

Audiovisual Archive Exploitation in the Networked Information Society

Roeland Ordelman
Netherlands Institute for Sound and Vision & University of Twente
Hilversum, The Netherlands
rordelman@beeldengeluid.nl

ABSTRACT

Safeguarding the massive body of audiovisual content, including rich music collections, in audiovisual archives and enabling access for various types of user groups is a prerequisite for unlocking the social-economic value of these collections. Data quantities and the need for specific content descriptors however, force archives to re-evaluate their annotation strategies and access models, and incorporate technology in the archival workflow. It is argued that this can only be successfully done provided that user requirements are studied well and that new approaches are introduced in a well-balanced manner, fitting in with traditional archival perspectives, and by bringing the archivist in the technology loop by means of education and by deploying hybrid work-flows for technology aided annotation.

Categories and Subject Descriptors

H.3 [Information Systems]: Information Storage and Retrieval; D.3.7 [Digital Libraries]: [Systems issues, Collection, User issues]

General Terms

Management, Performance, Reliability

Keywords

Audiovisual content annotation; annotation strategies

1. INTRODUCTION

There is common agreement that audiovisual collections within archives throughout the world have large social and economical potential. One of the largest audiovisual archives in Europe, The Netherlands Institute for Sound and Vision (NISV) comprises over 750.000 hours of radio, television, documentaries, films and music, over 2 million photographs, 20.000 objects like cameras, televisions, radios, costumes and pieces of scenery. The Dutch National Music Depot (294K hours) is the largest music collection and originates from the thirties in which the Dutch Public Radio started a record library. Since then the collection developed into a massive body of gramophone records, LPs, singles, CDs, DvDs and Blu-rays, and is still being supplemented with new releases from record companies. The collection has a

significant number of Dutch popular songs from the first half of the twentieth century. Another collection is the Music Recordings from Broadcast License-holders (85K hours), an exceptional collection with concert and studio registrations that have either or not been broadcast. The cultural-historical value of the collection is significant as it reflects the Dutch musical life in all possible genres. Traditionally, NISV is also in charge of the collection Sounds and Background music (4466 hours), historical recordings that go back to 1923, generally for the purpose of radio plays. Finally, NISV has an ethnomusicology collection (1650 hours) consisting of for instance field recordings and concert registrations from the former Dutch Colonies including copies of wax cylinder recordings. Safeguarding this massive and rich body by means of restoration, conservation and digitization is a prerequisite for unlocking its social-economic value.

2. ANNOTATION STRATEGIES

A key component for the ‘business model’ that allows the significant investment in restoration, conservation and digitization, is enabling *access* to these collections, *geared towards specific user groups*. The traditional approach towards bridging the semantic gap between low level descriptors of audiovisual content and access needs of users, has been the manual annotation by professional archivists. It is evident however that the data quantities we are dealing with limit the scope of the manual approach. The needs for specific content descriptors, needs that vary depending on the access requirements of different user groups in our networked information society, are an additional reason that forces archives to re-evaluate their annotation strategies and access models with respect to audiovisual content.

To support this re-evaluation process, it is crucial to develop ‘hands-on’ expertise with respect to new approaches in (semi) automatic content annotation and access—for instance via research projects and collaborations with research institutes—and to understand its opportunities and implications. In parallel, the needs of potential user groups with respect to accessing these data need to be mapped, not only as it allows to select *which* data need annotations the most, but also to be able to make decisions on *annotation levels* and *automation*. For example, as NISV is the business archive of the Dutch Public Broadcasters, broadcast professionals—typically searching the archive for material for reuse—are an important stakeholder. Analysis of the NISV transaction logs revealed that this user group is hardly interested in radio material for re-use. Nonetheless, archivists spend valuable time on its annotation. As research pilots

showed that automatic speech recognition provides reasonably accurate representations of the spoken word in news and actualities type of radio content, re-allocating sparse manual annotation capacity from radio to another type of content seems an obvious thing to do.

3. TECHNOLOGY IN THE ARCHIVE

One can argue that deploying audiovisual analysis and access technology and crowd-sourcing strategies inevitably needs to be part of modern-age audiovisual archives. However, its implementation can only be successful provided that user requirements are studied well (technology-pull) and that new approaches are introduced in a well-balanced manner, first and foremost fitting in with the traditional archival perspective (soft-launch), to progressively develop into a significant support for sustainable audiovisual access (consolidation).

One of the key issues with respect to the use of automatic annotation or crowd-sourcing is the uncertainty about the correctness of automatically generated labels that may interfere with the principals of reliability and integrity an archive adheres to. For certain collection this uncertainty may be acceptable. As the chances are extremely low that content with only marginal amounts of descriptive metadata will ever pop-up in search requests, deploying automatic annotation tools without manual supervision has much to recommend: it can't do any harm and it may do some good. A complementary approach is to start with selecting wisely those collections that are suitable for automatic annotation given both the collection characteristics and the requirements of the end-user. For instance, given the well understood domain dependencies of technologies such as speech

recognition or visual analysis, a critical selection of data to process makes sense. At the same time, user requirements such as fast access to recent broadcasts may put less constraints on accuracy levels.

4. MEDIA MANAGER

Eventually, in the archival work-flow, it is the archivist—and not the technology expert—who will decide whether or not to deploy a certain technology for annotation. Proper education and training of archivists becomes therefor a crucial element for the successful application of the technology. Moreover, by connecting multimedia technology with the human expertise of the professional archivist, being made concrete currently at NISV within a so called 'Archivist Support System', some of the restraints with respect to the use of automatic annotation in an archival environment where authority plays such an important role, can, at least partly, be overcome. Clearly, the traditional archivist in the scenario sketched here, is moving away from traditional roles and starts to become more of a *media manager*.

In the talk, the intertwining fields of access technology and audiovisual archiving will be discussed on the basis of practical examples from the NISV practice with respect to user studies, automatic annotation by means of speech recognition technology and visual concept detection, crowd-sourcing, and the Archivist Support System.

5. ACKNOWLEDGMENTS

This work has been funded by the EU Project FP7 AXES (ICT-269980) and Dutch government funded project COMMIT Infiniti.