Mobile laser Doppler vibrometry motion tracking and vibration compensation for in-field vehicular deployments

by Abdel Darwish

Thesis submitted in fulfillment of the requirements for
the degree of

Doctor of Philosophy

Under the supervision of:
A/Prof. Benjamin Halkon
A/Prof. Sebastian Oberst
Prof. Robert Fitch

University of Technology Sydney
Faculty of Engineering and Information Technology

December 2022

Certificate of Original Authorship

I, Abdel Darwish, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signed: Production Note: Signature removed prior to publication.

Date: 09/08/2023

Acknowledgements

First and foremost, I would like to thank my principal supervisor, Benjamin Halkon, and my co-supervisors, Sebastian Oberst and Robert Fitch, for their continued guidance and help throughout my candidature. I have greatly benefited from not only all of your expertise but also the life skills I have acquired over the last few years.

I would like to acknowledge the support of UTS for the International Research Scholarship and the Faculty of Engineering and Information Technology Scholarship. Along with the NSW Department of Industry, Innovation and Science and Robotic Systems Pty Ltd (Newcastle, Australia) for the Innovation Connections grant which I was a part recipient of. And thank you to Sebastian Oberst, who facilitated this additional funding through his tireless work.

I would like to thank my sister Haleimah who has always made time for me and patiently listened, despite being on the other side of the world. I would also like to thank my foster family, Sarah, Marwan, Amir and Nasser Kishawi, who opened up their home to me when I was a child, and all these years later still remain a supportive and caring part of my life. And last but definitely not least, a special thank you goes to my partner, Georgia Benson, for her unwavering love and support, particularly during these final few months. I dedicate this thesis to my late mother, Wafaa, who would have been overjoyed by this accomplishment.

Mobile laser Doppler vibrometry motion tracking and vibration compensation for in-field vehicular deployments

by

Abdel Darwish

A thesis submitted in partial fulfilment of the requirements for the
degree of Doctor of Philosophy

Abstract

The laser Doppler vibrometer (LDV) has become an indispensable tool in vibration engineering, boasting high bandwidths and spatial resolutions unrivalled by traditional contacting accelerometers - all whilst doing so via non-contact means. Since their widespread adoption, their application envelope as vibration transducers has been steadily expanding, encompassing increasingly impactful and interesting areas. This work focuses specifically on the significant potential of deploying LDV from mobile platforms, for example, terrestrial or airborne vehicles, since it has the potential to substantially increase land coverage rates whilst simultaneously enabling access to hazardous or remote areas. This field has already received some interest, with research taking place into mobile buried landmine detection, intelligence gathering from drones and structural health monitoring from drones. However, these represent only a small fraction of potential applications. The first portion of this work explores and mitigates the effects of any instrument motion on mobile LDV deployment, while the second portion explores a novel mobile application for a robust LDV: non-contact vibro-acoustic object recognition and enhanced point cloud perception for autonomous systems.

In the reference frame of the sensor head, the motion can be divided into two broad categories: translational motion along the beam axis *or* translational motion in the two orthogonal axes, plus the three rotational degrees of freedom. Since the underlying physics of the LDV dictate that the sensor head is as sensitive to self-vibration in the beam axis as it is to the target vibration in the beam axis, sensor head vibration could compromise the measurement. Whereas motion in the five non-beam axes would cause the measurement beam to stray from the intended measurement location. Two separate approaches can mitigate the effects of these two distinct phenomena, referred to herein as *measurement correction* and *arbitrary tracking.*

Since the sensor head is as sensitive to self-vibration as it is to target vibration, measurement correction is required when self-vibration is present. Despite the success of the existing, single accelerometer-based measurement correction technique, there were two main improvements to be made. Firstly, the technique was fundamentally limited to stationary signals due to the frequency domain-based signal processing employed, requiring a time domain alternative for mobile deployments. Secondly, the technique lacked proper accelerometer signal handling, leading to sub-optimal performance, therefore requiring a revised technique. As such, the first portion of this thesis focuses on the development of a new time domain and revised frequency domain-based processing techniques that display up to eight-fold improvements in performance over the previous technique. This is accompanied by a rigorous analytical model describing the effects of synchronisation quality on both techniques, delving into the nuances of time domain-based signal synchronisation. Finally, the thesis covers the extension of these signal processing techniques to be compatible with existing theoretical *scanning* LDV measurement correction work, with the first experimental validation of the technique on the Multi-Axis Simulation Table taking place.

While arbitrary path-tracking LDV solutions exist, current techniques cannot meet the strict weight requirements of a drone-mounted LDV system. As such, the development of a novel tracking system specifically tailored for the hovering drone is described; specifically focused on correcting small pitch and roll adjustments that drones make while holding their position in a hover. The proposed system employs a standard galvanometer steering mirror setup found in scanning LDV systems to counter-rotate the beam rather than the entire instrument. This technique's performance is then assessed on the Multi-Axis Simulation Table, configured to simulate a hovering drone in extreme conditions. Results show that the beam motion was reduced by 68%.

With a comprehensive framework established for mitigating the effects of both beam axis and non-beam axis motion, the focus shifts onto applications for such a system, specifically, the possibility of LDV integration into autonomous systems. Initially, this portion of the work describes a novel vibro-acoustic object recognition technique utilising convolutional neural networks to classify the LDV measurements. A rigorous five-fold cross-validation showed it is possible to recognise acoustically excited objects with up to 99.8% accuracy. Finally, the thesis explores the possibilities of merging point clouds with LDV scans for enhanced perception applications for autonomous systems in a first-of-its-kind proof-of-concept system, allowing autonomous systems to "see" the surrounding acoustic world.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the 1800s, it was generally believed that there existed a medium that permeated all space and through which light waves propagated, known as the luminiferous aether. To prove this, the Michelson interferometer was invented by A. Michelson and E. Morley [1], seen pictured in Figure 1.1. The Michelson interferometer takes advantage of the wave-like properties of light to detect changes in optical path length smaller than a fraction of the wavelength being used. While the pair were unable to detect any sign of the luminiferous aether, Michelson interferometers have become an invaluable tool in science and are the basis of many modern-day devices. Some examples of these devices include optical coherence tomography scanners, commonly used in ophthalmology; the Laser Interferometer Gravitational-Wave Observatory, used to detect the first gravitational wave based in the United States of America; and the laser Doppler vibrometer (LDV), which is the focus of this doctoral program.

