Maulik C. Madhavi

Research Fellow, Department of Electrical and Computer Engineering, NUS, Singapore
Email: maulik.madhavi@nus.edu.sg, maulikmadhavi@gmail.com, (Ph): +65-85443713
Personal website: https://sites.google.com/site/maulikcmadhavi
Github: https://github.com/maulikmadhavi/

Research and Technical Experience

|	
Dec, 2017-Present	**Designation**: Research Fellow National University of Singapore (NUS) **Project activities**: Chatbot for autonomous vehicle, Text-dependent speaker verification, AI for healthcare, Form filling system. **PI**: Prof. Haizhou Li **Others**: Mentor and co-supervise the final year undergrad students for their projects.
April, 2012-June, 2014	**Designation**: Project Personnel, DA-IICT, Gandhinagar **Project activities**: Prosodically Guided Phonetic Engine Development for Gujarati and Marathi languages Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT) **PI**: Prof. Hemant Patil
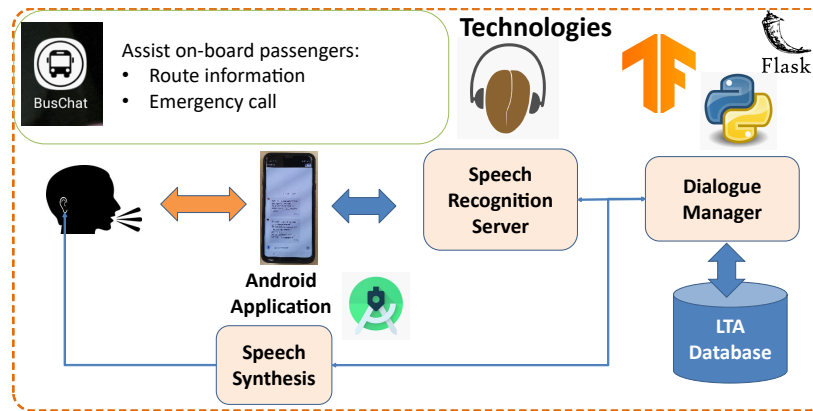
Figure 1: Chatbot for autonomous vehicle

Figure 2: Wakeup word ASR (integration to chatbot)

Education

2011-2017	**Ph.D., Information and Communication Technology (ICT)** Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT), Gandhinagar. - CPI 9.75/10
2009-2011	**M.Tech Information Communication Technology (ICT)**, Communication Systems specialization. Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT), Gandhinagar. - CPI 8.26/10
2005-2009	**B.E. Electronics and Communication (EC)** V. V. P. Engg. College, Rajkot (Saurashtra University) Percentage (69.73%)(aggregate)
2004-2005	**H.S.C** Gujarat Secondary and Higher Secondary Education Board - Percentage 85.80 % (aggregate), 70.77 %
2002-2003	**S.S.C.** Gujarat Secondary and Higher Secondary Education Board - Percentage 82.14%

Skills

Research area:	Speech signal processing, voice search, speaker verification, spoken language understanding, deep learning for spoken language processing.
Tools:	Bash, C/C++, Android (basic), Python, MATLAB, Git, Docker, kaldi, Deep learning frameworks (Tensorflow, PyTorch)
OS:	Windows and Linux
Google Scholar:	Citation: 157, h-index: **7**, i10-index : 5
Research Gate (RG) score:	**10.25**

Research Publications

International Journals

1. **M. C. Madhavi**, and H. A. Patil, "Vocal Tract Length Normalization using a Gaussian mixture model framework for query-by-example spoken term detection," in *Computer, Speech and Language, Elsevier*, vol. 58, pp. 175-202, November 2019.

2. **M. C. Madhavi**, and H. A. Patil, "Design of mixture of GMMs for query-by-example spoken term detection," in *Computer, Speech and Language, Elsevier*, vol. 52, pp. 41-55, November 2018.

3. H. A. Patil, and **M. C. Madhavi**, "Combining evidences from magnitude and phase information using VTEO for person recognition using humming," on *special issue of recent advances in speaker and language recognition and characterization Computer Speech and Language, Elsevier*, vol. 52, pp. 225-256, November 2018.

4. **M. C. Madhavi**, and H. A. Patil, "Partial matching and search space reduction for QbE-STD," in *Computer Speech & Language, Elsevier*, vol. 45, pp. 58-82, September 2017.

5. H. A. Patil, **M. C. Madhavi**, K. K. Parhi, "Static and dynamic information derived from source and system features for person recognition from humming," *Int. Journal of Speech Technology, Springer* , vol. 15, no. 3, pp. 393-406, 2012.

