ZONE-BASED ACTIVE NOISE CONTROL FOR AN AIRCRAFT PASSENGER SEAT

by

Wintta Ghebreiyesus

Bachelor of Engineering, Ryerson University (2014)

A thesis

Presented to Ryerson University

In partial fulfilment of the

Requirements for the degree of

Master of Applied Science

In the Program of

Aerospace Engineering

Toronto, Ontario, Canada, 2017

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A THESIS

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

ZONE-BASED ACTIVE NOISE CONTROL FOR AN AIRCRAFT PASSENGER SEAT

Wintta Ghebreiyesus

Master of Applied Science, Aerospace Engineering, Ryerson University (2017)

ABSTRACT

The goal of this research is to improve zone-based local active noise control for an aircraft passenger seat using head tracking and virtual sensing methods. Broadband diffuse sound fields are analyzed in order to determine the level of attenuation around the passenger's ears. The virtual sensing methods which were evaluated from literature include the virtual microphone technique, the forward difference prediction technique, and the adaptive LMS moving virtual microphone techniques. In addition to virtual sensing, a new methodology for integrating zone-based technologies with existing local ANC techniques has been developed. The virtual sensing simulation and head tracking measurements can be used to verify this methodology.

ACKNOWLEDGEMENTS

Table of Contents

List of Tables

List of Figures

Nomenclature

CSM cross spectral matrix

FFT Fast Fourier Transform

$HRTF$ head-related transfer function

l_2 width of boundary vertical to the boundary i

l_3 distance between receiver and cross-section with P_{IN}

P_{IN} input power

$P_{S0,i}$ power flow at the receiver caused by the boundary i

PWL sound power level

SH spherical harmonic

SPL sound pressure level

y $= \alpha_i l_3 / 2 l_2$

α_i absorption coefficients of the boundary i

$si(y), ci(y)$ $-\int_y^\infty (sint/t)dt, \quad -\int_y^\infty (cost/t)dt,$

TL transmission loss

1. Introduction

1.1. Background

There is a need for business jet cabin noise reduction research. The current methods for business aircraft cabin noise reduction or cancellation are not sufficient as passengers continue to be dissatisfied with loud flying environments. Generally, aircraft noise control can be classified as being environmental (outside of the aircraft) or cabin-related (inside of the aircraft). Cabin-related aircraft noise control can further be classified as being overall or zone-based. Most methods of business jet cabin noise cancellation occur outside the cabin and within the cabin trim or inner lining. Unfortunately, these methods do not target vibration-borne noise from the cabin itself or noise that propagates from inside of the cabin. In addition to this, the dominant sources of noise in an aircraft, which are the engine harmonic tones, still remain undamped by these noise reduction techniques. Noise reduction inside the cabin of a business jet is critical for providing passengers with optimal sound comfort. This thesis presents zone-based active noise control in a business aircraft passenger seat. This chapter is a brief overview of the subsequent chapters presented in this thesis.

According to Guigou et al., Turbofans, propellers and turbulent boundary layers are some causes of noise fields which propagate sound waves through the fuselage and create extremely high sound levels inside the aircraft cabin [23]. Major noise sources originating from the airframe of an aircraft include horizontal stabilizers, spoiler, flaps, leading and trailing edge devices, the wings, landing gears, nacelles, and the fuselage. Aircraft engine noise sources include fan exhaust, turbine and core, and fan inlet. Both passive and active cabin noise control methods have been investigated in recent years.

Moreover, the concept of head-tracking for localized active noise control creates a quiet zone in close vicinity to a single passenger. The head-tracking system has a fast face detector. The applications of a fast face detector are diverse and include user interfaces, image databases, and teleconferencing [68]. Most head tracking technologies impose the use of physical and expensive head gear. This thesis will present a method that does not require the user to wear any equipment for both the head tracking and noise reduction.

The work in the following chapters will attempt to solve this problem: How can a business jet passenger's sound comfort be improved using zone-based active noise control in a seat? The main objective is to reduce business aircraft cabin noise levels near a passenger's seat using an integrated active noise control (ANC) system. The system will consist of a headrest, microphone sensors, microcontroller, speakers, and a head-tracking device. Since no human participants will be used in the study, an acoustic dummy head will be the main subject. The following is a list of specific objectives of this research:

- To assess overall sound pressure levels in business jet cabins

- To analyse active noise control techniques near passenger's head

- To create a head tracking system while unrestricting passenger movement within personal space

- To use the head tracking system to design a novel method for determining optimal ANC zone for passenger

1.2. Scope

The purpose of this study is to improve zone-based active noise control for an aircraft passenger seat using virtual sensing and head tracking methods. The effects of head tracking on active noise control (ANC) localization will be studied in order to enhance personal sound comfort in aircraft cabins. Figure 1.1 below demonstrates the conceptual framework of the head tracking part of the study. The general process starts with the dummy head in a neutral front-facing configuration. The original pose of the smart structure is used for the first iteration of the ANC algorithm. Then, the dummy head is rotated or translated such that the head tracking subsystem is able to output a tracked posed to the ANC subsystem. The results from before and after head tracking are compared to determine the effects of head tracking on ANC localization. In this thesis, ANC localization represents an active noise cancelling technique that produces cancellation close to a user's head regardless of its pose.
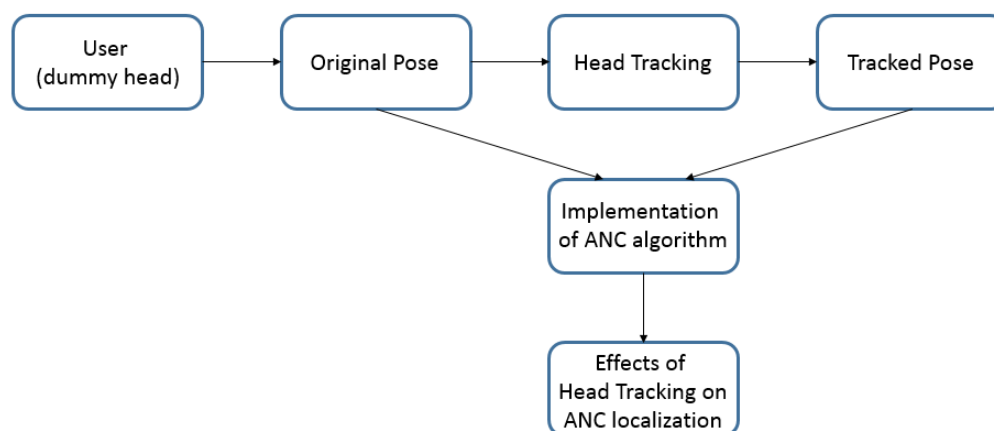
Fig. 1.1 Diagram of head tracking ANC framework.

The subject of this quantitative study is an acoustic dummy head that represents a human head and is used for the head tracking portion of the thesis. Since the dummy head is capable of housing

microphones, it can be used to validate the active noise control localization methodology in the future. The results of this study can be used for mock-up cabin testing with similar dummy heads and may not directly apply to human participants. More information about the dummy head subject can be found in Chapter 4. This work has used an extensive list of instrumentation including the National Instruments signal analyzer, the Texas Instruments digital signal processor, Arduino Uno microcontrollers, the Intel F200 camera motion sensing device, the Microsoft Kinect motion sensing device and the Thinkpad X1 Carbon PC. An assumption that was made was that the dummy head exhibited humanistic features such as eyes, ears, a nose, and a mouth. These features were exaggerated with multicolored-tape markers in order to use cascades for the head tracking. This was an important assumption to make because if human participants are ever used for such a study, a similar method of creating several positive and negative images for cascade selection would have to be employed.

The scope of this work entails research conducted in the department of aerospace's manufacturing and robotics laboratory facility at Ryerson University. A miniature cabin mock-up was used for several of the experiments. The miniature cabin mock-up consists of a wooden, elevated floor housed in an extruded aluminum frame (floor and hollow sides for structural support). The Bombardier Global Express 5000 aircraft seat was securely mounted onto a marine swivel base, which was then secured to the floor via custom-made extruded aluminum mounts. The scope can be further explained in four main points as follows:

- Analyze existing active noise control methods pertaining to passenger seat (i.e. headrest)

- Develop methodology for zone-based integrated ANC system

- Develop head tracking tool to improve ANC system

This thesis contains 5 chapters. Chapter 2 will discuss the literature review which was conducted. This review will go over existing noise control and head tracking algorithms and technologies. More specifically, Chapter 2 will look at the following topics: current research in cabin noise control, both overall cabin control and zone-based cabin control; active noise control and noise measurement; smart materials for noise vibration reduction; modeling and simulation of near-field sound propagation; modeling and measurement of cabin noise distribution; passive noise isolation and sound absorbent materials; modeling human hearing range; visual object detection; image registration; dummy head for room acoustic measurements. Chapter 3 will introduce the concept and effectiveness of Zone of Quiet (ZoQ). It will also discuss the filtered-x LMS algorithm as well as three virtual sensing methods: (1) virtual microphone technique, (3) forward difference prediction technique, and the (4) adaptive LMS moving virtual microphone technique. Chapter 3 will also demonstrate the practical implementation of zone-based local ANC systems. Chapter 4 will introduce the concept of using head tracking to improve the zone-of-quiet in addition to virtual sensing algorithms. Chapter 4 includes: the acoustic dummy head model; production of the model's humanistic facial features; the HSV scale and masking processes; and the rotation and translation methods used to obtain the best head-tracking results. Finally Chapter 5 is the conclusion of the work and will also discuss future work on zone-based active noise control for aircraft passenger seats.

2. Literature Review

This literature review consists of six miniature chapters which pertain to the broad field of active and passive noise control and reduction. These main themes include: active noise control and noise measurement, smart materials for noise and vibration, modeling and simulation of near-field sound propagation, modeling and measurement of aircraft cabin distribution, and modeling human hearing range including frequency and amplitude. This report describes the methods and evidences reviewed from existing and published literary texts.

2.1. Active noise control and noise measurement

The principle of active noise control is to create a zone of silence, and at the tympanic membrane, canceling the effect of noise leakage in a standard open-fitting hearing aid [59]. Adaptive filtering is one of many advanced processing techniques. In adaptive filtering, the filter coefficients are allowed to change in time. Consider an active noise control optimisation problem where noise is fed into a room and collected by a microphone. The noise is also fed to the adaptive filter which is designed to process the signal in exactly the same way as the room system. The error is essentially the difference between the microphone signal and the output of the adaptive filer. In theory, this error should be zero. If it is not zero, this error is used to update the adaptive filter, so it makes a better estimation next time. The adaptive filter is either a finite impulse response (FIR) or an infinite impulse response (IIR) filter. The major difference between the two is their corresponding filter coefficients. In a larger sense, this is similar to a numerical optimisation problem which can be tackled by a genetic algorithm (GA). However, in adaptive filtering the Least Mean Square (LMS) approach is more popular. This is important because in a

speech communication system, the intelligibility of the spoken messages must be maintained when threatened by noise. This noise may be created acoustically or electrically, or it may be the reverberant effects of a long system impulse response generating a signal which resembles noise (i.e. room reverberation). [34] Luckily, due to linearity consequences, such noise may be cancelled using signal processing techniques.

Consider an active noise control system. The speech signal s is corrupted by the addition of the noise signal n at the first summing node, generating the observable signal d. At the second summing node, a signal y is subtracted from d. Again, the error signal e is the result of this subtraction.

- If the signal y is a copy of signal n, then the noise corruption on the signal s is removed, $e = s$.

- If y is a reasonable approximation of n, then some of the noise contamination is removed, $e \approx s$.

- If y is largely uncorrelated with n, then the second summing node represents an additional source of noise, further corrupting the speech component in e.

The cancelling signal y is derived by filtering operations (through the filter block W which is an adaptive filter) on the reference signal x. All the signal variables and the filter responses are complex functions of frequency. Averaging of the mean squared error signal can be used to deal with the non-deterministic signal types. This is done using the expected value operator $E[|e|^2]$.

The following assumptions pertain to the statistical relationship between the signals:

(1) The noise n is uncorrelated with the speech s.

(2) The reference x is correlated with the noise n (and so, by 1), is uncorrelated with the speech s.

The mean square error will have a unique minimum value when the complex filter W is correctly adjusted to W_{opt} which minimizes the error, which will be shown next. W_{opt}, by definition, is the transfer function between the reference signal x (interpreted as input) and the noise signal n (interpreted as output). Another way of putting this is that the optimal configuration of the cancelling filter W_{opt}, is the inverse of the filter relating n (input) and x (output). In the idealised case of perfect correlation between n and x of a simple noise canceller, the transfer function e/s is easily found to be 1. Consequently, the noise added at the first summing node is perfectly cancelled at the second summing node. The noise is attenuated by ∞ dB. In reality, this is never the case. In other words, in practice perfect performance cannot ever be achieved for many reasons such as:

(1) Imperfect implementation of the cancelling filter.

(2) Imperfect correlation between the noise n and the reference x.

The second reason will be analyzed. Imperfect correlation between n and x can be modelled by the system presented by *Lecture Notes on the Mathematics of* Acoustics [34]. The additional noise n_2 represents those components of x which are not correlated with n (it can further be assumed that n_2 is independent of s). The linear time invariant (LTI) transfer function relating the correlated components of x to n is relabelled H for generality.

It should be noted that the introduction of the uncorrelated noise does not affect the phase of the optimal filter (defined by H-inverse), it only scales the optimal canceller gain by the coherence. It should also be noted that the coherence function is purely real. The attenuation of noise is a function of the coherence between the noise and reference signals (and the magnitude of the

transfer function H). The attenuation of noise increases as the coherence increases. However, the coherence function is bounded in magnitude: $0 < |\gamma_{n,x}^2| < 1$). In order to produce sufficient levels of noise attenuation, high coherence between the reference and the noise signal to be cancelled must be present. However, this is difficult in practice.

2.2. Smart materials for noise vibration reduction

For the past decade, more and more researchers are becoming interested in the reduction of sound and/or vibrations using hybrid active-passive techniques. The main goal of hybrid active-passive devices for noise control is to increase performance at lower frequencies; however, some by-products include increased robustness, increased stability and decreased control spillover [20]. According to Gentry et al., there are two general types of active-passive noise control approaches: the adaptive-passive and active-passive techniques [20]. The adaptive-passive methods are passive devices whose static properties are adapted or changed to optimize their performance [46]. Technologies involving smart structures have found their place in fields pertaining to actively-controlled vibrations, noise, and deformations [42]. For aircraft, so-called smart skin is designed to reduce sound through passive absorption of acoustic foam and the active component which is the piezoelectric polymer called polyvinylidene fluoride (PVDF) [23]. The passive component is most effective at higher frequencies, while the active component is most effective at lower frequencies. Active control is facilitated at lower frequencies because the signal processing speed demand is less which requires fewer actuators and sensors [22]. Active control techniques can ameliorate sound reduction of passive sound-absorbing materials at low frequencies [10]. A smart structure consists of four main components: structural material, distributed actuators and sensors,

control strategies, and power conditioning electronics [42]. A special feature of smart materials is their capability to change their mechanical properties (viscosity, stiffness, shape) based on changes in temperature or an induced field (electric or magnetic). A distinct physical characteristic of the PVDF actuator is that it is engineered to be curved in order to couple the predominantly in-plain strain associated with the piezoelectric effect and the vertical motion needed to eventually radiate sound away from the foam's surface [23]. The piezoelectric element translates an electrical input signal into a motion that radiates sound out of the surface of the smart foam element [22].

The passive component is a sound-absorbing material called partially-reticulated polyurethane foam. Guigou and Fuller describe the passive foam as being able to dissipate incident acoustic wave energy through friction associated with the coupling of the liquid and solid phase of the foam itself [23]. It is an acoustic grade, open cell, flexible ester based urethane foam designed to provide maximum sound absorption per given thickness. For the active component in Guigou and Fuller's research, a 28 µm Ag (silver-electroded) metallized PVDF film was employed and embedded in the foam [23]. This was chosen because it has an excellent tolerance of high voltage amplitudes required for actuator applications.

Guigou and Fuller's research looks at an active-passive foam-PVDF smart skin which was mounted in the cockpit of the mid-size business jet Cessna Citation III located at VPI&SU [23]. The idea behind the active-passive surface coating is to cover the inner surface of the aircraft skin with adjacent independent active tiles. The fuselage crown panels are excited with a single speaker located external to the cockpit. The speaker is driven by a band-limited random excitation. A multiple-input-multiple-output (MIMO) feedforward Filtered-x LMS controller is used for error minimization of the error sensor signals provided by close-by microphones. Three different reference signals are used by the feedforward controller and are compared (compensated) to the

achievable interior noise attenuation. The main objective of the smart foam-PVDF skin is to reduce interior cabin noise. As can be seen in Figure 2.1, each smart skin element is composed of three active cells driven in phase, and then mounted on a fuselage panel. A similar LMS algorithm (feedforward) is applied to the smart foam application developed by Gentry et al. in to minimize the signal from a far-field error microphone [20], [45].
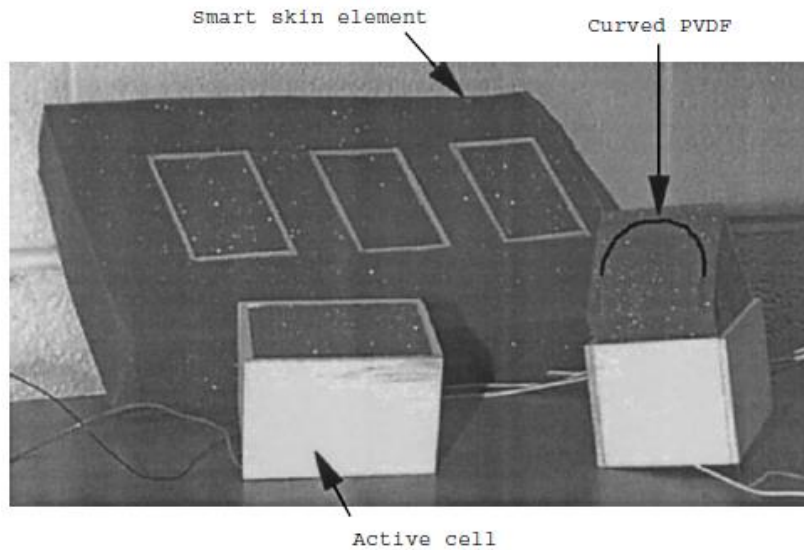
Fig. 2.1 Smart foam set-up used in [1].

The general physicality of the control setup is four crown panels seen below in Figure 2.2. Each treated crown panel is associated with its own control channel (labelled in Figure 2.2 as channels C1 through C4). In addition, four error microphones were located at pilot's ear level which was approximately 7 inches below the cockpit fuselage ribs. In order to simulate the excitation of flow separation and associated turbulence, band-limited and broadband random excitation of the crown panels was achieved by the use of a speaker mounted on the cockpit exterior. The disturbance speaker was located at the centre of treated area (1 inch from fuselage structure above top ribs). A traverse (-90° to 90°) consisting of 12 microphones (1-5 inches apart)

11

was used to assess control performance of planes located at pilot's ear and shoulder level (7 and 14 inches below cockpit ceiling).
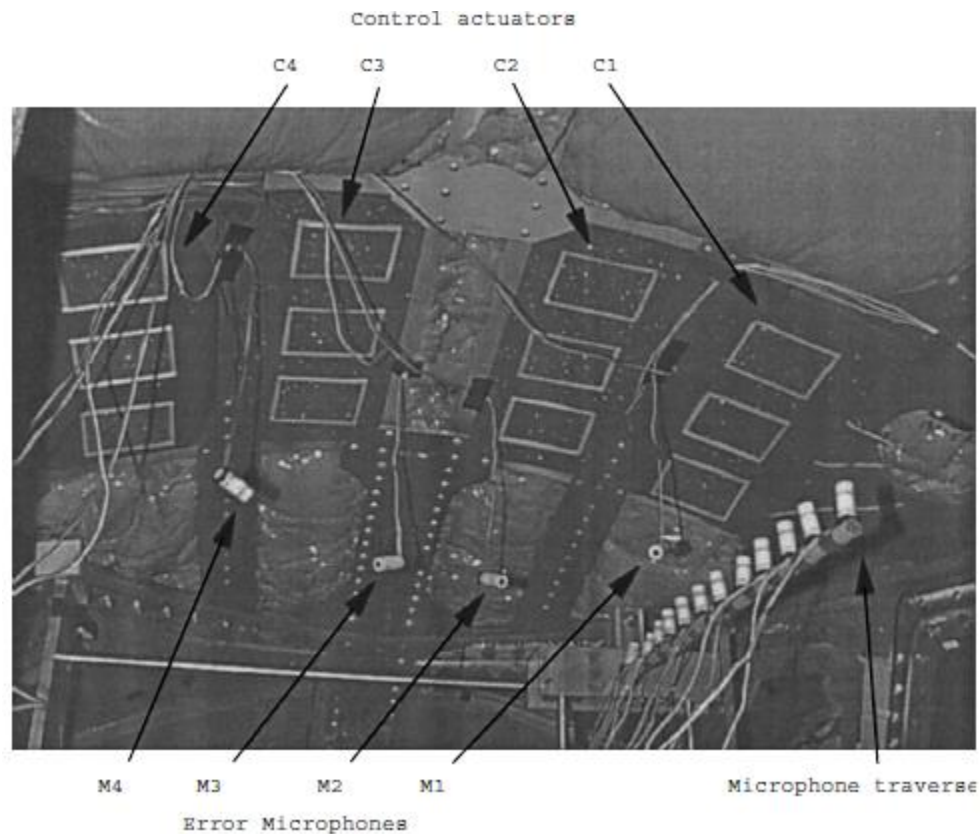
Fig. 2.2 Cessna crown panels control setup [1].

The experiment was able to produce 108 SPLs in each of the aforementioned measurement planes. The global attenuation (global sound level) was determined as being the ratio of the sum of the 108 measured square pressures summed over the excitation frequency range, before and after control. The software control consisted of a MIMO feedforward Filtered-x LMS algorithm which was driven by a TMS320 digital signal processor (DSP). This algorithm was used to minimize the acoustic pressure at the error microphones. The controller had a sampling frequency of 2-8 kHz and 90 and 100 coefficient finite impulse response (FIR) filters were used for error and control paths respectively. All control results are presented in Guigou and Fuller's paper [23].
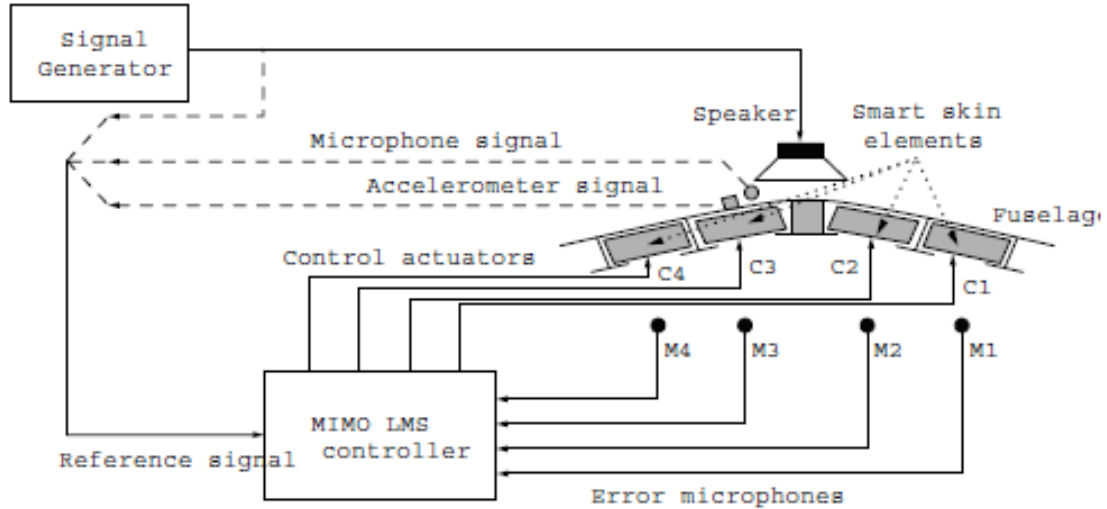
Fig. 2.3 Schematic of control setup in [1].

The smart skin is composed of a cylindrically curved PVDF piezoelectric film embedded in partially reticulated polyurethane acoustic foam. For the research methods presented above, the foam-PVDF smart skin was mounted under the crown panels in the cockpit of a Cessna Citation III fuselage. The main objective was to conduct performance testing as each smart foam element (4 in total) was used to control the effective acoustic source of an individual fuselage panel. It was found that increasing the number of control channels (from 2 to 4: depicted as M1 and M4 and M1-M4 respectively in Figure 2.3) led to an increase in global attenuation. This is equivalent to extending the smart skin to control a larger number of fuselage panels.

The best means of providing control performance was using a microphone located close to the fuselage as a reference signal. For band-limited excitation (200 Hz bandwidth), up to 13 dB global passive/active attenuated was attained at the pilot's ear level. When the excitation frequency bandwidth is increased to 800 Hz, the global passive/active attenuation is approximately 7 dB. PVDF and its copolymers can also induce electrostrictive (a dielectric material property caused by a slight displacement of ions in the crystal lattice upon being exposed to an external electric field)

strains of 2% when applying large alternating-current electric fields (200 kV/mm) [42]. Since PVDF can provide extremely high active strains with high operation frequencies, this type of electroactive polymer (EAP) could be used for several passive-active noise and vibration reduction tasks [42].

The film PVDF is extremely lightweight and flexible when compared to other piezoelectric materials. Its flexibility permits it to be shaped into any form thus customized for each specific application. In this case, this feature led to the determination of the optimum PVDF contour to maximize its far-field radiation efficiency. Original PVDF actuator configurations were first suggested by Tibbets [64]. A major advantage in terms of convenience and cost of PVDF is that it can be bonded or glued using commercial adhesives. Lastly, the use of PVDF as the active smart foam component is its similar acoustic impedance to that of air. This is key because the closer the acoustic impedance of an actuator to that of the medium on which it operates (i.e. air, water) the more efficient the source sound radiation. Consequently, the actuator will require less control energy to operate. The nonlinear behaviour of smart foam is assessed using a measure known as harmonic distortion. It is a measure of the pressure amplitude distortion and refers to the deviation from correspondence between the acoustic output wave and the electrical input wave that is caused by nonlinear effects in the smart foam [7]. Figure 2.4 below is a property comparison table of PVDF versus a conventional piezoelectric ceramic material. [22]

Property	Units	PVDF Film	PZT
Density	(10^3) kg/m^3	1.78	7.5
Relative Permittivity	$\varepsilon/\varepsilon_0$	12	1200
d_{31} Constant	(10^{-12}) C/N	23	110
g_{31} Constant	(10^{-3}) V·m/N	216	10
k_{31} Constant	% at 1 kHz	12	30
Acoustic Impedance	(10^6) kg/m^2·sec	2.7	30

Fig. 2.4 Comparison of piezoelectric material properties [22].

14

2.3. Modeling and simulation of near-field sound propagation

For near-field measurements, the measurements are made close to the noise-making object (usually at 1 m) on the surface of a similar shape to the object and corrections made for the effect of room reverberation and the near field. One suitable measurement locations (generally between 5 and 16) are determined, the measured average sound pressure level value L_p around the object is used to then determine the sound power level of the object using the equation from [25], where L_w is the sound power level of the object, S is the area of the test surface and Δ_1 and Δ_2 are correction terms. In the near field of an object, sound propagation is not always normal to the randomly chosen measurement surface. It can be seen that the equation suggested by Hansen implicitly assumes propagation normal to the measurement surface [25]. This is why the correction factor Δ_2 is introduced to compensate for possible tangential (non-normal) sound propagation. Values of Δ_2 are given in Table 3.3 of Hansen's work as a function of the ratio of the area of the measurement surface S divided by the area of the smallest parallelepiped, S_m, which snugly encloses the source [25]. On the other hand, the correction factor Δ_1 accounts for the absorption characteristics of the test room and can be determined from measurements on two test surfaces around the object. If measurements on two test surfaces are used, then Δ_1 is given by another equation suggested by Hansen [25]. Where the subscripts 1 and 2 refer respectively to the measurement surfaces near and remote from the object.

Source directivity has a major impact on sound propagation and the acoustics of its immediate environments [67]. The research paper by Mehra et al. presents a method to model dynamic, data-driven source and listener directivity for interactive wave-based sound propagation in virtual environments [39]. The experiment consisted of several spherical harmonic (SH) decompositions

of varying source directivity interactively and computations of the total sound field at the listener position as a weighted sum of precomputed SH sound fields. The general approach here was to, given a scene and a source position, precompute a set of pressure fields due to elementary SH sources using a frequency-domain wave-based sound propagation technique.

These pressure fields are then encoded in basis functions (e.g. multipoles) and stored for runtime use. With available runtime dynamic source directivity, a SH decomposition of the directivity is performed in order to obtain the corresponding SH coefficients. Next, the final pressure field is computed by combining (summation) the pressure fields due to SH sources determined at the listener positions weighted by the appropriate and previously-determined SH coefficients. Acoustic responses are processed for both ears at runtime by using a plane-wave decomposition and the head-related transfer function (HRTF)-based listener directivity. Lastly, to compute spatial sound, a dot product of SH coefficients of the HRTF and the plane wave decomposition of pressure field is performed. The general process of the complete approach is demonstrated in Figure 2.5 below.
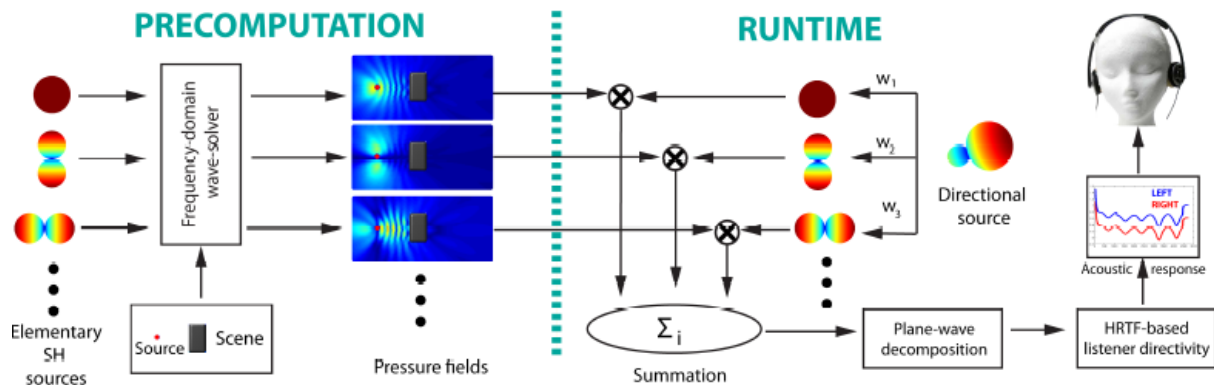
Fig. 2.5 Overview of source and directivity approach [39].

