

Defeating ambient noise: practical approaches for noise reduction and suppression

ICASSP 2006 tutorial

Ivan Tashev, Microsoft Research

Tutorial references

I. Introduction

II. Basics

- [1] D. F. Hoth, "Room noise spectra at subscribers' telephone locations." *Journal of the Acoustical Society of America*, 12:499-504, 1941.
- [2] H. S. Malvar, *Signal Processing with Lapped Transforms*. Boston, MA: Artech House, 1992.
- [3] H. S. Malvar, A modulated complex lapped transform and its applications to audio processing, *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, pp. 1421–1424, March 1999.

III. Single Channel Noise Suppression

- [4] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications*, Principles of Electrical Engineering Series. MIT Press, Cambridge, MA, 1949.
- [5] R. J. McAulay and M. L. Malpass, "Speech enhancement using a soft-decision noise suppression filter," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-28, no. 2, pp. 137–145, Apr. 1980.
- [6] D. L. Wang and J. S. Lim, "The unimportance of phase in speech enhancement," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-30, no. 4, pp. 679–681, Aug. 1982.
- [7] Y. Ephraim, D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-32, no. 6, pp. 1109–1121, Dec. 1984.
- [8] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-33, no. 2, pp. 443–445, Apr. 1985.
- [9] O. Cappé, "Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor," *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 2, pp. 345–349, Apr. 1994.
- [10] P. J. Wolfe and S. J. Godsill, "Simple alternatives to the Ephraim and Malah suppression rule for speech enhancement," in *Proc. 11th IEEE Worksh. Stat. Signal Processing*, 2001, pp. 496–499.

- [11] Martin, R.: "Speech Enhancement Using MMSE Short Time Spectral Estimation with Gamma Distributed Speech Priors ", IEEE ICASSP'02, Orlando, Florida, May 2002.
- [12] R. Martin, C. Breithaupt, "Speech Enhancement in the DFT Domain Using Laplacian Speech Priors." Proc. Intl. Workshop Acoustic Echo and Noise Control (IWAENC), 2003.
- [13] J. Sohn and W. Sung, "A voice activity detector employing soft decision based noise spectrum adaptation," Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, vol. 1, pp. 365–368, 1998.
- [14] J. Sohn, N. S. Kim, W. Sung. „A Statistical Model-Based Voice Activity Detection“, IEEE Signal Processing Letters, vol.6, no. 1. January 1999.
- [15] R. Martin, "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics," IEEE Trans. Speech and Audio Processing, vol. 9, no. 5, pp. 504–512, July 2001.
- [16] Y. Ephraim. "Statistical-Model-Based Speech Enhancement Systems", Proceedings of IEEE, vol. 80, no.10, pp 1526-1555, October 1992.
- [17] H. L. Van Trees, Detection, Estimation and Modulation Theory: Part 1, Detection, Estimation and Linear Modulation Theory, John Wiley & Sons, Inc., New York, 1968.

IV. Directional Microphones

- [18] J. Eargle. *The Microphone Book*. Focal Press, 2001, ISBN 0-240-80445-7.

V. Microphone Arrays

- [19] H. Cox, R. M. Zeskind, T. Kooij, "Practical Supergain", IEEE Trans. Acoust. Speech, Signal Processing, vol. ASSP-34, pp. 393-398, June 1986.
- [20] P. L. Chu. "Superdirective microphone array for a set-top videoconferencing system". IEEE International Conference on Acoustics, Speech, and Signal Processing, April 1977. Volume 1, pp 235-238.
- [21] G. Elko, "Superdirective Microphone Arrays", in Acoustic Signal Processing for Telecommunication, S. L. Gay and J. Benesty, eds. Ch. 10, pp. 181-235, Kluwer Academic Press, 2000.
- [22] R. Martin, A. Petrovsky, T. Lotter, "Planar Superdirective Microphone Arrays for Speech Acquisition in the Car," Proc. Euro. Conf. Speech Communication and Technology (EUROSPEECH), 2001.
- [23] Iv. Tashev, H. S. Malvar. "A new beamformer design algorithm for microphone arrays". Proceedings of International Conference of Acoustic, Speech and Signal Processing ICASSP 2005, Philadelphia, PA, USA, March 2005.
- [24] Iv. Tashev. "Gain Self-Calibration Procedure for Microphone Arrays". Proceedings of International Conference for Multimedia and Expo ICME 2004, Taipei, Taiwan, June 2004.
- [25] Iv. Tashev. "Beamformer Sensitivity to Microphone Manufacturing Tolerances". 19th International Conference Systems for Automation of Engineering and Research SAER 2005, St. Konstantin Resort, Bulgaria, September 2005.

