.e., size, colour, texture, orientation, shape, coordinates in the space (Bertin, 1983).

- *IsOfType(x, y)* means that *x* is a visual symbol of type *y*.
- *IsOfCat(x, y)* means that *x* is an instance of a visual category *y*.
- *ISA(x, y)*, where *x* and *y* are visual categories, means that the visual symbols that are instances of *x* are also instances of *y*.
- *HasAttr(x, y)* means that *y* is a visual attribute of *x*. Note that if *x* is a visual symbol type (or a visual category), then *y* is a visual attribute, whereas, if *x* is a visual symbol, then *y* is a value of a visual attribute.
- *IsAValue(x, y)* means that *x* is a value of the domain of the visual attribute *y*.

5.2 Data Rules

Data rules define the characteristics of the data model used, of the data schema and the data instances. In the following, we list the basic predicates used in the data rules, with the associated intended meaning.

- *DataType(x)* means that *x* is a data type, e.g., entity, relationship, class, etc.
- *IsOfType(x, y)* means that *x* is a data category of type *y*.
- *HasArg(x, n)* means that the data category *x* has *n* arguments. Generally speaking, *n* is equal to 1 for classes and greater than 1 for relationships.
- *IsInst(x, y)* means that *x* is an instance of the data category *y*.
- *Comp(x, y, z)* means that *z* is the *y*-component of the relationship instance *x*.
- *ISA(x, y)* where *x* and *y* are data categories, means that the instances of *x* are also instances of *y*.
- *DataAttr(x)* means that *x* is a data attribute, such as age, name, SSN, address, etc.. Data attributes characterise the properties of data elements.
- *HasAttr(x, y)* means that *y* is a data attribute of *x*. Note that if *x* is a data category, then *y* is a data attribute, whereas, if *x* is an instance, then *y* is a value of a data attribute.
- *IsAValue(x, y)* means that *x* is a value of the domain of the data attribute *y*.

5.3 Mapping Rules

The basic predicate is *Rep(x, y)*, meaning that *x* is the visual representation of *y*, where *x* is an element of the visual realm, and *y* is an element of the data realm. For example, the following rule

Rep(Zcoordinate, blood-pressure)

states that blood pressure is associated with the Z axis.

5.4 Perceptual Rules

In our work, we are not interested in defining specific perceptual rules (there is a branch of psychology that studies the visual impact that the objects have on the human psyche (Cornsweet, 1970; Arnheim, 1954). Instead, we aim at defining a framework in which generic perceptual rules may be formalised. To do that, we define a basic predicate, *GoodVis(x, y)*, which means that *x* is a good visual representation for *y*. Note that *GoodVis* is a polymorphic predicate, i.e. *x* is a generic element of the visual realm, as well as *y* is a generic element of the data realm.

As a matter of fact, human beings have a better perception of some objects rather than others. The form of such objects and the distance among them influence the perceptual capabilities of the human brain. The psychology of gestalt studies the perceptual capability of human beings. In our experiments we used some feedback coming from the end users to set the perceptual rules we need in the biomedical context.

5.5 3D Visualisation Prototype

In specialising Dare to our medical context, we decided to start designing and implementing a first prototype focusing on a precise and critical subproblem very common in the biomedical context: to allow the user to dynamically interact with a large amount of data, a situation in which usual browsing techniques are totally useless. Several attempts have been made to deal with such a problem, most of them focusing on the idea of visually restricting the data set by allowing the user to change the value of the attributes involved in his/her query through suitable widgets (Ahlberg and Shneiderman, 1994; Williamson and Shneiderman, 1992), in order to restrict the final result.

The case study we discuss in the following aims to solve the problem in a very general way: given a generic huge query result, how possible is it to automatically generate a visual representation that allows the user to dynamically and effectively rephrase her/his query in order to reduce the size of the outcoming data? Moreover, it is very common to look at large medical data sets to get some statistical information and not for selecting single data values. To deal with this problem, we use only a portion of the overall framework described so far, namely the visualisation of the database instances. Moreover we assume that the user is interacting with a single relational table and that the number of involved attributes is little (e.g., five).

It is worth noting that the above restrictions are not so severe: it is a well known fact that select-project queries represent the majority of the queries asked by users (Chandra and Merlin, 1977) and it is very seldom that the number of the attributes involved in the selection is greater than five.

The problem can be defined more precisely as follows: having a relational table representing the result of a selection involving few attributes through range conditions, which are the most suitable visual symbols and visual variables able to effectively convey to the user the distribution of the tuples across the selection attribute domains?

In the following, we detail the choices and the strategies devised to cope with such a problem.

Concerning the visual symbols, it is quite evident that the point is the best candidate to represent hundreds or thousands of tuples on the same screen. The visual attributes available for the point are the three coordinates, the colour hue, and the colour lightness, resulting in five different visual attributes. In the following we give more details about the rules we use in the prototype. Again, although we mentioned user interaction in the above example, we do not formally consider dynamic aspects of the visual representation in the following.

