

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

APPLE INC.,
Petitioner,

v.

Zentian Limited
Patent Owner.

Case IPR2023-00033
Patent No. 7,587,319

**DECLARATION OF DELIANG WANG, Ph.D., IN SUPPORT OF
PATENT OWNER'S PRELIMINARY RESPONSE**

TABLE OF CONTENTS

I.	Introduction.....	1
A.	Engagement.....	1
B.	Background and qualifications.....	1
C.	Materials considered.....	3
II.	Relevant legal standards	4
A.	Person of ordinary skill in the art.....	4
B.	Burden of proof	6
C.	Claim construction	6
D.	Obviousness.....	7
III.	Overview of the '319 Patent.....	8
IV.	The '319 Patent's limitations 46(b) and (d)	9

I, DeLiang Wang, Ph.D., do hereby declare as follows:

I. Introduction

A. Engagement

1. I have been retained by Patent Owner Zentian Limited (“Zentian” or “Patent Owner”) to provide my opinions with respect to Zentian’s Preliminary Response to the Petition in *Inter Partes* Review proceeding IPR2023-00033, with respect to U.S. Pat. 7,587,319. I am being compensated for my time spent on this matter. I have no interest in the outcome of this proceeding and the payment of my fees is in no way contingent on my providing any particular opinions.

2. As part of this engagement, I have also been asked to provide my technical review, analysis, insights, and opinions regarding the materials cited and relied upon by the Petition, including the prior art references and the supporting Declaration of Mr. Schmandt.

3. The statements made herein are based on my own knowledge and opinions.

B. Background and qualifications

4. My full qualifications, including my professional experience and education, can be found in my Curriculum Vitae, which includes a complete list of my publications, and is attached as Ex. A to this declaration.

5. I have spent my professional and academic career as a researcher in the field of speech processing and machine learning (including deep learning). I am currently a University Distinguished Scholar and Professor in the Department of Computer Science and Engineering at Ohio State University.

6. I received B.S. in 1983 and M.S. in 1986 from Peking (Beijing) University, both in computer science. I received a Ph.D. in computer science in 1991 from the University of Southern California.

7. I have received numerous awards and honors, including the U.S. Office of Naval Research Young Investigator Award, the Best Paper Awards from the Institute of Electrical and Electronics Engineers (“IEEE”) Computational Intelligence Society and the IEEE Signal Processing Society, and the Helmholtz Award from the International Neural Network Society. I am Co-Editor-in-Chief of *Neural Networks*, a premier journal in the field of neural networks and deep learning, and also served as President of the International Neural Network Society.

8. I am an IEEE Fellow and an ISCA Fellow. I have published 175 articles in major scientific journals and more than 250 papers in leading conference proceedings. In addition, I have supervised 29 graduate students who earned their PhDs in computer science and engineering, including those currently employed by

leading IT companies performing ASR and related work. More details are given in the attached *curriculum vitae*.

9. I am a recognized expert in the field of robust ASR (automatic speech recognition) technology, including scientific methods, and algorithm development and testing. Robust ASR aims to develop ASR algorithms that can suppress, or remain unaffected by, background interference (such as noise). ASR algorithms developed in my laboratory have been recognized as some of the best in the world; our algorithms achieved the highest recognition rate in the CHiME-2 challenge in 2016, in the CHiME-4 challenge in 2020, and the LibriCSS challenge in 2021. My research contributions and achievements in the fields of speech processing were featured in the March 2017 issue of *IEEE Spectrum*, the most circulated technical magazine in the world. I am one of the most published authors in peer-reviewed scientific journals in the fields of speech and audio processing.

C. Materials considered

10. In the course of preparing my opinions, I have reviewed and am familiar with the '319 patent, including its written description, figures, and claims. I have also reviewed and am familiar with the Petition in this proceeding, the supporting Declaration of Mr. Schmandt, and the relied upon prior art, including Thelen, Bailey, and Chen. I have also reviewed the materials cited in this declaration. My opinions

are based on my review of these materials as well as my 30 years of experience, research, and education in the field of art.

II. Relevant legal standards

11. I am not an attorney. I offer no opinions on the law. But counsel has informed me of the following legal standards relevant to my analysis here. I have applied these standards in arriving at my conclusions.

A. Person of ordinary skill in the art

12. I understand that an analysis of the claims of a patent in view of prior art has to be provided from the perspective of a person having ordinary skill in the art at the time of invention of the '319 patent. I understand that I should consider factors such as the educational level and years of experience of those working in the pertinent art; the types of problems encountered in the art; the teachings of the prior art; patents and publications of other persons or companies; and the sophistication of the technology. I understand that the person of ordinary skill in the art is not a specific real individual, but rather a hypothetical individual having the qualities reflected by the factors discussed above.

13. I understand that the Petition applies a priority date of February 4, 2002, for the challenged claims, Pet. 5, and I apply the same date.

14. I further understand that the Petition defines the person of ordinary skill in the art at the time of the invention as having had a master's degree in computer

engineering, computer science, electrical engineering, or a related field, with at least two years of experience in the field of speech recognition, or a bachelor's degree in the same fields with at least four years of experience in the field of speech recognition. The Petition adds that further education or experience might substitute for the above requirements. I do not dispute the Petition's assumptions at this time, and my opinions are rendered on the basis of the same definition of the ordinary artisan set forth in the Petition.

15. I also note, however, that an ordinarily skilled engineer at the time of the invention would have been trained in evaluating both the costs and benefits of a particular design choice. Engineers are trained (both in school and through general experience in the workforce) to recognize that design choices can have complex consequences that need to be evaluated before forming a motivation to pursue a particular design choice, and before forming an expectation of success as to that design choice. In my opinion, anyone who did not recognize these realities would not be a person of ordinary skill in the art. Thus, a person who would have simply formed design motivations based only on the premise that a particular combination of known elements would be possible would not be a person of ordinary skill regardless of their education, experience, or technical knowledge. Likewise, a person who would have formed design motivations as to a particular combination of known

elements based only on the premise that the combination may provide some benefit, with no consideration of the relevance of the benefit in the specific context and in relation to the costs or disadvantages of that combination, would also not have been a person of ordinary skill in the art, regardless of their education, experience, or technical knowledge. In my opinion, a person of ordinary skill in the art would have been deliberative and considered, rather than impulsive.

16. Throughout my declaration, even if I discuss my analysis in the present tense, I am always making my determinations based on what a person of ordinary skill in the art (“POSA”) would have known at the time of the invention. Based on my background and qualifications, I have experience and knowledge exceeding the level of a POSA, and am qualified to offer the testimony set forth in this declaration.

B. Burden of proof

17. I understand that in an *inter partes* review the petitioner has the burden of proving a proposition of unpatentability by a preponderance of the evidence.

C. Claim construction

18. I understand that in an *inter partes* review, claims are interpreted based on the same standard applied by Article III courts, *i.e.*, based on their ordinary and customary meaning as understood in view of the claim language, the patent’s description, and the prosecution history viewed from the perspective of the ordinary artisan. I further understand that where a patent defines claim language, the

definition in the patent controls, regardless of whether those working in the art may have understood the claim language differently based on ordinary meaning.

D. Obviousness

19. I understand that a patent may not be valid even though the invention is not identically disclosed or described in the prior art if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious to a person having ordinary skill in the art in the relevant subject matter at the time the invention was made.

20. I understand that, to demonstrate obviousness, it is not sufficient for a petition to merely show that all of the elements of the claims at issue are found in separate prior art references or even scattered across different embodiments and teachings of a single reference. The petition must thus go further, to explain how a person of ordinary skill would combine specific prior art references or teachings, which combinations of elements in specific references would yield a predictable result, and how any specific combination would operate or read on the claims. Similarly, it is not sufficient to allege that the prior art could be combined, but rather, the petition must show why and how a person of ordinary skill would have combined them.

21. I understand that where an alleged motivation to combine relies on a particular factual premise, the petitioner bears the burden of providing specific

support for that premise. I understand that obviousness cannot be shown by conclusory statements, and that the petition must provide articulated reasoning with some rational underpinning to support its conclusion of obviousness. I also understand that skill in the art and “common sense” rarely operate to supply missing knowledge to show obviousness, nor does skill in the art or “common sense” act as a bridge over gaps in substantive presentation of an obviousness case.

III. Overview of the ’319 Patent

22. U.S. Patent 7,587,319, titled “Speech recognition circuit using parallel processors,” is directed to an improved speech recognition circuit that “uses parallel processors for processing the input speech data in parallel.” Ex. 1001, 1:4-6. The ’319 patent teaches multiple processors “arranged in groups or clusters,” with each group or cluster of processors connected to one of several “partial lexical memories” that “contains part of the lexical data.” Ex. 1001, 2:64-3:3. “Each lexical tree processor is operative to process the speech parameters using a partial lexical memory and the controller controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.” Ex. 1001, 3:3-7. The ’319 patent further teaches that the invention “provides a circuit in which speech recognition processing is performed in parallel by groups of processors operating in parallel in which each group accesses a common memory of lexical data.” Ex. 1001, 3:44-47.

IV. The '319 Patent's limitations 46(b) and (d)

23. Limitation 46(b) recites “a plurality of lexical memories containing in combination complete lexical data for word recognition, each lexical memory containing part of said complete lexical data.” A person of ordinary skill would understand limitation 46(b) to require (1) multiple lexical memories; (2) with each lexical memory containing **part** of the “complete lexical data,” and (3) the lexical memories together containing “complete lexical data for word recognition.”

24. Limitation 46(d) recites: “said [plurality of] processors being arranged in groups of processors, each group of processors being connected to a lexical memory.” A person of ordinary skill would understand limitation 46(d) to require (1) multiple processors; (2) arranged in groups of processors; (3) with multiple groups each containing said multiple processors; and (4) with each group of multiple processors being connected to a lexical memory.

25. Together, limitations 46(b) and 46(d) require multiple groups of processors, each group containing a plurality of processors, and each group respectively connected to one of multiple lexical memories, with each lexical memory containing part of the complete lexical data, and all of the lexical memories combined containing all of the complete lexical data.

26. Figure 2 of the patent, annotated below, illustrates this architecture by showing two groups of lexical tree processors, with each group containing multiple

processors 1- k , and each group of processors connected to a dedicated “acoustic model memory,” such that there are at least two acoustic model memories for at least two groups of processors. Fig. 2 (annotated).

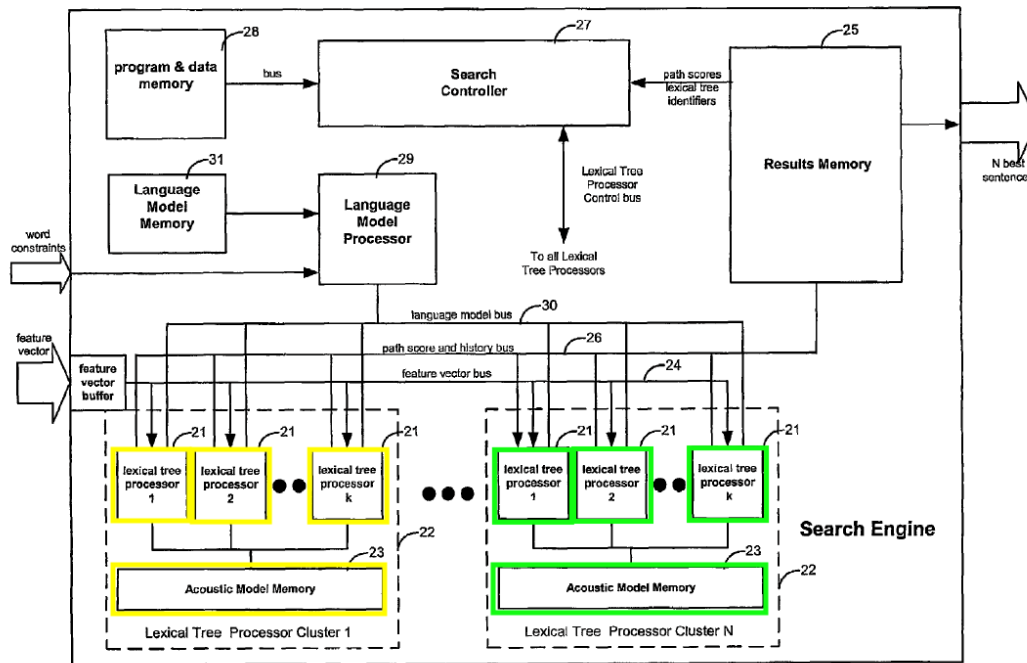


Fig 2

27. I also note that the '319 patent distinguishes that said architecture from two prior known alternative designs, which the patent describes as less advantageous. The '319 patent teaches that: “By providing a plurality of processors in a group with a common memory, flexibility in the processing is provided without being bandwidth limited by the interface to the memory that would occur if only a single memory were used for all processors. The arrangement is more flexible than the parallel processing arrangement in which each processor only has access to its

own local memory and requires fewer memory interfaces (i.e. chip pins).” Ex. 1001, 3:50-58. I understand from those teachings that the patent distinguishes its design from (1) a one-memory-to-all-processors design, which it describes as bandwidth (access) limited in the processor to memory interface; and (2) a one-memory-to-one-processor design, which would require more memory interfaces and afford less flexibility in controlling the extent of parallel processing of speech parameters.

28. I understand that in order to meet the processor to memory architecture of the challenged claims, the Petition relies on the combination of Thelen, Bailey, and Chen. Pet. 23-33, 38-46.

29. The Petition acknowledges that Thelen’s recognizers 1-3 all access a single memory storage 340. Pet at 24, 30. Thelen’s Figure 3 illustrates a single storage 340 serving each of the recognizers (alleged processors) labeled as Rec 1, Rec 2, Rec 3, and Testing Recognizer, and the single storage 340 containing all of the “models” 341-348. Ex. 1030, Fig. 3 (Petition’s annotations).

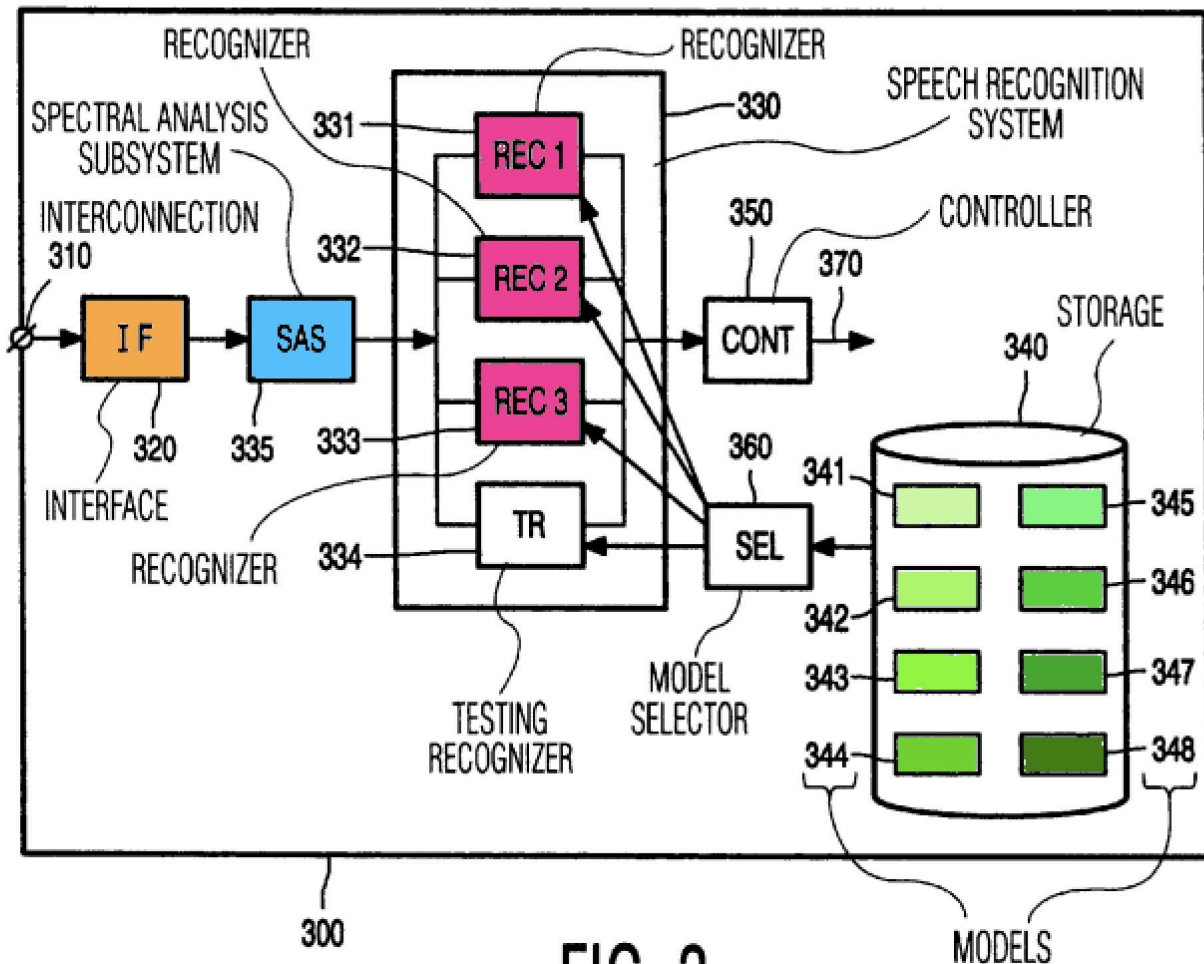


FIG. 3

30. As Thelen's Fig 3 illustrates, Thelen neither teaches limitation 46(b) nor 46(d), much less the two limitations combined.

