EigenNews: A Personalized News Video Delivery Platform

Matt C. Yu, Péter Vajda, David M. Chen, Sam S. Tsai, Maryam Daneshi
André F. Araujo, Huizhong Chen, and Bernd Girod
Department of Electrical Engineering, Stanford University, Stanford, CA 94305, U.S.A.
{mattcyu, pvajda, dmchen, sstsai, mdaneshi, afaraujo, hchen2, bgirod}@stanford.edu

ABSTRACT

We demonstrate EigenNews, a personalized television news system. Upon visiting the EigenNews website, a user is shown a variety of news videos which have been automatically selected based on her individual preferences. These videos are extracted from 16 continually recorded television programs using a multimodal segmentation algorithm. Relevant metadata for each video are generated by linking videos to online news articles. Selected news videos can be watched in three different layouts and on various devices.

Categories and Subject Descriptors

H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems

Keywords

story segmentation; story linking; personalization; news aggregation; mobile

1. INTRODUCTION

EigenNews is a personalized video news system which aggregates multiple sources of television news, automatically selects relevant videos based on a user's interests, and offers the user flexibility in how to watch the videos.

The features of this system can be summarized as follows:

- *Personalized Content*: Different stories are selected for each user based on her preferences.
- *Aggregation of Sources*: Different programs are chosen to meet a user's diverse interests.
- *Universal Access*: EigenNews is accessible on several devices including computers, smartphones, and tablets.

In this technical demo, we present EigenNews which applies story segmentation, linking, and personalization in a live, real-time system. Please see our demonstration video[1] for additional information and also visit the EigenNews website[2].

---

[1] http://blackhole1.stanford.edu/EigenNews.mov
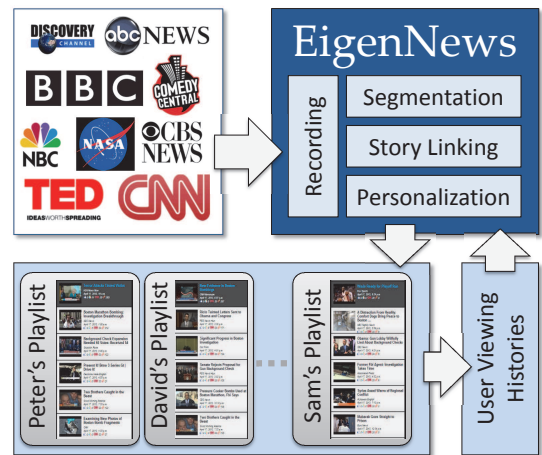[2] http://eigennews.stanford.edu

Figure 1: The EigenNews system architecture.

2. SYSTEM ARCHITECTURE

An overview of the system is shown in Fig. 1. We use several PCs and TV tuners to continually record television news programs from over-the-air (OTA) broadcasts and cable channels. These programs include national morning and evening news (e.g., *TODAY Show*, *ABC World News*), investigative journalism (e.g., *60 Minutes*), and late-night talk shows (e.g., *The Tonight Show*). We record the closed captioning (CC) along with each program.

We extract individual news segments from each recorded program by considering CC, audio, and visual cues [2]. CC cues such as story boundary markers as well as transition phrases are strong indicators of a change in topic. However, using the CC alone produces unreliable results due to the variable delay between the CC and the video stream. Thus, we perform automatic speech recognition (ASR) on the audio track using the CMUSphinx toolkit [4] to generate time-stamps for ASR-detected words. Matching these ASR-detected words with CC words via dynamic time warping provides accurate time-stamps for the CC and mitigates the undesirable delay. In addition to analyzing the CC and audio, keyframes are extracted from the video once per second to detect visual indicators of story boundaries. One important visual indicator is the presence of a program logo in a keyframe. To detect a logo, we utilize the image matching algorithm developed by Chen et al. [1]. Story boundary indicators in each modality are combined to perform segmentation.

Figure 2: Three ways to view EigenNews: *Playlist* (top), *Word Cloud* (bottom left), and *Grid* (bottom right).