- **Visual rules –** We assume that the set of visual rules includes a characterisation of both point and a subset of 3D figures (e.g., little cubes) to be used instead of points when the number of tuples to be represented on the screen is not so big. The visual rules contain the description of the available visual attributes as well.

- **Data rules –** Concerning data rules we need just the subset of data rules able to specify the characteristics of the relational model, so handling the simple concepts of relation, tuple, attribute, and domain. Here we report, as an example, the basic set of data rules chosen for capturing the relational model constructs and the relation used in the above example.

  DataType(relation).
  DataType(domain).
  IsOfType(x, relation) → HasArg(x, 1).
  IsOfType(x, domain) → HasArg(x, 1).
  IsOfType(patient, relation).
  IsOfType(total – order – domain, domain).

IsOfType(string, domain).
IsOfType(real, domain).
ISA(real, total – order – domain).
ISA(string, total – order – domain).
DataAttr(name).
DataAttr(weight).
DataAttr(height).
HasAttr(patient, name).
HasAttr(patient, weight).
HasAttr(patient, height).
IsOfType(name, string).
IsOfType(weight, real).
IsOfType(height, real).

Note that, as stated by the ISA rules, the available domains being implemented on a computer, it is always possible to define a total order on them.

- **Mapping Rules –** Here, being tuples represented by points or little 3D figures, the emphasis is on attribute domain representations. The available set of rules specify the association among the relational attributes and the available set of visual attributes. As an example, we show the basic associations between the domains defined for the relation *patient*.

  Rep(colour, weight).
  Rep(Ycoordinate, height).
  Rep(Xcoordinate, weight).

- **Perceptual Rules –** The general, challenging objective of perceptual rules, as described above is here tailored to give us information about the best representation of relational attributes. Being all attribute domains of the type total-order-domain, it is possible to associate any attribute with any of the 3D coordinates or with the size. Moreover, stating a suitable order on non ordered visual attribute domains (e.g., from low frequency to high frequency for colours), it is possible to associate any relational attribute with any visual attribute. However, it is quite evident that the height of a person is better represented by the Y coordinates than by colour. The goal of the perceptual rules is to provide us with such information. To do that, we slightly modify the rule *GoodVis(x,y)*, adding a third numeric field, giving a measure of the goodness of the representation. As an example, considering a patient height, we can state the following rules:

  GoodVis(height, colour, 0.2).
  GoodVis(height, Zcoodinate, 0.3).
  GoodVis(height, Xcoordinate, 0.6).
  GoodVis(height, Ycoordinate, 0.8).
  GoodVis(height, size, 0.8).

  Through such rules we state that even if in principle colour, size, and one of the 3D coordinates are suitable to visually represent the height of a person, size and Y coordinate are the best candidates to do that. Note that, once a visual representation has been selected, it associates each data attribute with a different visual attribute, according to the general rule structure.

Using the above rules, the system, given a user query and the related result, can automatically choose the best visual representation. First of all, through the result cardinality, it discovers the suitable visual elements to adopt to represent tuples (points and/or little cubes). After that, the system tries to associate the relational attributes appearing in the selection part of the query with the available visual attributes (i.e., colour and 3D coordinates) maximising the goodness of the representation. Note that even if the problem is, in general, computationally intractable, because of the assumption on the number of attributes involved in the query, it is possible to explore the whole space of solutions, getting an exact result.

6. A WORKING EXAMPLE

The overall system has been evaluated against data coming from the Nephrology Centre of the Civil hospital of Vigevano. Data have been collected in order to evaluate the centre's efficiency, in terms of successful hemodialyses. In particular, among other more technical pieces of information, synthesis data have been arranged in three relational tables described below. The first one collects summary data about hemodialyses:

Table 1:	PAZIENTE	Patient ID
	DIALISI	Haemodialysis #
	FLUSSO_PRESCRITTO	Prescribed blood flux
	T_INIZIO	Haemodialysis starting time
	T_FINE	Haemodialysis ending time
	T_PRESCRITTO	Prescribed duration
	P_INIZIO	Initial patient weight
	P_FINE	Final patient weight
	P_PRESCRITTO	Prescribed weight loss
	NI	Nurse intervention

The second relational table collects measures taken during hemodialyses (e.g., arterial and venous blood pressures):

Table 2:	PAZIENTE	Patient ID
	DIALISI	Haemodialysis #
	SEGNALE	Signal code
	TEMPO	Time
	VALORE	Value

The last one contains an overall haemodialysis evaluation and it is computed comparing the prescribed values (PTP) against theactual ones:

Table 3:	PAZIENTE	Patient ID
	DIALISI	Haemodialysis #
	NI	Nurse intervention
	NIR	Boolean nurse intervention(0,1)
	QB	Blood flux
	QBR	Boolean blood flux (0,1)
	T	Haemodialysis duration

