June 24, 2019

Kazuyoshi Yoshii, Ph.D.

Speech and Audio Processing Group

Department of Intelligence Science and Technology

Graduate School of Informatics, Kyoto University

Address: Room #412, Research Building #7,

Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan

Tel: 075-753-5386 / Fax:075-753-5977

Mail: yoshii@kuis.kyoto-u.ac.jp

Web: http://sap.ist.i.kyoto-u.ac.jp/members/yoshii/

Education

2005–2008: **Ph.D. student**, Audio Media Processing Group, Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University

2003–2005: **Master student**, Audio Media Processing Group, Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University

1999–2003: **Undergraduate student**, Undergraduate School of Informatics and Mathematical Science, Faculty of Engineering, Kyoto University

Employment

2018–now: **Associate Professor**, Speech and Audio Processing Group, Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University

2017–now: **Team Leader**, Sound Scene Understanding Team, Center for Advanced Intelligence Project (AIP), RIKEN

2014–2018: **Senior Lecturer**, Speech and Audio Processing Group, Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University

2013–2014: **Senior Researcher**, Media Interaction Research Group, Information Technology Research Institute (ITRI), National Institute of Advanced Industrial Science and Technology (AIST)

2008–2013: **Research Scientist**, Media Interaction Research Group, Information Technology Research Institute (ITRI), National Institute of Advanced Industrial Science and Technology (AIST)

Publications

Journal papers

1. Kazuki Shimada, Yoshiaki Bando, Masato Mimura, Katsutoshi Itoyama, Kazuyoshi Yoshii, Tatsuya Kawahara: "Unsupervised Speech Enhancement Based on Multichannel NMF-Informed Beamforming for Noise-Robust Speech Recognition," IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 27, No. 5, pp. 960–971, 2019.

2. Yuta Ojima, Eita Nakamura, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>: "Chord-Aware Automatic Music Transcription Based on Hierarchical Bayesian Integration of Acoustic and Language Models," APSIPA Transactions on Signal and Information Processing, Vol. 7, No. e14, pp. 1–14, 2018.

3. Eita Nakamura, <u>Kazuyoshi Yoshii</u>: "Statistical Piano Reduction Controlling Performance Difficulty," APSIPA Transactions on Signal and Information Processing, Vol. 7, No. e13, pp. 1–12, 2018.

4. Hiroaki Tsushima, Eita Nakamura, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>: "Generative Statistical Models with Self-Emergent Grammar of Chord Sequences," Journal of New Music Research, pp. 1–23, 2018.

5. Kousuke Itakura, Yoshiaki Bando, Eita Nakamura, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, Tatsuya Kawahara: "Bayesian Multichannel Audio Source Separation Based on Integrated Source and Spatial Models," IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 26, No. 4, pp. 831–846, 2018.

6. Yoshiaki Bando, Katsutoshi Itoyama, Masashi Konyo, Satoshi Tadokoro, Kazuhiro Nakadai, <u>Kazuyoshi Yoshii</u>, Tatsuya Kawahara, Hiroshi G. Okuno: "Speech Enhancement Based on Bayesian Low-Rank and Sparse Decomposition of Multichannel Magnitude Spectrograms," IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 26, No. 2, pp. 215–230 2018.

7. Eita Nakamura, <u>Kazuyoshi Yoshii</u>, and Simon Dixon: "Note Value Recognition for Rhythm Transcription Using a Markov Random Field Model for Musical Scores and Performances of Piano Music," IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 25, No.9, pp. 1846–1858, 2017.

8. Eita Nakamura, <u>Kazuyoshi Yoshii</u>, and Shigeki Sagayama: "Rhythm Transcription of Polyphonic Music Based on Merged-Output HMM for Multiple Voices," IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 25, No. 4, pp. 794–806, 2017.

9. Karim Youssef, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Simultaneous Identification and Localization of Immobile and Moving Speakers Based on Binaural Sound Acquisition," Journal of Robotics and Mechatronics, Vol. 29, No. 1, pp. 59–71, 2017.

10. Misato Ohkita, Yoshiaki Bando, Yukara Ikemiya, Eita Nakamura, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Audio-Visual Beat Tracking Based on a State-Space Model for a Robot Dancer Performing with a Human Dancer," Journal of Robotics and Mechatronics, Vol. 29, No. 1, pp. 125–136, 2017.

11. Kouhei Sekiguchi, Yoshiaki Bando, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Layout Optimization of Cooperative Distributed Microphone Arrays Based on Estimation of Source Separation Performance," Journal of Robotics and Mechatronics, Vol. 29, No. 1, pp. 83–93, 2017.