16

A HRTF describes the effect of the listener's outer ear, head and body on incoming omnidirectional sound. Mehra et al. [39] and Ravni [5] performed detailed studies using several metrics including inter-aural time difference (ITD), inter-aural level difference (ILD) and inter-aural time cross-correlation coefficient (IACC). These results demonstrated that a SH order of 1-2 is insufficient for spatial perception of frequencies up to 1 kHz; a SH order of 3 is good for frequencies up to 2 kHz; and a SH order of 3-6 is good for frequencies up to 8 kHz. The concept is that higher SH orders result in better spatial resolution, but computation of their derivatives of pressure field for plane-wave decomposition is computationally expensive. This means that an appropriate trade-off between the SH order and performance-accuracy must be carefully considered. The general consensus from Mehra et al.'s work is that sound directivity becomes more prominent with increasing frequency [39]. Consequently, this requires higher-order SH basis functions. Future work consisted of the exploration of other basis functions such as wavelets to handle sharper directivities. The source formulation can handle both near- and far-field sound radiation by directional sources. Near-field directivity requires a set of dense measurements of complex frequency responses very close to the source at twice the Nyquist rate (which was not available at the time the paper was published). A final area for further exploration entails hybridization of wave-based techniques with geometric approaches to handle directional sources over the entire human-hearing frequency range.

Wave theory looks at the details associated with the cross-sectional variations in the sound pressure field. With diffraction, Kang tells us that two cases must be considered: highly absorbent materials wherein which all the energy of a plane wave is absorbed effectively after only two reflections; or relatively hard boundary materials wherein which the absolute sound field consists of a small perturbation of the rigid boundary case [31]. In the past, Mehra et al. have developed a

novel algorithm that accurately solves the wave equation for dynamics sources and listeners using a combination of precomputation methods and graphics processing unit (GPU-based) runtime evaluation [38]. Wave-based techniques, in general, can accurately perform sound propagation at all frequencies and can model all acoustic effects (including wave effects) [65, 57]. These techniques numerically solve the acoustic wave equation [52]. The research by Mehra et al. presented WAVE (Wave-based Acoustics for Virtual Environments), an interactive wave-based sound propagation system for efficiently generating accurate and realistic sound for virtual-reality (VR) applications [38]. Wave solvers are generally classified as frequency-domain and time-domain methods. In frequency-domain approaches, a popular method is the finite element method (FEM) [63] and the boundary element (BEM) [14]. Of the time-domain approaches, the most-used method is the finite difference time domain (FDTD) method [62].

The Helmholtz equation is used to study sound wave propagation in the frequency domain. It does this by expressing sound wave propagation as a boundary value problem [39]. Where $p(x)$ is the complex-valued pressure field (sound field) at frequency v, $w = 2\pi v$ is the angular frequency, Ω is the propagation domain, and ∇^2 is the Laplacian operator. This equation can be solved using any frequency-domain, wave-based propagation technique including BEM, FEM, or equivalent source method (ESM). The linearity of the Helmholtz equation implies that the pressure field of a linear combination of sources is a linear combination of their respective pressure fields [52]. The boundary element method [35] is a numerical method used to solve the 3D Helmholtz equation which accurately models sound propagation in indoor and outdoor spaces. This technique converts the Helmholtz equation into the boundary integral equation, then solves for pressure and velocity on the boundary; thus the pressure at any point in the domain is resolved. Mehra et al.'s data is magnitude-only, averaged over the frequencies in each octave band, and for all octave bands within

18

the frequency range of the sound sources [39]. The SH order that was used is $L = 3 - 4$ for the source representation. This resulted in errors less than 10-15%, which is reasonable and widely-acceptable error threshold for auralization purposes in interactive applications [28, 40, 57]. The pressure fields produced by the two techniques used by Mehra et al. [39] agree within the error of $< 5\ to\ 10\%$.

2.4. Modeling and measurement of cabin noise distribution

A primary source of aircraft interior cabin noise is propeller noise and can be characterized by discrete tones at the fundamental blade passage frequency (BPF) of the engines and their subsequent harmonics [22]. In the business jet application addressed in the CREDO project, the primary noise source is the stochastic structural TBL excitation, which creates also a stochastic (diffuse) background sound field [24].

One common area of noise research in aircraft and helicopter cabins is cabin panels. There are two perspectives of noise in an aircraft: radiated and absorbed intensity components on cabin panels [24, 44]. The research conducted by Hald et al. [24] presents two methods; one of measuring the surface absorption coefficient and the other of measuring the surface admittance. Both of these methods use what is called a dual layer array (DLA). The DLA is placed in positions close to the surfaces under investigation to sample the near-field sound pressure [44]. The theory presented by Hald et al. is that a surface segment such an aircraft cabin panel can radiate sound energy due to external forcing [24]. This in turn can cause the surface to vibrate and absorb sound energy from an incident sound field because of its finite surface acoustic impedance. In order to measure the total sound intensity I_{tot}, the radiated sound intensity I_{rad} that would exist with no incident field and the sound intensity I_{abs} due to absorption are simply summated.

Based on directionality, the radiated intensity I_{rad} is generally a positive value, while the absorption component I_{abs} is negative. This is typically why the total sound intensity is sometimes smaller in comparison to the radiated intensity (which can be high). From a noise perspective, the most useful factor is usually the total sound intensity, but the aforementioned components are important for energy-based modelling of an environment where the flow of energy from one subsystem to another becomes essential. These intensity components may be needed for calibration of said models of perhaps vehicle interior noise [44]. Figure 2.6 depicts a setup of three incoherently excited speakers to create an incident field similar to the field incident under for example flight conditions in an aircraft. The DLA measurement can provide the total and incident sound field components on the panel surface. The surface admittance can then be found as it is the ration between the normal velocity and the pressure of the total field on the surface. The principal components are averaged, and since (in this case) the absorbed intensity is equal to the total intensity, the absorption coefficient is calculated as the ration between the total and the incident intensities.
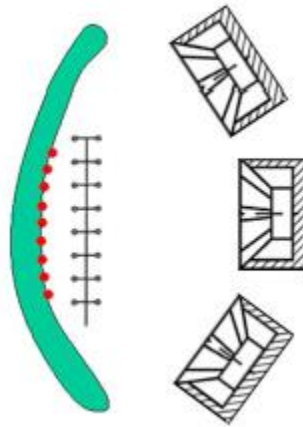
Fig. 2.6 *Left*: Cabin panel with surface calculation points covered by array. *Right*: Speakers [24].

In the case of aircraft cabin applications, the surface property measurement is conducted on the ground, while the operational measurement is obviously conducted during flight. Both measurements are required for a fair and accurate set of calculation points on the panels. The energy method error is shown in Figure 2.7 for the cases of 2, 5, 16, and 144 plane waves used in the measurement of absorption coefficient. It is intuitive that 2 plane waves are not enough, while 5 waves seem reasonable. For the 5-wave setup, one wave was situated at close to normal incidence and 4 distributed around, far off-axis. Again from Figure 2.7, it can be seen that the admittance method errors are larger at low frequencies (dotted black line). The five loudspeakers are also used to create the masking incident field that has to be suppressed in order to obtain the radiating intensity.
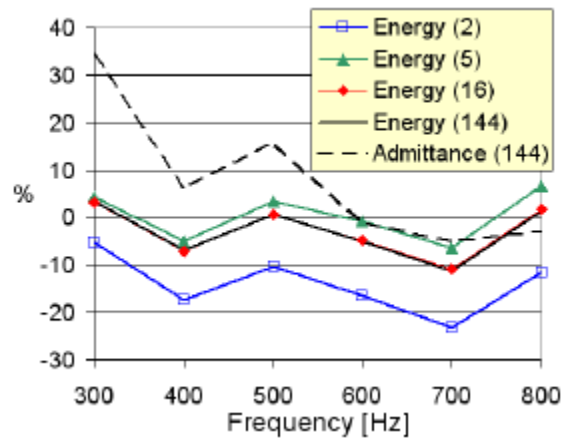
Fig. 2.7 Sound power error vs. number of plane waves for measuring surface properties [24].

The test area is separated into two sections, as seen in Figure 2.8: the trim panel and the window itself. It should be noted that the test room was a reverberant room at ambient temperature of 20°C and the spectrum of excitation was not representative of true flight conditions. This means that the data presented here was not an accurate presentation of the reality.
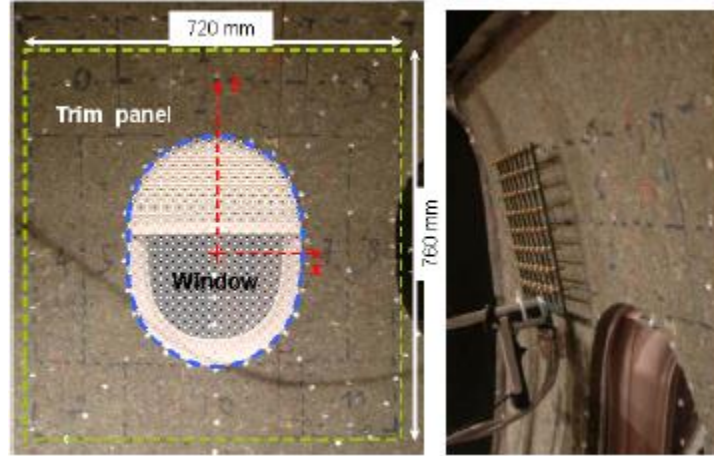
Fig. 2.8 *Left:* Measured area (panel) around left window. *Right:* DLA during measurement [24].

There were three configurations that were considered during this test:

1. Only TL noise: Reference free-field measurement

2. Only loudspeakers: Allows for measurement of surface absorption coefficient and admittance across the area.

3. TL noise and loudspeakers: In order to obtain radiated intensity in presence of background noise.

The DLA method and a normal intensity probe were used for sound intensity measurement of each configuration. Also, swept measurement of average intensity was performed with a two-microphone sound intensity probe over 12 small areas roughly 5 cm from the panel: 11 on trim panel and 1 covering the window. The specific methodology consisted of a set of overlapping array positions which were measured across the test area at 1-2 cm distance. This is done in order to cover the investigated surfaces patch by patch [44]. A DLA created by Morkholt et al. [44] is depicted in Figure 2.9. The 2x8x8 element DLA was mounted on a xy-robot, but its position and orientation was measured by an InterSense IS-900 system integrated into the handle of the array.

In the research conducted by Morkholt et al. [44], the method is described in more detail but the general process is explained here. The first step is to calculate the full cross spectral matrix (CSM) for all signals using Fast Fourier Transform (FFT). The frequency responses are corrected using the correction data from the individual microphones. Next, a principal decomposition (PCS) is conducted to determine the most incoherent components. Then, Statistically Optimized Near-field Acoustical Holography (SONAH) was used to perform calculations across a mesh at the panel surface. The results for total intensities as well as estimations of the absorption coefficient and radiated intensities can be found in Hald et al.'s work [24]. The results for the measured DLA versus classic probe techniques are typically within 1 dB.
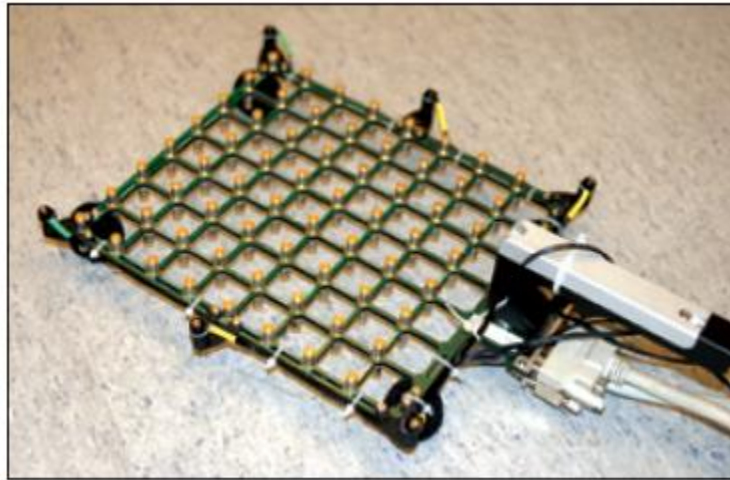
Fig. 2.9 Double-layer array (DLA) mounted with six IR LEDS [44].

2.5. Passive noise isolation and sound absorbent materials

Porous materials have many applications, one of which is sound absorption. Porous materials come in many forms; they can be fibrous, cellular or granular. Fibrous materials can be in the form of mats, board or preformed elements manufactured of glass, mineral or organic fibers including felts and felted textiles. Interestingly, cellular materials include polymer foams of varying degrees

of rigidity. The high weight-specific stiffness, good crash-energy absorption ability, and fire-resistance of porous metals makes them suitable for sound absorption panels in aircraft and automatic industries. Some examples of porous materials used for sound absorption include polyurethane foam, horse hair, glass wool, lead shot (3.8 mm), porous aluminum, gravel (7.5 mm), and porous concrete. [34]

In the case of long enclosures such as aircraft cabins, underground stations, and corridors, one dimension (length) is much greater than others. According to Kang, the most important factors of acoustic design of long enclosures are cross-sectional size and form, and absorption amount and arrangement. For low absorbent boundaries, as proved by Kang, increasing the cross-sectional size or area reduces the attenuation [31]. This is a result of the reduced number of reflections in a period of time for a larger section. This also means that the total energy in the whole section is less, although the attenuation of energy per unit area is greater. This can be expressed as a reduction in relative attenuation and an increase in the absolute attenuation with respect to PWL. A considerably practical method of predicting acoustic indices in long enclosures is physical scale modelling. The only limitation here is cost. However, loudspeaker directionality and other factors can be challenging to implement in the physical scale models. The model that will be reviewed consists of a basic geometry in order to roughly determine the acoustic indices of a single source using models stemming from existing theory or computer simulations. The general idea is that with the data from the single source, the acoustic indices of several other sources can be obtained simply by superimposing each source.

The inputs for the computer program called acoustics in long enclosures (ACLE) that was developed by Kang et al. [30] include geometry (dimensions, major obstructions), boundary conditions (absorption coefficient, absorber arrangements, diffuse situation, etc.), source

24

conditions (temporal and spatial distribution of sources, source directionality). Additional inputs include ambient noise for a PA system (input can be the temporal and spatial distribution of noise), proposed language(s) for said PA system, and unit costs of acoustic treatments [30]. This particular paper presents research pertaining to public address (PA) systems in underground stations. The sources can be noise sources or loudspeakers as well as directional or omni-directional. The systematic diagram below depicts the flow of calculations for determination of reverberation (prolongation of sound) and SPL. The perceived reverberation at the receiver can be found from a decay curve or an energy response. Also, if the decay curve is almost linear, the reverberation can be found from the early decay time (EDT) or reverberation time (RT).

In the complicated cases, the aforementioned simplifications become inaccurate. For instance, the presence of diffusers along boundaries produce more sound attenuation and are known to decrease reverberation. Also, in reality, absorbers are strategically placed along the length of the enclosures and they too behave similarly to diffusers. The absorbers are generally placed evenly in order to obtain a higher attenuation. Corrections related to this situation need to be taken into consideration in ACLE. Kang et al. found that a ribbed diffuser and a Schroeder diffuser are appropriate for reducing sound attenuation for the model with a circular cross-section [30]. This means that corrections for the variations in diffuser-effectiveness with respect to absorption conditions in a long enclosure have been taken into account. The extra attenuation caused by Schroeder diffusers is higher. This is important because this type of diffuser has a lower absorption coefficient than model absorbers (possess high absorption coefficients), but it still manages to produce more sound energy loss. In comparison to a ribbed diffuser, the Schroeder diffuser causes significantly extra attenuation. The consensus is that the Schroeder diffuser has an optimal diffusing and attenuating characteristic as well as extra absorption.

25
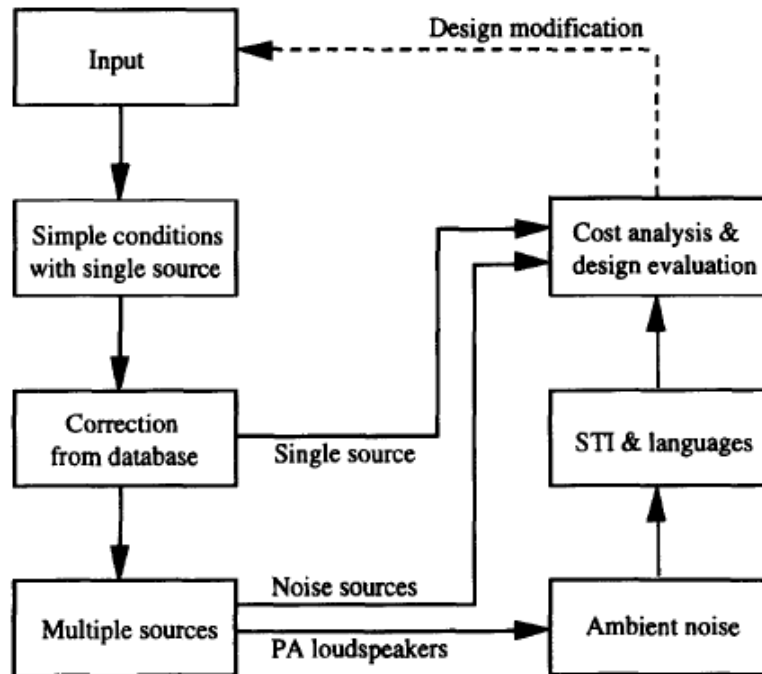
Fig. 2.10 Diagram of ACLE (method for predicting acoustic indices in long enclosures) [30].

The ACLE program also considers train noise and ventilation noise. The train noise is modelled as separate sound sources from the train model's various sections. The prediction subprogram is called train noise in stations (TNS) and is used to predict the temporal and spatial distribution of train noise in underground stations. Below is plan and cross-section of a real underground train station. It should be noted that the sound characteristics of the long enclosure determine the ideal reverberation and SPL locations and thus the receiver spacing.
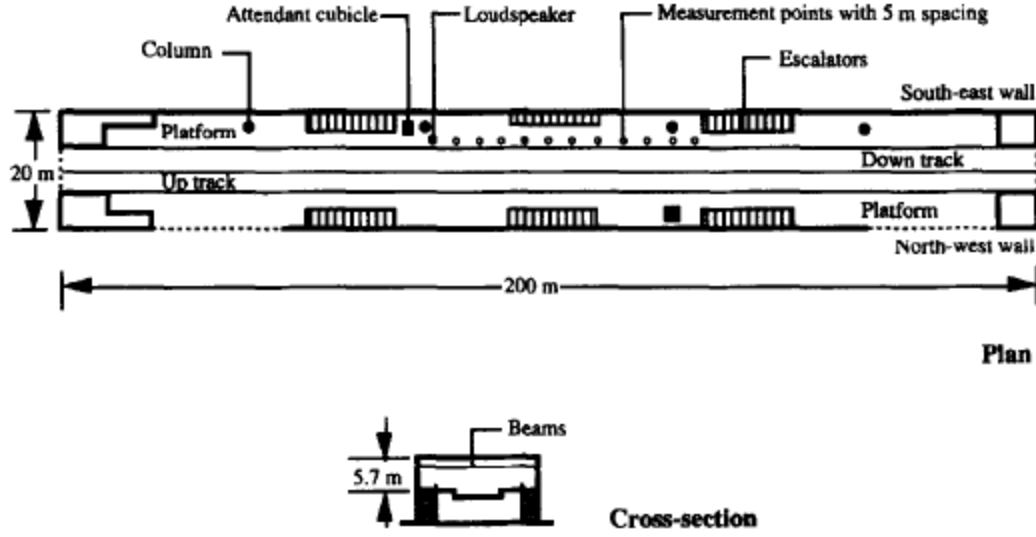
Fig. 2.11 Cross-section of Hong Kong MTR KWF station showing measurement arrangements [30].

The medium used by Hansen [25] explains the basic principles of sound propagation in porous media is characterized in dimensionless variables by a complex density, ρ_m, and complex compressibility, κ. These quantities are then used to define a complex impedance and propagation constant. The characteristic impedance of the porous material is expressed using the gas density, ρ, the gas speed of sound c, the complex density ρ_m, and complex compressibility, κ as described by Hansen [25], where $\omega = 2\pi f$ is the angular frequency in radians per second (rad/s) of the sound wave. The quantities ρ_m and κ can be determined using the procedure outlined in [8]. This particular method for fibrous porous materials yields results that are within 4% of the mean of published data. It also leads to the correct limits at both high and low values of the dimensionless frequency $\rho f / R_1$ which are thoroughly defined by Bies [8].

2.6. Modeling human hearing range: frequency and amplitude

The human auditory receives large directional cues from the minor differences in the sound received by the left and the right ear. These differences can be attributed to the scattering of sound around the human head and they are directly affected by the direction of incoming sound [38]. The minimum, healthy acoustic pressure audible to the young human ear is approximately 20×10^{-6} Pa, or 2×10^{-10} atmospheres. The minimum audible level occurs at approximately 4 kHz and is a physical limit. This means that lower sound pressure levels would be audibly eliminated by thermal noise due to molecular motion in air. It becomes quite painful for the normal human ear to experience sound pressures of the order of 60 Pa or 6×10^{-4} atmospheres.

The A-weighted decibel scale is a sound level metric used to measure noise for the human ear which is typically more sensitive to sound at higher frequencies (especially between 1 and 4kHz). The A-weighted Equivalent Continuous Noise Level is achieved by A-weighting the noise (applying inverse of 40 dB curve equal-loudness curve for human ear at 1 kHz) and then averaging the sound pressure squared over a period of time. This kind of averaging is known as energy averaging. This noise level is used as a descriptor of both occupational and environmental noise and for an average over time, T [25]. During some cases, the sound pressure is averaged solely (no A-weighting) and it is known as the Equivalent Noise Level, $L_{eq,T}$. Hearing loss is usually determined using pure tone audiometry (measurement of the range and sensitivity of a person's sense of hearing) in the frequency range from about 100 Hz to 8000 Hz. Essentially, hearing loss is considered to be the differences in sound pressure levels of a series of tones that are barely audible compared with reference sound pressure levels for the same series of tones. The two main causes of hearing loss are exposure to excessive noise and aging. Losses caused by excessive noise occur first in the frequency range from approximately 4000 Hz to 6000 Hz. This particular range

is known to be the range of greatest sensitivity of the human ear. At first, the loss occurs in the aforementioned range. However, according to Hansen [25], as the deterioration of hearing sensitivity increases, the maximum loss generally remains near 4000 Hz. Sufficient speech recognition usually requires a frequency range from 500 Hz to 2000 Hz. This is important because hearing loss pertains directly to this range. A hearing loss of 25 dB will allow speech to be barely understood, while a loss of 92 dB is regarded as total hearing loss. Secondly, if a person undergoes a hearing loss between 25 dB and 92 dB, that person's hearing is said to be impaired. The degree of impairment is calculated as a percentage at the rate of 1.5 percentage points for each decibel loss greater than 25 dB.

The IEEE technical paper by Serizel et al. [59] presents integrated active noise control and noise reduction techniques for hearing aids for the purpose of eliminating secondary path effects of noise leakage through an open fitting. The use of a noise-reduction (NR) algorithm and an active noise control (ANC) system in a cascade form may be efficient as long as the so-called causality margin of said system is large enough. When the two subsystems are positioned in parallel and then integrated, this has found to be a more robust algorithm. In terms of active control, a Filtered-x Multichannel Wiener Filter (MWF) is presented and applied to provide the ANC and integrate it with the NR. The paper also experimentally compares the cascaded scheme and the integrated scheme through the use of the aforementioned filter in a classic noise reduction framework (with active noise control). Here, the integrated scheme was found to provide the best performance. Since many hearing aids have open fittings, the noise leakage through the fittings is a major concern. One accurate way of eliminating the undesired noise leakage is to use active noise control [16, 32]. The Filtered-x version of the MWF algorithm (FxMWF) in both the cascaded and integrated cases has been investigated by Serizel et al. [60]. In a hearing aid, the location of the

noise cancellation is at the tympanic membrane, and in any ANC system, the secondary path is crucial for the algorithm. Since the introduction of said path can create several instabilities, it is suitable to use filtered-x algorithms [9, 11 and 69] based on the estimate of the secondary path (derivation produced by Serizel et al. [59]).

The experimental setup for all three algorithms (cascade, parallel, integrated) is described in great detail by Serizel et al. [59]. The simulations contained a mannequin head and torso with artificial ears and a two-microphone behind-the-ear (BTW) hearing aid. At the time of the research, hearing aids did not have ear canal microphones so the artificial ear eardrum microphone was used to generate the error signal. The tests were run on 22-s-long signals. The speech consisted of three sentences from the HINT database in combination with silence periods. The noise was a 16 kHz multitalker babble from Auditec [47].
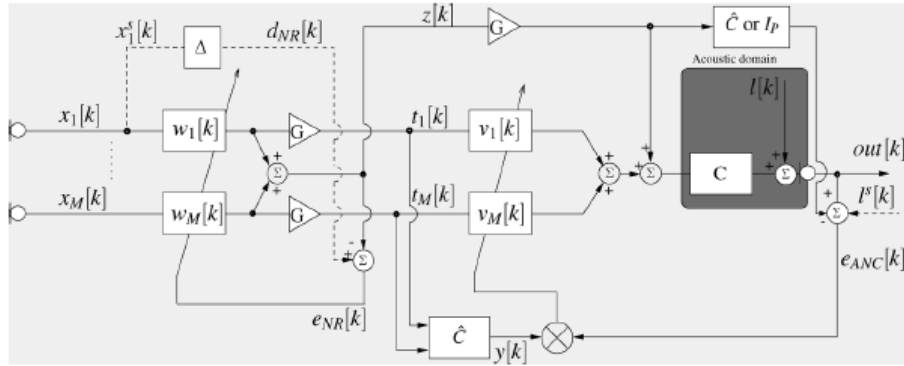
Fig. 2.12 Multichannel noise reduction and active noise control systems in cascade [59].
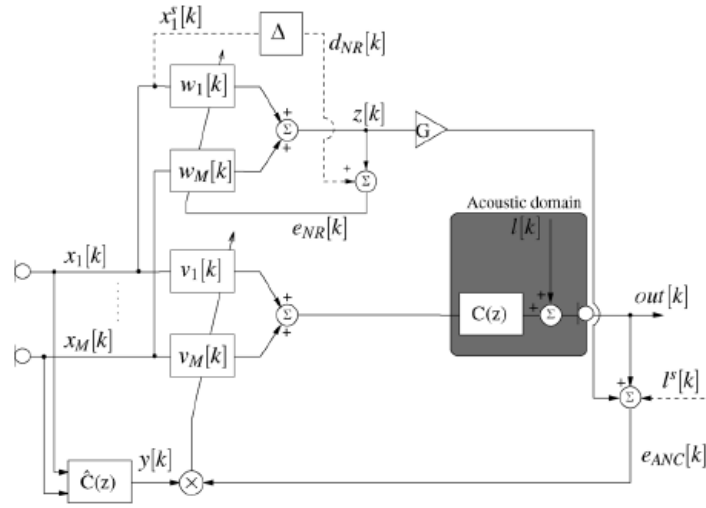
Fig. 2.13 Active noise control and noise reduction system in parallel [59].
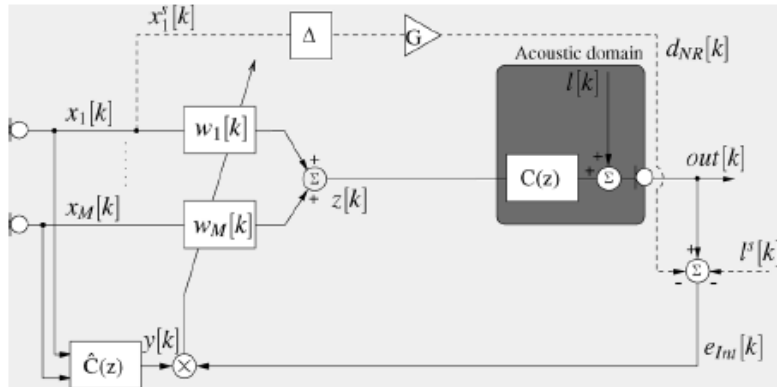
Fig. 2.14 Integrated multichannel active noise control and noise reduction system [59].
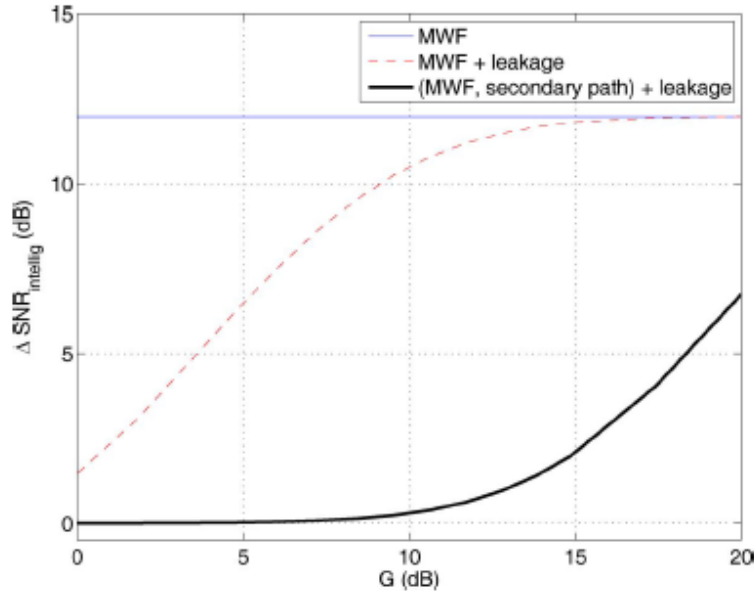
Fig. 2.15 Noise reduction performance comparison: with vs. without active noise control [59].

It has been concluded by Serizel et al. [59] that open fittings in hearing aids cannot be ignored and can negatively affect the NR performance. In response, ANC can be used to reduce the impact of the noise leakage. ANC has been found to provide signal to noise (SNR) improvements between 4 dB and 12 dB (varies with approach used). The general consensus is that a cascaded approach is not a realistic one because hearing aids have latency margins close to zero (cascaded technique uses high latency factor). For the integrated approach, the SNR improvement range is approximately 12 dB for low hearing gains (between 0 dB and 20 dB). This is the case only when the system is causal. However, when the system becomes non-causal, the integrated approach still has the upper-hand as it takes the secondary path into consideration in the speech enhancement and amplification process. This reduces the impact of the leakage on the output signal. Future work consisted of developing hearing aids with ear canal microphones. Consequently, the addition of said microphone might create concerns related to bone conduction to the ear canal microphone during user speech.

2.7. Visual Object Detection

The second and third contributions of the paper by Viola and Jones [68] represent two similar learning algorithms. The first of which selects a small number of critical visual features from a larger set and produces highly efficient classifiers. The latter is a method for combining increasingly more complex classifiers in a cascade. The idea behind the cascade is to efficiently and selectively remove background regions of an image in order to allocate more time to compute reliable object-like regions. The simple classifier method uses a technique called AdaBoost to successfully discard a large number of Harr-like (digital image) features and concentrate on a small set of critical ones. The method for feature selection constrains a weak learner in order for the weak classifier to result in a single feature. This process is reiterated such that every new classifier inherently results in a feature selection. Adaboost enables this strong learning algorithm and puts hard bounds on the overall system performance of object detection by focusing on promising regions of the image. From a conceptual stand-of-point, it is possible to effectively determine where an object is inside of an image [68]. An important factor of measurement called the "false negative" determines the rate of the attentional or focusing process. In the field of face detection, it is feasible to obtain less than 1% false negatives and 40% false positives through the use of a classifier created from two Harr-like features. This filter results in a one half reduction of the number of locations where the final detector should be examined. The weak learning algorithm is constructed to choose only one rectangle which best divides the positive and negative examples. Essentially, the weak learner's criterion yields the misclassification of the minimum number of examples via the optimal threshold classification function. A weak classifier $h_j(x)$ contains a feature f_j, a threshold θ_j and a parity p_j which is the direction of the inequality sign. In this case, x represents a 24x24 pixel sub-window of an image. Since this is an increasingly complex process

with the addition of more features, the classification task yields progressively higher errors. For example, features selected in later rounds or iterations result in error rates between 0.4 and 0.5 while features selected early on yield rates as low as between 0.1 and 0.3.