TR	Boolean haemodialysis duration(0,1)
WL	Weight loss
WLR	Boolean weight loss(0,1)
PA	Arterial pressure
PV	Venous pressure
PAR	Boolean arterial pressure(0,1)
PVR	Boolean venous pressure(0,1)
SUCCESSO	Overall success (0,1,2)
QTOT	Haemodialysis quality

We recall (see Section 3) that six parameters are used to evaluate an haemodialysis: nurse intervention (i.e., manual parameter adjustment), blood flux, haemodialysis duration, weight loss, arterial and venous pressures. Such values (i.e., their medians) are stored in the third relational table together with a boolean indication (0=success, 1=failure). An haemodialysis can be globally successful without manual parameter adjustments (0), successful with manual adjustments (1), or unsuccessful (2). Moreover, a global evaluation is computed (QTOT), weighting the difference between the pre-scribed values and the actual ones, in order to have an overall estimation of the haemodialysis quality. In the following, we show some possible interactions with the system using real data coming from more than 400 hemodialyses over four patients. The typical user of the system is a director interested in having an overall vision of the efficiency of the Nephrology Centre and/or a doctor zooming in on data about a little set of hemodialyses in order to understand reasons for failure or success.

Directors usually deal with very high level representations of the data collected from haemodialyses. As an example, consider the screen-shot shown in Figure 2. Looking at the bar charts, one can immediately get some aggregate pieces of information about global successful and unsuccessful haemodialyses. Moreover, it is possible to have overall views about single patients or on a single haemodialysis machine.

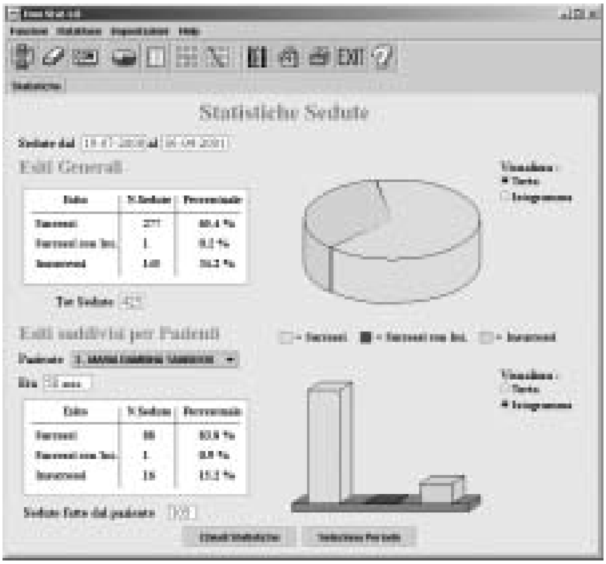

Figure 2: Overall statistics about haemodialysis

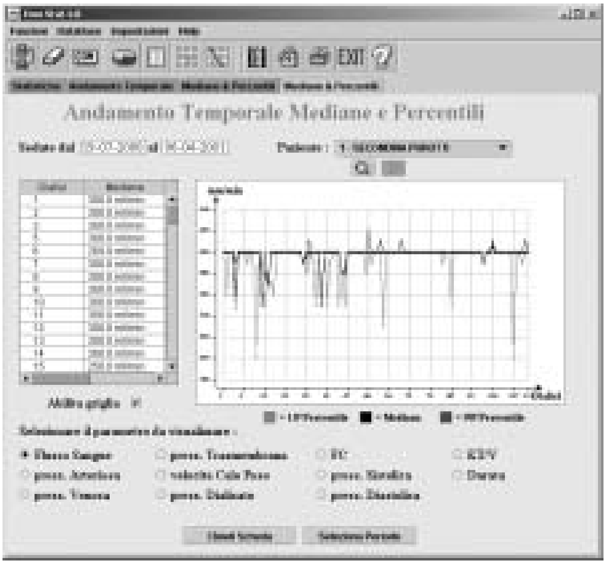

Figure 3: Detailed data about blood flux of a single patient

A doctor is likely to be interested in inspecting details about more specific data such as, e.g., medians about weight loss and blood flux. As an example, a doctor may activate the 2D visualisation shown in Figure 3, displaying detailed information about the blood flux of a single patient across all his/her haemodialyses.