Although the original Michelson interferometer employed white light, modern-day versions of the optical arrangement employ a laser. The laser was invented in 1960 by T. Maiman [2] by taking advantage of the stimulated emission of ruby, with its name originating from Light Amplification by Stimulated Emission of Radiation. They have become an indispensable tool in many optical applications since they output monochromatic coherent light, i.e. light of a single frequency with all wavelengths in phase [3]. Like the

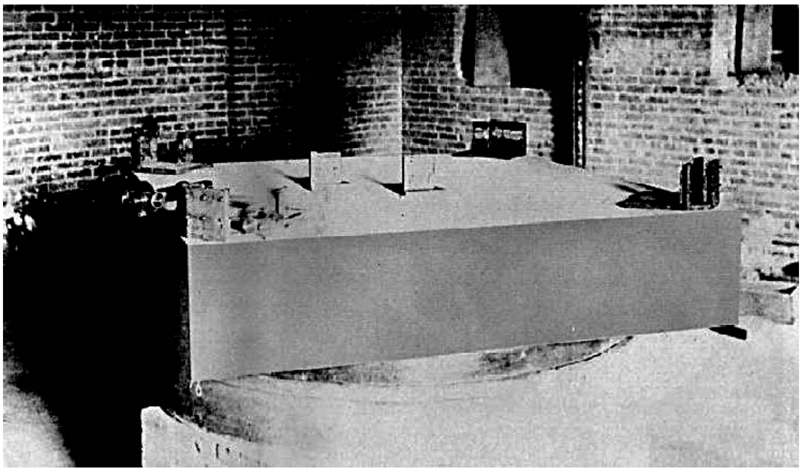


Figure 1.1: Photograph of the original Michelson-Morley interferometer [1]. Ⓢ Public domain, no permission required.

Michelson interferometer, the LDV takes advantage of the wave-like nature of light and the Doppler effect to make non-contact surface velocity measurements in the direction of the measurement beam. An example of a fairly popular LDV can be seen in Figure 1.2.

The optical nature of the LDV leads to a high spatial resolution and bandwidth compared to traditional contacting accelerometers, making them uniquely suited for vibration measurements [4]. This has led to their widespread adoption in a vast range of areas, a few examples include, mechanical engineering, where they have been crucial in studying vibrations, especially in rotating machinery [5, 6, 7, 8, 9]; electrical engineering, where they have been used on microelectromechanical systems as there is no mass loading [10, 11, 12, 13]; civil engineering, as they can make remote non-contact measurements over a dense grid of points for defect detection [14, 15, 16]; biology, where they have been applied to study vibrations in plants, ants and humans [17, 18, 19]; and medicine, where they have been used to assess cardiovascular risk and as a non-contact means to measure the respiration and heart rates in preterm infants [20, 21].

When an LDV is mounted to a moving platform, referred to as *mobile* laser Doppler vibrometry, it has the potential to substantially increase area coverage rates and enable access to remote or hazardous areas. Despite this, literature describing mobile LDV is sparse, currently including buried landmine detection from terrestrial vehicles [22, 23,

Figure 1.2: Photograph of the Polytec PDV-100 portable laser Doppler vibrometer.

24], seismic measurements from a moving ground vehicle [25], and railway bridge measurements with a drone-mounted LDV [26]. Other promising, yet-to-be-realised proposed mobile LDV deployments include orbital seismology [27, 28, 29], intelligence gathering [30], structural health monitoring [31, 32, 33] and LDV-based buried landmine detection from drones [34]. Despite the numerous benefits and interest surrounding mobile LDV deployment, progress has been largely hampered by the fundamental nature of LDV.

1.1 Challenges surrounding mobile laser Doppler vibrometer deployment

In the reference frame of the sensor head, the motion experienced during a mobile deployment can be divided into two broad categories: translational motion along the beam axis *or* translational motion in the two orthogonal axes, plus the three rotational degrees of freedom; referred collectively here as *non-beam DOF motion*. The underlying physics of LDV dictate that any motion along the beam axis, from the target or the sensor head[1], would equally lead to a velocity measurement; making instrument motion and target motion along the beam axis indistinguishable from one another. Whereas, instrument motion in the non-beam DOFs - which the LDV is not sensitive to - will cause the beam to stray from the intended measurement location. Therefore, two separate approaches are required

---

[1] Or the vibration of the entire instrument for portable models such as the Polytec PDV-100 or the Polytec VibroGo.

to mitigate the effects of these two distinct phenomena, referred to herein as *measurement correction* and *arbitrary motion tracking.*

1.1.1 Motion in the measurement beam axis

The innate sensitivity of the LDV to self-vibration normally limits its use to scenarios with low ambient vibration levels or requiring adequate isolation; typically a tripod with compliant feet. Some active vibration isolation techniques exist with improved performance [25]; however, these can be too heavy or costly for some applications. Differential LDVs can be configured to optically subtract the motion of the sensor head by mixing the measurement beam returning from the target with a second external beam returning from a static surface [35]. However, this approach is practically constrained in the context of mobile LDV deployment. Previous research has demonstrated that instrument vibration measurements must be taken along the beam axis to enable measurement correction during six-DOF vibration; failure to do so results in coupling between the corrected measurement and the non-beam axis instrument motion [31]. Therefore, requiring the reference beam to be anti-parallel and on the beam axis, focused on a static reference surface behind the instrument. While this may be possible in the laboratory, for mobile, field-based applications, this is practically impossible due to the simultaneous positioning and focusing requirements.