12. Yoshiaki Bando, Hiroshi Saruwatari, Nobutaka Ono, Shoji Makino, Katustoshi Itoyama, Daichi Kitamura, Masaru Ishimura, Moe Takakusaki, Narumi Mae, Kouei Yamaoka, Yutaro Matsui, Yuichi Ambe, Masashi Konyo, Satoshi Tadokoro, <u>Kazuyoshi Yoshii</u>, and Hiroshi G. Okuno: "Low-Latency and High-Quality Two-Stage Human-Voice-Enhancement System for a Hose-Shaped Rescue Robot," Journal of Robotics and Mechatronics, Vol. 29, No. 1, pp. 198–212, 2017.

13. Yukara Ikemiya, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Singing Voice Separation and Vocal F0 Estimation Based on Mutual Combination of Robust Principal Component Analysis and Subharmonic Summation," IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 24, No. 11, pp. 2084–2095, 2016.

14. Tomoyasu Nakano, <u>Kazuyoshi Yoshii</u>, and Masataka Goto: "Musical Similarity and Commonness Estimation Based on Probabilistic Generative Models," International Journal of Semantic

Computing, Vol. 10, No. 1, pp. 27–52, 2016.

15. Izaya Nishimuta, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, and Hiroshi G. Okuno: "Toward a Quizmaster Robot for Speech-Based Multiparty Interaction," Advanced Robotics, Vol. 29, No. 18, pp. 1205–1219, 2015.

16. Akira Maezawa, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, and Hiroshi G. Okuno: "Nonparametric Bayesian Dereverberation of Power Spectrograms Based on Infinite-Order Autoregressive Processes," IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 22, No. 12, pp. 1918–1930, 2014.

17. Matthew E. P. Davies, Philippe Hamel, <u>Kazuyoshi Yoshii</u>, and Masataka Goto: "AutoMashUpper: Automatic Creation of Multi-Song Music Mashups," IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 22, No. 12, pp. 1726–1737, 2014.

18. <u>Kazuyoshi Yoshii</u> and Masataka Goto: "A Nonparametric Bayesian Multipitch Analyzer Based on Infinite Latent Harmonic Allocation," IEEE Transactions on Audio, Speech and Language Processing, Vol. 20,
No. 3, pp. 717–730, 2012.

19. <u>Kazuyoshi Yoshii</u>, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno: "An Efficient Hybrid Music Recommender System Using an Incrementally Trainable Probabilistic Generative Model, " IEEE Transactions on Audio, Speech and Language Processing, Vol. 16, No. 2, pp. 435–447, 2008.

20. <u>Kazuyoshi Yoshii</u>, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno, "Drumix: An Audio Player with Real-Time Drum-Part Rearrangement Functions for Active Music Listening," IPSJ Digital Courier, Vol. 3, pp. 134–144, 2007.

21. <u>Kazuyoshi Yoshii</u>, Masataka Goto, and Hiroshi G. Okuno: "Drum Sound Recognition for Polyphonic Audio Signals by Adaptation and Matching of Spectrogram Templates with Harmonic Structure Suppression," IEEE Transactions on Audio, Speech and Language Processing, Vol. 15, No. 1, pp. 333–345, 2007.

Conference papers

22. Go Shibata, Ryo Nishikimi, Eita Nakamura, <u>Kazuyoshi Yoshii</u>: "Statistical Music Structure Analysis Based on a Homogeneity- and Repetitiveness-Aware Hierarchical Hidden Semi-Markov Model," International Society for Music Information Retrieval Conference (ISMIR 2019), Accepted, November 2019.

23. Adrien Ycart, Andrew McLeod, Emmanouil Benetos, <u>Kazuyoshi Yoshii</u>: "Blending Acoustic and Language Model Predictions for Automatic Music Transcription," International Society for Music Information Retrieval Conference (ISMIR 2019), Accepted, November 2019.

24. Yiming Wu, Tristan Carsault, <u>Kazuyoshi Yoshii</u>: "Automatic Chord Estimation Based on a Frame-wise Convolutional Recurrent Neural Network with Non-Aligned Annotations," European Signal Processing Conference (EUSIPCO 2019), Accepted, September 2019.

25. Kouhei Sekiguchi, Aditya Arie Nugraha, Yoshiaki Bando, <u>Kazuyoshi Yoshii</u>: "Fast Multichannel Source Separation Based on Jointly Diagonalizable Spatial Covariance Matrices," European Signal Processing Conference (EUSIPCO 2019), Accepted, September 2019.

26. Mathieu Fontaine, Aditya Arie Nugraha, Roland Badeau, <u>Kazuyoshi Yoshii</u>, Antoine Liutkus: "Cauchy Multichannel Speech Enhancement with a Deep Speech Prior," European Signal Processing Conference (EUSIPCO 2019), Accepted, September 2019.

27. Ryo Nishikimi, Eita Nakamura, Satoru Fukayama, Masataka Goto, <u>Kazuyoshi Yoshii</u>: "Automatic Singing Transcription Based on Encoder-Decoder Recurrent Neural Networks with a Weakly-Supervised Attention Mechanism," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2019), pp. 161–165, May 2019.