If the doctor is looking for some data correlations or significant data patterns s/he can activate the 3D visualisation module putting together up to five parameters at a time. As an example, assume that the doctor, starting from the visualisation shown in Figure 3, needs more details about blood

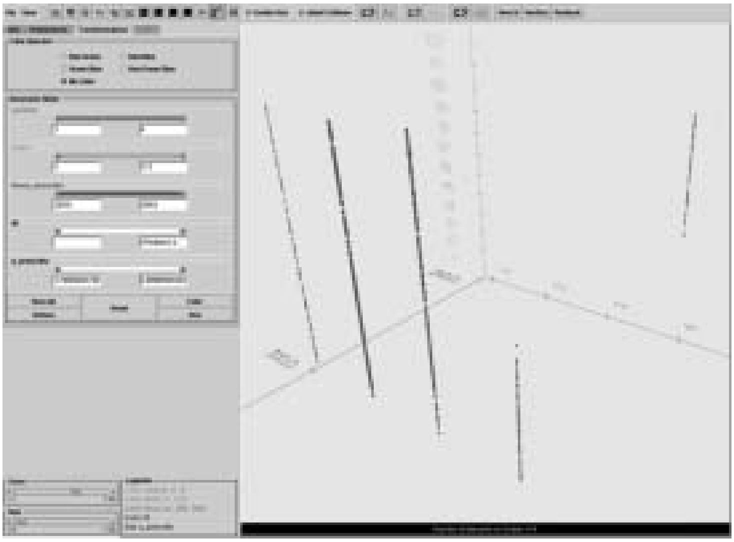

Figure 4: Blood flux and weight loss prescriptions

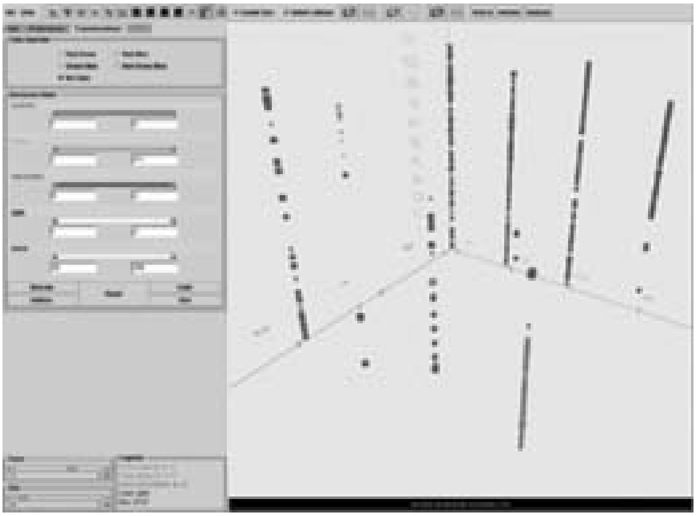

Figure 5: Overall haemodialyses evaluation

flux and starts weight loss prescriptions for each patient together with the presence/absence of nurse intervention. The system, harnessing the data on the first relational table, activates the 3D module producing the representation shown in Figure 4, where each cube represents a single haemodialysis. The picture shows data about four patients (associated with the X axis), each of them subject to several haemodialyses (about 100 and associated with the Y axis), a prescribed blood flux (associated with the Z axis), a prescribed weight loss (associated with the cube size), and the presence or absence of nurse interventions (associated with the cube color). It is quite easy to see that (1) the prescribed blood flux for Patient 4 has been reduced from 300 ml/minute to 250 ml/minute after the first set of haemodialyses and (2) that the prescribed weight loss (the bigger the cube the higher the prescription) for Patient 2 presents some quick variations in the most recent haemodialyses. Moreover, exploiting colors (red and blue) it is clear that while Patient 1, Patient 3, and Patient 4 performed their haemodialyses without manual parameter adjustments, Patient 2 required nurse interventions in four haemodialyses (red cubes).

In order to deepen her/his analysis, the doctor activates a new visualisation based on the third relational table. The result is shown in Figure 5, where, again, each cube represents a single haemodialysis, and the user focuses her/his attention on patients (X axis), haemodialysis # (Y axis), the overall success (Z axis), the local success about blood flux (color), and the global haemodialysis quality (size). The user can easily grasp that (1) most of hemodialyses have been successful without nurse intervention and (2) that Patient 4 presents an initial sequence of unsuccessful hemodialyses in which there is a failure in the blood flux (red cubes) followed by a successful sequence (blue cubes). Comparing Figure 4 with Figure 5, it easy to see that the unsuccessful sequence has been broken, not by performing better hemodialyses but by reducing the prescribed blood flux!

Still considering the third relational table, the user has partitioned the set of haemodialysis of each patient in intervals, where the bar dimension represents the average weight loss and the color represents the average haemodialysis quality (see Figure 6). It is clear that (1) Patient 1 presents a low average weight loss in the first 18 hemodialyses and that (2) hemodialyses between 55 and 73 present a very high quality.

