Technological principles of Digital Museum

Alina Gutnova

Master's thesis August 23, 2007
Department of Computer Science
University of Joensuu

Technological principles of Digital Museum

Alina Gutnova

Department of Computer Science
P.O. Box 111, FI-80101 Joensuu, FINLAND.

Master's thesis

Abstract

Digital museums appeared in the Internet many years ago. First virtual museums have emerged since 1991. However, despite the identity and quality of exhibitions their exhibits are comparable with traditional art museum publications albums, catalogues and books, while the Internet can create a qualitatively new displays, where previously impossible or inconsistent methods of display historical and artistic heritage have been set up. What is the identity of the museum, in which you enter? The main feature is that here are collections, which actually stored in different museums in different countries or even on different continents or does not exist at all.

This thesis presents the architecture developed for the Digital Museum Project under University of Joensuu, involving North Karelian Museum. The tools of creating the Digital Museum project are a set of special utilities and automated information system implemented as a Web application.

It is necessary to solve some fundamentally new challenges under creation of digital museum: 1) Define concept of selection of information documents (museum exhibits and materials), as well as a virtual museum collection; 2) Define a set of interfaces to access the museum; 3) Develop a distributed architecture information system that will be capable to act in a global network fixedly and steadily; 3) Develop software components to support the operation of a distributed system.

The main task of our project is the development of technologies for constructing the distributed information systems for the storage and display of heterogeneous information. Review of existing virtual museums was performed in the thesis, variety of technological tools was considered and optimal for our project ones were chosen and described carefully.

Key words:
Digital museum, virtual museum, content management system, multimedia databases, client-server architecture, connections.

Contents

Chapter 1

Introduction

1.1 Motivation

Small museums and galleries all around the world are not able to organize large-scale, permanent exhibitions where inquisitive people can see artifacts of cultural heritage in its elaboration. In a typical museum we can not see the whole collection of exhibits because of restriction of the space and some other reasons. For example, some museum objects can not be displayed because of their state to avoid the damage. Some other objects need a lot of space to show and for explanation or even do not have a shape at all. The exposed objects are not allowed to touch or rotate. One of the advantages of a virtual museum is the opportunity to see the objects from different sides and foreshortenings and to request a personal guide that will give the information about the exhibit off screen. Moreover, we can visit the real museum just while the exhibition is opened. Not all the people are able to afford a trip to a country on another side of the Earth to walk in the Louver's chambers, or to visit the Dresden gallery, or to see the collection of the Hermitage. That is why the constructing of the virtual museum is the only one way to make the worldwide treasures accessible to the public. Our Digital Museum project aims to study these problems.

Even if the virtual museum seems like the common museum it is undoubtedly a new reality that is beyond the scope of traditional thinking about a real museum with its permanent exposition and temporary exhibitions. The exposition of the virtual museum is constant just in its evolution, the life time of an exhibition can be amounted up to the several years, and the count of layouts depends not on the space but on the conceptual ideas about interesting projects.

What is the digital museum? There is no fixed definition for it yet and it varies

according to the content of the museum. For example, the digital museum of art gallery can be defined as follows. "The virtual museum is the electronic reproduction of an art gallery that contains a collection of electronic artifacts and information resources for public exploration and is subject to the spatial perception of the art gallery to the visitors" [SK99]. Based on it we can define the digital museum in general as follows. The digital museum is a computer aided museum consisted of web-pages not necessarily stored on one web-server which consist of catalogs and albums of exhibits, their description, multimedia etc. Virtual museum can be both flat that seems like simple catalog and three-dimensional that allows us to walk through chambers and see exhibits from different angle of view.

People visit the virtual museum from their home computers. Of course, it is not a Sunday stroll to the Prada museum or to the State Treyakov Gallery. The emergence of a theatre and later of a cinematograph did not recede a book into the background, because in spite of materiality and publicity of the cinema and theatre, just with a book a man soaks himself in a new reality that he reconstructs in his mind. The same situation is with virtual museum, with which man intercommunicates face to face. He is the owner of the world treasures.

1.2 Problem formulation

The goals of this thesis can be formulated as follows:

1. Creating the system that would be able to combine the scattered description of the world heritage into unified informational resource.

2. Designing the optimal way of storing the multimedia information.

3. Development the system for maintenance of information about museum objects, unique heritage of past.

4. Making review of case-studies necessary for the project.

5. Using different devices to deliver the information to the users.

6. Image preprocessing and comparison between spectral images and RGB-images.

1.3 Structure of the thesis

In Chapter 2 the operating principles of modern technologies are described. Here client-server architecture and different types of connections are considered. Nowadays museums should use all ways to deliver the information to the humanity. Review of existing types of connections and analysis of their advantages and disadvantages are represented.

For last 10 years mobile phones become essential part of our life. In UN's (United Nations) report was published the information about the count of mobile phone owners in the 2006 year. There were 2,17 billion people. According to the Worldwide GSM Association this showing will be increased till 3 billion people to the end of this year and will spread over 90% of the world population to the 2010 year ([20007]). Thus, we can see that mobile phone communications are very wide channel for deliverance the information to the users. In Chapter 2 content-based design of mobile applications for museums are represented [RTA05]. In the thesis we are also considering the design that specifies an XML-based style sheet to make a new exhibition from extracted multimedia objects. This would save a lot of time and attract more visitors.

In Chapter 3 basic concepts of digital museum are represented. Nowadays there are many existing virtual museums [HCH00] different by the structure, range of services and support. This thesis analyzes existing digital museums and describes an integrated solution of designing a walkthrough environment system for virtual museum with a client-server architecture supporting multi user's communion [US96], developing the optimal database structure, fast content management system with friendly user interface etc.

New multimedia applications are located on the top of informational infrastructures and require new databases. That is, next-generation database systems should provide users the efficient and easy development of the multimedia applications. During writing the thesis we were focused on development of a database system based on analysis of existing ones which facilitates adaptable and efficient acquisition, storage, retrieval, access and performance of the information.

In the chapter we are focused on the multimedia, its storing, combining and retrieving. Museum multimedia is the object that can not be seen by eyes and touched by hands. There are sound, digital and network objects. They exist just in virtual space. Museum multimedia consists not only some video work that become usual in the museums, but also the digital sculpture [LPC+00] and sound-visual composition. For example, one Dutch artist took a photo of the common pedes-

trian way to the camera, and then he marked it out like a music-paper in such way that any dirty spots and leaves on the asphalt turned into notes. As a result real "street music" appears.

New technologies allow us to get the true colors and real shapes of the museum objects and in present time the virtual visitors can see exhibits as they are in real life. The visitor can acquire more information about objects with high-resolution images from many directions and using different sources of lights and so force. The image processing and advantages of spectral approach are described in Chapter 4.

Chapter 2

The operating principles of modern technologies

Last time in the world often argue about the effectiveness of the application of computer technology in a museum, consider a phenomenon of a virtual museum [UYF+98] and a virtual exhibit, their pros and cons.

Firstly note that the virtual museums let to visit museum without leaving home or even the country. Secondly, they are a great way of advertising the real museum. And finally in the third, it is popular. It is irrationally not to take advantage of the computer technology that can facilitate and speed the work of the museum in dozens of times and for small money.

Museums visiting, dipping into its environment, the direct immediate reaction is effects that are not associated with virtual reality. Virtual museum will never claim for a taste and smell, but it is not limited by space! The vase, placed in the Hermitage, offers an opportunity to enjoy the work of art, the same vase in virtual reality make possible to find the history, examine it closely, and even its sectional view, find similar vessels stored in other museums. Virtual museum first of all is a source of full information.

Real museum becomes an information center, providing access for user to structured collection of information. Virtual museum is a global information field, where you can find the answer to any question. Real museum is still limited by funds and territorial situation. It is easy to locate in Internet images of the exhibits, but put them in real exposure can be difficult because of limited space, or in view of complexity of creating the necessary conditions for storage. In Web museum you can create a guide tour, accommodate various information associated with the whole epoch, or with a certain exhibit. In Web museum you can even ask a question and

receive an answer in the forum of the web site.

Modern technologies can combine digital images of people and objects with the virtual reality of sightseeing along the tourist's way. In this chapter we describe the basic principles of modern mobile technologies, which we used in development of our system of Digital Museum.

2.1 Client-server architecture

Usually included into informational system computers and applications are not equal. Some of them own resources like file system, processor, printer, database etc, and others can access them. A computer system that manages and delivers information for client computers is called *Server* [Glo07]. For example, file server, database server, computational server etc. Server is a set of programs that control different processes execution. The main function of server is executing some process by client request and sending the results to the client. A computer that has access to services over a computer network is called *Client* [Glo07]. The main function of client is application running and setting connection with server when it is necessary. In other words, client should provide user by application interface, implement its logic and send requests to server when it is needed.

The interaction between server and client is started on the initiative of client. Client requests the query from the server, establishes session, gets needed results and reports about completion of a task. Usually one server is used by several clients at the same time. Therefore each server should have enough high productivity and provide data security. Client and server can be both under same computation system and on different computers connected by network with each other.

The main principle of client-server technology is division the application functions into three groups:

- Data input and output, interaction with user;
- Applied functions specific for the subject field;
- Resource management functions (file system, database etc).

Thus every application consists of following components:

- Data representation component;

- Applied component;
- Resource management component.

Component interaction is realized by certain rules that are called interfacing protocol.

2.1.1 Client-server interfacing models

The Gartner Group Company that is specialized on IT research offered the following classification of two-tier model of client-server interfacing that is represented on the Figure 2.1 [AB95]. It is called two-tier, because three application components share diversely between two nodes.

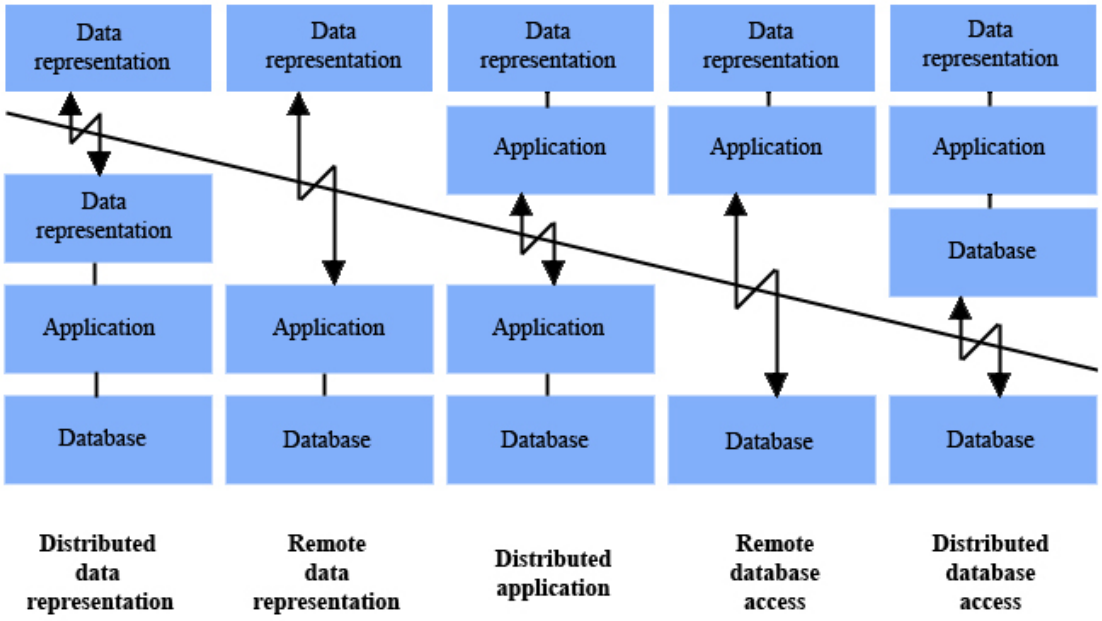

Figure 2.1: Classification of two-tier model of client-server interfacing.

Historically the first client-server interfacing model was distributed data presentation model. It was implemented on general IBM with non-intelligence terminals. Data management and user interaction were combined into one application. Later remote database access models were implemented as PC and network emerged. The main PC architecture for that time was file server architecture. In this case one of the computers is file server and others are clients that contain data representation

component and applied component (DBMS and applied program). Data communications protocol here is a set of low-level calls of file system operations. Such architecture has apparent disadvantages, for example, high rate traffic and unified resource access.

As first special-purpose database server emerged, another implementation of remote database access became possible. In this case DBMS runs on the server and communications protocol is supported by SQL language. In comparison with file server system the approach leads to network loading decrease and unification of client-server interface. But rate traffic was still high and application management was unsatisfied, because applications contained different services.

Later active server concept that used stored procedures mechanism was developed. Based on the concept, part of applied component was moved to the server. Such a model is called distributed applications model. Procedures are stored in the data dictionary, divided between several clients and executed on the same computer as SQL server. The advantages of such approach are single point administration of applied functions and network traffic decrease as we are working here not with SQL queries, but with stored procedures calls. The disadvantage is finiteness of development tools for stored procedures.

At present in practice there is used the following approach:

- Primitive applied functions are hold by stored procedures on the server;
- More complicated functions are implemented in the application directly on the client.

Lately the distributed applications model becomes very popular. Its specific feature is division an application logic into two or more parts, each of which can be run on individual computer. Different parts of application logic interact with each other through messages in special format. In this case, two-tier client-server architecture becomes three-tier (Figure 2.2) or even more tier architecture.

2.1.2 Client-server advantages over File-server technology

With file-server system all the responsibility for the safety and integrity of the database devolves on the program and the network operating system. All the data is processed on the workplaces, and the server is only used as a shared drive [MK02]. Each user directly uses information and edits data and index files. Since the more

Figure 2.2: Three-tier client-server architecture [Jaw01].

users the higher requirements for separation of data, then processing speed is significantly lower in multi-user mode. In addition, databases can become damaged. For example, at the time of writing to the file there may be network or power failure. In this case, the user's computer interrupts work, the database can be damaged, and the index file can be destroyed. Reindexing should be carried out after such halting and could last several hours.

Client-server system avoids these problems, because all the work with database occurs at the server and does not depend on the failures of workstations [Rom97]. All requests for file writing are intercepted by server. The file is modified just when the server receives a message that the editing of the file is complete. This excludes damage of index files and significantly increases system performance.

Besides the computing system speed and system reliability the client-server architecture gives many advantages also in a part of hardware. Firstly, the server optimizes the functions of data processing, avoiding the need to optimize workstations. A workstation can be equipped with not very fast processor, and yet the server will make it easy to get the results of the query. Secondly, as the workstations do not handle all intermediate data, the network loan significantly decreases. There is an opportunity to maintain an activity log, which would automatically register all the transactions. It would help to recover the system rapidly after hardware crash.

