國 立 清 華 大 學

碩 士 論 文

使用卷積傳遞函數進行聲學傳遞函數盲估計以應用於去混響、聲源分離以及增強

Blind estimation of acoustic transfer functions with application to dereverberation, separation and enhancement using convolutive transfer functions

系級：動力機械工程學系碩士班

組別：電機控制組

學號姓名：111033537 袁安志 Anchi Yuan

指導教授：白明憲 博士 (Dr. Mingsian R. Bai)

中 華 民 國 一 一 三 年 六 月

# 摘要

雖然聲學傳遞函數 (Acoustic Transfer Functions) 在各種陣列信號處理應用中通常優於相對傳遞函數 (Relative Transfer Functions)，但無法獲得源信號會對獲得可靠的聲學傳遞函數估計構成重大挑戰。為了解決這一問題，我們提出了一種基於卷積傳遞函數 (Convolutive Transfer Functions) 的創新盲估計方法。最初，使用加權預測誤差 (Weighted Prediction Error) 算法和基於到達時間差 (Time Difference of Arrival) 預測出之源位置的延遲和加總 (Delay and Sum) 波束形成器來獲取目標源信號的初始估計。隨後，使用維納濾波器或卡爾曼濾波器計算卷積傳遞函數係數，並通過粒子群優化 (Particle Swarm Optimization) 優化其參數。為了獲取聲學傳遞函數的房間脈衝響應，將單位脈衝序列的短時傅立葉變換 (Short-time Fourier Transform) 與卷積傳遞函數係數進行卷積，然後應用逆短時距傅立葉變換（Inverse Short-time Fourier Transform）來獲得它。為了展示我們所提出的聲學傳遞函數估計技術的有效性，我們對估計的脈衝響應與最先進的盲系統識別方法——自適應多通道時域最小均方（Adaptive Multi-channel Time Domain Least Mean Square） 和真實脈衝響應進行了比較分析。此外，本文展示了幾個需要精確聲學傳遞函數估計的應用，包括使用多輸入/輸出逆定理 (Multiple Input/Output Inverse Theorem) 進行去混響、通過吉洪諾夫正則化（Tikhonov regularization）進行聲源分離以及使用最小功率無失真響應 (Minimum Power Distortionless Response) 波束形成器進行聲源增強。在真實房間環境中進行的實驗結果表明，我們的方法能夠準確地估計聲學傳遞函數。

***關鍵詞 ― 卷積傳遞函數，加權預測誤差算法，到達時間差，延遲和加總波束成形器，維納濾波器，卡爾曼濾波器，粒子群優化，自適應多通道時域最小均方，多輸入/輸出逆定理，吉洪諾夫正則化，最小功率無失真響應波束形成器***

# ABSTRACT

While Acoustic Transfer Functions (ATFs) often outperform Relative Transfer Functions (RTFs) in array signal processing, obtaining reliable ATF estimates without the source signal is challenging. To address this, we propose a novel blind ATF estimation approach based on Convolutive Transfer Functions (CTFs). Initially, the Weighted Prediction Error (WPE) algorithm and Delay and Sum (DAS) beamformer with source location predicted by Time Difference of Arrival (TDOA) provide an initial estimate of the target source signal. Then, CTF coefficients are calculated using either the Wiener filter or Kalman filter, with parameters optimized via Particle Swarm Optimization (PSO). To obtain room impulse responses (RIR) of ATFs, the short-time Fourier transform (STFT) of a unit pulse sequence is convolved with the CTF coefficients, followed by the inverse STFT. We demonstrate the efficacy of our ATF estimation technique through a comparative analysis with the state-of-the-art Adaptive Multi-channel Time Domain Least Mean Square (MCLMS) method against the ground-truth RIR. Additionally, this paper highlights applications requiring precise ATF estimates, including dereverberation with MINT, source separation using Tikhonov Regularization (TIKR), and source enhancement with Minimum Power Distortionless Response (MPDR) beamformer. Realistic room environment experiments validate our method's accuracy in estimating ATFs.

***Index Terms — convolutive transfer functions, weighted prediction error, delay and sum beamformer, Time Difference of Arrival, Wiener filter, Kalman filter, particle swarm optimization, adaptive multi-channel time domain least mean square, multiple input/output inverse theorem, Tikhonov regularization, minimum power distortionless response beamformer***