

# Dmitry N. Zotkin

Institute for Advanced Computer Studies (UMIACS)  
AVW Building, University of Maryland, College Park MD 20742 USA  
TEL: (301) 405-8753 E-MAIL: dz@umiacs.umd.edu

## *Curriculum Vitae*

### Education

**09/1996 – 08/2002:** Department of Computer Science, University of Maryland at College Park

- M.S. degree in Computer Science, June 1999. GPA: 4.0.
- Ph.D. degree in Computer Science, August 2002. GPA: 4.0.

**09/1990 – 06/1996:** Moscow Institute of Physics and Technology, Department of RadioTechnics and Cybernetics, Dolgoprudny, Moscow reg., Russia (the leading engineering university in Russia)

- A combined M.S./B.S. degree (cum laude) in Applied Mathematics and Physics, June 1996. Diploma of Engineer-Physicist. Overall GPA: 3.96; Major GPA: 4.0.

### Work experience

**05/2010 – present:** Associate Research Scientist, UMIACS, University of Maryland

- Currently continuing work on acoustic and multimodal scene capture, analysis, and visualization

**01/2004 – 05/2010:** Assistant Research Scientist, UMIACS, University of Maryland

- Developed a fast algorithm for the real-time room impulse response computation.
- Built several spherical/hemispherical/cylindrical microphone array systems, including one with custom embedded FPGA/ADC controller and USB 2.0 digital interface
- Designed a theoretical framework for auditory scene capture, decomposition, representation, and efficient playback in terms of plane waves
- Implemented the plane-wave representation framework in practice in an array and accompanying audio rendering software delivered to Veterans Administration Rehab Center in Atlanta, GA
- Took part in various other projects, including HRTF computation using numerical methods; audio signal enhancement in spherical array framework; acoustic imaging and reverberation control with audio camera; multimodal authentication; and efficient hardware implementations

**08/2002 – 12/2003:** Research Associate, UMIACS, University of Maryland

- Designed and developed hierarchical framework for acoustic signal source localization in rooms.
- Designed and built a prototype system for fast head-related transfer function (HRTF) acquisition, cutting time necessary to measure HRTF from hours to minutes.

- Explored various methods for fast HRTF customization and interpolation using anthropometric body parameters.

**05/2002 – 08/2002:** Faculty Research Assistant, UMIACS, University of Maryland

- Designed and developed a suite of neuromorphic algorithms for independent manipulation of perceptual characteristics of sound (such as pitch and timbre).

**06/1998 – 05/2002:** Research Assistant, UMIACS, University of Maryland

- Developed a real-time acoustic tracking system that was used in combination with video input as a prototype smart video conferencing system.
- Developed a real-time multi-modal tracking system to perform an audiovisual bat tracking in an anechoic room (joint work with Biology and EE Dept.)
- Performed research on human location and orientation tracking for the creation of virtual auditory spaces and on joint audio-video processing for tracking and multi-modal sensor fusion.

**01/2000 – 03/2000:** Internship, ATR Media Integration and Communication Labs, Kyoto, Japan.

- Developed the audio-visual tracking system using a microphone array and an active camera for a “Virtual Metamorphosis” (clone agents) attentive environment system.

**09/1996 – 06/1998:** Teaching Assistant, Department of Computer Science, University of Maryland.

- Duties include holding office hours, designing and grading assignments, projects, and exams, and holding recitation sessions.

**09/1994 – 06/1996:** Research Assistant, Institute for Information Transmission Problems, Russian Academy Of Sciences, Moscow, Russia.

- Designed and implemented several data compression algorithms, investigating algorithm performance on different data sets. The project was undertaken as a part of thesis paper research.

## **Publications**

### **Book Chapters:**

1. "Multimodal tracking for smart videoconferencing and video surveillance", D. N. Zotkin, V. C. Raykar, R. Duraiswami, and L. S. Davis. In *Multimodal Surveillance: Sensors, Algorithms, and Systems*, ed. by Zhigang Zhu and Thomas S. Huang, Artech House Publishers, Norwood, MA, 2007, pp. 141-175.
2. "Signal processing for audio HCI", D. N. Zotkin and R. Duraiswami. In *Handbook on Signal Processing Systems*, ed. by S. S. Bhattacharyya, E. Deprettere, R. Leupers, and J. Takala, Springer, New York, NY, 2010, pp. 243-266.
3. "Capturing and recreating auditory virtual reality", R. Duraiswami, D. N. Zotkin, and A. E. O'Donovan. In *Principles and Applications of Spatial Hearing*, ed. by Y. Suzuki et al., World Scientific Publishing Company, London, UK, 2011, pp. 337-356.

