

Papers in Speech Communication: Speech Processing

Bishnu S. Atal, Joanne L. Miller, Raymond D. Kent, Editors

Published in 1991

TABLE OF CONTENTS

Preface

Acknowledgments

SPEECH ANALYSIS AND SYNTHESIS

Commentary

Paper 1. J. Allen (1976), Synthesis of speech from unrestricted text. Proc. IEEE 64, 433-442

Paper 2. B.S. Atal and S. L. Hanauer (1971), Speech analysis and synthesis by linear prediction of the speech wave.

Journal of the Acoustical Society of America 50, 637-655

Paper 3. J.N. Holmes (1983), Formant synthesizers: Cascade or parallel? Speech Communication 2, 251-273

Paper 4. D.H. Klatt (1987), Review of text-to-speech conversion for English. Journal of the Acoustical Society of America 82, 737-793

Paper 5. J. Makhoul (1975), Linear prediction: A tutorial review. Proc. IEEE 63, 561-580

Paper 6. S. Nakajima and H. Hamada (1988), Automatic generation of synthesis units based on context oriented clustering. Proc. IEEE-ICASSP, New York, NY, 659-662

Paper 7. Y. Sagisaka and H. Sato (1986) Composite phoneme units for the speech synthesis of Japanese. Speech Communication 5, 217-223

SPEECH CODING

Commentary

Paper 8. B.S. Atal and J.R. Remde (1982), A new model of LPC excitation for producing natural-sounding speech at low bit rates. Proc. IEEE-ICASSP, Paris, France, 614-617

Paper 9. B.S. Atal and M.R. Schroeder (1979), Predictive coding of speech signals and subjective error criteria. IEEE Trans. ASSP ASSP-27, 247-254

Paper 10. A. Buzo, A.H. Gray, Jr., R. M. Gray, and J.D. Markel (1980), Speech coding based upon vector quantization. IEEE Trans. ASSP ASSP-28, 562-574

Paper 11. J.H. Chen (1990), A robust low-delay CELP speech coder at 16 kb/s. In B.S. Atal, V. Cuperman, and A. Gersho (Eds.), Advances in Speech Coding, pp. 25-35. Norwell, MA: Kluwer Academic

Paper 12. I.A. Gerson and M.A. Jasiuk (1990), Vector sum excited linear prediction (VSELP). In B.S. Atal, V. Cuperman, and A. Gersho (Eds.), Advances in Speech Coding, pp. 69-79. Norwell, MA: Kluwer Academic

Paper 13. R.M. Gray, A. Buzo, A.H. Gray, Jr., and Y. Matsuyama (1980), Distortion measures for speech processing. IEEE Trans. ASSP ASSP-28, 367-376

Paper 14. A.H. Gray, Jr. and J.D. Markel (1976), Distance measures for speech processing. IEEE Trans. ASSP, ASSP-24, 380-391

Paper 15. J.S. Lim and A.V Oppenheim (1979), Enhancement and bandwidth compression of noisy speech. Proc. IEEE 67, 1586-1604

Paper 16. J. Makhoul, S. Roucos, and H. Gish (1985), Vector quantization in speech coding. Proc. IEEE 73, 1551-1588

Paper 17. M.R. Schroeder and B.S. Atal (1985), Code-excited linear prediction (CELP): High-quality speech at very low bit rates. Proc. IEEE-ICASSP, Tampa, FL, 937-940

Paper 18. M.R. Schroeder, B.S. Atal, and J.L. Hall (1979), Optimizing digital speech coders by exploiting masking properties of the human ear. Journal of the Acoustical Society of America 66, 1647-1652

Paper 19. S. Singhal and B.S. Atal (1989), Amplitude optimization and pitch prediction in multipulse coders. IEEE Trans. ASSP 37, 317-327

Paper 20. I.M. Trancoso and B.S. Atal (1990), Efficient search procedures for selecting the optimum innovation in stochastic coders. IEEE Trans. ASSP 38, 385-396

Paper 21. J.M. Tribolet and R.E. Crochiere (1970), Frequency domain coding of speech. IEEE Trans. ASSP ASSP-27, 512-530

AUTOMATIC SPEECH RECOGNITION

Commentary

Paper 22. L.R. Bahl, F. Jelinek, and R.L. Mercer (1983), A maximum likelihood approach to continuous speech recognition. IEEE Trans. Patt. Anal. Machine Intell. Pami-5, 179-190.

