Marc Delcroix
December 5th 1980 - Belgian – Married

Speech signal processing researcher with 10 years+ experience in speech enhancement and speech recognition research

1-7-18 Kizugadai, kizugawa-shi, Tel: +81 (0)80-6216-1898
Kyoto-fu, 6190225, Japan marc.delcroix@lab.ntt.co.jp

EXPERIENCE

NTT Communication Science Laboratories Kyoto

Senior Research scientist at Media Information Laboratory, Signal Processing Research Group	2015-present
Research scientist at Media Information Laboratory, Signal Processing Research Group	2012-2015
Research Associate at Media Information Laboratory, Signal Processing Research Group	2010-2012

Research on noise robust automatic speech recognition, including speech enhancement, acoustic modelling, acoustic model adaptation

Waseda University Tokyo

Visiting Lecturer 2015-present
Teaching Information theory

Pixela Corporation Osaka 2008 - 2010
Software development engineer

Software development for satellite and terrestrial digital TV applications for PCs.

NTT Communication Science Laboratories Kyoto

Research Associate at the Media Information Laboratory, Signal Processing Research Group 2007-2008

Research on efficient interconnection of speech enhancement pre-processor and speech recogniser

Intern at the Media Information Laboratory, Signal Processing Research Group 2004-2007

Research on speech enhancement algorithms for reducing the effect of reverberation on speech signals

IBM Montpellier summer 2002
Intern.

Equant Japan Tokyo summer 2001
Intern.

OTHER ACADEMIC AND PROFESSIONAL ACTIVITIES

ASRU 2017, publication chair
REVERB Challenge organization
REVERB Workshop publication chair
Senior Member of IEEE and Acoustic Society of Japan (ASJ)

EDUCATION

PhD Hokkaido University 2004-2007
Graduate School of Information Science and Technology, Division of Media and Network Technologies
in collaboration with NTT Communication Science Laboratories.

Ph.D. Dissertation on *Blind dereverberation of speech based on multi-channel linear prediction*,
Design of algorithms for suppressing the effect of reverberation on speech signals captured by multiple microphones. Proved the efficiency of the developed method for acoustic and speech recognition applications.

European grant, Vulcanus in Japan training program 2003-2004
Intensive Japanese study and internship at NTT Communication Science Laboratories.

Master of Telecommunications Engineering Ecole Centrale Paris 2000-2002
2-years exchange program leading to a double degree.
Main project: *Synthesis of carbon nano-tubes by chemical vapour deposition.*

Master of Telecommunications Engineering Brussels Free University 1998-2003
Graduated with Great honours.
Master Thesis: *Charging, billing and security in UMTS*, in collaboration with Alcatel.

SKILLS

Research: Speech and Audio signal processing, automatic speech recognition, speech enhancement, deep learning

Computer : Programming in C, C++, Python, Matlab. Linux/Windows OS.

Languages : French (mother tongue), English (fluent), Japanese (fluent, JPLT and BJT, level 2)

MAIN PUBLICATIONS/PRESENTATIONS

For the full list of publications, please see my Google scholar page at
https://scholar.google.co.jp/citations?user=QG8aWfIAAAAJ&hl=en

