M a t h e m a t i c a
B a l k a n i c a

New Series Vol. 24, 2010, Fasc.3-4

Organization and Security of the Audio and Video Archive for Unique Bulgarian Bells

Nikolay G. Noev

The purpose of this investigation is to study and to identify the most valuable unique bells as well as to develop a digital archive with the help of advanced technologies. The main tasks are: development of the audio archive of information gathered from artifacts, analysis, optimization and addition of metadata for indexing of digital data, compression and data protection, prevention of data loss, design, organization and maintenance of archive. We investigate the methods of protection with watermarking which can be used against illegal use of data. We create and protect samples for additional applications and web sites.

AMS Subj. Classification: H.3.7 Digital Libraries, K.6.5 Security and Protection

Key Words: digital archives, indexing of digital data, data protection

1. Introduction

The key to prosperity in today's world is access to digital content and skills to create new content. Digitization of analog materials and the creation of digital resources in the field of cultural heritage is a major contributor to e-Europe. The aim of this article is to study and identify several dozens of the most valuable bells in Bulgarian churches, monasteries and museums as well as to develop an audio, photo and video digital archive (with the help of advanced technologies) for analysis, preservation and protection of the data. Main tasks in the research are:

- Development of audio, photo and video archive with information collected by artifacts;
- Analysis and indexing the digital data;
- Design and maintenance of digital archive;
- Compression and optimization of the archive;
- Preventing data loss;
- Development of software for adding watermark against illegal use of data;
- Development of functions for creating, formatting and protection of samples for additional applications and web-sites.

In the second chapter we consider modern methods for digitization and data collection. In the third chapter we provide an analysis of the objects in order to determine the metadata of selected artifacts from selected collections and problem areas. Fourth chapter contains research of advanced technologies and methods for protecting intellectual property. In the fourth chapter has examined the organization of the digital archive with unique Bulgarian bells.

2. **Digitization and data collection**

Modern technologies have changed the way the information was presented in the archives and have made possible new services, unthinkable a few years ago.

Digitization is creation of an object, image, audio, document, or a signal (usually an analog signal) from a discrete set of its points or samples. The result is called "digital image", or more specifically "digital images" of the object and "digital form of the signal". The tasks of digitalization can be synthesized in certain key areas:

- Retention of funds and records - many of digitalized objects are fragile or brittle structure is influenced by weather conditions and over time their digitization is the only hope for preservation;
- Simultaneous access to materials - most objects are subject to the digitization of rare and unique items of historical past and have a priceless value, the process of digitization will allow more users involved to touch them;
- Conservation funds in digital formats - archives, websites, digital libraries. Strengthening international exchange and promotion;
- Providing access via computer networks - easy access to digital archives, access to records of persons with disabilities;
- Providing new opportunities to work with digitalized materials funds - all the functionality is available to users of web space to be copied, multiplied, forwarded, etc., without jeopardizing the its integrity and strength;
- Full text search - digital archives organization contributes to easier detection of the searched object among all the crowd, advanced search, unification of search results;
- Classification of digital funds via metadata - entire photo meta data wealth funds may be accompanied by important information about copyright, creation date, identification number, etc.

Development of digital technology hit the storage and processing of information. Today, almost every unit of information is created digitally: digital photography, digital sound recording, digital communications, text, saved in a file, videos, movies, multimedia presentations stored on digital media, etc.

Storage of such digital multimedia data in digital archives became subject to the same challenges, such as archives, we know before the invention of digital computing devices.

3. **Analysis, optimization and indexing the digital data**

3.1. **Essence of metadata.** Metadata are text fields, built-in media files or additional text files (XML, XMP) for recording information on the nature of the digital resource. In other words, meta data is "data about data" describing any electronic or non-electronic source of information on pre-established standard. Meta data contribute to finding and sharing information. In other words, this is the last stage of information management. Once digitized (implemented in electronic form) and structured (arranged in a specific sequence and line), information is visualized in an appropriate form. To be fully justified the efforts of both phases scheduled for managing the information it needs to be found and used by as many users. This is possible thanks to the meta data [3]. The presence of meta data with correctly placed points of connection ensures speed and accuracy of the application, and interactive user interaction. The shared experience of developers and users for the metabase warrants defining them as a new generation of data, integrating new technologies and therefore requires a new legal and economic regulation.

3.2. **Analysis of the objects in order to determine the metadata of selected artifacts collections and problem areas.** Following the studies and consultations with specialists, it was concluded the following organization of metadata:

- Title (name of subject);
- Creator (name of digitalizer);
- Description (additional data);
- Date (date of creation);
- Type (type of media);
- Format (file format, codec and parameters);
- Identifier (geographic coordinates);
- Rights (owner of property rights).

3.3. **Adding text metadata fields.** The process of creating the meta data to digitize files could be done by several basic approaches:

- Using the software that serves as a system for collecting and managing the meta data for individual sites in the process of digitization. The system can be either web-oriented and desktop application depending on team needs for digitization. Data can be stored in a database and easily accessible for further processing;

- Using the application to scan individual files and embed and extract from them the necessary meta data. For this purpose, one may use different standards for meta data and software tools for working with them. An example is the standard Adobe XMP and means of work with it Adobe XMP ToolKit;
- Presentation of metadata in a text editor, subject to certain rules and the standard XMP templates and languages used to define meta fields. The name of the file should completely coincide with the name of the digital resource, but with a different extension;
- Using embedded resources to work with meta data used to support the digitization software or operating systems. They keep the information that is specified in the study as necessary meta data in this digitization project.