## **Journal Articles:**

1. "Attacking the bottlenecks of backfilling schedulers", P. J. Keleher, D. N. Zotkin, and D. Perkovic. *Cluster Computing*, vol. 3(4), 2000, pp. 245-254.
2. "Audio-visual tracking system for multi-modal interface", D. N. Zotkin, K. Takahashi, T. Yotsukura, S. Morishima, and N. Tetsutani. *The Journal of the IIEEJ (Institute of Image Electronics Engineers of Japan)*, vol. 30(4), 2001, pp. 452-463.
3. "Joint audio-visual tracking using particle filters", D. N. Zotkin, R. Duraiswami, and L. S. Davis. *EURASIP Journal on Applied Signal Processing*, vol. 2002(11), 2002, pp. 1154-1164.
4. "Rendering localized spatial audio in a virtual auditory space", D. N. Zotkin, R. Duraiswami, and L. S. Davis. *IEEE Transactions on Multimedia*, vol. 6(4), 2004, pp. 553-564.
5. "Accelerated speech source localization via a hierarchical search of steered response power", D. N. Zotkin and R. Duraiswami. *IEEE Transactions on Speech and Audio Processing*, vol. 12(5), 2004, pp. 499-508.
6. "Neuromimetic sound representation for percept detection and manipulation", D. N. Zotkin, T. Chi, S. A. Shamma, and R. Duraiswami. *EURASIP Journal on Applied Signal Processing*, vol. 2005(9), 2005, pp. 1350-1364.
7. "Processing of reverberant speech for time-delay estimation", B. Yegananarayana, S. R. M. Prasanna, R. Duraiswami, and D. N. Zotkin. *IEEE Transactions on Speech and Audio Processing*, vol. 13(6), 2005, pp. 1110-1118.
8. "Fast head-related transfer function measurement via reciprocity", D. N. Zotkin, R. Duraiswami, E. Grassi, and N. A. Gumerov. *Journal of the Acoustical Society of America*, vol. 120(4), 2006, pp. 2202-2215.
9. "Fast evaluation of the room transfer function using multipole expansion", R. Duraiswami, D. N. Zotkin, and N. A. Gumerov. *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15(2), 2007, pp. 565-576.
10. "Plane-wave decomposition of acoustical scenes via spherical and cylindrical microphone arrays", D. N. Zotkin, R. Duraiswami, and N. A. Gumerov. *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18(1), 2010, pp. 2-16.
11. "Computation of the head-related transfer function via the fast multipole accelerated boundary element method and its spherical harmonic representation", N. A. Gumerov, A. E. O'Donovan, R. Duraiswami, and D. N. Zotkin. *Journal of the Acoustical Society of America*, vol. 127(1), 2010, pp. 370-386.
12. "Biomimetic multi-resolution analysis for robust speaker recognition", S. K. Nemala, D. N. Zotkin, R. Duraiswami, and M. Elhilali. *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2012(22), 2012.
13. "A symmetric kernel partial least squares framework for speaker recognition", B. V. Srinivasan, Y. Luo, D. Garcia-Romero, D. N. Zotkin, and R. Duraiswami. *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21(7), pp. 1415-1423, 2013.

## **Conference Papers:**

1. "Experimental study of some Cleary-Witten algorithm modifications", Y. M. Shtarkov and D. N. Zotkin. *Proc. 7th International Workshop on Information Theory, St. Petersburg, Russia, June 1995*, pp. 211-214.

