6th International Conference on Applied Informatics
Eger, Hungary, January 27–31, 2004.

The Music Library of the Future

György Barna Iszály

Department of Computer Graphics and Library and Information Science,
University of Debrecen
e-mail: gyiszaly@inf.unideb.hu

Abstract

In the last 20 years the Library and Information Science developed remarkably. The result of this is that in our present days most of the libraries use integrated library system and new electronic libraries were established. These latter included not only the bibliographic description but the full text of the documents.
However, music libraries were left out of this development. Although, the digitally acoustic formats have already been available for ten years, nothing happened in this domain. The hour has come to apply this knowledge. With the new result of the hardware and software, we can create a new, totally digitally music library system.
In the music library system of the future the acoustic documents would be stored in digitally format for instance: MP3 format. The user could reach the system from the library through the local network or from home with password through the Internet. The user could search not only on the basis of the title and author, but the album title, the music style or the participants too. This system could serve any number of users simultaneously, and at any time.
In this paper I will show, how we can build up such a system. I will acquaint the reader with the digitalization process of the acoustic documents. I will outline how we can create MP3 audio compressed format from the digitalized documents and why we can use this format in this system. At last, I will show what advantages this new system has compared to the existing systems.

Key Words and Phrases: Library, Library Integrated System, Digitalization, Mp3

1. Library Integrated System

In our time big libraries are unimaginable without developed integrated library systems. They serve as efficient help for the librarians in their job. The systems main functions are to keep records of documents and give bibliographic service

to the users. They do not store the full-text electronic format of documents in contrast with the full-text databases. Their main functions are the following:

1) Managing the electronical bibliography data of the printed documents in the MARC format of the country

- **a)** Entering new record in the database,
- **b)** deleting,
- **c)** modifying.

2) The users can reach the data:

- **a)** bibliographic records,
- **b)** data of the document's storage place.

3) Store the recording of the loan data.

The sketch diagram of a typical library with an integrated library system is the next:

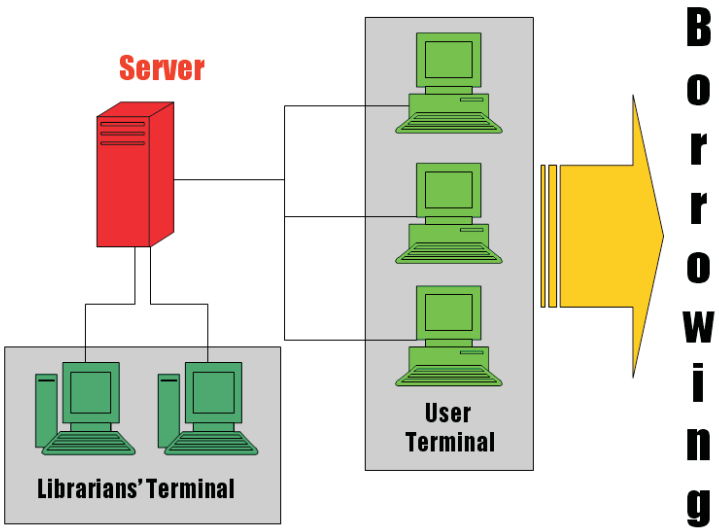

Figure 1: The present library integrated system

Whit this system the users can reach the bibliographic data of the documents and they can see if the document is available or not. With this information the user fills out the form which the librarians use to search for documents in the holdings, and give them to the reader.