Recently, one of the most important questions of development the network applications is the following. Which one of client-server and Web-based architecture to prefer?

Client-server architecture is more popular. The main thing amongst these approaches is how each of them performs the further development and responds to the market trends of computing technologies.

For example, the client-server systems cannot use the advantages of network

computers, intranet and Java language to the full. Web applications based on Java and ActiveX are a new centralized architecture, in contrast to the distributed approach to computing, implemented by the emergence of desktop computers. Let us consider how both approaches work today.

2.1.3 Client-server and Web

The criteria for network informational system assessment are performance, reliability, the intensity of network traffic, the occupance rate of administrator's work and security control. Moreover, all of them are closely interrelated.

It is difficult to give a precise definition of the two architectural approaches. In the classic client-server architecture the client part is performed by the application developed by using Visual Basic or Delphi. In addition to information representation level it contains also business logic of an application task. The client works both with relational database and file-server system. Standard Web application is usually used only for data representation. In this case the business logic is executed on the server that is working with DBMS.

But there is the client-server architecture, where some business logic is extended to selected computers. Such systems are already close to the classic Web architecture. There can be constructed Web applications, where some business logic is based on JavaScript or Java code.

2.1.4 The administrator's occupance

Network administrators usually have a range of duties, such as client and server applications support, help desk, security administration, software distribution etc.

The administrator of client-server system should control the usage of file server, configure drives and settings of server databases. He is responsible for security control that involves the setting of user names and passwords, regular monitoring of log files and the usage of security devices that are accepted in the organization.

The administrator of Web-based system has almost the same responsibilities. However, customer support in the situation of applications deadlock takes much less time, because Web-browsers, such as Microsoft Internet Explorer or Netscape Navigator, "freeze" much rarer than programs implemented by using development tools like Visual Basic or Delphi.

2.1.5 Comparison of reliability

Web-approach generally proved to be more reliable than client-server one. Program support makes the client less stable. Debugging and testing client-server applications are quite complex process, because they contain a large number of closely interacting modules. On the other hand, it is easier to find errors, because Windows output the corresponding message on the screen.

2.1.6 Disadvantages of client-server technology

Classic client-server system has some disadvantages. Firstly, there is high demand for communication channels capacity to the server, which prevents the use of client stations except in the local network. Secondly, there is weak data protection against burglary, especially from unscrupulous users of the system. Thirdly, it is a very high degree of complexity of administration. Fourthly, there is a need to use quite high-powered computers on clients. Fifthly, it is the complexity of integration with legacy systems. Sixthly, it is the high complexity of the system development because of the need to execute business logic and to support interface with the user in one program.

It may be noted that most of the disadvantages of classical or two-tier client-server architecture result from the use of client stations as a performer of IT business logic. The obvious step in further evolution of IT architectures was an idea of division algorithms of data processing into parts related with business functions execution and display of information in a user-rights presentation. The real functionality of the system is on the server. The client machine contains only the second part related to the initial validation and display information.

Recently there are often used standard web browsers as client programs. It becomes obligatory to use databases with all their advantages. Server part is written mostly in specialized languages, using a stored procedure of server database. Thus, the IT part of client-server architecture split into three layers: data layer; business functions, or stored procedures, layer and the presentation layer. Very often some business functions are implemented in the client systems. Partly so, in part because physically such an IT consists of two components, the architecture is often referred to as 2.5-tier client.

Unlike two-tier architecture the 2.5-tier system usually does not require a high-speed communication between the client and server parts of the system, because almost all the computations are done on the server side. But along with advantages of a unitary approach 2.5-architecture gets also all its disadvantages, such as limited scalability, dependence of the software platform, limited use of network

computing resources. All this reduces the system performance, increases working hours and affect the most negative on the cost of hardware needed for the operation.

To solve these problems there was proposed the so-called three-layer client-server architecture. Its main difference from 2.5-architecture is the physical separation of the programs that are responsible for the data storage (DBMS) from the application server (AS). This separation of the program components optimizes load both on the network and computing machinery.

Let us consider first the client side. Using the architecture allows to use low-power computers or terminals in the workplace. Slow work of one client device cannot lead to a slowdown in the whole system as in the case with two-layer architecture. There is transferred just minimum necessary data between the client and server applications, i.e. calling function arguments and return values from them. This is a theoretical limit to the effectiveness of communication lines. The application server can be launched in one or more copies on one or more computers, which allows the use of computing power so efficiently and safely as IT administrator wishes. Traffic between application server and databases can be great, but it is always the LAN traffic, and their capacity is large enough and cheaper. In an extreme case, you can always run AS and DBMS on the same computer that automatically bring network traffic to zero. And finally, the database handles just requests for data retrieval and modification and does not carry out complex business functions. That can reduce the load on the data server and increase the speed of the system as a whole.

Systems scalability implemented in three-layer architecture is impressive. The same system can work on one free-standing computer with AS and DBMS as well as in network of hundreds or thousands of computers. The only obstacle endless scalability is the only requirement to keep a single database.

Apart from the requirement of increasing the system performance with the growth of scale of operation, the important factor is the expansion of its functionality. Three-layer scheme satisfy the requirement. It does not necessary to change the entire system to increase the functionality. It is enough to set a new application server with the required function.

What are the disadvantages of three-layer model and why this technology has not received such widespread as its predecessors? In our view, the main disadvantage of this technology consists of system design difficulty. Expert designers know how difficult is combining interlinked modules of programs, especially written by different groups. Changes in one of the modules can often bring on avalanche-type

changes in other parts of the system. From this point of view, even primitive system performed on three-layer scheme is two times more complex than systems with two-level and 2.5-level architectures.

As we can see, the client-server systems have significant advantages over file-server systems. In spite of this, the file-server systems are still being in use. This is prevented by quite common myths about the architecture. The usual view is that client-server is slow and expensive.

The view that the client-server architecture is slow has been widespread not only among users. Some developers, commenting on the work of their systems, explain the low functional rate by basic disadvantages of the architecture. In our view this issue has two aspects.

Firstly, the database provides transaction mechanism changes in the database and self-consistency of properly built relational database. Naturally, the performance of these functions requires cpu time costs and, therefore, client-server system can work more slowly simply because they are carrying out work that should be done to ensure reliable operation of information systems.

Secondly, it is very often happened that the cause of slow operation of the client-server system belongs to the system application architecture. Programs that are written on the built-in database procedural programming language often work very inefficiently.

2.1.7 Conclusions

After analyzing the strengths and weaknesses of client-server, file-server architectures and Web approach, we have come to the following conclusion. From the point of view of the advantages of working in a network environment Web-application reduces the complexity of software, increases productivity and makes it possible to use advanced technologies, such as networked computers. However, client-server application provides all the power of desktop computer for business logic performance and gives the user a clearer understanding of results of data processing.

2.2 Types of connections

Mobile telephony and communication data system are the most dynamic telecommunications industry. The industry of mobile data transmission grows the most intensively. It includes wireless Internet access. Naturally, we could not get around

these technologies in our development of digital museum. People are using mobile technology everywhere. That is why we have created the Digital Museum application for mobile phones and digital TV, which use wireless protocols such as Bluetooth, WAP, GPRS. We consider in details the principle of their work in this chapter.

From luxury item the mobile communications are slowly becoming commonplace. Users of cellular phones no longer are concerned about the very existence of a mobile phone with which to contact the appropriate subscriber anywhere, at any time, but also new services by using the Internet.

Currently, people are looking to the availability of information services as and when the need arises, regardless of the time of day and location. Cellular phones, laptops (PDA) and wireless communication protocols such as Bluetooth, WAP, GPRS make this possible. Along with marking objects with the help of Radio Frequency Identification, libraries, museums and archives can provide users a special mobile device, which will guide visitors to an exhibit or a book shelf. If the mobile device is connected to the Internet or LAN of an organization, then personalization of user maintenance is provided by access to different levels of information. If we are talking about places of cultural heritage located outdoors, the number of information services, provided by wireless technology, permit to dispense without human guides.

Let us consider several wireless communication protocols that we used in our Digital Museum project.

2.2.1 WAP

Wireless Application Protocol (WAP) is a protocol for wireless access to information and Internet service resources directly from mobile phones [Ris02]. The main advantage of WAP is that the Internet subscriber does not need any additional equipment, such as a computer and modem. He needs just a mobile device that supports WAP.

WAP-sites are located on the web-servers and are written in a special format WML (Wireless Markup Language). This markup language is adapted to the phone opportunities, such as monochromatic graphic, small screen and small memory.

In most cases, WAP-pages are brief summaries of important information because of restrictions on the file size and the small area of the mobile device screen.

Information is transmitted between WAP-client and WAP-server (Figure 2.3). A normal mobile WAP-phone can act as WAP-client. By using mini-browser queries

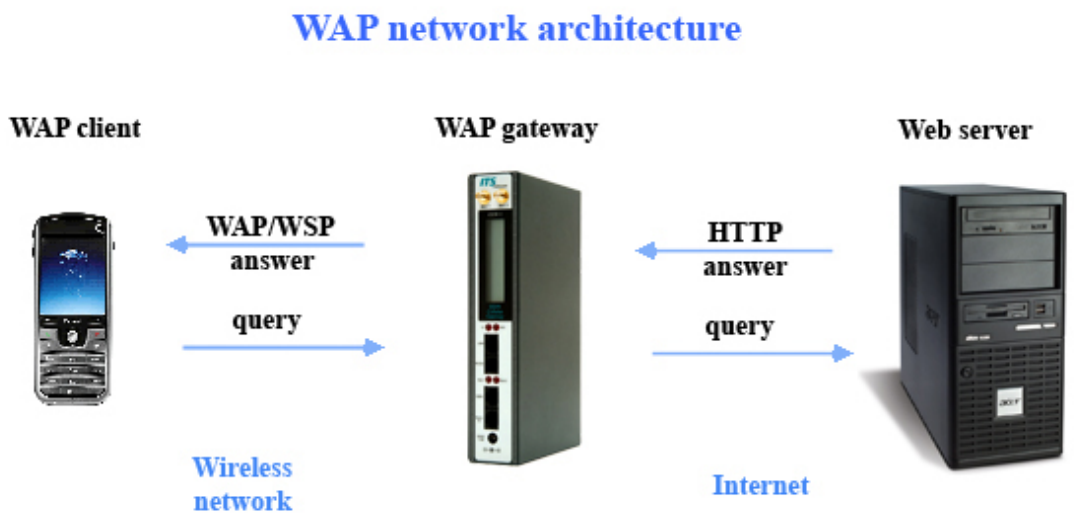

Figure 2.3: WAP network architecture.

are sending via network of wireless access, which is accepted by WAP-gateway. In turn, WAP-gateway sends URL-query to the corresponding websites, using the HTTP protocol. The requested web-pages are usually written in WML language. This site creates a response in the WML format, transmits it to WAP-gateway. Then the requested information is transmitted to customer's mobile phone.

We should give thought to two following questions before organizing mobile services in the museum. What is the possible scope of potential users for this service, and how these services will be available on the financial capabilities of users?

Certain characteristics conditioned by principles of site accessibility for mobile devices should be taken into consideration while creating WML-pages of WAP-site. The principles are:

- Keep in mind the size of a screen, look for a simple structure of content presentation on a small screen, do not clutter the pages;
- Minimize text size if it is possible, be concise;
- Avoid graphics if it is possible, because it reduces readability, increases page load time;
- Minimize hierarchy levels in the menu;
- Use a flat structure of the site;

- If it is possible minimize input/output operations when the user works with the website.

It is worthy of note that we can browse WAP-sites through an ordinary web. All new browsers have inbuilt WAP-browser (Figure 2.4).

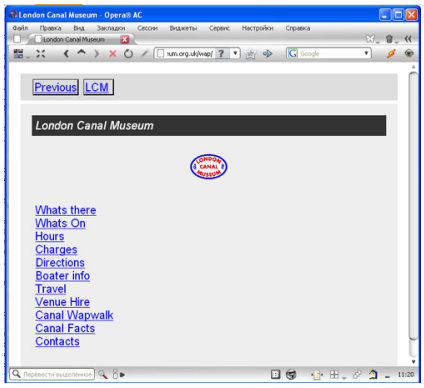

Figure 2.4: WAP-site of London Canal Museum displayed through browser Opera.

Future development of WAP

Along with adoption of WAP there were opened its advantages and disadvantages. As a result of analysis we can say that WAP in its current form requires serious improvement.

WAP interface leaves much to be desired because of the small size of displays. Short messages can be read fairly quickly, but messages containing hundreds of characters should be scrolled many times. The speed of data transfer from mobile terminals is also relevant, because we have to pay for work with WAP. The problem is eliminated by GPRS, a new data transmission technology, which increases the speed of the transfer several times and also lets to pay for the traffic, not for the time.

Professionals serving WAP face the following challenges. Sites that users can visit from mobile phones must be constantly adapted. In other words, the WAP-site, which can be viewed at 8-line display, was no longer useful for 4-line display and should be rewritten in a special version.

In its current form WAP-technology is a transitional solution. The developers have been updating it, excluding the initial disadvantages. The general point of supplying information from the Internet to mobile phone will only grow. The person

may need access to the e-mail, his financial data, and information on institutions in the area where the user is located during travel.

In our project Digital Museum we had used the WAP protocol to support as much different mobile phones as possible. Some old models support just WAP technology. GPRS and Bluetooth protocols were the emphasis of out work with mobile technologies.

2.2.2 GPRS

GPRS (General Packet Radio Service) is a technology of cellular digital packet data that does not require permanent occupation channel. It also supports significantly faster data traffic speed, up to 115 Kbps, at the expense of simultaneous use of several time slots. GPRS-based phone can always be connected with provider and always be ready to receive or send information.

The principle of GPRS

GPRS is an ETSI standard (European Telecommunications Standards Institute) for packet switching in GSM systems. Currently, the world most widely distributed networks are based on GSM networks. They are called second-generation networks (2G). GSM technology uses a variation of TDMA (time division multiple access). GSM is the most widely used of the three major digital wireless technologies TDMA, GSM and CDMA.

GSM digitizes and compresses data, then sends them through the channel. It is running at a frequency of either 900 MHz or 1800 MHz [ZJK98]. Since the majority of GSM network operators have entered into revenue agreements with foreign operators, users continue to use their mobile phones while traveling in other countries.

GPRS is the technology available for GSM, CDMA and TDMA networks. This technology uses a new method of effective transfer of packet data through radio networks. The technology is based on IP and X.25 packet switching techniques, both of which are very popular and widely used in many networks. GPRS packet switching works as well as IP packet switching, i.e. the data is split into the packages and dispatched in different ways through the network, then collected on the end point. GPRS packet switching allows any existing IP or X.25 traffic for sending data over the GPRS network.

