

Alexander G. Hauptmann

Home Address

339 Stone Church Rd.
Finleyville, PA 15332
Phone: (724)348-8848

Citizenship: Germany

Work Address

School of Computer Science
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, Pennsylvania 15213-3890 USA
Phone: (412)268-1448
Fax: (412)268-5576
Electronic mail: alex@cs.cmu.edu

Education

- 1984 - 1991 Carnegie Mellon University
PhD (Computer Science) May 1991
Thesis Title: *Meaning from Structure in Natural Language Processing*
- 1982 - 1984 Technische Universitaet Berlin
Diplom (Computer Science)
- 1978 - 1982 Johns Hopkins University
BA (Psychology), MA (Psychology), Minor (French)

Research Experience

Carnegie Mellon University,

Senior System Scientist, School of Computer Science, 1998 - present.
System Scientist, School of Computer Science, 1994 - 1998.
Research on the Informedia Project News-on-Demand Application
Research on speech interfaces, speech synthesis and
Language Instruction That Speech Technology ENables (LISTEN)

Carnegie Mellon University,

Research Associate, School of Computer Science, 1992 - 1994.
Research on speech interfaces, speech synthesis and
Language Instruction That Speech Technology ENables (LISTEN).

Carnegie Mellon University,

Research in the CMU speech group, 1984 - 1992.
Research concentrated on speech interfaces, parsing and grammars for speech recognition.
Also worked on a knowledge-based machine translation system for the
International Center for Machine Translation at CMU.

Teaching assistant for "Fundamental Structures of Computer Science II", spring 1987.

Electrotechnical Laboratory, Agency of Industrial Science and Technology, MITI
(Tsukuba Science City, Japan).

Research fellowship in the speech processing section of ETL, January 1990 - February 1990.
Designed and implemented an air traffic controller simulation with speech interface
for a NeXT workstation using the SPHINX speech recognition technology.

Siemens Research and Technology Laboratories (Princeton, NJ).

Summer internship. June 1985 - August 1985
Worked on the speaker-dependent continuously spoken digit recognition project.

Technische Universitaet Berlin,

Teaching assistant for "Question Answering Systems", Spring 1984.

Johns Hopkins University,

Assistant to Prof. A. Chapanis, 1979 - 1982,
Research on man-computer interaction, telecommunication,
voice communication, telereference libraries and task analysis.

Teaching assistant for "Introductory Psychology", Fall 1982.

Assistant to Prof. S.H. Hulse, 1978 - 1979,
Work involved serial pattern learning and memory.

Honors and Awards

Allen Newell Award for Research Excellence, for the Informedia Digital Library,
with H. Wactlar, M. Christel, T. Kanade and S. Stevens, 2003

AAAI Outstanding Paper Award, Twelfth National Conference on Artificial Intelligence AAAI-94,
for J. Mostow, S. Roth, A. Hauptmann, and M. Kane, "A Prototype Reading Coach that Listens", 1994.

PRIX ARS ELECTRONICA 93 Honorable Mention for "Edge of Intention"
by the OZ/Animation Group (with J.Bates, A.Witkin, J.Altucher, A.Hauptmann, M.Kantrowitz, A.B.Loyall,
K.Murakami, P.Olbrich, Z.Popovich, W.S.Reilly, P.Sengers, W.Welch and P.Weyrauch.)

Judge, 40th International Science and Engineering Fair, Pittsburgh, 1989.

Sigma Xi, 1988.

Phi Beta Kappa, 1982.

Johns Hopkins University, Beneficial - Hodson Scholarship 1978 - 1982.

Bayerische Begabtenfoerderung (Scholarship) 1978.

Published Chapters in Books

Yan, R., Hauptmann, A.G. and Jin, R., Pseudo-Relevance Feedback for Multimedia Retrieval, in Video Mining, Rosenfeld, A., Doermann, D., and DeMenthon, D. (eds), Kluwer, Boston, pp. 309 - 338, 2003.

Hauptmann, A.G., Jin, R. and Wactlar, H.D., Data Analysis for a Multimedia Library in Text and Speech-Triggered Information Access, Renals, S and Grefenstette, G. (eds), Springer, Berlin, pp. 6 - 37, 2003.

Lin, W.-H. and Hauptmann, A.G., Meta-classification: Combining Multimodal Classifiers, in Zaiane, O.R., Simoff, S.J., and Djeraba, C. (eds), Mining Multimedia and Complex Data, Springer Lecture Notes in Artificial Intelligence, LNAI2798, pp. 217-231, Berlin: Springer, 2003.

A.G. Hauptmann and M.J. Witbrock, *Informedia: News-on-Demand - Multimedia Information Acquisition and Retrieval*, in Maybury, M. (ed). *Intelligent Multimedia Information Retrieval* AAAI Press/MIT Press, 1997.

A.I Rudnicky and A.G. Hauptmann, *Multimodal interaction in speech systems* in Blattner, M.M. and Dannenberg, R.B. (eds), *Multimedia Interface Design*, pp. 147 - 172, New York, NY: ACM Press, 1992.

S.R. Young, A.G. Hauptmann, W.H. Ward, E.T. Smith and P. Werner, *High Level Knowledge Sources in Usable Speech Recognition Systems*, in Waibel,A. and Lee, K.-F. (eds) *Readings in Speech Recognition*, San Francisco: Morgan Kaufmann, 1990. Previously published in **Communications of the ACM**, vol. 32, no. 2, pp. 183 - 194, 1989.