Journal papers (peer-reviewed)
1. Delcroix, M., Hikichi, T. and Miyoshi, M., "Blind dereverberation algorithm for speech signals based on multi-channel linear prediction," Acoustical Science and Technology, vol. 26, no. 5, pp. 432-439, 2005.
2. Delcroix, M., Hikichi, T. and Miyoshi, M., "On a blind speech dereverberation algorithm using multi-channel linear prediction," IEICE Trans. Fundamentals, vol. E89-A, no. 10, pp.2837-2846, 2006.
3. Hikichi, T., Delcroix, M. and Miyoshi, M., "Speech dereverberation algorithm using transfer function estimates with overestimated order," Acoustical Science and Technology, vol. 27, no. 1, pp. 28-35, 2006.
4. Hikichi T., Delcroix M. and Miyoshi, M., "Inverse filtering for speech dereverberation less sensitive to noise and room transfer function fluctuations," EURASIP J. APS, vol.2007, Article-ID 34013, 2007.
5. Delcroix, M., Hikichi, T. and Miyoshi, M., "Precise dereverberation using multi-channel linear prediction," IEEE Trans. ASLP, vol. 15, no. 2, pp 430-440, 2007.
6. Delcroix, M., Hikichi, T. and Miyoshi, M., "Dereverberation and denoising using multi-channel linear prediction," IEEE Trans. ASLP, vol. 15, no. 6, pp 1791-1801, 2007.
7. Nakatani, T., Juang, B. H., Yoshioka, T., Kinoshita, K., Delcroix, M., Miyoshi, M., "Speech dereverberation based on maximum likelihood estimation with time-varying Gaussian source model,"IEEE Trans. ASLP, vol. 16, no. 8, pp. 1512-1527, 2008.
8. Miyoshi, M., Delcroix, M. and Kinoshita, K., "Calculating Inverse filters for speech dereverberation,"IEICE Trans. Fundamentals, vol. E91-A, no. 6, pp.1303-1309, 2008.
9. Delcroix, M., Nakatani, T. and Watanabe, S., "Static and dynamic variance compensation for recognition of reverberant speech with dereverberation preprocessing," IEEE Trans. ASLP, vol. 17, no. 2, pp 324-334, 2009.
10. Kinoshita, K., Delcroix, M., Nakatani, T. and Miyoshi, M., "Suppression of late reverberation effect on speech signal using long-term multiple-step linear prediction,"IEEE Trans. ASLP, vol. 17, no. 4, pp. 534-545, 2009.
11. Souden, M., Delcroix, M., Kinoshita, K., Yoshioka, T. and Nakatani, T., "Noise power spectral density tracking: A maximum likelihood perspective," IEEE Signal Processing Letters, vol. 19, no. 8, pp. 495-498, Aug. 2012.
12. Yoshioka, T., Sehr, A., Delcroix, M., Kinoshita, K., Maas, R., Nakatani, T., and Kellermann, W., "Making machines understand us in reverberant rooms: Robustness against reverberation for automatic speech recognition," IEEE Signal Processing Magazine, vol. 29, no. 6, pp. 114-126, Nov. 2012.
13. Delcroix, M., Watanabe, S., Nakatani, T. and Nakamura, A., "Cluster-based dynamic variance adaptation for interconnecting speech enhancement pre-processor and speech recognizer," Computer Speech and Language, Elsevier, vol. 27, no. 1, pp. 350-368, 2013.
14. Delcroix, M., Kinoshita, K., Nakatani, T., Araki, S., Ogawa, A., Hori, T., Watanabe, S., Fujimoto, M., Yoshioka, T., Oba, T., Kubo, Y., Souden, M., Hahm, S.-J. and Nakamura, A., "Speech recognition in living rooms: Integrated speech enhancement and recognition system based on spatial, spectral & temporal modeling of sounds," Computer Speech and Language, Elsevier, vol. 27, no. 3, pp. 851-873, 2013.
15. Nakatani, T., Araki, S., Yoshioka, T., Delcroix, M. and Fujimoto, M. "Dominance Based Integration of Spatial and Spectral Features for Speech Enhancement," IEEE Trans. ASLP, vol.21, no.12, pp.2516-2531, Dec. 2013.
16. Souden, M., Kinoshita, K., Delcroix, M. and Nakatani, T., "Location Feature Integration for Clustering-Based Speech Separation in Distributed Microphone Arrays," IEEE/ACM Trans. ASLP, vol.22, no.2, pp.354-367, Feb. 2014.
17. Delcroix, M., Yoshioka, T., Ogawa, A., Kubo, Y., Fujimoto, M., Ito, N., Kinoshita, K., Espi, M., Araki, S., Hori, T., and Nakatani, T., "Strategies for distant speech recognition in reverberant environments," EURASIP Journal on Advances in Signal Processing, 2015.
18. Delcroix, M., Ogawa, A., Hahm, S.-J., Nakatani, T. and Nakamura, A., "Differenced maximum mutual information criterion for robust unsupervised acoustic model adaptation," Computer Speech and Language, Elsevier, vol. 36, pp. 24-41, 2016.
19. Kinoshita, K., Delcroix, M., Gannot, S., Habets, E., Haeb-Umbach, R., Kellermann, W., Leutnant, V., Maas, R., Nakatani, T., Raj, B., Sehr, A. and Yoshioka, T., "A summary of the REVERB challenge: state-of-the-art and remaining challenges in reverberant speech processing research," EURASIP Journal on Advances in Signal Processing, 2016.
20. Higuchi, T., Ito, N., Araki, S., Yoshioka, T., Delcroix, M. and Nakatani, T., "Online MVDR Beamformer

Based on Complex Gaussian Mixture Model With Spatial Prior for Noise Robust ASR," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 25, no. 4, pp. 780-793, April 2017.