Using a technique somewhat similar to the differential LDV configuration, alternative techniques remove the requirement for an external static reference surface by using a secondary transducer to estimate the sensor head vibration and subsequently subtract it from the measurement [31, 36, 37, 38, 39]. Of these techniques, referred to herein collectively as *measurement correction*, the most accessible and accurate utilises a single rear-mounted accelerometer to estimate the motion of the sensor head such that it may be subtracted using a frequency domain-based signal processing technique [31]. Its performance and ease of implementation have also led it to be replicated by other international research groups [40]. However, the signal processing implemented before the work presented herein could have been improved in two main aspects. Firstly, due to the frequency domain analysis employed, the technique used was limited to stationary signal types only. Secondly, the frequency domain signal processing lacked proper handling of the accelerometer signal, which significantly hampered the total obtainable performance.

Since it is likely that the vibration experienced during these mobile deployments is *transient* in nature, it is imperative that any such system have time domain-based measurement correction capabilities. As such, the first portion of this thesis describes the development and performance assessment of a novel time domain-based signal processing technique which is compatible with the hardware arrangement of the existing single-accelerometer frequency domain-based technique. A rigorous assessment of the new time domain-based technique's performance in comparison to the established technique is conducted for stationary signal types. This assessment shows a significant improvement in performance, highlighting the importance of proper accelerometer signal handling. In some scenarios with stationary signal types, it may still be convenient to use frequency domain-based signal processing despite the development of this new time domain technique. As such, the focus of this thesis then shifts to the development of an improved frequency domain-based approach by introducing proper accelerometer signal handling. A second rigorous assessment of this improved technique shows that the modifications made to the technique successfully closed the performance gap, obtaining similar results to the new time domain-based approach.

In either domain, the quality of the measurement correction is related to the accuracy of the signal synchronisation from the two transducers since any error therein will adversely affect the quality of the corrected velocity estimate. This is often accounted for by estimating and subsequently removing the time delay with post-processing [38, 31]. The accuracy of this time delay estimate is a fundamental factor dictating achieved performance, however, the implementation of the synchronisation is constrained to integer multiples of the time step in the time domain, therefore, even an exact delay estimate is unlikely to lead to perfect synchronisation. Therefore, the development of a model which relates synchronisation error to the velocity estimate error is essential. As such, a rigorous analytical model describing the effects of sampling frequency and errors in the time delay estimate on the measurement correction performance is developed for both domains, focusing on the time domain. Using the insights gained from this model, a holistic framework is developed to assist the user with the selection of both the correction technique and the acquisition parameter selection, ensuring optimal measurement correction.

For the measurement correction described, it is vital that the measurement of the instrument motion be colinear *and* on the beam axis. While this is straightforward for a single-point LDV, it becomes challenging when the beam is directed independently of the sensor head, such as on a scanning LDV (SLDV). An SLDV is fundamentally no different to a single-point LDV but includes the addition of steering optics and the appropriate control circuitry such that the mirrors may be precisely positioned. This is most often implemented using a pair of orthogonal galvanometric steering mirrors; a commercial system can be seen pictured in Figure 1.3. SLDVs are often used for modal analysis due to their ability to automatically take measurements in a dense grid of points [4]. As such, there has similarly been interest in applying measurement correction to these systems. This was shown to be theoretically possible with a rigorous vector analysis if a total of three accelerometers are specifically positioned relative to the final steering mirror [41]. If the angles of the mirrors are known, then the appropriate component of the three accelerometers can be taken to find the velocity in the beam direction. This work is then integrated with the new time and frequency domain signal processing described herein, followed by the first experimental validation of SLDV measurement correction. This was performed on the Multi-Axial Simulation Table (MAST) whilst undergoing six degrees of freedom (DOF) vibration.

The work presented herein has defined a new gold standard for measurement corrections with the introduction of the two time and frequency domain techniques and their extension to SLDV systems, substantially improving the expected performance for both single-point LDVs and SLDVs while also introducing transient signal compatibility. Similarly, the development of the accompanying analytical model allows for the further optimisation of these techniques by informing the user of the necessary acquisition parameters *a priori*; further expanding the capability of the LDV towards truly mobile deployments. However, the aforementioned measurement correction in this thesis has focused specifically on sensor

FIGURE 1.3: Photograph of a Polytec PSV-500-Xtra scanning head.

head motion along the measurement beam axis. However, the mobile deployment of an LDV likely also requires the adequate handling of instrument motion in the two non-beam axes *and* any rational motion. This motion is distinctly different to motion along the beam axis since this will cause the beam to deviate from the intended target; as such, the beam will likely need to be directed independently from the mobile vehicle. An LDV system outfitted to handle motion in the five non-beam axes, combined with the measurement correction techniques described herein, would be sufficiently robust to enable high-quality measurements to be taken from mobile platforms with previously unacceptable levels of transient vibration and beam motion.

1.1.2 Motion in the non-beam axes

Independently aiming the measurement beam from a mobile vehicle is straightforwardly achieved using the aforementioned SLDV setup; however, since the vehicle motion during a mobile deployment will likely be unpredictable with no *a priori* knowledge of the expected trajectory, then such a system must be embedded with some form of sensor. Until the work presented herein, the only viable solutions were a range of various image-based tracking LDV systems (iTLDV) [42, 43]. These techniques utilise an SLDV with an additional in-line high-speed camera to observe the target. Using a sufficiently powerful computer running a tracking algorithm, the mirrors are driven in a closed control loop, therefore, maintaining a fix on the target. A benefit of iTLDV is that it is compatible with either instrument motion in the five non-beam axes, a mobile target, or both. However, for most cases discussed here, the target is likely to be stationary. Furthermore, since iTLDV also requires the necessary onboard computational and power resources, the additional weight can be relatively large, resulting in a system that may be more suitable for terrestrial deployment than drone deployments.