Journal Papers

- M. Christel, A. Hauptmann, and H. Wactlar, *Improving Access to Digital Video Archives through Informedia Technology*, Journal of the Audio Engineering Society, July/August 2001, vol.48, no. 7/8, 2001.
- H.D. Wactlar, A.G. Hauptmann, M.G. Christel, R.A. Houghton and A.M. Olligschlaeger, *Complementary Video and Audio Analysis for Broadcast News Archives*, **Communications of the ACM**, Vol 43, No. 2, February 2000, pp. 42 - 47.
- P.J. Jang and A.G. Hauptmann, *Learning How to Recognize Speech By Watching Television*, IEEE Intelligent Systems, vol. 14, no.5, pp. 51 - 58, September/October 1999.
- A.G. Hauptmann, *Integrating and Using Large Databases of Text, Images, Video and Audio*, IEEE Intelligent Systems, vol. 14, no. 5, pp. 34 - 35, September/October 1999.
- Wactlar, H.D., Christel, M.G., Hauptmann, A.G., and Gong, Y., *Informedia Experience-on-Demand: Capturing, Integrating and Communicating Experiences across People, Time and Space.*, ACM Computing Surveys 31(2es), June 1999, No.9 (electronic publication) <http://www.acm.org/pubs/articles/journals/surveys/1999-31-2es/a9-wactlar/a9-wactlar.pdf>
- Wactlar, H.D., Christel, M.G., Gong, Y., and Hauptmann, A.G. *Lessons Learned from the Creation and Deployment of a Terabyte Digital Video Library*, IEEE Computer 32(2): 66-73, 1999.
- M.J. Witbrock and A.G. Hauptmann, *Speech Recognition in a Digital Video Library*, JASIS: Journal of the American Society for Information Science, 1997
- A.I. Rudnicky, A.G. Hauptmann and K.-F. Lee, *Survey of Current Speech Technology*. Communications of the ACM, vol. 37, number 3, pp. 52-57, March, 1994.
- A.G. Hauptmann and P. McAvinney *Gestures with Speech for Graphic Manipulation* in International Journal of Man-Machine Systems, vol. 38, pp. 231 - 249, 1993.
- A.G. Hauptmann and A.I. Rudnicky, *Talking to computers: an empirical investigation*, in International Journal of Man-Machine Studies, vol. 8, no. 6, pp. 583 - 604, 1988.
- A.G. Hauptmann and B.F. Green, *Comparing Command, Menu and Natural Language Computer Systems*, In Behaviour and Information Technology, vol. 2, no. 2, pp. 163 - 178, 1983.

Refereed Conference/Workshop Papers

- R. Yan, J. Yang, A. Hauptmann Automatically Labeling Data Using Multi-class Active Learning To appear in 9th International Conference on Computer Vision (ICCV2003), Nice, France, 2003.
- R. Yan, A. Hauptmann and R. Jin Negative Pseudo-Relevance Feedback in Content-based Video Retrieval ACM Multimedia (MM2003), Berkeley, CA, USA, 2003.
- R. Yan and A. Hauptmann, The Combination Limit of Video Retrieval ACM Multimedia (MM2003), Berkeley, CA, USA, 2003.
- Mark Derthick, Michael G. Christel , Alexander G. Hauptmann, Howard D. Wactlar. Constant Density Displays Using Diversity Sampling. Proceedings of the IEEE Information Visualization Conference (InfoVis'03).
- Yan, R., Hauptmann, A.G. and Jin, R., Pseudo-Relevance Feedback for Multimedia Retrieval, in Video Mining, Rosenfeld, A., Doermann, D., and DeMenthon, D. (eds), Kluwer, Boston, pp. 309 - 338, 2003.
- Lin, W.H., Jin, R. and Hauptmann, A.G., Web Image Retrieval Reranking with a Relevance Model, 2003 IEEE/WIE International Conference on Web Intelligence (WI2003), Halifax, Canada, October, 2003.
- M. Derthick, M. G. Christel, A. G. Hauptmann, D. Ng, S. Stevens, H. D. Wactlar. A Cityscape Visualization of Video Perspectives. Poster presented at the National Academy of Sciences Arthur M. Sackler Colloquium on Mapping Knowledge Domains Colloquium, Irvine, CA, 2003.
- Jin, R., Liu, Y., Si, L., Carbonell, J., and Hauptmann, A. A new boosting algorithm using input dependent regularizer, The Twentieth International Conference on Machine Learning (ICML 2003) August 21-24, 2003 Washington, DC USA.
- Zhang, J., Jin, R., Yang, Y. and Hauptmann, A., Modified Logistic Regression: An Approximation to SVM