28. Kentaro Shibata, Ryo Nishikimi, Satoru Fukayama, Masataka Goto, Eita Nakamura, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>: "Joint Transcription of Lead, Bass, and Rhythm Guitars Based on a Factorial Hidden Semi-Markov Model," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2019), pp. 236–240, May 2019.

29. Shun Ueda, Kentaro Shibata, Yusuke Wada, Ryo Nishikimi, Eita Nakamura, <u>Kazuyoshi Yoshii</u>: "Bayesian Drum Transcription Based on Nonnegative Matrix Factor Decomposition with a Deep Score Prior," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2019), pp. 456–460, May 2019.

30. Eita Nakamura, Kentaro Shibata, Ryo Nishikimi, <u>Kazuyoshi Yoshii</u>: "Unsupervised Melody Style Conversion," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2019), pp. 196–200, May 2019.

31. Andrew McLeod, Eita Nakamura, <u>Kazuyoshi Yoshii</u>: "Improved Metrical Alignment of MIDI Performance Based on a Repetition-Aware Online-Adapted Grammar," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2019), pp. 186–190, May 2019.

32. Aditya Arie Nugraha, Kouhei Sekiguchi, <u>Kazuyoshi Yoshii</u>: "A Deep Generative Model of Speech Complex Spectrograms," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2019), pp. 905–909, May 2019.

33. Kouhei Sekiguchi, Yoshiaki Bando, <u>Kazuyoshi Yoshii</u>, Tatsuya Kawahara: "Bayesian Multichannel Speech Enhancement with a Deep Speech Prior," Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2018), pp. 1233–1239, November 2018.

34. Eita Nakamura, Ryo Nishikimi, Simon Dixon, <u>Kazuyoshi Yoshii</u>: "Probabilistic Sequential Patterns for Singing Transcription," Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2018), pp. 1905–1912, November 2018.

35. Yusuke Wada, Ryo Nishikimi, Eita Nakamura, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>: "Sequential Generation of Singing F0 Contours from Musical Note Sequences Based on WaveNet," Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2018), pp. 983–989, November 2018.

36. Hiroaki Tsushima, Eita Nakamura, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>: "Interactive Arrangement of Chords and Melodies Based on a Tree-Structured Generative Model," International Society for Music Information Retrieval Conference (ISMIR 2018), pp. 145–151, September 2018.

37. <u>Kazuyoshi Yoshii</u>, Koichi Kitamura, Yoshiaki Bando, Eita Nakamura, Tatsuya Kawahara: "Independent Low-Rank Tensor Analysis for Audio Source Separation," European Signal Processing Conference (EUSIPCO 2018), pp. 1671–1675, September 2018.

38. Eita Nakamura, Emmanouil Benetos, <u>Kazuyoshi Yoshii</u>, Simon Dixon: "Towards Complete Polyphonic Music Transcription: Integrating Multi-Pitch Detection and Rhythm Quantization," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018), pp. 101–105, April 2018.

39. Yoshiaki Bando, Masato Mimura, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, Tatsuya Kawahara: "Statistical Speech Enhancement Based on Probabilistic Integration of Variational Autoencoder and Nonnegative Matrix Factorization," IEEE International Conference on Acoustics, Speech, and Signal

Processing (ICASSP 2018), pp. 716–720, April 2018.

40. <u>Kazuyoshi Yoshii</u>: "Correlated Tensor Factorization for Audio Source Separation," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018), pp. 731–735, April 2018.

41. Kazuki Shimada, Yoshiaki Bando, Masato Mimura, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, Tatsuya Kawahara: "Unsupervised Beamforming Based on Multichannel Nonnegative Matrix Factorization for Noisy Speech Recognition," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018), pp. 5734–5738, April 2018.

42. Hirofumi Inaguma, Masato Mimura, Koji Inoue, <u>Kazuyoshi Yoshii</u>, Tatsuya Kawahara: "An End-to-End Approach to Joint Social Signal Detection and Speech Recognition," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018), pp. 6214–6218, April 2018.

43. Izaya Nishimuta, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, Hiroshi G. Okuno: "Multi-Party Interactions by Robot Quiz Master in Speech-Based Jeopardy!-Like Games," Annual Conference on Computational Science & Computational Intelligence (CSCI 2017), December 2017.

44. Ryo Nishikimi, Eita Nakamura, Masataka Goto, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Scale- and Rhythm-Aware Musical Note Estimation for Vocal F0 Trajectories Based on a Semi-Tatum-Synchronous Hierarchical Hidden Semi-Markov Model," International Society for Music Information Retrieval Conference (ISMIR 2017), pp. 376–382, October 2017.

45. Hiroaki Tsushima, Eita Nakamura, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Function- and Rhythm-Aware Melody Harmonization Based on Tree-Structured Parsing and Split-Merge Sampling of Chord Sequences," International Society for Music Information Retrieval Conference (ISMIR 2017), pp. 502–508, October 2017.

