

人工電子耳聲音訊號處理：通往人工智慧的創新旅程

Sound Processing for Cochlear Implants: The Journey of Innovation Toward Artificial Intelligence

Enoch Hsin-Ho Huang^{1,2}, Chao-Min Wu¹, Yu Tsao²

¹National Central University, Taiwan

²Academia Sinica, Taiwan

{enoch.huang, yu.tsao}@citi.sinica.edu.tw wucm@ee.ncu.edu.tw

摘要

本文回顧並簡介人工電子耳(Cochlear Implant)創新的過程，包括歷史發展、聲音處理器架構、效果評估的方法、轉向人工智慧的趨勢，並提出未來的前景，幫助讀者鑑往知來。隨著科技的發展，電子耳的訊號處理演進，從早期在硬體上增加電極數量開始，經過前處理方法的精進與聲音編碼策略的改良，專家學者們的貢獻造福了許多電子耳使用者，而近期人工智慧的發展，也將帶來下一波的改變。在當代的科技浪潮下，電子耳的研究將朝著人工智慧繼續它創新的旅程。

Abstract

This article briefly reviews the innovation process of the cochlear implant (CI), covering the historical development, sound processor architecture, evaluation methods, and the trend towards artificial intelligence (AI). Understanding the past helps illuminate the path towards the future. With advances in technology, the innovation of CI signal processing in the early days started with hardware, including an increased number of electrodes. Through the improvement of the pre-processing approaches and sound coding strategies, contributions of many experts have benefited numerous CI listeners. Within the current wave of technology, research in CI will continue its journey of innovation toward AI.

關鍵字：人工電子耳、聲音訊號處理、人工智慧
Keywords: Cochlear Implant, Sound Signal Processing, Artificial Intelligence.

1 緒論

人工電子耳又稱為人工耳蝸 (Cochlear Implant, CI)，是近半個世紀以來革命性的聽障輔具，它能幫助無法透過助聽器改善聽力的重度聽障者重獲新聲，目前在全球已有超過一百萬名使用者(Zeng, 2022)，在台灣則超過三千位(Huang, 2022)。透過跨領域專家的合作(Clark et al., 2015)，CI 已是相當成功的神經義體輔具 (Neuroprosthetic Assistive Device)，使用者大多在安靜環境下可有不錯的語音辨識表現，超越人工視覺及人工觸覺等其他感覺植入系統。對於耳蝸中毛細胞(Hair Cell)嚴重受到破壞的重度至極重度感音神經性聽損(Severe-to-Profound Sensorineural Hearing Loss)，電子耳植入手術已是標準治療方式，不但可幫助成年人重新具備溝通能力，也可幫助因新生兒聽力篩檢而重度聽損確診的兒童及早進行聽能創建(Aural Habilitation) (Chen, 2015)。

電子耳的系統架構如圖 1，包括體外的聲音處理器(Sound Processor)，及體內的植入體(Implant)。聲音由麥克風接收，處理後採用無線電波同時傳輸訊號及電能至體內，以植入於耳蝸內的電極陣列(Electrode Array)產生電流刺激聽神經，由聽覺路徑(Auditory Pathways)傳導至大腦而產生聽覺。相較於一般的聲音訊號處理研究，CI 植入體必須由耳科醫師以手術置入使用者的頭部，再由聽力師協助做個人化的電流值設定，因此電子耳研究需要與醫學等方面的專家保持密切的合作。



圖 1. 人工電子耳系統

目前電子耳在安靜環境中的聆聽及改善使用者的生活品質的成效是有目共睹的，然而尚有不少的挑戰，可歸納為以下兩大類：

- 電子耳聆聽的挑戰：使用者在噪音環境下的語音辨識、華語等聲調語言(Tonal Languages)的聲調分辨、音樂的聆聽，以及新冠疫情中口罩的配戴及線上溝通，均較正常聽力者困難許多。
- 電子耳研究的挑戰：除了臨床實驗過程比較繁瑣，電子耳使用者在聆聽表現上有巨大的變異性(Variability) (Wilson, 2019)而非接近一致的結果，都有相當的挑戰。

本文以聲音訊號處理的角度，簡介人工電子耳的發展歷史、聲音訊號處理架構、電子耳的效果評估，以及目前人工智慧(Artificial Intelligence, AI)在此領域的現況，並提出相關的探討與未來展望。