Book Chapters

6. **M. C. Madhavi**, H. A. Patil and N. Bhendawade, "Spoken Keyword Retrieval using Source and System Features," Pattern Recognition and Machine Intelligence. PReMI 2017, B. Uma Shankar et. al. (Eds), Lecture Notes in Computer Science (LNCS), pp. 333-341, 2017.

7. **M. C. Madhavi**, S. Sharma, H. A. Patil, "VTLN Using different warping functions for template matching," *Machine Intelligence and Big Data in Industry*, Springer International Publishing, D. Ryżko, P. Gawrysiak, M. Kryszkiewicz, and H. Rybiński, Henryk (Eds.), pp. 111-121, 2016.

8. **M. C. Madhavi**, S. Sharma, and H. A. Patil, "Vocal tract length normalization features for audio search," in *Int. Conf. Text, Speech, and Dialogue, TSD*, Pavel Král and Václav Matousek (Eds.), Pilsen,Czech Republic, pp. 387-395, 2015.

9. Y. Gaur, **M. C. Madhavi**, and H. A. Patil, "Speaker recognition using sparse representation via superimposed features," in P. Maji et. al. (Eds.) PReMI, Lecture Notes in Computer Science (LNCS), vol 8251, pp.140-147, Springer-Verlag, Berlin Heidelberg, Germany, 2013

10. H. A. Patil, **M. C. Madhavi**, R. Jain, and A. Jain, "Combining evidences from temporal and spectral features for person recognition from humming," in Malay K. Kundu et al. (Eds.) PerMIn, Lecture Notes in Computer Science (LNCS), vol. 7143, pp. 321-328, Springer-Verlag, 2012.

International Conferences

11. Y. Ong, **M. Madhavi**, K. Chan, "OpenNLU: Open-Source Web-Interface NLU Toolkit for Development of Conversational Agent," (accepted) in Proc. Asia-Pacific Signal and Information Processing Association (APSIPA) Annual Summit and Conference (ASC) 2020, Auckland, New Zealand, December 2020.

12. N. Shah, Sreeraj R, **M. Madhavi**, N. Shah, and H. Patil, "Query-by-Example Spoken Term Detection using Generative Adversarial Network," in Proc. Asia-Pacific Signal and Information Processing Association (APSIPA) Annual Summit and Conference (ASC) 2020, Auckland, New Zealand, December 2020.

13. W. Lin, **M. Madhavi**, R. Das and H. Li, "Transformer-based Arabic Dialect Identification," (accepted) in Proc. International Conference on Asian Language Processing (IALP), Kuala Lumpur, Malaysia, December 2020.

14. T. Liu, R. Das, **M. Madhavi**, S. Shen and H. Li, "Speaker-Utterance Dual Attention for Speaker and Utterance Verification," in Interspeech, Shanghai, China, October 2020, pp. 4293-4297.

15. R. Sheelvant, B. Sharma, **M. Madhavi**, R. Das, S.R.M. Prasanna and H. Li, "RSL2019: A Realistic Speech Localization Corpus," in Proc. Oriental COCOSDA 2019, Cebu City, Philippines, October 2019, pp. 1-6.

16. T. Liu, **M. Madhavi**, R. Das and H. Li, "A Unified Framework for Speaker and Utterance Verification," in Interspeech, Graz, Austria, September 2019, pp. 4320-4324.

17. **M. Madhavi**, T. Zhan, H. Li and M. Yuan, "First Leap Towards Development of Dialogue System for Autonomous Bus," in Int. Workshop on Spoken Dialogue Systems Technology (IWSDS), Sicily, Italy, April 2019, pp. 1-6.

18. R. Das, **M. C. Madhavi**, and H. Li, "Compensating Utterance Information in Fixed Phrase Speaker Verification," in Proc. Asia-Pacific Signal and Information Processing Association (APSIPA) Annual Summit and Conference (ASC) 2018, Honolulu, Hawaii, USA, November 2018, pp. 1708-1712.

19. P. A. Tapkir, M. R. Kamble, H. A. Patil, and **M. C. Madhavi**, "Replay Spoof Detection using Power Function Based Features," in Proc. Asia-Pacific Signal and Information Processing Association (APSIPA) Annual Summit and Conference (ASC) 2018, Honolulu, Hawaii, USA, November 2018, pp. 1019-1023.

20. N. Shah, **M. C. Madhavi**, and H. Patil,"Unsupervised Vocal Tract Length Warped Posterior Features for Non-Parallel Voice Conversion," *Proc. Interspeech*, Hyderbad, India, Sep. 2-6, 2018, pp. 1968-1972.