The learning results yield a detection rate of 95% with a false positive rate of 1 in 14084 using a frontal face classifier built from 200 features. However, it is the case that the most intuitive method for bettering detection performance is also the most computationally expensive because of the continual addition of features. When the Adaboost technique uses rectangles, the initial rectangles used for the selection process are usually more useful and easier to analyze. More specifically, the first selected feature concentrates on the region of the eyes and their relative darkness with respect to the region of the nose and cheeks (see Figure 2.16 below). This feature is selected because it is a large region with respect to the detection sub-window. In addition, this feature is relatively impartial to the face's size and location, which makes it easier to determine. The second feature selected concentrates on the region of the eyes and their relative darkness with respect to the region of the bridge of the nose.
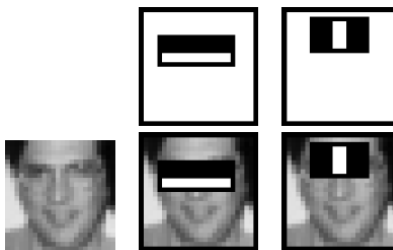
Fig. 2.16 Adaboost features. *Top row:* Two features; *Bottom row:* Same features on training face [68].

Viola and Jones' [68] results for the number of features in the first five layers of the detector is 1, 10, 25, 25, and 50 features respectively. As the layers continue from here, the number of features within each also grows. The total number of features in all the layers is 6061. The training process

34

contained 4916 training faces for each classifier in the cascade including their corresponding vertical mirror images. This resulted in a total of 9,832 training faces and 10,000 non-face sub-windows per layer via the Adaboost training process.

2.8. Image Registration

Image registration has many applications pertaining to computer vision such as image matching for stereo vision, pattern recognition, and motion analysis. The IJCAI paper by Lucas and Kanade [36] presents an image registration method that uses the spatial-intensity gradient of images to find a match (detection) via a Newton-Raphson type iteration. The particular technique suggested by Lucas and Kanade [36] can handle rotation or similarly-related distortions of the images, making it extremely powerful. The technique involves random linear distortions of the image as well as rotation to make use of a stereo vision system which creates a greater understanding of stereo imaging.

In order to solve the translational image registration problem, two functions $F(x)$ and $G(x)$ are used. These functions determine the respective pixel values at each location x in two consecutive images. In this case, x is simply a vector. The disparity vector h minimizes the difference between the functions of $F(x + h)$ and $G(x)$, for x in some region of interest R (see Figure 2.17 below). Typical measures of the difference between $F(x + h)$ and $G(x)$ are L1 norm, L2 norm, and the negative of the normalized correlation which can be determined from Lucas and Kanade's work [36]. From these three measures, the L1 norm is commonly used because it is a relatively inexpensive approximation of the L2 norm.
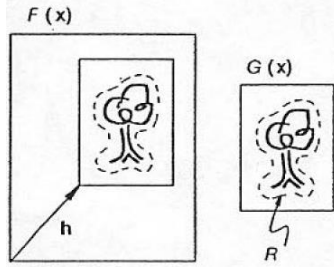
Fig. 2.17 The image registration problem [36].

Normally, the image registering technique for two images finds the difference between the images at all exhaustive values of the disparity vector *h*. This is a tedious process and the computation time is $O(M^2N^2)$, where the size of the picture is $N \times N$, and the region of possible *h* values is $M \times M$. The algorithm determines the order of searching the space of possible *h* values [36]. This is done by using an initial estimate of *h* and the spatial intensity gradient at each point of the image to update the current estimate of *h*. In this way, each current *h* estimate is matched or replaced by a better one. This process is repeated in a Newton-Raphson-like iteration approach. If and when the process converges, it will usually do so in $O(M^2 log\ N)$ steps. The generalized registration algorithm can be used to determine depth information from stereo images.

Depth information can be found from a stereo pair can be classified into four main areas: finding objects in the pictures, matching the objects in the two views, finding camera parameters, and calculating the distance between the camera and the objects [36]. In order to use the registration algorithm described above, the stereo vision systems must work at the pixel level since the technique uses pixel-level matching. Also, another important consideration is that several stereo vision systems are not concerned with the relative positions of the cameras which are never known to a high degree of accuracy. This means that the camera parameters in addition to the distances of the objects to the camera should be considered for any real applications.

2.9. Dummy Head for Room Acoustic Measurements

The concept of using dummy or mannequin heads for room measurements of sound fields has been developed extensively over the years. The research conducted by Norcross et al. [48] looks at the errors produced using a dummy head for room acoustics measurements in auditoriums. The halls used in the study include Mechanics Hall in Worcester, Massachusetts, Massey Hall in Toronto, and John Aird Centre Recital Hall in Waterloo, Ontario. It was discussed that omni-directional microphone impulses and binaural impulses are computer based measurement systems developed at the National Research Council of Canada (NRCC). Both these types of impulses create responses for a maximum length signal via a Fast Hadamard transform process. When the receiver of a sound source (for example, a dummy) is moved by only 30 cm, this can create a variation in early decay time (EDT) of 0.1 s at low frequencies and 0.05 s as high frequencies. At the higher frequencies, the head width adversely affects the wavelength, and so in this case, the directionality of the head becomes a factor of the aforementioned differences in EDT.

In addition to this, the measurements made using the dummy head are the average of two positions that are 15 cm apart (spacing between left and right ear). The energies of the two ears were superimposed and it was found that the head is less directional at lower frequencies. This is an indication that the energy for the binaural case at lower frequencies should be twice of the omni-directional microphone case. The dummy head was found to be more sensitive to sound originating from the side, which resulted in higher differences in relative level, G (level relative free field level at 10 m). Dummy head repositioning errors were about 1 dB for source/receiver differences of 30 cm. These errors were halved (0.5 dB) at higher frequencies, however the directionality of the head remains a factor to consider. A suggestion from Norcross et al. [48] addresses the systematic differences arising from head directionality is to use an average correction.

2.10. Summary of Literature Review

The various techniques, algorithms, and evidences introduced in this literature review have touched upon necessary aspects of noise control and reduction, head tracking and face detection, and dummy head applications for acoustic measurements. The most recurring themes were those pertaining to the combination or hybridization of active and passive means of control. Issues such as latency and accuracy of signals were reported. This also led to the discussion of several filters (mainly LMS-based) to alleviate these concerns. The cascade method for face detection was investigated, however its complexity requires the use of a large library of faces or images. After conducting this literature review, the general consensus was that active noise control systems are not extremely efficient due to the high level of signal processing required. A smart structure relating active noise control and head tracking is yet to be tested in great detail. Chapter 4 will look at viable head tracking methods for headrest active noise control betterment.

3. Zone of Quiet (ZoQ)

A Zone of Quiet (ZoQ) is created when a user is isolated from the outside acoustical environment. One of the major goals of the ZoQ is to focus the noise control strictly within the local zone of quiet. From a practical standpoint, the ZoQ is limited by the spatial distribution of sound sources and the geometry of the desired location of cancellation. A common local ANC application of a ZoQ is a noise reducing headrest in a passenger seat, attenuating noise around the passenger's ears. The physical microphone (located near the aircraft headrest) used to generate the error signal in the ANC scheme cannot be located inside a human ear. Active noise control permits the generation of a zone-of-quiet based on destructive interference. The size and shape of this zone-of-quiet depend on the type of the sound sources and the frequency components of the signal to be canceled [61].

3.1. Concept

The concept of zone-of-quiet can be used to define a design criterion, and in this case a mean squared error (MSE) criterion, at one particular point away from the physical microphone. Then a ZoQ-based approach can be used to derive a filter from the minimization of this MSE criterion to control the noise at the particular remote point. The research presented in this work predicts the spatial extent of zones-of-quiet at the error sensors when controlling pressure at a single secondary source in a pure tone diffused sound field. Minimization of sound pressure will produce noise cancellation of 10 dB from a secondary source over a spherical diameter of one-tenth of the acoustic wavelength, λ, of a compact primary source. This is an indication that the spherical diameter of the predicted zone-of-quiet is never larger than one-tenth of an acoustic wavelength.

The virtual microphone technique, which will be discussed later, projects the zone-of-quiet away from the physical microphone and closer to the desired location (ear) using the assumption of equal primary pressure at the physical and virtual locations.

A zone-of-quiet is defined by a sinc function, with the primary sound pressure level reduced by 10 dB over a sphere of diameter one tenth of the excitation wavelength, $\lambda/10$. The main benefit of the zone-of-quiet ANC technique versus conventional ANC methods for controlling the sound field in rooms is that the far field sound pressure level remains mostly unaffected, a typical increase of less than one dB [15]. Figure 3.1 below provides a rendered illustration of a person enclosed in a zone-of-quiet around his head and aircraft headrest area.

Fig. 3.1 Illustration of Zone of Quiet concept.

3.2. Performance Measure

Studies have shown that the 10 dB zone of quiet is extended to about one-tenth of an acoustic wavelength for pure tone diffuse fields [43]. However, in order to predict the performance of active headrests attenuating low-frequency tonal noise, the noise is broadband and not pure tone. For a broadband local ANC system a useful performance measure is the spatial extension of the overall sound attenuation which requires the analysis of broadband sound fields. Similar to zones of quiet for pure tone diffuse sound fields, spatial correlation of diffuse sound fields can be used to analyse broadband diffuse sound fields. This section and the next will introduce the diffuse sound field and spatial correlation for zones of quiet in broadband diffuse sound fields. It will be demonstrated that an initial approximation of a broadband zone of quiet is a ZoQ of tones at the mid-frequency of the broadband noise bandwidth.

3.2.1. Diffuse Sound Field

A diffuse sound field is one where sound pressure is the same throughout. A perfect diffuse field is almost unachievable, however, it is the model commonly used for reverberant sound field analysis. In these cases, the sound field is assumed to be sufficiently diffuse. This means that the field is diffuse above the Schroeder frequency. This is the frequency above which there exists at least three room acoustic modes within the 3 dB bandwidth of any one mode. For most enclosures, the Schroeder frequency is between 100 and 200 Hz. The equations presented in this section and in section 3.2.2 have been modified from Rafaely [55]. The pressure in a perfect diffuse field can be expressed as a function of space and time in spherical coordinates $\mathbf{x} = (x, \theta, \phi)$ as

$$e(\mathbf{x}, t) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} e_n(\mathbf{x}, t) \tag{3.1}$$

where $e(\mathbf{x}, t)$ is the total pressure at position $\mathbf{x}$ and time t, N is the number of plane waves which approaches infinity, and e_n is the nth plane wave. The autocorrelation (where observations are not independent) in the pure tone diffuse field can be expressed using a sinc function as

$$\rho(\Delta \mathbf{x}, t) = \frac{\text{sinc} (k \, \Delta \mathbf{x})}{k \, \Delta \mathbf{x}} \tag{3.2}$$

where k is the wave number and ρ is the correlation coefficient. The acoustic wave number, k, can be determined by dividing 2π times the natural frequency, f (in Hz), by c, the speed of sound (in m/s). The correlation coefficient, ρ, can be defined in space and time (cross-correlation) assuming the sound field is stationary over both as

$$\rho(\Delta \mathbf{x}, \Delta t) = \frac{E[e(\mathbf{x}_1, t_1) e(\mathbf{x}_0, t_0)]}{E[e^2]} \tag{3.3}$$

where $\Delta \mathbf{x}$ represents the distance between the two points, Δt represents the time delay, $E[\,\cdot\,]$ is the expectation operation which is determine by averaging several samples of diffuse sound fields, and $E[e^2]$ is the variance of the pressure which is independent of $\mathbf{x}$ and t because of the assumption that the sound field is stationary over space and time. The introduction of a time delay makes Equation 3.3 applicable to broadband diffuse sound fields. Now that the diffuse sound field has been explained, the following section will show how to determine the autocorrelation for zones of quiet in broadband diffuse sound fields.

3.2.2. Near-field Broadband Active Sound Control

Local active noise control in a diffuse sound field can be achieved by using a secondary source (often times a speaker) and canceling the total pressure in the near field of the source of noise. A theoretical model which addresses this approach models a monopole secondary source in a primary diffuse sound field. A derivation of the spatial extension of the zone of quiet in a broadband diffuse sound field for local active noise control will be presented.

The primary diffuse sound field is denoted by $d(\mathbf{x}, t)$. When a secondary monopole source is placed at the origin of a spherical coordinate system $\mathbf{x} = (x, \theta, \phi)$, the resulting pressure is expressed as $y(\mathbf{x}, t)$. The total pressure is found by superimposing the primary and secondary pressures as follows

$$e(\mathbf{x}, t) = d(\mathbf{x}, t) + y(\mathbf{x}, t) \tag{3.4}$$

The pressure cancellation is assumed to be at a point $\mathbf{x_0} = (x_0, \theta_0, \phi_0)$ such that

$$d(\mathbf{x_0}, t) + y(\mathbf{x_0}, t) = 0 \tag{3.5}$$

This shows that the primary and secondary fields are equal with opposite phase at the cancellation point $\mathbf{x_0}$, Figure 3.2 below is a visual representation of this concept.
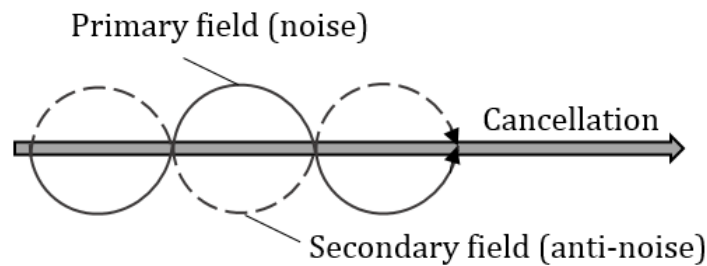
Fig. 3.2 Visual representation of primary and secondary fields at the cancellation point $\mathbf{x_0}$.

The position $\mathbf{x_0}$ can now be assumed to be in the near field of the secondary source, such that the indirect secondary sound field caused by reflections is negligible. The reverberation distance is defined as the distance from the source at which the direct field dominates. The reverberation distance depends on the room volume and reverberation time (length of time required for sound to decay 60 dB from its original level). The size of the zone of quiet is directly related to the effectiveness of the primary pressure attenuation around the cancellation point. The averaged squared total pressure at position $\mathbf{x_1} = (x_1, \theta_1, \phi_1)$ near the cancellation point using the expectation operation $E[\,\cdot\,]$ over several samples of diffuse sound fields. The variance of the total pressure at point $\mathbf{x_1}$ can now be determined using

$$E[e^2(\mathbf{x_1}, t)] = E[d^2(\mathbf{x_1}, t)^2] + E[y^2(\mathbf{x_1}, t)] + 2E[d(\mathbf{x_1}, t)y(\mathbf{x_1}, t)] \qquad (3.6)$$

The variance of the total pressure at $\mathbf{x_1}$ depends on the variance of the primary and secondary fields at the same point as well as on the correlation between the primary and secondary fields at $\mathbf{x_1}$. If the diffuse primary field is assumed to be stationary such that the variance of the pressure the same for all $\mathbf{x}$ and t, the following equality holds

$$E[d^2(\mathbf{x_1}, t)] = E[d^2(\mathbf{x}, t)] = E[d^2] \qquad (3.7)$$

The secondary source has been modelled as a monopole point source. Under some assumptions, the pressure produced by a monopole is similar to that produced by a piston in a baffle. Under one assumption, the source radius is significantly smaller than a wavelength λ and the source can be considered omni-directional. Under another assumption, only pressure further away than one source radius is considered for cancellation. These assumptions are sufficient for practical local active noise control systems such as active headrests, so the monopole model can be used successfully to obtain the behaviour of such systems. The secondary sound field produced

44

by a monopole point source in the near field is assumed to generate spherical waves which propagate away from the source. The spherical waves also decay in amplitude. The secondary sound field can be expressed as

$$y(\mathbf{x}, t) = \frac{\rho_0}{4\pi r} \dot{q}\left(t - \frac{x}{c}\right) \tag{3.8}$$

which is now dependent only on the distance from the source $\mathbf{x}$, with q representing the source strength (volume velocity per unit volume) and $\dot{q}$ its derivative with respect to time. The source strength q can be determined from the surface area and the amplitude of vibration of the spherical monopole point source. The secondary pressure at $\mathbf{x_1}$ can now be expressed in terms of the secondary pressure at $\mathbf{x_0}$ using Equation 3.8 as

$$y(\mathbf{x_1}, t) = \frac{x_0}{x_1} y\left(\mathbf{x_0}, t - \frac{\Delta x}{c}\right) \tag{3.9}$$

where Δx is the difference in the distances of the two points x_1 and x_0 to the source. The averaged squared secondary pressure at x_1 can now be written using Equations 3.9, 3.7 and 3.5 as

$$E[e^2(\mathbf{x_1}, t)] = \left(\frac{x_0}{x_1}\right)^2 E[d^2] \tag{3.10}$$

The last term in Equation 3.6 can be expressed using Equations 3.9, 3.5, and 3.3 as

$$E[d(\mathbf{x}, t)y(\mathbf{x_1}, t)] = -\frac{x_0}{x_1} \rho\left(\Delta\mathrm{x}, \frac{\Delta x}{c}\right) E[d^2] \tag{3.11}$$

The variance of the total pressure at position $\mathbf{x_1}$ in Equation 3.6 can now be expressed in terms of the variance of the primary pressure by substituting Equations 3.11, 3.10 and 3.7 into Equation 3.6 yielding

$$E[e^2(\mathbf{x_1}, t)] = E[d^2] + \left(\frac{x_0}{x_1}\right)^2 E[d^2] - 2\frac{x}{x_1} \rho\left(\Delta\mathbf{x}, \frac{\Delta x}{c}\right) E[d^2] \tag{3.12}$$

45

The final expression for the sound attenuation ϵ at $\mathbf{x_1}$ (measurement point) can now be obtained by dividing Equation 3.12 by the variance of the primary pressure. Assuming cancellation at $\mathbf{x_0}$, this will yield

$$\epsilon(\mathbf{x_1}, \mathbf{x_0}) = \frac{E[p^2(\mathbf{x_1}, t)]}{E[d^2]} = 1 + \left(\frac{x_0}{x_1}\right)^2 - 2\frac{x_0}{x_1}\rho\left(\Delta\mathbf{x}, \frac{\Delta x}{c}\right) \tag{3.13}$$

where the sound attenuation in dB can be found by $10\log_{10}\epsilon$. It should be noted that in Equation 3.13, $\Delta\mathbf{x}$ is the distance from $\mathbf{x_1}$ to the cancellation point $\mathbf{x_0}$, whereas Δx is the difference between the distances of the two points $\mathbf{x_1}$ and $\mathbf{x_0}$ to the secondary source (refer to Figure 3.3). The cross-correlation function in a diffuse field derived in the previous section, along with Equation 3.13 can be used to determine the zone of quiet with respect to the speaker or secondary source location. Generally, the size of the zone of quiet decreases at higher frequencies and at closer cancellation points, as can be seen in Figure 3.4. Figure 3.4 demonstrates the 10dB quiet zone size in mm with respect to frequencies between 0 and 100 Hz at different $\mathbf{r_0}$ values (10, 100 and 200 mm). In this case, $\mathbf{x_0}$ is equivalent to $\mathbf{r_0}$, which was the variable Elliott [15] used to represent the cancellation point. As the distance from the speaker to the cancellation point, x_0 is increased, the distance from $\mathbf{x_1}$ to the cancellation point $\mathbf{x_0}$ is also increased. This is a performance metric that can be used to determine the optimal zone of quiet.

46
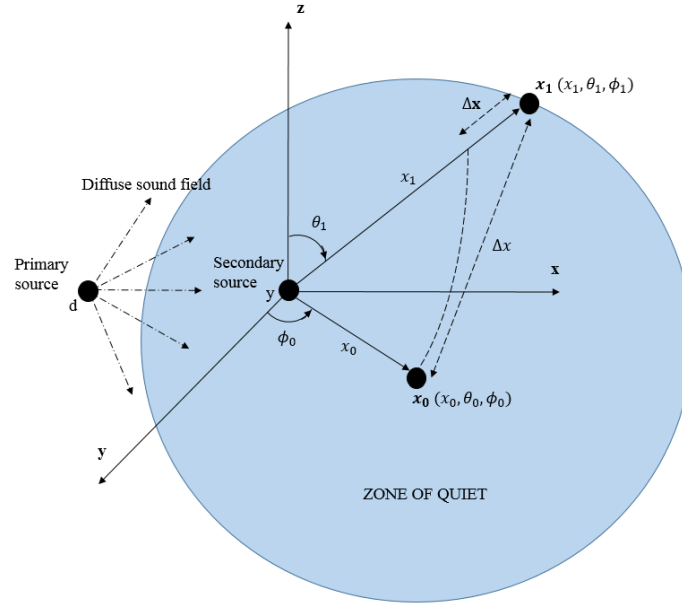
Fig. 3.3 Diagram of cancellation point x_0 and a position x_1 near it, relative to the secondary source.
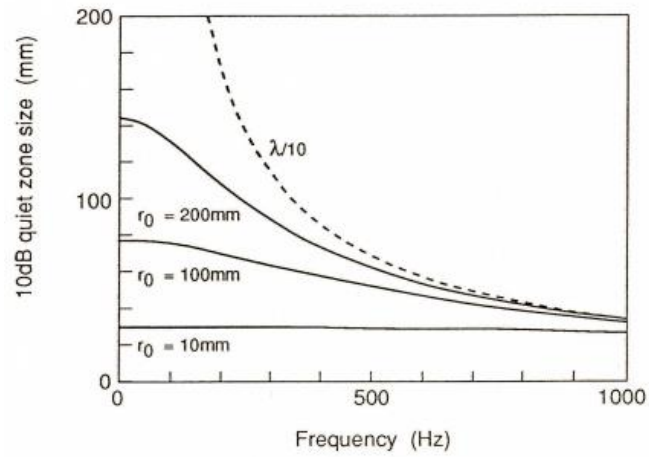
Fig. 3.4 10dB ZoQ reduction for 100mm dia. speaker adjustment w.r.t. distance r_0 on-axis [15].

3.3. Feedforward Filtered-x LMS Control

The general purpose of a feedforward control system is to reduce the levels of a disturbance through the addition of a secondary control signal that is determined by a computational algorithm. This algorithm is often tasked with minimizing sound radiation from structures when used in active noise control applications. One such feedforward control algorithm is the Filtered-x LMS algorithm, a modified version of the least-mean-square (LMS) algorithm commonly used for signal cancellation.

3.3.1. Single Channel Control (SISO)

The simplest form of active noise control is Single Input Single Output (SISO). Figure 3.5 is a block diagram of a single input single output (SISO) filtered-x LMS control system. The plant output $e(n)$, or error signal, can be defined as

$$e(n) = d(n) + y(n) \qquad (3.14)$$

where $y(n)$ is the plant response to the control input (which is a speaker), $d(n)$ is the plant response to the disturbance input, and n is the time step, in relation to Equation 3.5. However, here the error signal is not assumed to be equal to zero because it is a direct measurement. The plant response to the control input $y(n)$, in Equation 3.14 can be rewritten as

$$e(n) = d(n) + T_{ce}(z)u(n) \qquad (3.15)$$

where $T_{ce}(z)$ is the z-transform of the transfer function between the system response $y(n)$ and the controller output $u(n)$.
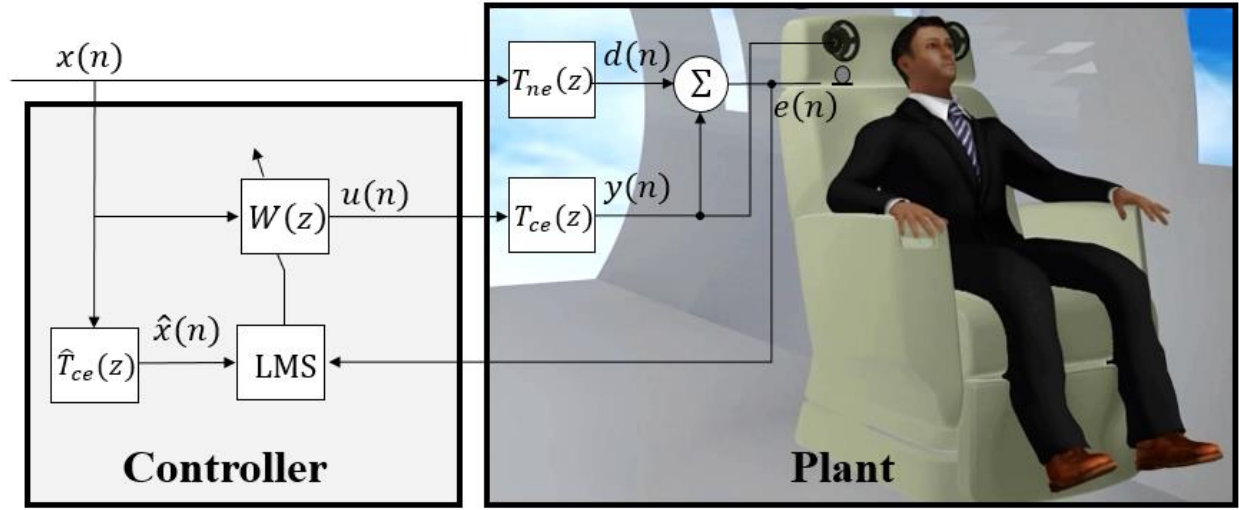
Fig. 3.5 Block diagram of single channel Filtered-x LMS feedforward control system.

The output signal from the controller, *u(n),* is obtained from the adaptive Finite Impulse Response (FIR) filter represented as *W(z).* The adaptive FIR filter is widely used in control algorithms and is fundamental component used in the implementation of many control algorithms and is unaffected by signal impulse (zero impulse response) after convergence of its coefficients. A FIR filter's output is the weighted sum of previous inputs, such that

$$y(n) = \sum_{i=0}^{I-1} W_i(n)x(n-i) \tag{3.16}$$

where $W_i(n)$ is the ith filter coefficient and I is the filter order. The output signal from the controller, $u(n)$, can be expressed as

$$u(n) = W(z)x(n) \tag{3.17}$$

In order for the adjusted plant input signal *u(n)* to minimize the plant response *e(n),* a cost function has to be determined.

It is characterized by a quadratic function of the filter coefficients. The LMS algorithm uses the gradient descent method as an adaptation criterion for adjusting the filter coefficients W_i ($i = 0, ..., I - 1$) so that the minimum mean-square error is found. The adaptation equation for the weights is

$$W_i(n + 1) = W_i(n) - \mu \frac{\partial C}{\partial W_i} \tag{3.18}$$

where μ is the convergence coefficient and controls the step size or rate of convergence in the minimization process. The update process is known as system identification. The update equation or control law for the filter coefficients can now be determined,

$$W_i(n + 1) = W_i(n) - 2\mu \hat{x}(n - i)e(n) \tag{3.19}$$

The convergence coefficient μ controls the step size of the LMS algorithm during iteration towards the global minimum of the quadratic cost function. As the convergence coefficient μ increases, the time for convergence decreases. Large μ values can, however, cause the algorithm to become unstable where the control inputs increase without bounds. The stability of the LMS algorithm can be monitored using the following expression which contains the input autocorrelation matrix R (signal correlation over time):

$$0 < \mu < \frac{1}{Trace[R]} \tag{3.20}$$

The autocorrelation matrix R can be determined in Matlab using the convolution between the reference signal and the reversed version of the conjugate of the signal. A general rule of thumb, is that the convergence coefficient, μ, be set to a large value initially and then later decreased. This is done so that convergence is reached, but also so that the cost function is able to find the minimum value.

3.3.2. Multiple Channel Control (MIMO)

It may not always be possible to obtain global noise control with the introduction of a single error sensor and secondary source. Often the primary disturbance or the radiated sound field is such that it requires the introduction of multiple sources and error sensors to achieve global control or control over a large volume. The filtered-x LMS algorithm has been developed for use with multiple secondary control sources and error sensors. This type of control system is known as a multiple-input-multiple-output or MIMO system. The modified update equation for each control filter coefficient with L actuators and M sensors is

$$W_{li}(n+1) = W_{li}(n) - 2\mu \sum_{m=1}^{M} \hat{x}_{ml}(n-i)e_m(n) \qquad (3.21)$$

where $\hat{x}_{ml}(n)$ is the reference signal filtered with a model of the system path from the lth actuator and the mth error sensor at sample n. It should be noted that Equation 3.21 is equivalent to the SISO system of Equation 3.19 when $L = M = 1$. Moreover, the level of control that can be obtained by a control system also depends on the quality of the reference signal used to capture the dynamics of the disturbance. For a spatially incoherent disturbance, as provided by a turbulent boundary layer, multiple references (K) may be required to achieve adequate control levels. The multiple-channel controller now contains MK finite impulse response filters each updated by the control algorithm. The zone of quiet is also affected by a MIMO system, such that two individual zones of quiet (one on either side of the head) can be coupled to yield a larger and more localised zone of quiet around the passenger's head. Figure 3.6 demonstrates a MIMO zone-based active noise control system around a head. A similar graphical setup to Figure 3.3 is used to illustrate that the combined zone of quiet would change with respect to the changing and coupled distances from the speakers to the cancellation points on either side of the head. A more detailed representation

using the adaptive LMS moving virtual microphone method and head tracking is presented in Figure 3.10.
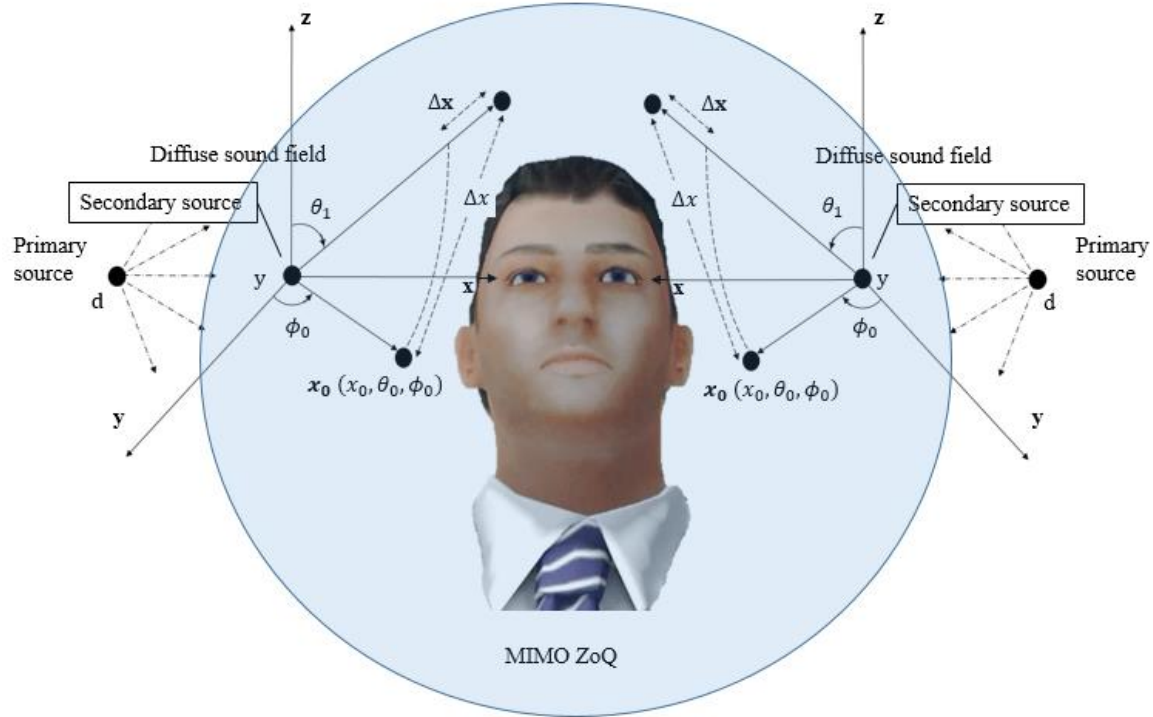
Fig. 3.6 MIMO graphical representation of zone-based active noise control.