46. Eita Nakamura, <u>Kazuyoshi Yoshii</u>, and Haruhiro Katayose: "Performance Error Detection and Post-Processing for Fast and Accurate Symbolic Music Alignment," International Society for Music Information Retrieval Conference (ISMIR 2017), pp. 347–353, October 2017.

47. <u>Kazuyoshi Yoshii</u>, Eita Nakamura, Katsutoshi Itoyama, and Masataka Goto: "Infinite Probabilistic Latent Component Analysis For Audio Source Separation, "IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2017), pp. 1–6, September 2017.

48. Antoine Liutkus, and <u>Kazuyoshi Yoshii</u>: "A Diagonal Plus Low-Rank Covariance Model For Computationally Efficient Source Separation," IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2017), pp. 1–6, September 2017.

49. Masaya Wake, Yoshiaki Bando, Masato Mimura, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, and Tatsuya Kawahara: "Semi-Blind Speech Enhancement Based On Recurrent Neural Network For Source Separation And Dereverberation," IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2017), pp. 1–6, September 2017.

50. Masato Mimura, Yoshiaki Bando, Kazuki Shimada, Shinsuke Sakai, <u>Kazuyoshi Yoshii</u>, and Tatsuya Kawahara: "Combined Multi-Channel NMF-Based Robust Beamforming for Noisy Speech Recognition," Annual Conference of the International Speech Communication Association (Interspeech 2017), pp. 2451–2455, August 2017.

51. Yuta Ojima, Tomoyasu Nakano, Satoru Fukayama, Jun Kato, Masataka Goto, and Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>: "A Singing Instrument for Real-Time Vocal-Part Arrangement of Music Audio Signals," Sound and Music Computing Conference (SMC 2017), pp. 443–449, July 2017.

52. Yusuke Wada, Yoshiaki Bando, Eita Nakamura, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "An Adaptive Karaoke System that Plays Accompaniment Parts of Music Audio Signals Synchronously

with Users' Singing Voices," Sound and Music Computing Conference (SMC 2017), pp. 110–116, July 2017.

53. Kousuke Itakura, Yoshiaki Bando, Eita Nakamura, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, and Tatsuya Kawahara: "Bayesian Multichannel Nonnegative Matrix Factorization for Audio Source Separation and Localization," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2017), pp. 551–555, March 2017.

54. Yoshiaki Bando, Hiroki Suhara, Motoyasu Tanaka, Tetsushi Kamegawa, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, Fumitoshi Matsuno, and Hiroshi G. Okuno: "Sound-Based Online Localization for an In-Pipe Snake Robot," International Symposium on Safety, Security, and Rescue Robotics (SSRR 2016), pp. 207–213, October 2016.

55. Kouhei Sekiguchi, Yoshiaki Bando, Keisuke Nakamura, Kazuhiro Nakadai, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Online Simultaneous Localization and Mapping of Multiple Sound Sources and Asynchronous Microphone Arrays," IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2016), pp. 1973–1979, October 2016.

56. Koichi Kitamura, Yoshiaki Bando, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Student's t Multichannel Nonnegative Matrix Factorization for Blind Source Separation," IEEE International Workshop on Acoustic Signal Enhancement (IWAENC 2016), pp. 1–5, September 2016.

57. Eita Nakamura, <u>Kazuyoshi Yoshii</u>, and Shigeki Sagayama: "Rhythm Transcription of MIDI Performances Based on a Merged-Output HMM for Multiple Voices," Sound and Music Computing Conference (SMC 2016), pp. 338–343, September 2016.

58. Eita Nakamura, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Rhythm Transcription of MIDI Performances Based on Hierarchical Bayesian Modelling of Repetition and Modification of Musical Note Patterns," European Signal Processing Conference (EUSIPCO 2016), pp. 1946–1950, August 2016.

59. Kousuke Itakura, Yoshiaki Bando, Eita Nakamura, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "A Unified Bayesian Model of Time-Frequency Clustering and Low-Rank Approximation for Multi-Channel Source Separation," European Signal Processing Conference (EUSIPCO 2016), pp. 2280–2284, August 2016.

60. Yoshiaki Bando, Katsuyoshi Itoyama, Masashi Konyo, Satoshi Tadokoro, Kazuhiro Nakadai, <u>Kazuyoshi Yoshii</u>, and Hiroshi G. Okuno: "Variational Bayesian Multi-Channel Robust NMF for Human-Voice Enhancement with a Deformable and Partially-Occluded Microphone Array," European Signal Processing Conference (EUSIPCO 2016), pp. 1018–1022, August 2016.

61. Yuta Ojima, Eita Nakamura, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "A Hierarchical Bayesian Model of Chords, Pitches, and Spectrograms for Multipitch Analysis," The 17th International Society for Music Information Retrieval Conference (ISMIR 2016), pp. 309–315, August 2016.