Theoretically, the speed of GPRS is 115 kbps, but in the real world the most likely estimate is 48 Kbps [GPZ99]. This result is much better than existing mo-

bile communications devices can offer. They give just 9.6 kbps [GPZ99]. Another important aspect of internet via GPRS is that the internet connection continuous, always on-line, and at the same time it does not have to use resources from access points at a time when it was not used, because the data is transmitted only when there is a need. Receiver requests the information and the device begin using radio-resources at this time. Then it stays out of use until receiving the requested information.

GPRS network

GPRS network is located on top of GSM infrastructure (Figure 2.5). There are following key components of the GPRS network [Bat01]:

- *PCU*, the package management block, providing an opportunity for GSM stations to send and receive packets through GPRS communications.
- *SGSN*, the part of GSM infrastructure, which is responsible for sending and receiving packets from subscribers in its service area. The unit also produces authenticated contact with the server and checks information about user. In addition, it monitors the route of the caller movements to be able to allocate resources properly.
- *GGSN*, the component of GSM network, responsible for interaction with the Internet and other public networks that transmit data and voice. The component stores database of addresses and filter database.
- *GTP*, tunneling protocol GPRS, based on TCP/IP protocols.

GPRS advantages

GPRS provides high-speed data connection, permanent, stable connection, widely supported applications and powerful security mechanisms.

GPRS restrictions

It is already clear that GPRS is an important new service for mobile devices. It not just transfers the data, but also significantly improves the efficiency, capacity and functionality. But GPRS has also some limitations, which are:

- Limited capacity;
- Real speed is much lower;

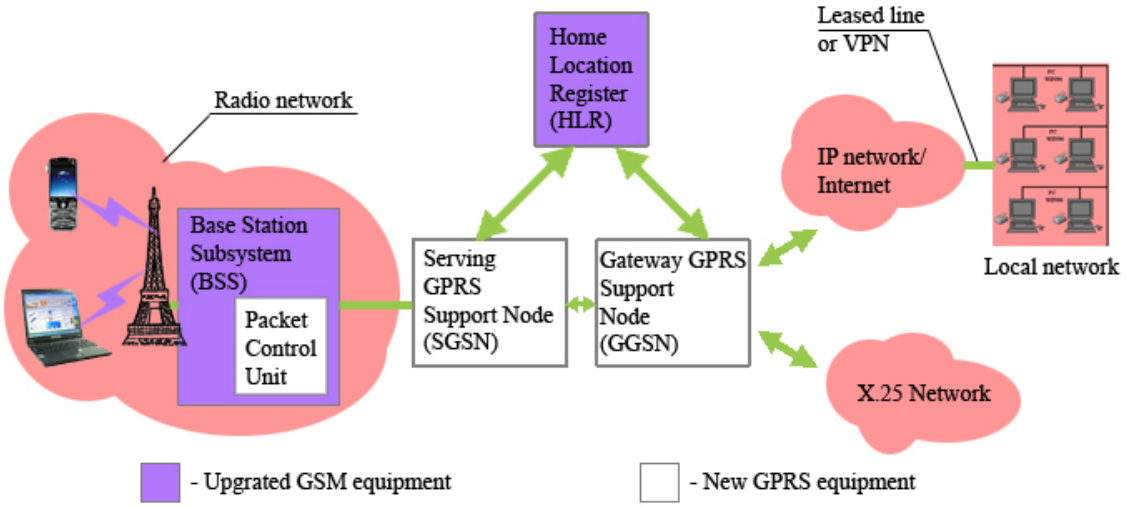

Figure 2.5: GPRS network architecture.

- Non-optimal modulation;
- Transmission delay.

More information about GPRS and some optimization of links are described in article [CCP03].

2.2.3 Wireless communication standards

IRDA

IrDA is an international organization that develops and promotes wireless communications standards. IrDA provides wireless connectivity between devices that might use cable for such purposes otherwise.

Devices should be in the direct vision of each other to connect to IrDA and the distance between devices should be not more than one meter. The infrared beam is emitted as a 30-degree cone. Transfer rate is ranging from 9600 bps to 16 Mbps [Rho01]. The standard connection is popular because of its cheap price and usability. You do not need any cables. The implementation of such a port in the device is cost just one dollar for producer.

Advantages and disadvantages of this method of connection are obvious. Probably the main disadvantage is that the devices must be in the direct line of sight to

connect. Moreover, some types of illumination, such as sunlight, can hinder communication, because there is used luminous flux. The next limitation is the range of communication. One meter distance is not always enough.

Price plays a major role in the calculations of producers and the popularization of this format. Direct visibility can be positive too. First, this condition is a simple physical data protection. It is quite problematic to intercept the information because of all the restrictions of IrDA. The only way is set the receiver anywhere within the 30-degree cone. Secondly, the communication is required between the two devices that brilliantly provides by infrared communication.

Infrared port is an ideal device for mobile phones and pocket PC by virtue of price, low energy consumption, compact, and the universality. But it already looks on the big computers like an unforgivable survival of times past. It is too slow and is not very comfortable, given the condition of direct line of sight.

Bluetooth

Bluetooth is a technology of data transferring through the radio for a short distance, which can connect cordless phones, computers etc in the absence of direct line of sight.

The development of Bluetooth was started by Ericsson company in 1994. The original aim was to get a new radio interface with low energy consumption and low cost, which would establish a link between cell phones and wireless headsets. Moreover, according to the Ericsson concept, the new interface was designed for the transmission of data and voice from anywhere in the world. To provide more support for young technology in such market niche as desktop systems, pocket computers and mobile phones Ericsson company organized a consortium to develop and promote a new technology called Bluetooth SIG (*Special Interest Group*) in February 1998 [PR06]. Today it consists of more than 2000 companies, including such major companies as 3COM, Nokia, Intel, National Semiconductor etc.

Bluetooth technology is currently the best achievement of modern microelectronics both in hardware and software. Bluetooth systems belong to the class of interacting open systems. Bluetooth devices physically are microchips that provide connection within the range of 2.4 GHz. The range in most European countries and the United States refers to band for industrial, scientific and medical gages ISM. The devices provide connectivity over distances up to 10 meters. By increasing capacity to 100 MW Bluetooth devices become more long range, up to 100 m.

Bluetooth technology is completely open. Any company that has signed license agreements may be part of the Bluetooth SIG and begin to build products on the basis of it. To avoid the incompatibility of devices there were developed detailed specifications, which include an explicit description of how to use the new standard features and data transmission protocols characteristics. The simplified block skeleton diagram of Bluetooth connection is shown in Figure 2.6.

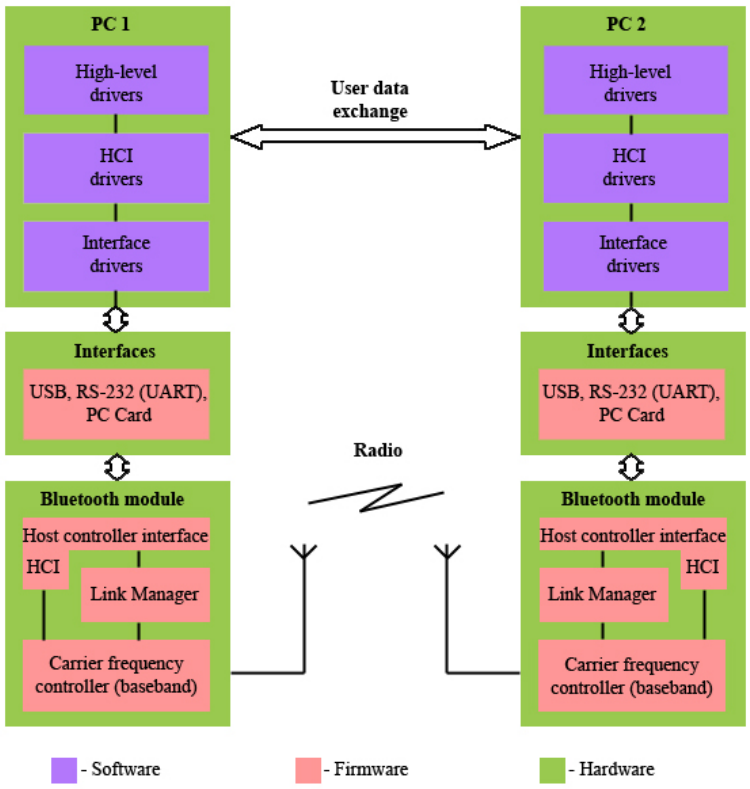

Figure 2.6: Bluetooth connection between software, firmware and hardware [BTs].

Bluetooth technology involves two types of communication: Synchronous Connection Oriented SCO and Asynchronous Connectionless ACL. SCO serves mainly for voice transmission. The speed of SCO information transfer is 64 kbps. ACL is designed for packet data transmission. It supports symmetric and asymmetric connections such as one-to-many. The transfer speed of packet data transmission of ACL is about 721 kbps. Data packets have a fixed format. In the beginning of the block is 72-bit access code. It may be used in particular to synchronize devices. Then there is 54-bit package title containing a checksum of the package and information about its characteristics, for example, about retransmission of data unit. In

the end of package there is area containing the transmission data. The size of this area varies from 0 to 2745 bits [PR06].

When any two Bluetooth devices are connected, they form pico-net. The device that initiates the communication is called the *master*, or the host, and the rest devices are called *slaves*. Usually the master is the module that is storing in the most powerful device, such as a personal computer. The number of modules in pico-net is unlimited, but at any given time there should be no more than eight active devices. There is no difference in the hardware and software part of the master and slave devices. Any of them can be both. Master device forms pico-net and controls traffic in full. There is only one master device in each network. Slave devices can send messages only in the range of slave-to-master and only after they had got earlier message from master-to-slaves. If there are more than 8 devices in the network, then there will be formed a second pico-net etc.

Many pico-networks, that are able to interact with each other, form a distributed network Scatternet. The Bluetooth pico-net and scatternet are shown in the Figure 2.7.

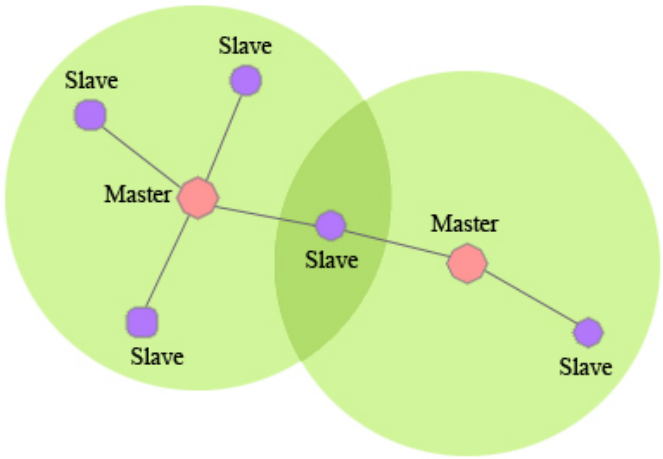

Figure 2.7: Bluetooth pico-net and scatternet [BTs].

Bluetooth is already regarded by many developers as partner technology of universal wireless technology for local area networks. Bluetooth is already active in the global market of new technology. Along with Jini it is proposed at the 1999 for the prize for the best technological innovation, annual premium of Discover journal.

2.3 Bluetooth API in J2ME

In this section we will consider the capabilities of JSR-82 Java API for Bluetooth. JSR-82 is an advanced J2ME package that implements the Java Community Process, which provides a standard API for Bluetooth connections.

To start the development of Bluetooth applications, we will need the following tools:

- Java 2 Standard Edition (J2SE) SDK 1.4.1 or later;
- Java 2 Runtime Environment (JRE) Standard Edition 1.4;
- Sun Wireless Toolkit version 2.0 or later.

Bluetooth is a low-cost technology that lets electronic devices to exchange information via radio. Bluetooth transmitter range is normally not exceeding 15 meters. There is used the frequency of 2.45 GHz for connection. Currently, there is available the Bluetooth technology, which includes radio communication technology, a set of programs and profiles.

2.3.1 Protocol stack

Bluetooth protocol stack consists of two parts: a controller, which is usually implemented as a hardware, and on-host stack that applications interact with. Bluetooth protocols are shown on the Figure 2.8.

The upper level of software protocol stack consists of the following [PR06]:

- Host Controller Interface (HCI). This is the lowest level of software protocol stack. It interacts directly with Bluetooth adapter;
- Logical Link Control and Adaptation Layer (L2CAP). This level contains segmentation and packages assembling, protocol multiplexing, provides control information quality;
- Service Discovery Protocols (SDP). The protocol is used to find available Bluetooth devices.
- RFCOMM. It provides step-by-step transfer of data via Bluetooth like a normal serial port (COM).
- Object Exchange Protocol. Protocol is taken from the Infrared Data Association (IrDA). It makes it easy to share objects and synchronize data.

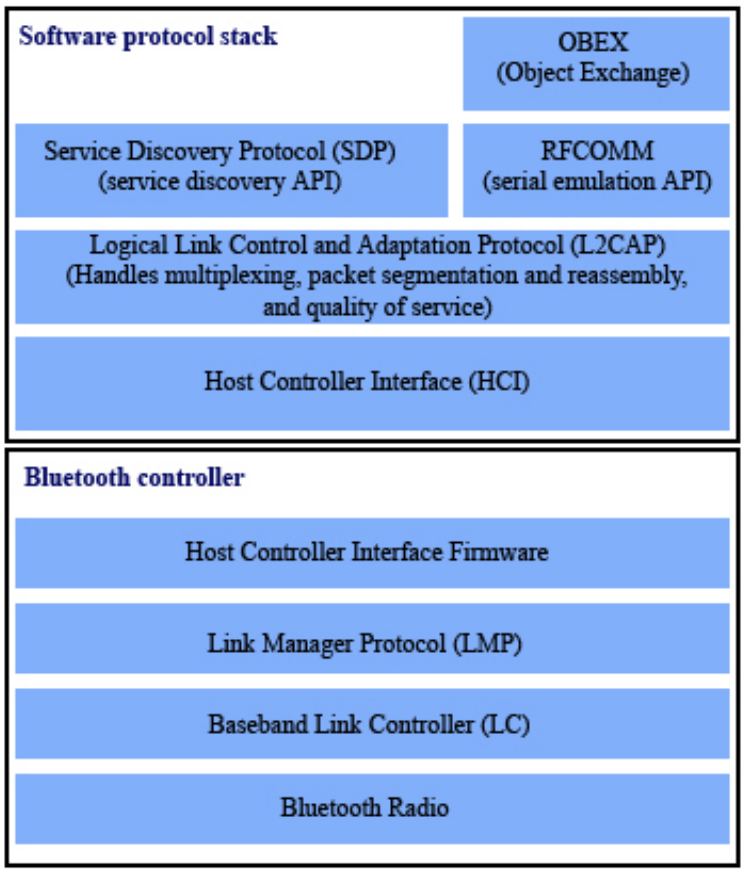

Figure 2.8: Bluetooth protocol stack [ove07].