Known drone-based LDV applications have been gaining popularity with the aforementioned research into intelligence gathering and structural health monitoring [26, 30, 31], however, many yet-to-be realised applications might hold even greater promise. For example, drone-based LDV buried landmine detection can dramatically increase area coverage rates whilst simultaneously removing any risk of accidental detonations, protecting assets from damage or destruction and potentially saving lives. Ground penetrating radar buried landmine detection has already been successfully implemented from a drone [44], although it has been shown that LDV-based detection techniques have significantly higher accuracy due to their ability to detect plastic land mines [45]. Therefore, this work proposes a novel tracking solution specifically tailored towards the hovering drone and a stationary target scenario - as would be expected during intelligence gathering, structural health monitoring or buried landmine detection from drones. Hovering drones have a remarkable ability to hold their location in space by utilising their myriad of sensors to measure and correct for any unwanted translational movement. These corrections come in the form of small adjustments in pitch and roll, therefore, an LDV tracking system intended for hovering drone deployment could simply act as a camera gimbal. Camera gimbals maintain a fix on a specific angle by counter-rotating the contained device, typically camera [46]. However, a gimbal-mounted LDV would either be sluggish or heavy due to the responsiveness and weight trade-off of the actuators which rotate instrument.

The novel system proposed here overcomes the challenges presented by traditional gimbals by counter-rotating the beam using an SLDV, rather than the entire instrument. This is shown to be more accessible, lightweight, and cheaper than the vision-based alternatives, making the system ideal for drone applications. The system's performance is assessed on the MAST while replicating a drone hovering in extreme conditions. The data is acquired using a vision-based system such that both the input of the MAST and the output of the gimbal are known in perfect synchronisation. The acquired data is analysed in both the time and frequency domain to fully characterise the system's performance. The results showed that even in these extreme conditions, the beam motion could be reduced by an average of 68%. The work on both the beam axis motion and non-beam axis motion conducted during this doctoral program, combined with existing work on the topic demonstrates a comprehensive framework for the components of a truly robust LDV ready for mobile deployment. Therefore, following these topics, the focus shift onto new applications of a mobile LDV.

1.2 Mobile laser Doppler vibrometer autonomous system applications

As mentioned, a truly robust LDV ready for mobile deployment has applications in many fields. This thesis specifically focuses on improving acoustic object recognition for autonomous vehicles, although the potential applications are by no means limited to these. Up until now, existing acoustic object recognition approaches described in the literature involve excitation techniques that are *contact* in nature, using either a simple actuator [47, 48] or a multiple degrees-of-freedom robotic arm [49, 50] to excite an audible response in the object which is recorded using a microphone. These recorded sounds are then processed and classified using a range of signal processing and machine learning techniques in order to classify the sound, hence recognising the object. Therefore, imbuing these robots with similar acoustic object recognition skills to a person. Despite recognition rates in the range of 85.5% [50] and 98.2% [49], these existing acoustic object recognition techniques have a number of shortcomings. Firstly, these techniques require physical access to the object, meaning the robot must move towards the object, making the task slower and more complex. Secondly, the object must be excited with sufficient force to produce an audible response. Therefore, some fragile objects could be damaged during the excitation, for example, an object made from glass. Therefore, this thesis presents a new entirely non-contact *vibro*-acoustic object recognition technique.

This new *non-contact* vibro-acoustic object recognition technique substitutes the contacting actuator of the robot with an acoustic excitation emitted by a loudspeaker. However, the response generated as a result of this acoustic excitation is orders of magnitude lower in amplitude than those previously excited via direct contact, therefore, generating little to no detectable sound. As such, the microphone previously employed is similarly substituted with an LDV directly measuring the low-amplitude vibrational response of the object itself, rather than measuring the sound the object generates as a result of the excitation. These recorded responses are processed with a range of spectrogram-based techniques and then used to train a convolutional neural network via transfer learning to classify the responses with accuracies of over 99.7%. This near-perfect classification performance is higher than the existing *contact* techniques, whilst simultaneously introducing valuable improved *non-contact* functionality, therefore, positioning it as a viable object recognition technique for autonomous systems and potentially even other machine automation tasks. Finally, the thesis explores what this practical LDV deployment onto an autonomous vehicle may look like. Here, LDV and point cloud data are used in conjunction with one another together in a proof-of-concept system, establishing the groundwork to enable objects of interest to be identified within a point cloud for further vibro-acoustic interrogation with the LDV. Therefore imbuing the autonomous system with the unique ability to perceive its surroundings as a spatially associated vibrational map and, if combined with the previously mentioned vibro-acoustic object recognition technique, meaningfully interpret this vibrational data.

1.3 Aims and Objectives

To summarise, the broad aim of this thesis is to enable the mobile deployment of LDVs and explore their integration within autonomous systems. The following objectives were pursued to achieve this aim:

1. **Measurement correction:** The objective is to improve the existing measurement correction techniques for LDVs. This includes developing new time domain signal processing techniques to handle non-stationary or transient vibration signals. Additionally, enhancing frequency domain-based techniques to outperform the original technique.
2. **Arbitrary tracking:** This objective involves developing a novel instrument motion-tracking solution suitable for downward-facing drone applications. The aim is to reduce beam motion during extreme flight conditions, thus, enabling high-quality measurements.
3. **Explore potential in-field mobile applications:** This objective focuses on investigating the application of LDVs in autonomous systems. The specific application explored herein is non-contact vibro-acoustic object recognition by utilising remote acoustic excitation of objects and advanced signal processing using convolutional neural networks.

By achieving these aims and objectives, this thesis aims to contribute to the advancement of LDV technology and its integration within mobile and autonomous systems, expanding its application envelope and enabling enhanced capabilities in various domains such as vibration engineering, autonomous systems, and hazard detection.

1.4 Thesis overview

This thesis can be divided into three main areas. The first part of this thesis is concerned with instrument motion in the beam axis, whereas the second is concerned with motion in the non-beam axes. The combination of these two research areas develop the groundwork for a truly robust LDV capable of obtaining high-quality vibration data during mobile deployment scenarios. Therefore, the third and final focus of the thesis is on new applications of this robust LDV system, specifically focusing on autonomous vehicles. A visual representation of these three main topics and their inter-dependencies is shown in

Figure 1.4.