Segmented videos are then matched with news articles from online sources (e.g., cnn.com, nytimes.com) to obtain descriptive titles, thumbnails, and categories. Several different comparisons are performed between videos and articles. First, the CC of a segmented video is compared with the text of the online news article using bag-of-words histograms. A match is declared if the *tf-idf* histogram intersection score exceeds a threshold. Second, named entities are extracted using the Stanford Named Entity Recognizer [3] and Open-Calais. A video and news article are paired if the video's CC contains more than 5 named entities in common with the article's text. Third, keyframes of videos are compared with the images of articles using the same image matching algorithm used for story segmentation. Note that these comparisons can also be performed between segmented videos to find related or duplicate videos. Once a video is matched with a news article, the titles, thumbnails, and categories of the article are automatically associated with the video.

Finally, a personalized playlist is generated for each user. Since each user has a limited amount of time to watch news videos, it is desirable for chosen videos to span the major news stories of the day. We cast this scenario as a maximum coverage problem where elements are the segmented videos and segmented videos about the same story belong to the same set [2]. Furthermore, news videos in the objective function of the maximum coverage problem are weighted by a linear combination of several personalization factors. These factors include a user's settings page, where a user indicates preferences for specific news programs or news categories, and the number of related videos a user has watched. Additionally, more recent videos receive higher weighting. Intuitively, this formulation chooses videos of trending stories, which originated from news programs the user prefers, and are also associated with categories the user is interested in.

3. USER INTERFACE

We also demonstrate three different user interfaces that we have developed and that are shown in Fig. 2. These three interfaces, *Playlist*, *Grid*, and *Word Cloud*, cater to different tastes that the user may have towards exploring the day's most relevant stories. The playlist view shows all the videos as rearrangeable tiles, grid view plays all videos simultaneously, and words in the word cloud indicate how prominent a word is in the day's news corpus.

In addition to the standard website, which is compatible across major browsers (Chrome, Firefox, Safari, Internet Explorer), we also developed a mobile version of EigenNews which works on devices such as the iPhone and iPad. The mobile version includes the same original viewing capabilities but can also extend to be a secondary device when paired with another display device. This includes allowing a user to treat the mobile device as a remote control (play, pause, choose next story) or to view additional information about the story they are watching.

4. IMPLEMENTATION

The EigenNews system has been recording content since July 2012. Statistics about the recordings are shown in Table 1. A 30 minute recording is processed in about 15 minutes where story linking and transcoding are performed in parallel, as shown in Fig. 3. Transcoding segmented videos to MP4 and OGV and to multiple bitrates is necessary to retain compatibility with our HTML5 video player and to accommodate users with different bandwidths, respectively. Subsequently, links to segmented videos and their associated metadata are stored in a MySQL database.

Table 1: EigenNews System

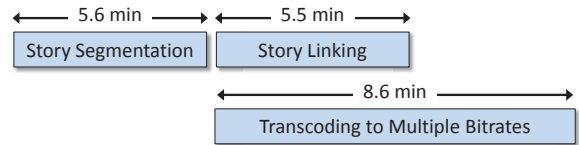|	
Number of different news programs	42
Number of news categories	28
Number of news channels	16 (4 OTA 12 cable)
Number of segmented videos (per day/total)	355/20,000+
Total time of recordings per day	16 hours
Storage used (per day/total)	40GB/2TB+

Figure 3: Breakdown of the video processing time.

5. REFERENCES

[1] D. Chen, P. Vajda, S. Tsai, M. Daneshi, M. Yu, A. Araujo, H. Chen, and B. Girod. Analysis of Visual Similarity in News Videos with Robust and Memory-Efficient Image Retrieval. *ICME MMIX*, July 2013.

[2] M. Daneshi, P. Vajda, D. Chen, S. Tsai, M. Yu, A. Araujo, H. Chen, and B. Girod. EigenNews: Generating and Delivering Personalized News Videos. *IEEE BRUREC*, July 2013.

[3] J. Finkel, T. Grenager, and G. Manning. Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. *ACL*, pages 363–370, 2005.

[4] W. Walker, P. Lamere, P. P. Kwok, B. Raj, R. Singh, E. Gouvea, P. Wolf, and J. Woelfel. Sphinx-4: a flexible open source framework for speech recognition. Technical report, 2004.