3.3.3. Virtual Microphones near Passenger's Ears

One of the challenges for active noise cancelling applications is to create quiet zones at the locations of virtual sensors (such as the passenger's ears). When only the pressure is estimated at a remote location the transducer is referred to as a "virtual microphone" [12]. A common issue with virtual sensors is that the desired location of attenuation is not spatially fixed in some cases. For example, if the virtual sensor is situated near the passenger's ears it will have to be relocated when the passenger moves his head. A comparison of active noise control at a physical sensor and

at a virtual sensor is demonstrated in Figure 3.7 below. The performance of a local active headrest system using the virtual microphone arrangement has been experimentally analyzed by several authors [18, 19, 26, 50, 51, 54, and 56]. According to Garcia-Bonito et al. [18, 19], below 500 Hz, the 10 dB zone of quiet generated at the virtual microphone extends approximately 8 cm forward and 10 cm side to side. During their experiments, the virtual microphones were located 2 cm from the ears of a mannequin and 10 cm from the physical microphones. Minimizing both pressure and pressure gradient along a single axis can also significantly increase the noise reduction and ZoQ along the axis of pressure gradient measurement [43].
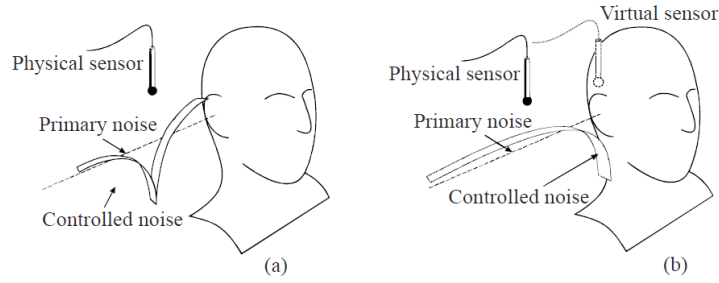
Fig. 3.7 ANC Comparison. *(a): at a physical sensor; (b): at a virtual sensor* [43].

3.4. Virtual Sensing Problem

The virtual sensing problem and notation presented here will be used for the virtual microphone technique which was developed by Daniel Moreau et al. [43]. The physical microphones will be denoted as M, the spatially fixed virtual microphones as M_v, and the secondary sources (speakers) as L. The vector of total pressures at the M physical microphones, $\mathbf{e}(n)$, is defined as

$$e(n) = \begin{bmatrix} e_{p1}(n) & e_{p2}(n) & ... & e_M(n) \end{bmatrix}^T \tag{3.22}$$

The total pressures at the M physical microphones, $\mathbf{e}(n)$, is the sum of the sound fields produced by the primary and secondary sound sources at the physical microphone locations and can be expressed as

$$e(n) = d(n) + y(n) = d(n) + T_{ce}u(n) \qquad (3.23)$$

where $\mathbf{d}(n)$ is the vector of primary pressures at the M physical microphones, $\mathbf{y}(n)$ is the vector of the secondary pressures at the M physical microphones, $\mathbf{T}_{ce}$ is the matrix of size $M \times L$ whose elements are the transfer functions between the secondary sources and the physical microphones, $\mathbf{u}(n)$ is the vector of the control sequence (output signal from controller) and n is the time step.

Similarly, the vector of total pressures at the M_v spatially fixed virtual locations, $\mathbf{e}_v(n)$, is defined as

$$e_v(n) = \begin{bmatrix} e_{v1}(n) & e_{v2}(n) & ... & e_{vM_v}(n) \end{bmatrix}^T \qquad (3.24)$$

The total pressures at the M_v virtual microphones, $\mathbf{e}_v(n)$, is the sum of the sound fields produced by the primary and secondary sound sources at the virtual locations and can be expressed as

$$e_v(n) = d_v(n) + y_v(n) = d_v(n) + T_{ve}u(n) \qquad (3.25)$$

where $\mathbf{d}_v(n)$ is the vector of the primary pressures at the M_v virtual microphones, $\mathbf{y}_v(n)$ is the vector of the secondary pressures at the M_v virtual locations and $\mathbf{T}_{ve}$ is the matrix of size $M_v \times L$ whose elements are the transfer functions between the secondary sources and the virtual microphones. A virtual sensing technique aims at estimating the pressures, $\mathbf{e}_v(n)$, at the spatially fixed virtual locations using the physical error signals, $\mathbf{e}(n)$. The estimated pressures are

minimised (instead of the physical error pressures) with the ANC system to create zone of quiet at the virtual locations.

3.5. Virtual Microphone Technique

The virtual microphone technique was the first ANC virtual sensing method. This technique used the assumption of equal primary sound pressure at the physical and virtual microphone locations. Figure 3.8 below is a block diagram of the virtual microphone technique. This method can be carried out using equal numbers of physical and virtual microphones ($M_v = M$). The microphones are located in M_v pairs, with one physical microphone and one virtual microphone in each pair. The primary sound pressure is assumed to be equal at the physical and virtual microphones in each pair such that $\mathbf{d}_v(n) = \mathbf{d}(n)$. This assumption is valid only if the primary sound field does not change drastically between the physical and virtual microphones in each pair.
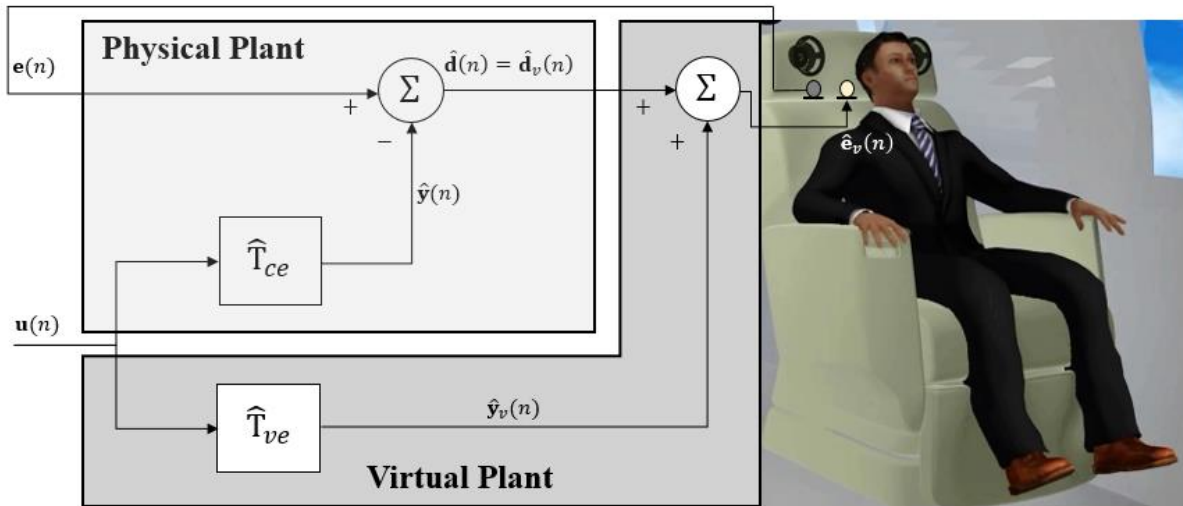
Fig. 3.8 Block diagram of the virtual microphone technique.

An initial identification process is required for the virtual microphone technique, where the matrices of the transfer functions $\widehat{\mathbf{T}}_{ce}$ and $\widehat{\mathbf{T}}_{ve}$, are modelled as matrices of FIR filters. Once this preliminary stage is accomplished, the microphones at the virtual locations are removed. Estimates, $\hat{\mathbf{e}}_v(n)$, of the total error signals at the virtual locations can be expressed as

$$\hat{\boldsymbol{e}}_v(n) = \boldsymbol{e}(n) - (\widehat{\boldsymbol{T}}_{ce} - \widehat{\boldsymbol{T}}_{ve})\boldsymbol{u}(n) \tag{3.26}$$

In a pure tone diffuse sound field and at low frequencies, the zone of quiet created at a virtual microphone using the virtual microphone technique is similar to that achieved by directly minimising the measured pressure of a physical microphone at the virtual location. However, at frequencies above 500 Hz, the 10 dB zone of quiet is significantly smaller when using virtual microphones compared to physical microphones at the same locations. This is because the equal primary pressure (at the physical and virtual microphone locations) assumption holds less as the wavelength decreases. A reason for the differences in the zone of quiet generated by minimising the physical and virtual microphone signals is that the physical microphone is drastically closer to the secondary source (speaker) than the virtual microphone. Consequently, the virtual plant inherits a longer delay, which decreases the level of attenuation.

3.6.   Adaptive LMS moving virtual microphone technique

The adaptive LMS moving virtual microphone technique developed by Daniel Moreau et al. [43] employs the adaptive LMS virtual microphone technique to estimate the virtual error signals at the moving virtual locations. The adaptive LMS moving virtual microphone technique

determines estimates of the virtual error signals, $\hat{\mathbf{e}}_v(n)$, at the moving virtual locations, $\mathbf{x}_v(n)$. A block diagram of the adaptive LMS moving virtual microphone technique is shown in Figure 3.9.

In this moving virtual sensing algorithm, the adaptive LMS virtual microphone technique is first used to obtain estimates of the virtual error signals, $\hat{\bar{\mathbf{e}}}_v(n)$ at the spatially fixed virtual locations $\bar{\mathbf{x}}_v$. The primary component of the physical error signals is first calculated using the matrix of physical secondary transfer functions $\hat{\mathbf{T}}_{ce}$ and is given as

$$\hat{\mathbf{d}}(n) = \mathbf{e}(n) - \hat{\mathbf{y}}(n) = \mathbf{e}(n) - \hat{\mathbf{T}}_{ce}\mathbf{u}(n) \tag{3.27}$$

Matrices of the primary and secondary weights, $\bar{\mathbf{w}}$ and $\bar{\mathbf{w}}_e$, of size $M \times \bar{M}_v$, at the $\bar{M}_v$ spatially fixed virtual locations, $\bar{\mathbf{x}}_v$, are then estimated separately. Estimates, $\hat{\bar{\mathbf{e}}}_v(n)$, of the total virtual error signals at the spatially fixed virtual locations, $\bar{\mathbf{x}}_v$, can then be calculated as

$$\hat{\bar{\mathbf{e}}}_v(n) = \hat{\bar{\mathbf{d}}}_v(n) + \hat{\bar{\mathbf{y}}}_v(n) = \bar{\mathbf{w}}^T\hat{\mathbf{d}}(n) + \bar{\mathbf{w}}_e^T\hat{\mathbf{y}}(n) \tag{3.28}$$

Figure 3.9 estimates, $\hat{\mathbf{e}}_v(n)$, of the virtual error signals at the moving virtual locations, $\mathbf{x}_v(n)$, are now obtained by spatially interpolating the virtual error signals, $\hat{\bar{\mathbf{e}}}_v(n)$, at the spatially fixed virtual locations, $\bar{\mathbf{x}}_v$.
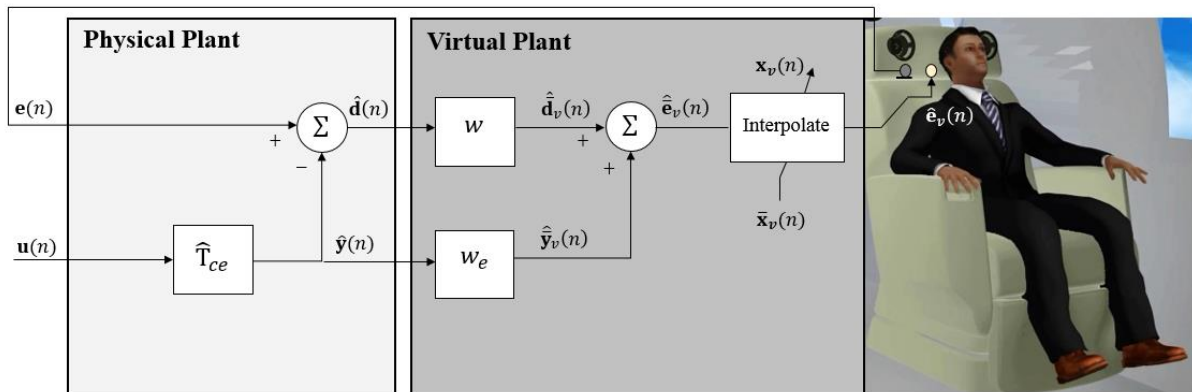
Fig. 3.9 Block diagram of the adaptive LMS moving virtual microphone technique.

The performance of the adaptive LMS moving virtual microphone technique has been investigated extensively by Moreau et al. [43]. In one of their studies, the virtual microphone was moved sinusoidally between a virtual distance of $v = 2$ cm and 12 cm in a time frame of 10 seconds. Experimental results have shown that further applying the feedforward control approach to this technique yields up to an additional 18 dB of attenuation at the moving virtual location compared to minimising the error signal at a fixed physical microphone at virtual distance $v = 0$ m or a fixed virtual microphone at virtual distance of 2 cm. In lieu with head tracking, the adaptive LMS moving virtual microphone technique can be used to create the optimal zone of quiet around an aircraft passenger's head. Referring to Figure 3.3, the zone of quiet representation can be modified to include the adaptive LMS moving virtual microphone elements. This can be done to create a local coordinate system for each ear of the passenger as its tracked using the head-tracking tool developed in this work. This modified representation of the zone of quiet can be seen in Figure 3.10 below. The cancellation point $\mathbf{x_0}$ now becomes $\bar{\mathbf{x}}_{vr}$ and $\bar{\mathbf{x}}_{vl}$, the spatially fixed virtual microphone locations for the right and left ear, respectively. These fixed microphone locations are interpolated to yield the moving virtual microphone locations for the right and left ear, $\mathbf{x}_{vr}$ and $\mathbf{x}_{vl}$, respectively. No matter if the head is rotated or translated about any axis, this coordinate system can be used to relate the moving virtual locations to the physical microphone locations near the right and left ears, $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively which will be needed to determine the error, $\hat{\mathbf{e}}_v$.
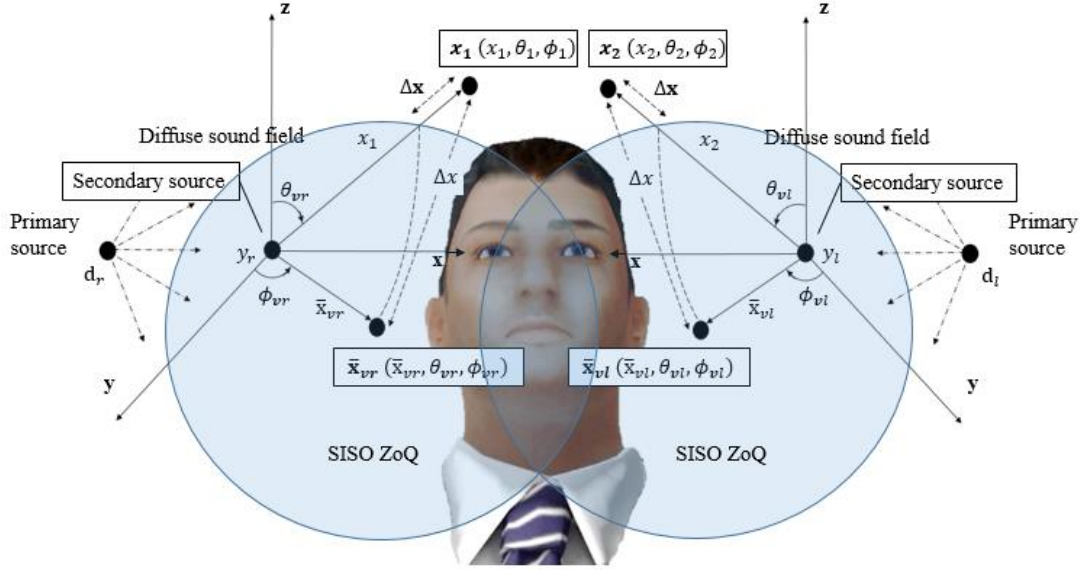
Fig. 3.10 Coordinate system for LMS moving virtual microphone technique with head tracking.

3.7. Forward Difference Prediction Technique

For the purpose of this analysis, only first order and second-order virtual microphone methods will be explored. A first-order and a second order forward difference extrapolation virtual microphone are known to have 2 and 3 physical microphones, respectively, associated with their design. This will be explained after their corresponding derivations. The reasoning behind fitting a straight line or a curve between pressures measured at fixed locations is that the spatial rate of change of the sound pressure between the locations is small at low frequencies. This makes it highly predictable. The locations should be relatively closely spaced (with respect to the wavelength of the sound) regardless of the surrounding environment. The fitted straight line or curve can then be used to estimate the pressures at other locations via interpolation or extrapolation depending on the location of the observer. The equations that will be used in the following derivations use extrapolation since it is desired that the observer be remote from the physical sensors. It should be noted that the symbols which will be used to denote pressure can be related

to the previous symbols used for the virtual microphone methods. For example, the pressure at the virtual location for the forward difference prediction method is p_x, which is equivalent to e_v and similarly p_1 is equivalent to e_1.
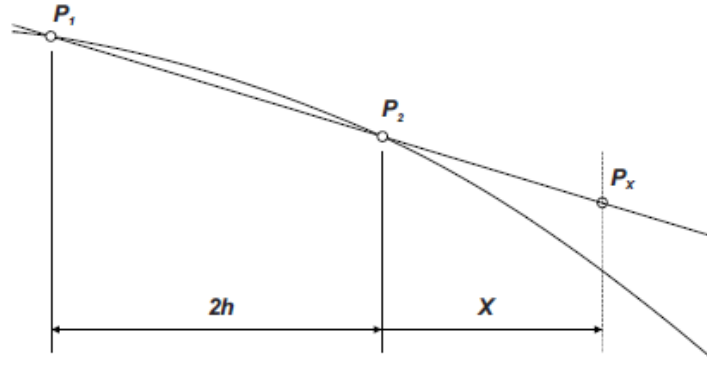
Fig. 3.11 First-order forward prediction.
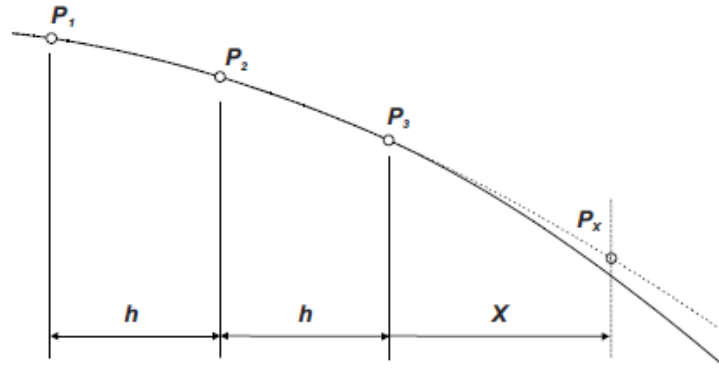
Fig. 3.12 Second-order forward prediction.

3.7.1. Two-microphone first-order pressure prediction

The pressure at location x (see Figure 3.11) can be determined by the first-order finite difference estimate from two remote microphones. The remote microphones are separated by a distance of $2h$. The pressure at location x can be found by measuring pressures p_1 and p_2 at two microphone locations by applying:

$$p = \left(\frac{dp}{dx}\right)x + c, \quad \text{where } \frac{dp}{dx} \text{ is constant} \tag{3.29}$$

$$p_x = \frac{(p_2 - p_1)}{2h}x + p_2 \tag{3.30}$$

So now, if the separation distance x between the observer and the nearest sensor is equal to h then Equation 3.30 becomes:

$$p_x = \frac{1}{2}(3p_2 - p_1) \tag{3.31}$$

Similarly, if the separate distance is increased to $2h$ then:

$$p_x = 2p_2 - p_1 \tag{3.32}$$

It is worth noting that when the separation distance $x = 0$, or if $x = -2h$ that Equation 3.30 reduces to $p_x = p_2$ or p_1 respectively.

3.7.2. Three-microphone second-order pressure prediction

The rate of change of the pressure gradient can be estimated by placing a third microphone between the other two physical microphones and the virtual locations. This three-microphone second-order method allows a more accurate prediction of the pressure at the virtual locations (see Figure 3.12). In this case, the second derivative of the pressure gradient change, $\frac{d^2p}{dx^2}$, is constant. This constant can be integrated to obtain the relationship between the pressure at the virtual location and the measured pressures at the three physical microphone locations as follows:

$$p_x = \iint \frac{d^2p}{dx^2} \cdot (dx)(dx)$$

$$p_x = \iint k_1 . (dx)(dx)$$

$$p_x = \int (k_1 x + k_2) . (dx)$$

$$p_x = \frac{k_1 x^2}{2} + k_2 x + k_3 \tag{3.33}$$

Now using the location of each of the three microphones at $x_3 = 0$, $x_2 = -h$ and $x_1 = -2h$, the constants of integration k_1, k_2 and k_3 can be found by applying above Equation 3.33. This yields:

$$\begin{bmatrix} p_3 \\ p_2 \\ p_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -h & h^2/2 \\ 1 & -h & 2h^2 \end{bmatrix} \begin{bmatrix} k_3 \\ k_2 \\ k_1 \end{bmatrix} \tag{3.34}$$

Solving the above system provides the integration constants, which can now be substituted into Equation 3.33 to obtain the pressure at a separation distance x as follows:

$$p_x = \left(\frac{p_1 - 2p_2 + p_3}{h^2} \right) \frac{x^2}{2} + \left(\frac{p_1 - 4p_2 + 3p_3}{2h} \right) x + p_3 \tag{3.35}$$

However, a more useful equation for the pressure at the location x can be determined by grouping together like terms to yield the weighting factors for each physical microphone as follows:

$$p_x = \frac{x(x+h)}{2h^2}p_1 + \frac{x(x+2h)}{-h^2}p_2 + \frac{(x+2h)(x+h)}{2h^2}p_3 \qquad (3.36)$$

So now, if the separation distance x between the observer and the nearest sensor is equal to h then Equation 3.36 becomes:

$$p_x = p_1 + 3(p_3 - p_2) \qquad (3.37)$$

Similarly, if the separate distance is increased to $2h$ then:

$$p_x = p_1 - 2(p_2 - p_1) + 6(p_3 - p_2) \qquad (3.38)$$

Again, it is worth noting that when the separation distance $x = 0$, or $x = -h$, or $x = -2h$ that Equation 3.36 reduces to $p_x = p_3$, p_2 or p_1 respectively. Now with adequate head tracking methods which will be discussed in Chapter 4, the adaptive LMS moving virtual microphone technique can be implemented for the active headrest system.

3.8. Integrated ANC System Building Blocks

Zone-based ANC technologies are currently being used by several world-wide markets including but not limited to home armchairs, train seats, vehicle interiors and aircraft interiors. The purpose of such technologies is to obtain environmental noise and in turn to create a zone of quiet. This can be achieved by capturing the physical characteristics of the unwanted noise field using a multichannel algorithm. This technique allows for sound field control via a loudspeakers array. A widely-used ZoQ ANC technology provides up to 14 dB of active noise reduction from a range of 100-1500 Hz. The benefit of the technology is that it delivers active noise cancellation without the need to wear a headset.

Before any zone-based ANC technology is implemented in an aircraft passenger seat, a theoretical approach must be taken in order to find the optimal locations for sensors and secondary sources. The ZoQ technology uses four main transfer functions for local control:

- MTF – Transfer function between the reference sensors and the errors sensors

- STF – Transfer function between the speaker input and the error sensors

- EC – Transfer function between the speaker input and the reference sensors

- PF – Prediction Filter that is used to predict future sampling data from the reference microphone

This technology measures the transfer functions between the desired locations of maximum reduction and the error sensor locations in order to create the appropriate control algorithms. The method that will be presented in this work uses up to five microphones to estimate the sound pressure level at a remote location using a forward-difference prediction method. Similar to the

ZoQ technology, the simulation uses a prediction filter to predict the sound pressure levels using the physical reference microphones.
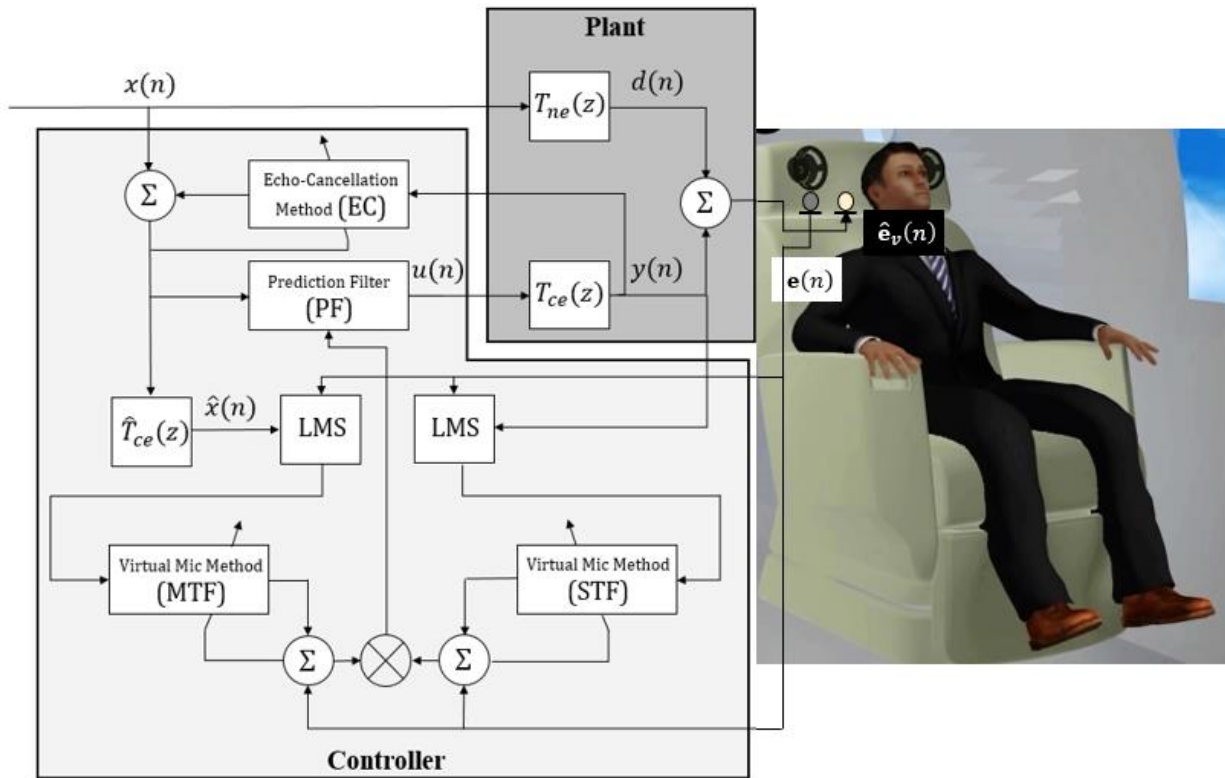
Fig. 3.13 Integrated ANC Building Block diagram.

3.8.1. MTF Transfer Function

The MTF is the transfer function between the reference sensors and error sensors. It is used to perform the Virtual Microphone Method. The system is capable of using 8 reference sensors, four of which can be microphones, and the other four are accelerometers. In some cases, it may be beneficial to use only 8 accelerometers. The system typically consists of 4 physical error sensors, which are microphones. In the case that an active/passive approach is used, the reference signals pass through a passive element before determining the plant response to the disturbance input, *d(n)*. Once a perfect electrical model (filter) of the acoustic medium between the reference sensors and the error sensors exists, the reference sensors signal can be fed to the model. This will result in an estimation of the undesired noise at the error sensors location. The MTF is estimated by applying the adaptive filter technique to the reference sensors and the error sensors. The input *x(n)* is connected to the reference sensors, and the input *d(n)* is connected to the error sensors. Generally-speaking, the reference sensors are accelerometers and the error sensors are microphones.

Fig. 3.14 ANC MTF Block diagram.

3.8.2. STF Transfer Function

The STF is the transfer function between the speaker input and the error sensors. It is mostly used to perform the Virtual Microphone Method. The system requires 4 outputs for the speakers. There are usually 8 reference sensors, 4 of which are microphones and 4 are accelerometers. It is also common to use 8 reference accelerometers to accurately capture the primary sound field. STF training uses the ZoQ relationship depicted in Figure 3.4 between the speaker and cancellation point, x_0 (virtual microphone location) is used. For a desired bandwidth (frequency range of cancellation) and the distance between the speaker and cancellation point, an optimal zone of quiet and electrical filter are determined. Once a perfect electrical model (filter) of the acoustic medium between the speaker and the error sensors exists, the speaker input signal can be fed to the model. This will result in an estimation of the reductive noise at the error sensors location. The STF is estimated by applying the adaptive filter technique to the speaker input and the error sensors, as the speaker is fed with white noise via a signal generator. The input $x(n)$ is connected to the speaker (which connects to the signal generator), and the input $d(n)$ is connected to the error sensors.
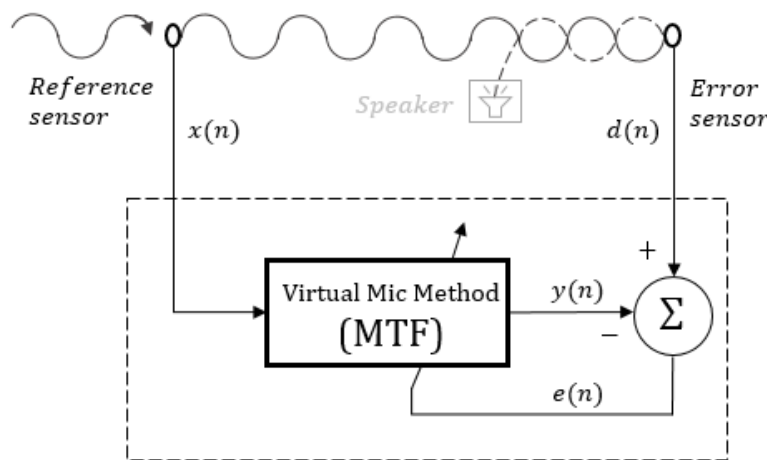

Fig. 3.15 ANC STF Block diagram.