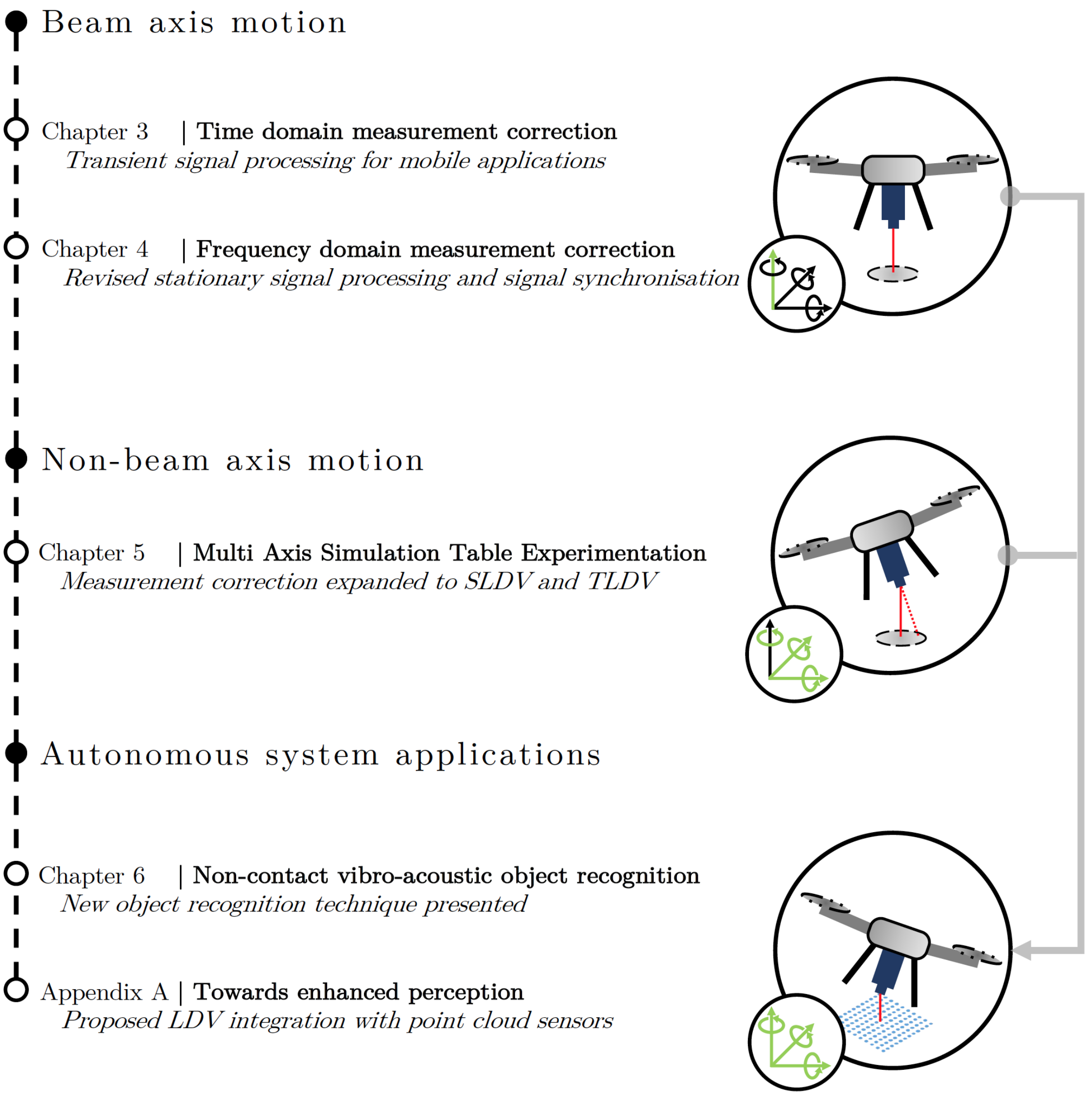


FIGURE 1.4: Conceptual diagram of this thesis highlighting the three main research areas and their inter-dependencies.

These specific topics have translated into the following thesis structure:

Chapter 2 presents an overview of the fundamental physical concepts surrounding laser Doppler vibrometry vibration measurements.

Chapter 3 describes the development of a time domain-based measurement correction technique, comparing it with the existing frequency domain counterpart. An accessible time domain correction technique is essential for mobile LDV scenarios where vibrations of the instrument will likely be transient in nature. Results show that this time domain-based technique outperforms the existing frequency domain-based technique by a significant margin for stationary signal types.

Chapter 4 focuses on the development of an improved frequency domain-based technique with proper accelerometer signal handling. Resulting in a significant performance increase, almost matching that of the time domain-based technique presented herein. Following this, it explores the fundamental differences of signal synchronisation in the frequency and time domains, presenting an analytical model relating the quality of the measurement correction to signal synchronisation and sampling frequency, therefore, optimising performance *a priori*.

Chapter 5 is a two-part chapter employing the MAST for both measurement correction and arbitrary path tracking for SLDV systems. First, the chapter explores the means by which motion along the beam axis of an SLDV can be removed based on previous theoretical work, but with the new signal processing techniques presented herein. This is then experimentally validated on the MAST, displaying significant improvements. Following which, the focus shifts to motion in the five non-beam axes. By considering the specific case of a hovering drone, a novel type of gimbal is presented and tested, which uses an unmodified SLDV to counter-rotate the beam and maintain a fix target of interest. This is similarly validated on the MAST.

Chapter 6 describes a novel non-contact vibro-acoustic object recognition technique intended for autonomous systems. This technique utilises a custom set-up involving a synchronised loudspeaker and scanning LDV to simultaneously, remotely solicit and record responses an acoustic excitation in various objects. Using a range of spectrogram-based pre-processing techniques and convolutional neural networks trained via transfer learning, objects could be recognised with accuracies of over 99%, including tests with never-before-seen instances of objects. Therefore, positioning this as a viable object recognition technique suitable for various machine automation tasks

Chapter 7 is the final chapter which presents a summary of the conclusions gained from this doctoral program and explores some areas for further work.