2. Electronic Library

The situation is different in the electronic libraries. In these databases not only bibliographical data are stored, but the full text of the documents. Therefore, when the users find the document in the electronic catalogue, they can read the whole document.
The electronic library system is simplified, because there is no necessity for the librarian's mediatory function anymore. In these libraries the documents are not present physically, for that reason the users need not to walk in the libraries. In this case the availability must be provided via Internet. The sketch diagram of an electronic library is the following:

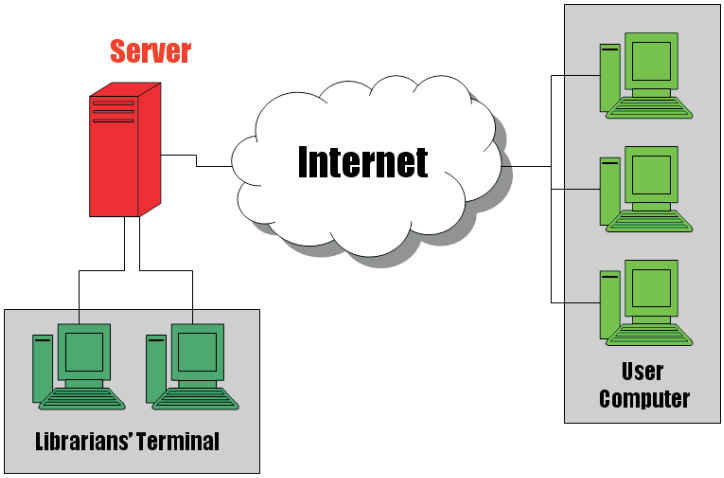

Figure 2: Electronic library

However, there is a big disadvantage to the text documents which are stored on a computer. Reading a long document from the monitor is very tiring for most of the users. The reason for this is that the human's eyes are easily tired when looking a CRT monitor for a long time. When we read a book our eyes blink 2425 times per minute, but when we look the monitor this number is reduced to 34 per minute. Hopefully this problem would be improved when the TFT monitors will be spread.
The other problem is that most of the users do not like to read Plato for example from the monitor, because it does not provide the same enjoyable experience as reading a book does. Naturally we can print the documents, but it is still a big expensive.

3. The Electronic Music Library

These problems do not occur when speaking about sound and motion picture mediums. My search topic is limited to sound, as an electronic medium. This is why I will introduce an integrated music library system, which will hopefully serve the users in the future. The system sketch diagram is similar to the previous diagram:

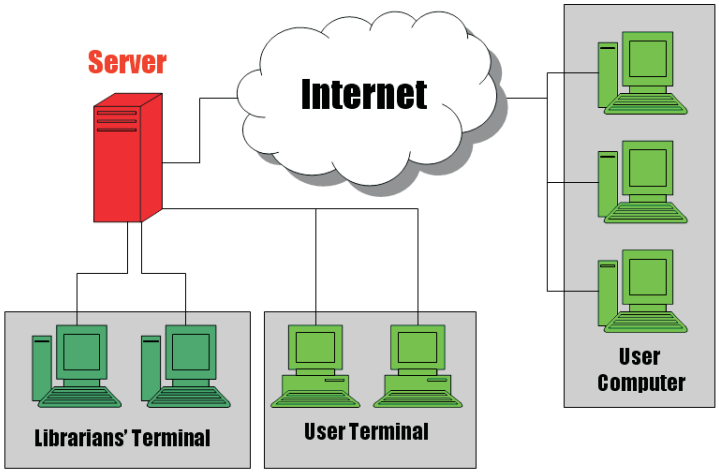

Figure 3: The practicable new music library system

The type of data which is stored basically determines the structure of the system. For this reason the first thing that we have to examine is the type of stored data.
The system could manage the records of digitalized music. One five minute music record reserves approximately 50MB size on the storage device, which is a considerable size even if we have large storage capacity. Therefore, these records should be compressed with compression programs into the MP3 format or the less-known VQF format.

4. Digitalization

First, we have to digitalize the sounds. The music consists of notes, and the notes consist of sound-waves. Therefore, music consists of waves which have to be digitalized. (see Fig. 4)

When we digitalize sound we have to take samples from it in well determined periods. This queue of samples unequivocally describes the series of original signs.