There are three ways to view the meta data:

- Main menu interface: For example, from MS Word document a user can choose File from the menu option Properties, then on General, Summary, Statistics, Contents, Custom to find the main categories of meta data embedded document;
- Computer mouse: By positioning the mouse pointer on the version in MS Word document, which has enabled feature Track Changes, a dialog box informs who made and when the change was made;
- Software for viewing metadata. Windows Vista and Windows 7 operating system has built in tools for meta data. Thus facilitate the process of organizing files with tags "metadata" which indicate that the information belongs to a project or category.

FotoWare FotoStation [4] application language uses XML technology to define META text boxes. Definitions can be saved by RDF language for ontologies. META text fields are added to the digital archive unit as follows:

- Indicate the object(s) for META data adding;
- The Metadata button is selected from the toolbar or right-click the selected object is selected, then the menu is selected metadata edit, and which is made by pressing keys (Ctrl + T);
- From the window to edit the META fields cause textual data.

4. **Modern techniques for intellectual property protection**

With the development of digital technologies increasing part of the audio, video and any other information is available for fast, easy and high quality copying. This fact entails the problem of protecting information from unauthorized distribution. Research in this area is considered in several aspects. One of the most important of these is steganography [5], [6]. Steganography deals

with the concealment of information, steganography hides the message which should remain hidden. Like steganography, watermark protection aims carrying hidden information. However, there are significant differences between the two techniques. Digital watermark is visible or preferably invisible to the identification code that is permanently embedded into digital data and maintained a presence in them after extracting it [2].

4.1. **Methods for image watermarking in the spatial region.** In these methods data are incorporated directly into the original image. The main advantage is that the key is not necessary to do any preconditions transformations. The watermark is embedded by changing the illumination or color components. The main disadvantage is the low resistance. An example of this method is the method of Kutter [7]. To derive the value of the embedded bit is calculated assumed difference between value and actual value of the pixels. The sign of the difference determines the value of the embedded bit. Extracting bits is done without the knowledge of the original message. The method is robust to filtering, JPEG compression and geometric transformations.

4.2. **Methods for audio watermarking using low-bit coding.** By replacing the least significant bit of each sampling point by a coded binary string, we can encode a large amount of data in an audio signal The major disadvantage of this method is its poor immunity to manipulation. Encoded information can be destroyed by channel noise, resembling etc. We improve robustness using by error-correcting codes.

4.3. **Visual watermarking.** Visible mark added in digital picture and video records.

5. Approaches and tools for building digital archive of unique Bulgarian bells

In our research we also consider some previous experience for creation of digital archives [1], [8].

We use the fooling software tools in order to digitize and secure objects:

- Photo image processing - Adobe Photoshop, Fotoware Fotostation;
- Sound processing - specialized acoustic software from Bruel and Kjer;
- Text include - MS word;
- Video - Adobe Premiere, Virtual Dub;
- Watermarking - our own software;
- Media file organizer - Fotoware Fotostation, Fotoware Camelion.

We also use specialized hardware tools as follows:

- Photo-camera - Sony Alpha DSLR-A100;
- Video-camera - Sony HDR-SR8E HD AVCHD Camcorder;

FIGURE 1. Visual watermarking example.

- Audio-system - PULSE11 from company Bruel and Kjer.

Archive creation and organization is performed using Fotoware Fotostation, Camelion. This software package has following main characteristics:

- Media file organizer;
- Automating repetitive tasks - actions;
- Share files via network or upload to internet;
- Create web pages from workflows;
- User-friendly interface that is available in 10 different languages;
- Sophisticated full text search;
- Import from digital cameras, scanners, DVD, CD or any other source;
- Storage capacity of up to 200 000 files.

Organization of the archive can be separated in following steps:

- Creating an archive;
- Creating own functions - actions, macros;
- Adding objects to a digital archive;
- Converting file format, codec, size;
- Adding META data to objects in a digital media collection;
- Adding watermarking - visible and invisible.

Acknowledgement

For this work I want to thank Galina Bogdanova and Todor Todorov for collaboration and colleagues connected with the project "BELL - Research and Identification of Valuable Bells of the Historic and Culture Heritage of Bulgaria and Development of Audio and Video".

References

[1] G. Bogdanova, R. Pavlov, G. Todorov, V. Mateeva,, Technologies for Creation of Digital Presentation and Significant Repositories of Folklore Heritage. *Advances in Bulgarian Science Knowledge, National Centre for Information and Documentation*, **3**, 2006,7–15.

[2] I.Cox, J. Kilian, T. Leighton, G. Shamoon, Secure spread spectrum watermarking for multimedia. In *Proceedings of the IEEE International Conference on Image Processing*, **6**, 1997.

[3] Dublin Core Metadata Initiative - *http://dublincore.org/groups/education.*

[4] Fotoware FotoStation - *http://www.fotoware.com/en/Products/FotoStation/.*

[5] G.Gribunin, I., Okov, I. Turincev, *Cifrovaia steganographia.* Solon-Press, 2002.

[6] A. Karasev, Komputernaia tainopis, grafika i zvuk priobretaut podtekst. *Mir PK.* **1/97**, 1997, 132-134.

[7] M. Kutter, Digital Signature of Color Images using Amplitude Modulation. *Journal of Electronic Imaging*, 1998, 326–332.

[8] D. Paneva, K. Rangochev, D. Luchev, Knowledge Technologies for Description of the Semantics of the Bulgarian Folklore Heritage. In *Proceedings of the Fifth International Conference Information Research and Applications* i.Tech, Varna, Bulgaria, **1**, 2007, 19–25.

Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Acad. G. Bonchev str., block 8,
1113 Sofia, Bulgaria

E-Mail:nickey.noev@gmail.com