2.3.2 Profiles

The term profile in itself refers to the set of functions and capabilities, which Bluetooth uses as a transportation unit. Profiles guarantee the possibility of an exchange of information between devices of different manufacturers. Bluetooth SIG defines several standard profiles [PR06]:

- Generic Access Profile (GAP) provides the use of lower-level protocol stack, including functions of device management. All the implementations of Bluetooth are made by GAP;
- Service Discover Application Profile (SDAP) describes the specific applications and use of SDP, accessibility and user interface aspects of the search for other devices, use of L2CAP and lower layers for the search;
- Serial Port Profile (SPP) determines interconnection requirements and opportunities for emulation of serial cable for RFCOMM, L2CAP, SDP and other

layers of the low-level interface;

- Dial-up Networking Profile (DUNP) defines the interconnection support requirements for the GAP, SPP and also for telephone calls;
- Generic Object Exchange Profile (GOEP) determines requirements for interaction and the capabilities of OBEX for file transfer, object location and synchronization for OBEX, SPP and GAP;
- Object Push Profile (OPP) designs the user interface, defines use of OBEX, SDP and the ability to locate objects for content in vCard, vCalendar, vNote and vMessage formats;
- File Transfer Profile (FTP) designs the user interface and defines the interaction and the use of GOEP, OBEX and SDP;
- Synchronization Profile (SP) designs the user interface and defines the interaction and the use of GOEP, OBEX and SDP in a similar to IrMC synchronization mode.

2.3.3 Java API for Bluetooth

All the low-level Bluetooth implementations should satisfy to a set of requirements, which comes down to support of a number of standard protocols and profiles that are shown in the Table 2.1.

Levels of Bluetooth version 1.1	**Profiles**
RFCOMM	Generic Access Profile
Service Discovery Protocol	Service Discovery Application Profile
L2CAP	Serial Port Profile

Table 2.1: Levels and profiles of Bluetooth version 1.1.

Moreover, Bluetooth devices should support the so-called Bluetooth Control Center (BCC), which can configure the local parameters of Bluetooth devices. BCC is not part of the Java Bluetooth API, but Bluetooth Java API uses it for installing the working level and for security settings.

Java API for Bluetooth is an additional package for the Java Community Process (JSR-82). This additional package introduces a common API for work with Bluetooth. The Figure 2.9 illustrates the relationship between Java API for Bluetooth and J2ME platform using Mobile Information Device Profile (MIDP) and

Connected Limited Device Configuration (CLDC).

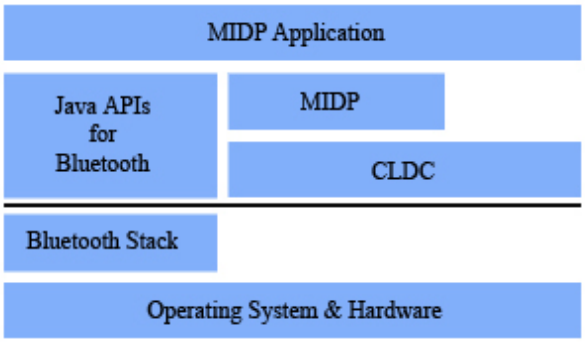

Figure 2.9: The relationship between Java API for Bluetooth and J2ME platform using Mobile Information Device Profile (MIDP) and Connected Limited Device Configuration (CLDC).

The hardware, operating system and Bluetooth stack are in the bottom, then above there are CLDC, MIDP and additional packages (Table 2.2). On the top there is MIDP application MIDlet.

Package	**Description**
javax.microedition.io	The kernel CLDC Generic Connection Framework
javax.bluetooth	The kernel Bluetooth API that includes Discovery, L2CAP, interfaces and classes of devices and data.
javax.obex	The kernel Object Exchange (OBEX) API. This package is optional and is not supported by all devices.

Table 2.2: Java APIs for Bluetooth

In Java Bluetooth API there is a new connection protocol for GCF and Object Exchange (OBEX) API based on specifications of IrDA Data Association.

The Figure 2.10 shows all interfaces and classes that can be used in the JSR-82 MIDlet.

Using Java APIs for Bluetooth consists of several distinct phases:

- Bluetooth stack initialization;
- Search of devices;

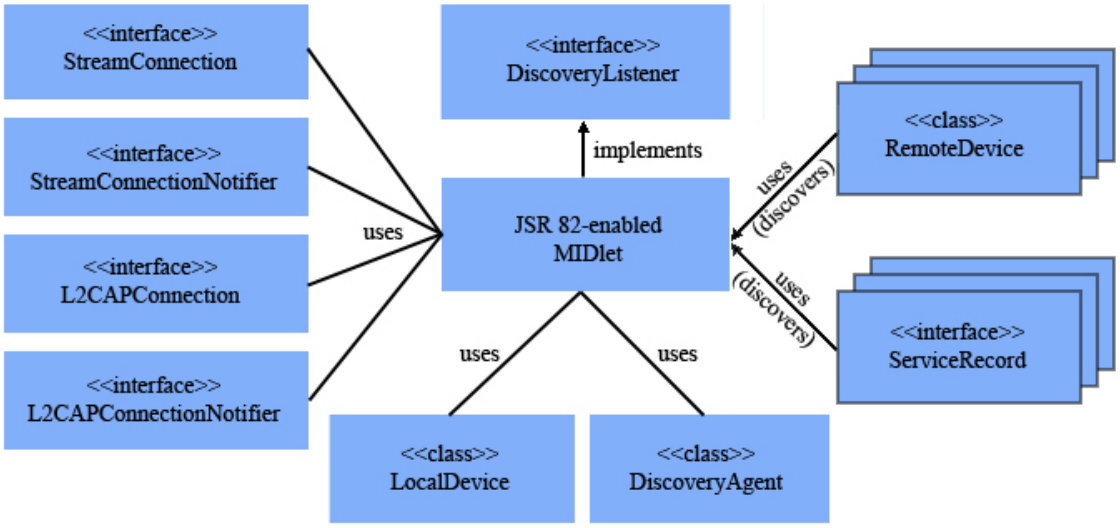

Figure 2.10: Interfaces and classes of the JSR-82 MIDlet.

- Search of services;
- Connection opening;
- Connection closing;
- Waiting for connection;
- Connection initialization;
- The I/O operations.

The Figure 2.11 schematically shows application that uses Bluetooth.

Bluetooth initialization usually entails the parameters installation. In particular, it is necessary to set device name, security settings, enable or disable the Bluetooth radio channel. All these you can do using Bluetooth Control Center (BCC), which is a set of control panels that are the main mechanism of managing Bluetooth devices.

Using Bluetooth Control Center you can set some additional settings, in particular, name, which will be seen to other devices, or the availability. You can also search for other Bluetooth devices, set join conditions, connect and disconnect from the device.

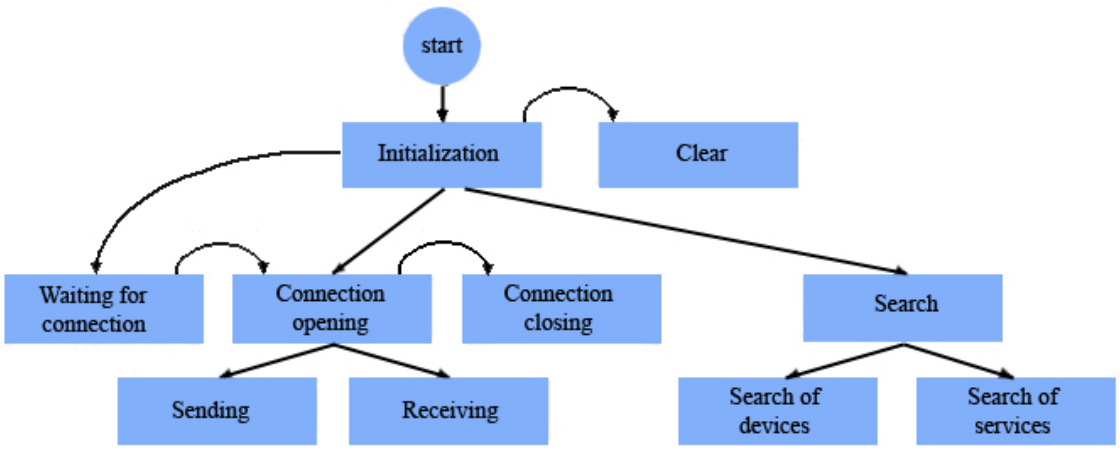

Figure 2.11: Schema of application that uses Bluetooth.

Using Java API for Bluetooth, you can create a Bluetooth connection that uses Serial Port Profile (RFCOMM), L2CAP and OBEX. OBEX is not supported by all devices, in particular, it is not supported by P900/P908.

2.3.4 Information about settings of Bluetooth device

Java APIs for Bluetooth is a set of properties for local Bluetooth device. These properties can be get by method LocalDevice.getProperty(). The list of properties of Bluetooth API is shown in Table 2.3.

More information about Java APIs for Bluetooth you can find in [Har98].

2.4 XML and XML-based language

In our project Digital Museum we had used XML language. XML (*Extensible Markup Language*) language is developed the Working Group, formerly known as SGML Editorial Review Board. It was formed with the assistance of the World Wide Web Consortium (W3C) in 1996. At the beginning of 1998 the W3C approved specification Extensible Markup Language (XML) 1.0. Current version is XML 1.1 Third Edition, W3C Recommendation, 4 February 2004 [Dic02].

Similarly HTML the XML language uses tags, but unlike HTML, XML language is designed for data processing. It was designed not for display the elements of a hypertext document, but the contents of these elements. The structure of XML is not a hypertext markup language, but the so-called meta language that is used to

Property	Description
bluetooth.api.version	The number of supported version of Java APIs for Bluetooth
bluetooth.l2cap.receiveMTU.max	The maximal size of ReceiveMTU in bytes supported by L2CAP
bluetooth.connected.devices.max	The maximal number of connected devices, including the stopped ones
bluetooth.connected.inquiry	Are queries allowed during the connection?
bluetooth.connected.page	Is the work with pages allowed during the connection?
bluetooth.connected.inquiry.scan	Is scanning of requests allowed during the connection?
bluetooth.connected.page.scan	Is scanning of pages allowed during the connection?
bluetooth.master.switch	Is it possible to switch master/slave modes?
bluetooth.sd.trans.max	The maximal number of parallel search services
bluetooth.sd.attr.retrievable.max	Maximum number of attributes of the service, which can be obtained from a service record

Table 2.3: List of properties of Bluetooth API with description.

describe other languages at a lower level.

The standardization of XML had launched the creation of technology platform for a new generation of the Web. XML language is used to specify the structure of the document, and for other purposes there are other languages, such as CSS1, CSS2 and XSL. Set of W3C standards, defining XML-based languages, is developing. These XML-based languages enhance its functionality and form technology platform of a new generation of Web, the XML platform.

There are two ways to describe the structure of XML documents and data within it: by using DTD (Document Type Definition) and using semantic patterns [Sim00]. The first method is quite outdated, but it still actively used. That is why in more detail, we will consider semantic patterns, but also give attention and DTD.

Documents description is used to impose restrictions on the document: limits of the values of elements and attributes, determination of the structure of the document, accessibility of element attributes etc.

DTD is a special language that is used to specify the structure of XML documents. By using DTD we can specify the attributes of the elements, types of attributes and elements and many other structural characteristics of the document. Here is DTD that we had used in our project:

<!DOCTYPE category [
<!ELEMENT category (module — card) >*
<!ATTLIST category type CDATA #IMPLIED>
<!ELEMENT module ANY>
<!ATTLIST module id ID #REQUIRED>
<!ELEMENT card ANY>
<!ATTLIST card id ID #REQUIRED>
]>

The DTD is called *category* (<! DOCTYPE category []). The name is used to refer in XML document. The root element of the document contains elements *module* and *category* in random sequence of. Moreover, the symbol "*" indicates that the root element may both not contain sub-elements at all and contain them a lot.

The root element *category* has the attribute *type*. This attribute has the type of CDATA and is optional. Elements *category* and *module* contain obligatory attribute *id*, which is the type of ID.

As already mentioned, the second way of describing an XML document is called XML schemes or XSD (XML Schema Definition). The main advantage of XML schemes is that for the creation of XML scheme it is not necessary to study new language. XML schemes are written in the XML.

Now let us consider the benefits that XML scheme gives for describing document:

- Number of available types for the description is increased in comparison with the DTD;
- XML schemes support namespace;
- The support the new contextual model *open*. This contextual model allows you to add attributes or child nodes to the element, other than those described the scheme;
- Also there are added many other capabilities, such as support of maxOccurs, minOccurs, number of sub-elements in the element.

As an example let us consider the part of XML scheme:

<Schema xmlns="urn:schemas-microsoft-com:xml-data" xmlns:dt="urn:schemas-microsoft-com:datatypes" >
<AttributeType name="Name" dt:type="string" required="yes"/ >
<AttributeType name="ID" dt:type="id" required="yes"/ >
<ElementType name="Module" content="eltOnly" model="open" order="many" >
<attribute type="Name"/ >
<attribute type="ID"/ >
<element type="Card" minOccurs="0" maxOccurs=""/ >*
<element type="Module" minOccurs="0" maxOccurs=""/ >*
</ElementType >
...
</Schema >

In the scheme are described accessible attributes *Name* and *ID*. There is described the element *Module*, which contain obligatory attributes *Name* and *ID*, and can contain also sub-elements of *Module* and *Card* types. The number of sub-elements is unlimited. There can be other elements inside the element *Module*. It is set out by *content= "eltOnly"*. The element *Module* has a model open. The order="many" indicates that the elements *Module* can appear in any order. More detailed information about the syntax of XML schemes is available on the web site `www.w3c.org` or in the *MSDN* library ([Lib07]).

Note that use of DTD or XML schemes is not necessarily in the XML documents. DTD and XML scheme are used only for the document validation. The scheme and DTD may be external and internal, i.e. to be stored in separate files or inside the XML document.

We had embedded tools for data export to XML packages into our project. The XML-package can be formed from the entire database, data selection or some object description. It contains museum exhibits description, full information about each object, such as attributes, images, audio and video materials. XML package is designed for development of multimedia products for users, and also for exchange the information between databases, union of descriptions made on different local PCs into one database.