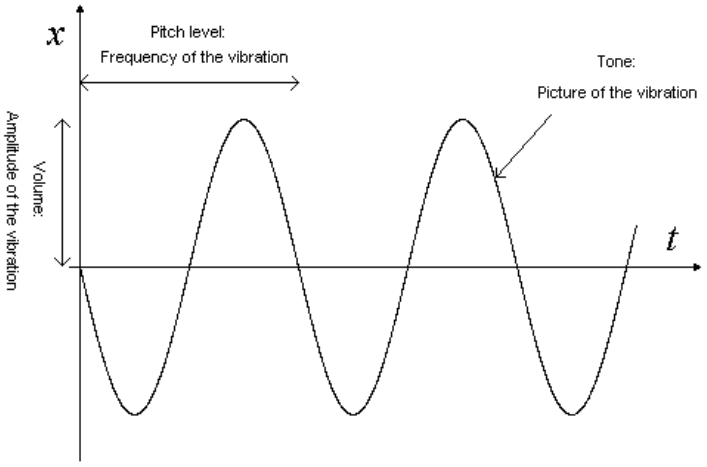

Figure 4: The sound

The sampling rate is as follows according to the formula of Shannon:

$$f_m \geq 2f_{max}$$

From this follows that different sounds need different sampling rate frequencies:

- human speech 11, 025 KHz-el,
- mediocre quality music 22, 05 KHzel,
- HIFI quality music 44, 1 KHzel

The other important factor, which we have to decide on is, how detailed should the stored amplitude samples be. This value – which shows the division of coordinate axis – is called quantification scale. The quantification scale must be given in bytes. (see Fig. 5) When we digitalize a HIFI quality sound, this value must be two bytes large, but if we increase this value, we can improve the quality level of stored sound.

If we stored the amplitudes in a file, we get a WAV format file. Its size is very large therefore it is practically unusable. This problem was solved by audio compress programs, which can create compressed audio formats. One of them is the most popular MP3 format.

The essence of MP3 compressor is that it takes advantage of the characteristic of human hearing – based on the psycho acoustic model –, and removes parts of sounds, which the human ear can not hear.

The human ear, do to its anatomical build can only hear sounds in the 20Hz-20kHz frequency domain. Those sounds which fall out of this frequency domain

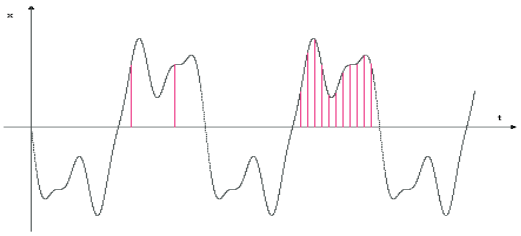

Figure 5: The digitalization

can be abandoned. We can also abort the sounds, which are filtered by the covering symptom. It is based on the fact that we can not hear a given pitch sound, if there is relatively close pitch sound near it, which is more intensive. (see Fig. 6)

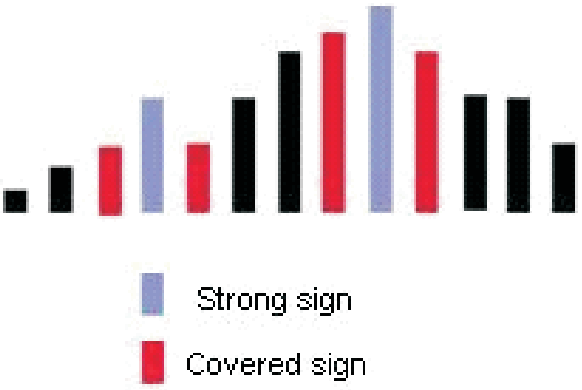

Figure 6: Covering symptom by frequency

The first diagram tries to demonstrate this symptom. The columns illustrate amplitudes of closely equal frequency signals. The purple columns show the strong signals, which cover the less intensive and relatively close – red columns –pitch sounds. The human ear can only hear the sounds marked in black and purple colour, so only these need to be stored. The covering symptom is a function of time also. When a loud sound ends some time passes before the covered sounds become audible again. (see Fig. 7) The information loss is quite substantial, however, do to the imperfection of the human ear, the recording does not deteriorate significantly in quality.