21. **M. C. Madhavi**, and H. A. Patil, "VTLN-Warped Gaussian posteriorgram for QbE-STD," in 25th *European Signal Process. Conf., EUSIPCO*, Kos island, Greece, Aug. 28-Sep. 2, 2017, pp. 593-597.

22. **M. C. Madhavi**, and H. A. Patil, "Two Stage Zero-resource Approaches for QbE-STD,"in Ninth International Conference on Advances in Pattern Recognition, *ICAPR 2017*, Bangalore, India, December 28-30, 2017.

23. **M. C. Madhavi**, and H. A. Patil, "Combining Evidences from Detection Sources for Query-by-Example Spoken Term Detection," accepted in Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, *APSIPA 2017*, Kuala Lumpur, Malaysia, December 12-15, 2017.

24. A. Rajpal, T. B. Patel, H. B. Sailor, **M. C. Madhavi**, Hemant A. Patil, Hiroya Fujisaki, "Native language identification using spectral and source-based features," in 17th Proc. Annual Conf. of Int. Speech Communication Association (ISCA), INTERSPEECH, San Francisco, USA, 8-12 Sept., 2016, pp. 369-372, 2383-2387.

25. **M. C. Madhavi**, and H. A. Patil, "Modification in sequential dynamic time warping for fast computation of query-by-example spoken term detection task," in *Int. Conf. on Signal Processing and Communications (SPCOM)* , IISc Bangalore, India, June 12-15, 2016, pp. 1-6.

26. **M. C. Madhavi**, H. A. Patil, and B. B. Vachhani, "Spectral transition measure for detection of obstruents," in 23rd *European Signal Process. Conf., EUSIPCO*, Nice, France, Aug 31 - Sept. 4, 2015, pp. 330-334.

27. H. B. Sailor, **M. C. Madhavi**, and H. A. Patil, "Significance of phase-based features for person recognition using humming," in 2nd Int. Conf. on Perception and Machine Intelligence (PerMin), C-DAC, Kolkata, Feb. 26-27, 2015, pp. 99-103.

28. B. Vachhani, K. D. Malde, **M. C. Madhavi** and H. A. Patil, "A spectral transition measure based MELCEPSTRAL features for obstruent detection," in *Int. Conf. on Asian Lang. Process. (IALP '14)*, Kuching, Malaysia, 2014, pp. 50-53.

29. S. Sharma, **M. C. Madhavi** and H. A. Patil, "Vocal Tract Length Normalization for Vowel Recognition in Low Resource Languages," in *Int. Conf. on Asian Lang. Process. (IALP '14)*, Kuching, Malaysia, 2014, pp. 54-57.

30. **M. C. Madhavi**, S. Sharma, and H. A. Patil, "Development of language resources for speech application in Gujarati and Marathi," in *Int. Conf. on Asian Lang. Process., (IALP)*, Kuching, Malaysia, 2014, pp. 115-118.

31. A. Undhad, H. Patil, and **M. C. Madhavi**, "Exploiting speech source information for vowel landmark detection for low resource language," in the 9th *Int. Symp. on Chinese Spoken Language Processing, ISCSLP'14*, Singapore, Sep. 12-14, 2014, pp. 546-550.

32. N. Shah, H. Patil, **M. Madhavi**, H. Sailor and T. Patel, "Deterministic annealing EM algorithm for developing TTS System in Gujarati," in the 9th *Int. Symp. on Chinese Spoken Language Processing, ISCSLP'14*, Singapore, Sep. 12-14, 2014, pp. 526-530.

33. **M. Madhavi**, and H. A. Patil, "Exploiting variable length Teager energy operator in melcepstral features for person recognition from humming," in the 9th *Int. Symposium on Chinese Spoken Language Processing, ISCSLP'14*, Singapore, Sep. 12-14, 2014, pp. 624-628.

34. S. Sharma, **M. C. Madhavi** and H. A. Patil, "Development of vocal tract length normalized phonetic engine for Gujarati and Marathi languages," in *The 17th Oriental COCOSDA'14*, Phuket, Thailand, Sept. 10-12, 2014.

35. K. D. Malde, B. B. Vachhani, **M. C. Madhavi**, N. H. Chhayani, and H. A. Patil. "Development of speech corpora in Gujarati and Marathi for phonetic transcription," in *Int. Conf. Oriental COCOSDA held jointly with 2013 Conf. on Asian Spoken Lang. Research and Evaluation (O-COCOSDA/CASLRE)*, 2013, Gurgaon, India, pp. 1-6. 2013.