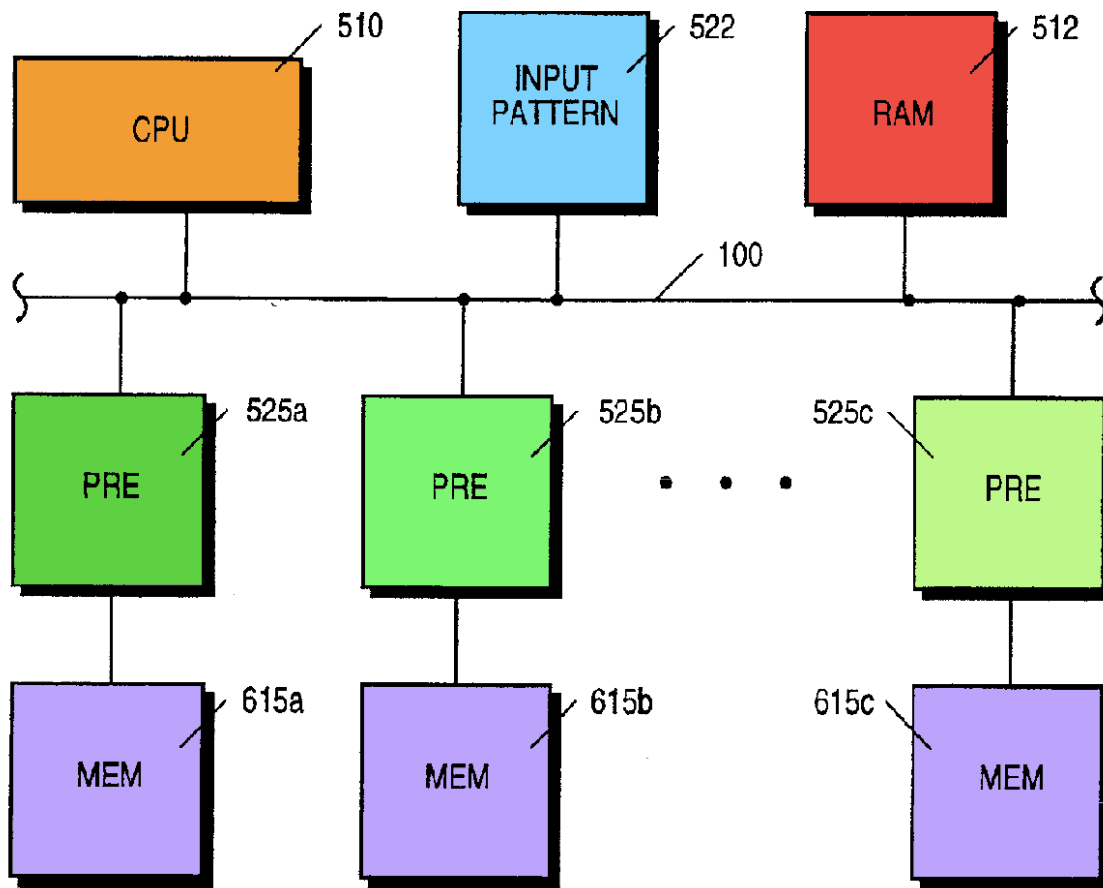
31. With respect to limitation 46(b), Thelen does not teach multiple lexical memories, with each lexical memory containing part of the complete lexical data, and the lexical memories together containing "complete lexical data for word recognition." Instead, Thelen teaches the one-memory-to-*all*-processors (or one-

memory-to-all-recognizers) design that the '319 patent expressly distinguished from and improved upon.

32. With respect to limitation 46(d), Thelen does not teach a plurality of processors “arranged in groups of processors, each group of processors being connected to a lexical memory.” Rather, Thelen teaches individual “recognizers,” which are never arranged in groups, and which are all connected to a single complete storage memory 340, rather than one of several lexical memories containing partial lexical data, as the claims require. It is useful to keep in mind that Thelen’s teachings mainly targeted huge vocabulary speech recognition, and the main innovation was to employ a plurality of large vocabulary speech recognizers, each targeting a subset of the huge vocabulary.

33. I understand that the Petition also relies on two embodiments from Bailey—the embodiments of Bailey’s Figures 3A and 3B. Pet. 30-33, 36-38.

34. In the embodiment of Bailey’s Fig. 3A (below), each individual processor (or pattern recognition engine, labeled PRE) is connected to its own dedicated memory (labeled MEM) such that there is one processor to one memory.



Ex. 1031, Fig. 3 (Petition’s annotated version).

35. Bailey teaches, “each PR engine 525a-525c [is] interfaced with its own private memory 615a-615c.” Ex. 1031, 11:1-3. The embodiment of Bailey’s Fig. 3A is thus the one-memory-to-one-processor design distinguished by the ’319 patent. Accordingly, while Bailey’s Fig. 3A teaches a “plurality” of memories, it does not teach the memory architecture of the ’319 patent, but rather teaches the prior art design that the ’319 patent explicitly distinguished as inferior.

36. The design of Bailey's Figure 3B is identical to the other design distinguished by the '319 patent. Bailey's Figure 3B illustrates a design in which all of the processors PRE are connected to a single memory MEM 615. Ex. 1031, Fig. 3B, 11:1-4.

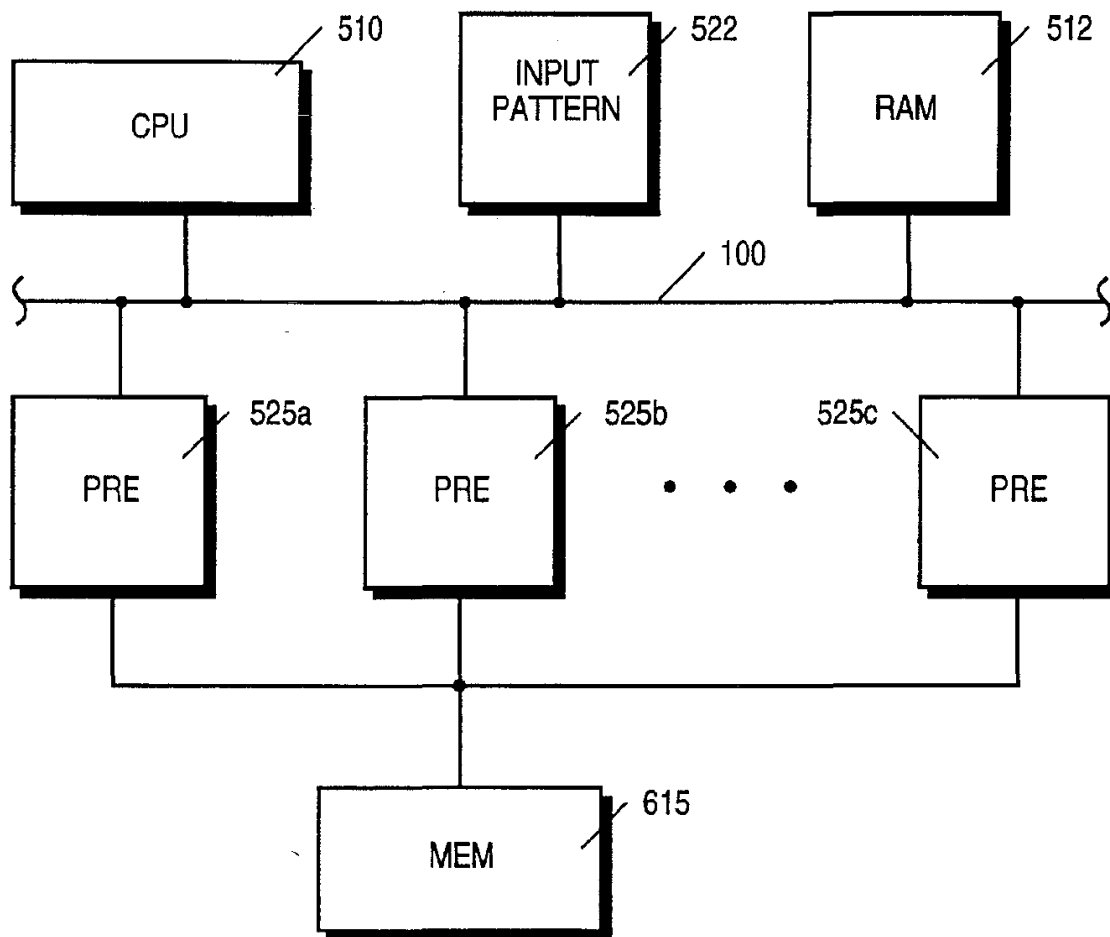


FIG. 3B

37. The design of Bailey's Figure 3B is thus the same as Thelen's, *i.e.*, one in which "only a single memory [is] used for all processors," as distinguished by the '319 patent.

38. As the '319 patent explained, the design of Thelen and Bailey's Fig. 3B is bandwidth limited compared to the '319 patent's claimed architecture. Ex. 1001, 3:44-58. Thus, while Bailey's Fig. 3B (like Thelen) teaches connecting a plurality of processors to one memory, Bailey's does not teach the design of the challenged claims. Bailey's Fig 3B does not teach multiple groups (clusters) of processors, each group containing multiple processors, and each group of processors connected to a shared lexical memory that contains a portion of the overall lexical data, with the combined lexical memories containing all of the lexical data.

39. Thus, neither of Bailey's two embodiments teaches the claimed innovative processor-to-memory architecture of the challenged claims.

40. I understand the Petition also relies on Chen for the disclosure of multiple clusters of processors, each cluster containing multiple processors connected to a cluster shared memory. Pet. at 38-40.

41. I note that Chen is not a reference in the field of speech recognition. Rather, Chen teaches a general-purpose parallel-computing architecture, designed to provide high performance for all sorts of computational tasks. Chen has no teachings

directed to speech recognition systems; in particular, Chen provides no teachings regarding “lexical memories,” much less “lexical memories containing in combination complete lexical data for word recognition,” with “each lexical memory containing part of said complete lexical data.” Pet. 38-40 (admitting Chen does not teach lexical memories or lexical data). Thus, Chen also does not teach the claimed design of limitations 46(b) and (d).

42. I understand that the Petition relies on a series of modifications to each of Thelen, Bailey, and Chen to arrive at what the Petition calls “a common sense extension” of the actual teachings of those three references. Pet. 44. According to the Petition, Thelen, Bailey, and Chen would have been reconfigured such that:

- (1) Thelen’s recognizers would be modified to be in “groups,” Pet. 40;
- (2) Each “group” of Thelen’s recognizers would consist of multiple recognizers for the same subject matter context, e.g., multiple recognizers for “sports,” and multiple recognizers for “health,” etc., Pet. 42-43;
- (3) Each newly created “group” of Thelen’s recognizers would be connected to a dedicated shared memory, such that there would be multiple dedicated shared memories in total, Pet. 40;
- (4) Each of the dedicated shared group memories would be a lexical memory, Pet. 40;

(5) Each dedicated shared group memory would only store the recognition model for the same specific “context” as the recognizers in the group, Pet. 42-43; and

(6) The shared group memories together would contain complete lexical data, Pet. 40, 43.

43. I note, however, that each modification in the above theory relies on an arrangement that is not taught in Thelen, Bailey, or Chen.

44. First, Thelen does not teach “groups” of multiple recognizers for the same subject matter, as the Petition’s steps (1) and (2) above require. Rather, Thelen teaches an embodiment in which “the number of recognition models corresponds to the number of recognizers; each recognizer being associated with an exclusive recognition model in a fixed one-to-one relationship.” Ex. 1030, 7:43-46, 7:30-33. In this embodiment, there is only one recognizer for each subject matter, not multiple recognizers per subject matter, as the Petition’s theory requires.

45. Alternatively, Thelen teaches a “preferred” embodiment in which “the system comprises more models than active recognizers[.]” Ex. 1030, 7:47-48. In this “preferred” embodiment, there is one recognizer for multiple recognition models, not multiple recognizers for one subject matter, as the Petition’s theory requires. Thus, Thelen’s preferred embodiment teaches the opposite of what the Petition’s

theory would require, i.e., an architecture in which multiple recognition models are grouped to one processor (recognizer), rather than multiple processors being grouped to one recognition model (handling a single subject matter) as described in the Petition's step (2).

46. Second, as I stated earlier, none of Thelen, Bailey, or Chen teaches a design in which multiple groups of speech recognition processors, with multiple processors per group, are each connected to one dedicated lexical memory, as steps (3) and (4) of the Petition's theory (above) require. Rather, Thelen and Bailey merely teach the prior art designs the '319 patent expressly distinguished, and Chen provides no teachings regarding speech recognition or lexical memories at all.

47. Third, none of Thelen, Bailey, or Chen teaches storing one recognition model per lexical memory, with multiple lexical memories combined storing all recognition models, as steps (5) and (6) of the Petition's theory (above) require. Rather, Thelen teaches storing all recognition models in one storage memory 340. With respect to the embodiment of Fig. 3B, Bailey also teaches storing all lexical data in a single storage memory. As for the embodiment of Fig. 3A, Bailey also teaches storing all of its lexical data in each of the dedicated private memories, such that each private memory 615a, b, and c contains all of the lexical data, not part of the lexical data. Ex. 1031, 13:12-13 (describing "the private memory 615 which

holds all the prototype patterns.”); *id.* 11:22-23 (“Each PR engine contains its own library of prototype patterns 615*a* to 615*c* respectively.”); *id.* at 11:34-37 (“Refer to FIG. 3(B) which illustrates the system substantially as described with reference to FIG. 3(A) bus [*sic*] wherein each separate PR engine shares the same prototype library 615.”). Bailey never teaches that the private memories in Fig. 3A contain portions of the overall lexical data. And Chen teaches nothing regarding lexical memories or lexical data at all.

48. Accordingly, each aspect of the Petition’s combination theory relies on arrangements and modifications that are not taught or suggested by any of Thelen, Bailey, or Chen.

49. Furthermore, I disagree with the Petition’s statement that arriving at the design claimed in limitations 46(b) and (d) would have been a “common sense extension” of Thelen, Bailey, and Chen. As explained above, the Petition’s theory requires modifying numerous aspects of each of Thelen, Bailey, and Chen to arrive at a design that is not disclosed in any of those references. In my experience, modifying the teachings of three different references to arrive at a design that is not taught by any of them is not typically the product of “common sense.”

50. For the same reasons, the Petition’s combination would not have been expected to yield “predictable result,” as persons of ordinary skill in the art would

not have found “predictable” the outcome of rearranging and redesigning Thelen, Bailey, and Chen in ways that are not taught or even hinted in any of those references.

51. I am aware that the Petition contends that its proposed combination would have improved Thelen’s “efficiency.” I note, however, that the proposed modification would have required far more hardware—more processors and more memories—than what either Thelen or Bailey teaches. An ordinary artisan would have known that this would have come at significant additional cost and complexity compared to Thelen’s and Bailey’s existing systems. The Petition does not account for that consideration—or indeed, any costs or trade-offs implied by its combination.

52. Finally, in my opinion the Petition’s rearrangement of Thelen, Bailey, and Chen goes well beyond the level of ordinary skill in the art. In the Petition’s combination, the artisan would have needed the skill and a reason to (1) dedicate multiple recognizers to each of Thelen’s recognition models, creating multiple groups each containing multiple recognizers, contrary to Thelen’s own teachings and motivations; (2) associate each group of recognizers with a dedicated memory, which was not taught by Thelen or Bailey; (3) ensure that each dedicated memory contained part of the lexical data, contrary to Bailey and Thelen’s teachings; and (4) ensure that the sum of the dedicated memories together contained the complete

lexical data, which was not even remotely taught in Thelen or Bailey. In my experience, an ordinary artisan with the qualifications and work experience described by the Petition would not have arrived at those choices through ordinary creativity. In particular, I note that the Petition identifies no logical path or motivation that would have led the ordinary artisan, using ordinary creativity, to start from Thelen, Bailey, and Chen and arrive at the design of limitations 46(b) and (d).

53. I hereby declare that all statements made herein of my own knowledge are true and that all opinions expressed herein are my own; and further that these statements were made with the knowledge that willful false statements and the like are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.



DeLiang Wang, Ph.D.

Executed on March 15, 2023

EXHIBIT A

DeLiang Wang

Department of Computer Science and Engineering

The Ohio State University

Columbus, OH 43210-1277

Phone: (614) 292-6827; Fax: (614) 292-2911

Email: dwang@cse.ohio-state.edu

URL: "https://web.cse.ohio-state.edu/~dwang"

Research Interests:

Computational audition and vision, speech and audio processing, and deep learning.

Education:

1/88 – 8/91 **Ph.D.**, *Computer Science*, University of Southern California, Los Angeles.
Dissertation Advisor: Prof. Michael Arbib.

9/83 – 7/86 **M.S.**, *Computer Science*, Beijing (Peking) University, Beijing.
Thesis Advisor: Prof. Zhuoqun Xu.

9/79 – 7/83 **B.S.**, *Computer Science*, Beijing University, Beijing.

Professional Experience:

10/91 – Professor (10/01 -)/Associate Professor (10/97 - 9/01)/Assistant Professor
(10/91 - 9/97), Department of Computer Science & Engineering and Center for
Cognitive and Brain Sciences (also Participating Faculty in Department of
Biomedical Engineering), The Ohio State University, Columbus, OH.