67

3.8.3. EC Transfer Function

The EC is the transfer function between the speaker input and the reference sensors. It is used to perform the Echo-Cancellation Method. Once a perfect electrical model (filter) of the acoustic medium between the speaker and the reference sensors exists, the speaker input signal can be fed to the model. This will result in an estimation of the reductive signal at the reference sensors location. This estimated signal is subtracted from the reference sensors signal to obtain and estimation of the undesired noise. Sometimes one speaker produces higher energy than the other speakers. The system has built-in speaker stability settings which can be used to address this concern. The EC is estimated by applying the adaptive filter technique to the speaker input and the reference sensors, while feeding the speaker with white noise through an intrinsic signal generator. The EC transfer function acts as a balance between the STF transfer function. This means that there is a compromise between the size of the 10 dB zone of quiet and the speaker sound quality which is reflected and recorded by the error sensors. The input $x(n)$ is connected to the speaker (which connects to signal generator), and the input $d(n)$ is connected to the reference sensors.
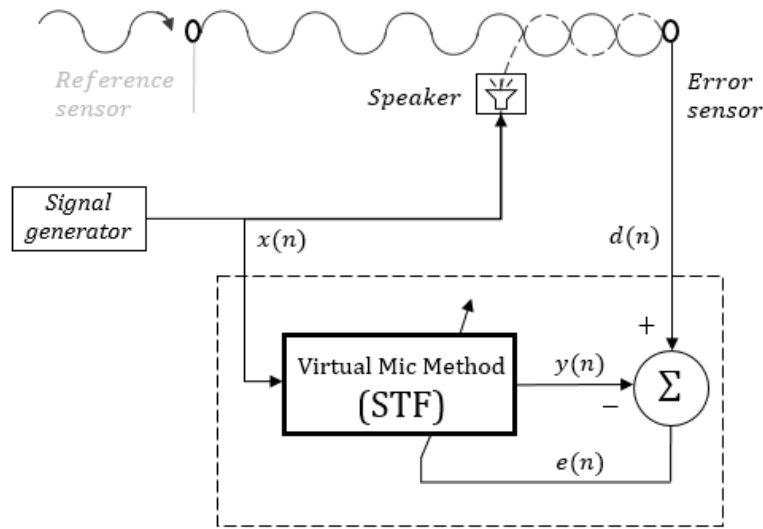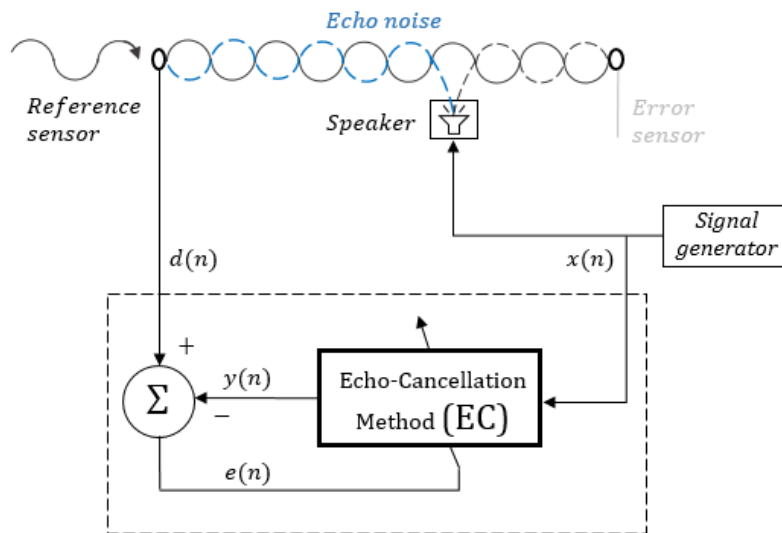

Fig. 3.16 ANC EC Block diagram.

68

3.8.4. Prediction Filter (PF)

 PF represents the prediction filter that is used to predict future sampling data from the reference sensors, in order to shorten the delay of the system and to reduce the overall size of the system. PF can be calculated by using MTF, STF, and the signals of the reference sensors and the error sensors that were recorded during the MTF estimation. This indicates that PF is computed after the MTF and the STF have been determined. The PF is an important aspect in the active noise control performance of this system. The quality of the PF depends on the bandwitch of the noise to be reduced and the predictability of the signal. ANC performance is enhanced if the frequency-band of the noise to be reduced is narrower. Signal predictability is critical because two signals with the same bandwidth can be different in their mathematical models, resulting in different ANC performance. There are 3 types of prediction filters available: (1) High, (2) Middle, and (3) Low. Each one is for a different situation or application. For example, the system can be designed to store 3 filters for 3 difference noise sources or 3 different passenger head configurations (desired heights, angles, etc.). The user has the ability to switch between the filter types offline.
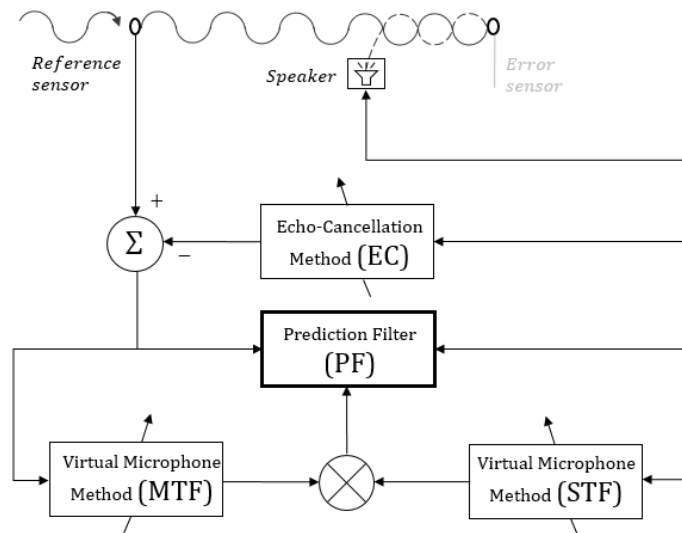
Fig. 3.17 ANC Prediction Filter Block diagram.

69

3.9. Model

The simulation uses Matlab's SIMULINK and consists of 5 equally-spaced microphones. The microphones are identical and precisely equally-spaced in order to create the same time delays and frequency responses for the noise cancellation. The source was located at the end of the duct (right-side of cabin). The total distance covered by the microphones was $2h = 0.05$m with the fifth microphone being the furthest from the source at a distance of 1m. The cabin was modelled as a rectangular duct for simplicity in determining the cabin' axial acoustic modes. In reality, the cabin consists of complex wall and seating geometries. The duct length of 5.6873 m, height of 1.91 m, and width of 2.49 m are approximately equal to that of the cabin entertainment zone length, height and width of the Global 7000 business jet.
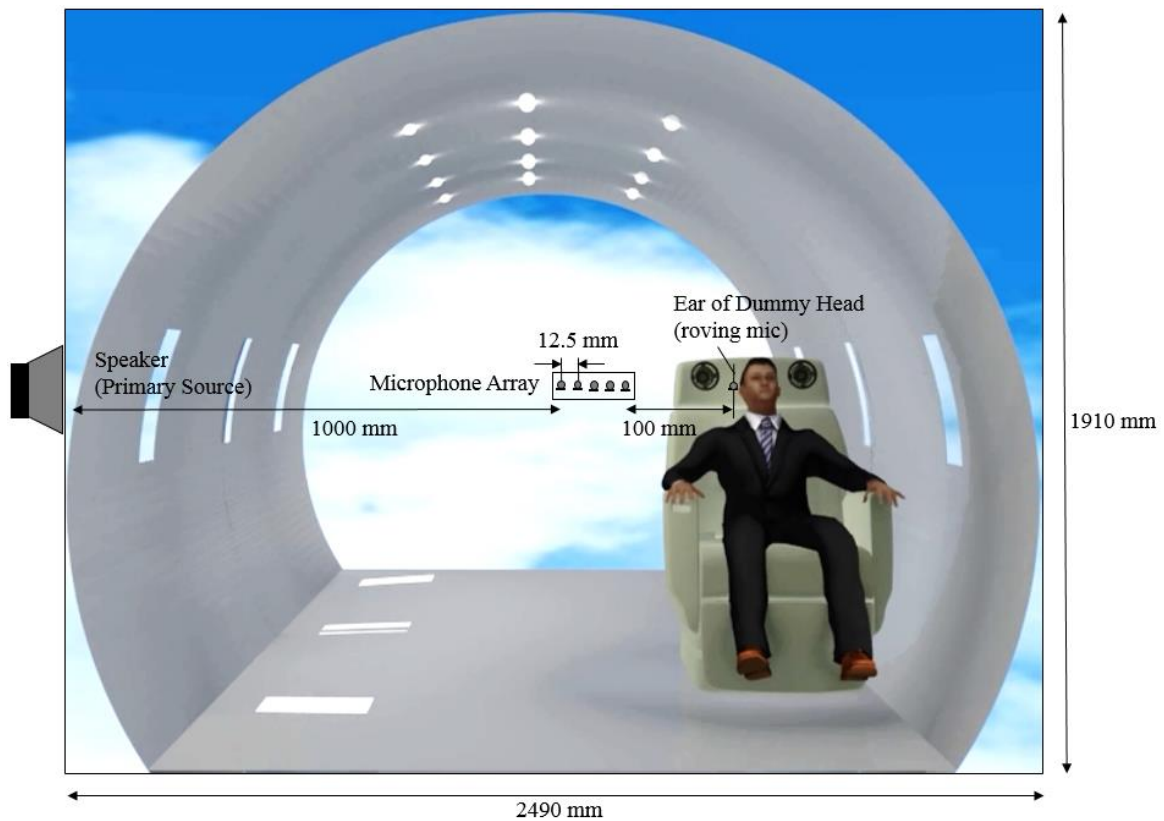
Fig. 3.18 Representation of experiment setup.

The Simulink system model was excited with broadband noise up to 300 Hz to simulate engine-induced and airframe-induced cabin noise. The Simulink signal generator function block was used to generate a sine waveform at a frequency of 300 Hz. An 8th-order Butterworth low-pass filter was used to shape the frequency spectrum. The model also contains the disturbance or additional noise Simulink block to understand the effect of uncorrelated noise on the adaption process. The simulation is only modeling a single SISO system, primarily the one closest to the primary source (right side of cabin). In future work, Figure 3.20 below of the MIMO setup will be implemented, which uses the hardware from two SISO systems to model a localized and consistent zone of quiet around the passenger's ears. Modeling two separate single unit ANC setups (one for each ear of the passenger) was not practical because a MIMO approach will always yield better results in this case. The difference between the right and left sides of the cabin is that for the first case, there is a greater distance between the passenger and the primary source of noise. For the latter case, there is a small distance between the passenger and the primary source (if it were relocated to the left side of the cabin). Also, the assumption here is that the passenger seat is closer to the left side of the cabin particularly close to the window and cabin trim. The MIMO implementation would factor in the various types of sound pressure reflections with respect to the room walls and linings as well as the windows, floor, and ceiling. This would also require an elaborate acoustic finite element model of the cabin interior in order to visualize the natural modes of vibration and points of dominant pressures.
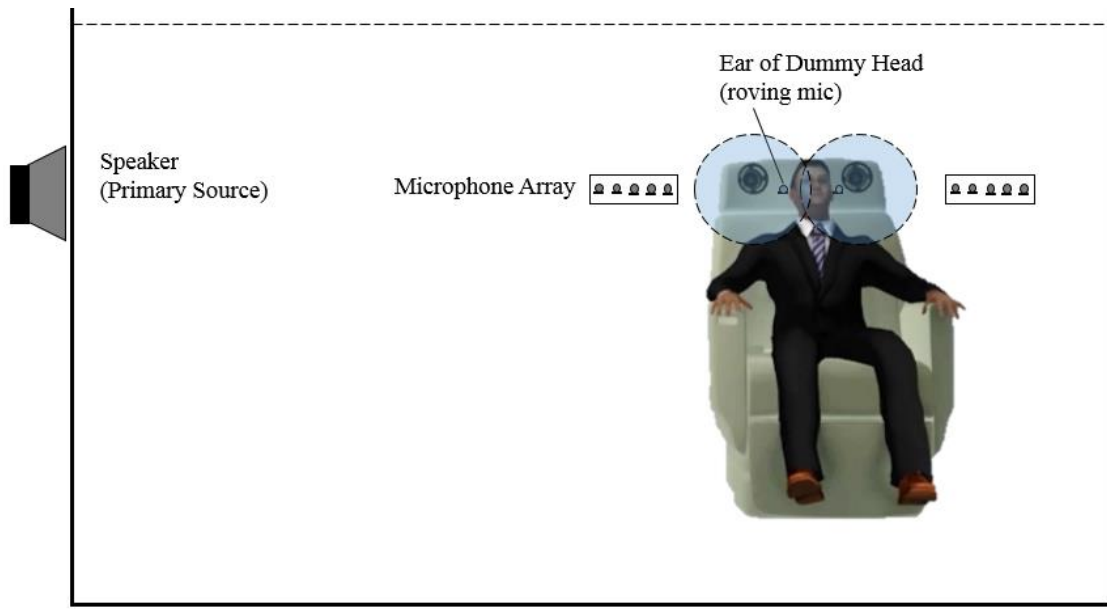
Fig. 3.19 SISO ANC setup (uncoupled ANC units near headrest).
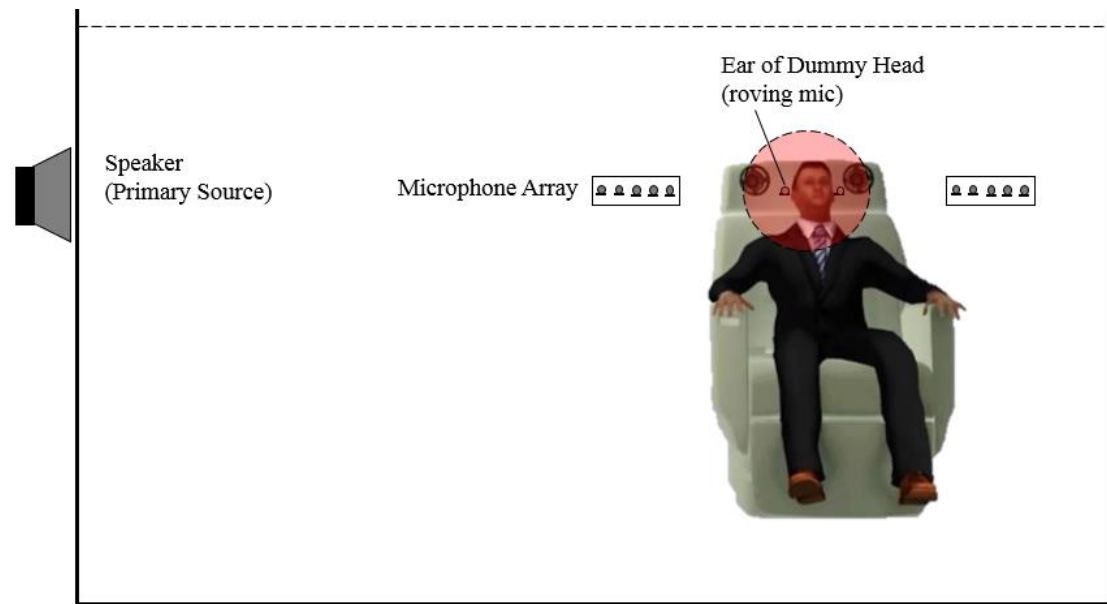
Fig. 3.20 MIMO ANC setup (coupled ANC units near headrest).

3.10. Results

In order to produce the results below, a logarithmically decreasing convergence coefficient, μ, was used (see Figure 3.21). Initially, the value of μ was 2.5e-08 at time zero and decreased to 2.5e-11 at the final simulation time of 0.5 seconds. The altering of the convergence coefficient was to ensure that convergence was reached in a short amount of time in order to yield accurate weight results. The final weights along with the corresponding normalised RMS errors for the *'2h'* and *'4h'* virtual locations are presented in Tables 3.1 and 3.2 below. These errors are the estimates of the sound pressures or desired signals of cancellation at the virtual location x. In comparing the errors amongst the virtual sensing methods for both virtual locations, the 5-microphone quadratic forward difference prediction sensor yielded the best results. As expected, all virtual sensors yielded significantly more accurate estimates than the closest physical microphone (microphone 5). For the adaptive LMS method, it seems beneficial to use more than 2 microphones, however additional sensors for the second and third order forward difference configurations do not have an effect on the system performance. Future work will include the adaptive LMS moving virtual microphone technique.

Fig. 3.21 Plot of decreasing convergence coefficient used for simulation.

Table 3.1 Final weights and normalised RMS errors of the adaptive virtual microphone vs. fixed weight virtual microphones for a virtual location of '*2h*'.

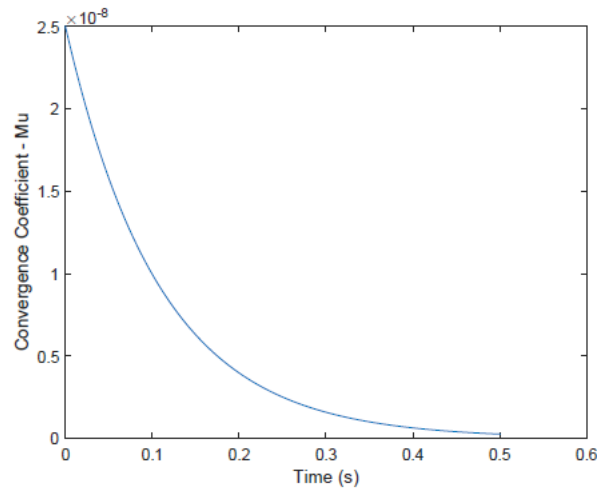| Virtual Sensing Method	Number of Mics	Weights					RMS Error
		1	2	3	4	5	
Adaptive	2	-0.1215	-	-	-	-0.1554	0.0497
	3	-0.1514	-	-0.1742	-	-0.1966	0.0479
	5	0.1932	0.2114	0.2273	0.2407	0.2509	0.0449
Fixed – Closest Mic	1	-	-	-	-	1	0.2652
Fixed – Linear	2	-2	-	-	-	3	0.0352
	3	-2.17	-	0.33	-	2.83	0.0395
	5	-1.8	-0.8	0.2	1.2	2.2	0.0395
Fixed – Quadratic	3	10	-	-24	-	15	0.0199
	5	12.2	-7.8	-13.8	-5.8	16.2	0.0199

Table 3.2 Final weights and normalised RMS errors of the adaptive virtual microphone vs. fixed weight virtual microphones for a desired virtual location of '*4h*'.

Virtual Sensing Method	Number of Mics	Weights					RMS Error
		1	2	3	4	5	
Adaptive	2	-0.1039	-	-	-	-0.1665	0.0509
	3	-0.1385	-	-0.1777	-	-0.2163	0.0497
	5	0.2037	0.2116	0.2173	0.2205	0.2211	0.0469
Fixed – Closest Mic	1	-	-	-	-	1	0.3437
Fixed – Linear	2	-2	-	-	-	3	0.0484
	3	-2.17	-	0.33	-	2.83	0.0424
	5	-1.8	-0.8	0.2	1.2	2.2	0.0424
Fixed – Quadratic	3	10	-	-24	-	15	0.0198
	5	12.2	-7.8	-13.8	-5.8	16.2	0.0198

3.11. Verification

3.11.1.   Practical Implementation

The zone-based MIMO ANC experiment will implement the filtered-x LMS algorithm. The Texas Instruments TMS320C6713 DSP was selected as a suitable controller to satisfy the control requirements. The DSP uses C code to determine the control parameters for the hardware configurations. The following is a methodology first suggested by Jason R. Griffin [22] which can be implemented on the controller.

1. Select the number of reference sensors, error sensors, and actuators to be used in the experiment.

2. Select the sampling frequency, system ID coefficients, and control path coefficients. The system ID coefficients are the number of FIR filter coefficients that are used to build the model of the plant $\hat{T}_{ce}(z)$, through which the reference signal $x(n)$ is filtered. The control path coefficients are the number of coefficients present in the adaptive FIR filter $W(z)$ that are updated by the LMS algorithm to generate the secondary control signal.

3. Set the system coupling. A fully coupled system will model the control paths from a speaker to each of the error sensors, whereas an uncoupled system will model the control path from an actuator to a single error sensor. Typically the control results are best when the system is fully coupled.

4. The controller builds the plant models $\hat{T}_{ce}(z)$, or a matrix of transfer functions, by sequentially driving each speaker and measuring the response at all of the error

microphones (for example, $\hat{T}_{ce}(z)$ will be a 4 x 4 matrix if a system has 4 reference sensosrs and 4 speakers).

5. The controller then drives the speakers such that the error signals are minimized. The user can control the rate of convergence by adjusting the magnitude of the convergence coefficient μ.

6. The user can monitor the error signals and adaptive filter coefficients to access when maximum control is achieved. Also, the level of the control signals and reference signals can be monitored to assure that the signals do not clip.

7. A Control ON/OFF button can be toggled so that the control performance can be monitored and acquired using a data acquisition system.

Figure 3.22 is a block diagram of the physical implementation of the integrated ANC control system. The A/D (analog-to-digital converters) and D/A (digital-to-analog converters) are actually embedded in the TMS320C6713 controller, so no additional hardware is required. All the control channels must use low-pass filters in order to avoid aliasing (distortion or misidentification of desired signals) and high frequency noise. The reference sensors x_K are mounted on the physical system in order to provide the controller with signals which yield the disturbance characteristics. The reference sensors are usually accelerometers for vibrating structures and depending on the amplitudes of vibration, the signals may need amplification before being fed to the controller. The error sensors e_M measure the superposition of the radiated sound field created by the disturbance and the secondary sound field created by the controller-driven speakers.

The controller uses the error signals and the filtered references $\hat{x}_{ML}(n)$ to adapt a set of control filters that drive the speakers. The error sensors are microphones in ANC applications

76

since the objective is to reduce radiated sound. The filtered-x LMS controller first controls the dominant portion of disturbance spectrum (noise). The signals at the error sensors then start to adapt the controller, so their measured response can be filtered to target a particular bandwidth (see Figure 3.22 below with two back-to-back low-pass and high-pass filters for the error microphones) however, the signals at the error sensors adapt the controller, so their measured response can be filtered to target a particular bandwidth. The speaker signals u_L generated by the controller also require amplification because the input and output limit of the DSP are $\pm$ 5.0 volts. All of the error sensors, speakers, amplifiers and filters used in the sound field measurement and generation of control signals are components of the plant model. The following section will look at the ANC experiment using the controller and internally-generated signals. The physical implementation is not in the scope of this work.
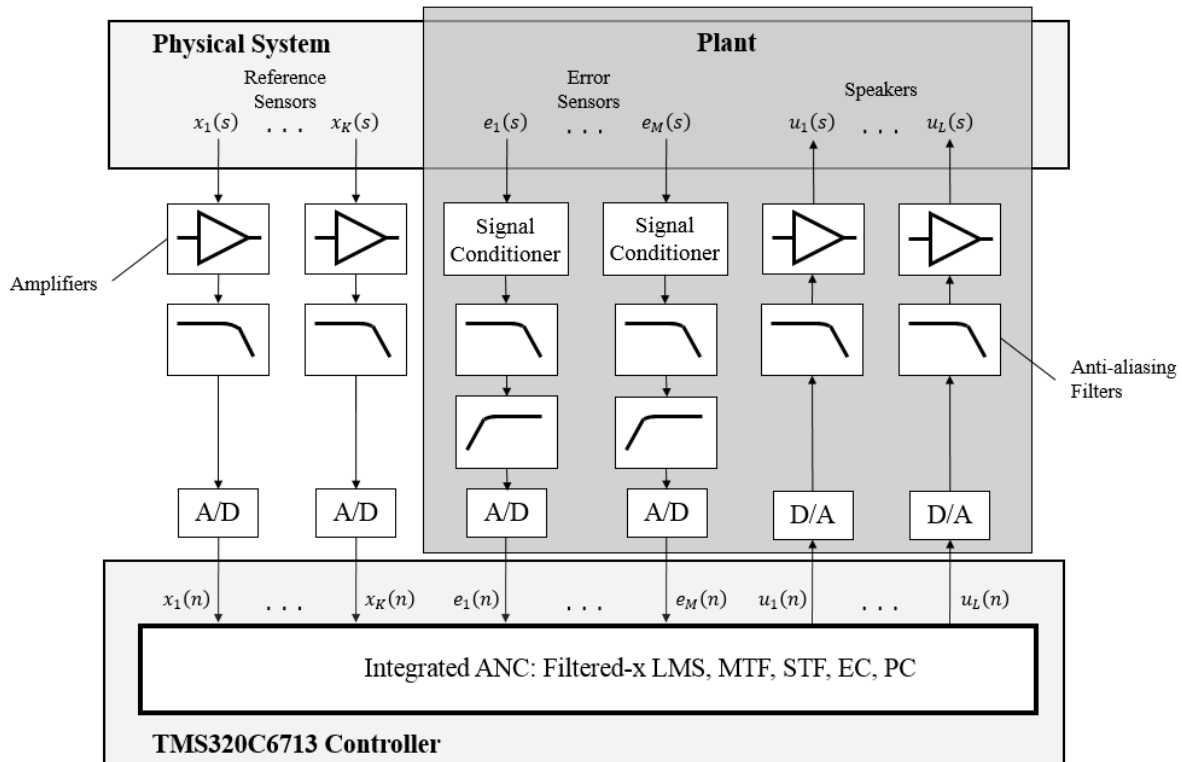

Fig. 3.22 Block diagram of MIMO control setup.

3.11.2. Active/Passive Noise Control Experiment

Referring to Equation 3.5, if the primary source pressure, **d**, is zero, then there is little need for the introduction of a secondary source pressure, **y**, into the control system. The primary source pressure, **d** can be driven to zero or close to it by employing conventional active/passive noise control methods. Initially, a smart-foam active/passive approach was to be used for noise control. The difference between this and conventional active noise control is that the active actuation uses piezoelectric films underneath the acoustic foam. The films are composed of polyvinylidene fluoride (PVDF) which is a highly reactive thermoplastic. This material is capable of generating an electric charge when excited. Similar to changing the location of the virtual microphones in the virtual microphone technique, the contour of the PVDF film can be changed to minimize the sound pressure error between the primary and secondary sources. The active element of this smart foam method was not implemented, because the application to the headrest and passenger seat was found to be intrusive to the overall personal passenger comfort experience. Figure 3.23 below demonstrates the use of the smart foam between the reference signal, $\mathbf{x}(n)$, and the primary source pressure response, $\mathbf{d}(n)$. The reference signal is usually obtained from accelerometers, but microphones can be used as well. In the passive control experiment which was conducted, electret microphones were used to obtain the reference signals on either side of the head.
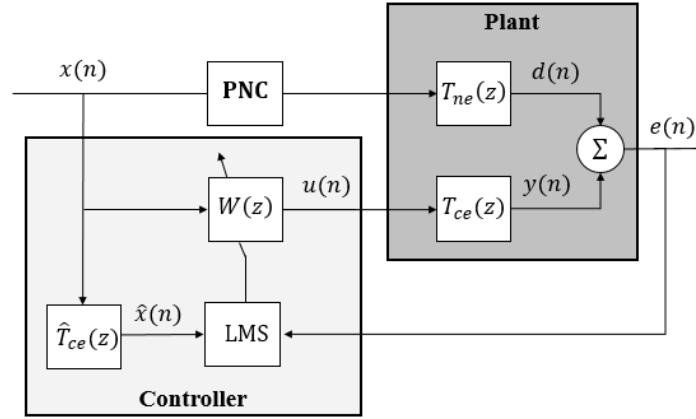
Fig. 3.23 Block diagram of adaptive LMS method with passive noise control element.

Two foam blocks were fixed to the sides of the headrest, wherein the microphones and necessary wiring were embedded safely and securely (see Figure 3.24). For the first reference model configuration the microphones were placed at 90 degrees from each, and both microphones were facing forward for the second configuration. This was done to ensure the concept of SPL omni-directionality of the microphones. These two cases are referred to perpendicular microphone setup and parallel microphone setup, respectively. The seat system was rotated 360 degrees (counter-clockwise) at 45 degree increments. The starting or reference position is 0 degrees which corresponds to a forward-facing setting similar to a normal flight condition. A diagram of the noise reference model setup is provided in Figure 3.25.
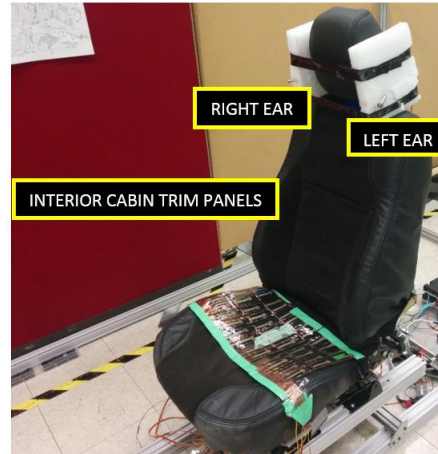
Fig. 3.24 Initial setup of Passive Noise Control.
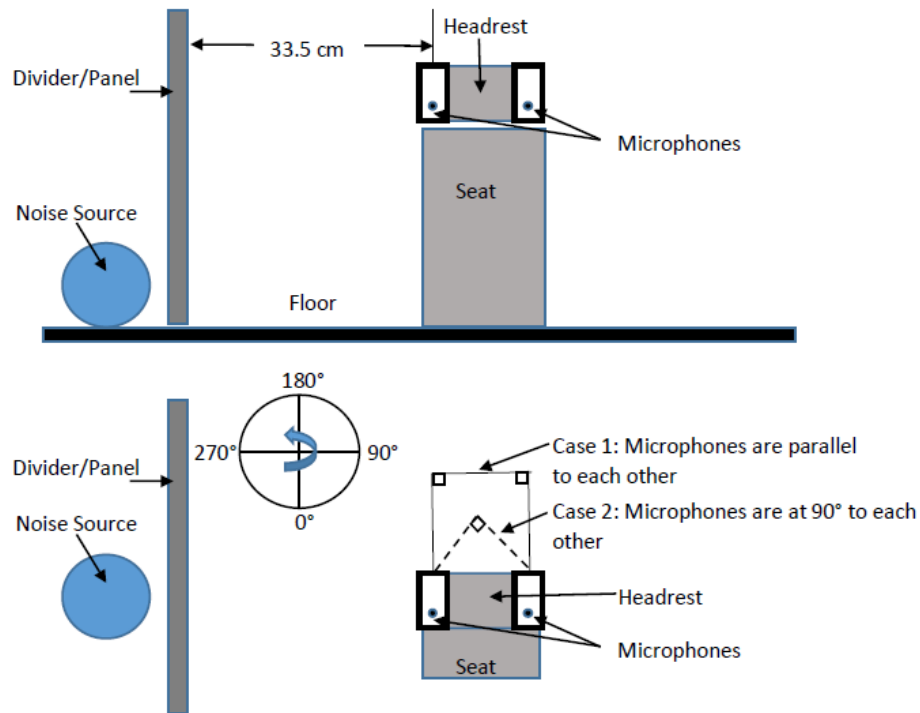
Fig. 3.25 *Top:* Front view of setup; *Bottom:* Top view of setup showing rotation about vertical axis.

A horizontal Campbell Hausfeld 3-Gallon air compressor was used to generate the noise near

the seat (see Figure 3.26 below). The air compressor was situated on the right side of the seat

directly behind a padded vertical panel and metal grid-frame. The initial tests were conducted

during 20 second time frames. This noise source was fixed, while the seat system was rotated or swiveled about the vertical or z-axis.

Fig. 3.26 Compressor used to generate experimental noise.

