陳致生 (Chih-Sheng Chen), Ph.D.

經理 | AI 語音與音訊處理專家

台灣新竹 • LinkedIn • 技術部落格

簡介

語音 AI 工程師與研發經理，14+ 年產業經驗 (自 2012 年起)——從聯發科的定點化語音喚醒，到在意騰科技 (intelliGo) 帶領 AI 語音演算法團隊。我與 5–6 位碩博士工程師組成的團隊，設計、壓縮並部署神經網路語音模型 (降噪、回音消除/抑制、波束成形、頻寬擴展) 到自研 AI 晶片，在嚴格的功耗與延遲限制下運作。成果落地於 Nintendo Switch 2 GameChat、通過 Microsoft Teams (Open Office) 認證的無吊桿耳機，以及數以百萬計的其他裝置。

除了工作上的語音領域之外，有興趣的研究包含: 模型壓縮、LLM 推論加速、強化學習、和生成模型。 同時持續經營技術部落格 (100+ 篇)，主題涵蓋量化、剪枝與生成式模型 (Diffusion、Flow Matching)，以及背後的隨機過程、貝氏方法與最優傳輸等數學。

工作經歷

意騰科技 intelliGo Technology Inc. — **新竹**
2019-08 – 至今

經理 (Manager) (2023-01 – 至今)

- 帶領演算法團隊 (5–6 位碩博士)，開發 NN-based 語音前端與音訊演算法：降噪、AEC/AES (聲學回音抑制)、shotgun/bubble 波束成形 (2–5 顆麥克風)、DoA/DoV/DoK、語者日誌與頻寬擴展 (BWE)。
- Nintendo Switch 2 GameChat：帶領團隊開發語音前端模型 (降噪、AEC、AES、遠場拾音)，並部署於自研晶片，在低功耗下達到高音質。
- 無吊桿耳機：帶領團隊開發「bubble mic」AI 波束成形——波束成形僅使用 3 顆麥克風——通過 Microsoft Teams (Open Office) 認證，收音品質媲美有吊桿麥克風。
- 筆電：團隊開發「shotgun」3 麥克風波束成形，以第 3 顆麥克風解決螢幕 2 麥克風陣列的前後鏡像收音問題。
- 帶領開發以 GAN 為基礎的頻寬擴展 (BWE)，並於手機產品量產。
- 以量化與剪枝將 OpenAI Whisper 改為串流推論並加速。
- 技術量產於數以百萬計的裝置，涵蓋筆電、TWS 耳機、會議喇叭與遊戲主機。
- 與 FW 團隊緊密合作把演算法部署上晶片；規劃技術 roadmap；招募人才；把 Sales/Marketing 的客戶回饋轉為技術目標；定期舉辦 AI 論文與工程 review。

副理 (技術) Deputy Technical Manager (2021-01 – 2022-12)

- 帶領開發 NN-based 降噪、AEC、波束成形與 DoA；與 FW 團隊合作晶片部署；規劃 roadmap；招募人才。

資深工程師 (Staff Engineer) (2019-08 – 2020-12)

- 打造極低 MIPS/記憶體的 ASR 與指令辨識系統：聲學模型訓練與量化、支援使用者自訂指令文法的解碼圖，以及部署於公司 AI 晶片的定點化 decoder。同時追蹤最新 ASR 研究與語音 UI roadmap。

聯想研究院 Lenovo Research (AI Lab) — **中國北京**
2017-08 – 2019-07

研究員 (Researcher)

- 以深度學習開發關鍵字喚醒 (trigger) 並結合聲紋驗證。
- 實作聲學回音消除與波束成形，提升 trigger 與雲端 ASR 在雜訊下的穩健度。
- 此期間為三件中國專利的共同發明人 (見專利)。

聯發科技 MediaTek — **新竹**
2013-09 – 2017-07

資深研發工程師 (Senior R&D Engineer)

- 在從零打造極低 MIPS、低記憶體語音喚醒系統中扮演關鍵角色，已應用於多款聯發科晶片 (含 Helio X20/X25)，並獲魅族 Pro 6 手機採用。
- 設計並實作語音喚醒演算法 (訓練與解碼)，含定點化最佳化，並整合 CM4 與 HAL。
- 以 Kaldi 訓練聲學模型，並建置 HMM/DNN 工具鏈。
- 參與多項 ML 專案：超音波手勢辨識、以動態規劃做發音解碼、音訊事件偵測 (如嬰兒哭聲偵測)。
- 以 C++、bash/perl 自動化並優化效能測試流程；應用 ML 與降維方法進行問題分類與視覺化。
- 低延遲關鍵字偵測美國專利的共同發明人 (見專利)。

華晶科技 Altek Corporation — 新竹
2012-09 – 2013-08

研發工程師 (R&D Engineer)