1/15 – 6/19 Visiting Scholar, Center of Intelligent Acoustics and Immersive Communications,
Northwestern Polytechnical University, Xi'an, China

10/14 – 12/14 Visiting Scholar, Starkey Hearing Technologies, Eden Prairie, MN

10/06 – 6/07 Visiting Scholar, Oticon A/S, Copenhagen, Denmark

10/98 – 9/99 Visiting Scholar, Department of Psychology, Harvard University, Cambridge,
MA

Summer 1995 Visiting Research Fellow, Department of Computer Science, University of
Sheffield, Sheffield, U.K.

1/88 – 8/91 Research Assistant, Department of Computer Science, University of Southern
California, Los Angeles, CA.

Summer 1990 Visiting Scholar, Department of Neurobiology, University of Kassel, Kassel,

Germany.

8/86 – 12/87 Assistant Research Fellow, Institute of Computing Technology, Academia Sinica, Beijing.

Spring 1985 Teaching Assistant, Computer Science Department, Beijing University, Beijing.

Awards and Honors:

National Science Foundation Research Initiation Award, 1992

Office of Naval Research Young Investigator Award, 1996

One of 34 Young Investigators selected from 416 contenders across science and engineering disciplines for "exceptional promise for doing creative research"

OSU College of Engineering Lumley Research Award, 1996, 2000, 2005, 2010

IEEE Fellow, 2004

"For contributions to advancing oscillatory correlation theory and its application to auditory and visual scene analysis"

IEEE Computational Intelligence Society Outstanding Paper Award, 2007

For Paper J54 published in the 2005 *IEEE Transactions on Neural Networks*

International Neural Network Society Helmholtz Award, 2008

IEEE Distinguished Lecturer, 2010-2012

OSU Distinguished Scholar Award, 2014

Starkey Signal Processing Research Award, 2015

For Paper C155 presented in the 2015 International Conference on Acoustics, Speech, and Signal Processing

IEEE Signal Processing Society Best Paper Award, 2019

For Paper J109 published in the 2014 *IEEE/ACM Transactions on Audio, Speech, and Language Processing*

International Neural Network Society Ada Lovelace Service Award, 2020

ISCA Fellow, 2021

"For seminal work in the area of speech separation and speech segregation"

Professional Services:

Co-Editor-in-Chief: *Neural Networks* (Elsevier, 2011-)

Editorial Board Membership:

- *Cognitive Neurodynamics* (Springer, 2006-2017)
- *IEEE Transactions on Neural Networks* (1998-2006)
- *IEEE Transactions on Audio, Speech, and Language Processing* (2012-2015)
- *Neurocomputing* (Elsevier Science, 1995-2010)
- *Neural Computing & Applications* (Springer, 1997-2017)
- *EURASIP Journal on Audio, Speech, & Music Processing* (2006-2016)
- *Neural Networks* (Elsevier, 2009-2010)
- *Frontiers in Electrical and Electronics Engineering* (Springer, 2010-2013)
- *Cognitive Computation* (Advisory, Springer, 2012-)

Guest Editor for *Photogrammetric Engineering & Remote Sensing*, the official journal of the American Society for Photogrammetry and Remote Sensing, for a special issue on Target Recognition (1999).

Editorial Advisory Board Membership: *Handbook of Brain Theory and Neural Networks*, 2nd Ed., 2003, MIT Press.

Guest Co-Editor for *Neural Networks* Special Issue on Progress in Neural Network Research, 2003.

Guest Editor for *IEEE Transactions on Neural Networks* Special Issue on Temporal Coding for Neural Information Processing, 2004.

Organizer: *International Neural Network Society Special Interest Group on Neurodynamics and Chaos* (1994-1996).

Co-Chair: Working Group III/5: "Remote sensing and vision theories for automatic scene interpretation" of the *International Society for Photogrammetry and Remote Sensing* (1996-2000).

Chair: *IEEE Computational Intelligence Society Neural Networks Technical Committee* (2004).

President: *International Neural Network Society* (2006).

Member: *Governing Board of the International Neural Network Society* (2003-2005; 2008-2013; 2020-2022); *IEEE Computational Intelligence Society Fellows Committee* (2008-2009); *IEEE Computational Intelligence Society Neural Networks Technical Committee* (2005-2006, 2012-2018); *IEEE Neural Network Society Neural Networks Technical Committee* (2000-2003); *IEEE Signal Processing Society Machine Learning for Signal Processing Technical Committee* (2002-2004); *IEEE Signal Processing Society Speech and Language Processing Technical Committee* (2009-2011); *International Neural Network Society Executive Committee* (2019-2020)

Co-Chair: *NIPS Workshop on Connectionist Modelling of Auditory Scene Analysis* (Dec. 1996).

Organizer: *USC Symposium on Computational and Cognitive Neuroscience* (Aug. 2000).

Organizer: *AFOSR-OSU Workshop on Computational Audition* (Aug. 2002).

Program Co-Chair: *International Joint Conference on Neural Networks* (July 2003).

Program Co-Chair: *International Conference on Neural Information Processing* (Oct. 2006).

Co-Organizer: *NSF Workshop on Perspectives on Speech Separation* (Oct. 2003).

Co-Organizer: *AFOSR-NSF Workshop on Speech Separation and Comprehension in Complex Acoustic Environments by Humans and Machines* (Nov. 2004).

Tutorials Chair: *World Congress on Computational Intelligence* (July 2006).

General Co-Chair: *International Conference on Intelligent Computing* (Aug. 2011).

Program/Advisory Committee Member:

- *International Conference on Pattern Recognition and Machine Intelligence* (2021)
- *AAAI Conference on Artificial Intelligence* (2019)
- *Annual Conference of the Cognitive Science Society* (2006, 2008)
- *CVPR Workshop on Computer Vision for the Nano-Scale* (2003)
- *ICASSP Multi-channel Multi-party Meeting Transcription Challenge* (2022)
- *IEEE ChinaSIP* (2014)
- *IEEE International Conference on Acoustics, Speech, and Signal Processing* (2003-2005, 2008-2012, 2015-2016)
- *IEEE International Conference on Cognitive Informatics* (2006-2011)
- *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (2017)
- *IEEE Workshop on Machine Learning for Signal Processing* (2004-2011)
- *IEEE Workshop on Neural Networks for Signal Processing* (2003)
- *IFIP World Computer Congress: Intelligent Information Processing* (2000)
- *International Conference for Young Computer Scientists* (1993, 1995)
- *International Conference on Artificial Intelligence and Information Processing* (2004)
- *International Conference on Artificial Intelligence and Soft Computing* (2008, 2010)
- *International Conference on Independent Component Analysis and Blind Source Separation* (2006, 2007, 2009)
- *International Conference on Cognitive Neurodynamics* (2007, 2009, 2011, 2013)
- *International Conference on Intelligent Computing* (2005-2007)
- *International Conference on Latent Variable Analysis and Signal Separation* (2015, 2017)
- *International Conference on Neural Information Processing* (1995, 1997, 2000, 2016, 2017)
- *International Conference on Neural Networks and Associative Memory* (2006- 2007)
- *International Conference on Neural Networks and Brain* (1998)
- *International Conference on Pattern Recognition and Machine Intelligence* (2021)
- *International Conference on Soft Computing and Machine Intelligence* (2014-2015)
- *International Joint Conference on Artificial Intelligence* (2020)

- *International Joint Conference on Neural Networks (1999-2002, 2005, 2007-2011, 2013-2016, 2019)*
- *International Symposium on Neural Networks (2005, 2006, 2009)*
- *ISCA Workshop on Statistical and Perceptual Audio Processing (2004, 2006)*
- *ISPRS Symposium on Object Recognition and Scene Classification from Multispectral and Multisensor Pixels (1998)*
- *Midwest Artificial Intelligence and Cognitive Science Conference (1996)*
- *REVERB Workshop (2014)*
- *Symposium on Computation and Cognitive Neuroscience (2020)*
- *Workshop on Auditory Attention and Scene Analysis (2014)*
- *World Congress on Neural Networks (1995)*
- *Journal Reviewer: AI Magazine; Acta Acustica united with Acustica; Applied Acoustics; Biological Cybernetics; Cognitive Neurodynamics; Computer Speech and Language; Digital Signal Processing; EURASIP Journal on Audio, Speech, & Music Processing; Europhysics; Hearing Research; IEE Electronics Letters; IEEE Computer; IEEE Intelligent Systems Magazine; IEEE Journal of Selected Topics in Signal Processing; IEEE Signal Processing Letters; IEEE Trans. on Biomedical Engineering; IEEE Trans. on Neural Networks; IEEE Trans. on Signal Processing; IEEE/ACM Trans. on Audio, Speech, & Language Processing (formerly IEEE Trans. on Speech and Audio Processing); IEEE Trans. on Signal Processing; IEEE Trans. on Systems, Man, and Cybernetics; International Journal of Artificial Intelligence and Pattern Recognition; International Journal of Bifurcation and Chaos; International Journal of Neural Systems; JASA Express Letters; Journal of Machine Learning Research; Journal of Physics: A; Journal of Signal Processing Systems; Network; Journal of Speech Language & Hearing Research; Journal of the Acoustical Society of America; Neural Networks; Neural Computation; Neurocomputing; NeuroImage; Pattern Recognition; Pattern Recognition Letters; Photogrammetric Engineering & Remote Sensing (PE & RS); Physica D; Physical Review Letters; PLOS Computational Biology; Science Advances; Sensors; Signal Processing; Signal Processing Magazine; Speech Communication.*
- *Conference Reviewer: ACM CHI Conference on Human Factors in Computing Systems; Conference of the Cognitive Science Society; Conference on Neural Information Processing Systems; IEEE International Conference on Acoustics, Speech, and Signal Processing; IEEE International Conference on Cognitive Informatics; IEEE International Conference on Neural Networks; IEEE Workshop on Applications of Signal Processing to Audio and Acoustics; IEEE Workshop on Machine Learning for Signal Processing; IEEE Workshop on Neural Networks for Signal Processing; IEEE Workshop on Statistical Signal Processing; ICSA Workshop on Statistical and Perceptual Audition; International Conference on Independent Component Analysis and Blind Source Separation; International Conference on Latent Variable Analysis and Signal Separation; International Joint Conference on Artificial Intelligence; International Joint Conference on Neural Networks; International Conference for Young Computer Scientists; International Conference on Pattern Recognition; International Conference on Robotics and Automation; ISPRS Symposium on Object Recognition and Scene Classification from Multispectral and Multisensor Pixels; Midwest Artificial Intelligence and Cognitive*

Science Conference; Midwest Cognitive Science Conference; World Congress on Neural Networks; REVERB Workshop.

Book Reviewer: *Cambridge University Press, De Gruyter, Kluwer Academic Publishers, MIT Press, Springer, Wiley*

- Proposal Reviewer: *Air Force Office of Scientific Research, Canadian Foundation for Innovation, Danish Council for Independent Research, King Fahd University (Saudi Arabia), Leverhulme Trust Early Career Fellowship, Mathematics of Information Technology and Complex Systems (Canada), National Research Foundation of Korea, National Science Foundation, Natural Sciences and Engineering Research Council of Canada, Human Frontier Research Program (Japan), National Science Foundation, Ohio Supercomputer Center, Oversea Review Board of the Chinese Academy of Sciences, Research Foundation of Belgium, Science Foundation Ireland, Qatar National Research Fund.*

Panelist: *National Science Foundation ITR Program; National Science Foundation Career Award Program; National Science Foundation POWRE Program; Canadian Foundation for Innovation; Ohio Supercomputer Center*

Professional Membership:

International Neural Network Society (Senior Member), IEEE Computational Intelligence Society, IEEE Signal Processing Society, IEEE (Fellow), and ISCA (Fellow)

Publications:

Journal Articles

- J1. Wang D.L. (1986): "A neuron model based on information processing of the nervous system," *Cognitive Science* (in Chinese), January 1986, pp. 81-88.
- J2. Wang D.L. and Hsu C.C. (1988): "A neuron model for computer simulation of neural networks," *Acta Automatica Sinica* (both in Chinese and in English), vol. 14, pp. 424-430.
- J3. Hsu C.C. and Wang D.L. (1988): "SLONN: A simulation language for neural networks and its implementation," *Journal of Computers* (in Chinese), vol. 11, pp. 741-749.
- J4. Wang D.L., Buhmann J., and von der Malsburg C. (1990): "Pattern segmentation in associative memory," *Neural Computation*, vol. 2, pp. 94-106. Also reprinted as B4.
- J5. Wang D.L. and Hsu C.C. (1990): "SLONN: A simulation language for modeling of neural networks," *Simulation*, vol. 55, pp. 69-83.
- J6. Wang D.L. and Arbib M.A. (1990): "Complex temporal sequence learning based on short-term memory," *Proceedings of the IEEE*, vol. 78, pp. 1536-1543.
- J7. Wang D.L. and Arbib M.A. (1991): "How does the toad's visual system discriminate different worm-like stimuli?" *Biological Cybernetics*, vol. 64, pp. 251-261.
- J8. Wang D.L. and Ewert J.-P. (1992): "Configurational pattern discrimination responsible for dishabituation in common toads *Bufo bufo* (L.): behavioral tests of the predictions of a neural model," *Journal of Comparative Physiology A*, vol. 170, pp. 317-325.
- J9. Wang D.L. and Arbib M.A. (1992): "Modeling the dishabituation hierarchy in toads: the role of the primordial hippocampus," *Biological Cybernetics*, vol. 67, pp. 535-544.
- J10. Wang D.L. (August 1993): "Pattern recognition: Neural networks in perspective," *IEEE Expert*, vol. 8, pp. 52-60.
- J11. Wang D.L. and Arbib M.A. (1993): "Timing and chunking in processing temporal order," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, pp. 993-1009.
- J12. Wang D.L. (1993): "A neural model of synaptic plasticity underlying short-term and long-term habituation," *Adaptive Behavior*, vol. 2, pp. 111-129.
- J13. Wang D.L. (1994): "Modeling neural mechanisms of vertebrate habituation: Locus specificity and pattern discrimination," *Journal of Computational Neuroscience*, vol. 1, pp. 285-299.
- J14. Wang D.L. and Terman D. (1995): "Locally excitatory globally inhibitory oscillator networks," *IEEE Transactions on Neural Networks*, vol. 6(1), pp. 283-286.

- J15. Wang D.L. and Yuwono B. (1995): "Anticipation-based temporal pattern generation," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 25, pp. 615-628.
- J16. Terman D. and Wang D.L. (1995): "Global competition and local cooperation in a network of neural oscillators," *Physica D*, vol. 81, pp. 148-176.
- J17. Wang D.L. (1995): "Emergent synchrony in locally coupled neural oscillators," *IEEE Transactions on Neural Networks*, vol. 6, pp. 941-948.
- J18. Campbell S.R. and Wang D.L. (1996): "Synchrony and desynchrony in a network of Wilson-Cowan oscillators," *IEEE Transactions on Neural Networks*, vol. 7, pp. 541-554.
- J19. Wang D.L. (1996): "Primitive auditory segregation based on oscillatory correlation," *Cognitive Science*, vol. 20, pp. 409-456.
- J20. Wang D.L., Liu, X.M., and Ahalt, S.C. (1996): "On temporal generalization of simple recurrent networks," *Neural Networks*, vol. 9, pp. 1099-1118.
- J21. Wang D.L. and Yuwono B. (1996): "Incremental learning of complex temporal patterns," *IEEE Transactions on Neural Networks*, vol. 7, pp. 1465-1481.
- J22. Wang D.L. and Terman D. (1997): "Image segmentation based on oscillatory correlation," *Neural Computation*, vol. 9, pp. 805-836 (For errata see *Neural Computation*, vol. 9, pp. 1623-1626, 1997).
- J23. Brown G.J. and Wang D.L. (1997): "Modelling the perceptual segregation of double vowels with a network of neural oscillators," *Neural Networks*, vol. 10, 1547-1558.
- J24. Campbell S.R. and Wang D.L. (1998): "Relaxation oscillators with time delay coupling," *Physica D*, vol. 111, pp. 151-178.
- J25. Linsay P.S. and Wang D.L. (1998): "Fast numerical integration of relaxation oscillator networks based on singular limit solutions," *IEEE Transactions on Neural Networks*, vol. 9, pp. 523-532.
- J26. Shareef N., Wang D.L., and Yagel R. (1999): "Segmentation of medical images using LEGION," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 74-91.
- J27. Liu X. and Wang D.L. (1999): "Range image segmentation using a relaxation oscillator network," *IEEE Transactions on Neural Networks*, vol. 10, pp. 564-573.
- J28. Wang D.L. and Brown G.J. (1999): "Separation of speech from interfering sounds based on oscillatory correlation," *IEEE Transactions on Neural Networks*, vol. 10, pp. 684-697.
- J29. Wang D.L. (1999): "Object selection based on oscillatory correlation," *Neural Networks*, vol. 12, pp. 579-592.

