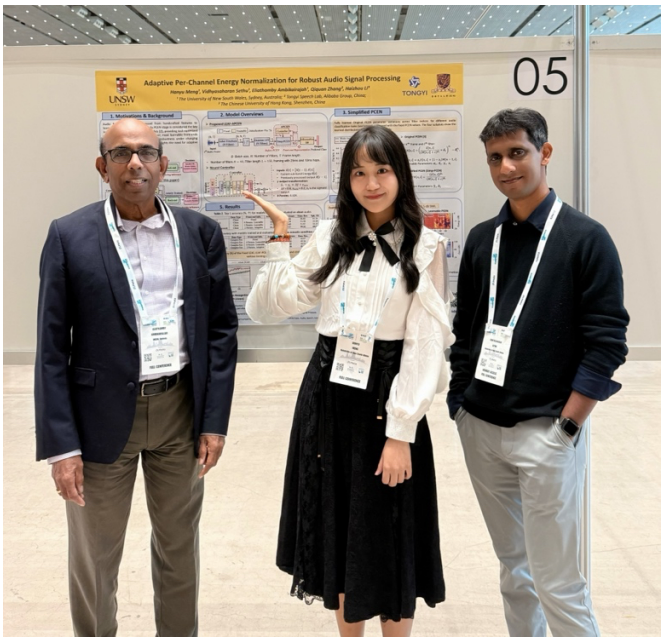
ASSTA Travel Grant Report

Hanyu Meng, PhD Candidate, University of New South Wales, Sydney, Australia

I had the great honour of attending the 2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2026), held in Barcelona, Spain, from 4 to 8 May 2026. ICASSP is one of the most prestigious international conferences in the broad field of signal processing, bringing together researchers from speech, audio, music, image, video, machine learning, and multimedia signal processing communities. I am sincerely grateful to the Australasian Speech Science and Technology Association (ASSTA) for supporting my travel through the Conference Travel Award, which enabled me to present my research, receive valuable feedback, and engage with the international research community.

At ICASSP 2026, I presented my first-authored paper titled "**Adaptive Per-Channel Energy Normalization Front-End for Robust Audio Signal Processing**" in the Audio and Acoustic Signal Processing poster session from 2:00 to 4:00 pm on 5 May 2026. This work proposes LEAF-APCEN, an adaptive audio front-end designed to improve the robustness of audio representations under complex and dynamic acoustic conditions. Conventional learnable front-ends can optimise task-specific representations during training, but their parameters remain fixed during inference, limiting their ability to adapt to changing acoustic environments. In contrast, our proposed framework introduces a lightweight neural feedback controller that dynamically adjusts the parameters of a simplified Per-Channel Energy Normalization (PCEN) module during inference. This enables input-dependent and time-varying dynamic range compression, inspired by adaptive mechanisms in the human auditory system.



Together with my supervisors (left: Prof Eliathamby Ambikairajah and right: A/Prof Vidhyasaharan Sethu) in my poster session.

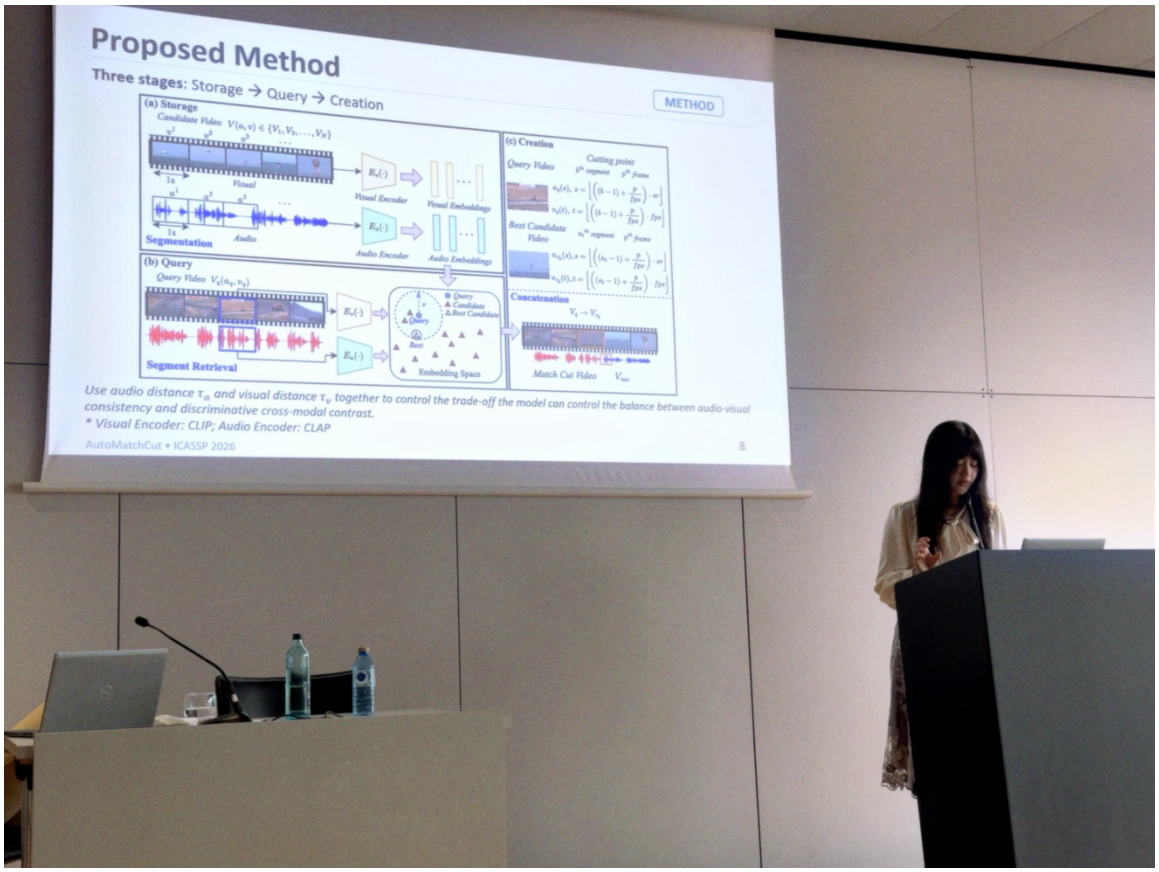
A key contribution of this work is the simplification of the original PCEN formulation. Instead of using four learnable parameters, we reformulate PCEN into a two-parameter version, named **SimpPCEN**, while maintaining strong performance. Building on this simplified structure, LEAF-APCEN uses a neural adaptive controller to regulate the two SimpPCEN parameters dynamically. The controller takes both current and previously processed subband energies as input, allowing the model to capture temporal energy variation and perform adaptive gain control during inference.