There are many XML-based formats. One of the most popular is RSS format. RSS is a format for syndicating news, ranging from leading news sites such as Wired, Slashdot, to the personal weblogs. But in fact, it not only can publish news.

Almost any material which can be divided into separate parts can be published through RSS, for example, the announcements about changes in museum multimedia archive, new exhibit description. Once the information is converted into RSS format you can get information on the changes, and depending on the result, for example, automatically take some actions by using special program for RSS feeds reading.

Programs that able to work with the RSS are called news aggregators. Aggregator allows you to collect all the publications together, and you can simultaneously monitor the emergence of new news at all sites at once and read them an outline without going to each site individually.

Today there are several different versions of at least two different RSS formats. The original RSS version 0.90 was developed by Netscape company. Currently there are seven different formats and all of them are called RSS (Table 2.4) ([RSS07]).

Version	**Owner**	**Advantages**	**Status**
0.90	Netscape	-	Obsoleted by 1.0
0.91	UserLand	It is very simple	Officially obsoleted by 2.0. It is still popular
0.92, 0.93, 0.94	UserLand	Here are more capabilities than in version 0.91	Obsoleted by 2.0
1.0	RSS-DEV Working Group	It based on RDF language and can be extended by using modules.	Stable core. There is active module development
2.0	UserLand	It can be extended by using modules. Can be easily transferred from versions 0.9	Stable core. There is active module development

Table 2.4: RSS formats.

For our project we had chosen RSS version 2.0, because we needed format for publishing news of general purpose. Here is an example of RSS 2.0 that we used in Digital Museum:

<rss version="2.0" xmlns:dc="http://purl.org/dc/elements/1.1/" >

In the Table 2.5 is the list of required and optional channel elements with a brief description ([RSS07]).

Element	Description
title	The name of the channel. It's how people refer to your service. If you have an HTML website that contains the same information as your RSS file, the title of your channel should be the same as the title of your website.
link	The URL to the HTML website corresponding to the channel
description	Phrase or sentence describing the channel.
language	The language the channel is written in. This allows aggregators to group all Italian language sites, for example, on a single page. A list of allowable values for this element, as provided by Netscape, is here. You may also use values defined by the W3C.
copyright	Copyright notice for content in the channel.
managingEditor	Email address for person responsible for editorial content.
webMaster	Email address for person responsible for technical issues relating to channel.
pubDate	The publication date for the content in the channel. For example, the New York Times publishes on a daily basis, the publication date flips once every 24 hours. That's when the pubDate of the channel changes. All date-times in RSS conform to the Date and Time Specification of RFC 822, with the exception that the year may be expressed with two characters or four characters (four preferred).
lastBuildDate	The last time the content of the channel changed.
category	Specify one or more categories that the channel belongs to. Follows the same rules as the <item>-level category element.
generator	A string indicating the program used to generate the channel.
docs	A URL that points to the documentation for the format used in the RSS file. It's probably a pointer to this page. It's for people who might stumble across an RSS file on a Web server 25 years from now and wonder what it is.
cloud	Allows processes to register with a cloud to be notified of updates to the channel, implementing a lightweight publish-subscribe protocol for RSS feeds.
ttl	ttl stands for time to live. It's a number of minutes that indicates how long a channel can be cached before refreshing from the source.
image	Specifies a GIF, JPEG or PNG image that can be displayed with the channel.
rating	The PICS rating for the channel.
textInput	Specifies a text input box that can be displayed with the channel.
skipHours	A hint for aggregators telling them which hours they can skip.
skipDays	A hint for aggregators telling them which days they can skip.

Table 2.5: The list of required and optional channel elements.

Chapter 3

Concept of the digital museum

3.1 Review of the virtual museum

With acceleration of creating and accumulation of information looking up for new case tools that can provide fast and effective access to different knowledge from different sources become necessary. One of such tools is digital collections that become widely in use from early 1990's [OBMM00]. Digital collection is distributed informational system for safe storing and efficient usage of dissimilar variety of objects such as text, graphics, audio, video etc. The idea of the digital collections is more close to the traditional museums. We have the same idea of fusion of different types of information, the same hierarchy of sections and exhibits. That's why nowadays there are so much virtual museums in the Internet. For example, virtual museum of Nobel (`www.nobel.se`, Javascript), museum of Louvre (`www.louvre.fr`, HTML), Hermitage museum (`www.hermitage.ru`, CGI). There are 97 400 000 web pages for the query "digital museum" and 43 200 000 web pages for the query "virtual museum" on the beginning of 2007 year. Virtual museums are different by their structure, direction and content. Analysis of existing museums was very important for constructing our virtual museum. The inappreciable experience of these museum's authors was used as a basis of our project.

In spite of apparent variety there are just few amounts of methodologies of developing the digital data storages nowadays. Usually designers use one of the following tools. By our analysis there is the following percentage distribution of using the tools for developing the digital collections in Table 3.1.

The absolute leader is HTML technology. Despite the appearance of new promising technologies, such as Java, XML etc, the basis of most sites is the model hopelessly laggard from the modern requirements. Even if there are other technologies

HTML	DHTML	JAVA	CGI	PHP(API)	XML(XHTML)	Other
40	10	15	18	10	5	2

Table 3.1: Percentage distribution of using the tools for developing the digital collections.

they are used just as some small insertions into HTML-pages. Other extremity is writing the huge and complicated programs using CGI/API technology. All the methods have one fundamental disadvantage: there is no unambiguous formalized informational model and hence the sites built using just these technologies, without using any CASE tools [MH06], have non-flexible structure. In consequence of it the upholding of the similar projects demands high span time and price costs. CASE-tools allow minimizing the outlays on development and supporting of software. One of the postulates of CASE-tools is the possibility to build the visual model of the application environment of the informational system. Visual models of informational systems vividly show the operated system in its entirety. In such a manner even in the early stages of designing the risk of error occurring that may carry the reconsideration of the whole model is minimal. CASE-tools support all main stages of development and supporting of informational system including the analysis and formulation of requirements, designing the application software and databases, code generation, testing, documentation, configuration management and project management and etc.

Despite the decisive superiority of CASE-modeling there are not a tithe of digital collections created with CASE-tools. It tells about necessity of construction the formalized model that is able to create the virtual museum of any direction and any structure for the minimal time.

Here was created by us the computer model of the virtual museum satisfied following properties:

- Easiness and simplicity of modification;
- Suitability for modern requirements to the online museum interface;
- Platform independency;
- Independency of programming language and DBMS (data base management system) [12006];
- Full automatization of development process and museum supporting.

Nowadays there are many CASE-tools able to model different aspects of application work. Some of them are suitable for creating the functional model, or description of program's functional capabilities (BPWin, PLATINUM technology Company), others are used for building the object model of application environment (Rational Rose, Rational Software Company), and by using the last tools it is easy to design logical and physical model of databases (ERwin) [MM87]. It should be noted that functional and object models of building the application are competitive somehow while database modeling is necessary in both cases. In many cases the CASE-environment supports several model types at the same time. For example, Rational Rose 2002 Enterprise, a product for object oriented analysis, modeling, designing and development tools, entirely supports UML 1.4 standard, contains tools for code generation for main programming languages, has friend interface, includes possibilities for building the model from source code, renewal the model from the code and rebuilding the code from the model.

The object model is built by using UML. It represents formalized structure of any virtual museum. Model contains both interface and application logic. Interface includes the authentification, selection of museum section, selection of exhibit etc. The application logic provides the museum work: selection of information from database, its visualization, editing the database etc.

The building of the object model consists of three stages. The main diagram (use case diagram Figure 3.1) was built on the first stage.

This diagram describes the functionality of informational system, its main actions and capabilities. The use case diagram is the basis for the entire model. For diagram construction was used the principle of decomposition which is described below. At first after previously done analysis the task is divided into several big blocks, aka "use cases". Each of them reflects part of program logic of application. In our case the diagram contains several actors that represent museum visitors with different access rights: guest, administrator and operator of database. Each type of visitors can execute some certain actions, for example, actor "Visitor" can use the case "Authentication", and actor "Administrator" can apply the case "Set access rights" etc. All possible actions of the museum visitors are described in the use case diagram of the model, but it do not reflect entirely the work of the program. For full description by using the principle of decomposition the logic of each use case should be detailed by extended embedded diagrams. They can be both use case diagrams and other diagrams. The procedure can be repeated until the work of the application become clear enough. Such a modeling principle allows constructing the software as "top-down programming". It can significantly cut down the elapsed time and mistakes in development. Consequently we have got functional model that

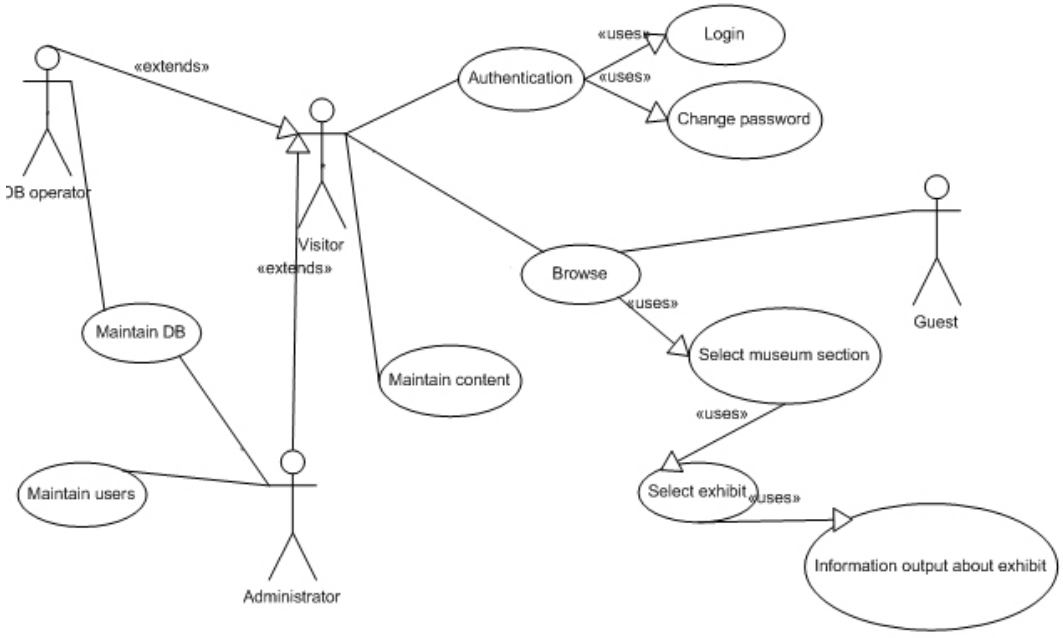

Figure 3.1: Use-case diagram for the class "Instance screen".

reflects the museum capabilities to the full.

After constructing the functional structure of the application in the form of use case diagram the class diagram was built in the next stage (Figure 3.2). The class diagram is the basis for program code generation on the selected programming language. The diagram is the central part of the museum model. It contains all required classes: actors, museum screens, databases etc. Each class contains attributes and operations on data. Also different connections between classes are represented on the diagram. As an example let us consider the class "Instance screen". After getting ID of the selected instance it initializes the output screen and prints the information about the instance to the screen.

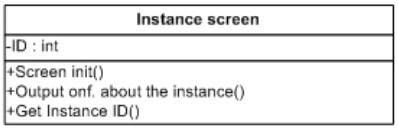

Figure 3.2: The class "Instance screen".

In virtue of each operation (in our case there three operations) program code is generated. But the code is not a complete application; it is just templates of

operations. The operations themselves should be completed by hand using suitable programming language. However, the generated skeleton significantly simplifies the coding since it contains the logic of the application in the form of templates and comments.

PHP was used as a programming language for our digital museum, because it is suitable first of all for web applications and it does not require any software for installing on the client's computer. However, if it would be necessary it is easy to change to another programming language.

On the second stage also was built the logical model of the database which will be considered carefully in the next section.

The class diagram does not show entirely the interaction between classes and the principle of the application work. There are other types of diagrams [MH06] for this purpose that shows the model dynamics:

- Collaboration diagram;
- Sequence diagram.

These diagrams detail the collaboration of classes and sequence of actions for the selected use case.

For example, let us consider the sequence diagram for the use case "Instance selection" (Figure 3.3).

At first the class "Section selection screen" forces the execution of the class "Instance selection screen" of the operation of initialization of the screen. Then the query to the database "Instance" which is the list of all instances of the selected section of the museum is sent. In this stage there is possible further decomposition. After selecting the number of the instance by the actor "Visitor" the initialization of the class "Instance screen" takes place. Basically collaboration and sequence diagrams serve the same purposes. Therefore, there is no use constructing both of them.

On the final stage after similar description of all use cases was developed the structure of the program, i.e. classes were divided into modules (files) by using component diagram, the connections between classes were established and the main module was coded. On the coding level we can add new classes, change the attributes and operations on the model classes and after all synchronize the code and the model by applying the "back development". Repeating this procedure several

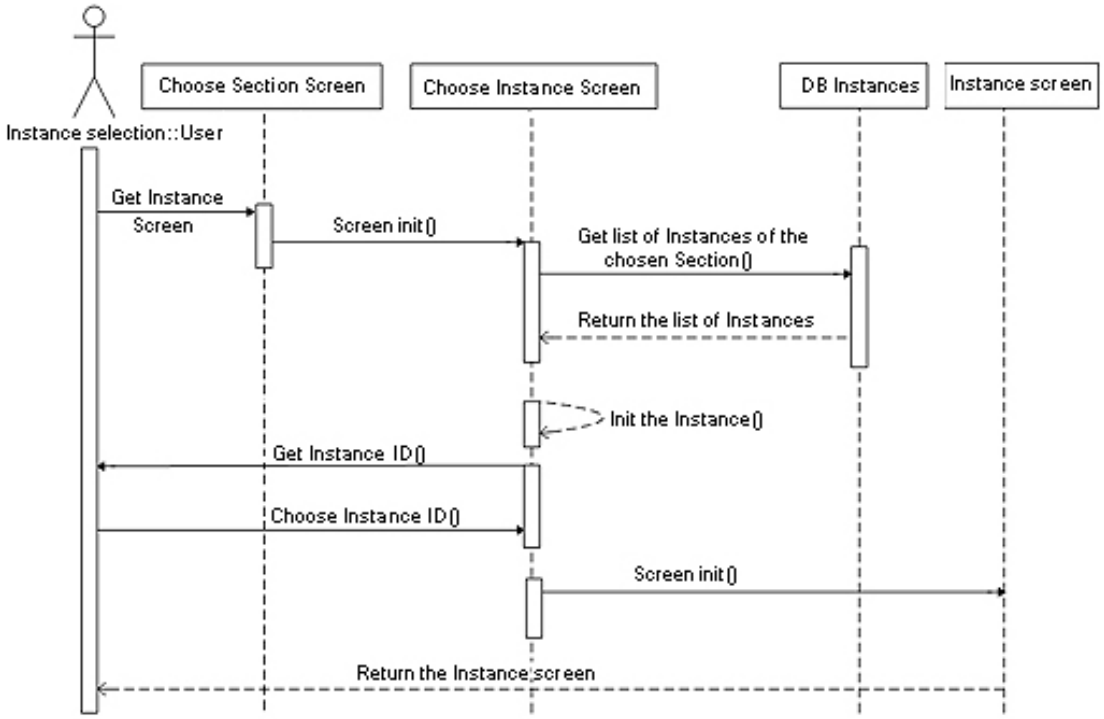

Figure 3.3: Sequence diagram for the use case "Instance selection".

times is called *round-trip modeling* that is the basis of step by step improvement of the tasking and coordination of requirements to the application.