For ease of installation, the parallel microphone configuration was used for demonstrating the extent of passive noise control. A semi-enclosed enclosure was built around the headrest and microphones. Two adjustable semi-rigid arms were fixed to wooden blocks and secured to the headrest. The other end of the arms were attached to thin curved sheets of aluminum via nuts and bolts. This setup is pictured in Figure 3.27 below.
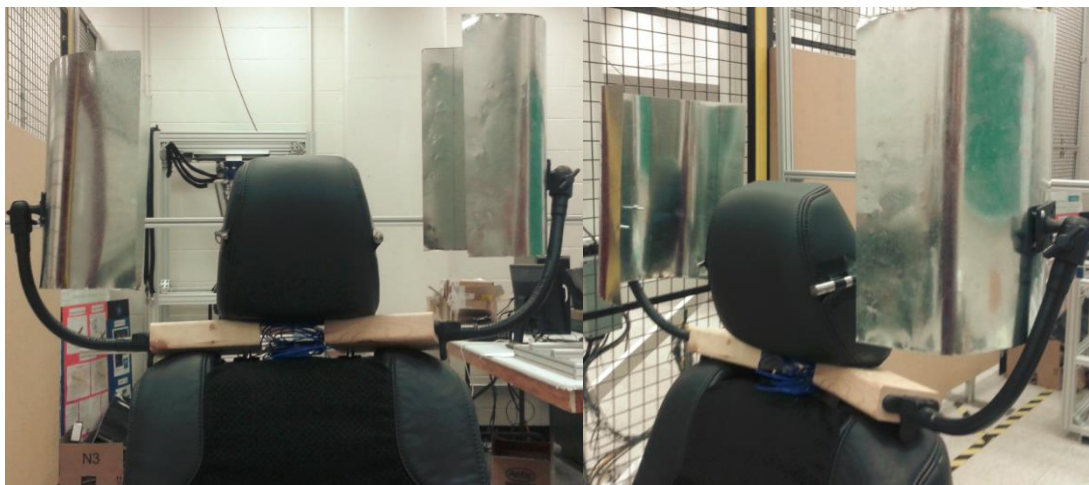
Fig. 3.27 Aluminum sheet curved panels for Passive Noise Control housing.

In order to provide adequate passive noise control in the immediate environment around the headrest, acoustic-grade foam wedges were implemented. The 2" (thick) x 2" (wide) Auralex Studiofoam® Wedges are meant to acoustically treat small areas such as vocal booths and control rooms. The anechoic wedge cut maximizes sound absorption while properly reducing standing waves (see figure 3.29). Each section was secured to the inside of the aluminum panels using 3M™ Supper 77™ Multipurpose Spray Adhesive. Similar to the noise reference model cases, the seat system was rotated 360 degrees (counter-clockwise) at 45 degree increments. The starting or reference position is 0 degrees which corresponds to a forward-facing setting similar to a normal flight condition. Images of the passive noise control setup are provided in figure 3.28 below.

Fig. 3.28 Passive Noise Control setup with curved acoustic-grade foam wedges.
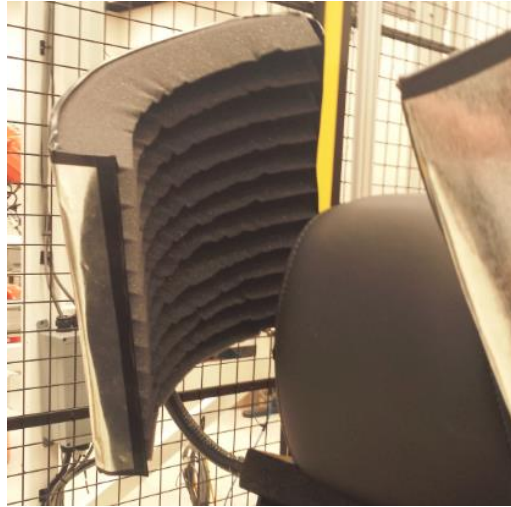
Fig. 3.29 Close-up view of acoustic foam wedges.

The omni-directionality results for the perpendicular and parallel microphone cases determined that the microphone angular anchor points do not make a significant difference in measuring the sound pressure level of the surrounding environment. The sound level plots for the left and right ears for the perpendicular and parallel microphone configurations in Figures 3.30 and 3.31 validate this assumption. The means of the differences for the left ear and right ear reference model plots are approoixmately 3 dB and 2 dB, respectively. The passive noise control results yielded decibel reductions of up to 14 decibels, however the average of the results were 5 dB of reduction for the left ear and 8 dB of reduction for the right ear (see Figures 3.32 and 3.33). This was expected for passive noise control methods, which is why active noise control techniques were investigated next.

Fig. 3.30 Noise Plot for Left Ear reference model.

Fig. 3.31 Noise plot for Right Ear reference model.
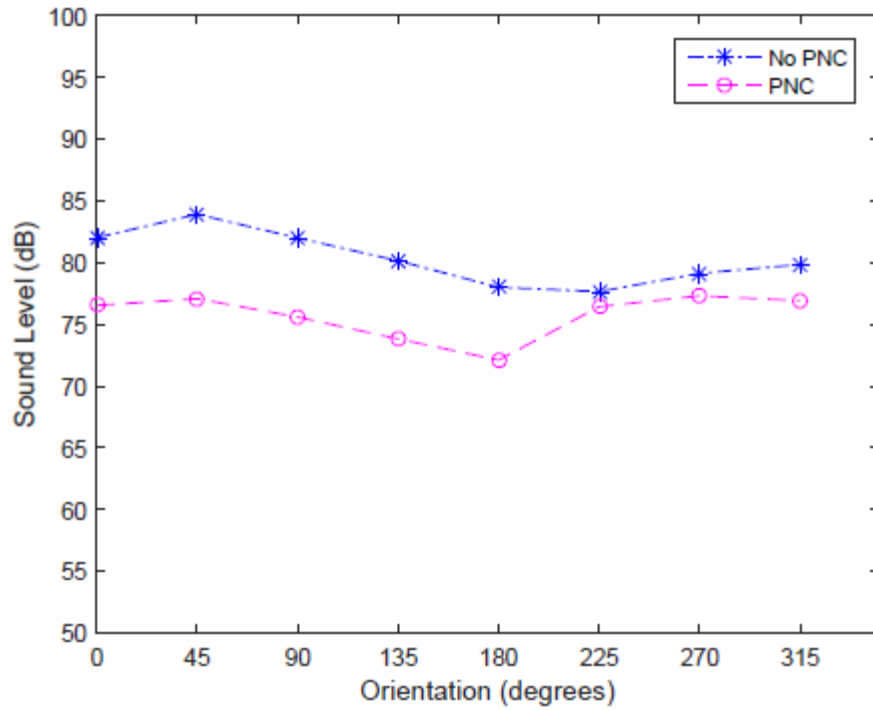
Fig. 3.32 Noise plot for Left Ear parallel ref model: without vs. with Passive Noise Control.
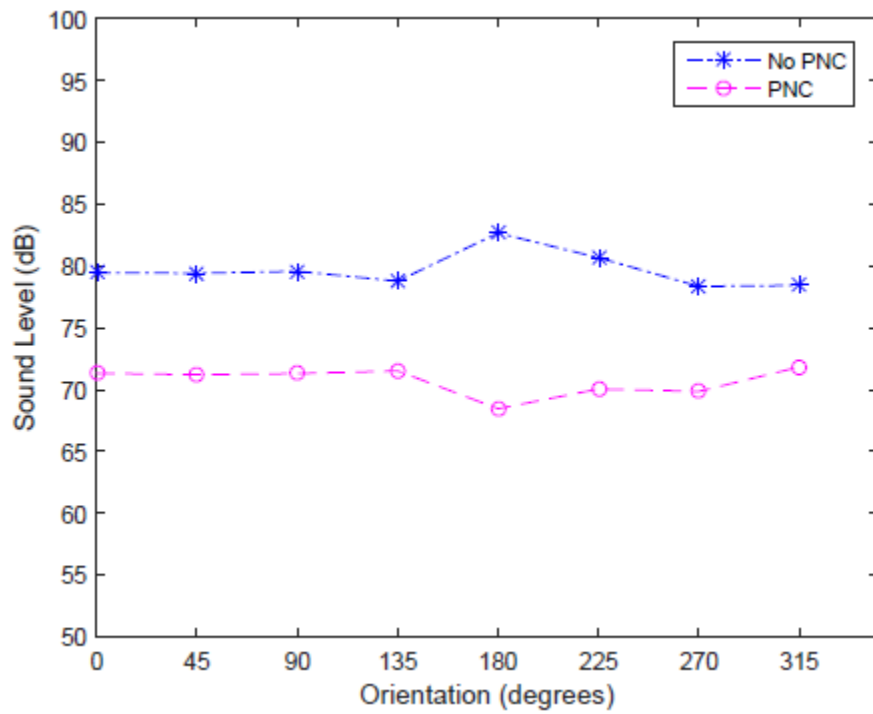
Fig. 3.33 Noise plot for Right Ear parallel ref model: without vs. with Passive Noise Control.

3.11.3.   Active Noise Control Experiment

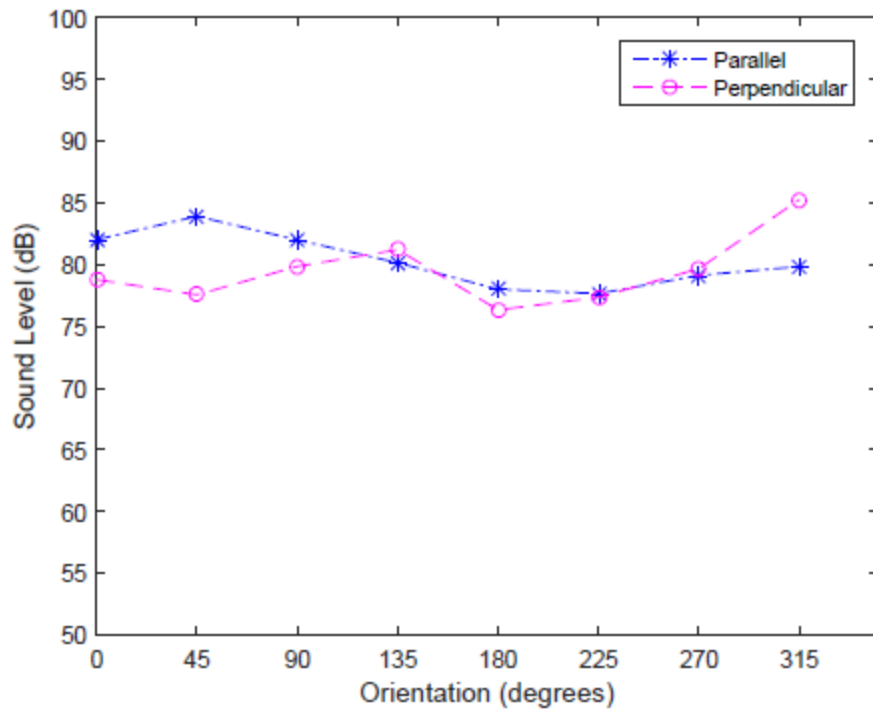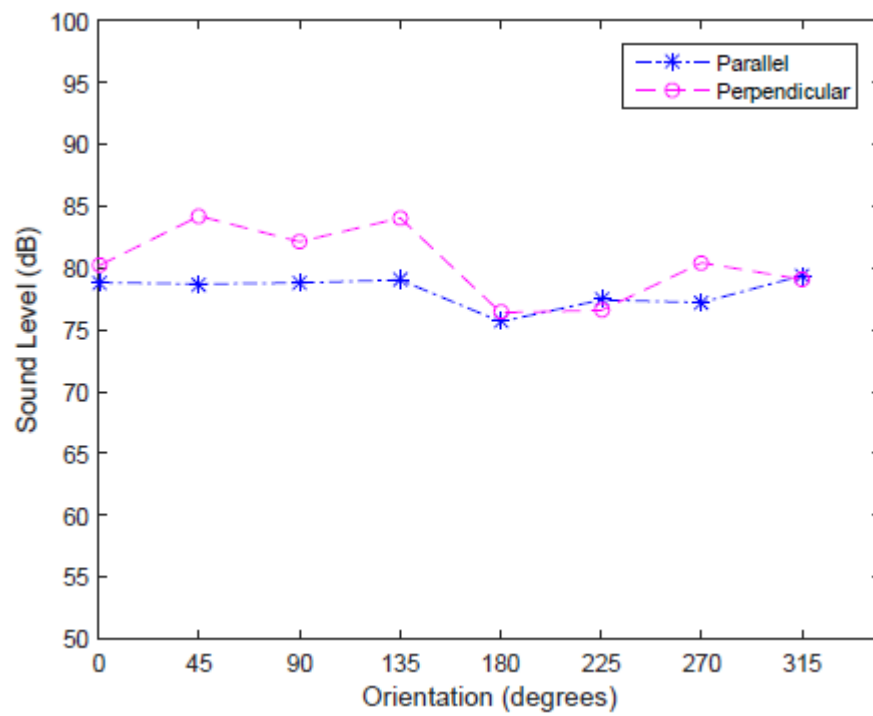The Texas Instruments TMS320C6713 DSP starter kit or DSK (refer to Figure 3.34) along with its audio interfaces, Line Out and Headphone Out, have been used to generate different waveforms and audio tones. Using its affiliated CCS software (Code Composer Studio), the TMS320C6713 DSP starter kit is a powerful tool that is not restricted by predefined signals, but can generate any desired signal by modifying the C-based source code. Some of the different features that have been tested include waveform generation, multitoned waveform generation, and pseudo random noise sequence generation. A highly relevant type of waveform generation is sine wave generation because it can be used to describe the behaviour of periodic, dominant anti-noise waves. On the other hand, multitoned waveform generation can be used to represent N1 & N2 aircraft engine tones. These topics are continually being investigated because CCS uses what is known as GEL (General Extension Language). GEL is an interpretive language, like C, which permits variable changes such as gain. Additionally, the variable values can be altered while the processor is running.

Fig. 3.34 Setup of TMS320C6713 DSK, multimedia speakers, and CCS software.

Adaptation Process using Adaptive Filter

The coefficients of an adaptive filter are updated to reflect the changes in the input signal, output signal, or system parameters. An adaptive filter is different from a fixed filter in the sense that it can self-learn the signal characteristics and any slow signal or system changes. A basic adaptive filter structure consists of the adaptive filter's output $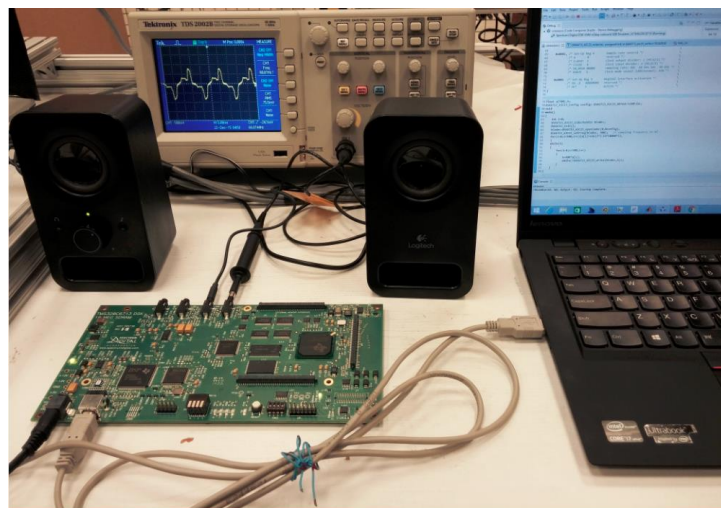y$ which is compared (or subtracted from) a desired signal d to calculate an error signal e, which is used as feedback to the adaptive filter. This is the least-mean-square (LMS) algorithm with a linear combiner (FIR filter).

here $w_k(n)$ represents N weights for a given time n. In order to determine the effectiveness of the adaptive filter, the error signal is used as the main performance measure. This error signal $e(n)$ is the result of the difference between the desired signal $d(n)$ and the adaptive filter's output $y(n)$. The minimization of the mean squared error function is used to adjust the weights or coefficients $w_k(n)$. The mean squared error function is $E[e^2(n)]$, where the expected value is denoted by E. The gradient of the mean squared error function is required, because there are k weights or coefficients. However, an estimate can be determined using the gradient $e^2(n)$ instead. The reference input to the adaptive filter is $x(n)$, and the rate of convergence or adaptation (adaptive step size) is represented by β.

During each time interval n, each weight or coefficient $w_k(n)$ is replaced by a new coefficient $w_k(n+1)$. However, if the error signal $e(n)$ is zero, the weight is not updated. The filter's output $y(n)$ and error signal $e(n)$ are also updated for a specific time n. After all of these updates occur for a specific time n, a new sample is extracted from the ADC, and the whole adaptation process is repeated for a different n. The linear combiner or finite-impulse response (FIR) filter is a common and useful adaptive filter structure partly because of its adjustability. The

coefficients of non-adaptive infinite-impulse/frequency-selective response (IIR) filters are fixed, whereas the coefficients of FIR filters can be altered using changes in signal environment such as the input signal. Adaptive IIR filters can also be used, however since their poles can be updated during the adaptive process to values outside of the unit circle, these filters can be somewhat unstable. The following steps are used for the LMS adaptation process using the adaptive filter structure:

1. Obtain a new sample for each, the desired signal $d(n)$, and the reference input to the adaptive filter $x(n)$, which represents a noise signal.

2. Determine the adaptive FIR filter's output $y(n)$.

3. Determine the error signal $e(n)$.

4. Update each weight or coefficient $w_k(n)$.

5. Update the input data samples for the next time n, with a data move scheme used in Chapter 4 of [13]. The data move scheme does exactly that – it moves the data instead of a pointer.

6. Repeat the adaptive process (steps 1 through 5) for the next output sample point.

There are two main differences with this adaptive filter structure: the desired signal d is corrupted by uncorrelated additive noise n, the input to the adaptive filter is a noise n' that is correlated with the additive noise n. The reference noise input n' could come from the same source as n but is altered by the surrounding environment. The adaptive filter's output y is adapted to the noise n (approaches n with time). In this case, the error signal e approaches the desired signal d. Similar to the general adaptive filter structure, the overall adaptive filter output is this error signal and not the adaptive filter's output y. There are other variants of the LMS algorithm such as the sign-error, sign-data, and the sign-sign LMS algorithms. However the TMS320C6713 Texas Instruments digit signal processor (controller) works fastest for the basic LMS algorithm. The

execution speed for the DSP using these variants is slower because of additional decision-type instructions required for testing conditions involving the sign of the error signal or the data sample.

Adaptive Filter for Noise Cancellation

N1 and N2 are aircraft jet engine RPMs or rotational speeds. These speeds have associated engine tones which sound at $40 - 100$ Hz and $100 - 200$ Hz respectively. In a business jet aircraft, these tones are generally the main sources of noise and the target for this active noise cancellation research. The first step in selecting the frequency range for cancellation was to start with a narrowband of a low-frequency tone. For demonstration purposes, the LMS algorithm was used to cancel an undesirable sinusoidal tone of 312 Hz frequency. A desired sine wave of 1500 Hz with an undesired sine wave noise of 312 Hz was one of two inputs to the adaptive filter structure. A reference cosine wave signal of 312 Hz (which is only correlated with the additive or undesired sine wave) is the second input to the adaptive filter structure. The FIR filter itself contains 30 coefficients. At each time interval n, the output of the adaptive filter is calculated and all the weights or coefficients are adjusted. The overall output of the adaptive filter is the error signal E which in this case is the difference between the desired signal and additive noise denoted $dplusn$, and the adaptive filter's output, $y(n)$.

Although the TMS320C6713 digital signal processor (controller) is used for this demonstration, all two of the input signals to the adaptive filter have been generated from a lookup table using MATLAB. There are three signals that were generated: the desired sine signal of 1500 Hz, the additive sine wave noise of 312 Hz, and the reference cosine signal of 312 Hz. These lookup tables were used to provide the DSP with the following two files: 1. $dplusn$, sine(1500

Hz) + sine(312 Hz); 2. *refnoise,* cosine(312 Hz). The constant *beta* is used to represent the rate

of convergence of the adaptive filter output. The following four plots in Figures 3.35 through 3.38

demonstrate the undesired 312 Hz sinusoidal signal being slowly reduced over time, while the

desired 1500 Hz is preserved. As part of the filter design process, different numbers of weights as

well as different betas were experimented with in order to determine the ideal adaptive filter. The

convergence rate beta can be increased for a faster rate of cancellation, however if it is too high

then the output becomes unobservable. In Figure 3.37, the beta used was 10 which was several

orders of magnitude higher than the first two cases. It can be seen that the output quickly converges

to the desired sinusoidal signal. Figure 3.38 shows the error plots of the two convergence rates

with 30 weights. The difference, in this case, between the two plots is on the order of milliseconds.
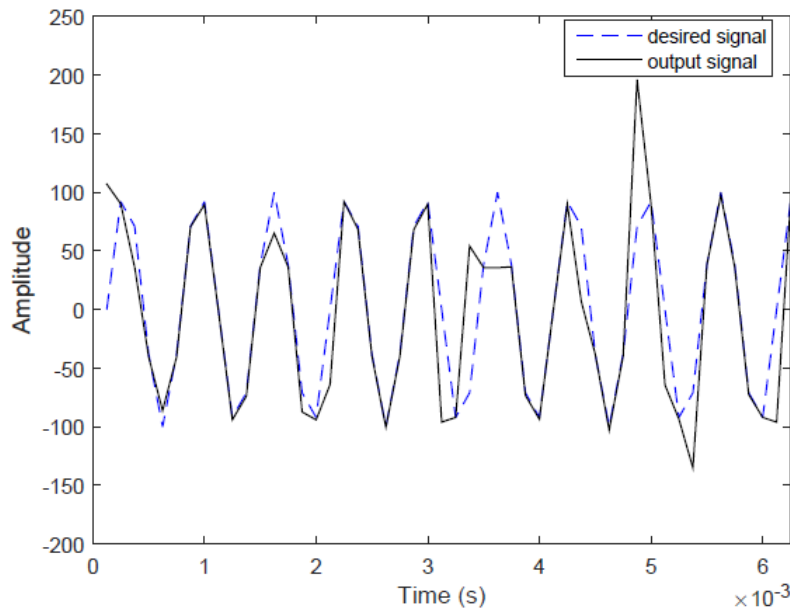

Fig. 3.35 Plot of adaptive filter's output converging to desired signal: 30 weights, beta 1e-10
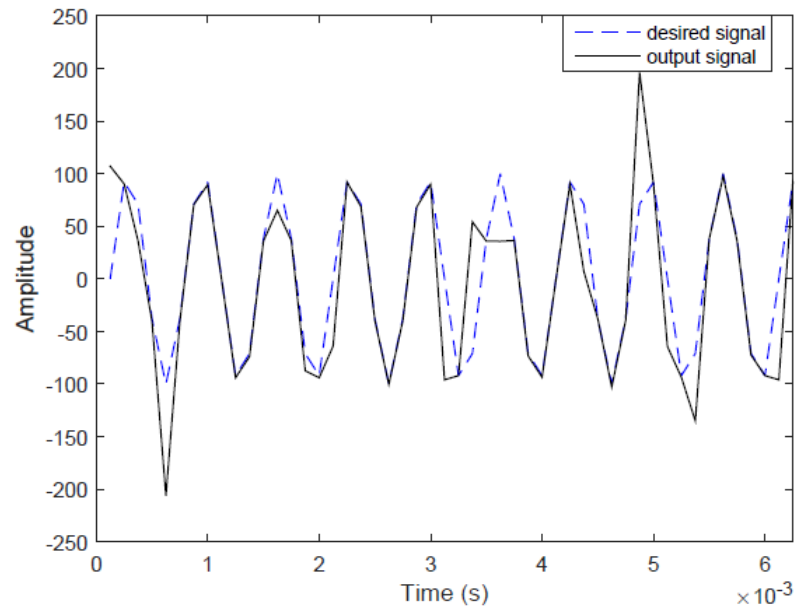
Fig. 3.36 Plot of adaptive filter's output converging to desired signal: 60 weights, beta 1e-10
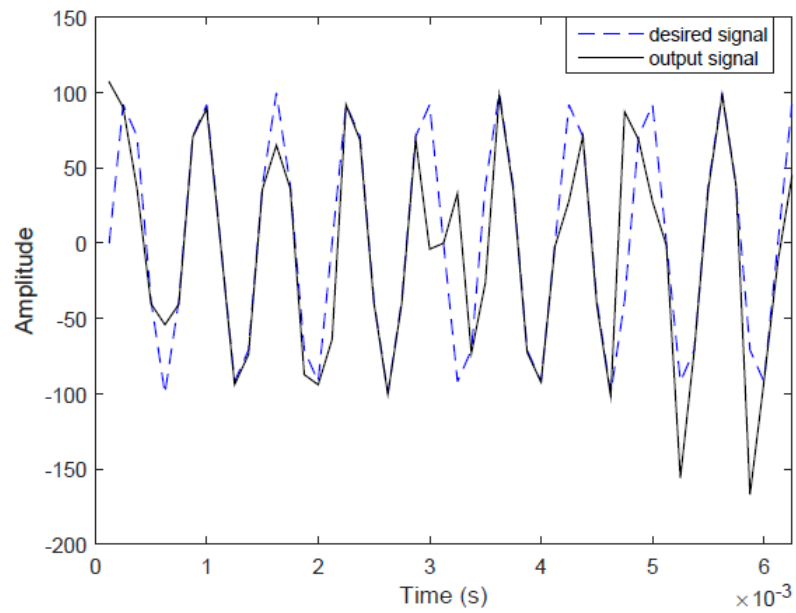
Fig. 3.37 Plot of adaptive filter's output converging to desired signal: 30 weights, beta 10

91
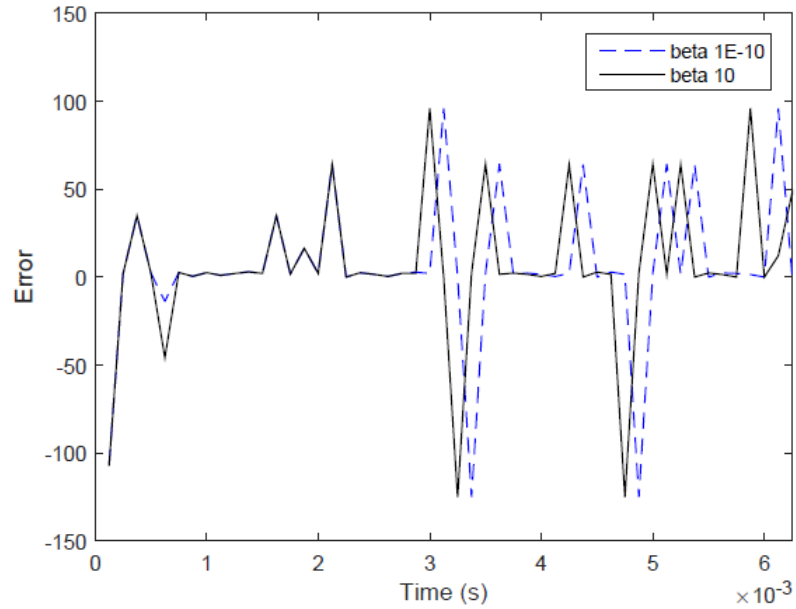
Fig. 3.38 Plot of adaptive filter's error signal for the different betas: 30 weights

Adaptive Weighted FIR for System Identification of Fixed FIR

The adaptive LMS method uses a fixed FIR bandpass filter with 55 coefficients centered at 150 Hz, which models the known fixed bandpass filter. It also used an adaptive FIR bandpass filter with 60 coefficients centered at 1 kHz. This adaptive FIR bandpass filter models the fixed unknown FIR bandpass filter that the adaptive scheme will try to identify. The coefficients of the FIR filters are determined using MATLAB's Filter Design & Analysis (FDA) Tool. Images of the MATLAB FDA tool windows used in the design of the two filters are provided in Figures 3.39 and 3.40 below. The filter coefficients are then exported to Code Composer Studio IDE header files (see Figure 3.41). The filter coefficients are extracted from the header files and stored in .cof CCS executable files. The first file with the fixed filter coefficients centered at 150 Hz is used in the main CCS program *adaptIDFIR.c* to initialize the weights of the adaptive FIR filter with the

92

weights of the known fixed filter. The program *adaptIDFIR.c* is a modified version of that provided in the TMS320C6713 digital signal processor starter kit or DSK.

Also provided in this DSK is a pseudorandom noise sequence *noise_gen.h*, which becomes the input to both the fixed and adaptive FIR filters. The noise input signal is referred to as a training signal. The error in this case is the feedback error signal which represents the difference between the outputs. The adaptation process is complete when the error is minimized. The two output variables are the output of the fixed unknown FIR filter *fir_out* and the output of the adaptive FIR filter *adaptfir_out*. Initially, during the first few iterations of running the program, the filter output exactly matches the known filter's output centered at 150 kHz. However, within milliseconds, the adaptive filter's output gradually converges or adapts to the desired filter centered at 1 kHz. Again, increasing the convergence rate beta was found to accelerate the adaptation process. The chosen weights or filter coefficients, 50 and 60 for the modelled known and unknown filters respectively, were found to successfully meet the noise cancellation requirements.

The spectrum of the adaptive filter's output can be seen in Figure 3.42. The original or fixed filter centered at 1 kHz is also in the figure and is represented by dashed lines. When comparing the two spectrums, it can be seen that the adaptive filter contains less of the known fixed filter centered at 1 kHz and significantly more of the unknown fixed filter centered at 150 Hz. The error or difference between the known fixed filter and adaptive filter outputs is plotted in Figure 3.43. The error is measured in decibels and has a maximum value of approximately 33 dB and a mean of 17 dB. This error represents the overall active noise reduction and falls within acceptable ANR ranges which are generally between 15 and 20 dB for state-of-the-art noise-cancelling technologies. For the desired frequency range of 40 to 200 Hz, the noise reduction is greater than 15 dB and peaks at 200 Hz. However, it should be noted that the reference noise signal and fixed

93

filters were internal signals to the DSP. This means that a real or external application case may result in a slight degradation of the adaptation scheme. Factors to consider are the environmental uncorrelated and unpredictable noise as well as the complexity of the filters, which will require a faster and more powerful DSP than the TMS320C6713.
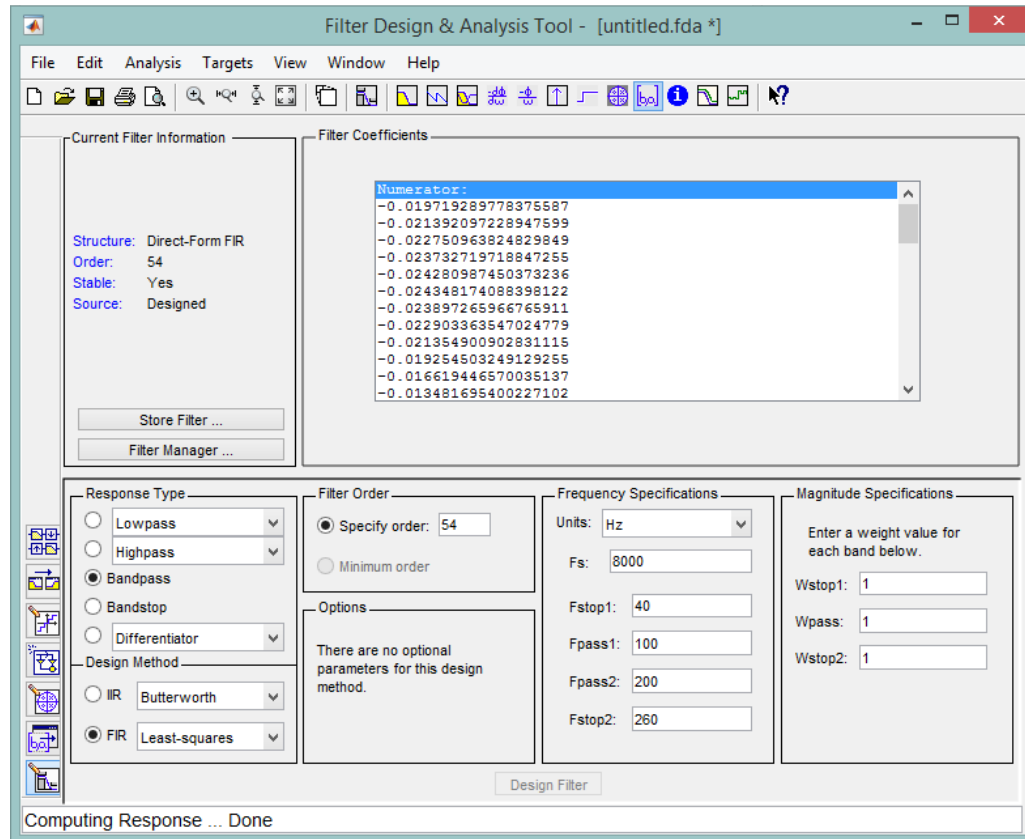
Fig. 3.39 Design of 150 Hz – centred bandpass filter using MATLAB's filter design and analysis tool.