62. Ryo Nishikimi, Eita Nakamura, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Musical Note Estimation for F0 Trajectories of Singing Voices Based on a Bayesian Semi-Beat-Synchronous HMM," The 17th International Society for Music Information Retrieval Conference (ISMIR 2016), pp. 461–467, August 2016.

63. Tomoyasu Nakano, Daichi Mochihashi, <u>Kazuyoshi Yoshii</u>, and Masataka Goto: "Musical Typicality: How Many Similar Songs Exist?," The 17th International Society for Music Information Retrieval Conference (ISMIR 2016), pp. 695–701, August 2016.

64. <u>Kazuyoshi Yoshii</u>, Katsutoshi Itoyama, and Masataka Goto: "Student's t Nonnegative Matrix Factorization and Positive Semidefinite Tensor Factorization for Single-Channel Audio Source

Separation," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2016), pp. 51–55, March 2016.

65. Eita Nakamura, Masatoshi Hamanaka, Keiji Hirata, and <u>Kazuyoshi Yoshii</u>: "Tree-Structured Probabilistic Model of Monophonic Written Music Based on the Generative Theory of Tonal Music," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2016), pp. 276–280, March 2016.

66. Masataka Goto, <u>Kazuyoshi Yoshii</u>, and Tomoyasu Nakano: "Songle Widget: Making Animation and Physical Devices Synchronized with Music Videos on the Web," IEEE International Symposium on Multimedia (ISM 2015), pp. 85–88, December 2015.

67. Tomoyasu Nakano, <u>Kazuyoshi Yoshii</u>, and Masataka Goto: "Musical Similarity and Commonness Estimation Based on Probabilistic Generative Models," IEEE International Symposium on Multimedia (ISM 2015), pp. 197–204, December 2015.

68. <u>Kazuyoshi Yoshii</u>, Katsutoshi Itoyama, and Masataka Goto: "Infinite Superimposed Discrete All-Pole Modeling for Source-Filter Decomposition of Wavelet Spectrograms," The 16th International Society for Music Information Retrieval Conference (ISMIR 2015), pp. 86–92, October 2015.

69. Yoshiaki Bando, Katsutoshi Itoyama, Masashi Konyo, Satoshi Tadokoro, Kazuhiro Nakadai, <u>Kazuyoshi Yoshii</u>, and Hiroshi G. Okuno: "Human-Voice Enhancement Based on Online RPCA for a Hose-Shaped Rescue Robot with a Microphone Array," IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR 2015), pp. 1–6, October 2015.

70. Akira Maezawa, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, and Hiroshi G. Okuno: "Unified Inter- and Intra-Recording Duration Model for Multiple Music Audio Alignment," IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2015), pp. 1–5, October 2015.

71. Karim Youssef, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Identification and Localization of One or Two Concurrent Speakers in a Binaural Robotic Context," IEEE International Conference on Systems, Man, and Cybernetics (SMC 2015), pp. 407–412, October 2015.

72. Misato Ohkita, Yoshiaki Bando, Yukara Ikemiya, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Audio-Visual Beat Tracking Based on a State-Space Model for a Music Robot Dancing with Humans," IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2015), pp. 5555–5560, September 2015.

73. Kouhei Sekiguchi, Yoshiaki Bando, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Optimizing the Layout of Multiple Mobile Robots for Cooperative Sound Source Separation," IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2015), pp. 5548–5554, September 2015.

74. Yoshiaki Bando, Kazuhiro Nakadai, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, Masashi Konyo, Hiroshi G. Okuno, and Satoshi Tadokoro: "Microphone-Accelerometer Based 3D Posture Estimation for a Hose-Shaped Rescue Robot," IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2015), pp. 5580–5586, September 2015.

75. Kousuke Itakura, Izaya Nishimuta, Yoshiaki Bando, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "Bayesian Integration of Sound Source Separation and Speech Recognition: A New Approach to Simultaneous Speech Recognition," The 16th Annual Conference of the International Speech Communication Association (Interspeech 2015), pp. 736–740, September 2015.

76. Ayaka Dobashi, Yukara Ikemiya, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "A Music Performance Assistance System Based on Vocal, Harmonic, and Percussive Source Separation and Content Visualization for Music Audio Signals," The 12th Sound and Music Computing Conference (SMC

2015), pp. 99–104, July 2015.

77. Tsubasa Fukuda, Yukara Ikemiya, Katsutoshi Itoyama, and <u>Kazuyoshi Yoshii</u>: "A Score-Informed Piano Tutoring System with Mistake Detection and Score Simplification," The 12th Sound and Music Computing Conference (SMC 2015), pp. 105–110, July 2015.

78. Tatsunori Hirai, Yukara Ikemiya, <u>Kazuyoshi Yoshii</u>, Tomoyasu Nakano, Masataka Goto, and Shigeo Morishima: "Automatic Singing Voice to Music Video Generation via Mashup of Singing Video Clips," The 12th Sound and Music Computing Conference (SMC 2015), pp. 153–159, July 2015.