Appendix A presents the practical extension of the work described in Chapter 6. Here, an additional point cloud mapping modality is used in conjunction with the SLDV, allowing an autonomous system to locate objects of interest within a scene for vibro-acoustic interrogation using an SLDV.

1.5 Contributions and publications

The main contribution of this thesis is the development of a truly robust LDV platform capable of deployment on highly mobile platforms, with a focus on autonomous system applications. Specifically:

1. A novel time domain-based correction measurement technique able to outperform the existing frequency domain-based signal processing technique, introducing transient signal compatibility and increasing performance by a factor of eight [51].
2. An improved frequency domain-based correction measurement technique able to outperform the existing frequency domain-based signal processing technique by a factor of seven, and an analytical model able to accurately predict the influence of signal synchronisation on measurement correction performance [52].

3. The extension and experimental validation of these new signal processing techniques to scanning laser Doppler vibrometer systems.

4. The development of a novel type of gimbal compatible with SLDV systems intended for drone applications.

5. A novel non-contact vibro-acoustic object recognition technique and enhanced spatially associated vibro-acoustic perception for autonomous systems [53]

List of papers published in the course of completing this research:

- Correction of laser Doppler vibrometer measurements affected by sensor head vibration using time domain techniques [51] - Published in the Proceedings of the International Conference on Structural Dynamics, Eurodyn.
- A comparison of time and frequency domain-based approaches to laser Doppler vibrometer instrument vibration correction [52] - Published in the Journal of Sound and Vibration.
- Non-contact vibro-acoustic object recognition using laser Doppler vibrometry and convolutional neural networks [53] - Published in Sensors Special Issue: Artificial Intelligence-Based Audio Signal Processing.
- Towards real-time vibro-acoustic classification, verification and tracking of in-flight UAVs [54] - Published in the Proceedings of the International Conference on Noise and Vibration Engineering (ISMA) 2022.

Chapter 2

Laser Doppler vibrometry background

This chapter presents an overview of the fundamental concepts required to understand laser Doppler vibrometry as this is the consistent theme throughout this thesis and requires the deepest understanding. In-depth reviews of related work pertaining to each chapter can be subsequently found at the beginning of each chapter.

2.1 Physical principles and limitations of laser Doppler vibrometry

To more easily understand the working principles of an LDV, it is best to start with the aforementioned Michelson interferometer. A Michelson interferometer is a precision instrument used to measure changes in optical path length on the order of the wavelength of light being used. This is made possible by taking advantage of the wave-like nature of light by observing the interference pattern produced when two coherent beams of the same wavelength are mixed.

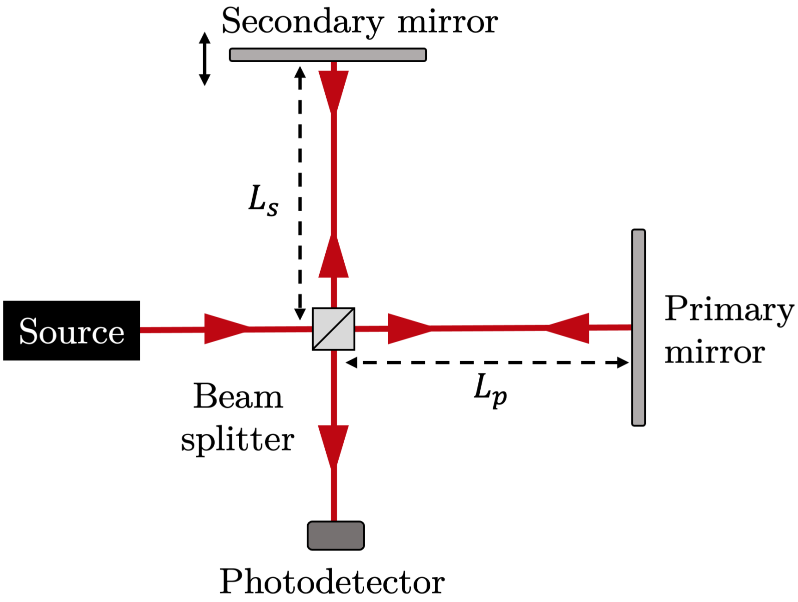


Figure 2.1: Schematic of a basic Michelson interferometer. Where L_s and L_p are the optical path lengths from the cube beam splitter to the secondary and primary mirror, respectively.

Practically, this is achieved using the optical arrangement shown in Figure 2.1. A monochromatic coherent light source, i.e. a laser, is directed onto a beam splitter. The role of the beam splitter is to allow the transmission of half of the light towards the primary mirror while reflecting the other half of the light towards the secondary mirror[1]. Upon reflection by their respective mirrors, the two beams are recombined at the beam splitter such that they may interfere with one another on the path towards the photodetector. If the optical path length difference of the two beams is an integer number of wavelengths, constructive interference will occur, and maximum brightness will be measured by the photodetector. Similarly, if the optical path length difference is a half-integer number of wavelengths, then destructive interference will occur and a minimum brightness will be measured. Therefore, by tracking the changing brightness at the photodetector, precise measurements can be made of any movement in a mirror, shown here with the secondary mirror undergoing motion.

---

[1] Here, a cube beam splitter is shown, however, a half-silvered mirror can similarly be used, typically with the addition of a compensator plate to make both optical path lengths equal.