2 電子耳的歷史發展

人工電子耳早期發展階段的重點，是植入耳蝸的電極數量，如圖 2。1960 年代的單電極/單頻道(Single Electrode/Single Channel)設計，期望只要用一顆植入在耳蝸內的電極，以速率編碼(Rate Coding)的方式改變刺激頻率，就可以達到輔助讀唇的效果。1978 年，首例多頻道(Multi-channel)裝置的手術成功。臨床實驗後，單頻道及多頻道的 CI 裝置分別在 1984 及 1985 年通過美國食品藥物管理局(FDA)許可，而台灣則是在 1993 年由當時的行政院衛生署許可多頻道 CI 系統。因為多頻道系統的電極陣列的設計核心理論是位置編碼(Place Coding)，也就是採用植入在耳蝸內不同深度的電極負責刺激不同的聽神經，以表示不同的聲音高低音頻率，所以可較單一電極傳遞更豐富的資訊，而臨床結果竟可超越原本的預期讓使用者能聽懂語音，於是後發先至而成為主流的 CI 裝置。

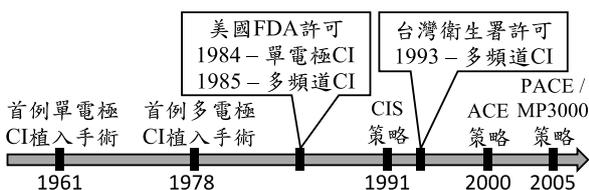


圖 2. 人工電子耳早期的發展歷史

隨著電子技術及半導體科技的發展，電子耳軟硬體上都持續演進。在硬體部分，聲音處理器的體積不斷縮小，從攜帶式(Body-Worn)、耳掛式(Behind-the-Ear, BTE)，到近年來出現的一體成形處理器(Off-the-Ear, OTE)。在 1980-2000 年的 CI 採用許多的類比元件，以致每一台處理器的聲音聽起來都不同。因此，數位訊號處理日益重要，除了能確保大量的 CI 聲音處理器產品能有一致的輸出表現，還能達到更佳的声音品質。此外，許多聲音編碼策略(Sound Coding Strategy)也不斷地推陳出新，特別是進階組合編碼 (Advanced Combination Encoder, ACE)策略 (Vandali et al., 2000)，成功地幫助多數 CI 使用者能在安靜環下理解 80%-90%的語音(Zeng et al., 2008)。

3 電子耳的聲音訊號處理

在圖 1 中，位於體外的聲音處理器是可以不斷更新升級的，只要相容於既有的無線通訊協定及體內裝置的硬體即可，因此是許多研究的重點。CI 的外部處理器早期稱為語音處理器(Speech Processor)，近年來各廠商一致地更名為聲音處理器，其企圖心不言而喻。聲音處理器的訊號處理架構說明如圖 3，可區分為前處理、聲音編碼策略、後處理三個訊號處理階段。

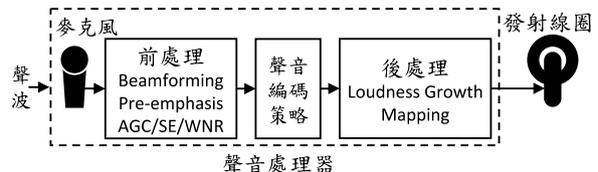


圖 3. CI 聲音處理器的訊號處理架構

CI 前處理有許多方法，包括了麥克風波束成形(Beamforming)、預強調(Pre-emphasis)、自動增益控制(Automatic Gain Control, AGC)、語音增強(Speech Enhancement, SE)、風聲消除(Wind Noise Reduction, WNR)等功能，目前這些採用傳統訊號處理的方法均已實現於產品中(Henry et al., 2023)。

聲音編碼策略又稱為聲音處理策略(Sound Processing Strategy)，是 CI 聲音訊號處理的核心(Wouters et al., 2015)，其功能是把聲音訊號翻譯成電訊號，也就是將經過前處理的聲音訊號，轉換為 M 個頻道(即 M 個電極)的電極刺激形式(Electrode Stimulation Patterns)。聲音編碼策略的主要演進可參考圖 2，連續交錯取