79. Yukara Ikemiya, <u>Kazuyoshi Yoshii</u>, and Katsutoshi Itoyama: "Singing Voice Analysis and Editing Based on Mutually Dependent F0 Estimation and Source Separation," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2015), pp. 574–578, May 2015.

80. Satoshi Maruo, <u>Kazuyoshi Yoshii</u>, Katsutoshi Itoyama, Matthias Mauch, and Masataka Goto: "A Feedback Framework for Improved Chord Recognition Based on NMF-Based Approximate Note Transcription," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2015), pp. 196–200, May 2015.

81. Yoshiaki Bando, Takuma Otsuka, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, Yoko Sasaki, Satoshi Kagami, and Hiroshi G. Okuno: "Challenges in Deploying a Microphone Array to Localize and Separate Sound Sources in Real Auditory Scenes," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2015), pp. 723–727, May 2015.

82. Yoshiaki Bando, Takuma Otsuka, Ikkyu Aihara, Hiromitsu Awano, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, and Hiroshi G. Okuno: "Recognition of In-Field Frog Chorusing Using Bayesian Nonparametric Microphone Array Processing," AAAI Workshop on Computational Sustainability, pp. 2–6, January 2015.

83. Izaya Nishimuta, <u>Kazuyoshi Yoshii</u>, Katsutoshi Itoyama, and Hiroshi G. Okuno: "Development of a Robot Quizmaster with Auditory Functions for Speech-Based Multiparty Interaction," IEEE/SICE International Symposium on System Integration (SII 2014), pp. 328–333, December 2014.

84. Izaya Nishimuta, Naoki Hirayama, <u>Kazuyoshi Yoshii</u>, Katsutoshi Itoyama, and Hiroshi G. Okuno: "A Robot Quizmaster that can Localize, Separate, and Recognize Simultaneous Utterances for a Fastest-Voice-First Quiz Game," IEEE-RAS International Conference on Humanoid Robots (Humanoids 2014), pp. 967–972, November 2014.

85. Yoshiaki Bando, Katsutoshi Itoyama, Satoshi Tadokoro, Masashi Konyo, Kazuhiro Nakadai, and <u>Kazuyoshi Yoshii</u>: "A Sound-Based Online Method for Estimating the Time-Varying Posture of a Hose-Shaped Robot," IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR 2014), pp. 1–6, October 2014.

86. Akira Maezawa, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, and Hiroshi G. Okuno: "Bayesian Audio Alignment Based on a Unified Model of Music Composition and Performance," The 15th International Society for Music Information Retrieval Conference (ISMIR 2014), pp. 233–238, October 2014.

87. Taro Masuda, <u>Kazuyoshi Yoshii</u>, Masataka Goto, and Shigeo Morishima: "Spotting a Query Phrase from Polyphonic Music Audio Signals Based on Semi-Supervised Nonnegative Matrix Factorization," The 15th International Society for Music Information Retrieval Conference (ISMIR 2014), pp. 227–232, October 2014.

88. Shoto Sasaki, <u>Kazuyoshi Yoshii</u>, Tomoyasu Nakano, Masataka Goto, Shigeo Morishima: "LyricsRadar: A Lyrics Retrieval System Based on Latent Topics of Lyrics," The 15th International Society for Music Information Retrieval Conference (ISMIR 2014), pp. 585–590, October 2014.

89. <u>Kazuyoshi Yoshii</u>, Hiromasa Fujihara, Tomoyasu Nakano, Masataka Goto: "Cultivating Vocal Activity Detection for Music Audio Signals in a Circulation-Type Crowdsourcing Ecosystem," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2014), pp. 624–628, May 2014.

90. Tomohiko Nakamura, Hirokazu Kameoka, <u>Kazuyoshi Yoshii</u>, Masataka Goto: "Timbre Replacement of Harmonic and Drum Components for Music Audio Signals," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2014), pp. 7520–7524, May 2014.

91. Tomoyasu Nakano, <u>Kazuyoshi Yoshii</u>, Masataka Goto: "Vocal Timbre Analysis Using Latent Dirichlet Allocation and Cross-Gender Vocal Timbre Similarity," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2014), pp. 5239–5243, May 2014.

92. <u>Kazuyoshi Yoshii</u>, Ryota Tomioka, Daichi Mochihashi, Masataka Goto: "Beyond NMF: Time-Domain Audio Source Separation without Phase Reconstruction," The 14th International Society for Music Information Retrieval Conference (ISMIR 2013), pp. 369–374, November 2013.

93. Satoru Fukayama, <u>Kazuyoshi Yoshii</u>, Masataka Goto: "ChordSequenceFactory: A Chord Arrangement System Modifying Factorized Chord Sequence Probabilities," The 14th International Society for Music Information Retrieval Conference (ISMIR 2013), pp. 457–462, November 2013.