That fact that object model does not depend on the usable programming language and the data model does not depend on DBMS is the main advantage of using CASE-tools.

As a result the formalized scheme of building the digital collections was constructed. By using our scheme it is easy to create any virtual museum in a short time. CASE-tools give us the opportunity to make unified model that is independent both of programming language and DBMS. It provides cross-platform system, easiness of information updating and low span time and price costs of collection development and supporting.

3.2 Multimedia database system

Every time during constructing multimedia products developers should choose a format of storing multimedia information. In our project the format is a key player

in system development.

There exist three methods of information storing: allocation of data directly in program body, storing information as files, using database technology. Each of the methods has both advantages and disadvantages. After their analysis we had chosen mixed data organization. The textual information we store in database and audio and video information we keep in files. We had used files, because it costs too much time and RAM memory space to manage audio and video information in databases.

Let us focus in the section on the capabilities of the databases [IKO+96].

Productivity optimization of databases for the web

For getting maximum of the productivity of open source databases you need deep knowledge. There hardly can be found web application that does not using database. If you do not have enough money or if you are adherent of open source code products then much more likely you will develop your application based on PHP Hypertext Processor and some open source database. In such case it would be useful to study methods that render possible to get everything that databases can do. In the thesis we consider some of productivity increasing methods that suit for almost all open source databases.

Optimization at a level of databases

The fastest method of productivity increasing of the database program code is substitution of embedded SQL operators [Mol05] with stored procedures.

Stored procedures usage

Stored procedures [Mol05] are subprograms contained in the database. Such procedures are preliminary compiled by a database processor. They substantially increase the database productivity by eliminating multiple references of calling application to the database core. Usually they are PHP pages. Moreover, stored procedures are much easier supported and supplied. The logic of the program is collected in one place and all the changes immediately come into operation. When the program code always call for the same procedure, then it is much easier to make sure that business logic stand in one-to-one correspondence with functional requirements. It would save a lot of time if the procedures are generalized and can be used in other projects.

Moreover, stored procedures allow to cut the network traffic intensity more significantly in comparison with SQL code execution in such environments as PHP, ASP (Active Server Page), JSP (Java Server Pages) etc. And finally, if you need scaling the application, it would be easier to extend the code to several servers, when the logic of database access is stored and executed within the database.

Storage of the multimedia data in file system

Multimedia data, whether they are static images, audio files or movies, are often considered as binary objects. There is special term for it: BLOB (Binary Large Object). BLOB fields can be stored in database or in file system. In the latter case the paths to the BLOB objects are stored in database. Storing BLOB in file system requires a bit more work, but it would allow to get higher productivity than in storage in database. With increase the amount of binary objects the productivity of database rapidly decreased. Deleting the objects are at risk of appearance the "dead zones" in database. When the entire information go through database processor then it is more difficult to support multitask operation. On the contrary, BLOB objects storing in file system make easy the creation of links to the downloaded from the Web objects. After information load the Web server serve the file request and the database processor is concerned on other tasks. Another advantage is that fact that administrator can easily catalogue and manage the multimedia files saved on the dick drive and make their backup copies.

Indexing usage

Indexing is one of the most accurate methods of database productivity increase. It is one of the main database tools that is usually insufficiently comprehensive considered. As a rule, database records are stored ordered by their creation time. For getting some information from the concrete field we have to scan all of records one by one. Index creates set of records ordered by the chosen field and contained the reference to the original ones. Indexed databases work much faster than unindexed one. But indexing occupies some additional disk space. Also it requires more time for reindexing after each change in the indexed table.

Usage of integer-valued key field

Sometimes you want to make table without integer-valued key field. For example, there is table with information about staff. It seems that there are enough character fields *last_name* and *first_name*, fields for address and contact information. We can use names of the fields for record combining, browse and other operations. But it is not optimal way. It is better to use integer key field, for example, *person_id*. If there is no such field in the data, then we have to create autoincrement field that

will not contain any real data, just represent the key field.

Integer fields have a lot of capabilities. The probability of wrong number is much less than probability of symbolic information. After name changing, for example, after marriage, we do not have to change all the references to it. Moreover, combining the records that have integer key value is much more efficient than combining the records with character key field. It should fall into habit to create integer primary key field in each new table building.

3.2.1 Program code optimization

For getting maximal database productivity increase there can be used several different strategies of application program code optimization. We consider some code optimization recommendations that can e applied to any Web application development language.

Usage of communication session

Communication session is supported by several Web application development environments. Usually communication sessions, being popular among applied service suppliers and PHP programmers, are implemented through cookies. Web is not an environment that uses information about state, and it does not provide for programmers the information about users actions made before entering the page. Using communication session programmers can follow the user navigation process. As side effect many programmers work for saving the whole information about it in the session variables. Saving the references to the database, such as connections or input the information, seems to be attractive, but it is bad habit. Saving the connection during the communication session prevent from their combination to the pool. Also such a practice leads to memory and CPU computational capability waste.

If communication session was not terminated correctly, then connections exist during the whole time-out. Time-out period can vary, but usually it is more than 20 minutes. During this period ram memory and CPU computational capability tick over. Opening and following closing the connection on the Web page seem to lead to the resource waste, but virtually it contributes to their efficient saving. It suffices to follow the next rule of thumb: create the connection as late as possible and close it as soon as possible. The same rule should be applied to the information typing.

Optimal query usage

Method of using the SQL operators has essential influence on your Web application productivity. In particular, retrieval the long list of records from the database for their future displaying on one static Web page is irrationally. It would be much better to retrieve just some part of records at a time, for example, 10 or 50, and after that browse the next group of records, using "Next 50" button or link on the Web page.

Try to use all advantages of the SQL language to the full during the coding. For example, keyword LIMIT restricts the number of records. Keyword OFFSET passes certain number of records and returns consequent records. For getting the third group of user records in the amount of 50 lines we should use the query of the following form:

SELECT exhibit_id, exhibit_name from exhibit ORDER BY exhibit_id LIMIT 50 OFFSET 100

If tables contain a large number of columns or you are using union operations in the query, then it is better to print necessary names of fields instead of using operator *SELECT **. Printing of necessary names of fields will save a lot of CPU cycles in each code startup.

Also using SELECT operator there should be used its expression WHERE with maximum benefit. If there are numbers of several columns in the WHERE expression, then the productivity will depend on their order. The number of column with minimal number of records should be on the first place, the number of column with the next minimal number of records should be on the second place and so on for other columns.

Select operator usage

For query composition depending on the results of performance we can use one of the two following ways. The most obvious way, but also the slowest, is to execute he query and check he results using additional code. The more qualified and efficient way is to use specialized SQL functions. The CASE operator in SQL language like in other programming languages can make choice according to input parameter. In such case SELECT operator result can be controlled using CASE operator as it is shown below:

SELECT exhibit_name,

CASE
when century <18 then 'very old'
when century >18 and century <20 then 'old'
else 'modern'
END as exhibit_century FROM Exhibits ORDER BY exhibit_name;

As a result we have set of records consists of two columns: first one contain name of the product and the second one contain selected by us price interpretation.

3.2.2 MySQL

By virtue of initial characteristics and history of MySQL database development it has some specific productivity problems. For example, the highest productivity of MySQL can be reached on the Intel processors working on the Linux OS [ZB04]. There are many reasons for that, but the main one consists in ram memory distribution method. So, if you have chosen MySQL, do not use Microsoft Windows NT and vice versa.

There is an opinion that MySQL is exceptionally high-performance database. Many a man even affirm that it outperforms any other database on the market by speed. The TcX Company, exploiting a MySQL to the market, organized web site `http://www.tcx.com`, where it compares the MySQL with other databases and print the productivity comparison result. According to these results the productivity of the MySQL database processor is higher than all other database productivities on 40% in average. But let us look at things in the context of scientific skepticism.

The fact is that such a productivity of MySQL is given us not for free, but to the extent that it does not support transactions. Transactions significantly decrease the database processor productivity. Moreover, they demand journal file supporting that allows rolling back some changes or entirely canceling them. System administrator should look after the registration file size and do not allow it become too large. Furthermore, along with database backup you should make also journal files backup.

Limitations

Limitations are used for forced interconnection establishment between tables and provide data consistency. The main advantage of limitations is that fact that they prevent many probable consequences of programming mistakes. Limitations are not used so much in MySQL. It is better to avoid them as far as possible. More reasonable to make the applied logic implement all functional requirements in full and

provide data consistency. In such case you will avoid the situation when you application crash down just because one of your program limitations was set incorrectly and therefore was offend.

Some programmers spend much more time for adding limitations into code and then debug them and solve problems with their failure than for development and program logic optimization. Moreover, if you do not use limitations, you have more chances that your application can be relocated on other platforms.

Table types used in MySQL

There are four table types in MySQL: static, dynamic, Heap and archived. According to MySQL manual static tables are the fastest from three types of tables stored on the disk drives. But they can not contain variable-length columns. If there is at least one such a column, MySQL creates dynamic table instead of static one. Dynamic tables contain much more information especially about table size, but they are much slower then static tables.

Two other table types are specific. Tables of Heap type exist just in ram memory and therefore they are highly quick operating. But such tables can be just small or medium size. Archived tables are intended just for reading the information and also quick operating. Additional information you can find in MySQL manual on the web-site `http://www.mysql.com/doc`.

MySQL dead spots

When data of variable-length columns is modified and the size of new data is less then there appear "dead spots" in MySQL data files [Ull05]. There is no program for solving the problem. But in MySQL software is tool *myisamchk* that allow cleaning dead spots and optimize indexes again. This program should be run periodically for the database.

3.2.3 POSTGRESQL

In distinction from MySQL the PostgreSQL database supplies stored procedures, limitations and transactions. The PostgreSQL database developers did not go on the way of reduction the functional capabilities for productivity increase. But rich assortment of PostgreSQL functions has its own advantages. There are two unique embedded functions that can be used for increase the database productivity: VACUUM and EXPLAIN commands [Ull05].

When PostgreSQL modifies the record, it saves the original record and creates the new one in the end of its internal file. The old record is marked as outdated and is used by other transactions that still use previous state of the database. The same process takes place in deleting the records. Command VACUUM deletes outdated records from database and packs it. For keeping database "clean", the command should be run periodically.

The important method of code debugging is optimization of method of query execution. If you will simply scan visually the query for identifying bottlenecks, then it will lead nowhere. For excluding such a primitive way many databases have function of analysis that do not execute the query, but analyze it for you. Such a function in PostgreSQL is EXPLAIN. Here is an example:

EXPLAIN SELECT exhibit_number from exhibits where id=930;
NOTICE: QUERY PLAN: Seq Scan on exhibits (year=1900..2007 rows=20 width=15)

Database report that batch scanning is required.

Integer values of cost are done just for example. Value rows is a supposed number of records that should be returned by the query. The rest value width is the string width in bytes. If you will run the command VACUUM for the database then most probably you will see productivity increase predicted by function EXPLAIN. Moreover, you can speed up the query by creating index for the column *exhibit_number* and in such a manner substitute batch scanning with index scanning.

3.3 The concept of the digital museum

In this section we summarize and explain the concept of a virtual museum, which is the result of collaborative work of University of Joensuu and North-Karelian museum.

The museum is a repository of cultural heritage. Every museum has two important functions:

1. Collecting, storing of information;

2. Delivering it to the public.

In the real world, these two functions are contradictory. It is better to keep exhibits in a secure place and limit the visitors from touching them, which leads to their destruction. On the other hand they should be presented to people. The

dilemma is easily resolved in the virtual world. Modern technologies make possible to scan exhibits without loss of information and even with redundancy, which will come in useful for future research. In cyberspace first function of museum is implemented as a digital data, which does not lose information and preserves the original exhibits for future descendants. The latter function is made as a virtual exhibition of WWW interface or networking 3D virtual environment.

We have developed a virtual museum that includes four concepts:

1. Digital archive;
2. Accessibility;
3. Individual approach;
4. Distributed museum.

Let us look at them in more detail. The first concept solves two problems: preserve the collection in its present state without damage; and conflict resolution between two main functions of museum, heritage preservation and exhibition. Review of technologies of digitizing data is very important, because the data with loss of information can be futile to future research, and their scan may not be possible again [HCD01, LPC+00].

The concept of the availability consists of three types. The first is access of exhibit in its entirety. Museum visitors would like to consider the object close, touch it to see the texture, shape, to feel and explore it better. But mostly exhibits are put under the glass to protect them from damage. One touch could not have destructive power to exhibit, but the museum visitors are billions and the overall harm from them is great. Another type of accessibility is access of all exhibits for all people. Not all the exhibits are presented; exhibited objects are accompanied by brief explanations on the same language, which is a barrier to foreign citizens. And finally the third type of accessibility is the access that is associated with the time and place. The heritage of humanity is great and the huge collections may not be exhibited all at once. In this case, real museums organize thematic exhibitions, but they are limited in time. After the exhibition objects become not available. Not everyone has the opportunity to get to the museum during the exhibition.

All these problems of the accessibility of exhibits can be solved in a virtual world easily and without extra costs. Even after the exhibition all objects can be viewed at any time in any place.

The next concept is the concept of individual approach to visitors. Up to now the museum has been unavailable or restricted for weak-eyed or colorblind people, or with other disabilities. Real museum could not provide differential material. Usually the information is provided the same for all. But it would be just fantastic to explain the history and heritage value to children in simple language, interest them; provide a detailed analysis to experts, give some interesting exhibit specifications; help teachers and students in education; provide information in many languages in order to reach a greater number of visitors. We developed a virtual museum, where all of these functions are implemented in a web interface for the user that visits the museum with a home computer, or using mobile phone or PDA, if user is in a real museum. User just once adjusts its communicator and then receives information according to the settings.