2. "Pictorial query trees for query specification in image databases", A. Soffer, H. Samet, and D. N. Zotkin. Proc. ICPR 1998, Brisbane, Australia, August 1998, pp. 919-921.
3. "Sloppiness as a virtue: Job-length estimation and performance in backfilling schedulers", D. N. Zotkin and P. J. Keleher. Proc. 8th HPDC, Redondo Beach, CA, August 1999, pp. 236-243.
4. "A real-time audio-video front-end for multimedia applications", D. N. Zotkin, R. Duraiswami, I. Haritaoglu, and L. S. Davis. [Abstract]. Proc. 138th meeting of the Acoustical Society of America, Columbus, OH [Journal of the Acoustical Society of America, vol. 106(4), p. 2271], October 1999.
5. "Exact solutions for the problem of source location from measured time differences of arrival", R. Duraiswami, D. N. Zotkin, and L. S. Davis. [Abstract]. Proc. 138th meeting of the Acoustical Society of America, Columbus, OH [Journal of the Acoustical Society of America, vol. 106(4), p. 2277], October 1999.
6. "Active source location and beamforming", R. Duraiswami, D. N. Zotkin, E. A. Borovikov, and L. S. Davis. [Abstract]. Proc. 139th meeting of the Acoustical Society of America, Atlanta, GA [Journal of the Acoustical Society of America, vol. 107(5), p. 2790], May 2000.
7. "Smart videoconferencing", D. N. Zotkin, R. Duraiswami, V. Philomin, and L. S. Davis. Proc. ICME 2000, New York, NY, August 2000, pp. 1597-1600.
8. "An audio-video front end for multimedia applications", D. N. Zotkin, R. Duraiswami, L. S. Davis, and I. Haritaoglu. Proc. SMC 2000, Nashville, TN, October 2000, pp. 786-791.
9. "Active speech source localization by a dual coarse-to-fine search", R. Duraiswami, D. N. Zotkin, and L. S. Davis. Proc. ICASSP 2001, Salt Lake City, UT, May 2001, vol. 5, pp. 3309-3312.
10. "Multimodal localization of a flying bat", K. Ghose, D. N. Zotkin, R. Duraiswami, and C. F. Moss. Proc. ICASSP 2001, Salt Lake City, UT, May 2001, vol. 5, pp. 3057-3060.
11. "Multimodal 3-D tracking and event detection via the particle filter", D. N. Zotkin, R. Duraiswami, and L. S. Davis. Proc. IEEE Workshop on Detection and Recognition of Events in Video (in association with ICCV 2001), Vancouver, Canada, July 2001, pp. 20-27.
12. "Multimodal tracking for smart videoconferencing", D. N. Zotkin, R. Duraiswami, H. Nanda, and L. S. Davis. Proc. ICME 2001, Tokyo, Japan, August 2001, pp. 37-40.
13. "Attentive toys", I. Haritaoglu, A. Cozzi, D. Koons, M. Flickner, D. N. Zotkin, and Y. Yacoob. Proc. ICME 2001, Tokyo, Japan, August 2001, pp. 1124-1127.
14. "Efficient evaluation of reverberant sound fields", R. Duraiswami, N. A. Gumerov, D. N. Zotkin, and L. S. Davis. Proc. IEEE WASPAA 2001, New Paltz, NY, October 2001, pp. 203-206.
15. "Creation of virtual auditory spaces", D. N. Zotkin, R. Duraiswami, and L. S. Davis. Proc. IEEE ICASSP 2002, Orlando, FL, May 2002, vol. 2, pp. 2113-2116.
16. "Customizable auditory displays", D. N. Zotkin, R. Duraiswami, and L. S. Davis. Proc. 2002 International Conference on Auditory Displays (ICAD 2002), Kyoto, Japan, July 2002, pp. 167-176.
17. "Virtual audio system customization using visual matching of ear parameters", D. N. Zotkin, R. Duraiswami, L. S. Davis, A. Mohan, and V. Raykar. Proc. ICPR 2002, Quebec City, August 2002, pp. 1003-1006.
18. "A multipole method for computation of acoustics of a convex polyhedral room", N. A. Gumerov, R. Duraiswami, D. N. Zotkin, and L. S. Davis. [Abstract]. Proc. 144th meeting of the Acoustical Society of America, Cancun, Mexico [Journal of the Acoustical Society of America, vol. 112(5), p. 2396], December 2002.
19. "Pitch and timbre manipulations using cortical representation of sound", D. N. Zotkin, S. A. Shamma, P. Ru, R. Duraiswami, and L. S. Davis. Proc. IEEE ICASSP 2003, Hong Kong, April 2003, vol. 5, pp. 517-520. (Paper is reprinted in Proc. ICME 2003, Baltimore, MD, July 2003, vol. 3, pp. 381-384, because of the cancellation of the ICASSP 2003 conference meeting).