- Paper 23. Y.L. Chow, M.O. Dunham, O.A. Kimball, M.A. Krasner, G.F. Kubala, J. Makhoul, P.J. Price, S. Roucos, and R.M. Schwartz (1987), BYBLOS: The BBN continuous speech recognition system. Proc. IEEE-ICASSP, Dallas, TX, 89-92
- Paper 24. S. B. Davis and P. Mermelstein (1980), Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. IEEE Trans. ASSP, ASSP-28, 357-366
- Paper 25. S. Furui (1986), Speaker-independent isolated word recognition using dynamic features of speech spectrum. IEEE Trans. ASSP ASSP-34, 52-59
- Paper 26. F. Itakura (1975), Minimum prediction residual principle applied to speech recognition. IEEE Trans. ASSP, ASSP-23, 67-72
- Paper 27. F. Jelinek (1976), Continuous speech recognition by statistical methods. Proc. IEEE 64, 532-556
- Paper 28. F. Jelinek (1985), The development of an experimental discrete dictation recognizer. Proc. IEEE 73, 1616-1624
- Paper 29. B.H. Juang and L.R. Rabiner (1985), Mixture autoregressive hidden Markov models for speech signals. IEEE Trans. ASSP, ASSP-33, 1404-1413
- Paper 30. K.-F. Lee, H.-W. Hon, and R. Reddy (1990), An overview of the SPHINX speech recognition system. IEEE Trans. ASSP-38, 35-45
- Paper 31. C.H. Lee, L.R. Rabiner, R. Pieraccini, and J.G. Wilpon (1990), Acoustic modeling for large vocabulary speech recognition. Computer Speech and Language 4, 127-165
- Paper 32. S.E. Levinson, L.R. Rabiner, and M.M. Sondhi (1983), An introduction to the application of the theory of probabilistic functions of a Markov Process to automatic speech recognition. Bell System Technical Journal 62, 1035-1074
- Paper 33. L.R. Rabiner (1989), A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE 77, 257-286
- Paper 34. L.R. Rabiner, S.E. Levinson, A.E. Rosenberg, and J.G. Wilpon (1979), Speaker-independent recognition of isolated words using clustering techniques. IEEE Trans. ASSP, ASSP-27, 336-349
- Paper 35. A.J. Robinson and F. Fallside (1989), A dynamic connectionist model for phoneme recognition. In L. Personnaz and G. Dreyfus (Eds.), Neural Networks from Models to Applications, pp. 541-550. Paris: IDSET
- Paper 36. H. Sakoe and S. Chiba (1978), Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. ASSP, ASSP-26, 43-49
- Paper 37. K. Shikano, K.-F. Lee, and R. Reddy (1986), Speaker adaptation through vector quantization. Proc. IEEE-ICASSP, Tokyo, Japan, 2643-2646

AUTOMATIC SPEAKER RECOGNITION

Commentary

- Paper 38. B.S. Atal (1974), Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification. Journal of the Acoustical Society of America 55, 1304-1312
- Paper 39. S. Furui (1981), Cepstral analysis technique for automatic speaker verification. IEEE Trans. ASSP, ASSP-29, 254-272
- Paper 40. A.E. Rosenberg and F.K. Soong (1987), Evaluation of a vector quantization talker recognition system in text independent and text dependent modes. Computer Speech and Language 22, 143-157
- Paper 41. F.K. Soong and A.E. Rosenberg (1988), On the use of instantaneous and transitional spectral information in speaker recognition. IEEE Trans. ASSP 36, 871-879

Index to current volume

Table of Contents and Index to companion volume on Speech Production

Table of Contents and Index to companion volume on Speech Perception

Preface

Over the past few decades there has been great progress in understanding the nature of human speech production and perception, and in applying this knowledge to problems of speech processing (coding, recognition, and synthesis). Given the interdisciplinary nature of the enterprise, important papers in these areas have appeared in a wide range of journals, proceedings, and books from such diverse fields as engineering, linguistics, physics, psychology, and speech and hearing science. The current volume forms part of a three-volume series whose purpose is to bring together a number of these important papers. The series is sponsored by the Acoustical Society of America and, following the classification system of the Society's journal, one volume focuses on speech production, one on speech perception, and one on speech processing.

The idea of the three-volume series originated within the Speech Technical Committee of the Society. The Committee discussed and enthusiastically endorsed the project at the Society's fall 1989 meeting in St. Louis, Missouri, and subsequently chose the editors and editorial boards. A formal proposal for the project was then drafted by the Chair of the Speech Technical Committee and was forwarded to the Executive Council of the Society. The Council gave final approval for the project at the Society's spring 1990 meeting in State College, Pennsylvania.

We have organized each of the three volumes into topical sections, with the papers within each section ordered alphabetically by author. To help guide readers—especially students and nonexperts—we have written editorial commentary for each section. The commentary is intended to provide a brief context for the individual papers, placing them within the history of the discipline. We have also included a topical subject index at the end of each volume, keyed to individual papers. Finally, because the three volumes are so closely interrelated, at the end of each volume we have included the table of contents and the index of each of the other two volumes.

We have worked closely with our editorial boards in selecting the papers that appear in these volumes. The members of the boards were involved in all stages of the selection process, from the initial generation of a list of potential papers to the final decisions on selection. In making the selections, we were guided by the goal of including papers that are important in their own right and, in addition, collectively reflect progress in the field and present a range of viewpoints, approaches, and methodologies. Given the vast literature on speech, and practical constraints on the size of the volumes, the choices were difficult, and many important papers are not included. We can only hope that the volumes, as constituted, will prove useful to the speech community as research on speech communication proceeds.

Bishnu S. Atal
Joanne L. Miller
Raymond D. Kent

© Acoustical Society of America