Abstract

MODERN speech communications are evolving towards a new direction which involves users in a more perceptive way. That is the *immersive experience*, which may be considered as the "last-mile" problem of telecommunications.

One of the main feature of immersive communications is the *distant-talking*, i.e. the hands-free (in the broad sense) speech communications without body-worn or tethered microphones that takes place in a multisource environment where interfering signals may degrade the communication quality and the intelligibility of the desired speech source.

In order to preserve speech quality *intelligent acoustic interfaces* may be used. An intelligent acoustic interface may comprise multiple microphones and loudspeakers and its peculiarity is to model the acoustic channel in order to adapt to user requirements and to environment conditions. This is the reason why intelligent acoustic interfaces are based on adaptive filtering algorithms.

The acoustic path modelling entails a set of problems which have to be taken into account in designing an adaptive filtering algorithm. Such problems may be basically generated by a linear or a nonlinear process and can be tackled respectively by linear or nonlinear adaptive algorithms.

In this work we consider such modelling problems and we propose novel effective adaptive algorithms that allow acoustic interfaces to be robust against any interfering signals, thus preserving the perceived quality of desired speech signals.

As regards *linear adaptive algorithms*, a class of adaptive filters based on the sparse nature of the acoustic impulse response has been recently proposed. We adopt such class of adaptive filters, named *proportionate adaptive filters*, and derive a general framework from which it is possible to derive any linear adaptive algorithm. Using such framework we also propose some efficient proportionate adaptive algorithms, expressly designed to tackle problems of a linear nature.

On the other side, in order to address problems deriving from a nonlinear process, we propose a novel filtering model which performs a nonlinear transformations by means of *functional links*. Using such nonlinear model, we propose *functional link adaptive filters* which provide an efficient solution to the modelling of a nonlinear acoustic channel.

Finally, we introduce robust filtering architectures based on *adaptive combinations of filters* that allow acoustic interfaces to more effectively adapt to environment conditions, thus providing a powerful mean to immersive speech communications.

Acknowledgments

The accomplishment of the Ph.D. degree is cause of huge gratification to me so I would like to share my contentment with all the people who help me to attain this goal.

First and foremost, I would like to express my sincere appreciation and gratitude to my advisor Prof. Aurelio Uncini for his valuable guidance and continuous support throughout the entire project. He transmitted his passion for research to me and taught me how to strive for excellence in my professional pursuits.

I am also in debt to Prof. Michele Scarpiniti for being my well-wisher. I would like to thank him for his invaluable advices, besides being an exemplary workmate he is firstly a friend.

Special thanks to Dr. Mauro Falcone, I am grateful to him for giving me an opportunity to work with him and to confront myself with an important corporate reality, such as the "Fondazione Ugo Bordoni".

My sincere thanks also go to my group member, Prof. Raffaele Parisi for the insightful discussions and help along the way.

I am also thankful with Prof. Jerónimo Arenas García who has incisively increased my professional skills during my visiting period at "Universidad Carlos III de Madrid", thus resulting decisive for my thesis work. It was a great pleasure to work with him both from a professional and a personal point of view. I would like to thank him and Dr. Luis Azpicueta Ruiz for their precious welcome to Madrid.

I would like to thank my dissertation committee members: Prof. Giovanni Iacovitti, Prof. Antonello Rizzi, Prof. Marco Listanti for taking the time to serve on my committee and for their helpful suggestions.

A fundamental role has been played by my flatmates and friends, whose presence is decisive not only in my work but above all in my life.

My best thanks are to Alessia who is continuously at my side, cheering me up in hard times and fortunately sharing with me joys and gratifications. This work is dedicated to her.

To make this acknowledgment complete, a special thanks to my family for their unconditional love and support, which added the best essence to my work.

CONTENTS

List of Figures

List of Tables

List of Acronyms

AEC Acoustic Echo Cancellation

AI Artificial Intelligence

AIR Acoustic Impulse Response

ANC Adaptive Noise Canceller

ANN Artificial Neural Network

APA Affine Projection Algorithm

AZK All-Zero Kernel

BCFLAF Block-Based Collaborative FLAF

BM Blocking Matrix

CANC Combined Adaptive Noise Canceller

CFLAF Collaborative FLAF

DSB Delay-and-Sum Beamformer

DSP Digital Signal Processing

DTD Double Talk Detector

EG± Exponentiated Gradient

EMSE Excess Mean Square Error

ERLE Echo Return Loss Enhancement

FEB Functional Expansion Block

FIR Finite Impulse Response

FLAF Functional Link Adaptive Filter

FLANN Functional Link Artificial Neural Network

GSC Generalized Sidelobe Canceller

IAI Intelligent Acoustic Interface

IIR Infinite Impulse Response

IPAPA Improved PAPA

IPNLMS Improved Proportionate NLMS

ISO International Organization for Standardization

LMS Least Mean Square

LSI Linear Shift-Invariant

MIMO Multiple-Input Multiple Output

MISO Multiple-Input Single-Output

MLP Multi-Layer Perceptron

MMSE Minimum Mean Square Error

MSE Mean Square Error

NAEC Nonlinear Acoustic Echo Cancellation

NAPA Natural Affine Projection Algorithm

NLMS Normalized Least Mean Square

NNG Normalized Natural Gradient

PAPA Proportionate APA

PBAPA Proportionate Block APA

PNLMS Proportionate NLMS

QAM Quadrature Amplitude Modulation

RBF Radial Basis Function

RLS Recursive Least Squares

SFLAF Split FLAF

SISO Single-Input Single-Output

SNR Signal to Noise Ratio

ULA Uniform Linear Array

VAD Voice Activity Detection

VF Volterra Filter

VSS Variable Step Size