20. "Using computer vision to generate customized spatial audio", A. Mohan, R. Duraiswami, D. N. Zotkin, D. DeMenthon, and L. S. Davis. Proc. ICME 2003, Baltimore, MD, July 2003, vol. 3, pp. 57-60 (Invited paper).
21. "Sonification of dynamic choropleth maps: geo-referenced data exploration for the vision-impaired", H. Zhao, C. Plaisant, B. Shneiderman, D. N. Zotkin, and R. Duraiswami. [Abstract]. Proc. ICAD 2003, Boston, MA, July 2003, p. 307.
22. "Extending the visual field via perceptually-processed audio in high-bandwidth user interfaces", R. Duraiswami, D. N. Zotkin, N. Mesgarani, S. A. Shamma, K. Grant, and L. S. Davis. Proc. WMSCI 2003 [Panel on Cognitive Systems Exploratory Efforts], Orlando, FL, July 2003.
23. "HRTF personalization using anthropometric measurements", D. N. Zotkin, J. Hwang, R. Duraiswami, and L. S. Davis. Proc. IEEE WASPAA 2003, New Paltz, NY, October 2003, pp. 157-160.
24. "Interpolation and range extrapolation of HRTFs", R. Duraiswami, D. N. Zotkin, and N. A. Gumerov. Proc. IEEE ICASSP 2004, Montreal, QC, Canada, May 2004, vol. 4, pp. 45-48.
25. "High order spatial audio capture and its binaural head-tracked playback over headphones with HRTF cues", R. Duraiswami, D. N. Zotkin, Z. Li, E. Grassi, N. A. Gumerov, and L. S. Davis. Proc. AES 119th convention, New York, NY, October 2005, preprint #6540.
26. "Plane-wave decomposition analysis for the spherical microphone arrays", R. Duraiswami, Z. Li, D. N. Zotkin, E. Grassi, and N. A. Gumerov. Proc. IEEE WASPAA 2005, New Paltz, NY, October 2005, pp. 150-153.
27. "Frequency independent flexible spherical beamforming via RBF fitting", A. Yerukhimovich, R. Duraiswami, N. A. Gumerov, and D. N. Zotkin. Proc. IEEE ICASSP 2006, Toulouse, France, May 2006, vol. 5, pp. 45-48.
28. "Capture and rendering of spatial sound over headphones", R. Duraiswami, D. N. Zotkin, and A. E. O'Donovan. [Abstract]. Proc. 152nd meeting of the Acoustical Society of America, Honolulu, HI [Journal of the Acoustical Society of America, vol. 120(5), p. 3094], November 2006.
29. "Spherical and hemispherical microphone arrays for capture and analysis of sound fields", R. Duraiswami, Z. Li, D. N. Zotkin, and E. Grassi. [Abstract]. Proc. 152nd meeting of the Acoustical Society of America, Honolulu, HI [Journal of the Acoustical Society of America, vol. 120(5), p. 3225], November 2006.
30. "Head-related transfer functions via the fast multipole accelerated boundary element method", N. A. Gumerov, R. Duraiswami, and D. N. Zotkin. [Abstract]. Proc. 152nd meeting of the Acoustical Society of America, Honolulu, HI [Journal of the Acoustical Society of America, vol. 120(5), p. 3342], November 2006.
31. "Efficient conversion of X.Y surround sound content to binaural head-tracked form for HRTF-enabled playback", D. N. Zotkin, R. Duraiswami, and N. A. Gumerov. Proc. IEEE ICASSP 2007, Honolulu, HI, April 2007, vol. 1, pp. 21-24.
32. "Fast multipole accelerated boundary elements for computation of the head related transfer function", N. A. Gumerov, R. Duraiswami, and D. N. Zotkin. Proc. IEEE ICASSP 2007, Honolulu, HI, April 2007, vol. 1, pp. 165-168.
33. "Room acoustics: Compact representation for multi-source binaural playback", D. N. Zotkin, R. Duraiswami, and N. A. Gumerov. Proc. ICA 2007 (19th International Congress on Acoustics), Madrid, Spain, September 2007, paper # RBA-11-007 (Invited paper).
34. "Sound field decomposition using spherical microphone arrays", D. N. Zotkin, R. Duraiswami, and N. A. Gumerov. Proc. IEEE ICASSP 2008, Las Vegas, NV, April 2008, pp. 277-280.