- J30. Campbell S.R., Wang D.L., and Jayaprakash C. (1999): "Synchrony and desynchrony in integrate-and-fire oscillators," *Neural Computation*, vol. 11, pp. 1595-1619.
- J31. Liu X., Wang D.L., and Ramirez J.R. (2000): "Boundary detection by contextual nonlinear smoothing," *Pattern Recognition*, vol. 33, pp. 263-280.
- J32. Wang D.L. (2000): "On connectedness: a solution based on oscillatory correlation," *Neural Computation*, vol. 12, pp. 131-139.
- J33. Cesmeli E. and Wang D.L. (2000): "Motion segmentation based on motion/brightness integration and oscillatory correlation," *IEEE Transactions on Neural Networks*, vol. 11, pp. 935-947.
- J34. Chen K., Wang D.L., and Liu X. (2000): "Weight adaptation and oscillatory correlation for image segmentation," *IEEE Transactions on Neural Networks*, vol. 11, pp. 1106-1123.
- J35. Liu X., Chen K., and Wang D.L. (2001): "Extraction of hydrographic regions from remote sensing images using an oscillator network with weight adaptation," *IEEE Transactions on GeoScience and Remote Sensing*, vol. 39, pp. 207-211.
- J36. Van der Kouwe A.J.W., Wang D.L., and Brown G.J. (2001): "A comparison of auditory and blind separation techniques for speech segregation," *IEEE Transactions on Speech and Audio Processing*, vol. 9, pp. 189-195.
- J37. Fox J.J., Jayaprakash C., Wang D.L., and Campbell S.R. (2001): "Synchronization in relaxation oscillator networks with conduction delays," *Neural Computation*, vol. 13, pp. 1003-1021.
- J38. Cesmeli E. and Wang D.L. (2001): "Texture segmentation using Gaussian Markov random fields and neural oscillator networks," *IEEE Transactions on Neural Networks*, vol. 12, pp. 394-404.
- J39. Chen K. and Wang D.L. (2001): "Perceiving geometric patterns: from spirals to inside/outside relations," *IEEE Transactions on Neural Networks*, vol. 12, pp. 1084-1102.
- J40. Wang D.L. and Liu X. (2002): "Scene analysis by integrating primitive segmentation and associative memory," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 32, pp. 254-268.
- J41. Chen K. and Wang D.L. (2002): "A dynamically coupled neural oscillator network for image segmentation," *Neural Networks*, vol. 15, pp. 423-439.
- J42. Kristjansson A., Wang D.L., and Nakayama K. (2002): "The role of priming in conjunctive visual search," *Cognition*, vol. 85, pp. 37-52.
- J43. Liu X. and Wang D.L. (2002): "A spectral histogram model for texton modeling and texture discrimination," *Vision Research*, vol. 42, pp. 2617-2634.

- J44. Cesmeli E., Lindsey D.T., and Wang D.L. (2002): "An oscillatory correlation model of visual motion analysis," *Perception & Psychophysics*, vol. 64, pp. 1191-1217.
- J45. Wu M., Wang D.L., and Brown G.J. (2003): "A multi-pitch tracking algorithm for noisy speech," *IEEE Transactions on Speech and Audio Processing*, vol. 11, pp. 229-241.
- J46. Liu X. and Wang D.L. (2003): "Texture classification using spectral histograms," *IEEE Transactions on Image Processing*, vol. 12, pp. 661-670.
- J47. Liu X., Srivastava A., and Wang D.L. (2003): "Intrinsic generalization analysis of low dimensional representations," *Neural Networks*, vol. 16, pp. 537-545.
- J48. Roman N., Wang D.L., Brown G.J. (2003): "Speech segregation based on sound localization," *Journal of the Acoustical Society of America*, vol. 114, pp. 2236-2252.
- J49. Hu G. and Wang D.L. (2004): "Monaural speech segregation based on pitch tracking and amplitude modulation," *IEEE Transactions on Neural Networks*, vol. 15, pp. 1135-1150.
- J50. Campbell S.R. and Wang D.L. (2004): "Synchronization rates in classes of relaxation oscillators," *IEEE Transactions on Neural Networks*, vol. 15, pp. 1027-1038.
- J51. Palomäki K.J., Brown G.J., and Wang D.L. (2004): "A binaural processor for missing data speech recognition in the presence of noise and small-room reverberation," *Speech Communication*, vol. 43, pp. 361-378.
- J52. Srinivasan S. and Wang D.L. (2005): "A schema-based model for phonemic restoration," *Speech Communication*, vol. 45, pp. 63-87.
- J53. Wang D.L., Kristjansson A., and Nakayama K. (2005): "Efficient visual search without top-down or bottom-up guidance," *Perception & Psychophysics*, vol. 67, pp. 239-253.
- J54. Wang D.L. (2005): "The time dimension for scene analysis," *IEEE Transactions on Neural Networks*, vol. 16, pp. 1401-1426.
- J55. Shao Y. and Wang D.L. (2006): "Model-based sequential organization in cochannel speech," *IEEE Transactions on Audio, Speech, & Language Processing*, vol. 14, pp. 289-298.
- J56. Wu M. and Wang D.L. (2006): "A two-stage algorithm for one-microphone reverberant speech enhancement," *IEEE Transactions on Audio, Speech, & Language Processing*, vol. 14, pp. 774-784.
- J57. Wu M. and Wang D.L. (2006): "A pitch-based method for the estimation of short reverberation time," *Acta Acustica united with Acustica*, vol. 92, pp. 337-339.
- J58. Liu X. and Wang D.L. (2006): "Image and texture segmentation using local spectral histograms," *IEEE Transactions on Image Processing*, vol. 15, pp. 3066-3077.

- J59. Srinivasan S., Roman N., and Wang D.L. (2006): "Binary and ratio time-frequency masks for robust speech recognition," *Speech Communication*, vol. 48, pp. 1486-1501.
- J60. Brungart D.S., Chang P.S., Simpson B.D., and Wang D.L. (2006): "Isolating the energetic component of speech-on-speech masking with ideal time-frequency segregation," *Journal of the Acoustical Society of America*, vol. 120, pp. 4007-4018.
- J61. Roman N. and Wang D.L. (2006): "Pitch-based monaural segregation of reverberant speech," *Journal of the Acoustical Society of America*, vol. 120, pp. 458-469.
- J62. Roman N., Srinivasan S., and Wang D.L. (2006): "Binaural segregation in multisource reverberant environments," *Journal of the Acoustical Society of America*, vol. 120, pp. 4040-4051.
- J63. Hu G. and Wang D.L. (2007): "Auditory segmentation based on onset and offset analysis," *IEEE Transactions on Audio, Speech, & Language Processing*, vol. 15, pp. 396-405.
- J64. Li Y. and Wang D.L. (2007): "Separation of singing voice from music accompaniment for monaural recordings," *IEEE Transactions on Audio, Speech, and Language Processing*, pp. 1475-1487.
- J65. Srinivasan S. and Wang D.L. (2007): "Transforming binary uncertainties for robust speech recognition," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, pp. 2130-2140.
- J66. Pedersen M.S., Wang D.L., Larsen J., and Kjems U. (2008): "Two-microphone separation of speech mixtures," *IEEE Transactions on Neural Networks*, vol. 19, pp. 475-492.
- J67. Wang D.L. and Chang P.S. (2008): "An oscillatory correlation model of auditory streaming," *Cognitive Neurodynamics*, vol. 2, pp. 7-19.
- J68. Roman N. and Wang D.L. (2008): "Binaural tracking of multiple moving sources," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, pp. 728-739.
- J69. Hu G. and Wang D.L. (2008): "Segregation of unvoiced speech from nonspeech interference," *Journal of the Acoustical Society of America*, vol. 124, pp. 1306-1319.
- J70. Wang D.L., Kjems U., Pedersen M.S., Boldt J.B., and Lunner T. (2008): "Speech perception of noise with binary gains," *Journal of the Acoustical Society of America*, vol. 124, pp. 2303-2307.
- J71. Srinivasan S. and Wang D.L. (2008): "A model for multitalker speech perception," *Journal of the Acoustical Society of America*, vol. 124, pp. 3213-3224.
- J72. Wang D.L. (2008): "Time-frequency masking for speech separation and its potential for hearing aid design," *Trends in Amplification*, vol. 12, pp. 332-353.

- J73. Li Y. and Wang D.L. (2009): "On the optimality of ideal binary time-frequency masks," *Speech Communication*, vol. 51, pp. 230-239.
- J74. Jin Z. and Wang D.L. (2009): "A supervised learning approach to monaural segregation of reverberant speech," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, pp. 625-638.
- J75. Wang D.L., Kjems U., Pedersen M.S., Boldt J.B., and Lunner T. (2009): "Speech intelligibility in background noise with ideal binary time-frequency masking," *Journal of the Acoustical Society of America*, vol. 125, pp. 2336-2347.
- J76. Li Y. and Wang D.L. (2009): "Musical sound separation based on binary time-frequency masking," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2009, Article ID 130567, 10 pages.
- J77. Li Y., Woodruff J., and Wang D.L. (2009): "Monaural musical sound separation based on pitch and common amplitude modulation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, pp. 1361-1371.
- J78. Brungart D.S., Chang P.S., Simpson B.D., and Wang D.L. (2009): "Multitalker speech perception with ideal time-frequency segregation: Effects of voice characteristics and speaker number" *Journal of the Acoustical Society of America*, vol. 125, pp. 4006-4022.
- J79. Shao Y. and Wang D.L. (2009): "Sequential organization of speech in computational auditory scene analysis," *Speech Communication*, vol. 51, pp. 657-667.
- J80. Kjems U., Boldt J.B., Pedersen M.S., Lunner T., and Wang D.L. (2009): "Role of mask pattern in intelligibility of ideal binary-masked noisy speech," *Journal of the Acoustical Society of America*, vol. 126, pp. 1415-1426.
- J81. Shao Y., Srinivasan S., Jin Z., and Wang D.L. (2010): "A computational auditory scene analysis system for speech segregation and robust speech recognition," *Computer Speech and Language*, vol. 24, pp. 77-93.
- J82. Srinivasan S. and Wang D.L. (2010): "Robust speech recognition by integrating speech separation and hypothesis testing," *Speech Communication*, vol. 52, pp. 72-81.
- J83. Woodruff J. and Wang D.L. (2010): "Sequential organization of speech in reverberant environments by integrating monaural grouping and binaural localization," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, pp. 1856-1866.
- J84. Narayanan A. and Wang D.L. (2010): "Robust speech recognition from binary masks," *Journal of the Acoustical Society of America Express Letters*, vol. 128, pp. EL217-222.
- J85. Hu G. and Wang D.L. (2010): "A tandem algorithm for pitch estimation and voiced speech segregation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, pp. 2067-2079.

- J86. Quiles M.G., Wang D.L., Zhao L., Romero R.A.F., and Huang D.-S. (2011): "Selecting salient objects in real scenes: An oscillatory correlation model," *Neural Networks*, vol. 24, pp. 54-64.
- J87. Jan T., Wang W., Wang D.L. (2011): "A multistage approach to blind separation of convolutive speech mixtures," *Speech Communication*, vol. 53, pp. 524-539.
- J88. Jin Z. and Wang D.L. (2011): "HMM-based multipitch tracking for noisy and reverberant speech," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, pp. 1091-1102.
- J89. Hu K. and Wang D.L. (2011): "Unvoiced speech segregation from nonspeech interference via CASA and spectral subtraction," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, pp. 1600-1609.
- J90. Jin Z. and Wang D.L. (2011): "Reverberant speech segregation based on multipitch tracking and classification," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, pp. 2328-2337.
- J91. Yuan J., Wang D.L., Wu B., Yan L., and Li R. (2011): "LEGION-based automatic road extraction from satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, pp. 4528-4538.
- J92. Yuan J., Wang D.L., and Li R. (2012): "Image segmentation using local spectral histograms and linear regression," *Pattern Recognition Letters*, vol. 33, pp. 615-622.
- J93. Hsu C.-L., Wang D.L., and Jang J.-S.R., Hu K. (2012): "A tandem algorithm for singing pitch extraction and voice separation from music accompaniment," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, pp. 1482-1491.
- J94. Woodruff J. and Wang D.L. (2012): "Binaural localization of multiple sources in reverberant and noisy environments," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, pp. 1503-1512.
- J95. Zhao X., Shao Y. and Wang D.L. (2012): "CASA-based robust speaker identification," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, pp. 1608-1616.
- J96. Narayanan A. and Wang D.L. (2012): "A CASA based system for long-term SNR estimation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, pp. 2518-2527.
- J97. Han K. and Wang D.L. (2012): "A classification based approach to speech segregation," *Journal of the Acoustical Society of America*, vol. 132, pp. 3475-3483.
- J98. Hu K. and Wang D.L. (2013): "An unsupervised approach to cochannel speech separation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, pp. 120-129.

J99. Han K. and Wang D.L. (2013): "Towards generalizing classification based speech separation," *IEEE Transactions on Audio, Speech, and Language Processing*, pp. 166-175.

J100. Wang Y., Han K., and Wang D.L. (2013): "Exploring monaural features for classification-based speech segregation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, pp. 270-279.

J101. Woodruff J. and Wang D.L. (2013): "Binaural detection, localization, and segregation in reverberant environments based on joint pitch and azimuth cues," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, pp. 806-815.

J102. Wang Y. and Wang D.L. (2013): "Towards scaling up classification-based speech separation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, pp. 1381-1390.

J103. Narayanan A. and Wang D.L. (2013): "The role of binary mask patterns in automatic speech recognition in background noise," *Journal of the Acoustical Society of America*, vol. 133, pp. 3083-3093.

J104. Hu K. and Wang D.L. (2013): "An iterative model-based approach to cochannel speech separation," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2013, Article ID 2013-14, 11 pages.

J105. Healy E.W., Yoho S.E., Wang Y., and Wang D.L. (2013): "An algorithm to improve speech recognition in noise for hearing-impaired listeners," *Journal of the Acoustical Society of America*, vol. 134, pp. 3029-3038.

J106. Hartmann W., Narayanan A., Fosler-Lussier E., and Wang D.L. (2013): "A direct masking approach to robust ASR," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, pp. 1993-2005.

J107. Yuan J., Wang D.L., and Li R. (2014): "Remote sensing image segmentation by combining spectral and texture features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, pp. 16-24.

J108. Narayanan A. and Wang D.L. (2014): "Investigation of speech separation as a front-end for noise robust speech recognition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, pp. 826-835.

J109. Zhao X., Wang Y., and Wang D.L. (2014): "Robust speaker identification in noisy and reverberant conditions," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, pp. 836-845.

J110. Williamson D.S., Wang Y., and Wang D.L. (2014): "Reconstruction techniques for improving the perceptual quality of binary masked speech," *Journal of the Acoustical Society of America*, vol. 136, pp. 892-902.

J111. Wang Y., Narayanan A. and Wang D.L. (2014): "On training targets for supervised speech separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, pp. 1849-1858.

J112. Chen J., Wang Y., and Wang D.L. (2014): "A feature study for classification-based speech separation at low signal-to-noise ratios," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, pp. 1993-2002.

J113. Jiang Y., Wang D.L., Liu R.S., Feng Z.M. (2014): "Binaural classification for reverberant speech segregation using deep neural networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, pp. 2112-2121.

J114. Han K. and Wang D.L. (2014): "Neural network based pitch tracking in very noisy speech," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, pp. 2158-2168.

J115. Healy E.W., Yoho S.E., Wang Y., Apoux F., and Wang D.L. (2014): "Speech cue transmission by an algorithm to increase consonant recognition in noise for hearing-impaired listeners," *Journal of the Acoustical Society of America*, vol. 136, pp. 3325-3336.

J116. Narayanan A. and Wang D.L. (2015): "Improving robustness of deep neural network acoustic models via speech separation and joint adaptive training," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, pp. 92-101.

J117. Han K., Wang Y., Wang D.L., Woods W.S., Merks I., and Zhang T. (2015): "Learning spectral mapping for speech dereverberation and denoising," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, pp. 982-992.

J118. Yuan J., Wang D.L., and Cheriyyadat A.M. (2015): "Factorization-based texture segmentation," *IEEE Transactions on Image Processing*, vol. 24, pp. 3488-3497.

J119. Zhao X., Wang Y., and Wang D.L. (2015): "Cochannel speaker identification in anechoic and reverberant conditions," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, pp. 1727-1736.

J120. Williamson D.S., Wang Y., and Wang D.L. (2015): "Estimating nonnegative matrix model activations with deep neural networks to increase perceptual speech quality," *Journal of the Acoustical Society of America*, vol. 138, pp. 1399-1407.

J121. Healy E.W., Yoho S.E., Chen J., Wang Y., and Wang D.L. (2015): "An algorithm to increase speech intelligibility for hearing-impaired listeners in novel segments of the same noise type," *Journal of the Acoustical Society of America*, vol. 138, pp. 1660-1669.

J122. Zhang X.-L. and Wang D.L. (2016): "Boosting contextual information for deep neural network based voice activity detection," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, pp. 252-264.

J123. Chen J., Wang Y., and Wang D.L. (2016): "Noise perturbation for supervised speech

separation,” *Speech Communication*, vol. 78, pp. 1-10.

J124. Williamson D.S., Wang Y., and Wang D.L. (2016): “Complex ratio masking for monaural speech separation,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, pp. 483-492.

J125. Wang Z.-Q. and Wang D.L. (2016): “A joint training framework for robust automatic speech recognition,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, pp. 796-806.

J126. Zhang X.-L. and Wang D.L. (2016): “A deep ensemble learning method for monaural speech separation,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, pp. 967-977.

J127. Chen J., Wang Y., Yoho S.E., Wang D.L., and Healy E.W. (2016): “Large-scale training to increase speech intelligibility for hearing-impaired listeners in novel noises,” *Journal of the Acoustical Society of America*, vol. 139, pp. 2604-2612.