- [26] C. Knapp, G. Carter. "The Generalized Method for Estimation of Time Delay". IEEE Transactions of Acoustics, Speech and Signal Processing, vol. ASSP-24, no. 4, August 1976.
- [27] M. Brandstein, J. Adcock, H. Silverman. "A Robust Method for Speech Signal Time-Delay Estimation in Reverberant Rooms". IEEE International Conference on Acoustics, Speech, and Signal Processing, May 1996.
- [28] S. Birchfield, D.K. Gillmor. "Acoustic source direction by hemisphere sampling". Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, May 2001, vol. 5, pp. 3053-3056.
- [29] Y. Rui, D. Florencio. "New direct approaches to robust sound source localization", Proceedings of International Conference on Multimedia and Expo, July 2003, vol. 1, pp. 737-740.
- [30] Iv. Tashev. "Improving Meetings with Microphone Array Algorithms". Machine Learning Meets the User Interface Workshop, Neural Information Processing Systems NIPS 2003, Whistler, Canada, December 2003.
- [31] O. Frost. "An Algorithm for Linearly Constrained Adaptive Array Processing", Proceedings of IEEE, August 1972, vol. 60, no. 8, pp 926-934.
- [32] L. J. Griffiths, C.W. Jim, "An alternative approach to linear constrained adaptive beamforming", IEE Trans. Antennas Propagat., vol. AP-30, no. 1, pp. 27-34, January 1982.
- [33] H. L. Van Trees, Detection, Estimation and Modulation Theory: Part 4, Optimum Array Processing, John Wiley & Sons, Inc., New York, 2002.
- [34] M. Brandstein, D. Ward. *Microphone Arrays*. Springer, 2001. ISBN 3-540-41953-5.

VI. Advanced Methods

- [35] I. Tashev, J. Droppo, A. Acero. "Suppression Rule for Speech Recognition Friendly Noise Suppressors". Proceedings of International Conference of Digital Signal Processing and Applications DSPA-06, March 2006, Moscow, vol. 2, pp. 634-636.
- [36] H. Drucker, "Speech processing in a high ambient noise environment", IEEE Trans. on Audio Electroacoustics, AU-16(2):165-8, June 1968.
- [37] Iv. Tashev, M. Seltzer, A. Acero. "Microphone Array for Headset with Spatial Noise Suppressor". Proceedings of Ninth International Workshop on Acoustic, Echo and Noise Control IWAENC 2005, Eindhoven, The Netherlands, September 2005.
- [38] Z. Liu, M. Seltzer, A. Acero, Iv. Tashev, Z. Zhang, M. Sinclair. "A Compact Multi-Sensor Headset for Hands-Free Communication". Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, USA, October 2005.
- [39] J. D. Johnston, "Transform coding of audio signals using perceptual noise criteria," IEEE Journal *Selected Areas in Communications*, vol. 6, no. 2, pp. 314-323, 1988.
- [40] D. Tsoukalas, M. Paraskevas, and J. Mourjopoulos, "Speech enhancement using psychoacoustic criteria," in Proc. IEEE ICASSP, 1993, vol. 2, pp. 359-362.
- [41] N. Virag, "Speech enhancement based on masking properties of the auditory system," in Proc. IEEE ICASSP, 1995, vol. 1, pp. 796-799.

- [42] A. Akbari Azirani, R. le Bouquin Jeannès, and G. Faucon, “Optimizing speech enhancement by exploiting masking properties of the human ear,” in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Detroit, MI, 1995, vol. 1, pp. 800–803.
- [43] N. Virag, “Single channel speech enhancement based on masking properties of the human auditory system,” *IEEE Trans. Speech and Audio Processing*, vol. 7, no. 2, pp. 126–137, Mar. 1999.
- [44] P. J. Wolfe and S. J. Godsill, “Towards a perceptually optimal spectral amplitude estimator for audio signal enhancement,” in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Istanbul, 2000, vol. 2, pp. 821–824.