

Facilitating Access to Large Digital Oral History Archives through Informedia Technologies

Michael G. Christel
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213 USA
+1 412 268 7799
christel@cs.cmu.edu

Julieanna Richardson
Executive Director, *The HistoryMakers*
1900 South Michigan Avenue
Chicago, IL 60616 USA
+1 312 674 1900
jlr@thehistorymakers.com

Howard D. Wactlar
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213 USA
+1 412 268 2571
wactlar@cs.cmu.edu

ABSTRACT

This paper discusses the application of speech alignment, image processing, and language understanding technologies to build efficient interfaces into large digital oral history archives, as exemplified by a thousand hour HistoryMakers corpus. Browsing, querying, and navigation features are discussed.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems. H.3.1, H.3.7 [Information Storage and Retrieval]: Content Analysis and Indexing, Digital Libraries.

General Terms: Design, Human Factors.

Keywords: Digital video library, oral histories, video browsing.

1. INTRODUCTION

Satisfying users with efficient, effective access to large oral history archives presents a number of challenges [1]. Contextual inquiry has found that while the interview recordings are considered a central historical artifact, providing a level of emotion and humanity not available in text transcripts, often “these recordings sit unused after a written transcript is produced” [2]. Collaboration between *The HistoryMakers* and the Informedia research group made use of speech alignment, image processing, and language understanding technologies to promote multiple levels of access and fuel the viewing of the actual video recordings in a large oral history corpus. The specifics of the technologies are overviewed elsewhere with respect to cultural archives [3]; this paper focuses on some accessibility issues.

The HistoryMakers is the world’s largest African American oral history archive of thousands of video interviews with accomplished African Americans across a variety of disciplines, including those who have played a role in African American led movements and organizations. The purpose of *The HistoryMakers* is to educate and to show the breadth and depth of this important American history as told by the first person, to highlight the accomplishments of individual African Americans across a variety of disciplines, and to preserve this material for years and generations to come. *The*

HistoryMakers is committed to creating and exposing its archival collection to the widest audience possible, making use of new technologies as appropriate.

2. SEARCH AND NAVIGATION

HistoryMakers archivists provided a transcript for each of the interviews fed into Informedia processing, identifying 18254 interview story segments across 400 interviewees. Automatic speech alignment time-tagged each transcript word to video time, enabling quick access to the video sections of interest. Fast, direct video access to oral histories was found to be critical in other studies as well [2]. The synchronized metadata allows results for a text query like “war protest Nixon” to be presented with the terms color-coded in the matching transcripts and color-coded time markers to be shown on video timelines. The user can click on a marker and seek the video immediately to that point, e.g., skipping over the first 2:30 of the clip shown in Figure 1 to view the last 30 seconds where “Nixon” gets discussed.

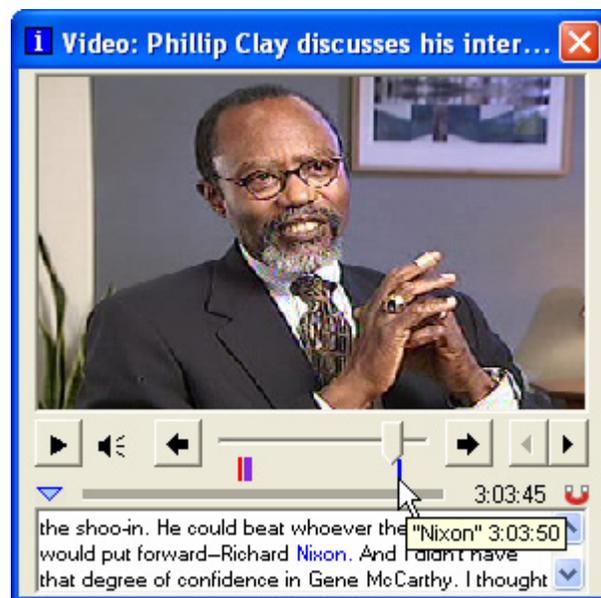


Figure 1. Time-aligned text for quick video access.

Retrieval as shown in Figure 1 is done using an inverted index of the entire transcripts of all interviews, and each indexed story segment is ranked by relevance based on the term frequency and inverse document frequency of the search words occurring within the transcript. With a reasonably sized library, most searches return hundreds, thousands, or more segments, necessitating a way for the user to navigate through the returned set, as well as a means to browse the corpus as a whole without having to resort to a text

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to publish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

JCDL'06, June 11–15, 2006, Chapel Hill, North Carolina, USA.
Copyright 2006 ACM 1-59593-354-9/06/0006...\$5.00.

query. Automated named entity extraction against the transcript identified people, organizations, time references, and places. Human annotators categorized the oral histories with additional date reference tags and subject headings drawn from Brown’s hierarchical taxonomy of terms for African American materials [4]. These sources of metadata enabled rich exploration interfaces supporting navigation and browsing of the videos.

3. FACILITATING BROWSING

Along with text search, a set of video segments can be returned from a color query, map search, or browsing action. Each set can be inspected through a number of views, including thumbnails, timeline scatter plot emphasizing time, visualization by example (VIBE) plot emphasizing query terms, map emphasizing location, common phrases, and named entity diagrams. Figure 2 shows a few of these views open for the 1381 segments returned by the “war protest Nixon” query. In practice, the user would check perhaps a few at once: here four views are opened at once for illustration.

