MECHATRONICS BOOK SERIES SYSTEM DESIGN AND SIGNAL PROCESSING

VOLUME 1

Editors
Asan G. A. Muthalif
Amir Akramin Shafie
Siti Fauziah Toha
Iskandar Al-Thani Mahmood

IIUM PRESS

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

MECHATRONICS BOOK SERIES: SYSTEM DESIGN AND SIGNAL PROCESSING - VOLUME 1

Editors

Asan G. A. Muthalif
Amir Akramin Shafie
Siti Fauziah Toha
Iskandar Al-Thani Mahmood

IIUM Press

Published by:
IIUM Press
International Islamic University Malaysia

First Edition, 2011

Perpustakaan Negara Malaysia Cataloguing-in-Publication Data

ISBN: 978-967-418-173-4

Member of Majlis Penerbitan Ilmiah Malaysia – MAPIM
(Malaysian Scholarly Publishing Council)

Printed by :
IIUM PRINTING SDN.BHD.
No. 1, Jalan Industri Batu Caves 1/3
Taman Perindustrian Batu Caves
Batu Caves Centre Point
68100 Batu Caves
Selangor Darul Ehsan
Tel: +603-6188 1542 / 44 / 45 Fax: +603-6188 1543
EMAIL: iiumprinting@yahoo.com

CONTENTS

CHAPTER 34

Sound Identification in Noisy Environment

Nahrul Khair Alang Md Rashid, Nor Hidayati Diana Nordin and Alim Sabur Ajibola

Department of Mechatronics Engineering

International Islamic University Malaysia

Jalan Gombak, Selangor D.E., Malaysia

nahrul@iium.edu.my

34.1 Introduction

Noise can be defined as an unwanted signal that carries no meaning. Noise from various sources such as lawn mower, drilling and other electrical devices are everywhere around us. Noise level that is above 80 dB is considered as hazardous sounds. Sound identification in noisy environment has gained interest among the researches since many years ago.

Sound recognition has applications in areas such as a speech recognition system implemented as a part of the Teaching and Learning Using Information Technology (TLIT) [1], speaker identification [2, 3] automatic speaker recognition [4], phoneme recognition [5], word recognition [6], voice recognition [7] and emotion recognition [8].

A lot of algorithms and methods have been developed in order to produce the most efficient and robust sound identification (SI) system. This system can be implemented anywhere including underwater. The main purpose of sound identification in noisy environment is to remove or filter any unwanted signal, so that only the desired signal can be passed through to the user. This is because some applications do not perform well in noisy environment. There is also the need for picking up specific sound buried in noise.

34.2 Sound Identification

Generally sound recognition is a two-stage process, viz. (1) Identification and subsequent extraction of relevant sound features, and (2) Use of the features for sound identification.

34.2.1 Feature Extraction: There are several methods available for extracting sound features that can be used for its classifications. This includes Mel Frequency Cepstral Coefficient (MFCC) and Linear Prediction Coefficient (LPC) that are the most widely used. The objective of feature extraction is to represent a speech signal by a finite number of characteristics of the signal. This is because all the information in the speech signal is too much to process and not all this information is relevant for the identification process [9]. Due to the essentiality of this process, quality features beget quality classification.

34.2.1.1 Mel Frequency Cepstral Coefficient (MFCC)- The main purpose of the Mel Frequency Cepstral Coefficient is to mimic the behavior of the human ears. In first step, the continuous speech signal is blocked into frames of N samples, with adjacent frames being separated by M ($M < N$). The next step in the processing is to window each individual frame so as to minimize the signal discontinuities at the beginning and end of each frame. Typically the Hamming window is used. The next processing step is the Fast Fourier Transform, which converts each frame of N samples from the time domain into the frequency domain. After that the scale of frequency is converted from linear to mel scale. In final step, the log mel spectrum is converted back to time domain. The result is called the mel frequency cepstrum coefficients (MFCC). The cepstral representation of the speech