- and its Applications in Large-Scale Text Categorization The Twentieth International Conference on Machine Learning (ICML 2003) August 21-24, 2003 Washington, DC USA.
- Jin, R., Yan, R., Zhang, J. and Hauptmann, A., A Faster Iterative Scaling Algorithm For Conditional Exponential Model, The Twentieth International Conference on Machine Learning (ICML 2003) August 21-24, 2003 Washington, DC USA.
- Yan R., Hauptmann, A., Jin, R., Multimedia Search with Pseudo-Relevance Feedback International Conference on Image and Video Retrieval (CIVR2003) July 24-25, 2003 Urbana, IL, USA,
- Jin, R., Hauptmann, A., Learning to Identify Video Shots with People Based On Face Detection 2003 IEEE International Conference on Multimedia & Expo (ICME 2003) July 6-9, 2003 Baltimore, MD, USA.
- Y. Qi, A. Hauptmann, and T Liu, Supervised Classification for Video Shot Segmentation, Proceedings of 2003 IEEE International Conference on Multimedia & Expo (ICME 2003) July 6-9, 2003 Baltimore, MD, USA.
- Yan R., Liu, Y., Jin, R., Hauptmann, A., On Predicting Rare Class with SVM Ensemble in Scene Classification IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2003) April 6-10, 2003 Hong Kong, China
- Ng, T.D., Christel, M., Hauptmann, A.G., and Wactlar, H., Collages as Dynamic Summaries of Mined Video Content for Intelligent Multimedia Knowledge Management, AAAI Spring Symposium Series Intelligent Multimedia Knowledge Management, Palo Alto, CA, March 24-26, 2003.
- Yan, R., Hauptmann, A.G. and Jin, R., Multimedia Search with Pseudo-Relevance Feedback AAAI Spring Symposium Series Intelligent Multimedia Knowledge Management, Palo Alto, CA, March 24-26, 2003.
- Jin, R., Hauptmann, A.G. and Yan, R., Image Classification Using a Bigram Model AAAI Spring Symposium Series Intelligent Multimedia Knowledge Management, Palo Alto, CA, March 24-26, 2003.
- Hauptmann, A.G., Jin, R., and Ng, T.D. Video Retrieval using Speech and Image Information, Storage and Retrieval for Multimedia Databases 2003, EI'03 Electronic Imaging, Santa Clara, CA, January 20-24th, 2003.
- Jin, R., Zhai, C., and Hauptmann, A.G., Information Retrieval for OCR Documents: A content-based Probabilistic Correction Model Electronic Imaging '03, Document Recognition and Retrieval Conference DRR(X), Santa Clara, CA, January 20-24th, 2003.
- Christel, M., Hauptmann, A.G., Wactlar, H., and Ng, T.D., Collages as Dynamic Summaries for News Video, Proceedings of ACM Multimedia 2002, Juan-les-Pins, France, December 1-6, 2002
- Lin, W.H. and Hauptmann, A., News Video Classification Using SVM-based Multimodal Classifiers and Combination Strategies, Proceedings of ACM Multimedia 2002, Juan-les-Pins, France, December 1-6, 2002
- Jin, R., and Hauptmann, A., Using a Probabilistic Source Model for Comparing Images, IEEE 2002 International Conference on Image Processing (ICIP'02), Rochester, New York, September, 22-25, 2002
- Jin, R., and Hauptmann, A., A New Probabilistic Model for Title Generation, The 19th International Conference on Computational Linguistics, Howard International House and Academia Sinica, August 14-September 1, 2002
- A.F. Smeaton, P. Over, C. Costello, A.P. de Vries, D.S. Doermann, A.G. Hauptmann, M.E. Rorvig, J.R. Smith, and L. Wu: The TREC2001 Video Track: Information Retrieval on Digital Video Information. ECDL 2002: 266-275, 2002
- Hauptmann, A.G. and Jin, R., Video Information Retrieval: Lessons Learned with the Informedia Digital Video Library, International Workshop on Digital Communication IWDC 2002: Advanced Methods for Multimedia Signal Processing, Capri, Italy, September 2002.
- Jin, R., Y. Qi, and Hauptmann, A., A probabilistic model for camera zoom motion detection, The sixteenth conference of the International Association for Pattern Recognition (ICPR 2002) QuØbec City, Canada August 11-15 2002

- A.G. Hauptmann, N. Papernick: Video-cuebik: adapting image search to video shots. JCDL 156-157, 2002
- A.G. Hauptmann, R.Jin, T.D. Ng: Multi-modal information retrieval from broadcast video using OCR and speech recognition. JCDL 160-161, 2002
- W. Lin, A.G. Hauptmann: A wearable digital library of personal conversations. JCDL 277-278, 2002
- A.G. Hauptmann, M.G. Christel, N.Papernick: Video retrieval with multiple image search strategies. JCDL 376, 2002
- Lin, W., Jin, R., and Hauptmann, A., Meta-classification of Multimedia Classifiers, International Workshop on Knowledge Discovery in Multimedia and Complex Data (KDMCD), Taipei, Taiwan, May 6, 2002
- Jin, R. , A.G. Hauptmann, C. Zhai, A Content-Based Probabilistic Correction Model for OCR Document Retrieval, the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval Workshop Program (SIGIR 2002), Tampere, Finland, August 11-15, 2002
- R. Jin, L. Si, A.G. Hauptmann, and J.P. Callan: Language model for IR using collection information. SIGIR 2002: 419-420, 2002
- R. Jin, C. Zhai, and A. G. Hauptmann: Title language model for information retrieval. SIGIR'02 5th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Tempere, Finland, p. 42-48, August 11-15 2002
- Lin,W., Jin, R., Hauptmann,A., Triggering Memories of Conversations using Multimodal Classifiers, AAAI-02 Workshop on Intelligent Situation-Aware Media and Presentation (ISAMP 2002),Edmonton, Alberta, Canada, July 28, 2002
- Hauptmann, A., and Lin, W., Video and Audio Analysis for Remembering Conversations, ASRU'01 Automatic Speech Recognition and Understanding Workshop, Madonna di Campiglio, Trento, Italy, December 9-13, 2001 (invited paper)
- R. Jin, C. Faloutsos, and A.G. Hauptmann, "Meta-Scoring: Automatically Evaluating Term Weighting Schema in IR without Precision-Recall", ACM SIGIR Conference on Research and Development in Information Retrieval, New Orleans, LA, September 9-13, 2001.
- R. Jin and A.G. Hauptmann, "Title Generation for Machine Translated Documents", 7th International Joint Conference on Artificial Intelligence (IJCAI'01), Seattle, WA, August 4-10, 2001.
- R. Jin and A.G. Hauptmann, "Learning to Select Good Title Words: A New Approach based on Reversed Information Retrieval", International Conference on Machine Learning (ICML'01), Bershires, MA, June 28-July 1, 2001.
- Palmer, C.R., Wactlar, H., Valdes-Perez, R.E., Pesenti, J., Ng, D., Hauptmann, A., and Christel, M., "Demonstration of Hierarchical Document Clustering of Digital Library Retrieval Results", Joint Conference on Digital Libraries (JCDL'01), Roanoke, VA, p. 451, June 24-28, 2001.
- R. Jin and A.G. Hauptmann, "Automatic Title Generation for Spoken News", Human Language Technology Conference (HLT'01), San Diego, CA, March 18-21, 2001.
- R. Jin and A.G. Hauptmann, "Headline Generation Using a Training Corpus", Second International Conference on Intelligent Text Processing and Computational Linguistics (CICLING'01), Mexico City, Mexico, February 18-24, 2001.
- R. Jin and A. Hauptmann, "Title Generation for Broadcast News using a Training Corpus", ICSLP-2000 6th International Conference on Spoken Language Processing, Beijing, China, October 16-20, 2000.
- Kennedy, P.E. and Hauptmann, A.G., "Automatic Title Generation using EM", ACM Digital Libraries 2000 Conference, San Antonio, TX, pp. 230-231, June 2000.
- Olligschlaeger, A. and Hauptmann, A.G., Multimodal Information Systems and GIS: The Informedia Digital Video Library", 1999 ESRI User Conference, July 27-30, 1999.
- Hauptmann, A.G. and Olligschlaeger, A.M., Using Location Information from Speech Recognition of Television News Broadcasts, European Speech Communication Association Workshop on Accessing Information in Spoken Audio, Cambridge, England, April 19-20, 1999.
- Hauptmann, A.G., and Wactlar, H.D., Informedia News-on-Demand: Automatic Information Extraction for

