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半盲估計以應用於去混響、聲源分離以及增強

Semi-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），但源信號的不可取用性對於獲得可靠的聲學傳遞函數估算構成了重大挑戰。為應對這一問題，我們提出了一種新穎的基於卷積傳遞函數（Convolutive Transfer Functions）的聲學傳遞函數盲估算方法。首先，使用加權預測誤差算法（Weighted Prediction Error）和延遲和加總（Delay and Sum）波束成形器來獲得聲源位置的目標源信號之初始估算。隨後，使用維納濾波器或卡爾曼濾波器計算卷積傳遞函數之係數，並通過粒子群優化（Particle Swarm Optimization）來最佳化濾波器中的參數。為了得到聲學傳遞函數的脈衝響應，將單位脈衝序列的短時距傅立葉變換（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）針對與真實脈衝響應之間的一致性進行了比較分析。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。此外，我們還在真實房間環境中進行了實驗。這些實驗結果表明，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，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 various array signal processing applications, the unavailability of the source signal poses a significant challenge for obtaining reliable ATF estimates. In response to this issue, we introduce an innovative approach for blind ATF estimation, which is based on Convolutive Transfer Functions (CTFs). Initially, the Weighted Prediction Error (WPE) algorithm and the Delay and Sum (DAS) beamformer are employed to acquire an initial estimate of the target source signal at the source's location. Subsequently, the CTF coefficients are calculated employing either the Wiener filter or the Kalman filter, with their respective parameters optimized by the application of Particle Swarm Optimization. (PSO). To retrieve the 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 application of the inverse STFT. In order to demonstrate the efficacy of our proposed ATF estimation technique, a comparative analysis was conducted between the estimated RIR and the state-of-the-art blind system identification method, Adaptive Multi-channel Time Domain Least Mean Square (MCLMS), with regard to the ground-truth RIR. Furthermore, this paper illustrates several applications that necessitate precise ATF estimates, including dereverberation using the Multiple Input/Output Inverse Theorem (MINT), source separation via Tikhonov regularization (TIKR), and source enhancement employing the Minimum Power Distortionless Response (MPDR) beamformer. Furthermore, experiments were conducted in realistic room environments. The outcomes of these experiments demonstrate that our method is capable of accurately estimating the ATF.

***Index Terms — convolutive transfer functions, weighted prediction error, delay and sum beamformer, 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***