IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Hershey PA 17033-1117, USA ITB7306
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

Chapter V

News On Demand

Mark T. Maybury
The MITRE Corporation, USA

Recently scientists have been focusing on a new class of application that promises on-demand access to multimedia information such as radio and broadcast news. This chapter describes how the synergy of speech, language and image processing has enabled a new class of information on demand news systems. We will describe the ability to automatically process broadcast video 7x24 and serve this to the general public in individually tailored personalcasts. We describe several systems and identify some remaining challenging research areas.

MULTIMEDIA INFORMATION ON DEMAND

Information on demand, the ability to provide information tailored to specific user needs, promises new capabilities for research, education and training, and electronic commerce (e.g., on-line information access, question answering and customer service). Whereas significant commercial activity has focused on providing access to documents, Web pages and structured data sources, less attention has been given to multimedia information on demand. To achieve effective multimedia information on demand, however, requires a confluence of capabilities from several fields including image, speech and language processing, information retrieval, information extraction, translation, summarization and presentation design.

Over the past several years, scientists have been exploring a range of systems to provide tailored, content-based access to multimedia including text, imagery, audio and video (Maybury, 1993). We describe two systems to illustrate this new functionality: BBN's Rough 'n Ready and MITRE's Broadcast News Navigator (BNN).

ROUGH 'N READY

BBN's Rough 'n Ready (Kubala et al., 2000) uses a 60,000-word automatic speech transcription system to enable real-time monitoring of broadcast news with approximately

This chapter appears in the book, Multimedia Networking: Technology, Management and Applications by Syed Mahbubur Rahman.

18.8% word error rates for broadcast news sources. As Figure 1 illustrates, Rough 'n Ready provides a "rough" translation of content that is "ready" for browsing. The left-most column in Figure 1 illustrates the system's ability to detect a particular speaker (e.g., Elizabeth Vargas) or at least a speaker change and their gender. The center column shows the transcribed words, with color highlighting of named entities (people, locations, organizations) detected by natural language processing. Finally, using topic detection and tracking software, the third column lists keywords indicative of the subject of the associated text segment. Using the system, a user can either browse or search news for relevant stories. Topic Detection and Tracking (TDT) has become a common task with progress measured by community-based evaluation across different tools (Allan et al., 1998; Wayne, 1998).

BROADCAST NEWS NAVIGATOR (BNN)

MITRE's Broadcast News Navigator (BNN) (Maybury et al., 1997) goes beyond spoken language processing and synergistically combines techniques for processing speech, language and imagery, providing more sophisticated presentation of and access to news. The Web-based BNN gives the user the ability to browse, query (using free text or named entities) and view stories or their multimedia summaries. For example, Figure 2 displays all stories about Princess Diana on CNN Prime News during August-September 1997. For each story, the user is given the ability to view its closed-caption text, named entities (i.e., people, places, organizations, time, money), a generated multimedia summary or view the full original video of a story. The user can also graph trends of named entities in the news for given sources and time periods. For example, Figure 3 graphs the onset and abatement of stories on Princess Diana and Mother Teresa, September 8-15, 1997.

Analyzing audio, video and text streams from digitized video, BNN segments stories extracts named entities, summarizes stories and designs presentations to provide the end user with content-based, personalized Web access to the news (Maybury et al., 1997). For example, within the video stream color histograms are used to classify frames and detect scene changes. In the audio stream, algorithms detect silence, speaker changes and transcribe

Figure 1: BBN's Rough 'n Ready ABC World News Tonight, January 31, 1998