- Analysis of Broadcast News, AIPA-99, 8th Annual Advanced Information Processing and Analysis Symposium, Washington, DC, March 23-24, 1999.
- Jang, P.J. and Hauptmann, A.G., "Improving Acoustic Models with Captioned Multimedia Speech", accepted to ICMCS'99 (IEEE International Conference on Multimedia Computing and Systems), June 1999, Florence, Italy (in press)
- Jang, P.J. and Hauptmann, A.G., "Selecting For Acoustic Coverage From Unlimited Speech Training Data", Eurospeech'99, 6th European Conference on Speech Communication and Technology, Budapest, Hungary, September 1999 (in press)
- Kennedy, P.E. and Hauptmann, A.G., Laughter extracted from Television Closed Captions as Speech Recognizer Training Data, Eurospeech'99, 6th European Conference on Speech Communication and Technology, Vol2., pp. 663-666, Budapest, Hungary, September 1999.
- A.G. Hauptmann, D. Lee and P.E. Kennedy, Topic Labeling of Multilingual Broadcast News in the Informedia Digital Video Library, ACM Digital Libraries/ SIGIR MIDAS Workshop, Berkeley, CA, June 1999.
- Christel, M.G., Hauptmann, A.G., Warmack, A.S., and Crosby, S.A. Adjustable Filmstrips and Skims as Abstractions for a Digital Video Library. Proceedings of the IEEE Advances in Digital Libraries Conference (Baltimore, MD, May 1999).
- Jang, P.J. and Hauptmann, A.G., "Hierarchical Cluster Language Modeling for Rescoring N-Best Hypotheses during Speech Decoding", Proceedings of ICSLP'98 (International Conference on Speech and Language Processing), Sydney Australia
- Jang, P.J. and Hauptmann, A.G., Hierarchical Cluster Language Modeling with Statistical Rule Extraction for Rescoring N-Best Hypotheses During Speech Decoding ICSLP '98 5th International Conference on Spoken Language Processing, Sydney, Australia, December, 1998.
- Hauptmann, A.G., and Lee, D., Topic Labeling of Broadcast News Stories in the Informedia Digital Video Library Digital Libraries '98 - The Third ACM Conference on Digital Libraries, Pittsburgh, PA, June, 1998.
- Hauptmann, A.G., The Informedia Digital Video Library: Learning in a Mixed Media Database, Conference on Automated Learning and Discovery, Workshop on Mixed Media Databases, Pittsburgh PA, June 11-13, 1998.
- Hauptmann, A.G., and Witbrock, M.J., Story Segmentation and Detection of Commercials in Broadcast News Video, ADL-98 Advances in Digital Libraries, Santa Barbara, CA, April, 1998.
- Hauptmann, A.G., Jones, R.E., Seymore, K., Siegler, M.A., Slattery, S.T., and Witbrock, M.J., Experiments in Information Retrieval from Spoken Documents, BNTUW-98 Proceedings of the DARPA Workshop on Broadcast News Transcription and Understanding Systems, Lansdowne, VA, February, 1998.
- Hauptmann, A.G., Scheytt, P., Wactlar, H.D., and Kennedy, P.E., Multi-Lingual Informedia: A Demonstration of Speech Recognition and Information Retrieval across Multiple Languages, BNTUW-98 Proceedings of the DARPA Workshop on Broadcast News Understanding Systems, Lansdowne, VA, February, 1998.
- M.J. Witbrock, and A.G. Hauptmann, *Using Words and Phonetic Strings for Efficient Information Retrieval from Imperfectly Transcribed Spoken Documents*, in DL97, The Second ACM International Conference on Digital Libraries, July 23 - 26, 1997, Philadelphia.
- A.G. Hauptmann, M.J. Witbrock, and M.G. Christel, *Artificial Intelligence Techniques in the Interface to a Digital Video Library*, Proceedings of the CHI-97 Computer-Human Interface Conference New Orleans LA, March 1997
- M.J. Witbrock and A.G. Hauptmann, *Improving Acoustic Models by Watching Television*, Working Notes for AAAI-97 Spring Symposium on Intelligent Integration and Use of Text, Image, Video and Audio Corpora. March 24-26 1997, Stanford, pp 61-63
- A.G. Hauptmann and M.J. Witbrock, *Informedia News-On-Demand: Using Speech Recognition to Create a Digital Video Library*, Working Notes for AAAI-97 Spring Symposium on Intelligent Integration and Use of Text, Image, Video and Audio Corpora. March 24-26 1997, Stanford, pp 120-126