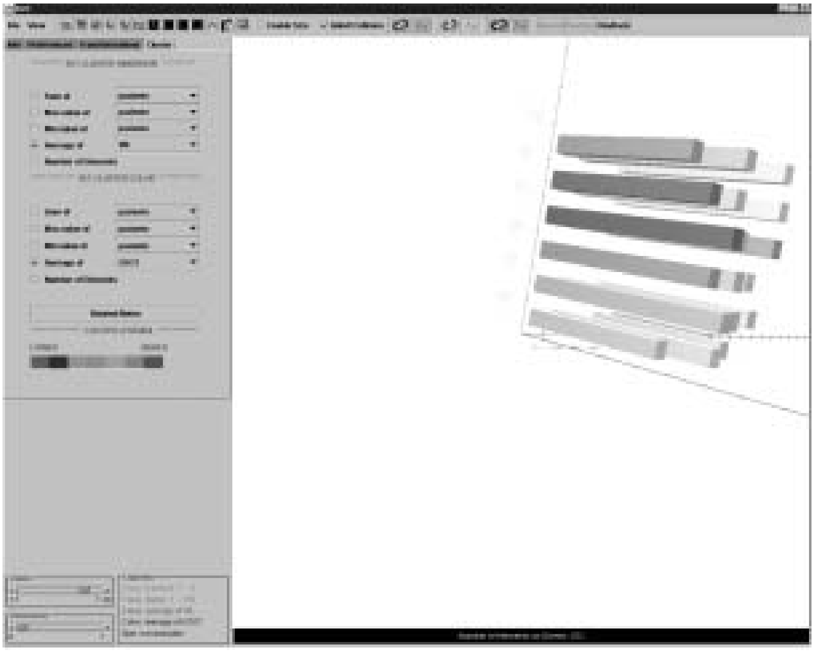

Figure 6: Statistics about quality and weight loss

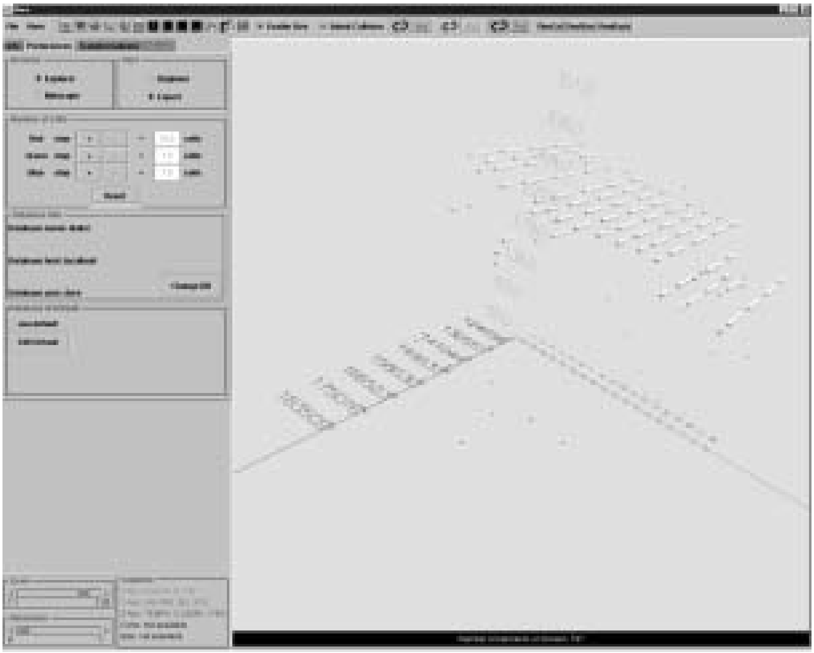

Figure 7: Blood flux of Patient 1 in his/her first 18 haemodialyses

The doctor inspecting such a visualisation zooms (using data coming from the second relational table) within the first 18 haemodialyses of Patient 1, detailing the blood flux values during each haemodialysis (see Figure 7). The X axis corresponds to the haemodialysis #, the Y axis to the blood flux value, and the Z axis to the time (an haemodialysis ranges about four hours). It is possible to figure out that some haemodialyses (i.e., 5, 15, 16, and 17) present strong irregularities in the blood flux (that is below the standard flux of 300 ml/minute). The doctor can inspect in a similar way other parameters, trying to find out some other possible reasons for the bad performance of the first 18 haemodialyses.

7. USABILITY TESTS

ISO 9241-11, Ergonomic requirements for office work with visual display terminals: Guidance on usability (1998), defines usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". ISO 9241 also identifies the following as the most useful indicators in measuring the level of usability of a product:

- Effectiveness in use, which encompasses accuracy and completeness through which users achieve certain results.
- Efficiency in use, which has to do with the resources utilised in relation to accuracy and completeness.
- Satisfaction in use, which includes freedom from inconveniences and positive attitude towards the use of a product.

The new ISO/IEC 9126-1 definition of "goal quality" bears resemblance to the foregoing definition. Consequently, usability has an impact on the quality of a product. The evaluation of goal quality includes the use of effectiveness, efficacy and satisfaction indicators, and security (Catarci, 2000).

To determine the usability of an interface, it is necessary to subject it to rigorous tests. It should be pointed out that our system primarily targets a specific type of users that are highly acquainted with the domain of interest. It would arguably be easier to design an interface for a specific type of users than for a mixed audience. Nonetheless, the need to carry out usability tests remains.

As a way of getting started, we carried out usability heuristics. The term "usability heuristics" refers to a more informal evaluation where the interface is assessed in terms of more generic features. This informal evaluation presents reasonably concise and generic principles that apply to virtually any kind of user interface. In the following discussion, we analyse how some of the principles have been applied in the design of the system.