Book and Book chapters:
1. Miyoshi, M., Delcroix, M., Kinoshita, K., Yoshioka, T., Nakatani, T. and Hikichi, T., "Inverse filtering for speech dereverberation without the use of room acoustics information," in Speech Dereverberation, Naylor, P. A. and Gaubitch, N. (eds.), Springer, pp. 271-310, 2010.
2. Delcroix, M., Nakatani, T. and Watanabe, S., "Variance Compensation for Recognition of Reverberant Speech with Dereverberation Preprocessing," in Robust Speech Recognition of Uncertain or Missing Data, Haeb-Umbach, R. and Kolossa, D. (eds.), Springer, pp. 225-255, 2011.
3. Delcroix, M., Yoshioka, T., Ito, N., Ogawa, A., Kinoshita, K., Fujimoto, M., Higuchi, T., Araki, S. and Nakatani, T. "Multi-channel speech enhancement approaches for DNN-based far-field speech recognition," in New Era for Robust Speech Recognition: Exploiting Deep Learning, Watanabe, S., Delcroix, M., Metze, F. and Hershey, J. (eds) Springer, 2017.
4. Watanabe, S., Delcroix, M., Metze, F. and Hershey, J. (eds), "New Era for Robust Speech Recognition: Exploiting Deep Learning," Springer, 2017.

Invited talk/Tutorials:
1. Delcroix, M., Yoshioka, T., Ogawa, A., Kubo, Y., Fujimoto, M., Ito, N., Kinoshita, K., Espi, M., Araki, S. and Nakatani, T., "Defeating reverberation: Advanced dereverberation and recognition techniques for hands-free speech recognition," Invited paper to 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP), pp. 522-526, Dec. 2014.
2. Delcroix, M., Watanabe, S., "Recent advances in distance speech recognition," Tutorial at Interspeech 2016
3. Watanabe, S., Xiao, X. and Delcroix, M., "Multi-Microphone Speech Recognition," Tutorial at APSIPA 2016.

Conference papers (peer-reviewed):
1. Delcroix, M., Hikichi, T. and Miyoshi, M., "Dereverberation of speech signals based on linear prediction," Proc. of International Conference on Spoken Language Processing (ICSLP'04), Vol. 2, pp. 877-880, 2004.
2. Delcroix, M., Hikichi, T. and Miyoshi, M., "Improvement of AR parameters estimation for blind dereverberation," Proc. of the 2005 Joint Workshop on Hands-Free Speech Communication and Microphone Arrays (HSCMA'05), pp. d–5–d–6, 2005.
3. Delcroix, M., Hikichi, T. and Miyoshi, M., "Improved blind dereverberation performance by using spatial information," Proc. Interspeech'05, pp. 2309-2312, 2005.
4. Delcroix, M., Hikichi, T. and Miyoshi, M., "On the use of LIME dereverberation algorithm in an acoustic environment with a noise source," Proc. of the 2006 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'06), I, pp. 825-828, 2006.
5. Kinoshita, K., Delcroix, M., Nakatani, T., and Miyoshi, M., "Multi-step linear prediction based speech enhancement in noisy reverberant environment," Proc. of Interspeech'07, pp.854-857, 2007.
6. Delcroix, M., Watanabe, S., and Nakatani, T. "Combined static and dynamic variance adaptation for efficient interconnection of speech enhancement pre-processor with speech recognizer," Proc. of the 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'08) pp. 4073-4076, 2008.
7. Kolossa, D., Araki, S., Delcroix, M., Nakatani, T. , Orglmeister, R. and Makino, S., "Missing Feature Speech Recognition in a Meeting Situation with Maximum SNR Beamforming," Proc. IEEE International Symposium on Circuits and Systems (ISCAS'08), pp. 3218 -3221, 2008.
8. Delcroix, M., Watanabe, S., Nakatani, T. and Nakamura, A., "Discriminative approach to dynamic variance adaptation for noisy speech recognition," Proc. Of Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA'11), pp. 7-12, 2011.
9. Delcroix, M., Kinoshita, K., Nakatani, T., Araki, S., Ogawa, A., Hori, T., Watanabe, S., Fujimoto, M., Yoshioka, T., Oba, T., Kubo, Y., Souden, M., Hahm, S.-J. and Nakamura, A., "Speech Recognition in the Presence of Highly Non-Stationary Noise Based on Spatial, Spectral and Temporal Speech/Noise Modeling Combined with Dynamic Variance Adaptation," Proc. of CHiME International Workshop on Machine Listening in Multisource Environments, pp. 12-17, 2011. (Best performance on the The PASCAL 'CHiME' Speech Separation and Recognition Challenge)
10. Delcroix, M., Ogawa, A., Watanabe, S., Nakatani, T. and Nakamura, A., "Discriminative feature transforms using differenced maximum mutual information," Proc. of the 2012 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'12) pp. 4753-4756, 2012.
11. Delcroix, M., Ogawa, A., Nakatani, T. and Nakamura, A., "Dynamic variance adaptation using differenced maximum mutual information," Proc. of MLSLP, 2012.
12. Delcroix, M., Ogawa, A., Hahm, S.-J., Nakatani, T. and Nakamura, A., "Unsupervised discriminative adaptation using differenced maximum mutual information based linear regression," Proc. of the 2013 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'13), pp. 7888-7892, 2013.
13. Delcroix, M., Kubo, Y., Nakatani, T. and Nakamura, A., "Is speech enhancement pre-processing still