- M.J. Witbrock, and A.G. Hauptmann, *Speech Recognition and Information Retrieval* Proceedings of the 1997 DARPA Speech Recognition Workshop, February 2-5, 1997.
- A.G. Hauptmann and H.D. Wactlar, *Indexing and Search of Multimodal Information*, International Conference on Acoustics, Speech and Signal Processing (ICASSP-97) Munich, Germany, April, 1997
- H.D. Wactlar, A.G. Hauptmann, and M.J. Witbrock, *Informedia: News-on-Demand Experiments in Speech Recognition*, ARPA Speech Recognition Workshop, February 18-21, 1996.
- A.G. Hauptmann, *Speech Recognition in the Informedia Digital Video Library: Uses and Limitations*, ICTAI-95 7th IEEE International Conference on Tools with AI, Washington, DC, November 6-8, 1995.
- A.G. Hauptmann, and M. Smith, *Text, Speech, and Vision for Video Segmentation: The Informedia Project*. AAAI Fall Symposium, Computational Models for Integrating Language and Vision. Boston, November 10-12, 1995
- A.G. Hauptmann, M.J. Witbrock, A.I. Rudnicky, and S. Reed, *Speech for Multimedia Information Retrieval*, UIST-95 Proceedings of User Interface Software and Technology, Pittsburgh, PA, November, 1995.
- H.D. Wactlar, A.G. Hauptmann, M. Smith and K. Pendyala, *Automated Video Segmentation for On-Demand Retrieval from Very Large Video Libraries*, The 138th SMPTE (Society of Motion Picture and Television Engineers) Technical Conference, Los Angeles, CA, 1995.
- H.D. Wactlar, M. Christel, A. Hauptmann, T. Kanade, M. Mauldin, R. Reddy, M. Smith, S. Stevens *Technical Challenges for the Informedia Digital Video Library* Proceedings of ISDL'95: International Symposium on Digital Libraries 1995. Tsukuba Science City, Ibaraki, Japan, August 1995. (Invited paper)
- A.G. Hauptmann and M.A. Smith, *Video Segmentation in the Informedia Project* IJCAI-95, Workshop on Intelligent Multimedia Information Retrieval - Working Notes, Montreal, Quebec Canada. August 19, 1995, pp. 17 - 22.
- L. Chase, R. Rosenfeld, A. Hauptmann, R. Mosur, E. Thayer, P. Placeway, W. Weide, and C. Lu, *Improvements in Language, Lexical and Phonetic Modeling in Sphinx-II*, Proceedings ARPA Spoken Language Technology Workshop, Austin TX January 1995, pp. 60-65.
- J. Mostow, S. Roth, A.G. Hauptmann, and M. Kane, *A Prototype Reading Coach that Listens*, Twelfth National Conference on Artificial Intelligence (AAAI94), Seattle, WA, August 1994. Recipient of the AAAI-94 Outstanding Paper Award.
- J. Mostow, S. Roth, A. Hauptmann, M. Kane, A. Swift, L. Chase, and B. Weide, *A Reading Coach that Listens* (6-minute video), Video Track of the Twelfth National Conference on Artificial Intelligence (AAAI94), Seattle, WA, August 1994.
- A. Hauptmann, J. Mostow, S.F. Roth, M. Kane and A. Swift, *A Prototype Reading Coach that Listens: Summary of Project LISTEN*, in Proceedings of the ARPA Workshop on Human Language Technology, Princeton, NJ: March 1994.
- A.G. Hauptmann, L.L. Chase and D.J. Mostow *Speech Recognition applied to reading assistance for children: A baseline language model*. in Eurospeech-93, 3rd European Conference on Speech Communication and Technology - Proceedings, vol. 3, pp. 2255 - 2258, Berlin, Germany: ESCA, September 1993.
- A.G. Hauptmann *SpeakEZ: A first experiment in concatenation synthesis from a large corpus* in Eurospeech-93, 3rd European Conference on Speech Communication and Technology - Proceedings, vol. 3, pp. 1701 - 1705, Berlin, Germany: ESCA, September 1993.
- D.J. Mostow, A.G. Hauptmann, L.L. Chase and S. Roth *Towards a reading coach that listens: Automated detection of oral reading errors* "Proceedings of the Eleventh National Conference on Artificial Intelligence (AAAI93)", pp. 392 - 397, Washington, DC: American Association for Artificial Intelligence, July 1993.
- A.G. Hauptmann, *Meaning from Structure in Natural Language Processing* in Proceedings of AAAI-91, National Conference on Artificial Intelligence, pp. 125 - 130, July, 1991.
- A.I. Rudnicky and A.G. Hauptmann, *Models for evaluating interaction protocols in speech recognition*, in

CHI-91: Proceedings of the Conference on Computer Human Interfaces, May 1991.

- A. Waibel, A.N. Jain, A.E. McNair, H. Saito, A.G. Hauptmann and J. Tebelskis, *JANUS: a speech-to-speech translation system using connectionist and symbolic processing strategies*, ICASSP-91, International Conference on Acoustics, Speech and Signal Processing, 1991.
- A.G. Hauptmann and A.I. Rudnicky, *A comparison of speech versus typed input*, in Proceedings of the Third DARPA Speech and Natural Language Workshop, Morgan Kaufmann, San Mateo, CA. June 1990.
- A.G. Hauptmann and A.I. Rudnicky, *Speaking, Typing and Gesturing* in CHI-90 Workshop on Multimedia and Multimodal Interfaces, International Conference on Human Factors in Computing Systems, Seattle WA, April 1990.
- A.I. Rudnicky and A. G. Hauptmann *Errors, Correction and Contrastive Emphasis in Speech Recognition Systems*, in AAAI Spring Symposium: Spoken Language Systems, March 1989.
- S.R. Young, A.G. Hauptmann and W.H. Ward, *Layering Predictions: Flexible Use of Dialog Expectations in Speech Recognition*, in IJCAI-89, Proceedings of the Eleventh Annual Joint Conference on Artificial Intelligence, August 1989.
- A.G. Hauptmann, *Speech and Gestures for Graphic Image Manipulation*, in CHI-89: Proceedings of the Conference on Computer Human Interfaces, May 1989.
- A.G. Hauptmann, S.R. Young and W.H. Ward, *Using Dialog-Level Knowledge Sources to Improve Speech Recognition*, in Proceedings of AAAI-88, The 7th National Conference on Artificial Intelligence, Gregg, L.W. and Steinberg, E.R.(eds), August 1988.
- W.H. Ward, A.G. Hauptmann, R.M. Stern and T. Chanak, *Parsing Spoken Phrases Despite Missing Words*, ICASSP-88, International Conference on Acoustics, Speech and Signal Processing, 1988.
- R.M. Stern, W.H. Ward, A.G. Hauptmann and J. Leon, *Sentence Parsing with Weak Grammatical Constraints*, in ICASSP-87, International Conference on Acoustics, Speech and Signal Processing, August 1987.
- P.J. Hayes, A.G. Hauptmann, J.G. Carbonell and M. Tomita, *Parsing Spoken Language: a Semantic Caseframe Approach*, COLING-86, International Conference on Computational Linguistics, August 1986.
- S. Nakagawa, A.G. Hauptmann and M. Tomita, *On Quick Word Spotting Techniques*, ICASSP-86, International Conference on Acoustics, Speech and Signal Processing, April 1986.

