PhD Studentship in Neural Architectures for Multimodal Speech Recognition

Position: PhD Studentship
Funding: EU/Non-EU Fees plus tax-free stipend of €22,000 per annum for 4 years
School/Location: EE Engineering, School of Engineering, Trinity College Dublin, Ireland
Supervisor: Prof Naomi Harte
Start Date: September 1st 2024

Background
It is known that many modalities (e.g. articulation, mouth movements, eye gaze, head nods, back channels and gestures) play a role in communication success in speech-based interaction. Lombard speech is one way we modify the sound of our voices to ensure we communicate clearly with others in noisy environments, but recent research has shown that Lombard speech in reality is multimodal. This PhD student will focus on approaches to speech recognition that incorporate knowledge from the visual modality of speech to support decoding in noisy conditions. By gaining deeper insights into what visual information is available, the research will look at approaches to multimodal optimisation in training that can yield superior multimodal neural architectures for speech recognition.

This PhD research is part of a larger project on multimodal speech in Prof. Harte’s lab. The funding available will cover fees and €22,000 per annum stipend, but also equipment and conference travel associated with the research.

Person Requirements
• Primary or Master’s degree in Electronic Engineering or Computer Science, with an interest in engaging in multidisciplinary research
• Must meet TCD University requirements at https://www.tcd.ie/study/apply/admission-requirements/postgraduate/
• Strong skills in coding, machine learning, deep learning, and signal processing with the willingness and motivation to learn new skills and packages
• Prior experience with speech-based interaction desirable
• Excellent communication skills, both spoken and written, and fluency in English
Application Process:

Interested candidates can submit their application via the following link: 
https://forms.gle/qSXuEFZn3vdfszeQ7

Applicants should provide the following information when applying:

1. A motivation statement outlining their interest and suitability for the position (approx. 300-600 words)
2. A comprehensive curriculum vitae
3. The names and contact details (e-mail) of two academic references

Note: candidates who do not address the application requirements above will not be considered for interview.

Further Information

Informal enquiries about this post should be made to Professor Naomi Harte (nharte@tcd.ie) but applications are only accepted through the procedure outlined above.

About the School of Engineering

The School of Engineering is ranked #1 in Ireland (QS Rankings) and is a proud recipient of a Bronze Athena Swan award, attained in 2021. As part of the School’s on-going actions in relation to equality, diversity and inclusion it welcomes all applications that meet the criteria below and particularly those from under-represented groups. The School offers a collegiate and supportive environment to all its staff and works to ensure that all its staff and students can perform at their best while putting in those steps that facilitate a healthy work/life balance.

About Sigmedia Research Group

The Sigmedia Research Group is part of the Discipline of Electronic and Electrical Engineering (https://www.tcd.ie/eleceng/). The group was founded in 1998 in Trinity College Dublin. Originally with a focus on video and image processing, the group today spans research in areas across all aspects of media – video, images, speech and audio. The group is widely recognised for its work in audio and video engineering, and is one of the few groups in the world to have had its technology acquired by Google in the areas of video engineering and spatial audio. Prof. Naomi Harte leads the Sigmedia research endeavours in human speech communication. Her team has active research in audio-visual speech recognition, evaluation of speech synthesis, multimodal cues in human conversation, and birdsong analysis. Her group is interested in all aspect of human interaction,
centred on speech. Much of the research is underpinned by signal processing and machine learning, but we also have researchers grounded in linguistic and psychology aspects of speech processing to give us the complete view of human interaction in all its richness.

About the ADAPT Centre

ADAPT is the world-leading SFI research centre for AI Driven Digital Content Technology, coordinated by Trinity College Dublin and based within Dublin City University, University College Dublin, Technological University Dublin, Maynooth University, Munster Technological University, Technological University of the Shannon, and the University of Galway. ADAPT’s research vision is to pioneer new forms of proactive, scalable, and integrated AI-driven Digital Content Technology that empower individuals and society to engage in digital experiences with control, inclusion, and accountability with the long-term goal of a balanced digital society by 2030. ADAPT is pioneering new Human Centric AI techniques and technologies including personalisation, natural language processing, data analytics, intelligent machine translation human-computer interaction, as well as setting the standards for data governance, privacy and ethics for digital content.

Equal Opportunities Policy

Trinity is an equal opportunities employer and is committed to employment policies, procedures and practices which do not discriminate on grounds such as gender, civil status, family status, age, disability, race, religious belief, sexual orientation or membership of the travelling community. On that basis we encourage and welcome talented people from all backgrounds to join our staff community. Trinity’s Diversity Statement can be viewed in full at https://www.tcd.ie/diversity-inclusion/diversity-statement.