94. Matthew E. P. Davies, Philippe Hamel, <u>Kazuyoshi Yoshii</u>, Masataka Goto: "AutoMashUpper: An Automatic Multi-Song Mashup System," The 14th International Society for Music Information Retrieval Conference (ISMIR 2013), pp. 575–580, November 2013.

95. Philippe Hamel, Matthew E. P. Davies, <u>Kazuyoshi Yoshii</u>, Masataka Goto: "Transfer Learning in MIR: Sharing Learned Latent Representations for Music Audio Classification and Similarity," The 14th International Society for Music Information Retrieval Conference (ISMIR 2013), pp. 9–14, November 2013.

96. Yoko Sasaki, <u>Kazuyoshi Yoshii</u>, and Satoshi Kagami: "A Nested Infinite Gaussian Mixture Model for Recognizing Known and Unknown Audio Events," The 14th International Workshop on Image and Audio Analysis for Multimedia Interactive services (WIA²MIS 2013), July 2013.

97. <u>Kazuyoshi Yoshii</u>, Ryota Tomioka, Daichi Mochihashi, and Masataka Goto: "Infinite Positive Semidefinite Tensor Factorization for Source Separation of Mixture Signals," The 30th International Conference on Machine Learning (ICML 2013), pp. 576–584, June 2013.

98. <u>Kazuyoshi Yoshii</u> and Masataka Goto: "Infinite Kernel Linear Prediction for Joint Estimation of Spectral Envelope and Fundamental Frequency," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013), pp. 463–467, May 2013.

99. <u>Kazuyoshi Yoshii</u> and Masataka Goto: "Infinite Composite Autoregressive Models for Music Signal Analysis," The 13th International Society for Music Information Retrieval Conference (ISMIR 2012), pp. 79–84, October 2012.

100. Masataka Goto, Jun Ogata, <u>Kazuyoshi Yoshii</u>, Hiromasa Fujihara, Matthias Mauch, and Tomoyasu Nakano: "PodCastle and Songle: Crowdsourcing-Based Web Services for Retrieval and Browsing of Speech and Music Content," The 1st Workshop on Crowdsourcing Web Search (CrowdSearch 2012), pp. 1–6, April 2012.

101. <u>Kazuyoshi Yoshii</u> and Masataka Goto: "Unsupervised Music Understanding Based on Nonparametric Bayesian Models," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), pp. 5353–5356, March 2012.

102. Masataka Goto, Jun Ogata, <u>Kazuyoshi Yoshii</u>, Hiromasa Fujihara, Matthias Mauch, and Tomoyasu Nakano: "PodCastle and Songle: Crowdsourcing-Based Web Services for Spoken Document

Retrieval and Active Music Listening," Information Theory and Applications Workshop (ITA 2012), pp. 298–299, February 2012.

103. Kazuyoshi Yoshii and Masataka Goto: "A Unified Probabilistic Model of Note Combinations and Chord Progressions," The 4th International Workshop on Music and Machine Learning (MML 2011), December 2011.

104. Kazuyoshi Yoshii and Masataka Goto, "A Vocabulary-Free Infinity-Gram Model for Nonparametric Bayesian Chord Progression Analysis," The 12th International Society for Music Information Retrieval Conference (ISMIR 2011), pp. 645–650, October 2011.

105. Masataka Goto, Kazuyoshi Yoshii, Hiromasa Fujihara, Matthias Mauch, and Tomoyasu Nakano: "Songle: A Web Service for Active Music Listening Improved by User Contributions," The 12th International Society for Music Information Retrieval Conference (ISMIR 2011), pp. 311–316, October 2011.

106. Matthias Mauch, Hiromasa Fujihara, Kazuyoshi Yoshii, and Masataka Goto: "Timbre and Melody Features for the Recognition of Vocal Activity and Instrumental Solos in Polyphonic Music," The 12th International Society for Music Information Retrieval Conference (ISMIR 2011), pp. 233–238, October 2011.

107. Kazuyoshi Yoshii and Masataka Goto: "Infinite Latent Harmonic Allocation: A Nonparametric Bayesian Approach to Multipitch Analysis," The 11th International Society for Music Information Retrieval Conference (ISMIR 2010), pp. 309–314, August 2010.

108. Kazuyoshi Yoshii and Masataka Goto: "Continuous pLSI and Smoothing Techniques for Hybrid Music Recommendation," The 10th International Society for Music Information Retrieval Conference (ISMIR 2009), pp. 339–344, October 2009.

109. Kazuyoshi Yoshii and Masataka Goto: "MusicCommentator: Generating Comments Synchronized with Musical Audio Signals by a Joint Probabilistic Model of Acoustic and Textual Features," The 8th International Conference on Entertainment Computing (ICEC 2009), pp. 85–97, September 2009.