Technical Reports

- Wactlar, H.D., Hauptmann, A.G., and Witbrock, M.J., *Informedia News-on-Demand: Using Speech Recognition to Create a Digital Video Library*, Technical Report, CMU-CS-98-109, Pittsburgh, PA, March 19, 1998.
- Witbrock, M.J., and Hauptmann, A.G., *Improving Acoustic Models by Watching Television*, Technical Report, CMU-CS-98-110, Pittsburgh, PA, March 19, 1998.
- A.G. Hauptmann, *Meaning from Structure in Natural Language Processing*, Ph.D. Dissertation, Technical Report CMU-CS-91-158, July 1991.
- A.G. Hauptmann, P. McAvinney, and S. Shepard, *Gesture Analysis for Graphic Manipulation*, in Technical Report CMU-CS-88-198, November, 1988.
- S.R. Young, A.G. Hauptmann and W.H. Ward, *An Integrated Speech and Natural Language Dialog System: Using Dialog Knowledge in Speech Recognition*, in Technical Report CMU-CS-88-128, April, 1988.

Other Publications

- A. Hauptmann, R.V. Baron, M.-Y. Chen, M. Christel, P. Duygulu, C. Huang, R. Jin, W.-H. Lin, T. Ng, N. Moraveji, N. Papernick, C.G.M. Snoek, G. Tzanetakis, J. Yang, R. Yan, and H.D. Wactlar, Informedia at TRECVID 2003: Analyzing and Searching Broadcast News Video, Proceedings of (VIDEO) TREC 2003, Gaithersburg, MD, November 2003.
- C.G.M. Snoek and A.G. Hauptmann, Learning to Identify TV News Monologues by Style and Context, , Technical Report CMU-CS-03-193, October 2003.
- Hauptmann, A.G. Yan, R., Jin, R., Christel, M., Derthick, M., Chen, M., Lin, W., Baron, R. and Ng, T.D. Video Classification and Retrieval with the Informedia Digital Video Library System TREC 2002, Gaithersburg, MD, November 2002.
- Alexander G. Hauptmann, Rong Jin, Norman Papernick, Dorbin Ng, Yanjun Qi, Ricky Houghton, Sue Thornton: Video Retrieval with the Informedia Digital Video Library System.TREC 2001, Gaithersburg, MD, November 2001.
- Siegler, M., Jin, R. and Hauptmann, A., CMU Spoken Document Retrieval in TREC-8: Analysis of the role of Term Frequency TF, Proceedings of TREC-8, The eighth Text REtrieval Conference, NIST, Gaithersburg, MD, November 1999.
- Siegler, M., Berger, A., Witbrock, M. and Hauptmann, A., Experiments in Spoken Document Retrieval at CMU, TREC-7, Proceedings of the 7th Text Retrieval Conference, Gaithersburg MD, November, 1998.
- Siegler, M.A., Witbrock,M.J., Slattery, S.T., Seymore, K., Jones, R.E., and Hauptmann, A.G., Experiments in Spoken Document Retrieval at CMU, TREC-6 Proceedings of the 1997 Text REtrieval Conference, Gaithersburg, MD, November 19-21, 1997.
- A.G. Hauptmann, M.J. Witbrock and M.G. Christel. *News-on-Demand: an Application of Informedia Technology*. D-Lib Magazine, Sept 1995, <http://www.cnri.reston.va.us/home/dlib.html>, cnri.dlib/september95-hauptmann
- K.-F. Lee, A.G. Hauptmann and A.I. Rudnicky, *The Spoken Word*, in BYTE, July 1990.

Invited Talks and Presentations

- September 2003., A. Hauptmann, Machine Learning for Video Classification and Retrieval, ECML/PKDD-2003, Workshop on Multimedia Discovery and Mining, Dubrovnik, Croatia,
- July, 2003, A. Hauptmann, Evaluation of Video Retrieval, ARDA VACE Challenge Workshop on Video Event Ontologies, San Diego, CA.
- January 2003, Video Analysis in the Informedia Project, Fraunhofer Institute for Computer Graphics, Workshop: Multimedia Information Retrieval in Business Applications, Darmstadt, Germany.
- November, 2002. Invited Presentation, Finding Information in a Digital Video Archive, DIMACS Workshop on Video Mining. DIMACS Center, Rutgers University, Piscataway, New Jersey
- September 2002, Panel on "Semantic Multimedia Indexing", International Workshop on Digital Communication IWDC 2002: Advanced Methods for Multimedia Signal Processing, Capri, Italy, September 2002.
- August 2002, Location Based Multimedia Information for Mobile Environments, Seminar on Emerging Challenges of Mobile Multimedia Services, Chinese University of Hong Kong, Hong Kong. July 2002, Information Retrieval and Video Summarization, CIVR2002 International Conference on Image and Video Retrieval, London, U.K., 2002 (keynote address)
- January 2002, The Informedia Digital Video Library System and the TREC Video Retrieval Evaluations, Dagstuhl Seminar 02021: Content-Based Image and Video Retrieval, Schloss Dagstuhl, Wadern, Germany.
- May 2001, Automatic Content Extraction and Indexing of Video Data, Invited Talk, Information Technology for NIMA 2010-2020, National Imagery and Mapping Agency, Washington, DC.
- December, 2000, Content-based Video Retrieval, Invited Talk at the Multi Media Mining Retrieval Workshop, DaimlerChrysler, Ulm, Germany.