Universal exhibitions are organized in the web by combining collections of different museums, which is called the distributed museum concept. Archive of the world heritage in one museum. The main task of this concept is development a common standard for storage or exchange of information. Standardization should include rigorous description of the structure of data storage (syntax) and their importance (semantics) so that the computer can understand and handle. It is impossible to search and display collections of various museums without a common standard.

We built a system that solves these problems or lays the foundation to their solution. In virtual world the space is almost infinite. Visitors of modern museums often would like to discuss the objects, share impression, ask some questions. In the real world there is no place for a number of consultants and interlocutors to satisfy all visitors. But it is possible in cyberspace. We had organized communication space, or forum, where people can communicate. It would be nice to do in the future so that they can track the movement of other visitors in the museum to see what the exhibits are free, which are popular. Now this function is carried out by statistics of visiting, which we count in the web space. By using results of statistics we can define the most popular exhibition and decide what to visit the next.

Our virtual museum has the following tools:

1. Explanation tool;

2. Zoom tool;

3. Rotation tool;

4. Virtual video tour.

It is also possible to make bookmarks for favorite articles.

Based on the above model of a virtual museum, we have implemented digital museum with two clients:

1. Digimuseo-web client;
2. WSS (*Wireless Syndication Service*) mobile client;

There is also used Semacode system in the museum. One click on the image and you get the information about the object quickly, easily, safely and according to your needs.

Chapter 4

Images on the path to the digital museum

4.1 Basics of colors

4.1.1 The light and its spectrum

The wavelength spectrum of color, or color spectrum, is the intensity distribution of light according to the wavelength. The light has wave nature. Discussing about the light, we will mean the visible portion of the electromagnetic waves. Each wave is characterized by its length. The electromagnetic spectrum that the human eye can perceive lies about 400 to 700 nanometers [sWtfe07] (Figure 4.1).

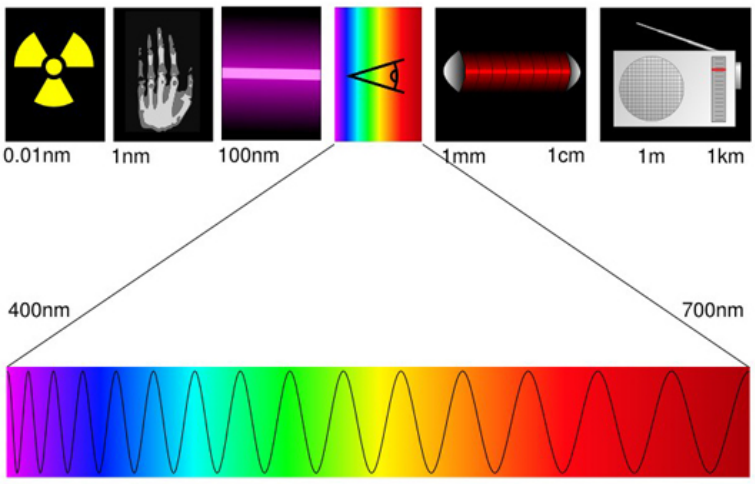

Figure 4.1: The visible spectrum of electromagnetic radiation [weWtfe07].

Inside the eye there are sensors that are sensitive to the different waves of the visible spectrum. When the electromagnetic waves reach these sensors, there is formed a signal, which then enters the brain. And the brain decides what the color of light person sees. The signal that will be formed in the human eye depends on the wavelengths that will go to its sensors. For instance, if all the wavelengths of visible spectrum will go on the sensors, then the brain will recognize that light as a white light, and if no wave of visible spectrum will go on the sensors, the signal will be understand as black [Pe01].

If we take a look through prism on the white light, we will see a rainbow (Figure 4.2). The rainbow will show what the white light consists of. And since the rainbow is the visible part of the electromagnetic spectrum, we can draw conclusions about how the eye sees different parts of the spectrum.

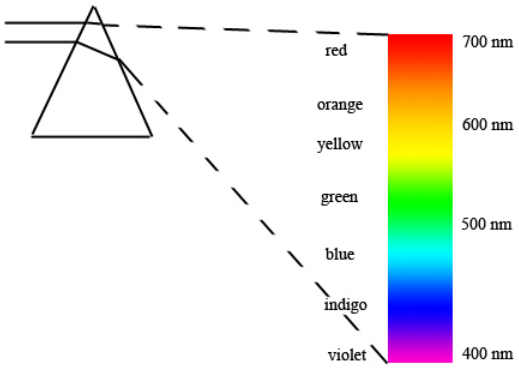

Figure 4.2: Rainbow, corresponding to the visible spectrum [oL07].

We can clearly see seven basic colors on the rainbow: violet, blue, azure, green, yellow, orange and red. For example, the wavelength of 700 nanometers our eye recognizes as red, a wavelength of approximately 450-500 nanometers it perceives as blue.

Different objects of the world can reflect, pass, or radiate light (Figure 4.3). Objects that reflect or pass light make adjustments to it. They change its spectral content depending on the properties of the materials that these objects consist of, the properties of the surface on which light falls, and many other factors. Objects, radiating light, also emit light waves of different length and intensity. The color of the object we call reflected, refracted or radiated light. The combination of wavelengths coming from the object is its spectral data. Spectral data can be obtained

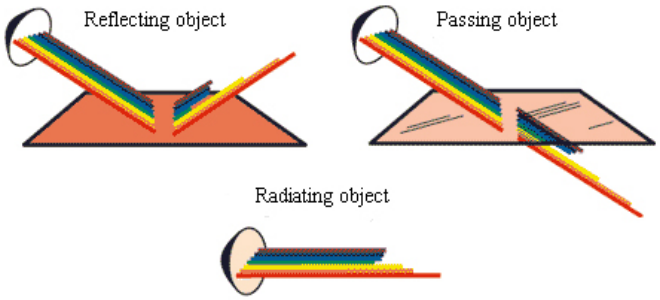

Figure 4.3: Reflecting, passing and radiating objects [the07].

as a result of careful analysis of various wavelengths that coming from the object.

To get spectral data of an object we used spectral camera in our laboratory (Figure 4.4). Spectral camera gives the spectrum of colors on the local scale, i.e. very close to the object, whose properties we want to explore.

Figure 4.4: Spectral camera and illumination.

One of the most important things that we should keep in mind in terms of color is that some things we see because they emit light, and others we see because they reflect it. When objects emit light, they become a color that we see. When they reflect light, for example, paper, their color is determined by the color of light falling on them and color that these objects reflect. Thus, the emitted light is a light going from the active source, such as the sun, light bulbs, the monitor screen; and the reflected light is a light reflected from the surface of the object.

The emitted light can contain all colors, the white light, any of their combinations, or only one color. The emitted light going directly from the source to our

eyes retains all the colors of which it was created. Some waves of radiated light are absorbed by the object, so our eyes can see just unabsorbed reflected waves.

The white paper is white because it reflects all the colors in white light, and no one is absorbed. If you highlight the white paper by a blue light, the paper will looks like blue. If you lighten a red paper, it will looks like red because it absorbs all colors except red. What will happen if light red paper by a blue light? The paper will appear black because it does not reflect the blue light falling on it.

Reflection and radiation of light remained no more than an interesting topic before computer color image processing started actively develop. Today totally opposite ways of generating the color of monitors and printers are the main cause of screen color distortions in printing. In order to produce color selection correctly, you must be aware of two opposite systems of description the color in computer: additive and subtractive [Pe01].

Additive color is produced by combining light rays of different colors. In this system, the lack of color is black, and the presence of all is white. The additive color system works with the radiated light, for example, from the computer monitor. The system uses three basic colors: red, green and blue (RGB). If they mix with each other in equal proportions, they are white or gray, and mixed in different proportions they form some other color (Figure 4.5).

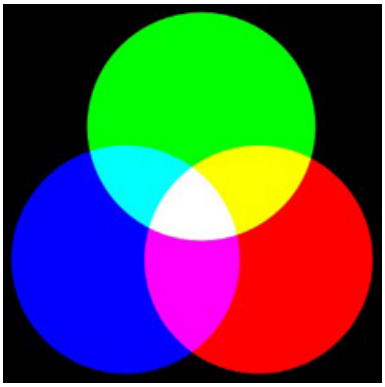

Figure 4.5: Additive color system.

The subtractive color system is the reverse: you get a color by eliminating other colors from the reflected light ray. The white in this system is got by absence of all colors, but their presence gives black color (Figure 4.6). The subtractive color system works with reflected light, such as a sheet of paper. The white paper reflects

all colors. The painted paper absorbs some colors and reflects the rest.

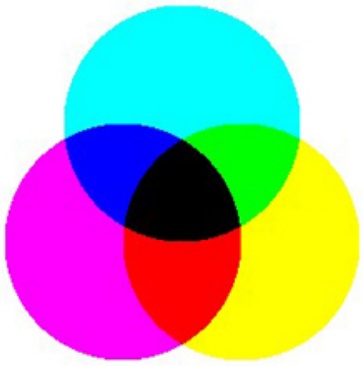

Figure 4.6: Subtractive color system.

The main colors in subtractive color system are cyan, magenta and yellow (CMY), that are opposite to the red, green and blue. When the colors are mixed on white paper in equal proportions, they produce black. Rather, it is expected that they should produce black. But in fact, inks do not absorb light fully, and therefore the combination of the all three basic colors looks like dark brown. To correct the inaccuracy, printers add a bit of black paint for a truly black color. The color system, based on a four-color printing process, is marked by CMYK.

4.1.2 RGB color system

Computer monitor creates color directly by radiation of light and therefore uses RGB color system. The surface of the monitor consists of tiny red, green and blue dots, pixels. The shape of points varies depending on the type of cathode-ray tubes CRT. CRT gives the signal of different power to screen pixels. Each point is one of the three colors. It is painted in a certain shade of its color, depending on the signal strength. Because the points are small, visually they are mixed with each other and are no longer apparent. You can create any shade of the more than 16 million that are available in RGB, by combining different values of basic colors. The RGB color system is similar to human color vision. Our eye receptors are also configured to red, green and blue colors.

4.1.3 CMYK color system

The CMYK color has been widely known long before computers become used for graphic images creating. Triad of major print colors (cyan, magenta and yellow) is an heir of three basic colors in painting (blue, red and yellow). Changing the shades of first two colors is due to the chemical composition of art inks that are distinct from the art ones, but the principle of mixing is the same. Both art and printing inks cannot produce very many shades of colors. Therefore the artists use additional paints based on pure pigments, and printers add at least black ink. The CMYK is designed and used for printing. All files for output to printers must be converted to CMYK. This process is called color selection.

4.2 Spectral approach

So, we have considered the fundamental methods of describing color. These methods can be divided into two categories.

- There is so-called spectral data, which describes the properties of surface of color object. They show how the surface affects the light: it reflects, transmits or radiate it. These surface properties do not depend on environmental conditions, such as luminance, the individuality of perception of each viewer and the differences in methods of color interpretation.
- There is so-called three-dimensional data, which in terms of three values simply describes how the color of object seems to the sensory device or viewer, or how the color will be displayed on a device, such as a monitor or printer. CIE color systems, such as XYZ and L*a*b*, position the color in color space through three-dimensional coordinates, while the color system, such as RGB and CMY (+K), describe the color in terms of the three variables, which set values of three ingredients that give some color.

The spectral data have certain advantages over three-dimensional formats, such as RGB and CMYK, as a format for the specification of colors and transfer of data about color. First of all spectral data is the only objective description of a real object, painted in some color. In contrast, the description in terms RGB and CMYK depends on conditions of object viewing, such as type of device reproducing color, and type of luminance, where the color is considered.

Comparing different color spaces, we can see that each color monitor has its own range of displayable colors, which it generates through RGB luminous colors. Even

monitors manufactured in the same year by the same producer, different from each other in this sense. The same situation is with printers and CMYK colors that have more limited range of colors than most monitors. To specify some color through RGB or CMYK values, it is also necessary to specify characteristics of the device on which the colors will be displayed.

Different light sources, such as incandescent or fluorescent lamps, have their own spectral characteristics. The color is very dependent on these characteristics. The same object appears differently under different type of illumination. To specify the color as accurately as possible through three values, it is also necessary to specify the characteristics of the light source.

Unlike all mentioned above, measurements of spectral data does not depend either on the device or the lighting. The spectra shows the light reflected from an object before it is interpreted by observer or device. Different light sources look different when their light is reflected from the object, as they contain different amount of spectrum for each wavelength. But the object always absorbs and reflects the same percentage of spectrum for each wavelength. Spectral data are measurements of this percentage.

Spectral camera captures the real nature of the color, even before it gets observer's eye. The color depends on the observer and the lighting conditions. But the way that an object affects the light does not depend on these factors.

4.3 Color exhibition

North-Karelian Museum under the auspices of in cooperation with the University of Joensuu, InFotonics Center Joensuu, the North-Karelian Polytechnics, and the National Museum of Japanese History (Rekihaku) organized exhibition *Shades of Colors* (Figure 4.7), where they introduced the visitors to the extensiveness of colors as a phenomenon, the variety of aspects it comprises.

The original purpose of the exhibition was to present colors in their physical sense using multimedia technology and Internet. It has enabled to meet the different requirements of visitors and enhance their knowledge in this field. Color is described from divergent angles, different phenomena are presented. For example, you can find there demonstration of getting rainbow using optical grating, explanation of why the sky is blue, pigments extracted from finnish nature, metamerism (Figure 4.8), eyes illusion room, demonstration about how colorblind people see

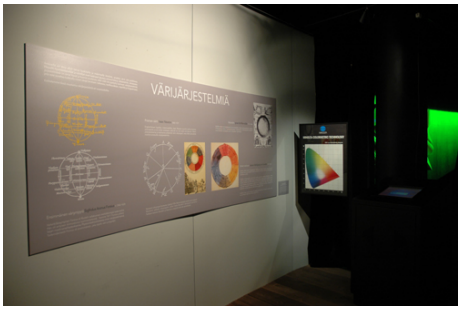

Figure 4.7: Board of color theories in exhibition *Shades of Colors* in North-Karelian Museum.

colors, rooms that show how color effects to space (Figure 4.9), etc. There are also represented four multimedia demonstrations in the exhibition about color temperature, color palette of icons, design of kimono and interpreting of colors depending on the time and place.

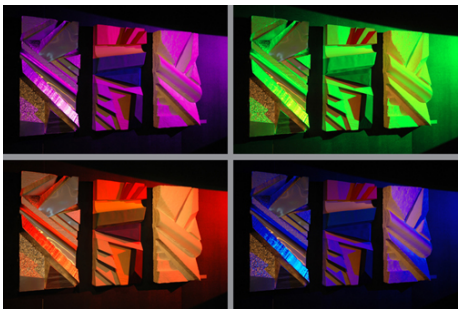

Figure 4.8: Demonstration of metamerism in exhibition *Shades of Colors* in North-Karelian Museum.