relevant when using deep neural networks for acoustic modeling?" Proc. of Interspeech 2013, pp. 2992 -2996, 2013.

14. Delcroix, M., Yoshioka, T., Ogawa, A., Kubo, Y., Fujimoto, M., Ito, N., Kinoshita, K., Espi, M., Nakatani, T. and Nakamura, A., "Linear prediction-based dereverberation with advanced speech enhancement and recognition technologies for the REVERB challenge," proc. of REVERB challenge workshop, 2014. (Best performance on the recognition task of the REVERB Challenge)
15. Delcroix, M., Kinoshita, K., Hori, T. and Nakatani, T., "Context adaptive deep neural networks for fast acoustic model adaptation," Proc. of the 2015 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'15), pp. 4535–4539, April 2015.
16. Yoshioka, T., Ito, N., Delcroix, M., Ogawa, A., Kinoshita, K., Fujimoto, M., Yu, C., Fabian, W. J., Espi, M., Higuchi, T., Araki, S., and Nakatani, T., "The NTT CHiME-3 system: advances in speech enhancement and recognition for mobile multi-microphone devices," Proc. of IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), 2015. (Best performance on the recognition task of the CHiME Challenge 3, Best paper honourable mention)
17. Delcroix, M., Kinoshita, Yu, C., Ogawa, A., Yoshioka, T. and Nakatani, T., "Context adaptive deep neural networks for fast acoustic model adaptation in noisy conditions," Proc. of 2016 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'16) pp. 5270-5274, March 2016.
18. Kundu, S., Mantena, G. V., Qian, Y., Tan, T., Delcroix, M. and Sim, K. C., "Joint acoustic factor learning for robust deep neural network based automatic speech recognition," Proc. 2016 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'16), pp. 5025-5029, March 2016
19. Zmolikova, K., Karafiat, M., Vesel, K., Delcroix, M., Watanabe, S., Burget, L. and Cernock, H., "Data selection by sentence summarization in mismatch condition training," Interspeech 2016.
20. Delcroix, M., Kinoshita, Ogawa, A., Yoshioka, T., Dung Tran and Nakatani, T., "Context adaptive neural network for rapid adaptation of deep CNN based acoustic models," Submitted to the Interspeech, 2016.
21. Ochiai, T., Delcroix, M., Kinoshita, K., Ogawa, A., Asami, T., Katagiri, S. and Nakatani, T., "Cumulative Moving Averaged Bottleneck Speaker Vectors for Online Speaker Adaptation of CNN-based Acoustic Models," ICASSP 2017.
22. Huemmer, C, Delcroix, M., Ogawa, A., Kinoshita, K., Nakatani, T. and Kellermann, W., "Online environmental adaptation of CNN-based acoustic models using spatial diffuseness features," ICASSP 2017.
23. Zmolikova, K., Delcroix, M., Kinoshita, K., Higuchi, T., Ogawa, A., Nakatani, T., "Speaker-aware neural network based beamformer for speaker extraction in speech mixtures," Interspeech 2017.
24. Zmolikova, K., Delcroix, M., Kinoshita, K., Higuchi, T., Ogawa, A. and Nakatani, T., "Learning speaker representation for neural network based multichannel speaker extraction," ASRU 2017.
25. Araki, S., Ono, N., Kinoshita, K. and Delcroix, M., "Meeting recognition with asynchronous distributed microphone array," ASRU 2017.

Awards:
1. 2005 presentation award for the poster presentation entitled "Multi-microphone dereverberation method based on linear prediction" presented at the young researcher meeting of the Kansai section of the Acoustic Society of Japan.
2. 2006 IEEE Kansai section student paper award for the paper entitled "On the use of LIME dereverberation algorithm in an acoustic environment with a noise source".
3. 2007 Sato Paper Awards from the acoustic society of Japan (ASJ) for the paper entitled "Speech dereverberation algorithm using transfer function estimates with overestimated order"
4. 2016 Acoustic Society of Japan, Awaya award.

Patents:
Marc Delcroix contributed to more than 25 patents