35. "Imaging concert hall acoustics using visual and audio cameras", A. E. O'Donovan, R. Duraiswami, and D. N. Zotkin. Proc. IEEE ICASSP 2008, Las Vegas, NV, April 2008, pp. 5284-5287.
36. "A spherical microphone array based system for immersive audio scene rendering", A. E. O'Donovan, D. N. Zotkin, and R. Duraiswami. Proc. ICAD 2008, Paris, France, June 2008.
37. "Plane-wave decomposition of a sound scene using a cylindrical microphone array", D. N. Zotkin and R. Duraiswami. Proc. IEEE ICASSP 2009, Taipei, Taiwan, R.O.C., April 2009, pp. 85-88.
38. "Imaging room acoustics with the audio camera", A. E. O'Donovan, R. Duraiswami, N. A. Gumerov, and D. N. Zotkin. [Abstract]. Proc. 157th meeting of the Acoustical Society of America, Portland, OR [Journal of the Acoustical Society of America, vol. 125(4), p. 2544], May 2009.
39. "Regularized HRTF fitting using spherical harmonics", D. N. Zotkin, R. Duraiswami, and N. A. Gumerov. Proc. IEEE WASPAA 2009, New Paltz, NY, October 2009, pp. 257-260.
40. "Capturing and recreating auditory virtual reality", R. Duraiswami, D. N. Zotkin, and A. E. O'Donovan. Proc. 2009 International Workshop on the Principles and Applications of Spatial Hearing, Sendai, Japan, November 2009.
41. "Automatic matched filter recovery via the audio camera", A. E. O'Donovan, R. Duraiswami, and D. N. Zotkin. Proc. IEEE ICASSP 2010, Dallas, TX, March 2010, pp. 2826-2829.
42. "Kernelized Renyi distance for speaker recognition", B. V. Srinivasan, R. Duraiswami, and D. N. Zotkin. Proc. IEEE ICASSP 2010, Dallas, TX, March 2010, pp. 4506-4509.
43. "Audio visual scene analysis using spherical arrays and cameras", A. E. O'Donovan, R. Duraiswami, and D. N. Zotkin. [Abstract]. Proc. 159th meeting of the Acoustical Society of America, Baltimore, MD [Journal of the Acoustical Society of America, vol. 127(3), p. 1979], April 2010.
44. "Partial least squares based speaker recognition system", B. V. Srinivasan, D. N. Zotkin, and R. Duraiswami. Proc. Computational and Biological Learning Society Workshop (The Learning Workshop) 2011, Fort Lauderdale, FL, April 2011.
45. "A partial least squares framework for speaker recognition", B. V. Srinivasan, D. N. Zotkin, and R. Duraiswami. Proc. IEEE ICASSP 2011, Prague, Czech Republic, May 2011, pp. 5276-5279.
46. "Kernel partial least squares for speaker recognition", B. V. Srinivasan, D. Garcia-Romero, D. N. Zotkin, and R. Duraiswami. Proc. INTERSPEECH 2011, Florence, Italy, August 2011, pp. 493-496.
47. "The UMD-JHU 2011 speaker recognition system", D. Garcia-Romero et al. Proc. IEEE ICASSP 2012, Kyoto, Japan, March 2012, pp. 4229-4232.
48. "Kernel regression for head-related transfer function interpolation and spectral extrema extraction", Y. Luo, D. N. Zotkin, H. Daume III, and R. Duraiswami. Proc. IEEE ICASSP 2013, Vancouver, BC, Canada, May 2013, pp. 256-260.
49. "Listening space measurements resolved by direction and spectro-temporally", A. E. O'Donovan, D. N. Zotkin, and R. Duraiswami. Proc. ICA 2013 (21st International Congress on Acoustics), Montreal, QC, Canada [Journal of the Acoustical Society of America, vol. 133(5), p. 3254], June 2013. Manuscript available in Proceedings of Meetings on Acoustics, vol. 19, doi:10.1121/1.4800008.
50. "Statistical analysis of head related transfer function data", Y. Luo, D. N. Zotkin, and R. Duraiswami. Proc. ICA 2013 (21st International Congress on Acoustics), Montreal, QC, Canada [Journal of the Acoustical Society of America, vol. 133(5), p. 3281], June 2013. Manuscript available in Proceedings of Meetings on Acoustics, vol. 19, doi:10.1121/1.4799872.
51. "Chain architecture: An efficient hardware solution for a large microphone array system", D. N. Zotkin, R. Adelman, A. E. O'Donovan, and R. Duraiswami. Proc. ICA 2013 (21st International Congress on Acoustics), Montreal, QC, Canada [Journal of the Acoustical Society of America, vol.