J128. Liu Y. and Wang D.L. (2017): “Speaker-dependent multipitch tracking using deep neural networks,” *Journal of the Acoustical Society of America*, vol. 141, pp. 710-721.

J129. Wang D.L. (March 2017): “Deep learning reinvents the hearing aid,” *IEEE Spectrum*, pp. 32-37 (cover story).

J130. Zhang X. and Wang D.L. (2017): “Deep learning based binaural speech separation in reverberant environments,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, pp. 1075-1084.

J131. Delfarah M. and Wang D.L. (2017): “Features for masking-based monaural speech separation in reverberant conditions,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, pp. 1085-1094.

J132. Healy E.W., Delfarah M., Vasko J.L., Carter B.L., and Wang D.L. (2017): “An algorithm to increase intelligibility for hearing-impaired listeners in the presence of a competing talker,” *Journal of the Acoustical Society of America*, vol. 141, pp. 4230-4239.

J133. Mayer F., Williamson D.S., Mowlae P., and Wang D.L. (2017): “Impact of phase estimation on single-channel speech separation based on time-frequency masking,” *Journal of the Acoustical Society of America*, vol. 141, pp. 4668-4679.

J134. Chen J. and Wang D.L. (2017): “Long short-term memory for speaker generalization in supervised speech separation,” *Journal of the Acoustical Society of America*, vol. 141, pp. 4705-4714.

J135. Williamson D.S. and Wang D.L. (2017): “Time-frequency masking in the complex domain for speech dereverberation and denoising,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, pp. 1492-1501.

J136. Chen J. and Wang D.L. (2018): “Supervised speech separation based on deep learning: An overview,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, pp. 1702-1726.

J137. Zhao Y., Wang D.L., Johnson E.M, and Healy E.W. (2018): “A deep learning based segregation algorithm to increase speech intelligibility for hearing-impaired listeners in reverberant-noisy conditions,” *Journal of the Acoustical Society of America*, vol. 144, pp. 1627-1637.

J138. Zhao Y., Wang Z.-Q., and Wang D.L. (2019): “Two-stage deep learning for noisy-reverberant speech enhancement,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, pp. 53-62.

J139. Wang Z.-Q., Zhang X., and Wang D.L. (2019): “Robust speaker localization guided by deep learning-based time-frequency masking,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, pp. 178-188.

J140. Tan K., Chen J., and Wang D.L. (2019): “Gated residual networks with dilated convolutions for monaural speech enhancement,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, pp. 189-198.

J141. Wang Z.-Q. and Wang D.L. (2019): “Combining spectral and spatial features for deep learning based blind speaker separation,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, pp. 457-468.

J142. Healy E.W., Delfarah M., Johnson E.M., and Wang D.L. (2019): “A deep learning algorithm to increase intelligibility for hearing-impaired listeners in the presence of a competing talker and reverberation,” *Journal of the Acoustical Society of America*, vol. 145, pp. 1378-1388.

J143. Pandey A. and Wang D.L. (2019): “A new framework for CNN-based speech enhancement in the time domain,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, pp. 1179-1188.

J144. Healy E.W., Vasco J.L., and Wang D.L. (2019): “The optimal threshold for removing noise from speech is similar across normal and impaired hearing – a time-frequency masking study,” *Journal of the Acoustical Society of America Express Letters*, vol. 145, pp. EL581-586.

J145. Delfarah M. and Wang D.L. (2019): “Deep learning for talker-dependent reverberant speaker separation: An empirical study,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, pp. 1839-1848.

J146. Liu Y. and Wang D.L. (2019): “Divide and conquer: A deep CASA approach to talker-independent monaural speaker separation,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, pp. 2092-2102.

J147. Wang P., Tan K., and Wang D.L. (2020): “Bridging the gap between monaural speech

enhancement and recognition with distortion-independent acoustic modeling,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 39-48.

J148. Tan K. and Wang D.L. (2020): “Learning complex spectral mapping with gated convolutional recurrent networks for monaural speech enhancement,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 380-390.

J149. Wang Z.-Q. and Wang D.L. (2020): “Deep learning based target cancellation for speech dereverberation,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 941-950.

J150. Taherian H., Wang Z.-Q., Chang J., and Wang D.L. (2020): “Robust speaker recognition based on single-channel and multi-channel speech enhancement,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1293-1302.

J151. Healy E.W., Johnson E.M., Delfarah M., and Wang D.L. (2020): “A talker-independent deep learning algorithm to increase intelligibility for hearing-impaired listeners in reverberant competing talker conditions,” *Journal of the Acoustical Society of America*, vol. 147, pp. 4106-4118.

J152. Zhao Y., Wang D.L., Xu B., and Zhang T. (2020): “Monaural speech dereverberation using temporal convolutional networks with self attention,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1598-1607.

J153. Wang Z.-Q., Wang P., and Wang D.L. (2020): “Complex spectral mapping for single- and multi-channel speech enhancement and robust ASR,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1778-1787.

J154. Liu Y. and Wang D.L. (2020): “Causal deep CASA for monaural talker-independent speaker separation,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 2109-2118.

J155. Delfarah M., Liu Y. and Wang D.L. (2020): “A two-stage deep learning algorithm for talker-independent speaker separation in reverberant conditions,” *Journal of the Acoustical Society of America*, vol. 148, pp. 1157-1168.

J156. Pandey A. and Wang D.L. (2020): “On cross-corpus generalization of deep learning based speech enhancement,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 2489-2499.

J157. Wang P., Chen Z., Wang D.L., Li J., and Gong Y. (2021): “Speaker separation using speaker inventories and estimated speech,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 537-546.

J158. Pandey A. and Wang D.L. (2021): “Dense CNN with self-attention for time-domain speech enhancement,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 1270-1279.

J159. Tan K. and Wang D.L. (2021): “Towards model compression for deep learning based speech enhancement,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 1785-1794

J160. Tan K., Zhang X., and Wang D.L. (2021): “Deep learning based real-time speech enhancement for dual-microphone mobile phones,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 1853-1863.

J161. Healy E.W., Tan K., Johnson E.M., and Wang D.L. (2021): “An effectively causal deep learning algorithm to increase intelligibility in untrained noises for hearing-impaired listeners,” *Journal of the Acoustical Society of America*, vol. 149, pp. 3943–3953.

J162. Zhang H. and Wang D.L. (2021): “Deep ANC: A deep learning approach to active noise control,” *Neural Networks*, vol. 141, pp. 1-10.

J163. Wang Z.-Q., Wang P., and Wang D.L. (2021): “Multi-microphone complex spectral mapping for utterance-wise and continuous speech separation,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 2001-2014.

J164. Wang H. and Wang D.L. (2021): “Towards robust speech super-resolution,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 2058-2066.

J165. Li H., Wang D.L., Zhang X., Gao G. (2021): “Recurrent neural networks and acoustic features for frame-level signal-to-noise ratio estimation,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 2878-2887.

J166. Healy E.W., Johnson E.M., Delfarah M., Krishnagiri D.S., Sevich V.A., Taherian H., and Wang D.L. (2021): “Deep learning based speaker separation and dereverberation can generalize across different languages to improve intelligibility,” *Journal of the Acoustical Society of America*, vol. 150, pp. 2526-2538.

J167. Healy E.W., Taherian H., Johnson E.M., and Wang D.L. (2021): “A causal and talker-independent speaker separation/dereverberation deep learning algorithm: Cost associated with conversion to real-time capable operation,” *Journal of the Acoustical Society of America*, vol. 150, pp. 3976-3986.

J168. Tan K., Wang Z.-Q., and Wang D.L. (2022): “Neural spectrospatial filtering,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 605-621.

J169. Wang H. and Wang D.L. (2022): “Neural cascade architecture with triple-domain loss for speech enhancement,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 734-743.

J170. Pandey A. and Wang D.L. (2022): “Self-attending RNN for speech enhancement to improve cross-corpus generalization,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 1374-1385.

Views and Reviews

V1. Wang D.L. (1997): "On the computational basis of synchronized codes," *Behavioral and Brain Sciences*, vol. 20, 700-701.

V2. Wang D.L. (September 2001): Book Review of "Supervised Learning - Foundations of Neural Computation" (MIT Press, 1999). *AI Magazine*, vol. 22, pp. 101-102.

V3. Wang D.L., Freeman W.J., Kozma R., Lozowski A., and Minai A. (2004): Guest Editorial for Special Issue on Temporal Coding for Neural Information Processing, *IEEE Transactions on Neural Networks*, vol. 15, pp. 953-956.

Edited Books and Conference Proceedings

E1. Wunsch D.C., Hasselmo M.E., Venayagamoorthy G.K., and Wang D.L. (ed.): *Advances in Neural Network Research*. Elsevier, Oxford U.K., 2003.

E2. Wunsch D.C., Hasselmo M.E., Wang D.L., and Venayagamoorthy G.K. (ed.): *Proceedings of International Joint Conference on Neural Networks*, IEEE, 2003.

E3. Wang D.L. and Brown G.J. (ed.): *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*. IEEE Press/Wiley, 2006.

E4. King I., Wang J., Chan L., and Wang D.L. (ed.): *Neural Information Processing*. Springer Lecture Notes in Computer Science, Springer, 2006.

Book Chapters

B1. Wang D.L. (1986): "Computer simulation," In: Mao Y.S. (ed), *Handbook for Modern Engineers* (in Chinese), Beijing Press, Beijing, China, pp. 615-617. This book won the 1986 National Award of Excellent and Best-selling Book of Science and Technology.

B2. Arbib M.A. and Wang D.L. (1991): "Computational models of visual pattern discrimination in common toads," In: Ewert J.-P. and Werner H. (eds), *Models of Brain Function and Artificial Neuronal Nets*, GhK University Edition, Kassel, Germany, pp. 67-97.

B3. Wang D.L., Arbib M.A., and Ewert J.-P. (1991): "Dishabituation hierarchies for visual pattern discrimination in toads: A dialog between modeling and experimentation," In: Arbib M.A. and Ewert J.-P. (eds), *Visual Structures and Integrated Functions*, Research Notes in Neural Computing, Springer-Verlag, Berlin, pp. 427-441.

B4. Wang D.L., Buhmann J., and von der Malsburg C. (1991): "Pattern segmentation in associative memory," In: David J.L. and Eichenbaum H. (eds), *Olfaction: A Model System for Computational Neuroscience*, MIT Press, Cambridge MA, pp. 213-224. Reprint of J4.

- B5. Wang D.L. (1995): "Temporal pattern processing," In: Arbib M.A. (ed.), *The Handbook of Brain Theory and Neural Networks*, MIT Press, Cambridge MA, pp. 967-971.
- B6. Wang D.L. (1995): "Habituation," In: Arbib M.A. (ed.), *The Handbook of Brain Theory and Neural Networks*, MIT Press, Cambridge MA, pp. 441-444.
- B7. Wang D.L. (1995): "An oscillatory correlation theory of temporal pattern segmentation," In Covey E., Hawkins H., and Port R.F. (eds.), *Neural Representation of Temporal Patterns*, Plenum, New York, pp. 53-75.
- B8. Wang D.L. (1996): "Synchronous oscillations based on lateral connections," In: Sirosh J., Miikkulainen R., and Choe Y. (eds.), *Lateral Connections in the Cortex: Structure and Function*, Electronic Book at "<http://www.cs.utexas.edu/users/nn/web-pubs/htmlbook96/>".
- B9. Wang D.L. (1998): "Stream segregation based on oscillatory correlation," In: Rosenthal D. and Okuno H.G. (eds.), *Computational Auditory Scene Analysis*, Lawrence Erlbaum, Mahwah NJ, pp. 71-86.
- B10. Wang D.L. (1999): "Relaxation oscillators and networks," In Webster J. (ed.), *Encyclopedia of Electrical and Electronics Engineering*, Wiley & Sons, New York, vol. 18, pp. 396-405.
- B11. Wang D.L. (2000): "Anticipation model for sequential learning of complex temporal patterns," In: Sun R. and Giles L. (ed.), *Sequence Learning: Paradigms, Algorithms, and Applications*, LNAI 1828, Springer, Berlin, pp. 53-79.
- B12. Wang D.L. (2003): "Temporal pattern processing," In: Arbib M.A. (ed.), *The Handbook of Brain Theory and Neural Networks*, 2nd Ed., MIT Press, Cambridge MA, pp. 1163-1167.
- B13. Wang D.L. (2003): "Visual scene segmentation," In: Arbib M.A. (ed.), *The Handbook of Brain Theory and Neural Networks*, 2nd Ed., MIT Press, Cambridge MA, pp. 1215-1219.
- B14. Liu X., Srivastava A., and Wang D.L. (2003): "Intrinsic generalization analysis of low dimensional representations," In: Wunsch D.C., Hasselmo M.E., Venayagamoorthy G.K., and Wang D.L. (ed.), *Advances in Neural Network Research*, Elsevier, Oxford U.K., pp. 537-545. Reprint of J47.
- B15. Wang D.L. (2005): "On ideal binary mask as the computational goal of auditory scene analysis," In: Divenyi P. (ed.), *Speech Separation by Humans and Machines*, Kluwer Academic, Norwell MA, pp. 181-197.
- B16. Brown G.J. and Wang D.L. (2005): "Separation of speech by computational auditory scene analysis," In: Benesty J., Makino S., and Chen J. (ed.), *Speech Enhancement*, Springer, New York, pp. 371-402.

- B17. Brown G.J. and Wang D.L. (2006): "Timing is of the essence: Neural oscillator models of auditory grouping," In: Greenberg S. and Ainsworth W. (ed.), *Listening to Speech: An Auditory Perspective*, Lawrence Erlbaum, Mahwah NJ, pp. 375-392.
- B18. Hu G. and Wang D.L. (2006): "An auditory scene analysis approach to monaural speech segregation," In: Hänsler E. and Schmidt G. (ed.), *Selected Methods for Acoustic Echo and Noise Control*, Springer, Berlin, pp. 485-515.
- B19. Wang D.L. and Brown G.J. (2006): "Fundamentals of computational auditory scene analysis," In: Wang D.L. and Brown G.J. (ed.), *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*, Wiley/IEEE Press, Hoboken NJ, pp. 1-44.
- B20. Wang D.L. (2006): "Feature-based speech segregation," In: Wang D.L. and Brown G.J. (ed.), *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*, Wiley/IEEE Press, Hoboken NJ, pp. 81-114.
- B21. Stern R.M., Brown G.J., and Wang D.L. (2006): "Binaural sound localization," In: Wang D.L. and Brown G.J. (ed.), *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*, Wiley/IEEE Press, Hoboken NJ, pp. 147-185.
- B22. Brown G.J. and Wang D.L. (2006): "Neural and perceptual modeling," In: Wang D.L. and Brown G.J. (ed.), *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*, Wiley/IEEE Press, Hoboken NJ, pp. 351-387.
- B23. Wang D.L. (2007): "Computational scene analysis," In: Duch W. and Mandziuk J. (ed.), *Challenges for Computational Intelligence*, Springer, Berlin, pp. 163-191.
- B24. Wang D.L. (2007): "Relaxation oscillators and networks," In Webster J. (ed.), *Encyclopedia of Electrical and Electronics Engineering* (2nd Edition), Wiley & Sons, New York.
- B25. Wang D.L. and Hu G. (2008): "Cocktail party processing," In: Zurada J.M., Yen G.G., and Wang J. (ed.), *Computational Intelligence: Research Frontiers*, Springer, Berlin, pp. 333-348.
- B26. Roman N. and Wang D.L. (2008): "Binaural speech segregation," In: Hänsler E. and Schmidt G. (ed.), *Speech and Audio Processing in Adverse Environments*, Springer, Berlin, pp. 525-549.
- B27. Narayanan A. and Wang D.L. (2013): "Computational auditory scene analysis and automatic speech recognition," In: Virtanen T., Raj B. and Singh R. (ed.), *Techniques for Noise Robustness in Automatic Speech Recognition*, Wiley, Chichester U.K., pp. 433-462.
- B28. Chen J. and Wang D.L. (2018): "DNN based mask estimation for supervised speech separation," In: Makino S. (ed.), *Audio Source Separation*, Springer, Cham Switzerland, pp. 207-235.