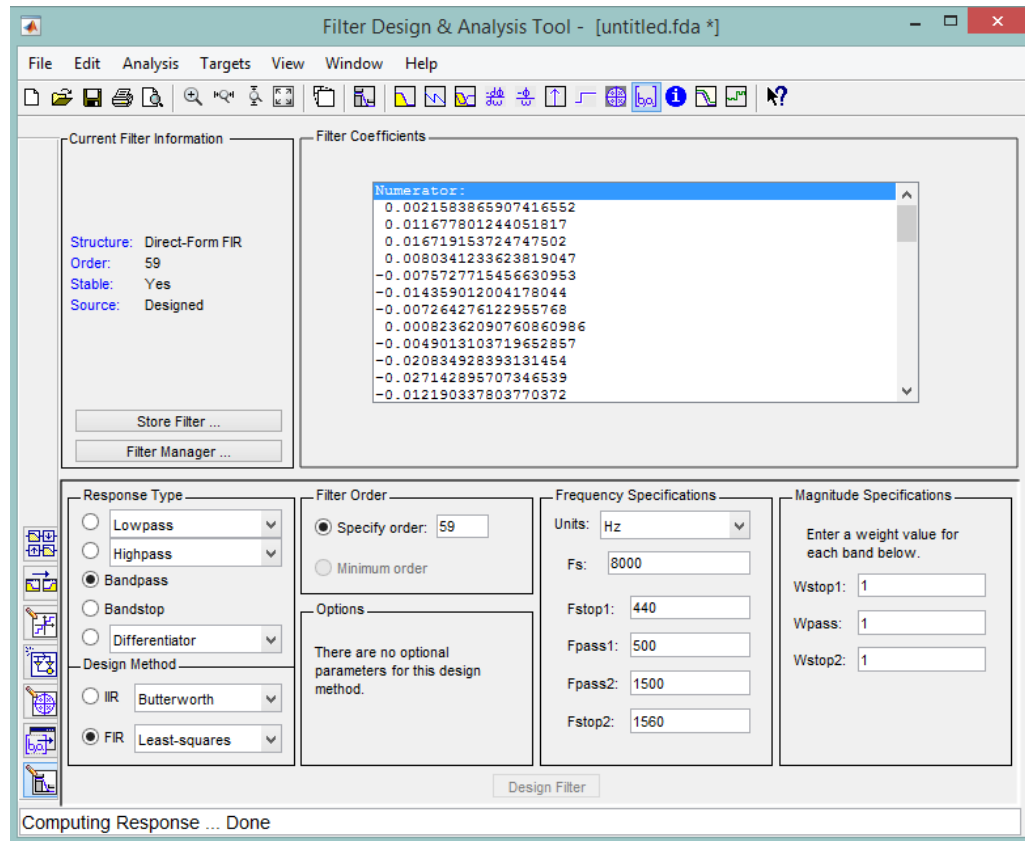**Fig. 3.40** Design of 1 kHz – centred bandpass filter using MATLAB's filter design and analysis tool.

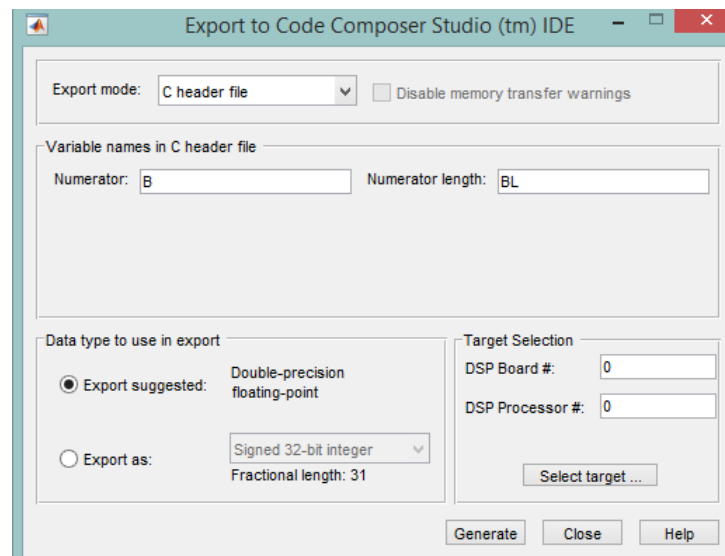

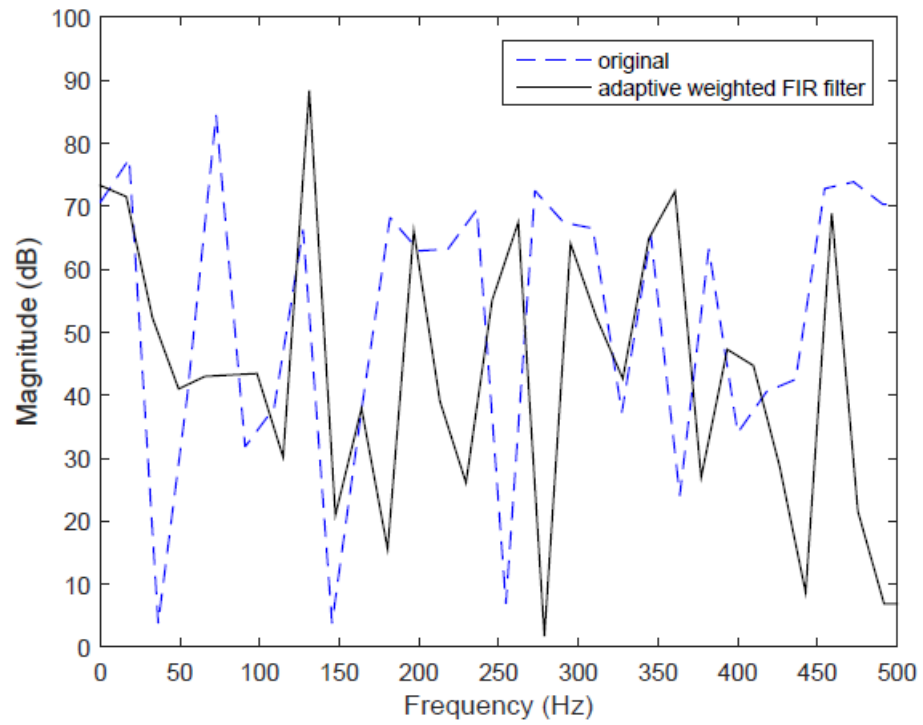**Fig. 3.41** MATLAB's filter coefficient export to CCS IDE feature.

Fig. 3.42 Spectrum plots of original fixed filter and adaptive filter outputs: beta 1e-13
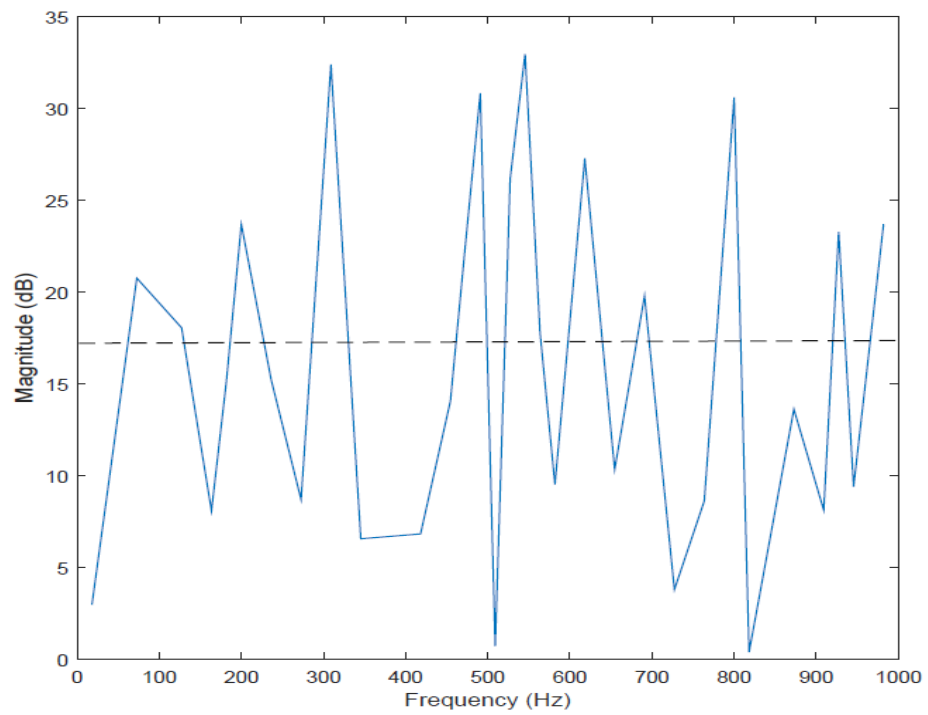

Fig. 3.43 Plot of adaptive filter's active noise reduction measured in dB.

Data Acquisition

The electret microphone amplifier MAX4466 with adjustable gain was used for audio sampling. The gain can be modified from 25x to 125x and this allows the output to range from 200 mV_{pp} and up to1 V_{pp}. The default value is a power gain of 25 times (approximately 28 dB) the power amplification of the input. The output pin can be connected directly to the Mic In jack on the TMS320C6713 board. An external power supply is used to power the sensor with up to 5 V. The best operating voltage for the quietest setting (least amount of noise) is 2.4 V.

The TMS320C6713 digital signal processor has a 16-bit analog-to-digital converter (ADC). The data is stored in the registers as Uint16 which means 16-bit unsigned integers. A 16-bit ADC will need a maximum value of 2^{16} for calibration and measurement functions. The maximum output voltage is created when the digital input word is equal to $2^N - 1$, which occurs when all 16 bits are equal to 1. The reference voltage V_{ref} refers to the step voltage or the voltage required to switch a single bit from 0 to 1. In order to convert these values into decibels, a voltage gain formula which incorporates the sensor's sensitivity or transfer factor as well as its gain value will be used.

The sensitivity or transfer factor demonstrates the use of microphones – conversion of sound pressure (Pa) to audio voltage (mV or V). The MAX4466 microphone sensor has a sensitivity of -44 dB and a corresponding transfer factor of 6.3096 mV/Pa. International standards have set 94 dB sound pressure level as the reference SPL equal to 1 Pa and this is usually denoted as 94 dB re. The symbol $'re'$ is an indication of said reference.

L_{eq} represents the equivalent continuous noise level. It is the most common method used to describe time-varying sound levels. This results in a single decibel value which considers the total sound energy over the desired period of time. The first step in calculating the equivalent

97

continuous noise level L_{eq}, is determining the output voltage. The output voltage denoted by V_{out} is found by dividing the digital-to-analog or DAC output of the TMS320C6713 DSP by the product of the maximum number of bits and the reference voltage V_{ref} (see Equation 3.39 below). The reference voltage V_{ref} was found to be 25 percent of the maximum voltage V_{max} (see Equations 3.40 and 3.41 below).

$$V_{out} = \frac{DAC_{out}}{2^N \cdot V_{ref}} \qquad (3.39)$$

$$V_{out} = \frac{DAC_{out}}{2^{16} \cdot 0.25 V_{max}} \qquad (3.40)$$

$$V_{out} = \frac{DAC_{out}}{65,536 \cdot 0.25 \cdot 5} \qquad (3.41)$$

$$V_{out} = \frac{DAC_{out}}{81,920} \qquad (3.42)$$

Once the output voltage V_{out} has been computed, the equivalent continuous noise level L_{eq} can be readily determined. The equation for calculating L_{eq} is presented in Equation 3.43. The first part of the equation determines the power change from the sensor, while the second two terms incorporate the sensor's sensitivity as well the reference dB level. The last term subtracts the power gain of the adjustable-gain sensor, which according to Equation 3.44 is approximately 28 dB. The final equation for quantifying the equivalent continuous noise level L_{eq} is presented in Equation 3.45 below.

$$L_{eq} = 20log\left(\frac{DAC_{out}}{81,920 \cdot 0.0063096}\right) - 44 + 94 - 20log(25) \qquad (3.43)$$

$$L_{eq} = 20log\left(\frac{V_{out}}{516.88243}\right) + 50 - 27.95880 \qquad (3.44)$$

$$L_{eq} = 20log(0.0019347 \cdot V_{out}) + 22.0412 \qquad (3.45)$$

3.12. Summary

The quiet bubble technology which was discussed earlier is one of the most advanced in its field, however it only uses fixed virtual sensor locations. A significant limitation of this ANC technology is that the transfer function parameters can only be set once during the calibration and initial setup stage. This means that if a passenger turns their head, a different transfer function may be needed to capture the change in the sound field from the speakers to the error sensors (STF) or from the speakers to the reference sensors (EC). The disadvantage of this technique is that it is not reliably robust enough to changes in the acoustic environment. In order to deliver such a system, a head tracking algorithm in lieu with forward-difference prediction must be used to constantly update the prediction filter with a changing virtual sensor-to-observer distance. This will allow for the creation of a zone of quiet which is localised to the passenger's head movement. Similar to this technology, the following simulation only uses a LMS-based virtual sensing method, however it has the potential to include moving virtual sensors. Chapter 4 will look at viable head tracking methods for zone of quiet-based virtual sensors.

4. Head Tracking

4.1. Dummy Head Model

An acoustic dummy head was purchased from Binuaral Enthusiast to model the human head and its rotation about the vertical axis. The purpose of the following experiments is to detect and track the head angle with respect to the frontal view (head facing straight forward). The fiberglass dummy is threaded on the bottom, so it was easily secured to an MDF plate which sits onto a custom protractor. This assembly is secured to an aluminum strut later and is free to rotate between two nuts. At the centre of the neck of the dummy, a copper-wire needle is secured in order to physically position the head at given angles between -90 to 90 degrees. The dummy head is also equipped with silicone ears which contain ear canals. This is resourceful because it has allowed for the installation of two microphones in the ears.

Fig. 4.1 Acoustic Dummy from Binaural Enthusiast.

4.1.1. Creation of Humanistic Facial Features

Each feature (eyes, nose, ears, and mouth) were stickered using pieces of various solid colours of electrical tape. The eyes are represented by blue rectangles, the right and left ears by white rectangles, the nose by a yellow rectangle, and the mouth by a green rectangle. This colour masking technique was created in order to isolate the individual facial features from the face itself as well as the surrounding environment (i.e. headrest, surrounding room). Each piece of tape represents a particular facial feature and is matched to its corresponding parameters which include hue, saturation, and value or intensity HSV. The significance of the HSV scale to the head-tracking task will be explained next.

Fig. 4.2 Acoustic Dummy with facial feature stickers.

4.1.2. HSV Scale and Masking Process

The HSV scale is representative of the purpose of RGB, however it organizes the geometry in such a way so that it becomes more meaningful in comparison to its Cartesian or cube counterpart. The HSV scale is a cylindrical system where the hue is represented by the angle around the central

vertical axis, which is the azimuth or theta coordinate. The saturation is represented by the distance from this vertical axis which is the radial coordinate, and the value is represented by the distance along the axis which is the height coordinate. The value parameter is sometimes referred to as lightness or brightness. Table 4.1 below demonstrates the use of the HSV parameters and their corresponding lower and upper bound pixel values for each facial feature.

Table 4.1 Pixel Ranges for HSV Parameters.

HSV Bounds	HSV Parameters ([Hue, Saturation, Value])			
	Eyes	Nose	Mouth	Ears
Lower Limit	[103, 168, 54]	[33, 53, 131]	[78, 152, 57]	[82, 23, 200]
Upper Limit	[114, 255, 153]	[46, 253, 235]	[96, 255, 147]	[112, 69, 240]

The first step of the masking process is to represent the selected feature as a binary one which shows up on the display as white. On the other hand, all of the negatives or non-selected regions that do not fall into the HSV classification or range are binary zeroes and show up on the display as black. The secondary step in the masking process is to combine the ones and zeroes from each mask (one mask per facial feature) into a matrix format. This information is then displayed on a computer monitor or screen in the form of a white and black image. The final step is to combine the original RGB image with the masked image to create a display of the detected and coloured facial features. A sample of this combined RGB and masked image can be seen in Figure 4.3 below.
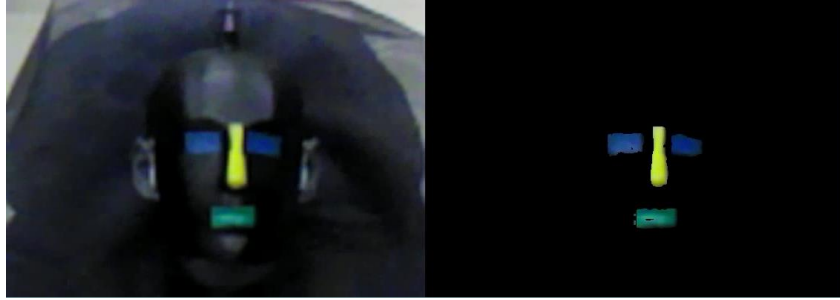
Fig. 4.3 Separated Colour Facial Feature Detection.

4.2. Rotation Method

4.2.1. Setup

The dummy assembly was fixed to an aluminum strut which was then secured to the seatback of a Bombardier Global Express 5000 aircraft seat. Directly in front of the seat, a camera tripod was mounted onto a desk to create a suitable and approximately-frontal field of view for the dummy head and the seatback. The Intel RealSense Camera F200 is mounted atop the tripod. The distance from the tripod to the nose of the dummy is 138 cm. Also, the distance from the floor to the top of the camera is 180 cm (see Figure 4.4 below).The F200 camera connects to a PC via a USB 3.0 port and cable.
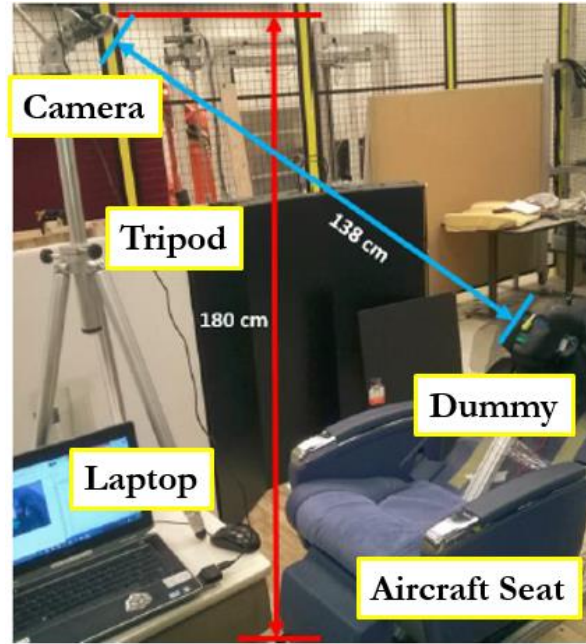
Fig. 4.4 Head-tracking setup with tripod, camera, and dummy placement.

Next, the dummy head is calibrated for each side or direction of rotation from the central or forward-facing view. Counter-clockwise rotation of up to 90 degrees is considered positive, whereas a clockwise rotation of up to 90 degrees is considered negative. This can be seen in Figure 4.5 below. For calibration purposes, the head is rotated in 5 degree increments from 0 to 90 degrees and similarly from 0 to -90 degrees. The visibility or the lack thereof of facial features will be discussed for the two aforementioned calibration cases.

Fig. 4.5 Dummy Head Rotation Schemes. *Left Turn:* 0 to +90°; *Right Turn:* 0 to -90°.

4.2.2. Calibration

The process of calibration for the dummy head rotation consists of obtaining pixel measurements for given and known angles of rotation. These values were tabulated and several polynomial models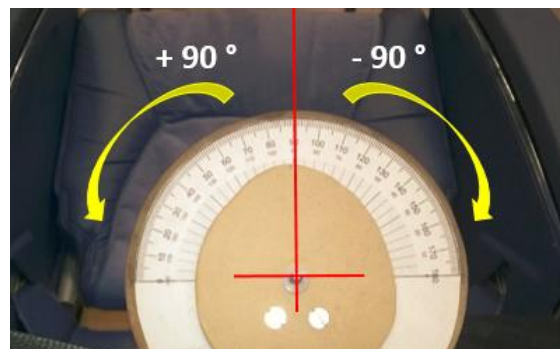 were fit to the data. The equations for the models are tabulated in Table 4.2. It can be seen that each feature is not necessarily plotted for all test angles. The reason for this is that each feature is not visible for certain angle ranges. For the left turn case, the right eye is visible between 0 and 70 degrees, the nose between 0 and 80 degrees, the mouth between 0 and 60 degrees, and the right ear between 35 and 90 degrees. The left ear is completely invisible for the left turn case, which is a valid observation. Another important factor of consideration is the accuracy of the facial feature when it is visible. This is determined when the polynomial models were fit to the calibration data and validated using Python code and the Intel RealSense Camera F200 DSK software. The validation data can be seen in Tables 4.3 and 4.6 and associated errors in Tables 4.4, 4.5, 4.7, and 4.8.

Figure 4.6 below is used to determine which section the head angle is in. For example, if the pixel value is roughly less than 300 then the head has been rotated anywhere between 0 and -90 degrees. These results have been modeled using a cubic function. This model is in accordance with the head angle sweep as it crosses the origin – representing the forward-facing reference position.
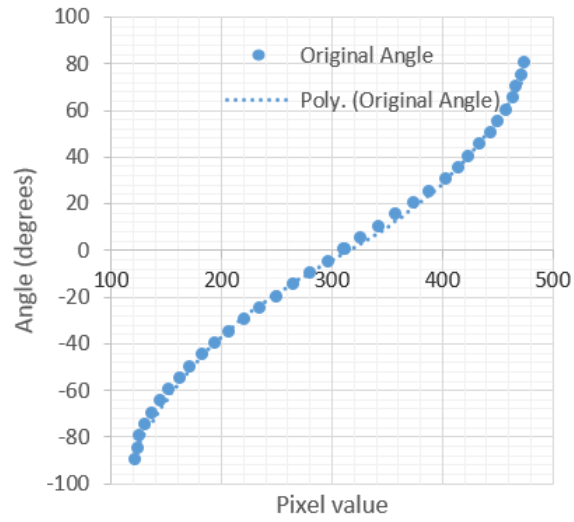
Fig. 4.6 Cubic model for head angle direction. *Left Turn:* 0 to +90°; *Right Turn:* 0 to -90°.
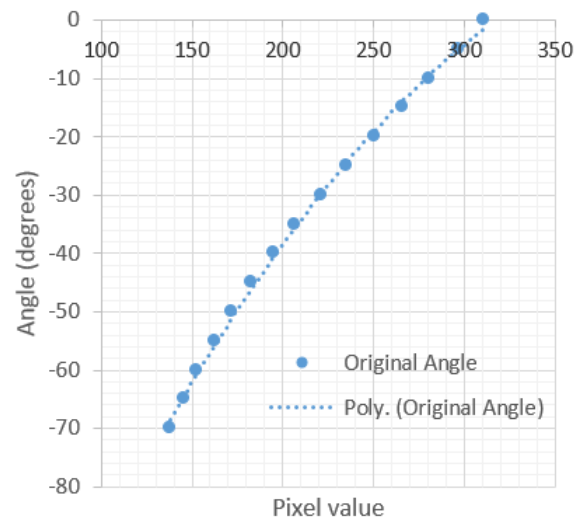
Fig. 4.7 Quadratic model for determining head angles less than -35° using the nose.
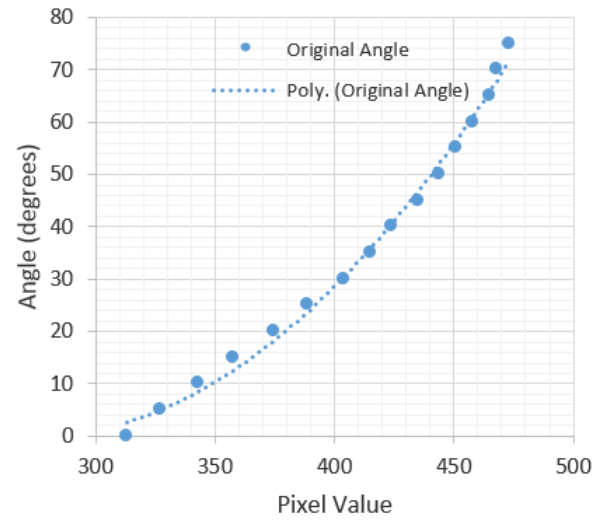
Fig. 4.8 Quadratic model for determining head angles greater than 35° using the nose.
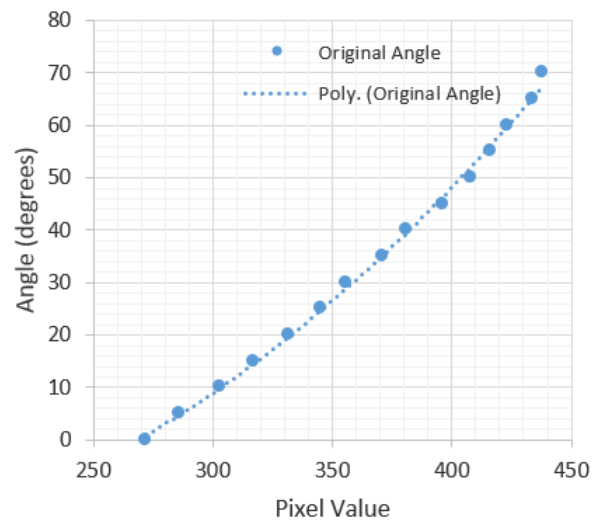
Fig. 4.9 Quadratic model for determining head angles greater than 0° using the right eye.
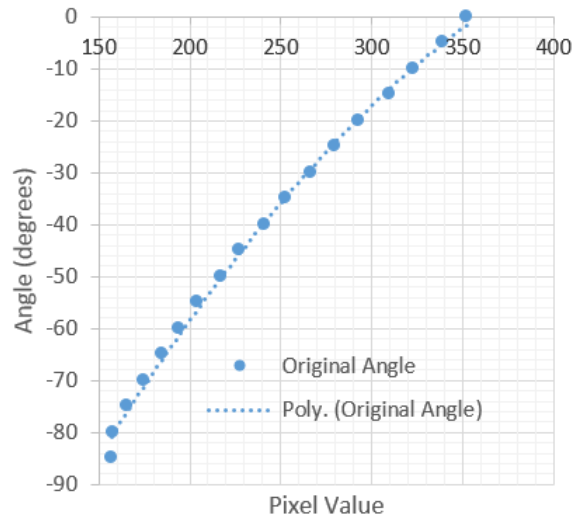
Fig. 4.10 Quadratic model for determining head angles less than 0° using the left eye.
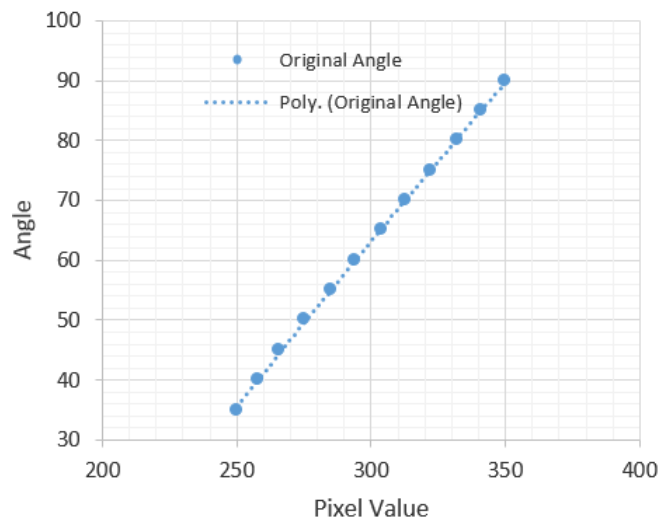
Fig. 4.11 Quadratic model for determining head angles greater than 35° using the right ear.
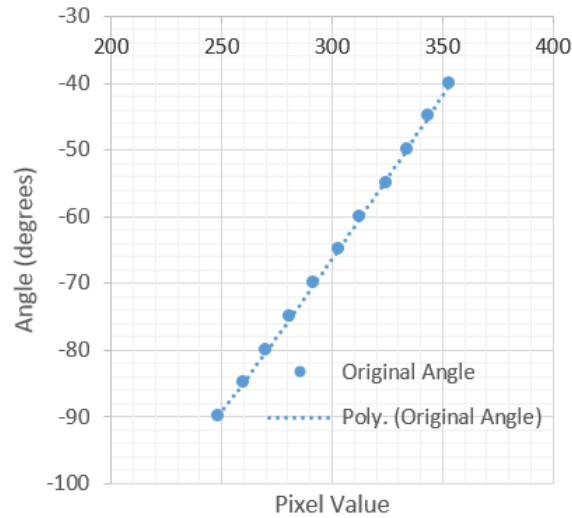
Fig. 4.12 Quadratic model for determining head angles less than -35° using the left ear.