- The interface dialogue should be simple and natural. Moreover, the interface design should be based on the user's language/terms. In general, there should be an effective mapping between the interface and the user's conceptual model. In our system, the interface primarily uses medical terms, familiar to the users.
- The interface should shift the user's mental workload from the cognitive processes to the perceptual processes. Our interface supplies various mechanisms to support the shift. For practically all inputs, the user does not have to supply the units of measurement. Moreover, the system offers interaction controls (e.g., sliders) for helping the user get familiar with the range of valid values and also for helping him/her input within the range.
- There should be consistent usage and placement of interface design elements. Consistency builds confidence in using the system and also facilitates exploratory learning of the system. In our interface, the same information is presented in the same location on all the screens. In fact, the overall system interface is uniform across the various environments for 2D and 3D visualisation.

- The system should provide continuous and valuable feedback. One of the mechanisms our system uses to provide feedback is realized by dynamically updating the visualisation as the user changes (or interacts with) the various parameters.
- There are many situations that could potentially lead to errors. Adopting an interface design that prevents error situations from occurring would be of great benefit. In fact, the need for error prevention mechanisms arises before (but does not eliminate) the need to provide valuable error messages. Our interface offers mechanisms to prevent invalid inputs (e.g., specification by selection, specification through sliders). It also provides some status indicators.

Moreover, we initially tested our interface in an informal way, presenting it to the doctors and to the researchers involved in the project. We got encouraging results from the tests and even suggestions on how to improve the interface. For instance, doctors suggested that the interface should be based on the 2D environment, with the 3D as a means of deepening the analysis. Consequently, the foregoing suggestion has been included, producing the interface described in the previous sections. The experts also suggested that the system should provide an optional interaction environment specifically designed for the expert user and still leave the user with the freedom to switch between the two.

We are planning to perform some more user tests on the current version of the prototype with selected users from the Nephrology Centre. The selected users will be provided with various documents to help them in the test. For instance: a case study, user tasks, and data schema. Each user will be expected to run and interact with the interface of the prototype with reference to the accompanying documents. After the experiment, the user will be expected to fill in a questionnaire.

The first part of the questionnaire will contain closed questions pertaining to the simplicity or complexity of carrying out user tasks. The second part of the questionnaire also will have closed questions, aimed at assessing interface design aspects, like consistency, intuitiveness, and interface elements organisation. The last part of the questionnaire will contain open questions pertaining to strengths, weaknesses, capability of the system/interface. It also will have room for extra/other comments.

We plan to modify the current version of the prototype by exploiting the results that will be collected from aforementioned planned user studies.

8. FUTURE WORK AND CONCLUSIONS

Associating an effective visual representation to medical databases, and, in general, to large data sets, is crucial for allowing different kinds of non-technical users to easily grasp the database information content. In this paper we describe a system dealing with biomedical data coming from haemodialysis, a domain in which “natural” data representations are hard to find. The proposed approach extends the usual 3D visualisations introducing more visual dimensions (e.g., colour, size, etc.).

We also show how the system would work on a realistic case study, in which simple data analysis and statistical data analysis are presented. We are collecting more feedback from doctors involved in the experiment and the next step will be the implementation of a new version of the system where the end user will be enabled to add her/his own rules, and equipped with an enhanced knowledge base including the rules defining the users' tasks as well.

Another extension is to allow our system to handle large data sets with high dimensionality (higher than five dimensions) and a large number of records that can display an excessive density of pixels (data points). Known techniques must be applied, such as volume rendering (Yang, 2000).

Finally, since the end-users, in general, are only interested in a subset of the entire database, we will enrich the system by using an approach that can combine visual query formulation (where the visualised query purpose is the database schema) and query result visualisation, in order to allow the end users indicate the data subset in a complex database schema (which is constituted by several relational tables).

ACKNOWLEDGEMENTS

We gratefully acknowledge the sponsoring of Italian Ministry of Education, University and Research (MIUR) through the project "ANALYSIS, INFORMATION VISUALISATION, AND VISUAL QUERY IN DATABASES FOR CLINICAL MONITORING". We are indebted to the other researchers participating in the above project for useful discussions and with the doctors and nurses of the Nefrology Centre of the Civil hospital of Vigevano for their extremely useful cooperation. Finally, we thank the anonymous reviewers for their comments and suggestions.

REFERENCES

AHLBERG, C. and SHNEIDERMAN, B. (1994): Visual information seeking: Tight coupling of dyamic query filters with starfield displays. *Proc. International ACM Conference on Human Factors in Computing Systems* (CHI'94), Boston MA, USA, 313–317.

AHLBERG, C. and WISTRAND, E. (1995): IVEE: An information visualization and exploration environment. *Proc. Int. Symp. on Information Visualization*, Atlanta GA, USA, 66–73.