36. H. A. Patil, **M. C. Madhavi**, K. D. Malde, and B. B. Vachhani,"Phonetic transcription of Fricatives and Plosives for Gujarati and Marathi Languages, " Int. Conf. on Asian Lang. Process. (IALP), Hanoi, Vietnam, November 13-15, 2012, pp. 177-180.

37. H. A. Patil, **M. C. Madhavi**, and N. H. Chhayani," Person Recognition Using Humming, Singing and Speech, " Int. Conf. on Asian Lang. Process. (IALP), Hanoi, Vietnam, November 13-15, 2012, pp. 149-152.

38. H. A. Patil, and **M. C. Madhavi**, "Significance of magnitude and phase information via VTEO for humming based biometrics," in *Proc. Int. Conf. on Biometrics (ICB)*, New Delhi, India, 2012, pp. 372-377.

39. H. A. Patil, **M. C. Madhavi**, and K. K. Parhi,"Combining evidence from spectral and source-Like features for person recognition from humming," in 12th Proc. Annual Conf. of Int. Speech Communication Association (ISCA), INTERSPEECH, Florence, Italy, August 27-31, 2011, pp. 369-372.

Thesis and Project Details

Ph.D. Thesis	**"Design of QbE-STD System: Audio Representation and Matching Perspective"**

Query-by-Example approach for speech retrieval has been an interesting problem among speech research community. The problem of retrieving and detecting the location of the query within speech databases using spoken example is known as "Query-by-Example Spoken Term Detection (QbE-STD)". The major outline of thesis are as follows:

- We present the representation and matching perspective in QbE-STD problem.
- Representation: We proposed vocal tract length normalization (VTLN) warping factor estimation in GMM framework in order to design posteriorgram.
- Representation: We proposed to the design of posteriorgram with the mixture of GMMs.
- Matching: We proposed different partial matching schemes to execute a non-exact audio matching task.
- Matching: We proposed feature reduction approach with phone boundaries into consideration and a segment-level bag of acoustic words approach.
- Representation & matching: The use of different detection cues such as term frequency vector, self-similarity matrix, depth of detection valley along warping path for matching. In addition, we explored pseudo-relevance feedback and weighted mean features as additional representations.

Supervisor: Prof. Hemant A. Patil

M.Tech Thesis	**"Person recognition from their hum"**

The objective of the thesis is to use hum of a person for recognition. The major outline of the thesis are as follows:

- In-house humming data collection from around 97 subjects on a voluntarily basis for the old Hindi songs.
- A novel feature-based approach for person recognition using humming is presented.

Supervisor: Prof. Hemant A. Patil

B.E. Project	**"RADAR Based Navigation"**

MATLAB simulation of cross-correlation to detect the enemy target was the part of the project. In addition, a hardware module, which detects the object's distance from that module, was also designed. Team Size: 4

Supervisor: Dr. Charmy Patel, V.V.P. Engg. College, Rajkot

Awards and Achievements

1. Received IAPR (International Association for Pattern Recognition) Travel Scholarship for presenting our joint paper in International Conference on Biometrics, ICB'12, Delhi, India. The grant has been approved by Dr. Salil Prabhakar (IEEE Fellow). Title of the paper is "Significance of Magnitude and Phase Information via VTEO for Humming Based Biometrics". (Grant reference:IAPR009)

2. Graduate Aptitude Test in Engineering (GATE)-2009 exam qualified in Electronics & Communication Engineering with All India Rank- **2142**

Invited Talks/ Presentations

1. Doctoral consortium talks on "Design of QbE-STD System: Audio Representation and Matching Perspective "at Summer School on Speech Signal Processing, **DA-IICT, Gandhinagar** during 8-12 July 2017.

2. Talk on, "An overview of Query-by-Example Spoken Term Detection (QbE-STD)", in ISCA supported Summer School on Advances in Speech and Audio Processing (ASAP) from representations and models to Applications, at **DA-IICT, Gandhinagar** during 19-20 July 2016.

3. *5* min Ph.D. competition, "Query-by-Example Spoken Term Detection - A template matching approach" at ISCA supported summer School on Speech Signal Processing (S4P) on "Speech Source Modelling and its Applications", at **DA-IICT, Gandhinagar** during 4-8 July 2016.

Conference and Workshop Attended

International Conferences

- **INTERSPEECH** 2019 at **Graz, Austria** during 14-18 September 2019. (presented 1 poster)

- **INTERSPEECH** 2018 at **Hyderabad, India** during 2-6 September 2018. (presented 1 poster with DA-IICT colleague)

- IEEE International Conference on Asian Language Processing (**IALP**), 2014 at **Kuching, Sarawak, Malaysia** during 20-22 Oct. 2014. (Presented 3 papers as well as 2 papers of speech research lab, DA-IICT colleagues).