- 開發低 MIPS 的定點化 LPC 特徵擷取 (Levinson–Durbin 遞迴) 與定點化單通道降噪方法。

技能

- **音訊與語音 AI：** 降噪 · AEC / AES · 波束成形 (NN-based 2–5 顆麥克風；DS/LCMV/MVDR/CGMM-MVDR) · 語音/關鍵字方向偵測 (DoV/DoK) · 語者日誌與驗證 · 頻寬擴展 · ASR (Kaldi、WeNet、Whisper) · 關鍵字偵測
- **模型壓縮與部署：** PTQ (AIMET、OBQ、OCTAV、GPTQ、AWQ、SmoothQuant) · QAT (LSQ+、EWGS、MAD、OCTAV) · 剪枝 (Movement、L0 regularization、LoRAPrune) · 串流化 (計算圖轉換與 OP fusion) · NAS
- **訓練與生成式模型：** SWA 與 SAM 提升訓練穩定性 · GAN · Transformer · Diffusion 與 Flow Matching
- **工具：** Python · PyTorch · PyTorch Lightning + Hydra · C/C++ (基礎) · Kaldi · git
- **管理：** 管理 5–6 位碩博士工程師 · 跨團隊專案 (8–9 人) · 技術 roadmap · 招募 · 與 Sales/Marketing 及 FW 團隊協作 · 從 0→1 概念到 1→100 量產

學歷

- **博士，資訊工程**, 國立清華大學 (NTHU) (2005–2011) — 多媒體資訊檢索實驗室；指導教授：張智星 (Jyh-Shing Roger Jang)
- **訪問學者，電機與計算機工程學院 (CSIP)**, 喬治亞理工學院 (Georgia Tech) (2010–2011) — 指導教授：Chin-Hui Lee
- **學士與碩士，資訊科學**, 國立交通大學 (NCTU) (M.S. 2003–2005) — 碩士就讀自動資訊處理實驗室

論文

1. Sung-Feng Huang, Chia-ping Chen, Zhi-Sheng Chen, Yu-Pao Tsai, Hung-yi Lee. “Personalized Lightweight Text-to-Speech: Voice Cloning with Adaptive Structured Pruning.” ICASSP, 2023.
2. Wei-Tsung Kao, Yuan-Kuei Wu, Chia-Ping Chen, Zhi-Sheng Chen, Yu-Pao Tsai, Hung-Yi Lee. “On the Efficiency of Integrating Self-supervised Learning and Meta-learning for User-defined Few-shot Keyword Spotting.” IEEE SLT, 2022.
3. Zhi-Sheng Chen, Jyh-Shing Roger Jang, Chin-Hui Lee. “A Kernel Framework for Content-Based Artist Recommendation System in Music.” IEEE Transactions on Multimedia, 2011.
4. M. Wu, Z. Chen, J. R. Jang, J. Ren, Y. Li, C. Lu. “Combining Visual and Acoustic Features for Music Genre Classification.” ICMLA, 2011.
5. Jia-Min Ren, Zhi-Sheng Chen. “On the Use of Sequential Patterns Mining as Temporal Features for Music Genre Classification.” ICASSP, 2010.
6. Zhi-Sheng Chen, Jyh-Shing Roger Jang. “On the Use of Anti-Word Models for Audio Music Annotation and Retrieval.” IEEE Transactions on Audio, Speech, and Language Processing, 2009.

專利

以下專利皆為共同發明人。日期為核發/公告日。

- US 12143782 B2 — 麥克風系統 (2024-11-12; Intelligo Technology Inc.)
- US 12219329 B2 — 適用於無吊桿耳機的波束成形方法與麥克風系統 (2025-02-04; Intelligo Technology Inc.)
- US 12231844 B2 — 麥克風系統及波束成形方法 (2025-02-18; Intelligo Technology Inc.)
- US 12256202 B2 — 立體聲增強系統及立體聲增強方法 (2025-03-18; Intelligo Technology Inc.)
- TWI831513 B — 適用於無吊桿的頭戴式耳機之麥克風系統及波束成形方法 (2024-02-01; Intelligo Technology Inc.)
- TWI835246 B — 麥克風系統及波束成形方法 (2024-03-11; Intelligo Technology Inc.)
- TWI861569 B — 麥克風系統 (2024-11-11; Intelligo Technology Inc.)
- TWI870681 B — 立體聲增強系統及立體聲增強方法 (2025-01-21; Intelligo Technology Inc.)
- CN 108986815 B — 语音控制方法、装置及电子设备 (2021-11-16; Lenovo (Beijing))
- CN 110334240 B — Information processing method and system, first device and second device (2021-10-22; Lenovo (Beijing))

- CN 109978145 B — Processing method and device (2021-09-14; Lenovo (Beijing))
- US 10032449 B2 — Keyword spotting system for achieving low-latency keyword recognition by using multiple dynamic programming tables reset at different frames of acoustic data input and related keyword spotting method (2018-07-24; MediaTek Inc.)

語言： 中文 (母語)、英文 (專業工作能力)