Conference Papers

- C1. Wang D.L. and King I. (1988): "Three neural models which process temporal information," In *Proceedings of the First Annual Conference of the International Neural Network Society*, pp.227, Boston, MA.
- C2. Wang D.L. (1989): "An extended model of the Neocognitron for pattern partitioning and pattern composition," *Proceedings of International Joint Conference of Neural Networks (IJCNN)*, pp. II.267-274. Washington, DC.
- C3. Wang D.L. and Arbib M.A. (1990): "Mechanisms of pattern discrimination in the toad's visual system. *IJCNN-90*, pp. II.477-482. San Diego, CA.
- C4. Wang D.L. and Arbib M.A. (1990): "A computational model of visual pattern discrimination in toads. *Proceedings of the 12th Annual Conference of the Cognitive Science Society*, pp. 598-605. Cambridge, MA.
- C5. Wang D.L. and Arbib M.A. (1991): "Hierarchical dishabituation of visual discrimination in toads," In: Meyer J.-A. and Wilson S. (eds), *Simulation of adaptive behavior: From animals to animats*, pp. 77-88, MIT Press, Cambridge, MA.
- C6. Wang D.L. and Arbib M.A. (1991): "A neural model of temporal sequence generation with interval maintenance," *Proceedings of the 13th Annual Conference of the Cognitive Science Society*, pp. 944-948. Chicago, IL.
- C7. Wang D.L. (1993): "Modeling global synchrony in the visual cortex by locally coupled neural oscillators," *Proceedings of the 15th Annual Conference of the Cognitive Science Society*, pp. 1058-1063. Boulder, CO.
- C8. Wang D.L. (1993): "Modeling stimulus specific habituation: The role of primordial hippocampus," In: Eeckman F.H. and Bower J.M. (eds.), *Computation and Neural Systems (CNS-92)*, pp. 103-107, Kluwer, Boston.
- C9. Wang D.L. (1993): "A neural architecture for complex temporal pattern generation," *Proceedings of the 3rd International Conference for Young Computer Scientists*, pp. 3.25-3.38. Beijing.
- C10. Wang D.L. (1993): "Global phase synchrony in the visual cortex: A local mechanism," *Proceedings of World Congress on Neural Networks (WCNN-93)*, pp. I29-I32. Portland, OR.
- C11. Wang D.L. (1994): "Unifying two forms of memory: A neural model," In: Eeckman F.H. (ed.), *Computation in Neurons and Neural Systems (CNS-93)*, pp. 281-286, Kluwer.
- C12. Wang D.L. (1994): "Modeling global synchrony in the visual cortex by locally coupled neural oscillators," In: Eeckman F.H. (ed.), *Computation in Neurons and Neural Systems (CNS-93)*, pp. 109-114, Kluwer.

- C13. Wang D.L. and Terman D. (1994): "Locally excitatory globally inhibitory oscillator networks: Theory and application to pattern segmentation," *Proceedings of IEEE International Conference on Neural Networks (ICNN-94)*, pp. 945-950. Orlando, FL.
- C14. Campbell S. and Wang D.L. (1994): "Synchronization and desynchronization in locally coupled Wilson-Cowan oscillators," *Proceedings of ICNN-94*, pp. 964-969.
- C15. Wang D.L. and Yuwono B. (1994): "Temporal pattern generation based on anticipation," *Proceedings of ICNN-94*, pp. 3148-3153.
- C16. Wang D.L. and Terman D. (1994): "Locally excitatory globally inhibitory oscillator networks," *Proceedings of WCNN-94*, pp. IV. 745-750. San Diego, CA.
- C17. Wang D.L. and Yuwono B. (1994): "Self-organization of temporal pattern generation based on anticipation," *Proceedings of WCNN-94*, pp. IV. 149-154.
- C18. Wang D.L. and Terman D. (1994): "Locally excitatory globally inhibitory oscillator networks: Theory and application to pattern segmentation," *Proceedings of IEEE Conference on Neural Networks for Signal Processing*, pp. 136-145. Ermioni, Greece.
- C19. Wang D.L. (1994): "Auditory stream segregation based on oscillatory correlation," *Proceedings of IEEE Workshop on Neural Networks for Signal Processing*, pp. 624-632.
- C20. Wang D.L. (1994): "An oscillation model of auditory stream segregation," *Proceedings of the International Conference on Pattern Recognition*, pp. C.198-200. Jerusalem, Israel.
- C21. Wang D.L. and Terman D. (1994): "Locally excitatory globally inhibitory oscillator networks: Theory and application to scene segmentation," *Proceedings of the 23rd Artificial Intelligence and Pattern Recognition Workshop*, SPIE Proceedings 2368, pp. 135-142. Washington, DC.
- C22. Wang D.L. and Terman D. (1994): "Synchrony and desynchrony in neural oscillator networks," *Advances in Neural Information Processing Systems (NIPS-94)*, pp. 199-206. Denver, CO.
- C23. Liu X.M., Wang D.L., and Ahalt S.C. (1995). On the temporal generalization capability of simple recurrent networks," *Proceedings of the 1995 SPIE Conference on Applications and Science of Artificial Neural Networks IV*, pp. 392-403. Orlando, FL.
- C24. Wang D.L. and Terman D. (1995): "Image segmentation based on oscillatory correlation," *Proceedings of WCNN-95*, pp. II. 521-525. Washington, DC.
- C25. Campbell S. and Wang D.L. (1995): "Relaxation oscillators with time delay coupling," *Proceedings of WCNN-95*, pp. I. 258-261.
- C26. Wang D.L. and Terman D. (1995): "Image segmentation by a neural oscillator network," *Proceedings of International Conference on Neural Information Processing*, pp. 722-726. Beijing.

C27. Wang D.L. and Terman D. (1996): "Image segmentation by neural oscillator networks," *Proceedings of ICNN-96*, pp. 1534-1539.

C28. Campbell S. and Wang D.L. (1996): "Loose synchrony in relaxation oscillator networks with time delays," *Proceedings of ICNN-96*, pp. 828-833.

C29. Wang D.L. and Yuwono B. (1996): "A neural model of sequential memory," *Proceedings of ICNN-96*, pp. 834-839.

C30. Campbell S. and Wang D.L. (1996): "Loose synchrony in networks of relaxation oscillators with time delays," *Proceedings of WCNN-96*, pp. 717-720.

C31. Shareef N., Wang D.L., and Yagel R. (1996): "Segmentation of medical data using locally excitatory globally inhibitory oscillator networks," *Proceedings of WCNN-96*, pp. 1245-1248.

C32. Wang D.L. and Yuwono B. (1996): "Incremental learning of complex temporal patterns," *Proceedings of WCNN-96*, pp. 757-762.

C33. Brown G.J. and Wang D.L. (1997): "Modelling the perceptual separation of concurrent vowels with a network of neural oscillators," *Proceedings of ICNN-97*, pp. 569-574.

C34. Campbell S. and Wang D.L. (1997): "Relaxation oscillator networks with time delays," *Proceedings of ICNN-97*, pp. 645-650.

C35. Cesmeli E. and Wang D.L. (1997): "Texture segmentation using Gaussian Markov random fields and LEGION," *Proceedings of ICNN-97*, pp. 1529-1534.

C36. Liu X.W. and Wang D.L. (1997): "Range image segmentation using an oscillatory network," *Proceedings of ICNN-97*, pp. 1656-1660.

C37. Wang D.L. (1997): "Object selection by a neural oscillator network," *Proceedings of International Conference on Neural Information Processing*, pp. 1137-1140.

C38. van der Kouwe A.J.W. and Wang D.L. (1997): "Temporal alignment, spatial spread and the linear independence criterion for blind separation of voices," *Proceedings of IEEE EMBS*, pp. 1994-1996.

C39. Wang D.L. (1998): "Object selection by oscillatory correlation," *Proceedings of International Joint Conference on Neural Networks (IJCNN-98)*, pp. 1182-1187.

C40. Campbell S. and Wang D.L. (1998): "Synchrony and desynchrony in integrate and fire oscillators," *Proceedings of IJCNN-98*, pp. 1498-1503.

C41. Liu X.W., Wang D.L., and Ramirez J.R. (1998): "Extracting hydrographic objects from satellite images using a two-layer neural network," *Proceedings of IJCNN-98*, pp. 897-902.

- C42. Chen K. and Wang D.L. (1998): "Perceiving spirals and inside/outside relations by a neural oscillator network," *Proceedings of IJCNN-98*, pp. 619-624.
- C43. Cesmeli E., Wang D.L., Lindsey D.L., and Todd J.T. (1998): "Motion segmentation using temporal block matching and LEGION," *Proceedings of IJCNN-98*, pp. 2069-2074.
- C44. Liu X., Wang D.L., and Ramirez J R. (1998): "A two-layer neural network for robust image segmentation and its application in revising hydrographic features," *International Archives of Photogrammetry and Remote Sensing*, vol. 32, part 3/1, pp. 464-472.
- C45. Liu X., Wang D.L., and Ramirez J.R. (1998): "Oriented statistical nonlinear smoothing filter," *Proceedings of the International Conference on Image Processing*, vol. 2, pp. 848-852.
- C46. Cesmeli E. and Wang D.L. (1998), "Gauss Markov Rasgele Alanlari ve Salingan Sinir Aglariyla Doku Bolutlemesi," SIU-98, Ankara TURKEY (in Turkish).
- C47. Chen K. and Wang D.L. (1999): "Perceiving without learning: from spirals to inside/outside relations," *Advances in Neural Information Processing Systems 11 (NIPS-98)*, MIT Press, pp. 10-16.
- C48. Liu X. and Wang D.L. (1999): "Perceptual organization based on temporal dynamics," *Proceedings of IJCNN-99*, 6 pages on CD-ROM.
- C49. Liu X. and Wang D.L. (1999): "A boundary-pair representation for perception modeling," *Proceedings of IJCNN-99*, 6 pages on CD-ROM.
- C50. Cesmeli E. and Wang D.L. (1999): "Image segmentation based on motion/luminance integration and oscillatory correlation," *Proceedings of IJCNN-99*, 6 pages on CD-ROM.
- C51. Brown G.J. and Wang D.L. (1999): "The separation of speech from interfering sounds: an oscillatory correlation approach," *Proceedings of IJCNN-99*, 6 pages on CD-ROM.
- C52. Chen K. and Wang D.L. (1999): "Image segmentation based on a dynamically coupled neural oscillator network," *Proceedings of IJCNN-99*, 6 pages on CD-ROM.
- C53. Brown G.J. and Wang D.L. (2000): "An oscillatory correlation framework for computational auditory scene analysis," *Proceedings of NIPS-99*, pp. 747-753.
- C54. Liu X. and Wang D.L. (2000): "Perceptual organization based on temporal dynamics," *Proceedings of NIPS-99*, pp. 38-44.
- C55. Liu X. and Wang D.L. (2000): "Spectral histograms for texton modeling and discrimination," *Proceedings of the Sixteenth World Computer Congress (Intelligent Information Processing)*, pp. 202-205.
- C56. Cesmeli E., Lindsey D.L., and Wang D.L. (2000): "An oscillatory correlation model of human motion perception," *Proceedings of IJCNN-00*, pp. IV.267-IV.272.

C57. Wu M., Wang D.L., and Brown G.J. (2001): "Pitch tracking based on statistical anticipation," *Proceedings of IJCNN-01*, pp. 866-871.

C58. Liu X. and Wang D.L. (2001): "A spectral histogram model for textures and texture discrimination," *Proceedings of IJCNN-01*, pp. 1083-1088.

C59. Hu G. and Wang D.L. (2001): "An extended model for speech segregation," *Proceedings of IJCNN-01*, pp. 1089-1094.

C60. Liu X. and Wang D.L. (2001): "Appearance-based recognition using perceptual components," *Proceedings of IJCNN-01*, pp. 1943-1948.

C61. Roman N., Wang D.L., and Brown G.J. (2001): "Speech segregation based on sound localization," *Proceedings of IJCNN-01*, pp. 2861-2866.

C62. Brown G.J., Barker J., and Wang D.L. (2001): "A neural oscillator sound separator for missing data speech recognition," *Proceedings of IJCNN-01*, pp. 2907-2912.

C63. Palomaki K., Brown G.J., and Wang D.L. (2001): "A binaural model for missing data speech recognition in noisy and reverberant conditions," *Web Proceedings of Workshop on Consistent and Reliable Acoustic Cues for Sound Analysis*, 4 pages.

C64. Hu G. and Wang D.L. (2001): "Speech segregation based on pitch tracking and amplitude modulation," *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 79-82.

C65. Liu X., Wang D.L., Srivastava A. (2001): "Image segmentation using local spectral histograms," *Proceedings of IEEE International Conference on Image Processing*, pp. 70-73.

C66. Chen K., Wang D.L., and Liu X. (2001): "Image segmentation by weight adaptation and oscillatory correlation," *Proceedings of International Conference on Neural Information Processing*, invited paper, 6 pages on CD-ROM.

C67. Hu G. and Wang D.L. (2002): "On amplitude modulation for monaural speech segregation," *Proceedings of IJCNN-02*, pp. 69-74.

C68. Roman N., Wang D.L., and Brown G.J. (2002): "Localization-based sound segregation," *Proceedings of IJCNN-02*, pp. 2299-2303.

C69. Wu M., Wang D.L., and Brown G.J. (2002): "A multi-pitch tracking algorithm for noisy speech," *Proceedings of IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP)*, pp. 369-372.

C70. Hu G. and Wang D.L. (2002): "Monaural speech segregation based on pitch tracking and amplitude modulation," *Proceedings of ICASSP-02*, pp. 553-556.

C71. Roman N., Wang D.L., and Brown G.J. (2002): "Localization-based sound segregation," *Proceedings of ICASSP-02*, pp. 1013-1016.

C72. Hu G. and Wang D.L. (2003): "Monaural speech separation," in *Advances in Neural Information Processing Systems (NIPS 2002)*, vol. 15, pp. 1221-1228, Cambridge MA: MIT Press.

C73. Hu G. and Wang D.L. (2003): "Separation of stop consonants," *Proceedings of ICASSP-03*, pp. II.749-752.

C74. Roman N. and Wang D.L. (2003): "Binaural tracking of multiple moving sources," *Proceedings of ICASSP-03*, pp. V.149-152.

C74. Shao Y. and Wang D.L. (2003): "Co-channel speaker identification using usable speech extraction based on multi-pitch tracking," *Proceedings of ICASSP-03*, vol. II.205-208.

C75. Wu M. and Wang D.L. (2003): "A one-microphone algorithm for reverberant speech enhancement," *Proceedings of ICASSP-03*, pp. I.844-847.

C76. Liu X., Srivastava A., and Wang D.L. (2003): "On intrinsic generalization of low dimensional representations of images for recognition," *Proceedings of IJCNN-03*, pp. 182-187.

C77. Srinivasan S. and Wang D.L. (2003): "Schema-based modeling of phonemic restoration," *Proceedings of EUROSPEECH-03*, pp. 2053-2056.

C78. Hu G. and Wang D.L. (2003): "Segregation of stop consonants from acoustic interference," *Proceedings of IEEE Workshop on Neural Networks for Signal Processing*, pp. 647-656.

C79. Roman N., Wang D.L., and Brown G.J. (2004): "A classification-based cocktail-party processor," in *Advances in Neural Information Processing Systems (NIPS 2003)*, vol. 16, Cambridge MA: MIT Press.

C80. Roman N. and Wang D.L. (2004): "Binaural sound segregation for multisource reverberant environments," in *Proceedings of ICASSP-04*, pp. II.373-376.

C81. Wang D.L. (2004): "A comparison of CNN and LEGION networks," *Proceedings of IJCNN-04*, pp. 1735-1740.

C82. Hu G. and Wang D.L. (2004): "Auditory segmentation based on event detection," *Proceedings of ISCA Tutorial and Research Workshop on Statistical and Perceptual Audio Processing (SAPA)*, 6 pages.

C83. Srinivasan S., Roman N., and Wang D.L. (2004): "On binary and ratio time-frequency masks for robust speech recognition," *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*, 4 pages.

C84. Shao Y. and Wang D.L. (2004): "Model-based sequential organization for cochannel speaker identification," *Proceedings of ICSLP-04*, 4 pages.

C85. Wu M. and Wang D.L. (2005): "A two-stage algorithm for enhancement of reverberant speech," in *Proceedings of ICASSP-05*, pp. I.1085-1088.

C86. Hu G. and Wang D.L. (2005): "Separation of fricatives and affricates," in *Proceedings of ICASSP-05*, pp. I.1001-1004.

C87. Li Y. and Wang D.L. (2005): "Detecting pitch of singing voice in polyphonic audio," in *Proceedings of ICASSP-05*, pp. III.17-20.

C88. Srinivasan S. and Wang D.L. (2005): "Robust speech recognition by integrating speech separation and hypothesis testing," in *Proceedings of ICASSP-05*, pp. I.89-92.

C89. Srinivasan S. and Wang D.L. (2005): "Modeling the perception of multitalker speech," *Proceedings of Interspeech-05*, pp. 1265-1268.

C90. Roman N. and Wang D.L. (2005): "A pitch-based model for separation of reverberant speech," *Proceedings of Interspeech-05*, pp. 2109-2112.

C91. Pedersen M.S., Wang D.L., Larsen J., and Kjems U. (2005): "Overcomplete blind source separation by combining ICA and binary time-frequency masking," *Proceedings of IEEE Workshop on Machine Learning for Signal Processing*, pp. 15-20.

C92. Pedersen M.S., Wang D.L., Larsen J., and Kjems U. (2006): "Separating underdetermined convolutive speech mixtures," *Proceedings of International Conference on Independent Component Analysis and Blind Source Separation, Springer Lecture Notes in Computer Science*, vol. 3889, pp. 674-681.

C93. Rand R.S. and Wang D.L. (2006). "A biologically-inspired neural oscillator network for geospatial analysis," *Proceedings of SPIE Defense Symposium*, 10 pages.

C94. Wang D.L. and Hu G. (2006): "Unvoiced speech segregation," *Proceedings of ICASSP-06*, pp. V.953-956.

C95. Shao Y. and Wang D.L. (2006): "Robust speaker recognition using binary time-frequency masks," *Proceedings of ICASSP-06*, pp. I.645-648.

C96. Srinivasan S. and Wang D.L. (2006): "A supervised learning approach to uncertainty decoding for robust speech recognition," *Proceedings of ICASSP-06*, pp. I.297-300.

C97. Roman N., Srinivasan S., and Wang D.L. (2006): "Speech recognition in multisource reverberant environments with binaural inputs," *Proceedings of ICASSP-06*, pp. I.309-312.