- International Conference **Oriental COCOSDA** (The International Committee for the Co-Ordination and standardization of Speech Databases and Assessment Techniques), at KIIT college of engineering, **Gurgaon**, 25-27 Nov. 2013. (Presented 1 poster)

- The 5th IAPR International Conference on Biometrics (**ICB**) at **New Delhi**, India during March 29 - April 1, 2012. (Presented 1 poster)

- The Indo-Japan conference on Perception and Machine Intelligence (**PerMIn'12**), Centre for Development of Advanced Computing (C-DAC), **Kolkata**, India during 11-12 January 2012. (Presented 1 paper)

Workshops

- International Workshop on Spoken Dialogue Systems Technology (**IWSDS**), at **Sicily, Italy**, during 24-26 April 2019. (Presented 1 poster)

- International Workshop on Spoken Dialog System Technology 2018, **Singapore**, May 14-16, 2018

- Summer School on Advances in Speech and Audio Processing (ASAP) from representations and models to Applications, at **DA-IICT, Gandhinagar** during 19-20 July 2016.

- Summer School on Speech Signal Processing (S4P) on "Speech Source Modelling and its Applications", at **DA-IICT, Gandhinagar** during 4-8 July 2016.

- Winter School on Speech and Audio Processing (WiSSAP) on "Production-Perception Based New Models of Speech Analysis", at **DA-IICT, Gandhinagar** during 4-7 January 2015.

- CEP workshop on, "Text-to-Speech (TTS) Synthesis", at **DA-IICT, Gandhinagar** during 16-18 June 2014.

- Deity progress review meeting and Workshop at **IIT-Kharagpur** during 8-9 March 2014.

- Winter School on Speech and Audio Processing (WISSAP) on "Deep Learning for Multi-Lingual Speech Processing" at **IIIT, Hyderabad**, 17-20 January 2014.

- One-day CEP workshop on "Speech Signal Processing and its application" at **DA-IICT, Gandhinagar**, 14 October 2013.

- Deity progress review meeting and Workshop on "Development of prosodically guided phonetic engine for searching speech databases in Indian languages" held at **Thapar University, Patiala** on 9-10 March, 2013.

- Workshop on "HMM-Based Speech Synthesis" at **SSN College of Engg., Kalavakkam, Chennai** along with IIT-Madras during 26-30 November 2012.

- Deity progress review meeting and and workshop on "Prosodic Labelling" at Speech and Vision Laboratory, **IIIT Hyderabad** during 25-28 October 2012.

- Deity progress review meeting and and workshop on "Phonetic Transcription" at Speech and Vision Laboratory, **IIIT Hyderabad** during 11-13 May 2012.

- Deity progress review meeting and workshop on "Phonetic Transcription" at Speech and Vision Laboratory, **IIIT Hyderabad** during 18-20 February 2012.

- "DIT Workshop on Speaker Verification: Development and Applications" at IIT Guwahati during February 13-17, 2012.

Other Activities

Organizing Committee:	Publication chair, Web content Manager, IEEE Automatic Speech Recognition and Understanding Workshop, 14-18 Dec. 2019, Singapore.
Reviewer:	INTERSPEECH 2018, INTERSPEECH 2019, Interspeech 2020, APSIPA 2018, IALP 2019, IALP 2020, Computer Speech Language, Speech Communication, Elsevier, IEEE Access.
Volunteer:	International Workshop on Spoken Dialog System Technology 2018 at Singapore, 14-16 May 2018.
Volunteer:	Summer School on Speech Signal Processing (S4P) on "Speech Source Modelling and its Applications", at DA-IICT, Gandhinagar 4-8 July 2016.
Volunteer:	Winter School on Speech and Audio Processing (WiSSAP) on "Production-Perception Based New Models of Speech Analysis", at DA-IICT, Gandhinagar during 4-7 January 2015.
Volunteer:	CEP workshop on, "Text-to-Speech (TTS) Synthesis" at DA-IICT, Gandhinagar 16-18 June 2014.
Volunteer:	One-day CEP workshop on "Speech Signal Processing and its application" at DA-IICT, Gandhinagar, 14 October 14 2013.

Personal Information

DOB	March 4, 1988
Current Address	Blk 59, Commonwealth Drive, #09-139, Singapore, 140059.
Other Address	"AUM", 1-Wankaner Soc., Jamanagar Road, Rajkot, Gujarat, India, 360006.
URL	https://sites.google.com/site/maulikcmadhavi/
Github	https://github.com/maulikmadhavi

Declaration

The above information is correct to the best of my knowledge.

Maulik C. Madhavi

December 1, 2020