樣(CIS, Continuous Interleaved Sampling)策略將各個電極的發射刺激脈衝的時間交錯，以避免相鄰電極的互相干擾(Wilson et al., 1991)。ACE 策略透過頻道選擇(Channel Selection, CS)，僅由振幅最大的 N 個頻道刺激，以減少訊號較小頻道的電流干擾，又稱為 N-of-M 策略。心理聲學的 PACE 策略(Psychoacoustic ACE)利用心理聲學遮蔽原理(Psycho-acoustic Masking)，用較少的電刺激可達到省電的效果。PACE 策略的語音理解度表現與 ACE 不分軒輊，由於其技術源自 MP3 音樂壓縮標準，因此又稱為 MP3000 策略(Nogueira et al., 2005)。近年來，不少學者提出了好些以聽覺生理(Auditory Physiology)或心理聲學為基礎的策略，不過它們的表現未有明顯的突破，因此 ACE 從 2000 年提出至今仍是主流的編碼策略(Huang et al., 2021; Huang, 2023)。

後處理包括響度成長函數(Loudness Growth Function, LGF)，以及個人化的電流圖值(Map)等設定。由於電刺激的動態範圍相當窄，加上每個人的耳蝸形狀大小及電極植入位置均不同，因此需要以 LGF 函數對應為適當的對數曲線，而 Map 則需要由聽力師來依照使用者在各個電極的反應，設定最大和最小的刺激電流，即舒適電位(Comfortable Level, C-Level)和閾值電位(Threshold Level, T-Level)，以達到個人化的電刺激效果(Loizou, 1998)。

4 電子耳的效果評估

要了解訊號處理方法在人工電子耳的表現，早期僅能直接進行 CI 個案實驗，如此既費時又費力。隨著聲碼器(Vocoder)的出現，研究人員可容易地合成電子耳的模擬語音(Simulated Speech/Vocoded Speech)，因此近年來普遍的 CI 評估方式轉變為三個步驟：

- 客觀評估：採用短時客觀理解度 (Short-Term Objective Intelligibility, STOI)或正規化共變異數測量 (Normalized Covariance Metric, NCM)等方法，計算 CI 模擬語音與原始語音的相關性，可快速估測語音理解度 (Speech Intelligibility, SI) (Falk et al., 2015)。
- 主觀模擬實驗：將模擬語音由正常聽力的個案戴耳機進行聽測實驗，從他們的

回答計算正確的字詞的比例，即 SI 分數。此方法較客觀評估更接近 CI 聆聽的效果，且個案招募較 CI 使用者容易。

- 主觀電子耳使用者實驗：當訊號處理方法通過以上的評估後，再進行 CI 使用者的實驗。實驗方法除了與正常聽力個案相似的步驟，還需要特別的研究設備 (Litovsky et al., 2017)，並與醫院密切合作，以確保個案的安全與舒適。

由於以上實驗方法的成熟，CI 訊號處理研究因此蓬勃發展，並開始轉向人工智慧。

5 當 CI 遇見 AI

近年來，人工智慧開始被運用在聽力醫療相關領域(Lesica et al., 2021)，包括人工電子耳。這些可稱為「AI+CI」的研究(Huang et al., 2023)，內容除了聲音訊號的處理，還包括了透過大數據預測手術的效果、塗藥植入體(Drug-Eluting Implant)的藥物釋放模擬、以影像處理輔助手術中電極的放置、機器人手術、術後的電流圖調整等多方面的應用(Crowson et al., 2020)。在聲音處理的部分，運用人工智慧的研究主要是針對語音增強及音樂的重新混音(Remix)，此外，聲音編碼策略也開始採用 AI。

5.1 語音增強

語音增強是 CI 前處理的重要部分，目前除了已有好些傳統方法實現在產品上，在人工智慧方面已有不少實驗性的研究(Crowson et al., 2020; Henry et al., 2023; Huang et al., 2023)。以不同的神經網路架構為例，均有相關的探討：

- 深度除噪自編碼器(Deep Denoising Auto Encoder, DDAE)：採用多層感知器(Multi-Layer Perceptron, MLP)的深度神經網路(Deep Neural Network, DNN)，在正常聽力模擬及電子耳使用者的實驗中，可以改善華語語音的理解(Lai et al., 2017)。
- 維納卷積神經網路(Wiener-CNN)：結合維納濾波器的 CNN 網路，可較單純的 CNN 網路及採用頻譜刪減法(Spectral Sub-traction, SS)的 SS-CNN 網路有更加的語音增強效果(Mamun et al., 2019)。