- October 2000, Multimedia Information Retrieval from a Digital Video Library", Invited Talk at the 6th International Workshop on Multimedia Information Systems, Chicago, IL, October 26-26, 2000.
- June 2000, Talk on Visualization and Summarization over Multiple Documents, DLI2 All-Projects Meeting, Stratford-upon-Avon, England, June 12-13, 2000.
- June 2000, Invited Talk on Multilingual Title Generation, Workshop on Cross Lingual Information Retrieval, Human Computer Interaction Laboratory (HCIL), University of Maryland, College Park MD.
- February 1999, Invited Talk on Information Retrieval from Video Documents, Center for Library and Information Science, University of Maryland, College Park MD.
- January 1999, Invited Talk on Information Extraction in the Informedia Digital Video Library, JustResearch Pittsburgh Research Center, Pittsburgh PA
- October 1998, Invited Presentation on Experience-on-Demand at the International Symposium for Wearable Computing ISWC, Pittsburgh PA.
- March 1997 Invited Talk on Multimedia Information Retrieval at ISI, University of Southern California, Los Angeles, CA.
- February 1997 - Invited Talk on Spoken Document Retrieval at AT&T Research Labs, Murray Hill, NJ.
- November 1996 - Invited Talk on Speech Recognition in Informedia at Lucent Technologies, NJ.
- October 1996 Invited Talk at the 10th Anniversary of the Center for Machine Translation, Carnegie Mellon University, Pittsburgh, PA.
- May 1996 Invited Talk at the University of Karlsruhe on Informedia News-on-Demand, Karlsruhe Germany.
- October 1995 - April 1996 Multiple Informedia presentations at ARPA in Washington DC to ARPA and military officials, including the Secretary of Defense. Washington, DC.
- August 1995 Informedia News-on-Demand System Demonstration at the ARPA STISS Symposium. Washington, DC.
- May 1995 Informedia News-on-Demand System Demonstration at the Association of American Universities Open House for Congressional members. Washington, DC.
- June 1994 Invited Presentation on Speech Synthesis at Apple Computer, Cupertino CA.
- February 1992 Invited Talk on Dialog in Speech Recognition System at the German Research Institute for Artificial Intelligence (DFKI), Saarbruecken, Germany.
- Panel on Glove-based Interfaces*, HFS-91, Annual Meeting of the Human Factors Society, San Francisco, September 1991.
- March 1991 NCR Human Interface Technology Labs, Talk on the Use of Dialog in Speech Recognition Systems. Atlanta, GA.

Professional Activities

- Spring 2004, program committee of PCM2004, Pacific-Rim Conference for Multimedia, Tokyo, Japan
- Spring 2004 Co-Organizer of the ARDA VACE Challenge Workshop: Large-Scale Ontologies for Multimedia
- Spring 2004, program committee, International Workshop Multimedia Data and Document Engineering, Washington, DC.
- Spring 2004, program committee, ACM Multimedia.
- 2003-2004 program committee, SPIE/EI Document Recognition and Retrieval Conference (DRR)
- 2003-2004 Advisory Panel to TREC Video Retrieval Track
- Spring 2003 Co-Organizer of the AAAI Spring Symposium on Intelligent Multimedia Knowledge Management, Palo Alto, CA.
- Spring 2003, program committee, SIGIR 2003.

Spring 2003, program committee, International Workshop Multimedia Data and Document Engineering, Berlin, Germany.

Spring 2002 Grant Reviewer for the NSF Information Technology Proposals.

Spring 2001 Grant Reviewer for the Institute of Museum and Library Services

2001 - present, Co-founder, Carnegie Interactive, Pittsburgh, PA.

2001 - present, Co-founder and Board Member, Medrespond, Pittsburgh, PA.

Spring 2001, program committee, SIGIR 2001, New Orleans, LA.

Spring 2001, program committee, International Workshop on Multimedia Data Mining, KDD 2001, San Francisco, CA.

Spring 2000, program committee, RIAO 2000, Paris, France

Spring 2000 Consultant on Synthetic Interview Technology for Westwind Media, Denver CO.

April 1999, program committee, ESCA Cambridge Robust Audio Information Extraction Workshop, Cambridge University, Cambridge, England

August 1998, program committee, SIGIR-98 Workshop on Multimedia Indexing and Retrieval, Melbourne, Australia.

May 1998, Co-organizer, Center for Automated Learning and Discovery, Workshop on Mixed Media Databases, Pittsburgh, PA

1998/1999 NIST SBIR Grant Reviewer

1999 Guest Editor of IEEE Expert Special Issue on the Integration of Text, Image, Video and Audio Information.

Sept 1997 Grant Reviewer for the National Endowment for the Humanities

March 1997 Co-Organizer of the AAAI Spring Symposium of intelligent integration of text, image, video and audio information.

March 1996-2000 Co-Founder and member of the board, Grand Illusion Studios, Inc., Pittsburgh PA.

1994 - present, Member of the editorial board of the International Journal of Speech Technology.

1992 - 1994, Member of the editorial board of the Journal of the American Voice Input/Output Society (AVIOS)

March 1993 - June 1994 Consultant on Speech Synthesis for EER Systems Corp, Vienna VA.

Video Productions

1995 *Informedia: News-on-Demand*, several versions of this video are available

1994 J. Mostow, S. Roth, A. Hauptmann, M. Kane, A. Swift, L. Chase, and B. Weide, *A Reading Coach that Listens* (6-minute video), Video Track of the Twelfth National Conference on Artificial Intelligence (AAAI94), Seattle, WA, August 1994.

1988 Young, S.R., Hauptmann, A.G. and Ward, W.H. *The Multimodal Interactive Dialog Systems (MINDS)*, several versions available.

Grants as Co-Principal Investigator

Howard Wactlar, Alex Hauptmann, Mike Christel, ENVIE - Extensible Video Information Environment, ARDA VACE-II \$1,599,326, 2004 - 2005;

Howard Wactlar, Scott Stevens, Ashok Bharucha, Mike Christel, Mark Derthick, Alex Hauptmann, CareMedia: Automated Video and Sensor Analysis for Geriatric Care, NSF IIS-0105219, \$2,000,000, 2002 - 2007.

Howard Wactlar, Alex Hauptmann, Question Answering from Errorful Multimedia Streams, ARDA MDA904-02-C-0451, \$1,115,833, 2002 - 2004.