ANDREWS, K., WOLTE, J. and PICHLER, M. (1997): Information pyramids: A new approach to visualizing large hierarchies. *IEEE Visualization '97, Late Breaking Hot Topics Proc*., Phoenix, Arizona, 49–52.

ANKERST, M. ESTER, M. and KRIEGEL, H. (2000): Towards an effective cooperation of the user and the computer for classification. *Proc. of the 6th ACM SIGMOD International Conf. on Knowledge Discovery and Data Mining (KDD 2000)*, Boston MA, USA.

ARNHEIM, R. (1954): *Art and Visual Perception: a Psychology of the Creative Eye*. The University of California Press.

ASIMOV, D. (1985): The grand tour: A tool for viewing multidimensional data. *SIAM Journal of Science and Statistical Computing* 6(1): 128–143.

BECKER, B. (1997): Volume rendering for relational data. *Proc. of the IEEE Symp. on Information Visualization (InfoVis'97)*, Phoenix, Arizona, USA.

BELLAZZI, R., LARIZZA, C., MAGNI, P., BELLAZZI R. and CETTA, S. (2001): Intelligent data analysis techniques for quality assessment of hemodialysis services. *Proc. of the Workshop on Intelligent Data Analysis and Pharmacology.*

BERTIN, J. (1983): *Semiology of Graphics.* The University of Wisconsin Press.

BRINCK, T. and YORK, G. (1998): User interfaces for computer-based patient records. *Proc. International ACM Conference on Human Factors in Computing Systems (CHI'98)*, 214.

CATARCI, T. (2000): What happened when database researchers met usability. *Information Systems* 25(3): 177–212.

CATARCI, T., COSTABILE, M. and SANTUCCI, G. (1999): Foundations of the DARE system for drawing adequate representations. *Int. Symposium on Database Applications in Non-Traditional Environments (DANTE 99).*

CHAN, A., JIANNONG, C., CHAN, H. and YOUNG, G. (2001): A web-enabled framework for smart card application in health services. *Communictions of the ACM* 44(9): 77–82.

CHANDRA, A. and MERLIN, P. (1977): Optimal implementation of conjunctive queries in relational data bases. *Proc. of 9th ACM Symposium on Theory of Computing.*

COOK, D., BUJA, A. and CABRERA, J. (1995): Grand tour and projection pursuit. *Journal of Computational and Graphical Statistics* 4(3): 155–172.

CORNSWEET, T. (1970): *Visual Perception*. Academic Press.

FURNAS, G. (1996): Generalized fisheye views. *Proc. International ACM Conference on Human Factors in Computing Systems (CHI'86)*, Boston MA, USA, 18–23.

GRIMSON, J., GRIMSON, W. and HASSELBRING, W. (2000): The SI challenge in health care. *Communications of the ACM* 43(6) 49–55.

HINNEBURG, A., KEIM, D. and WAWRYNUCK, M. (1999): HD-eye: Visual mining of high-dimensional data. *IEEE Computer Graphics and Applications* 19(5): 22–31.

HSU, W., LEE, M., LIU, B. and LING, T. (2000): Exploration mining in diabetic patients databases: Findings and conclusions. *Proc. of the 6th ACM SIGMOD International Conference on Knowledge Discovery and Data Mining (KDD 2000)*, Boston MA, USA, 430–436.

HUSTON, T.L. and HUSTON, J.L. (2000): Is telemedicine a practical reality: *Communications of the ACM* 43(6): 91–95.

INSELBERG, A. and AVIDAN, T. (2000): Classification and visualization for high-dimensional data. Proc. of the 6th ACM SIGMOD International Conference on Knowledge Discovery and Data Mining (KDD 2000), Boston MA, USA, 370–374.

INSELBERG, A. and DIMSDALE, B. (1990): Parallel coordinates: A tool for visualizing multi-dimensional geometry. *Visualization '90*, San Francisco CA, USA, 361–370.

JOHNSON, B. and SHNEIDERMAN, B. (1991): Tree-maps: A space-filling approach to the visualization of hierarchical, information structures. Proc. *IEEE Visualization '91*, San Diego CA, USA, IEEE Computer Society, 284–291.

KEIM, D.A. (1997): Visual techniques for exploring databases, invited tuturial. Proc. *Int. Conference on Knowledge Discovery in Databases (KDD '97)*, Newport Beach CA, USA.

KEIM, D.A. (2001): Visual exploration of large data sets. *Communications of the ACM* 44(8): 39–44.

KEIM, D. and KRIEGEL, H. (1994): VisDB: Database exploration using multidimensional visualization. *IEEE Computer Graphics and Applications*, 40–49.

KILMAN, D. and FORSLUND, D. (1997): An international collaboratory based on virtual patient records. *Communications of the ACM* 40(8): 111–117.

KLEIBERG, E. VAN DE WETERING, H. and VAN WIJK, J. (2001) Botanical visualization of huge hierarchies. *Proc. IEEE InfoVis 2001*, San Diego CA, USA, IEEE Computer Society, 87–94.