Table 4.2 Estimated Polynomial Equations for Each Facial Feature.

| Angle Range	Feature	Estimated Equations
Find Left/Right	Nose	angle = 5.9077e-06*(pixel)^3 - 0.005307*(pixel)^2 + 1.858*pixel - 243.532
35 to -35	Nose	same as above
< -35	Nose	angle = -0.0008467*(pixel)^2 + 0.7683*(pixel) - 158.29
> 35	Nose	angle = 0.001800*(pixel)^2 - 0.9838*pixel + 134.2948
> 0	Right Eye	angle = 0.000834*(pixel)^2 - 0.1908*pixel - 8.88629
< 0	Left Eye	angle = 0.0007575*(pixel)^2 + 0.7917*pixel -186.4473
> 35	Right Ear	angle = -0.000235*(pixel)^2 + 0.6826*pixel - 120.45
< -35	Left Ear	angle = 0.0003015*(pixel)^2 + 0.2933*pixel - 181.5046

In Figure 4.13 below, it can be seen that when the dummy head is rotated to 20 degrees, the models are used to estimate the angle. All four facial features are visible in this case, however only the right eye and nose models are used (averaged) to generate the best angle estimation.
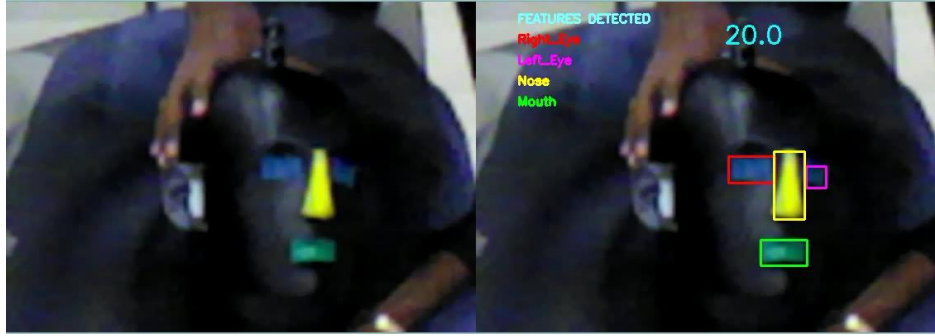
Fig. 4.13 Sample image of calibration results (*Left Turn:* 20°).

Table 4.3 Calibration Results (*Left Turn:* 0 to 90°).

Actual Angle (°)	Generated Angle (°)					
	Right Eye	**Left Eye**	**Nose**	**Mouth**	**Right Ear**	**Left Ear**
90	-	-	*	-	90.2	-
80	-	-	*	-	80.3	-
70	67.6	-	*	-	69.6	-
60	59.7	-	61.00	-	59.4	-
50	52.1	-	53.00	51.20	50	-
40	42.2	-	42.00	42.60	40.6	-
30	28.5	-	30.00	32.20	-	-
20	20.8	-	20.00	22.20	-	-
10	9.6	-	9.00	11.20	-	-
0	1.5	-	0.00	2.60	-	-

" - " Feature not available
" * " Feature not accurate

Table 4.4 Generated Angle Error (*Left Turn:* 0 to 90°).

Actual Angle (°)	Generated Angle Error (°)					
	Right Eye	**Left Eye**	**Nose**	**Mouth**	**Right Ear**	**Left Ear**
90	-	-	-	-	0.2	-
80	-	-	-	-	0.3	-
70	-2.4	-	-	-	-0.4	-
60	-0.3	-	1.00	-	-0.6	-
50	2.1	-	3.00	1.2	0	-
40	2.2	-	2.00	2.6	0.6	-
30	-1.5	-	0.00	2.2	-	-
20	0.8	-	0.00	2.2	-	-
10	-0.4	-	-1.00	1.2	-	-
0	1.5	-	0.00	2.6	-	-

" - " Feature not available
" * " Feature not accurate

Table 4.5 Average Angle Errors (*Left Turn*: 0 to 90°).

Actual Angle (°)	Average (°)	Error (°)	MSE (°)
90	90.2	0.2	0.04
80	80.3	0.3	0.09
70	69.6	-0.4	0.16
60	59.4	-0.6	0.36
50	50.0	0.0	0.00
40	40.6	0.6	0.36
30	29.3	-0.8	0.56
20	20.4	0.4	0.16
10	9.3	-0.7	0.49
0	0.8	0.8	0.56
		Left Turn RMSE	**0.53**

Table 4.6 Calibration Results (*Right Turn:* 0 to -90°).

Actual Angle (°)	Generated Angle (°)					
	Right Eye	**Left Eye**	**Nose**	**Mouth**	**Right Ear**	**Left Ear**
0	1.5	-	0.00	2.60	-	-
-10	-	-9.8	-9.00	-9.70	-	-
-20	-	-19.9	-	-20.60	-	-
-30	-	-29.8	-	-31.30	-	-
-40	-	-38.4	-	-39.7	-	-39.3
-50	-	-50.3	-	-52.7	-	-50.4
-60	-	-61.4	-61	-61.8	-	-60.2
-70	-	-71.1	*	-	-	-69.7
-80	-	-78.6	*	-	-	-79.9
-90	-	-	*	-	-	-89.8

" - " Feature not available
" * " Feature not accurate

Table 4.7 Generated Angle Error (*Right Turn:* 0 to -90°).

Actual Angle (°)	Generated Angle Error (°)					
	Right Eye	**Left Eye**	**Nose**	**Mouth**	**Right Ear**	**Left Ear**
0	1.5	-	0.00	2.6	-	-
-10	-	0.2	1.00	0.3	-	-
-20	-	0.1	0.00	-0.6	-	-
-30	-	0.2	0.00	-1.3	-	-
-40	-	1.6	0.00	0.3	-	0.7
-50	-	-0.3	-2.00	-2.7	-	-0.4
-60	-	-1.4	-1.00	-1.8	-	-0.2
-70	-	-1.1	-	-	-	0.3
-80	-	1.4	-	-	-	0.1
-90	-	-	-	-	-	0.2

" - " Feature not available

" * " Feature not accurate

Table 4.8 Average Angle Errors (*Right Turn*: 0 to -90°).

Actual Angle (°)	Average (°)	Error (°)	MSE (°)
0	0.8	0.8	0.56
-10	-9.4	0.6	0.36
-20	-20.0	0.1	0.00
-30	-29.9	0.1	0.01
-40	-39.3	0.7	0.49
-50	-50.4	-0.4	0.16
-60	-60.2	-0.2	0.04
-70	-69.7	0.3	0.09
-80	-79.9	0.1	0.01
-90	-89.8	0.2	0.04
Right Turn RMSE			**0.42**
Total RMSE			**0.46**

4.2.3. Rotation Results

Based on the validation analysis, it can be concluded that for the left turn case, when going from 0 to 35 degrees the combination of the right eye and nose models yield the best results. In addition, when going from 35 to 90 degrees solely using the right ear yields the best results within this range. For the right turn case, when going from 0 to -35 degrees the combination of the left eye and nose models yield the best results within this range. Again for the right turn case, when going from -35 to -90 degrees solely using the left ear yields the best results within this range. This means that the mouth resulted in the highest errors and was found to be insufficient in accurately tracking the angle for the whole range of -90 through 90 degrees. The root mean square error for the left turn case (0 to 90°), was determined to be 0.53°. Similarly, the root mean square error for the right turn case (0 to -90°), was determined to be 0.42°. The overall root mean square error for an angle sweep from -90 to 90° was found to be 0.46°. These results are impressive, however this method can be improved by using a 3D ellipsoidal model. The continuation of this work is the tracking and determination of the head angle for rotation about other axes, as well as translations.

4.3. Translation Method

The purpose of the translation analysis is to detect the position of the head for small head movements. An example of such movements is when a person is sitting in a chair. If the positions of the head and major facial features can be determined, then the head rotation schemes developed earlier can be implemented with ease. Generally, when the dummy head is translated along the y-axis or sideways, the 2d size of the head remains approximately constant. For the translation analysis, three sets of cascades were initially created. One set from 0 to 15 degrees, another from 15 to 35, and the last one from 30 to 45.

The first step in the creation of these cascades was to produce a set of positive images for each of the angles previously specified: 0, 15, 30, and 45. These sets of positive images were created using 10-second videos, where three frames were saved every second. Through trial-and-error, the ideal number of positive images was found to be 50 images per set. The size of the positive images was 25 x 35 pixels. Next, 4 sets of negative images were produced (for 0, 15, 30, and 45 degrees), where the size of the images varied because it was irrelevant to the cascading process. An important factor for the training process was the ratio 1 to 10 of positive to negative images. If this criteria is not met, the training can go wrong. This can result in an increased level of difficulty in detecting a face in a single frame. Over 600 negative images were obtained in a similar manner to the positive images.

4.3.1. Setup

After the positive and negative image sets have been created, the training process could be started. A difficult part in the training process is selecting the right number of stages. Typically, stages greater than 25 are considered computationally complex and expensive. The training process can be considered as a training tree. The first stage contains a simple training step. As each stage is developed, the process becomes more and more complex. Stages 1 through 5 contain 1, 10, 25, 25, and 50 features respectively. These features are the combination of edges, lines, and rectangles (refer to visual object detection section in Literature Review). Through this exhaustive selection process, the final number of training stages was chosen to be 10.

In order to calibrate the translation, a ruler was created using red tape. The right-leg of the tripod was positioned at 15 cm on the taped red marker. This position was used as the origin for the translation. The tripod was translated 5 cm in each direction along the y axis, in 0.5 cm increments. Alignment of the tripod legs was crucial in determining the true accuracy of the translation and rotation. The error in translation was found to be +/- 0.2 cm, which is acceptable. Ambient lighting was a great concern in 2D tracking analysis, whether it be for rotation or translation. The face pixel values may need to be constantly recalibrated if the lighting in the immediate or surrounding environment changes. This can be challenging because it is time consuming and inefficient. An IR projector and sensor with a range of at least 2 m would be ideal for better face detection and an increased depth field. The Intel RealSense Camera F200 has a range of 0.8 m which, based on the geometry of the current tripod and seat setup, is not enough to completely detect the face. This is why the Microsoft (MS) Kinect was investigated next. The MS

Kinect is a widely used and non-expensive line-of-motion sensing device. It has an admirable range of up to 4 m.

4.3.2. Translation and Rotation

Figures 4.14 and 4.15 demonstrate the process used to develop an integrated rotation and translation system for tracking the dummy head using the MS Kinect. In plain words, the first step is to capture the frame and then to obtain a window near the headrest. The size of the window for positive images (images of the face) is 25 x 35 pixels. Then the window is converted to a grayscale image. Then the Python program makes a decision to obtain the face from the window if it is detected, or to select another cascade to search for the face until it is found. A cascade is a special classifier that is trained using sets of positive and negative images. After the training process, the cascade becomes an object detector and in this case a face detector. Once the face has been selected, the translation and rotation schemes can be applied. The most computationally-challenging part is detecting the face. In order to determine the translated distance, the position of the face's centroid is calculated and then fed into a predetermine translation equation to finally output the translated distance in centimetres. The head rotation angle determination process begins with storing the selected face in HSV format, which uses cylindrical coordinates to express colour in terms of hue, saturation, and value. Then the centroids of the nose and ears are stored separately. Similar to the translation scheme, the various centroid values are fed into predetermined rotation equations, and averaged. This whole process is then repeated for the next frame of the live video feed.
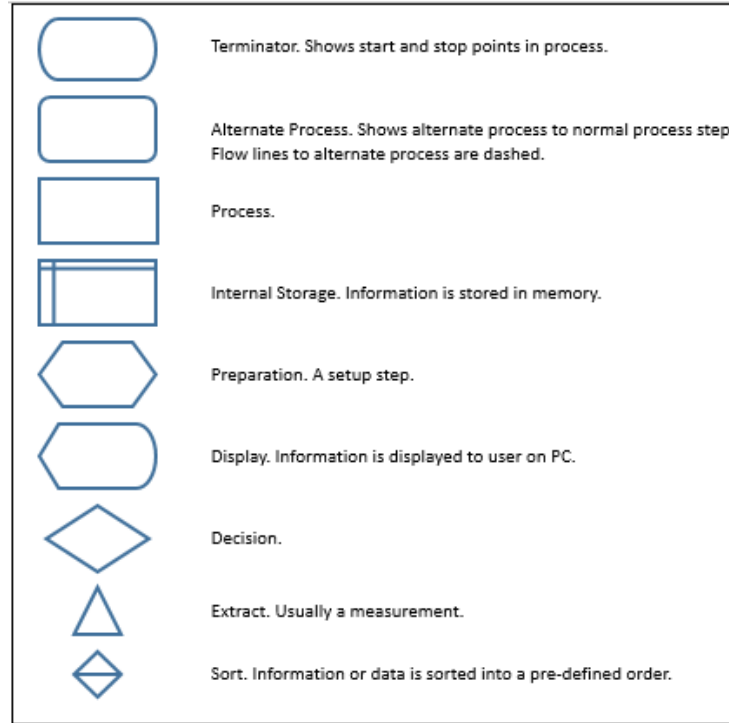
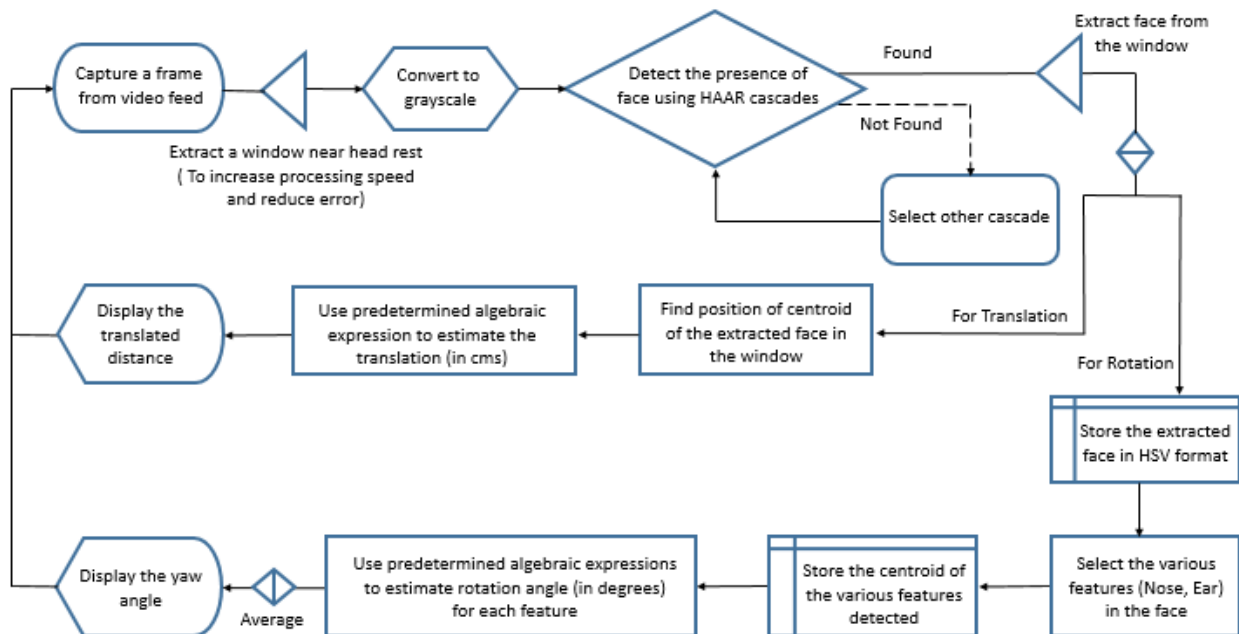**Fig. 4.14** Legend for translation and rotation head-tracking flowchart.

Fig. 4.15 Translation and rotation head-tracking flowchart process diagram.

4.3.3. Results

Sufficient prediction intervals of 0.5 cm for translation and 5 degrees for rotation were selected. At 17 degrees, there is a pixel range gap about 10 pixels wide, as can be seen by the red circled regions in Figure 4.17. The two cascades used for determining the yaw angle converge to the 17 degree output mark, so if the pixel value falls into this region on the graph, the angle is assumed to be 17. The equations for translation and rotation are also presented in Figures 4.16 and 4.17, respectively. It can be seen that the translation equation is first order, while the rotation equations are second and third order. Originally, translation using the Intel F200 camera yielded results within +/- 0.2 cm from the true values. When the MS Kinect was employed, the error decreased almost tenfold with results within +/- 3 mm. With only 2 cascades (0 to 17 degrees and 17 to 40 degrees), the final rotation results were within +/- 0.2 degrees. These results are sufficient, however, when 6 degree-of-motion tracking is implemented, it will be much more difficult to reproduce similar outcomes along each axis of rotation and translation.
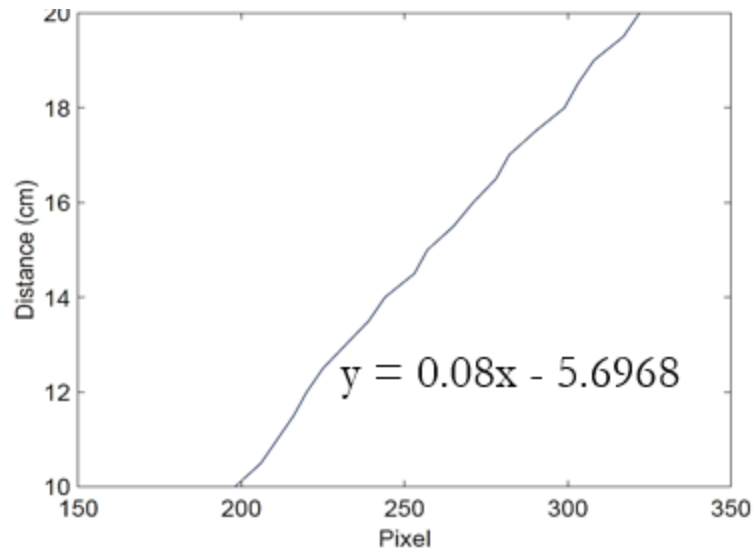
Fig. 4.16 Dummy head translation along y-axis versus pixel position.
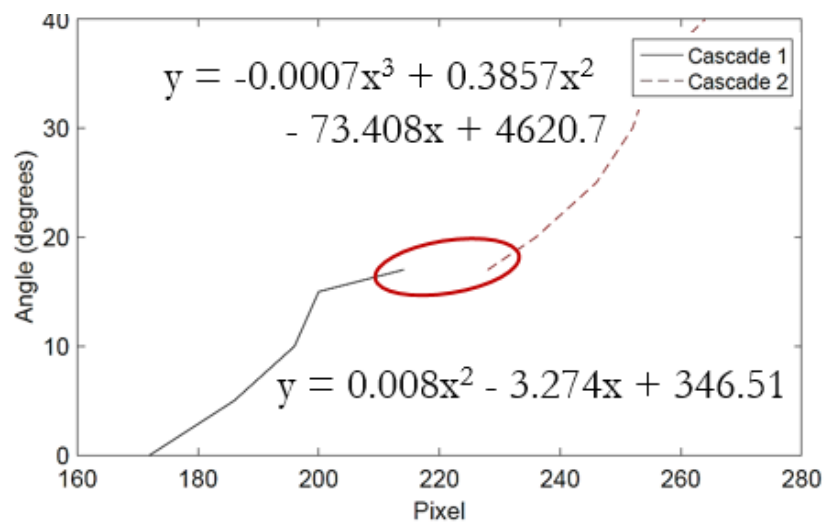

Fig. 4.17 Dummy head angle versus pixel position.

119

5. Conclusion

This research has aimed to improve zone-based local active noise control for an aircraft passenger seat using head tracking and virtual sensing methods. The main contributions of this work include the development of a zone-based integrated ANC system for an aircraft cabin seat which can be used with any virtual sensing method, the creation of a simulation tool in Matlab for determining the location of virtual sensors for ideal ZoQ, and the development of a head tracking tool which interfaces with Python programming language and MS Kinect. The zones of quiet for broadband diffused sound fields have been derived theoretically. The near-field zone of quiet was improved by developing a new method to combine zone-based technologies with existing local ANC virtual sensing techniques. The virtual sensing methods which were reviewed include virtual microphone technique, the forward difference prediction technique, and the adaptive LMS moving virtual microphone techniques. A detailed simulation of virtual sensing methods for an aircraft passenger using adaptive LMS and the forward difference prediction technique was generated. The results of the simulation showed that the desired virtual microphone location is 10 cm from the closest microphone in the array and that the quadratic forward difference prediction technique yielded the lowest RMS errors of 0.0198. Moreover, a head tracking scheme, which utilizes HSV parameters and face cascading methods for an acoustic dummy head, was developed in order to improve the near-field zone of quiet for broadband noise near an aircraft passenger seat. The head tracking measurements proved that the translation of the head could be tracked within +/- 0.2 cm, whereas the head rotation angle could be matched to the closest degree which is acceptable. The theory and methods developed in this work will aid in the accurate prediction of broadband zones of quiet for various MIMO local ANC configurations, which is suggested for future work.

References

1. Alvarez B., J. D. (2006). *In-situ measurements of the Complex Acoustic Impedance of materials for automobile interiors.* Lyngby: Technical University of Denmark.

2. Amit, Y., Geman, D., & Wilder, K. (1997). Joint induction of shape features and tree classifiers.

3. Ardekani, I. T., & Abdulla, W. H. (2014). Active Noise Control in Three Dimensions. *IEEE Transactions on Control Systems Technology*, 2150-2159.

4. Avni, A. (2010). *Spaciousness of sound fields captured by spherical microphone arrays.* Bengurion University of the Negev Faculty of Engineering Sciences.

5. Avni, A., & Rafaely, B. (2010). Interaural cross correlation in a sound field represented by spherical harmonics. *J. Acoust. Soc. Amer.*, 823-828.

6. Bailo, K. C., Brei, D. E., & Grosh, K. (2003). Investigation of Curved Polymeric Piezoelectric Active Diaphragms. *Journal of Vibrations and Acoustics*, 146-154.

7. Beranek, L. (1988). *Acoustical Measurements.* Cambridge: Acoust. Soc. Amer.

8. Bies, D. A. (1981). A unified theory of sound absorption in fibrous porous materials . *Australian Acoustical Society Annual Conference.* Cowes.

9. Bjarnason, E. (1995). Analysis of the filtered-x LMS algorithm. *IEEE Trans. Speech Audio Process.*, 504-514.

10. Bolton, J. S., & Green, E. R. (1993). Smart foams for active absorption of sound. *Second Conference on Recent Advanced in Active Control of Sound and Vibration*, (pp. 139-149). Blacksburg.

11. Burgess, J. C. (1981). Active adaptive sound control in a duct: a computer simulation. *J. Acoust. Soc. Amer.*, 715-726.

12. Cazzolato, B. (n.d.). *An adaptive LMS virtual microphone.* Adelaide: Department of Mehanical Engineering, University of Adelaide.

13. Chassaing, R. (2002). *DSP Applications Using C and the TMS320C6x DSK.* New York: John Wiley & Sons.

14. Cheng, A., & Cheng, D. (2005). Heritage and early history of the boundary element method. *Engineering Analysis with Boundary Elements*, 268-302.

15. Elliott, S. (2005). *Active Noise Control: Physical Principles and Practical Applications.* Southampton: Institute of Sound and Vibration Research (ISVR) University of Southampton.

16. Elliott, S. J., & Nelson, P. A. (1993). *Active Control of Sound.* Cambridge: Academic.

17. Federal Aviation Administration. (1985, March). *Aviation Noise Effects*. Retrieved August 27, 2015

18. Garcia-Bonito, J., Elliott, S., & Boucher, C. (1996). A virtual microphone arrangement in a practical active headrest. *In Proceedings of Inter-noise 96*, 1115-1120.

19. Garcia-Bonito, J., Elliott, S., & Boucher, C. (1997). Generation of zones of quiet using a virtual microphone arrangement. *Journal of the Acoustical Society of America 101(6)*, 3498-3516.

20. Gentry, C. A., Guigou, C., & Fuller, C. R. (1997). Smart foam for applications in passive-active noise radiation control. *J. Acoust. Soc. Amer.*, 1771-1778.

21. Green, R. (2003). Spherical Harmonic Lighting: The Gritty Details. *Archives of the Game Developers Conference*.

22. Griffin, J. R. (2006). *The Control of Interior Cabin Noise Due to a Turbulent Boundary Layer Noise Excitation Using Smart Foam Elements.* Blacksburg: Faculty of the Virginia Polytechnic Institute and State University.

23. Guigou, C., & Fuller, C. R. (1999). Control of aircraft broadband noise with foam-PVDF smart skin. *Sound and Vibration*, 541-557.

24. Hald, J., Morkholt, J., Hardy, P., Trentin, D., Bach-Andersen, M., & Keith, G. (2008). Array based measurement of radiated and absorbed sound intensity components. *J. Acoust. Soc. Amer.*, 919-924.

25. Hansen, C. (2005). *Noise Control: from concept to application.* Taylor & Francis Inc.

26. Holmberg, U., Ramner, N., & Slovak, R. (2002). Low complexity robust control of a headrest system based on virtual microphones and the interal model principle. *In Proceedings of Active 02, ISVR, Southampton, UK*, 1243-1250.

27. Itti, L., Koch, C., & Niebur, E. (1998, November). A model of saliency-based visual attention for rapid scene analysis. *IEEE Patt. Anal. Mach. Intell., 20*(11), 1254-1259.

28. James, D. L., Barbic, J., & Pai, D. K. (2006). Precomputed acoustic transfer: output-sensitive, accurate sound generation for geometrically complex vibration sources. *ACM SIGGRAPH* , pp. 987-995.

29. Joseph, P., Elliott, S. J., & Nelson, P. A. (1994). Near Field Zones of Quiet. *Journal of Sound and Vibration 172(5)*, 605-627.

30. Kang, J. (1996). Sound attenuation in long enclosures. *Building and Environment*, 245-253.

31. Kang, J. (1997). A method for predicting acoustic indices in long enclosures. *Applied Acoustics*, 169-180.

32. Kuo, S. M., & Morgan, D. R. (1999). Active noise control: a tutorial review. *IEEE*, (pp. 943-973).

33. Lavergne, A. (1999). *Computer Vision System for Head Movement Detection and Tracking.* The University of British Columbia.

34. *Lecture Notes on the Mathematics of Acoustics.* (2005). London: Imperial College Press.

35. Liu, Y. (2009). *Fast Multipole Boundary Element Method: Theory and Applications in Engineering.* Cambridge University Press.

36. Lucas, B. D., & Kanade, T. (1981, August 24-28). An Iterative Image Registration Technique wtih an Application to Stereo Vision., (pp. 674-679). Vancouver.

37. Maekawa, Z., & Osaki, S. (1985). A Simple Chart for the Estimation of the Attenuation by a Wedge Diffraction. *Applied Acoustics*, 355-368.

38. Mehra, R., Antani, L., Kim, S., & Manocha, D. (2014). Source and Listener Directivity for Interactive Wave-based Sound Propagation. *IEEE Transactions on Visualization and Computer Graphics*, 495-503.

39. Mehra, R., Raghuvanshi, N., Antani, L., Chandak, A., Curtis, S., & Manocha, D. (2013). Wave-based sound propagation in large open scenes using an equivalent source formulation. *ACM Trans. Graph.*

40. Mehra, R., Rungta, A., Golas, A., Lin, M., & Manocha, D. (2015). WAVE: Interactive Wave-based Sound Propagation for Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics*, 434-442.

41. Menounou, P., & Papaefthymiou, E. S. (2010). Shadowing of directional noise source by finite noise barriers. *Applied Acoustics*, 351-367.

42. Monner, H. P. (2005). Smart materials for active noise and vibration. *Noise and Vibration: Emerging Methods*, 1-17.

43. Moreau, D., Cazzolato, B., Zander, A., & Petersen, C. (2008). A review of virtual sensing algorithms for active noise control. *Algorithms 1*, 69-99.

44. Morkholt, J., Hald, J., & Gade, S. (2011). Sound Intensity Measurements in Vehicle Interiors. *Sound & Vibration*, 12-14.

45. Nelson, P. A., & Elliot, S. J. (1992). *Active Control of Sound.* London: Academic.

46. Ng, K. W. (1995). Applications of active control. *The 1995 Symposium on Active Control of Sound and Vibration.* Newport Beach.

47. Nilsson, M., Soli, S. D., & Sullivan, A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J. Acoust. Soc. Amer.*, 1085-1099.

48. Norcross, S., Bradley, J., & Halliwell, R. (n.d.). Errors Incurred Using a Dummy Head to Make Omni-directional Room Acoustics Measurements. *Architectural & Performance Acoustics*, 69-70.

49. Ochmann, M. (1995). The source simulation technique for acoustic radiation problems. *Acustica*, 512-527.

50. Pawelczyk, M. (2003). Multiple input-multiple output adaptive feedback control strategies for the active headrest system: design and real-time implementation. *Internationl Journal of Adaptive Control and Signal Processing, 17(10)*, 785-800.

51. Pawelczyk, M. (2003). Noise control in the active headrest based on estimated residual signals at virtual microphones. *In Proceedings of the 10th International Congress on Sound and Vibration*, 251-258.

52. Pierce, A. D. (1989). Acoustics: An Introduction to Its Physical Principles and Applications. *J. Acoust. Soc. Amer.*

53. Porada, W. (1975). Model Measurements on Noise Screening of Line Sources By Single and Double Barriers. *Applied Acoustics*, 271-280.

54. Rafaely , B., Elliott, S., & Garcia-Bonito, J. (1999). Broadband performance of an active headrest. *Journal of the Acoustical Society of America 106(2)*, 787-793.

55. Rafaely, B. (2001). Zone of quiet in a broadband diffuse sound field. *Journal of the Acoustical Societ of Americal 110 (1)*, 296-302.

56. Rafaely, B., Garcia-Bonito, J., & Elliott, S. (1997). Feedback control of sound in headrest. *In Proceedings of Active 97*, 445-456.

57. Raghuvanshi, N., Snyder, J., Mehra, R., Lin, M. C., & Govindaraju, N. K. (2010). Precomputed Wave Simulation for Real-Time Sound Propagation of Dynamic Sources in Complex Scenes. *Siggraph*.

58. Roure, A., & Albarrazin, A. (1999). The remote microphone technique for active noise control. *In Proceedings of Active* , 1233-1244.

59. Serizel, R., Moonen, M., Wouters, J., & Jensen, S. H. (2008). Combined active noise control and noise reduction in hearing aids. *Int. Workshop Acoust. Echo Noise Control (IWAENC).*

60. Serizel, R., Moonen, M., Wouters, J., & Jensen, S. H. (2010). Integrated Active Noise Control and Noise Reduction in Hearing Aids. *IEEE Transactions on Audio, Speech, and Language Processing*, 1137-1146.

61. Serizel, R., Moonen, M., Wouters, J., & Jensen, S. H. (2012). A Zone-of-Quiet Based Approach to Integrated Active Noise Control and Noise Reduction for Speech Enhancement in Hearing Aids. *IEEE Transactions on Audio, Speech, and Language Processing*, 1685-1697.

62. Taflove, A., & Hagness, S. C. (2005). *Computational Electrodynamics: The Finite-Difference Time-Domain Method, Third Edition.* Artech House Publishers.

63. Thompson, L. L. (2006). A review of finite-element methods for time-harmonic acoustics. *J. Acoust. Soc. Amer.*, 1315-1330.

64. Tibbets, G. C. (1977, November). *U.S. Patent No. 4,056,742.*

65. Tsingos, N., Dachsbacher, C., Lefebvre, S., & Dellepiane, M. (2007). Instant Sound Scattering. *Eurographics Symposium on Rendering.* Rendering Techniques.

66. Tsotos, J., Culhane, S., Wai, W., Lai, Y., Davis, N., & Nuflo, F. (1995, October). Modeling Visual-attention via Selective Tuning. *Artificial Intelligence Journal, 78*(1-2), 507-545.

67. Vigeant, M. C. (2008). Investigations of incorporating source directivity into room acoustics computer models to improve auralizations. *J. Acoust. Soc. Amer.*, 2664.

68. Viola, P., & Jones, M. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. *Accepted Conference on Computer Vision and Pattern Recognition*.

69. Widrow, B., & Stearns, S. D. (1985). *Adaptive Signal Processing.* Englewood Cliffs: Prentice-Hall.