C98. Srinivasan S., Shao Y., Jin Z., and Wang D.L. (2006): "A computational auditory scene analysis system for robust speech recognition," *Proceedings of Interspeech-06*, pp. 73-76.

C99. Li Y. and Wang D.L. (2006): "Singing voice separation from monaural recordings," *Proceedings of the International Symposium on Music Information Retrieval (ISMIR)*, pp. 176-179.

C100. Li Y. and Wang D.L. (2007): "Pitch detection in polyphonic music using instrument tone models," *Proceedings of ICASSP-07*, pp. II.481-484.

C101. Shao Y., Srinivasan S., and Wang D.L. (2007): "Incorporating auditory feature uncertainties in robust speaker identification," *Proceedings of ICASSP-07*, pp. IV.277-280.

C102. Srinivasan S., Roman N., and Wang D.L. (2007): "Exploiting uncertainties for binaural speech recognition," *Proceedings of ICASSP-07*, pp. IV.789-792.

C103. Jin Z. and Wang D.L. (2007): "A supervised learning approach to monaural segregation of reverberant speech," *Proceedings of ICASSP-07*, pp. IV.921-924.

C104. Li Y. and Wang D.L. (2008): "Musical sound separation using pitch-based labeling and binary time-frequency masking," *Proceedings of ICASSP-08*, pp. 173-176.

C105. Li Y. and Wang D.L. (2008): "On the optimality of ideal binary time-frequency masks," *Proceedings of ICASSP-08*, pp. 3501-3504.

C106. Shao Y. and Wang D.L. (2008): "Robust speaker identification using auditory features and computational auditory scene analysis," *Proceedings of ICASSP-08*, pp. 1589-1592.

C107. Boldt J.B., Kjems U., Pedersen M.S., Lunner T., and Wang D.L. (2008): "Estimation of the ideal binary mask using directional systems," *Proceedings of International Workshop on Acoustic Echo and Noise Control (IWAENC-08)*, 4 pages.

C108. Woodruff J., Li Y., and Wang D.L. (2008): "Resolving overlapping harmonics for monaural musical sound separation using pitch and common amplitude modulation," *Proceedings of ISMIR-08*, pp. 538-543.

C109. Hu K., Divenyi P., Ellis D.P.W., Jin Z., Shinn-Cunningham B.G., and Wang D.L. (2008): "Preliminary intelligibility tests of a monaural speech segregation system," *ISCA Tutorial and Research Workshop on Statistical and Perceptual Audition (SAPA)*, 6 pages.

C110. Jan T., Wang W., Wang D.L. (2009): "A multistage approach for blind separation of convolutive speech mixtures," *Proceedings of ICASSP-09*, pp. 1713-1716.

C111. Woodruff J. and Wang D.L. (2009): "On the role of localization cues in binaural segregation of reverberant speech," *Proceedings of ICASSP-09*, pp. 2205-2208.

C112. Hu K. and Wang D.L. (2009): "Incorporating spectral subtraction and noise type for unvoiced speech segregation," *Proceedings of ICASSP-09*, pp. 4425-4428.

C113. Shao Y., Jin Z., Wang D.L., and Srinivasan S. (2009): "An auditory-based feature for robust speech recognition," *Proceedings of ICASSP-09*, pp. 4625-4628.

C114. Jin Z. and Wang D.L. (2009): "Learning to maximize signal-to-noise ratio for reverberant speech segregation," *Proceedings of ICASSP-09*, pp. 4689-4692.

C115. Wu B., Zhou Y., Yan L., Yuan J., Li R., and Wang D.L. (2009): "Object detection from HS/MS and multi-platform remote sensing images by integrating biologically and geometrically inspired approaches," *Proceedings of ASPRS-09*, 7 pages.

C116. Quiles M., Wang D.L., Zhao L., Romero R.A.F., and D.-S. Huang (2009): "An oscillatory correlation model of object-based attention," *Proceedings of IJCNN-09*, pp. 2596-2602.

C117. Yuan J., Wang D.L., Wu B., Yan L., and Li R. (2009): "Automatic road extraction from satellite imagery using LEGION networks," *Proceedings of IJCNN-09*, pp. 3471-3476.

C118. Jin Z. and Wang D.L. (2010): "A multipitch tracking algorithm for noisy and reverberant speech," *Proceedings of ICASSP-10*, pp. 4218-4221.

C119. Woodruff J. and Wang D.L. (2010): "Integrating monaural and binaural analysis for localization of multiple sound sources in noisy and reverberant conditions," *Proceedings of ICASSP-10*, pp. 2706-2709.

C120. Yan L., Yuan J., Cheng L., Li R., and Wang D.L. (2010): "A biologically and geometrically inspired approach to target extraction from multiple-source remote-sensing imagery," *Proceedings of ASPRS-10*, 8 pages.

C121. Kjems U., Boldt J.B., Pedersen M.S., Lunner T., and Wang D.L. (2010): "Speech intelligibility of ideal binary masked mixtures," *Proceedings of EUSIPCO-10*, pp. 1909-1913.

C122. Hu K. and Wang D.L. (2010): "Unsupervised sequential organization for cochannel speech separation," *Proceedings of INTERSPEECH-10*, pp. 2790-2793.

C123. Hu K. and Wang D.L. (2010): "Unvoiced speech segregation based on CASA and spectral subtraction," *Proceedings of INTERSPEECH-10*, pp. 2786-2789.

C124. Woodruff J., Prabhavalkar R., Fosler-Lussier E., and Wang D.L. (2010): "Combining monaural and binaural evidence for reverberant speech segregation," *Proceedings of INTERSPEECH-10*, pp. 406-409.

C125. Woodruff J. and Wang D.L. (2011): "Directionality-based speech enhancement for hearing aids," *Proceedings of ICASSP-11*, pp. 297-300.

C126. Hsu C.-L., Wang D.L. and Jang J.R. (2011): "A trend estimation algorithm for singing pitch detection in musical recordings," *Proceedings of ICASSP-11*, pp. 393-396.

C127. Narayanan A. and Wang D.L. (2011): "On the use of ideal binary masks for improving phonetic classification," *Proceedings of ICASSP-11*, pp. 4632-4635.

C128. Hu K. and Wang D.L. (2011): "An approach to sequential grouping in cochannel speech," *Proceedings of ICASSP-11*, 4636-4639.

C129. Narayanan A., X. Zhao, Wang D.L., and Fosler-Lussier E. (2011): "Robust speech recognition using multiple prior models for speech reconstruction," *Proceedings of ICASSP-11*, pp. 4800-4803.

C130. Han K. and Wang D.L. (2011): "An SVM based classification approach to speech separation," *Proceedings of ICASSP-11*, pp. 5212-5215.

C131. Yuan J., Wang D.L., and Li R. (2011): "Image segmentation based on local spectral histograms and linear regression," *Proceedings of IJCNN-11*, pp. 482-488.

C132. Zhao X., Shao Y., and Wang D.L. (2011): "Robust speaker identification using a CASA front-end," *Proceedings of ICASSP-11*, pp. 5468-5471.

C133. Woodruff J. and Wang D.L. (2012): "Binaural speech segregation based on pitch and azimuth tracking," *Proceedings of ICASSP-12*, pp. 241-244.

C134. Han K. and Wang D.L. (2012): "On generalization of classification based speech separation," *Proceedings of ICASSP-12*, pp. 4541-4544.

C135. Hu K. and Wang D.L. (2012): "SVM-based separation of unvoiced-voiced portions in cochannel speech," *Proceedings of ICASSP-12*, pp. 4545-4548.

C136. Narayanan A. and Wang D.L. (2012): "On the role of binary mask pattern in automatic speech recognition," *Proceedings of INTERSPEECH-12*, pp. 1239-1242.

C137. Wang Y., Han K., and Wang D.L. (2012): "Acoustic features for classification based speech separation," *Proceedings of INTERSPEECH-12*, pp. 1532-1535.

C138. Wang Y. and Wang D.L. (2012): "Boosting classification based speech separation using temporal dynamics," *Proceedings of INTERSPEECH-12*, pp. 1528-1531.

C139. Wang Y. and Wang D.L. (2012): "Cocktail party processing via structured prediction," *Proceedings of NIPS-12*, pp. 224-232.

C140. Han K. and Wang D.L. (2013): "Learning invariant features for speech separation," *Proceedings of ICASSP-13*, pp. 7492-7496.

C141. Wang Y. and Wang D.L. (2013): "Feature denoising for speech separation in unknown noisy environments," *Proceedings of ICASSP-13*, pp. 7472-7476.

C142. Zhao X. and Wang D.L. (2013): "Analyzing noise robustness of MFCC and GFCC features in speaker identification," *Proceedings of ICASSP-13*, pp. 7204-7208.

C143. Williamson D.S., Wang Y. and Wang D.L. (2013): "A sparse representation approach for perceptual quality improvement of separated speech," *Proceedings of ICASSP-13*, pp. 7015-7019.

C144. Narayanan A. and Wang D.L. (2013): "Ideal ratio mask estimation using deep neural networks for robust speech recognition," *Proceedings of ICASSP-13*, pp. 7092-7096.

C145. Narayanan A. and Wang D.L. (2013): "Coupling binary masking and robust ASR," *Proceedings of ICASSP-13*, pp. 6817-6821.

C146. Wang Y. and Wang D.L. (2014): "A structure-preserving training target for supervised speech separation," *Proceedings of ICASSP-14*, pp. 6148-6152.

C147. Chen J., Wang Y. and Wang D.L. (2014): "A feature study for classification-based speech separation at very low signal-to-noise ratio," *Proceedings of ICASSP-14*, pp. 7089-7093.

C148. Narayanan A. and Wang D.L. (2014): "Joint noise adaptive training for robust automatic speech recognition," *Proceedings of ICASSP-14*, pp. 2523-2527.

C149. Zhao X., Wang Y., and Wang D.L. (2014): "Robust speaker identification in noisy and reverberant conditions," *Proceedings of ICASSP-14*, pp. 4025-4029.

C150. Williamson D.S., Wang Y., and Wang D.L. (2014): "A two-stage approach for improving the perceptual quality of separated speech," *Proceedings of ICASSP-14*, pp. 7084-7088.

C151. Han K., Wang Y., and Wang D.L. (2014): "Learning spectral mapping for speech dereverberation," *Proceedings of ICASSP-14*, pp. 4661-4665.

C152. Han K., Wang Y., and Wang D.L. (2014): "Neural networks for supervised pitch tracking in noise," *Proceedings of ICASSP-14*, pp. 1502-1506.

C153. Zhang X.-L. and Wang D.L. (2014): "Boosted deep neural networks and multi-resolution cochleagram features for voice activity detection," *Proceedings of INTERSPEECH-14*, pp. 1534-1538.

C154. Jiang Y., Wang D.L., and Liu R.S. (2014): "Binaural deep neural network classification for reverberant speech segregation," *Proceedings of INTERSPEECH-14*, pp. 2400-2404.

C155. Wang Y. and Wang D.L. (2015): "A deep neural network for time-domain signal reconstruction," *Proceedings of ICASSP-15*, pp. 4390-4394.

C156. Zhao X., Wang Y., and Wang D.L. (2015): "Deep neural networks for cochannel speaker identification," *Proceedings of ICASSP-15*, pp. 4824-4828.

C157. Williamson D.S., Wang Y., and Wang D.L. (2015): "Deep neural networks for estimating speech model activations," *Proceedings of ICASSP-15*, pp. 5113-5117.

C158. Chen J., Wang, Y., and Wang D.L. (2015): "Noise perturbation improves supervised speech separation," *Proceedings of LVA/ICA-15*, pp. 83-90.

C159. Han K., He Y., Bagchi D., Fosler-Lussier E., and Wang D.L. (2015): "Deep neural network based spectral feature mapping for robust speech recognition," *Proceedings of INTERSPEECH-15*, pp. 2484-2488.

C160. Zhang X.-L. and Wang D.L. (2015): "Multi-resolution stacking for speech separation based on boosted DNN," *Proceedings of INTERSPEECH-15*, pp. 1745-1749.

C161. Liu Y. and Wang D.L. (2015): "Speaker-dependent multipitch tracking using deep neural networks," *Proceedings of INTERSPEECH-15*, pp. 3279-3283.

C162. Wang Z.-Q. and Wang D.L. (2015): "Joint training of speech separation, filterbank and acoustic model for robust automatic speech recognition," *Proceedings of INTERSPEECH-15*, pp. 2839-2843.

C163. Zhao Y., Wang D.L., Merks I., and Zhang T. (2016): "DNN-based enhancement of noisy and reverberant speech," *Proceedings of ICASSP-16*, pp. 6525-6529.

C164. Wang Z.-Q. and Wang D.L. (2016): "Robust speech recognition from ratio masks," *Proceedings of ICASSP-16*, pp. 5720-5724.

C165. Wang Z.-Q., Zhao Y., and Wang D.L. (2016): "Phoneme-specific speech separation," *Proceedings of ICASSP-16*, pp. 146-150.

C166. Liu Y. and Wang D.L. (2016): "Robust pitch tracking in noisy speech using speaker-dependent deep neural networks," *Proceedings of ICASSP-16*, pp. 5255-5259.

C167. Williamson D.S., Wang Y., and Wang D.L. (2016): "Complex ratio masking for joint enhancement of magnitude and phase," *Proceedings of ICASSP-16*, pp. 5220-5224.

C168. Delfarah M. and Wang D.L. (2016): "A feature study for masking-based reverberant speech separation," *Proceedings of INTERSPEECH-16*, pp. 555-559.

C169. Chen J. and Wang D.L. (2016): "Long short-term memory for speaker generalization in supervised speech separation," *Proceedings of INTERSPEECH-16*, pp. 3314-3318.

C170. Liu Y. and Wang D.L. (2017): "Time and frequency domain long short-term memory for noise robust pitch tracking," *Proceedings of ICASSP-17*, pp. 5600-5604.

C171. Chang J. and Wang D.L. (2017): "Robust speaker recognition based on DNN/i-vectors and speech separation," *Proceedings of ICASSP-17*, pp. 5415-5419.

C172. Williamson D. and Wang D.L. (2017): "Speech dereverberation and denoising using complex ratio masks," *Proceedings of ICASSP-17*, pp. 5590-5594.

C173. Zhao Y., Wang Z.-Q., and Wang D.L. (2017): "A two-stage algorithm for noisy and reverberant speech enhancement," *Proceedings of ICASSP-17*, pp. 5580-5584.

C174. Wang Z.-Q. and Wang D.L. (2017): "Recurrent deep stacking networks for supervised speech separation," *Proceedings of ICASSP-17*, pp. 71-75.

C175. Wang Z.-Q. and Wang D.L. (2017): "Unsupervised speaker adaptation of batch normalized acoustic models for robust ASR," *Proceedings of ICASSP-17*, pp. 4890-4894.

C176. Zhang X., Wang Z.-Q., and Wang D.L. (2017): "A speech enhancement algorithm by iterating single- and multi-microphone processing and its application to robust ASR," *Proceedings of ICASSP-17*, pp. 276-280.

C177. Zhang X. and Wang D.L. (2017): "Binaural reverberant speech separation based on deep neural networks," *Proceedings of INTERSPEECH-17*, pp. 2018-2022.

C178. Delfarah M. and Wang D.L. (2018): "Recurrent neural networks for cochannel speech separation in reverberant environments," *Proceedings of ICASSP-18*, pp. 5404-5408.

C179. Liu Y. and Wang D.L. (2018): "A CASA approach to deep learning based speaker-independent co-channel speech separation," *Proceedings of ICASSP-18*, pp. 5399-5403.

C180. Liu Y. and Wang D.L. (2018): "Permutation invariant training for speaker-independent multi-pitch tracking," *Proceedings of ICASSP-18*, pp. 5594-5598.

C181. Pandey A. and Wang D.L. (2018): "On adversarial training and loss functions for speech enhancement," *Proceedings of ICASSP-18*, pp. 5414-5418.

C182. Tan K., Chen J., and Wang D.L. (2018): "Gated residual networks with dilated convolutions for supervised speech separation," *Proceedings of ICASSP-18*, pp. 21-25.

C183. Wang P. and Wang D.L. (2018): "Utterance-wise recurrent dropout and iterative speaker adaptation for robust monaural speech recognition," *Proceedings of ICASSP-18*, pp. 4814-4818.

C184. Wang P. and Wang D.L. (2018): "Filter-and-convolve: a CNN based multichannel complex concatenation acoustic model," *Proceedings of ICASSP-18*, pp. 5564-5568.

C185. Wang Z.-Q. and Wang D.L. (2018): "Mask weighted STFT ratios for relative transfer function estimation and its application to robust ASR," *Proceedings of ICASSP-18*, pp. 5619-5623.

C186. Wang Z.-Q. and Wang D.L. (2018): "On spatial features for supervised speech separation and its application to beamforming and robust ASR," *Proceedings of ICASSP-18*, pp. 5709-5713.

C187. Zhao Y., Wang D.L., Xu B., and Zhang T. (2018): "Late reverberation suppression using recurrent neural networks with long short-term memory," *Proceedings of ICASSP-18*, pp. 5434-5438.

C188. Zhang H. and Wang D.L. (2018): "Deep learning for acoustic echo cancellation in noisy and double-talk scenarios," *Proceedings of INTERSPEECH-18*, pp. 3239-3243.

C189. Wang Z.-Q., Le Roux J., Wang D.L., Hershey, J.R. (2018): "End-to-end speech separation with unfolded iterative phase reconstruction," *Proceedings of INTERSPEECH-18*, pp. 2708-2712.