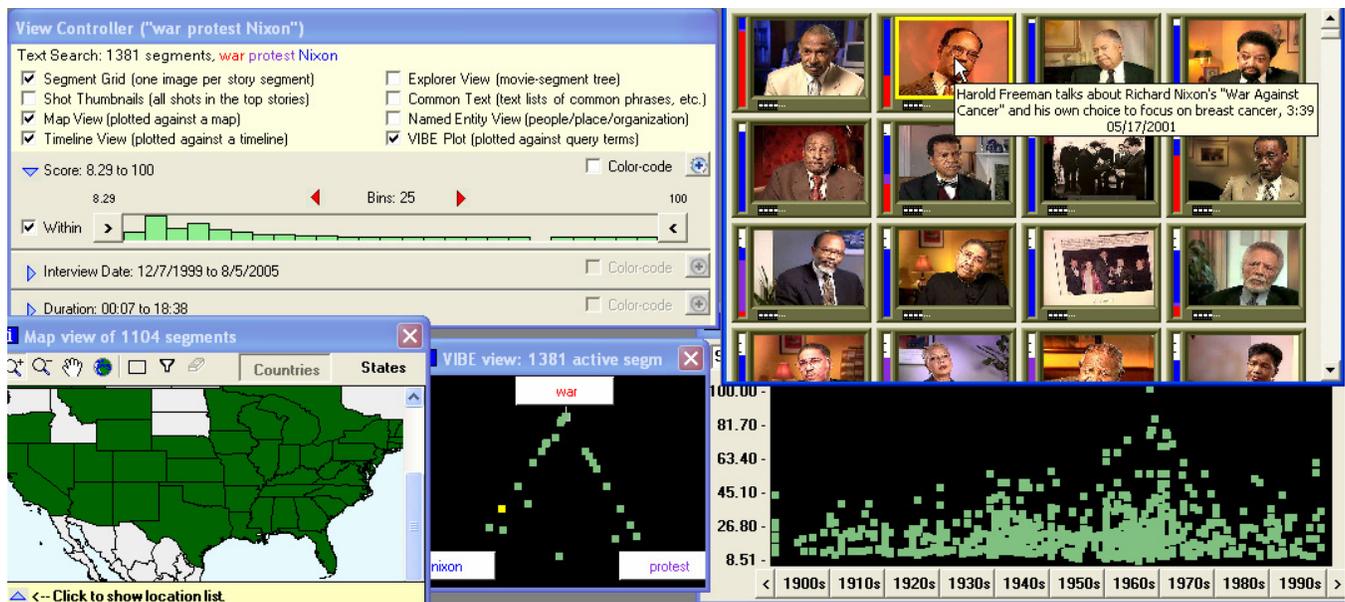


Figure 2. Multiple browsing views into oral histories, based on text, time, location, named entities, and synchronization metadata.

The same metadata that underlies these views can be used to support dynamic query previews [5]. The user can explore the oral history archive through histogram breakdowns of the whole set of 18254 segments to see how the stories map to different geographic breakdowns, across the decades, according to Brown’s subject headings, and according to 16 HistoryMakers general categories like “LawMakers”. For example, through previews the user can see that there are 1504 MusicMakers stories, 2116 stories with references to the 1960s, and 218 MusicMakers stories referring to the 1960s. Without issuing a text query, the user can manipulate query previews based on these types of partitions to produce a result set, like the 218 stories, which can then be loaded and further viewed with windows as illustrated in part in Figure 2.

The features discussed here expand the utility of oral history archives like HistoryMakers by exploiting their digital nature to deliver synchronized video quickly and provide interactive visualizations dynamically. These interfaces enable multiple investigative paths into the archive for witnessing, appreciating, and reflecting on prior historical times and experiences. This work was partially funded by IMLS Grant LG-03-03-0048-03.

Even visual processing serves a purpose, albeit slight in this genre: a view of shot thumbnails filtered down to just those automatically tagged as showing two or more people lets the user navigate to the handful of shots showing photographs of groups, drawn from the thousands of dominating head and shoulder shots.

The browsing views all interact with an underlying XML representation of the segment set. The user can affect all the views by interacting with controls in the interface, e.g., by dragging the dynamic query slider for relevance score to only keep segments scoring greater than 50, the result set of Figure 2 drops from 1381 segments to 59, with the map view for the filtered set showing only California remaining from the U.S. West, with all time references dropping out from 1900-1940, and with no stories remaining dealing with “Nixon” and “protest” but not “war”. Even without further interaction to navigate, the views as illustrated in Figure 2 communicate characteristics of the result set, e.g., the lack of protest stories with Idaho or Dakota references, and the dominance of “war” in the VIBE plot.

4. REFERENCES

- [1] Gustman, S., et al. Supporting Access to Large Digital Oral History Archives. In *Proc. JCDL* (Portland, 2002), 18-27.
- [2] Klemmer, S.R., et al. Books with Voices: Paper Transcripts as a Tangible Interface to Oral Histories. In *Proc. CHI 2003* (Ft. Lauderdale, FL, April 2003), 89-96.
- [3] Wactlar, H., and Chen, C. Enhanced Perspectives for Historical and Cultural Documentaries using Informedia Technologies. In *Proc. JCDL* (Portland, OR, 2002), 338-339.
- [4] Brown, L. B. *Subject Headings for African-American Materials*. Libraries Unlimited, Englewood, NJ, 1995.
- [5] Jones S. Dynamic Query Result Previews for a Digital Library. In *Proc. ACM DL* (Pittsburgh, PA, 1998), 291-292.