The exhibition *Shades of Colors* presents color expertise based in Joensuu (Figure 4.10). InFotonics Center Joensuu is an internationally recognized center for applied research, founded in the University of Joensuu in 2003. The center combines knowledge of photonics and information technology, the top research concentrating on wave-optical technology and spectral colors. The expertise Infotonics Center is based on high-level fundamental research in physics and computer science performed in the University of Joensuu. The Center also connects the university research and the industry which applies its results to practice. InFotonics Center is an active partner in international research networks as well.

The development of digital technology has introduced new requirements, new

Figure 4.9: Rooms in exhibition *Shades of Colors* in North-Karelian Museum that shows how color effects to space.

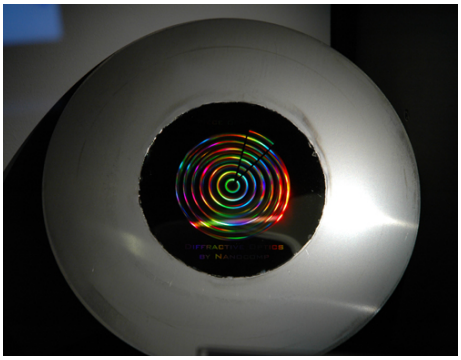

Figure 4.10: Color image on the metal plate produced by diffractive optical technology.

dimension in the study and interpretation of colors.

The exhibition is dedicated to colors, its history, cultural significance and color research. The exhibition as a whole consists of the exhibits presented in the museum Carelicum, and its virtual part is presented on our Web-site `http://digimuseo.joensuu.fi` (Figure 4.11).

The usage of colors can be considered from different points of view: the symbolism, traditional color palette, color used by artists, the history of pigments, etc. Our understanding of the color depends on the culture of country where we live. The colors are in all areas of our lives.

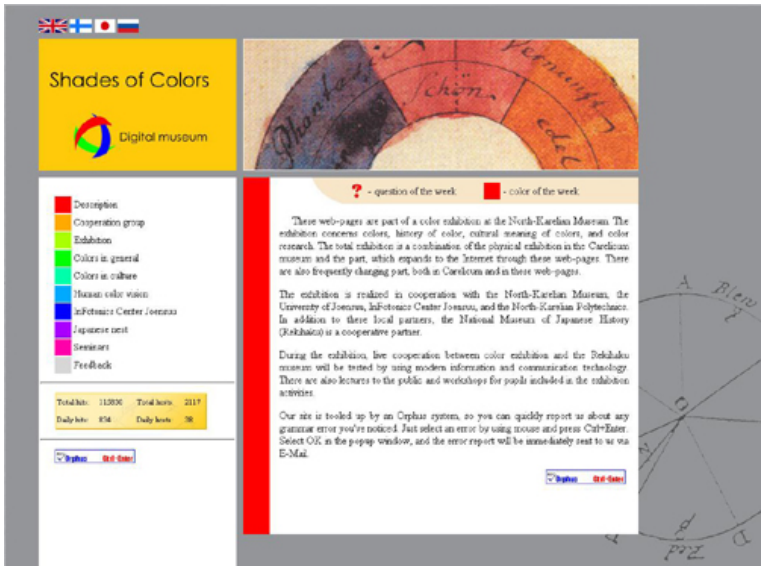

Figure 4.11: Web-site *http://digimuseo.joensuu.fi*.

The exhibition shows the difference in the perception of colors depending on the cultural differences in the case of the Japanese culture. Japanese nest shows how the achievements of digital color research are used in the National Museum of Japanese History Rekihaku (Figure 4.12).

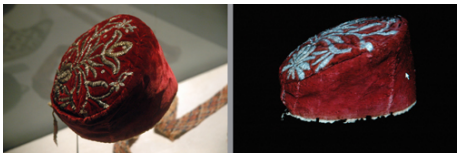

Figure 4.12: Exhibit and its digital three dimensional image from Rekihaku produced by using spectral approach.

The simulation of color objects of cultural heritage has become increasingly important area of study. InFotonics Center Joensuu and the National Museum of Japanese History, also known as Rekihaku, closely cooperate in this field. They presented applications of color analysis in the research of kimonos. The kimonos are part of the museum collection of Rekihaku and belong to Edo period (1603 - 1867). In Figure 4.13 you can see the real kimono that is presented in the exhibition.

A number of objects in the Museum of North Karelia, on view in this exhibition, have been shot for the first time by using the new digital technology. Applying the

Figure 4.13: The real kimono from Edo period represented in exhibition *Shades of Colors* in North-Karelian Museum.

spectral approach, which is described above, we have got images of icons (Figure 4.14) and different textiles.

Still, traditional museums will not disappear. Digital presentations increase interest in the physical holdings. By using modern technologies the information becomes more attractive, fully represented and easily spread over the world. Physical museums and their companion digital museums will give new possibilities to demonstrate the world from its best part. Sometimes they are located together; sometimes separately. We hope that our project will help museums research, organize and present their holdings.

Figure 4.14: The image of icon made by spectral camera.

Chapter 5

Discussion and Conclusions

Last years can be seen slow but steady growth of digital technologies, as well as the fact that the museum could not remain out of the technological trends, if they wish to attract audiences of the XXI century. Tomorrow, the museum visitors will be people to whom computers and multimedia already playing a significant role in their lives, education, recreation, and work. The interest in developing mobile, digital technologies to enhance the experience of visitors in museums is growing.

The most common way to implement such a system today is the WWW Internet service. It is the most simple and effective method of network access to the information resources of all kinds. As a technology platform of the Digital Museum we used the methodology of the electronic collections in distributed information systems. This technology provides a distributed storage and presentation of data, the access isolation of information resources and databases, created on the server. It allows implementing authorized access to data. However, there is no such an important component in currently used technologies as the global standardization of data organization and presentation formats. Recently, there were an appropriate approaches and technologies that are based on different ways of working with metadata, such as XML and RDF.

There are also thematic multimedia virtual tours in the museum. Walking in the halls of museum in cyberspace becomes exciting journey. Multimedia virtual tour is a new technology, a way of space performance in cyberspace, which plunges the user into the atmosphere of created world. All-round panoramic view creates an illusion of presence in a real architectural environment. Virtual tour intrigues and attracts visitors to the museum. User can interact with the virtual world: roam around the museum, observe different processes, examine artifacts, interact with active objects. Technically virtual tour is a series of interactive scenes, which are connected by passages. Each interactive scene is a panoramic picture of objects and

sounds. There are used video clips to animate the view.

During the exhibition *Shades of Colors* the North-Karelian museum was visited by 11 179 people, where 2176 people are pupils and students. In the virtual museum were 200350 visitors. Average number of visitors per day was 834 people. The exhibition lasted 8 months.

In the future it is nice to implement accessing museum collections through games. Millions of children and adults regularly engage in informal learning through game play. The digital museum was developed based on mechanism of digital archiving. It is currently completed a mechanism of automatic web production, content management system, support of multiple services and needs. The next step should be development of optimal search system in the multimedia database.

We hope that our work will help in the future for development of a global virtual museum that permit easy access to the cultural heritage stored in museums around the world. Nowadays the main task is development of unitary standard of data storage.

Bibliography

[12006] *Fundamentals of Database Management Systems*. Academic Internet Publishers, Inc., 2006.

[20007] GSM Association Press Release 2006. `http://www.gsmworld.com/news/press_2006/press06_51.shtml`, August 2007.

[AB95] Alasdair Arthur and Peter Burke. Mercury's experiences with 3 layer architectures. `http://ourworld.compuserve.com/homepages/alasdair_arthur/3LAexp.htm`, November 1995.

[Bat01] Regis J. Bates. *GPRS: General Packet Radio Service*. McGraw-Hill Professional, 2001.

[BTs] Bluetooth specification. v. 1.1. part h:1. host controller interface.

[CCP03] Rajiv Chakravorty, Andrew Clark, and Ian Pratt. Gprsweb: optimizing the web for gprs links. In *MobiSys '03: Proceedings of the 1st international conference on Mobile systems, applications and services*, pages 317–330, New York, NY, USA, 2003. ACM Press.

[Dic02] Kevin Dick. *Xml: A Manager's Guide*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2002.

[Glo07] PA Online Web Design A Beginner's Web Glossary. `http://www.paweb.net/webglossary.htm`, August 2007.

[GPZ99] Rajamani Ganesh, Kaveh Pahlavan, and Zoran Zvonar. *Wireless Multimedia Network Technologies*. Kluwer Academic Publishers, Norwell, MA, USA, 1999.

[Har98] Elliotte Rusty Harold. *Java I/O*. O'Reilly & Associates, Inc., Sebastopol, CA, USA, 1998.

[HCD01] Tim Hawkins, Jonathan Cohen, and Paul Debevec. A photometric approach to digitizing cultural artifacts. In *VAST '01: Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage*, pages 333–342, New York, NY, USA, 2001. ACM Press.

[HCH00] Jen-Shin Hong, Herng-Yow Chen, and Jieh Hsiang. A digital museum of taiwanese butterflies. In *DL '00: Proceedings of the fifth ACM conference on Digital libraries*, pages 260–261, New York, NY, USA, 2000. ACM Press.

[IKO+96] Hiroshi Ishikawa, Koki Kato, Miyoki Ono, Naomi Yoshikawa, Kazumi Kubota, and Akiko Kondo. A next generation industry multimedia database system. In *ICDE '96: Proceedings of the Twelfth International Conference on Data Engineering*, pages 364–371, Washington, DC, USA, 1996. IEEE Computer Society.

[Jaw01] Thota Neena Jawaharlal. Web-based application for aircrew. Master's thesis, Dept. of Mathematics, Computer Science and Mechanics, Warsaw University, Poland, `http://students.mimuw.edu.pl/SR/prace-mgr/jawaharlal/index.htm`, 2001.

[Lib07] MSDN Library. `http://msdn2.microsoft.com/en-us/library/default.aspx`, August 2007.

[LPC+00] Marc Levoy, Kari Pulli, Brian Curless, Szymon Rusinkiewicz, David Koller, Lucas Pereira, Matt Ginzton, Sean Anderson, James Davis, Jeremy Ginsberg, Jonathan Shade, and Duane Fulk. The digital michelangelo project: 3d scanning of large statues. In *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 131–144, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[MH06] Russ Miles and Kim Hamilton. *Learning UML 2.0*. O'Reilly Media, Inc., 2006.

[MK02] Chuck Musciano and Bill Kennedy. *HTML and XHTML: The Definitive Guide*. O'Reilly & Associates, Inc., Sebastopol, CA, USA, 2002.

[MM87] David A. Marca and Clement L. McGowan. *SADT: structured analysis and design technique*. McGraw-Hill, Inc., New York, NY, USA, 1987.

[Mol05] Anthony Molinaro. *SQL Cookbook (Cookbooks (O'Reilly))*. O'Reilly Media, Inc., 2005.

[OBMM00] Oguzhan Ozcan, Egeli B., Korzay M., and Ozturak M. The use of interactive media for cultural heritage. In *Technology Impact on Cultural Tourism Conference Proceedings*, pages 247–261. Bogazici University, June 2000.

[oL07] Dispersion of Light. `http://www.school-for-champions.com/science/light_dispersion.htm`, August 2007.

[ove07] Bluetooth overview. `http://developer.uiq.com/devlib/uiq_30/SDKDocumentation/sdl/guide/N1008E/BluetoothOverview.html`, August 2007.

[Pe01] K. N. Plataniotis and et al. BOOK REVIEW: Color Image Processing and Applications. *Measurement Science and Technology*, 12:222–+, February 2001.

[PR06] C. S. R. Prabhu and Prathap A. Reddi. *Bluetooth Technology and Its Applications with JAVA and J2ME*. Prentice-Hall of India Pvt.Ltd, 2006.

[Rho01] John Rhoton. *The Wireless Internet Explained*. Butterworth-Heinemann, Newton, MA, USA, 2001.

[Ris02] Ray Rischpater. *Wireless Web Development*. APress L. P., 2002.

[Rom97] Steven A. Roman. *Access Database Design and Programming*. O'Reilly & Associates, Inc., Sebastopol, CA, USA, 1997.

[RSS07] What Is RSS. `http://www.xml.com/pub/a/2002/12/18/dive-into-xml.html`, August 2007.

[RTA05] Dimitrios Raptis, Nikolaos Tselios, and Nikolaos Avouris. Context-based design of mobile applications for museums: a survey of existing practices. In *MobileHCI '05: Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*, pages 153–160, New York, NY, USA, 2005. ACM Press.

[Sim00] John E. Simpson. *Just XML*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2000.

[SK99] Narushige Shiode and Tomoko Kanoshima. Utilising the spatial features of cyberspace for generating a dynamic museum environment. In *VRML '99: Proceedings of the fourth symposium on Virtual reality modeling language*, pages 79–84, New York, NY, USA, 1999. ACM Press.

[sWtfe07] Visible spectrum Wikipedia the free encyclopedia. `http://en.wikipedia.org/wiki/Visible_spectrum`, August 2007.

[the07] Color theory. `http://www.xrite.com.ua/colorTheory1.php`, August 2007.

[Ull05] Larry Ullman. *PHP and MySQL for Dynamic Web Sites: Visual Quick-Pro Guide (2nd Edition) (Visual QuickPro Guides)*. Peachpit Press, Berkeley, CA, USA, 2005.

[US96] Tomonori Usaka and Ken Sakamura. A design and evaluation of the multi-user virtual environment server system for the digital museum. In *TRON '96: Proceedings of the 13th TRON Project International Symposium*, page 60, Washington, DC, USA, 1996. IEEE Computer Society.

[UYF+98] T. Usaka, S. Yura, K. Fujimori, H. Mori, and K. Sakamura. A multimedia mud system for the digital museum. In *APCHI '98: Proceedings of the Third Asian Pacific Computer and Human Interaction*, page 32, Washington, DC, USA, 1998. IEEE Computer Society.

[weWtfe07] Electromagnetic wave equation Wikipedia the free encyclopedia. `http://en.wikipedia.org/wiki/Electromagnetic_wave_equation`, August 2007.

[ZB04] Jeremy D. Zawodny and Derek J. Balling. *High Performance MySQL: Optimization, Backups, Replication, Load-balancing, and More*. O'Reilly & Associates, Inc., Sebastopol, CA, USA, 2004.

[ZJK98] Zoran Zvonar, Peter Jung, and Karl Kammerlander, editors. *GSM: Evolution Towards 3rd Generation Systems*. Kluwer Academic Publishers, Norwell, MA, USA, 1998.