

identifies speech turns, and an analyzer of nonverbal cues based on video analysis to rapidly review records of human interaction.

Multimedia data introduces several challenges to knowledge management systems, including the uncertainties associated with media analyzers and the need for good scalability and effective user interfaces. Architectures capable of handling system complexity will also play a crucial role in deploying multimedia-based KM solutions.

Nonetheless, the prospects for fully exploiting multimedia content are promising. The experience gained in developing multimedia retrieval systems such as SpeechBot shows that even with current limitations in speech recognition technology, analyzers can achieve good performance when searching multimedia sources. Given current trends in audio and video analysis, multimedia storage and distribution over the Internet, developments in XML representations, and integration with knowledge portals, we expect multimedia data to become truly pervasive and as important, if not more so, than textual sources in KM systems. ■

References

1. G. Salton and M.J. McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill, New York, 1983.
2. L.R. Rabiner and B.H. Juang, *Fundamentals of Speech Recognition*, Prentice Hall, Upper Saddle River, N.J., 1993.
3. J.S. Garafolo, C.G.P. Auzanne, and E.M. Voorhees, "The TREC Spoken Document Retrieval Track: A Success Story," *Proc. 6th RIAO Conf. (Content-Based Multimedia Information Access)*, Center for the Advanced Study of Information Systems, Paris, 2000, pp. 1-20.
4. M.G. Brown et al., "Open-Vocabulary Speech Indexing for Voice and Video Mail Retrieval," *Proc. ACM Multimedia*, ACM Press, New York, 1996, pp. 307-316.
5. F. Kubala et al., "Integrated Technologies for Indexing Spoken Language," *Comm. ACM*, vol.43, no. 2, 2000, pp. 48-56.
6. H.D. Wactlar et al., "Lessons Learned from Building a Terabyte Digital Video Library," *Computer*, Feb. 1999, pp. 66-73.
7. S. Srinivasan and D. Petkovic, "Phonetic Confusion Matrix-Based Spoken Document Retrieval," *Proc. 23rd Int'l Conf. Information Retrieval*, ACM Press, New York, 2000, pp. 81-87.
8. J-M. Van Thong et al., "SpeechBot: A Speech Recognition-Based Audio Indexing System for the Web," *Proc. 6th RIAO Conf. (Content-Based Multimedia Information Access)*, Center for the Advanced Study of Information Systems, Paris, 2000, pp. 106-115.
9. B. Logan and A. Salomon, "A Music Similarity Function Based on Signal Analysis," *Proc. IEEE Int'l Conf. Multimedia and Expo*, IEEE Press, Piscataway, N.J., 2001, pp. 952-955.
10. J. Ashley and M. Flickner et al., "The Query by Image Content (QBIC) System," *Proc. ACM Sigmod Conf.*, ACM Press, New York, 1995, p. 475.
11. E.W. Brown et al., "Towards Speech as a Knowledge Resource," *IBM Systems J.*, vol. 40, no. 4, 2001, pp. 985-1001.
12. A. Waibel et al., "Meeting Browser: Tracking and Summarizing Meetings," *Proc. DARPA Broadcast News Transcription and Understanding Workshop*, Morgan Kaufmann, Lansdowne, Va., 1998, pp. 281-286.

Pedro J. Moreno is a senior member of the technical staff at Compaq Computer Corp. Cambridge Research Laboratory. His research interests are multimedia indexing, speech recognition, and applications of machine learning. He received a PhD in electrical and computer engineering from Carnegie Mellon University. He is a member of the IEEE and ESCA. Contact him at pedro.moreno@compaq.com.

J-M. Van Thong is a senior member of the technical staff at Compaq Computer Corp. Cambridge Research Laboratory. His research interests are speech recognition, multimedia indexing, and user interfaces. He received a PhD in computer science from the Pierre and Marie Curie University in Paris. Contact him at jm.vanthong@compaq.com.

Beth Logan is a member of the technical staff at Compaq Computer Corp. Cambridge Research Laboratory. Her research interests are acoustic modeling and indexing of speech and music. She received a PhD in electrical engineering from the University of Cambridge. She is a member of the IEEE. Contact her at beth.logan@compaq.com.

Gareth J.F. Jones is a lecturer in the Department of Computer Science, University of Exeter. His research interests include multimedia, cross-language, and context-aware retrieval. He received a PhD in electrical and electronic engineering from the University of Bristol. He is a member of the Institute of Electrical Engineers. Contact him at G.J.F.Jones@exeter.ac.uk.