Howard Wactlar, Takeo Kanade, Michael Christel, Alex Hauptmann, Capturing, Coordinating and Remembering Human Experience, NSF IIS ITR/IM #0121642, \$2,000,000, 2001-2005,

Howard Wactlar, Michael Christel, Alex Hauptmann, Jianbo Shi, Informedia Video Information Summarization and Demonstration, MDA908-00-C-0037 Advanced Research and Development Agency (ARDA), \$800,000, 2000 - 2002.

Howard Wactlar, Alex Hauptmann, ECHO European Chronicles On-Line. NSF, \$300,005, 2001 - 2003.

Howard Wactlar, Michael Christel, Alex Hauptmann, Threading Information Pathways Through NSDL Video, National Science Mathematics, Engineering and Technology Education Digital Library, NSF National Science, Mathematics, Engineering and Technology Education Digital Library Program, DUE-0085834, \$500,000, 2000 - 2002.

Howard Wactlar, Michael Christel, Alex Hauptmann, Informedia-II: Integrated Video Information Extraction and Synthesis for Adaptive Presentation and Summarization from Distributed Libraries, NSF Digital Libraries Initiative Phase II, NSF IRI-9817496, \$4.150,000, 1999 - 2004.

Howard Wactlar, Michael Christel, Alex Hauptmann, Yihong Gong, Informedia Experience On Demand, DARPA CVIM, N66001-97-C-8517, 1997

Learning, Visualization and the Analysis of Large-Scale, Multiple-Media Data Stephen Fienberg, T. Mitchell, Jaime Carbonell, Bill Eddy, Clark Glymour, Alex Hauptmann, Marsha Lovett, Mike Meyer, Sebastian Thrun, Andrew Moore, Peter Spirtes, Larry Wasserman, Manuela Veloso NSF LIS 97-18, \$825,000, 1997 - 2001.

Interests

Over the years, my research interests have led me to pursue and combine several different areas of research: man-machine communication, natural language processing and speech understanding: In the area of man-machine communication, I am interested in the tradeoffs between different modalities, including gestures and speech, and in the intuitiveness of interaction protocols. In natural language processing, my desire is to break through the bottlenecks that are currently preventing larger scale natural language applications. The latter theme was also the focus of my thesis, which investigated the use of machine learning on large text samples to acquire the knowledge needed for semantic natural language understanding. The approach developed in my thesis research was quite successful at using example sentences and target meanings to learn rules for semantic interpretation from syntactic structures. In the field of speech understanding, I am intrigued by the potential of combining natural language technology with clever interfaces in speech recognition applications.

Earlier work: During my graduate research at CMU, I was centrally involved in the MINDS project (Multi-modal Interactive Dialog System) which became the first system to demonstrate the use of dialog constraints for improving speech recognition accuracy. Other research touched on speech synthesis using a large corpus of example data, and speech-to-speech machine translation in the form of a simple, but effective phrasebook translation system.

In more recent years, I worked with Jack Mostow to build the first prototypes for Project LISTEN (Language Instruction that Speech Technology Enables), where we explored the idea that that speech recognition can be used to detect reading mistakes and help children learn to read correctly.

Most recently, my research in the Infromedia Project has been to build the News-on-Demand application, which is an instantiation of the Infromedia Digital Video Library idea, based completely on automatic methods for processing television and radio news.

In each of these systems, there was one central aspect that made them all effective examples for their respective purposes. Through the combination of the strengths of speech recognition, natural language processing and interface design, the systems were able to overcome some of the shortfalls inherent in speech, natural language and interfaces.

My goal is to continue this type of research, at the boundaries of natural language processing, speech recognition and synthesis and man-machine communication.

Near term goals: In the immediate future, my research goals are closely connected to the current issues faced by the Infromedia project. One fundamental tenet of the Infromedia project is the use of speech recognition to generate transcripts which provide the basis for information retrieval of spoken content in video and audio material. My research efforts will be directed at establishing that speech recognition of video tape sound track is adequate for information retrieval. One aspect of this research involves looking at the relationship between speech recognition accuracy and information retrieval precision and recall, which is the standard metric in the information retrieval community to measure the quality of retrieval. I also plan to leverage the strengths of the Infromedia system to help improve speech recognition through information retrieval and through access to large amounts of transcribed or closed-captioned acoustic data.

Long term goals: In the long term, my research goals are to utilize large corpora of found data, or other sources of knowledge that already exist to improve speech and natural language processing by exploiting advantages across different modalities. It has become clear in recent years that large volumes of text, image, video and audio can be easily stored and made available for research and applications. However, most of these text, image, video and audio sources were not produced with computer processing in mind. My goal is to design and build intelligent, understanding programs that help process data from these sources and make the data useful for other applications. This data can be used to improve speech recognition, image understanding, natural language processing, machine learning as well as information retrieval. While data can always be created through manual labeling, transcription, or annotation, my goal is to exploit "found" data, that is data that is already available through the Internet or other readily accessible open sources. The challenge will be to find the right data, process it into suitable form for training, learning or re-use and build mechanisms that can successfully utilize this data.

For speech recognition, large amounts of "found" data can improve acoustic models, language models and pronunciation models. Training speech recognition systems is currently time-consuming and expensive. Much of the effort involved is devoted to collecting and transcribing speech for training. My desire is to produce a system capable of independently learning to recognize speech given only television or radio input and an Internet connection, without manually created meta-data for training.

Similarly, natural language processing aimed at producing summaries or categorizations could improve through large amounts of found data. In the field of computer vision, image processing can detect features in images that are correlated to the text, and over large amounts of data, this correlation can form the basis of knowledge about images and lead to a form of image understanding. Information

retrieval systems, and even speech synthesis, all stand to gain from methods that exploit large quantities of available data.

My research goals are to conduct research at the intersection of speech, natural language and man-machine communication. I believe it is through the crossing of boundaries and disciplines that the most dramatic research accomplishments will be achieved.

Multimedia technology is about to make a major impact on our daily interaction with computers. What is needed at this point are examples of applications of the technology that are easy to use and that provide clear demonstrations of the advantages of integrated speech and multimedia interfaces. Successful systems will probably also include a variety of knowledge sources that can complement each other. It is my goal is to help this breakthrough come about.