The Michelson interferometer arrangement was modified by Y. Yeh and H. Cummins in 1966 to enable velocity measurements of a flowing fluid, known as laser Doppler anemometry [55]. Rather than one of the beams being reflected from a mirror, it was made to be incident upon a flowing fluid containing monodispersed polystyrene spheres with a diameter of 0.6 µm. This beam, now known as the *measurement beam*, was backscattered by these particles with a Doppler-shifted frequency proportional to the velocity of the fluid, given by:

$$\Delta f = \frac{2n|\vec{v}|}{\lambda} \cos(\gamma) \sin\left(\frac{\varepsilon}{2}\right) \tag{2.1}$$

where Δf is the Doppler shifted frequency, n is the refractive index of the medium, $|\vec{v}|$ is the magnitude of the fluid velocity, λ is the wavelength of the source light, ε is the angle between the incidence and scattering directions, and γ is the angle between $\vec{v}$ and the bisector angle between the incidence and scattering directions. The setup can be seen depicted in Figure 2.2, where ε is 30° and $\gamma = \pi - \varepsilon/2$.

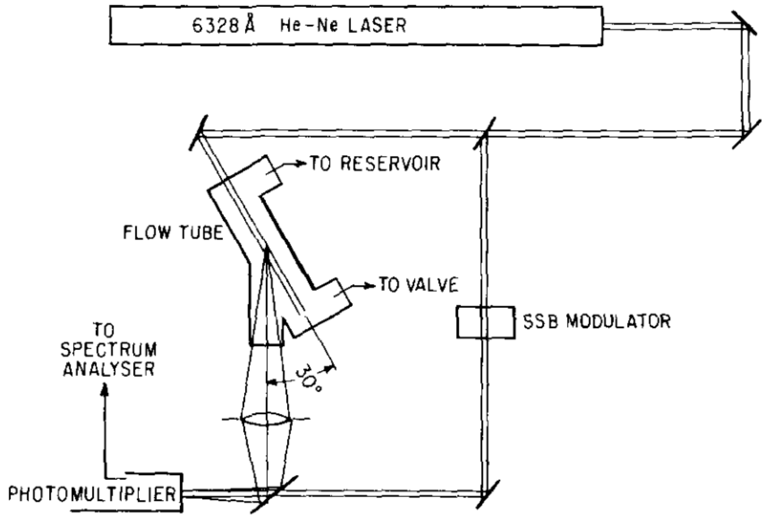


Figure 2.2: A depiction of the experimental arrangement used to develop laser Doppler anemometry [55]. © AIP Publishing, reprinted with permission.

Using Equation 2.1, it can be shown that the expected frequency shift for light of non-astronomical velocities is many orders of magnitude smaller than the frequency of light itself ($\approx 10^{15}$ Hz). Therefore, the inclusion of an interferometric arrangement is essential as the photodetector is not sufficiently responsive to directly measure the changing electric and magnetic fields of the light. Practically, this means mixing the backscattered measurement beam with the beam reflected from the internal mirror, known as the *reference beam.* Since the backscattered measurement beam frequency has been Doppler shifted, the recombined beam will periodically constructively and destructively interfere with itself, modulating the intensity with a frequency known as the *beat frequency.* The frequency of this modulation is directly proportional to the fluid's velocity and is measurable by the photodetector. Hence, the fluid's velocity could be determined by demodulating the signal from the photodetector using a spectrum analyser.

In the following years, it was quickly realised that this setup can be similarly used to take velocity measurements of solid surfaces since light-scattering surface elements can play the same role as scattering particles within a fluid. In this scenario, light is collected in direct-backscatter, meaning $\varepsilon = \pi$, and $n = 1$ for air; thus, by rearranging and simplifying Equation 2.1, the following is obtained:

$$v = \frac{\lambda \Delta f}{2\cos(\gamma)} \tag{2.2}$$

where v is the surface velocity and $\cos(\gamma) = 1$ if the measurement beam is normal to the surface. A depiction of this scenario is shown in Figure 2.3, related by $\Delta f = f - f_0$ and $v = |\vec{v}|$. Here, light is incident upon an optically rough surface such that light may still be collected in direct backscatter by the instrument when the vibrating surface is not normal to the beam.

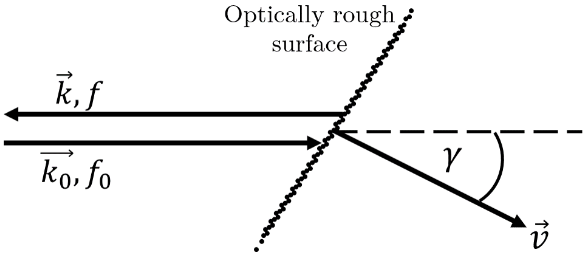


FIGURE 2.3: Depiction of light undergoing direct backscatter from an optically rough surface moving with a velocity vector of $\vec{v}$ and an angle of γ to the incoming light. The incoming light has a wave vector $\overrightarrow{k_0}$ and frequency f_0, whereas the backscattered Doppler-shifted light has a wave vector of $\overrightarrow{k}$ and frequency of f.

However, the optical arrangement previously described suffers from directional ambiguity, meaning positive and negative velocities cannot be distinguished. In order to distinguish positive and negative velocities, a known frequency shift is introduced to the reference beam [56, 57, 58]. This results in a beat frequency that increases if the scattering surface moves towards the instrument and decreases if the scattering surface moves away. While this frequency shift can be introduced in several ways, the most popular method utilises an acousto-optic modulator known as a Bragg cell [59]. Therefore, the surface velocity of the target is related to the beat frequency by the following relationship [60]:

$$f_{\text{beat}} = \left| f_R - \frac{2}{\lambda} v(t) \right| \tag{2.3}$$

where f_{beat} is the beat frequency, f_R is the frequency shift, λ is the light wavelength and $v(t)$ is the target surface velocity. Practically, the target surface velocity is recovered by treating the output of the photodetector as a typical frequency-modulated signal; with the Bragg cell frequency as the carrier frequency and the Doppler shift frequency as the modulation frequency. This can be demodulated to obtain the surface velocity by using a range of frequency demodulation techniques such as single-chip frequency modulated decoders [61], or quadrature detection [62]. LDV manufacturers include increasingly complex optical arrangements, demodulation and preprocessing techniques in order to increase the data quality under a range of measurement scenarios. However, the setup described here represents a standard basic arrangement for most commercial LDVs, depicted in Figure 2.4.