C190. Wang Z.-Q., Zhang X., and Wang D.L. (2018): "Robust TDOA estimation based on time-frequency masking and deep neural networks," *Proceedings of INTERSPEECH-18*, pp. 322-326.

C191. Wang Z.-Q. and Wang D.L. (2018): "All-neural multi-channel speech enhancement," *Proceedings of INTERSPEECH-18*, pp. 3234-3238.

C192. Wang Z.-Q. and Wang D.L. (2018): "Integrating spectral and spatial features for multi-channel speaker separation," *Proceedings of INTERSPEECH-18*, pp. 2718-2722.

C193. Tan K. and Wang D.L. (2018): "A two-stage approach to noisy cochannel speech separation with gated residual networks," *Proceedings of INTERSPEECH-18*, pp. 3484-3488.

C194. Tan K. and Wang D.L. (2018): "A convolutional recurrent neural network for real-time speech enhancement," *Proceedings of INTERSPEECH-18*, pp. 3229-3233.

C195. Pandey A. and Wang D.L. (2018): "A new framework for supervised speech enhancement in the time domain," *Proceedings of INTERSPEECH-18*, pp. 1136-1140.

C196. Chakrabarty S., Wang D.L., Habets E.A.P. (2018): "Time-frequency masking based online speech enhancement with multi-channel data using convolutional neural networks," *Proceedings of IWAENC-18*, pp. 476-480.

C197. Zhang H. and Wang D.L. (2018): "Deep learning for acoustic echo cancellation in noisy and double-talk scenarios," *Proceedings of INTERSPEECH-18*, pp. 3239-3243.

C198. Wang Z.-Q., Le Roux J., Wang D.L., Hershey, J.R. (2018): "End-to-end speech separation with unfolded iterative phase reconstruction," *Proceedings of INTERSPEECH-18*, pp. 2708-2712.

C199. Wang Z.-Q., Zhang X., and Wang D.L. (2018): "Robust TDOA estimation based on time-frequency masking and deep neural networks," *Proceedings of INTERSPEECH-18*, pp. 322-326.

C200. Wang Z.-Q. and Wang D.L. (2018): "All-neural multi-channel speech enhancement," *Proceedings of INTERSPEECH-18*, pp. 3234-3238.

C201. Wang Z.-Q. and Wang D.L. (2018): "Integrating spectral and spatial features for multi-channel speaker separation," *Proceedings of INTERSPEECH-18*, pp. 2718-2722.

C202. Tan K. and Wang D.L. (2018): "A two-stage approach to noisy cochannel speech separation with gated residual networks," *Proceedings of INTERSPEECH-18*, pp. 3484-3488.

C203. Tan K. and Wang D.L. (2018): "A convolutional recurrent neural network for real-time speech enhancement," *Proceedings of INTERSPEECH-18*, pp. 3229-3233.

C204. Pandey A. and Wang D.L. (2018): "A new framework for supervised speech enhancement in the time domain," *Proceedings of INTERSPEECH-18*, pp. 1136-1140.

C205. Pandey A. and Wang D.L. (2018): "A new framework for supervised speech enhancement in the time domain," *Proceedings of INTERSPEECH-18*, pp. 1136-1140.

C206. Pandey A. and Wang D.L. (2019): "Exploring deep complex networks for complex spectrogram enhancement," *Proceedings of ICASSP-19*, pp. 6885-6889.

C207. Pandey A. and Wang D.L. (2019): "TCNN: Temporal convolutional neural network for real-time speech enhancement in the time domain," *Proceedings of ICASSP-19*, pp. 6875-6879.

C208. Tan K. and Wang D.L. (2019): "Complex spectral mapping with a convolutional recurrent network for monaural speech enhancement," *Proceedings of ICASSP-19*, pp. 6865-6869.

C209. Tan K., Zhang X., and Wang D.L. (2019): "Real-time speech enhancement using an efficient convolutional recurrent network for dual-microphone mobile phones in close-talk scenarios," *Proceedings of ICASSP-19*, pp. 5751-5755.

C210. Xie J., Jin D., Zhang W., Zhang X.-L., Chen J., and Wang D.L. (2019): "Robust sparse multichannel active noise control," *Proceedings of ICASSP-19*, pp. 521-525.

C211. Wang Z.-Q., Tan K., and Wang D.L. (2019): "Deep learning based phase reconstruction for speaker separation: A trigonometric perspective," *Proceedings of ICASSP-19*, pp. 71-75.

C212. Zhang H., Tan K., and Wang D.L. (2019): "Deep learning for joint acoustic echo and noise cancellation with nonlinear distortions," *Proceedings of INTERSPEECH-19*, pp. 4255-4259.

C213. Taherian H., Wang Z.-Q., and Wang D.L. (2019): "Deep learning based multi-channel speaker recognition in noisy and reverberant environments," *Proceedings of INTERSPEECH-19*, pp. 4070-4074.

C214. Wang P. and Wang D.L. (2019): "Enhanced spectral features for distortion-independent acoustic modeling," *Proceedings of INTERSPEECH-19*, pp. 476-480.

C215. Wang P., Tan K., and Wang D.L. (2019): "Bridging the gap between monaural speech enhancement and recognition with distortion-independent acoustic modeling," *Proceedings of INTERSPEECH-19*, pp. 471-475.

C216. Delfarah M., Liu Y., and Wang D.L. (2020): "Talker-independent speaker separation in reverberant conditions," *Proceedings of ICASSP-20*, pp. 8723-8727.

C217. Tan K. and Wang D.L. (2020): "Improving robustness of deep learning based monaural speech enhancement against processing artifacts," *Proceedings of ICASSP-20*, pp. 6914-6918.

C218. Pandey A. and Wang D.L. (2020): "Densely connected neural network with dilated convolutions for real-time speech enhancement in the time domain," *Proceedings of ICASSP-20*, pp. 6629-6633.

C219. Liu Y., Delfarah M., and Wang D.L. (2020): "Deep CASA for talker-independent monaural speech separation," *Proceedings of ICASSP-20*, pp. 6354-6358.

C220. Wang H. and Wang D.L. (2020): "Time-frequency loss for CNN based speech super-resolution," *Proceedings of ICASSP-20*, pp. 861-865.

C221. Wang Z.-Q. and Wang D.L. (2020): "Multi-microphone complex spectral mapping for speech dereverberation," *Proceedings of ICASSP-20*, pp. 486-490.

C222. Zhang H. and Wang D.L. (2020): "A deep learning approach to active noise control," *Proceedings of INTERSPEECH-20*, pp. 1141-1145.

C223. Zhao Y. and Wang D.L. (2020): "Noisy-reverberant speech enhancement using DenseUNet with time-frequency attention," *Proceedings of INTERSPEECH-20*, pp. 3261-3265.

C224. Pandey A. and Wang D.L. (2020): "Learning complex spectral mapping for speech enhancement with improved cross-corpus generalization," *Proceedings of INTERSPEECH-20*, pp. 4511-4515.

C225. Li H., Wang D.L., Zhang X., and Gao G. (2020): "Frame-level signal-to-noise ratio estimation using deep learning," *Proceedings of INTERSPEECH-20*, pp. 4626-4630.

C226. Wang Z.-Q. and Wang D.L. (2021): "Count and separate: incorporating speaker counting for continuous speaker separation," *Proceedings of ICASSP-21*, pp. 11-15.

C227. Zhang Y., Liu Y., and Wang D.L. (2021): "Complex ratio masking for singing voice separation," *Proceedings of ICASSP-21*, pp. 41-45.

C228. Taherian H. and Wang D.L. (2021): “Time-domain loss modulation based on overlap ratio for monaural conversational speaker separation,” *Proceedings of ICASSP-21*, pp. 5744-5748.

C229. Tan K., Zhang X., and Wang D.L. (2021): “Real-time speech enhancement for mobile communication based on dual-channel complex spectral mapping,” *Proceedings of ICASSP-21*, pp. 6134-6138.

C230. Tan K. and Wang D.L. (2021): “Compressing deep neural networks for efficient speech enhancement,” *Proceedings of ICASSP-21*, pp. 8358-8362.

C231. Zhang H. and Wang D.L. (2021a): “A deep learning approach to multi-channel and multi-microphone acoustic echo cancellation,” *Proceedings of INTERSPEECH-21*, pp. 1139-1143.

C232. Zhang H. and Wang D.L. (2021b): “A deep learning method to multi-channel active noise control,” *Proceedings of INTERSPEECH-21*, pp. 681-685.

C233. Pandey A., Xu B., Kumar A., Donley J., Calamia P., and Wang D.L. (2022a): “TPARN: Triple-path attentive recurrent network for time-domain multichannel speech enhancement,” *Proceedings of ICASSP-22*, pp. 6497-6501.

C234. Pandey A., Xu B., Kumar A., Donley J., Calamia P., and Wang D.L. (2022b): “Multichannel speech enhancement without beamforming,” *Proceedings of ICASSP-22*, pp. 6502-6506.

C235. Taherian H., Tan K., and Wang D.L. (2022): “Location-based training for multi-channel talker-independent speaker separation,” *Proceedings of ICASSP-22*, pp. 696-700.

C236. Wang H. and Wang D.L. (2022): “Cross-domain speech enhancement with a neural cascade architecture,” *Proceedings of ICASSP-22*, pp. 7862-7866.

C237. Wang H., Zhang X., and Wang D.L. (2022): “Attention-based fusion for bone-conducted and air-conducted speech enhancement in the complex domain,” *Proceedings of ICASSP-22*, pp. 7757-7761.

C238. Wang H., Qian Y., Wang X., Wang Y., Wang C., Liu S., Yoshioka T., Li J., and Wang D.L. (2022): “Improving noise robustness of contrastive speech representation learning with speech reconstruction,” *Proceedings of ICASSP-22*, pp. 6062-6066.

C239. Wang Z.-Q. and Wang D.L. (2022): “Localization based sequential grouping for continuous speech separation,” *Proceedings of ICASSP-22*, pp. 281-285.

C240. Yu F., et al. (2022): “Summary on the ICASSP 2022 multi-channel multi-party meeting transcription grand challenge,” *Proceedings of ICASSP-22*, pp. 9156-9160.

C241. Zhang H. and Wang D.L. (2022): “Neural cascade architecture for joint acoustic echo and noise suppression,” *Proceedings of ICASSP-22*, pp. 671-675.

Abstracts

Invited:

AI1. Brown G.J. and Wang D.L. (1999): "An oscillatory correlation framework for the separation of speech from interfering sounds," *Journal of the Acoustical Society of America*, vol. 105, p. 1307. (Invited presentation at the Joint Meeting of the Acoustical Society of America and the European Acoustical Association, Berlin, Germany, March 1999.)

AI2. Wang D.L. and Hu G. (2003): "Monaural speech segregation," *Journal of the Acoustical Society of America*, vol. 113, p. 2229. (Invited presentation at the Meeting of the Acoustical Society of America, Nashville TN, April 2003.)

AI3. Wang D.L. (2006): "Cocktail party processing," In: *Advances in Artificial Intelligence – IBERAMIA-SBIA 2006* (Springer Lecture Series in Computer Science), p. 6. (Invited presentation at the International Joint Conference of IBERAMIA/SBIA/SBRN, Ribeiro Preto, Brazil, October 2006.)

AI4. Li Y. and Wang D.L. (2007): "Separation of singing voice from music accompaniment from monaural recordings," *Journal of the Acoustical Society of America*, vol. 122, p. 2989. (Invited presentation at the Meeting of the Acoustical Society of America, New Orleans LA, Nov. 2007.)

AI5. Jin Z. and Wang D.L. (2008): "Monaural segregation of reverberant speech," *Journal of the Acoustical Society of America*, vol. 123, p. 2979. (Invited presentation at the Meeting of the Acoustical Society of America, Paris France, June 2008.)

AI6. Wang D.L. (2008): "Computational auditory scene analysis and its potential application for hearing aids," *Proceedings of International Hearing Aid Research Conference (IHCON)* (Invited presentation at IHCON-08, Lake Tahoe CA, Aug. 2008.)

AI7. Wang D.L. (2015): "From auditory masking to supervised separation: A tale of improving intelligibility of noisy speech for hearing-impaired listeners," *Proceedings of Conference on Implantable Auditory Prosthesis (CIAP)* (Invited presentation at CIAP-15, Lake Tahoe CA, July 2015.)

Refereed:

AR1. Cesmeli E., Lindsey D.T., and Wang D.L. (1999): "Integration of static luminance and motion cues: a step towards solving the blank wall problem," *Proceedings of the Association for Research in Vision and Ophthalmology (ARVO) Annual Meeting*.

AR2. Wang D.L. (1999): "Object-based selection by a neural oscillator network," *Proceedings of the Third International Conference on Cognitive and Neural Systems*.

AR3. Kristjansson A., Wang D.L., and Nakayama K. (1999): "Perceptual organization in visual search," *Proceedings of the Object Perception and Memory (OPAM) Annual Workshop*.

AR4. Nakayama K., Wang D.L., and Kristjansson A. (2000): "Efficient visual search not explained by local feature mechanisms or top down guidance," *Proceedings of the Association for Research in Vision and Ophthalmology (ARVO) Annual Meeting*.

AR5. Brungart D., Simpson B.D., Chang P., and Wang D.L. (2005): "A binary masking technique for isolating energetic masking in speech perception," *Journal of the Acoustical Society of America*, vol. 117, p. 2484.

AR6. Wang D.L., Kjems U., Pedersen M.S., Boldt J.B., Lunner T. (2008): "Speech perception in noise with binary gains," *Journal of the Acoustical Society of America*, vol. 123, p. 3066.

AR7. Wang D.L., Kjems U., Pedersen M.S., Boldt J.B., Lunner T. (2008): "Speech intelligibility improvements in noise with ideal binary time-frequency masking," *Proceedings of IHCON-08*.

AR8. Divenyi P., Livingston N., Lammert A., Hu K., and Wang D.L. (2010): "Age effects in the understanding of noisy speech denoised by estimated ideal binary masks," *Journal of the Acoustical Society of America*, vol. 127, p. 1902.

AR9. Yuan J., Wang D.L., Liu J., Yan L., and Li R. (2011): "Texture region extraction from multispectral and hyperspectral images by classifying local spectral histograms," *Proceedings of ASPRS-11*.

AR10. Healy E.W., Yoho S.E., Wang Y., Apoux F.C., Youngdahl C.L., and Wang D.L. (2014): "An algorithm to improve speech recognition in noise for hearing-impaired listeners: Consonant identification and articulatory feature transmission," *Journal of the Acoustical Society of America*, vol. 135, p. 2412.

AR11. Healy E.W., Yoho S.E., Wang Y., and Wang D.L. (2014): "A classification algorithm to improve speech recognition in noise for hearing-impaired listeners," *Proceedings of International Hearing Aid Research Conference (IHCON)*.

AR12. Healy E.W., Yoho S.E., Chen J., Wang Y., and Wang D.L. (2015): "An algorithm that generalizes to novel noise segments to improve speech intelligibility for hearing-impaired listeners," *Journal of the Acoustical Society of America*, vol. 137, p. 2208.

AR13. Chen J., Wang Y., Yoho S.E., Wang D.L., and Healy E.W. (2016): "An algorithm to increase speech intelligibility for hearing-impaired listeners in entirely novel noises," *Journal of the Acoustical Society of America*, vol. 139, p. 1995.

AR14. Healy E.W., Delfarah M., Vasko J., Carter B., and Wang D.L. (2017): “Can a trained deep neural network be implemented into hearing technology?” *Journal of the Acoustical Society of America*, vol. 141, p. 3813.

AR15. Healy E.W., Delfarah M., Vasko J., Carter B., and Wang D.L. (2017): “An algorithm to increase intelligibility for hearing-impaired listeners in the presence of a competing talk,” *Journal of the Acoustical Society of America*, vol. 141, p. 3819.

AR16. Vasko J., Healy E.W., and Wang D.L. (2018): “The optimal noise-rejection threshold for normal and impaired hearing” *Journal of the Acoustical Society of America*, vol. 143, p. 1940.

AR17. Zhao Y., Wang D.L., Johnson E.M., and Healy E.W. (2018): “A deep learning based segregation algorithm to improve speech intelligibility of hearing-impaired listeners in reverberant-noisy conditions,” *Proceedings of the International Hearing Aid Research Conference (IHCON)*.

AR18. Healy E.W., Delfarah M., Johnson E.M., and Wang D.L. (2019): “A deep learning algorithm to increase intelligibility for hearing-impaired listeners in the presence of a competing talker and reverberation,” *Journal of the Acoustical Society of America*, vol. 145, pp. 1378.

AR19. Zhao Y., Wang D.L., Johnson E.M., and Healy E.W. (2019): “A deep learning based segregation algorithm to increase speech intelligibility for hearing-impaired listeners in reverberant-noisy conditions,” *Journal of the Acoustical Society of America*, vol. 145, p. 1874.

AR20. Healy E.W., Delfarah M., Johnson E.M., and Wang D.L. (2019): “A deep learning algorithm to increase intelligibility for hearing-impaired listeners in the presence of a competing talker and reverberation,” *Journal of the Acoustical Society of America*, vol. 145, pp. 1378.

AR21. Zhao Y., Wang D.L., Johnson E.M., and Healy E.W. (2019): “A deep learning based segregation algorithm to increase speech intelligibility for hearing-impaired listeners in reverberant-noisy conditions,” *Journal of the Acoustical Society of America*, vol. 145, p. 1874.