During the poster session, I had the opportunity to discuss this work with researchers working on audio representation learning, robust speech processing, and biologically inspired signal processing. Many discussions focused on the motivation for making audio front-ends adaptive, the relationship between PCEN and auditory gain control, and the potential extension of the proposed framework to multi-channel and spatial audio tasks. These conversations were particularly valuable for my PhD research, which aims to bridge human auditory mechanisms and machine hearing systems. The feedback I received helped me think more clearly about how adaptive front-end design can be positioned not only as an engineering technique, but also as a biologically inspired modelling approach for robust audio perception.

Photo taken together with my supervisors in front of the photo board of ICASSP2026.

The conference programme was broad and highly inspiring. I attended a range of oral and poster sessions related to sound event detection, speech enhancement, spatial audio production, music information retrieval, self-supervised learning, and audio-visual signal processing. These sessions gave me a clearer understanding of current trends in the field, especially the increasing interest in robust and generalisable audio representations. I was particularly inspired by works that explored how signal processing principles, machine learning models, and perceptual knowledge can be combined to build systems that are more reliable in real-world acoustic environments. This strongly aligns with the long-term goal of my PhD: to design interpretable and adaptive audio front-ends that can improve machine listening under challenging acoustic conditions.

In addition to my poster presentation, I also gave an oral presentation for another my ICASSP 2026 paper, "**Auto-MatchCut: An Audio-Visual Retrieval Framework for Seamless Match Cutting**." Presenting in both poster and oral formats was a very rewarding experience. The poster session allowed for detailed technical discussions, while the oral presentation helped me practise communicating research ideas clearly and concisely to a wider audience. Together, these two presentations greatly improved my confidence in presenting my work internationally and engaging with researchers from different signal processing communities.



My oral presentation took place at 10:00 am on 7 May 2026.

Attending ICASSP 2026 also provided valuable networking opportunities. I met PhD students, postdoctoral researchers, and senior academics from different countries and institutions. These conversations helped me broaden my academic perspective and introduced me to new research directions that may inform my future work. In particular, discussions around adaptive audio representation, spatial hearing, and robust machine listening encouraged me to think about extending my current monaural adaptive front-end framework towards multi-channel and binaural audio processing. This is closely related to my broader PhD direction on biologically inspired binaural modelling and selective hearing.

Overall, ICASSP 2026 was an extremely valuable experience for my academic development. It allowed me to present my work to an international audience, receive constructive feedback, learn about the latest advances in signal processing, and build connections with researchers in related fields. The experience also strengthened my motivation to continue developing adaptive and interpretable audio front-ends for robust machine hearing. I would like to express my sincere gratitude to the ASSTA for providing me the travel grant award, which made it possible for me to attend ICASSP 2026, share my research, and benefit from this important international conference.

Below are some other photos in Barcelona, Spain:

Tibidabo, where known as the "Magic Mountain," offering panoramic views of the city and Mediterranean Sea.

Sagrada Família (Outside)

Sagrada Família (inside).

Sagrada Família (inside)

Meet and connect with other researchers in my research field in ICASSP:


ICASSP2026
2026 IEEE International Conference on Acoustics,
IEEE
Signal
Processing
Society
No passar
No pasar
Do not pass
ICASSP26