- 循環神經網路(Recurrent Neural Network, RNN)：採用長短期記憶(Long Short-Term Memory, LSTM)的網路架構，在雜踏式噪音(Babble Noise)的干擾下，可有較佳的客觀模擬結果，CI 使用者也較偏好以此網路處理過的語音。(Goehring et al., 2019)。

語音增強除了採用一般常見的方法，還有一些適合電子耳的特別方法，例如針對尚有殘餘低聽力採用電聲混合刺激(Electro-acoustic stimulation, EAS)(Wang et al., 2021)，或是加上視覺線索的除噪方案(Tseng et al., 2021)。

5.2 重新混製音樂

人工電子耳在音樂的表現有相當大的限制，使用者普遍的反應是，對於音源的喜好與一般人不同，較偏好人聲(Vocals)。因此，由這個特性而發展出重新混製音樂的方法。

首先，AI 可用於聲源的分離，例如分開人聲及樂器的聲音，再針對這兩種聲源按照設計的比例重新混合為音樂。由於 AI 在聲源分離(Source Separation)有相當好的效果，因此更進階的做法可以將音樂分離為不同人聲、鼓聲、低音線(bassline) 以及其餘的聲音，再依照 CI 使用者的偏好給予不同的增益，然後重新混合為聲音，提供後端的編碼策略進一步處理，以促進更佳的音樂感知與體驗(Nogueira et al., 2019)。

5.3 ElectroNet 深度學習聲音編碼策略

有別於多數人工智慧聲音訊號處理方法是以 CI 訊號的前處理為主，比較少針對聲音編碼的部分，因此直接運用深度學習為聲音編碼策略相當值得探討。(Huang et al., 2023)提出了 ElectroNet 策略，它將 ACE 策略的包絡偵測(Envelope Detection)以 MLP/DNN 網路取代，其結果在中英文二種語言及四種噪音下與 ACE 策略均呈現高度的相關性。而同時涵蓋包絡偵測與頻道選擇功能的 ElectroNet-CS 策略(圖 4)，由於它的 CS 功能直接參與了網路模型的訓練過程，因此在輸出後經聲碼器模擬的語音，與 ACE 策略的表現非常接近，包括客觀評估及主觀的正常聽力個案模擬實驗，在結果上沒有統計上的顯著差異。上述的實驗僅針對聲音編碼策略的表現進行評估，其訓練條件未採用進行語音增強，可證

實以神經網路為基礎的聲音編碼策略能達到相近於 ACE 策略的結果，因此為相關研究開啟了嶄新的方向。

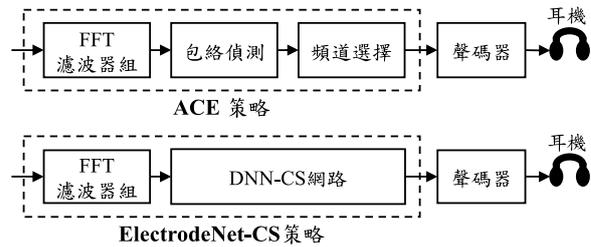


圖 4. ACE 與 ElectroNet-CS 聲音編碼策略

6 討論與未來展望

人工電子耳的研究，除了幫助聽損者有更好的聆聽效果並進而改善其生活品質，對於聲音感知的生理及心理機制也能促進更深入的認識。由於電聽覺(Electric Hearing) 可提供聲聽覺(Acoustic Hearing)較無法觀察到的一些面向，還有許多值得探索之處。

人工智慧相關的方法目前還在研究階段，距離應用在 CI 產品上仍有一段距離。依照相關領域的發展，目前已有部分助聽器產品開始採用 DNN 網路，而在自動語音辨識(Automatic Speech Recognition, ASR)領域也有晶片的出現，因此在電子耳方面，相似的 AI 應用是指日可待，而更複雜的網路如何在聲音處理器上以低功耗與低延遲的方式實現，尚有許多努力的空間。此外，更強大的網路架構，例如聯合訓練(Joint-Training)及端對端學習(End-to-End Learning)，還有許多可以探討(Huang et al., 2023)。