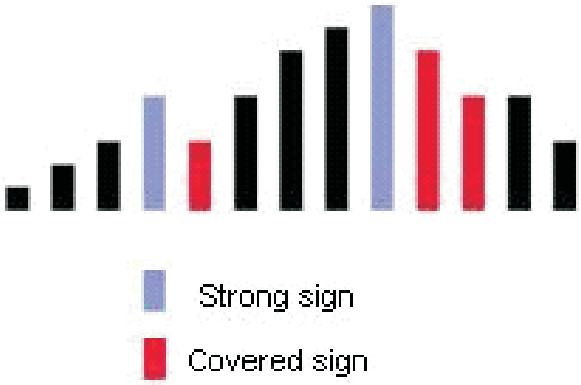

Figure 7: Covering symptom by time

This is not the only way we can reduce the size of a sound document. We can also reduce the depth of storage, which means the bite number. This procedure lowers the quality of the recording do to the quantization noise caused by rounding off the digits. During the quantization process sound samples are first rounded off –depending on storage depth– and then stored. Quantization noise is the distance between the original sound and its stored unit. For example if we wish to store a sound with an amplitude of $x = 249$ and the scale/quantization factor is $Q = 16$, then, a $y = x/Q = 15$ value is stored, which is just 4 bites in size. If we want to reset the amplitude, then we get an $x' = y \cdot Q = 240$ measure where the quantization noise is $x - x' = 9$. The size of the noise is statistically stable which means if the stored sample is played loud, its quality will be acceptably good, because in this case the useful sound oppresses the noise.

The core of compression is that the frequency range is subdivided into 32 frequency bands, which slightly overlap each other. These subbands are evenly divided according to the pitch level. Then the 32 subbands are each divided into 18 other subbands with the help of MDCT (Modified Discrete Cosine Transform). The importance of this division is that the 576 bands help in the precise distinction of relevant and irrelevant sounds, which accelerates the efficiency of the compression.

Simultaneously the sign goes through the psychoacoustic model, which separates the different covering tonal components from the noise like components. Then with the help of the hearing threshold and mutual covering, it determines the elements to be saved in the given bands and their level.

The bands are then separately strengthened and stored with distinct accuracy. To determine which band should be stored with which accuracy a complicated sound analytical algorithm is used, which takes into consideration the sensitivity of human hearing. It is unquestionable, that during compression the largest profit can be reached when the high sounds are coded with low accuracy.

Besides all these, the algorithm performs statistic compression with the help of the Huffman algorithm, which considerably enlarges the rate of the compression. The basis of statistical compression is that statistics are compiled based on the frequency of the appearance of the signs in a given section to be compressed. Based on the statistics bite series are ordered to the different signs, so that, the more frequent get a short and the less frequent are assigned a longer series. In the obtained encoded part the average word length will be decreased, but the statistic results must also be sent. What is interesting in this type of statistics, is that it is not the algorithm which compiles the statistics, it works with a set chart. Although this lowers the efficiency of the compression, it is not necessary to store a chart for each and every setting.

This form of compression is most effective when dealing with a recording of clear instrumental musical sounds, or where few instruments sound at once. The efficiency of the compression may decrease if "noisy" instruments are sounded such as cymbals or when the recording contains sound effects such like the roaring of the sea, or when many of the same type of instruments or voices hit the same note at the same time, or when not just the pure instrumental sounds are recorded, but for example their echo, which bounced back off the walls is also gathered.

The Layer-3 algorithm allows the problematic areas to be a signed (on a short term) a larger amount of the file, because in music the number of the simultaneously sounding instruments changes quickly. This notably decreases quality oscillation. This, which is known as variable bit quota, allows the compressor to work with a varied bit quota at certain time periods. The compression rate will be greater at the significant parts of music, which means the given part is stored on a smaller amount of space, and in order to keep the quality of the music good, the compression rate is accelerated only where the more demanding parts are.