110. Kazumasa Murata, Kazuhiro Nakadai, Kazuyoshi Yoshii, Ryu Takeda, Toyotaka Torii, Hiroshi G. Okuno, Yuji Hasegawa, and Hiroshi Tsujino: "A Robot Uses Its Own Microphone to Synchronize Its Steps to Musical Beats While Scatting and Singing," 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008), pp. 2459–2464, September 2008.

111. Takeshi Mizumoto, Ryu Takeda, Kazuyoshi Yoshii, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno: "A Robot Listens to Music and Counts Its Beats Aloud by Separating Music from Counting Voice," 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008), pp. 1538–1543, September 2008.

112. Kazuyoshi Yoshii and Masataka Goto: "Music Thumbnailer: Visualizing Musical Pieces in Thumbnail Images Based on Acoustic Features," The 9th International Conference on Music Information Retrieval (ISMIR 2008), pp. 211–216, September 2008.

113. Kouhei Sumi, Katsutoshi Itoyama, Kazuyoshi Yoshii, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno: "Automatic Chord Recognition Based on Probabilistic Integration of Chord Transition and Bass Pitch Estimation," The 9th International Conference on Music Information Retrieval (ISMIR 2008),
pp. 39–44, September 2008.

114. Kazumasa Murata, Kazuhiro Nakadai, Kazuyoshi Yoshii, Ryu Takeda, Toyotaka Torii, Hiroshi G. Okuno, Yuji Hasegawa, and Hiroshi Tsujino: "A Robot Singer With Music Recognition Based on

Real-Time Beat Tracking," The 9th International Conference on Music Information Retrieval (ISMIR 2008), pp. 199–204, September 2008.

115. Takehiro Abe, Katsutoshi Itoyama, <u>Kazuyoshi Yoshii</u>, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno: "Analysis-and-Manipulation Approach to Pitch and Duration of Musical Instrument Sounds without Distorting Timbral Characteristics," The 11th International Conference on Digital Audio Effects (DAFX 2008), pp. 249–256, September 2008.

116. <u>Kazuyoshi Yoshii</u>, Kazuhiro Nakadai, Toyotaka Torii, Yuji Hasegawa, Hiroshi Tsujino, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno: "A Biped Robot that Keeps Steps in Time with Musical Beats while Listening to Music with Its Own Ears," 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007), pp. 1743–1750, October 2007.

117. <u>Kazuyoshi Yoshii</u>, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno: "Improving Efficiency and Scalability of Model-Based Music Recommender System Based on Incremental Training," The 8th International Conference on Music Information Retrieval (ISMIR 2007), pp. 89–94, September 2007.

118. <u>Kazuyoshi Yoshii</u>, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno: "Hybrid Collaborative and Content-Based Music Recommendation Using Probabilistic Model with Latent User Preferences," The 7th International Conference on Music Information Retrieval (ISMIR 2006), pp. 296–301, October 2006.

119. <u>Kazuyoshi Yoshii</u>, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno: "An Error Correction Framework Based on Drum Pattern Periodicity for Improving Drum Sound Detection," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2006), Vol. V, pp. 237–240, May 2006.

120. <u>Kazuyoshi Yoshii</u>, Masataka Goto, and Hiroshi G. Okuno: "INTER:D: A Drum Sound Equalizer for Controlling Volume and Timbre of Drums," The 2nd European Workshop on the Integration of Knowledge, Semantic and Digital Media Technologies (EWIMT 2005), pp. 205–212, November 2005.

121. <u>Kazuyoshi Yoshii</u>, Masataka Goto, and Hiroshi G. Okuno: "Automatic Drum Sound Description for Real-World Music Using Template Adaptation and Matching Methods," The 5th International Conference on Music Information Retrieval (ISMIR 2004), pp. 184–191, October 2004.

122. <u>Kazuyoshi Yoshii</u>, Masataka Goto, and Hiroshi G. Okuno: "Drum Sound Identification for Polyphonic Music Using Template Adaptation and Matching Methods," ISCA Tutorial and Research Workshop on Statistical and Perceptual Audio Processing (SAPA 2004), October 2004.

Awards

1. IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR 2015), Most Innovative Paper Award, 2015/10/19.
2. IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR 2015), People's Choice Demo Award, 2015/10/19.
3. The 10th Annual Music Information Retrieval Evaluation eXchange (MIREX 2014), Winner of the Singing Voice Separation Track, 2014/10/14.
4. The 14th International Society for Music Information Retrieval Conference (ISMIR 2013) Best Oral Presentation Award, 2013/11/7.

5. Award for Entertainment Robots and Systems (IROS 2008 NTF Award) Nomination Finalist, 2008/9/24.
6. Award for Entertainment Robots and Systems (IROS 2008 NTF Award) Nomination Finalist, 2008/9/24.
7. The 1st Annual Music Information Retrieval Evaluation eXchange (MIREX 2005), Winner of the Audio Drum Detection Track (Best-in-Class Award), 2005/9/14.