CI 相關領域仍有許多未完全了解的問題，例如在 2000 年提出的 ACE 策略後缺乏有效提升聆聽效果的臨床聲音編碼策略、極少數的單電極電子耳使用者可理解語音的原因(House, 2011)、兒童使用者在音樂訓練後的表現較成人使用者更佳的神經可塑性(Neuroplasticity)原理(Ab Shukor et al., 2021)，以及如何克服使用者變異性以更佳地預測聆聽效果(Wilson, 2019)。這些問題除了有待專家學者的進一步探索，人工智慧的輔助或許能提供一部份的解答。

在「AI+CI」的浪潮下，本文透過回顧相關歷史發展的足跡，了解電子耳聲音處理的原理及評估方法，介紹目前 AI 在此領域的成果，期盼能提供相關領域一些啟發。

參考文獻

- Nor F. Ab Shukor, Jihyeon Lee, Young Joon Seo, and Woojae Han. 2021. Efficacy of music training in hearing aid and cochlear implant users: A systematic review and meta-analysis. *Clinical and Experimental Otorhinolaryngology*, 14(1), pages 15-28.
- Graeme M Clark. 2015. The multi-channel cochlear implant: Multi-disciplinary development of electrical stimulation of the cochlea and the resulting clinical benefit. *Hearing Research*, 322, pages 4-13.
- Jui-Ling Chen, Chi-Hsun Wang, Jiunn-Liang Wu, Jun-Yang Fan, Yu-Fu Chou, and Hung-Ching Lin. 2015. 台灣實施公費新生兒聽力篩檢之過去、現況與未來(The past, present and future of newborn hearing screening program in Taiwan) [in Chinese]. *Journal of Taiwan Otolaryngology-Head and Neck Surgery*. 50(2), pages 67-73.
- Matthew G. Crowson, Vincent Lin, Joseph M. Chen, and Timothy C. Y. Chan. 2020. Machine learning and cochlear implantation—a structured review of opportunities and challenges. *Otology & Neurotology*, 41(1), pages e36-e45.
- Tiago h. Falk, Vijay Parsa, João F. Santos, Kathryn Arehart, Oldooz hazrati, Rainer huber, James M. Kates, and Susan Scollie. 2015. Objective quality and intelligibility prediction for users of assistive listening devices: Advantages and limitations of existing tools. *IEEE Signal Processing Magazine*, 32(2), pages 114-124.
- Fergal Henry, Martin Glavin, and Edward Jones. 2023. Noise reduction in cochlear implant signal processing: A review and recent developments. *IEEE Reviews in Biomedical Engineering*, 16, pages 319-331.
- Tobias Goehring, Mahmoud Keshavarzi, Robert P. Carlyon, and Brian C. J. Moore (2019). Using recurrent neural networks to improve the perception of speech in non-stationary noise by people with cochlear implants. *The Journal of the Acoustical Society of America*, 146(1), pages 705-718.
- William F. House. 2011. The struggles of a medical innovator: Cochlear implants and other ear surgeries. *CreateSpace Independent Publishing Platform*, pages 179-193.
- Chi-Ming Huang. 2022, October 18. 人工電子耳在台三十年 助 3000 名重度聽損者重返有聲(30 Years of Cochlear Implants in Taiwan: Restoring Auditory Experience for 3,000 Profoundly Hearing-Impaired Individuals). [in Chinese]. *Economic Daily News*. accessed August 15, 2023, from <https://money.udn.com/money/story/11800/6696245>.
- Enoch Hsin-Ho Huang, Chao-Min Wu, and Hung-Ching Lin. 2021. Combination and comparison of sound coding strategies using cochlear implant simulation with mandarin speech. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29, pages 2407-2416.
- Enoch Hsin-Ho Huang, Rong Chao, Yu Tsao, and Chao-Min Wu. 2023. ElectrodeNet—A Deep Learning Based Sound Coding Strategy for Cochlear Implants. *IEEE Transactions on Cognitive and Developmental Systems*, pages 1-12. Accepted. <https://doi.org/10.1109/TCDS.2023.3275587>.
- Enoch Hsin-Ho Huang. 2023. Investigations of cochlear implant sound coding strategies based on auditory physiology and deep learning. PhD dissertation. National Central University, Taiwan.
- Ying-Hui Lai, Fei Chen, Syu-Siang Wang, Xugang Lu, Yu Tsao, and Chin-Hui Lee. 2017. A deep denoising autoencoder approach to improving the intelligibility of vocoded speech in cochlear implant simulation,” *IEEE Transactions on Biomedical Engineering*, 64(7), pages 1568-1578.
- Nicholas A. Lesica, Nishchay Mehta, Joseph G. Manjaly, Li Deng, Blake S. Wilson, and Fan-Gang Zeng. 2021. Harnessing the power of artificial intelligence to transform hearing healthcare and research. *Nature Machine Intelligence*, 3(10), pages 840-849.
- Ruth Y. Litovsky, Matthew J. Goupell, Alan Kan, and David M. Landsberger. 2017. Use of research interfaces for psychophysical studies with cochlear-implant users. *Trends in Hearing*, 21, pages 1-15.
- Philipos C. Loizou. 1998. Mimicking the human ear. *IEEE Signal Processing Magazine*, 15(5), pages 101-130.
- Nursadul Mamun, Soheil Khorram, and John H.L. Hansen, 2019. Convolutional neural network-based speech enhancement for cochlear implant recipients, In *Interspeech*, pages 4265-4269.
- Brian CJ. Moore. 2003. Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants. *Otology & Neurotology*, 24(2), pages 243-254.
- Waldo Nogueira, Andreas Büchner, Thomas Lenarz, and Bernd Edler. 2005. A psychoacoustic "NofM"-type speech coding strategy for cochlear implants. *EURASIP Journal on Advances in Signal Processing*, 18, pages 3044-3059.
- Waldo Nogueira, Anil Nagathil, and Rainer Martin. 2019. Making music more accessible for cochlear implant listeners: Recent developments. *IEEE Signal Processing Magazine*, 36(1), pages 115-127.