133(5), p. 3302], June 2013. Manuscript available in Proceedings of Meetings on Acoustics, vol. 19, doi:10.1121/1.4800868.

52. "Gaussian process data fusion for heterogeneous HRTF datasets", Y. Luo, D. N. Zotkin, and R. Duraiswami. Proc. IEEE WASPAA 2013, New Paltz, NY, October 2013.
53. "Virtual autoencoder based recommendation system for individualizing head-related transfer functions", Y. Luo, D. N. Zotkin, and R. Duraiswami. Proc. IEEE WASPAA 2013, New Paltz, NY, October 2013.
54. "Gaussian process active-learning for head-related transfer functions", Y. Luo, D. N. Zotkin, and R. Duraiswami. IEEE ICASSP 2014, under review.
55. "Gaussian process models for HRTF based 3-D sound localization", Y. Luo, D. N. Zotkin, and R. Duraiswami. IEEE ICASSP 2014, under review.

## **Service**

- Member of Paper Review Committee, ICAD 2003, Boston, MA, July 2003; WASPAA 2003, New Paltz, NY, October 2003; KBCS 2004, Hyderabad, India, December 2004; ICAD 2005, Limerick, Ireland, July 2005; WMSCI 2005, Orlando, FL, July 2005; WASPAA 2005, New Paltz, NY, October 2005; ICAD 2006, London, UK, June 2006; WMSCI 2006, Orlando, FL, July 2006; ICASSP 2007, Honolulu, HI, April 2007; ICAD 2007, Montreal, Canada, June 2007; WMSCI 2007, Orlando, FL, July 2007; WASPAA 2007, New Paltz, NY, October 2007; ICASSP 2008, Las Vegas, NV, April 2008; ICAD 2008, Paris, France, June 2008; WMSCI 2008, Orlando, FL, July 2008; WMSCI 2009, Orlando, FL, July 2009; WASPAA 2009, New Paltz, NY, October 2009; IMCIC 2010, Orlando, FL, April 2010; CISP 2010, Yantai, China, October 2010; WASPAA 2011, New Paltz, NY, October 2011; and CISP 2012, Chongqing, China, October 2012.
- Auxiliary Reviewer, WOMOT 2003 (workshop on multiple object tracking in association with CVPR 2003), Madison, WI, June 2003; ICCV 2003, Nice, France, October 2003; ICASSP 2005, Philadelphia, PA, March 2005; CVPR 2007, Minneapolis, MN, June 2007; and CVPR 2008, Anchorage, AK, June 2008.
- Reviewer for *ACM Transactions on Applied Perception*; *Computer Vision and Image Understanding*; *Virtual Reality*; *Journal of the Acoustical Society of America*; *JASA Express Letters*; *IEEE Transactions on Wireless Communications*; *Information Fusion*; *IEEE Transactions on Audio, Speech, and Language Processing* (formerly *IEEE Transactions on Speech and Audio Processing*); *IEEE Signal Processing Letters*; *IEEE Transactions on Circuits and Systems I*; *International Journal of Distributed Sensor Networks*; *IEEE Sensors Journal*; and *IEEE Transactions on Signal Processing*.
- Advisor on honor's project for: Joe Jakuta, undergraduate student in Computer Science, University of Maryland, Fall 2002 through Spring 2003, and Jane Hwang, undergraduate student in Computer Science, University of Maryland, Spring 2003.

## **References**

- Available upon request.