KREUSELER, M., LOPEZ, N. and SCHUMANN, H. (2000): A scalable framework for information visualization. *Proc. IEEE InfoVis 2000*, Salt Lake City, Utah, USA, IEEE Computer Society, 27–36.

KUMAR, V., FURUTA, R. and ALLEN, R. (1998): Metadata visualization for digital libraries: Interactive timeline editing and review. *Proc. of ACM Digital Libraries*, Pittsburgh, USA, 126–133.

LAMPING, J., RAO, R. and PIROLLI, P. (1995): A focus+context technique based on hyperbolic geometry for visualizing large hierarchies. *Proc. International ACM Conference on Human Factors in Computing Systems (CHI'95)*, Denver CO, USA, 401–408.

MACKINLAY, J. ROBERTSON, G. and CARD, S. (1991): The perspective wall: Detail and context smoothly integrated. *Proc. International ACM Conference on Human Factors in Computing Systems (CHI'91)*, New Orleans, Louisiana, USA, 173–179.

PLAISANT, C., MUSHLIN, R.,, SNYDER, A., LI, J., HELLER, D. and SHNEIDERMAN, B. (1998): Lifelines: using visualization to enhance navigation and analysis of patient records. *American Medical Informatic Association Annual Fall Symposium*, Orlando, USA, 76–80.

POWSNER, S. and TUFTE, E. (1994): Graphical summary of patient status. The Lancet 344(8919): 386–389.

RAO, R. and CARD, S. (1994): The table lens: Merging graphical and symbolic representation in an interactive focus+context visualization for tabular informtion. *Proc. International ACM Conference on Human Factors in Computing Systems (CHI'94)*, Boston MA, 318–322.

ROBERTSON, G., MACKINLAY, J. and CARD, S. (1991): Cone trees: Animated 3D visualization of hierarchical information. *Proc. International ACM Conference on Human Factors in Computing Systems (CHI'91*), New Orleans, Louisiana, USA, 189–194.

SHELFER, K.M. and PROCACCINO, J. (2002): Smart card evolution. *Communications of the ACM* 45(7): 83–88.

SHNEIDERMAN, B. (1983): Direct manipulation, a step beyond programming languages. *IEEE Computer* 16(8): 57–69.

SHNEIDERMAN, B. (1996): The eyes have it: A task by data type taxonomy for information visualization. *Proc. of IEEE Symposium on Visual Languages*, Boulder CO, USA, IEEE Computer Society, 336–343.

SHNEIDERMAN, B. (2001): Inventing discovery tools: Combining informtion visualization with data mining. *Proc. of Discovery Science 2001* Conference, Washington DC, USA.

SILVA, S. and CATARCI, T. (2000): Visualiztion of linear time-oriented data: a Survey. *Proc. of the First International Conference on Web Information Systems Engineering (WISE'00)*, Hong Kong, IEEE Computer Society.

WILLIAMSON, C. and SHNEIDERMAN, B. (1992): The dynamic homefinder: Evaluating dynamic queries in a real-estate information exploration system. *Proc. of ACM SIGIR'92 Conference*, 338–346.

WITTEN, I. and FRANK, E. (2000): *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann Publishers, San Francisco.

WRIGHT, W. (1995): Information animation applications in the capital markets. *Proc. Int. Symp. on Information Visualization*, Atlanta GA, USA, 19–25.

YANG, L. (2000): Interactive exploration of very large relational datasets through 3D dynamic projections. *Proc. of the 6th ACM SIGMOD International Conference on Knowledge Discovery and Data Mining (KDD 2000)*, Boston MA, USA, 236–243.

BIOGRAPHICAL NOTES

Tiziana Catarci received her PhD in Computer Science from the University of Roma "La Sapienza", where she is currently a Full Professor. She has published over 100 papers and 10 books on a variety of subjects comprising user interfaces for databases, 2D and 3D data visualisation, adaptive interfaces, visual metaphors, usability testing, query languages, cooperative database systems, database integration, web access.

Tiziana Catarci

Giuseppe Santucci graduated in Electrical Engineering from the University of Roma "La Sapienza", Rome, Italy, where he is presently an Associate Professor in Computer Engineering. His main research activity concerns both theoretical and practical aspects of visual query language and user interfaces. Other fields of research are data dictionaries and graphical environments. He authored several scientific papers on these topics published in international journals and conference proceedings.

Giuseppe Santucci

Sonia Fernandes Silva received her PhD in computer science from the Electric Engineering Department of the Federal University of Paraiba, Brazil. Currently, she works in cooperation with Tiziana Catarci and Giuseppe Santucci at the University of Roma "La Sapienza", Rome, Italy. Her main research interests include: visual formalisms for representation and querying of relational and historical databases, visual query languages, visual metaphors and information visualization.

Sonia Fernandes Silva