- Rung-Yu Tseng, Tao-Wei Wang, Szu-Wei Fu, Chia-Ying Lee, and Yu Tsao. 2021. A study of joint effect on denoising techniques and visual cues to improve speech intelligibility in cochlear implant simulation. *IEEE Transactions on Cognitive and Developmental Systems*, 13(4), pages 984-994.
- Andrew E. Vandali, Lesley A. Whitford, Kerrie L. Plant, and Graeme M. Clark. (2000). Speech perception as a function of electrical stimulation rate: Using the Nucleus 24 cochlear implant system. *Ear and Hearing*, 21(6), pages 608-624.
- Natalie Yu-Hsien Wang, Hsiao-Lan Sharon Wang, Tao-Wei Wang, Szu-Wei Fu, Xugan Lu, Hsin-Min Wang, and Yu Tsao. 2021. Improving the intelligibility of speech for simulated electric and acoustic stimulation using fully convolutional neural networks. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29, pages 184-195.
- Blake S. Wilson, Charles C. Finley, Dewey T. Lawson, Robert D. Wolford, Donald K. Eddington, and William M. Rabinowitz. 1991. Better speech recognition with cochlear implants. *Nature*, 352(6332), pages 236-238.
- Blake S. Wilson. 2019. The remarkable cochlear implant and possibilities for the next large step forward. *Acoustics Today*, 15(1), pages 53-61.
- Jan Wouters, Hugh Joseph McDermott, and Tom Francart. 2015. Sound coding in cochlear implants: From electric pulses to hearing. *IEEE Signal Processing Magazine*, 32(2), pages 67-80.
- Che-Ming Wu and Yung-Ting Tsou. 2015. 人工電子耳綜論(Cochlear Implantation) [in Chinese]. *Journal of Taiwan Otolaryngology-Head and Neck Surgery*. 50(4), pages 197-210.
- Fan-Gang Zeng, Stephen Rebscher, William Harrison, Xiaohan Sun, and Haihong Feng. 2008. Cochlear implants: System design, integration, and evaluation. *IEEE Reviews in Biomedical Engineering*, 1, pages 115-142.
- Fan-Gang Zeng. 2022. Celebrating the one millionth cochlear implant." *JASA Express Letters*, 2(7), 077201, pages 1-8.