It is undoubted that the compression of sound material causes the quality of the sound to decrease, but this deterioration is does not shift the recording below the level of enjoyment, but substantially decreases the amount of the recording to about $^{1}/_{10}$ or $^{1}/_{12}{}^{th}$ the size of the original. The other advantage to compression is that the amount of data which has to be processed through the network by the second is limited to 128 Kbits.

5. Advantage and disadvantage

Our system has to contain one or more computers that are capable of digitizing and compressing audio material. A better quality PC is capable to do so, if we supply it with a good quality sound card, and with the software needed for digitizing and compressing. Once the computer is connected to the player, digitizing may begin. The system could allow the digitization of the sound material to be processed during the day, and at night, when the librarians rest, the machines would convert the digitized collection into the MP3 format.

The collection prepared as such would then be transported to the descriptive librarians who would record the necessary data into the integrated system database

according to the MARC format based on the data of the original recording. The opportunity to scan in the cover of the original recording is also possible here at this phase. Collections in MP3 format have a very useful part called TAG.

This is the part where the most important things about the composer, the title and the genre. These are automatically added to the sound document by the MARC format.

The disclosed and prepared collection is then added to the database. From this point on they become available for the users. All yet needed is a terminal capable of playing MP3 collections. The users have many options on finding the music collection they want from the database. The system allows findings based on information on the composer, performer, contributors, or the title of album, the title of the composition, or according to its genre. This last option is helpful for those who are not looking for a concrete work of music, but would like to hear something from a specific genre.

The system would also provide the user with a 10 second introduction to the chosen work of art, this would help the users in their decision of which piece of music to choose. The users then choose the works they want and with the help of an MP3 player they can listen to the music.
The advantages of a system like this are:

- According to one of Ranganatan's rules, the librarian should economize with the user's time. This system eliminates the waiting time, because the user can immediately listen to the piece of music.
- As in regard to the care of the sound documents, the original piece of music is only played once at the time of digitalization. This is extremely important when we take in to consideration old material and the fact that it is not exposed to the danger of being damaged every time it would be played.
- The recorded material can be simultaneously listened to by any number of users, only a network connection is needed to the library's system.
- Statistics can automatically be compiled by the system about the system and the habits of the users.

One even more important advantage of this system is the opportunities it would provide once it is extended to the Internet. In this case the users would get a personalized library card with unique username and password. Using this specific card the users would have access to the system and its services from a remote computer with the help of a web surface.

Only legal protection would mean a problem for the system. The users should not be able to make a copy of the collection. This can be easily solved in the internal system, by using such terminals that ban the saving of elements. A bigger problem occurs when the collection is accessible via the internet, because in order to be able to listen to the audio material it has to be first downloaded. A possible solution would be if the documents would be available in a Real Audio format, which prohibits the saving of downloaded collections.

6. Discussion

As it can be seen from the information above, the music library of the future will be capable of supplying the user with the digital collection of the music they want. The system will be accessible from closed local network or via the Internet. It is also important that the music libraries not only depend on electronically stored collections, but they must place emphasis on continually gathering original sound documents also in order to be able to serve the public in any case.

References

[1] Horváth Tibor, Papp István: *Könyvtáosok kézikönyve 3.*, Osiris, Budapest 2001.
[2] Cedro, Jean-Michel: *Multimédia*, Alexandra, Pécs 1998.
[3] Tószegi Zsuzsanna: *Multimédia a könyvtárakban*, Akadémiai, Budapest 2001.
[4] Wischner, Stefan: Amit az MP3-ról tudni kell: könnye is lehet, Panem, Budapest 2001.
[5] Fraunhofer IIS: `http://www.iis.faunhofer.de`

Postal addresses

György Barna Iszály
Department of Computer Graphics and
Library and Information Science
University of Debrecen
Egyetem tér 1., 4032 Debrecen
Hungary