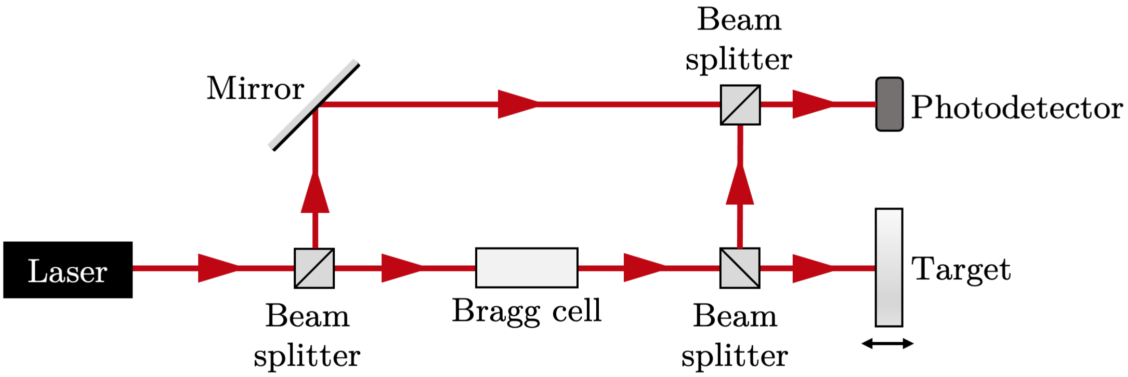


FIGURE 2.4: A schematic of a laser Doppler vibrometer system capable of distinguishing positive and negative velocities using a Bragg cell.

2.2 Signal level and speckle noise challenges

The functional principles of LDV require that a sufficiently high portion of the measurement beam returns from the target surface and is collected by the measurement optics. If an insufficient amount of light is collected, then the amplitude of the aforementioned modulation incident upon the photodetector will decrease in intensity, thus, increasing the uncertainty in the demodulation and increasing the signal-to-noise ratio, potentially compromising the measurement [63, 64]. Light can return from the target via two mechanisms, specular reflection from a surface normal to the beam or diffuse reflection, also referred to as *scattering*. Specular reflection can direct the vast majority of the light back towards the instrument if aligned correctly, giving high signal levels[2]. However, there are two main drawbacks to relying on specular reflection alone. Firstly, aligning the instrument such that it is normal to the reflecting surface in all six DOF is very difficult to achieve practically. Secondly, measurements are limited to those with reflecting surfaces.

More often, LDV measurements will rely on the collection of *diffusely* reflected light. Diffuse reflection occurs through two main mechanisms: *surface scattering* and *sub-surface scattering*. Surface scattering occurs when the target surface is rough on the scale of the optical wavelength being used, whereas sub-surface scattering occurs within the superficial

---

[2] The return of all the light would require a perfect mirror.

outer layer of the material [65, 66]. While metals do not exhibit any subsurface scattering, many materials which may seem opaque do in fact cause some subsurface scattering, such as human tissue and some plastics. However, regardless of the exact mechanism, the result in both cases is that some light is collected by the measurement optics in direct backscatter, making measurements of non-normal surfaces possible.

While the role of scattering is essential to most LDV measurements, it can also lead to the introduction of spurious signal content, known as *speckle noise*. This is due to the de-phasing of the scattered monochromatic coherent light, where returning wavelets interfere constructively and destructively with one another resulting in a chaotic distribution of high and low intensities [67], referred to as a *speckle pattern*; an example of which can be seen in Figure 2.5. The measurement optics generally collect light over several speckles mitigating these effects, however, sometimes small adjustments in the measurement location are required to avoid low optical signal levels caused by an unfavourable summation of speckles over the photodetector surface [68]. This phenomenon can also manifest as pseudo-vibration if there is off-axial target motion, such as in-plane vibration, whole body motion, or a continuous scanning scenario, where the measurement beam traverses the target surface throughout the measurement [69, 70]. This is due to the dynamic nature of the evolving speckle pattern, leading to two major issues. Firstly, signal dropouts can occur when the summation of speckles over the photodetector surface is particularly dark, similarly causing sharp deviations in the time trace [4]. Secondly, phase noise can occur during speckle transitions since the phases of the bright speckles are inhomogeneous and randomly distributed [71, 60].

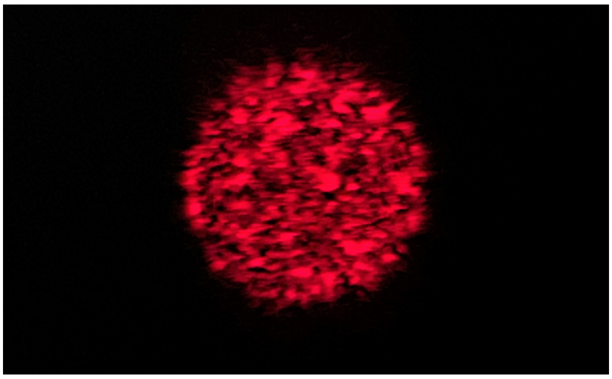


FIGURE 2.5: Photograph of a speckle pattern off black anodised aluminium.

State-of-art LDV systems include increasingly sophisticated attempts to overcome low optical signal and laser speckle challenges caused by the collection of backscattered light. Most recently, it was shown via simulation that the summation of multiple photodetector outputs, all perceiving unique speckle patterns, can mitigate the introduction of this pseudo-vibration [27, 28]. This fundamental concept, known as diversity reception [72], has been recently integrated into the Polytec QTec line of LDVs. However, when these advanced systems are inaccessible or when the techniques are not completely effective, there are several steps that can be taken to mitigate the effects of speckle noise on a measurement. When encountering low optical signal level from an uncooperative target, surface treatment can be used. Surface treatment most commonly comes in the form of retro-reflective tape applied to the measurement surface of the target[3]. Retro-reflective tape consists of many microscopic glass beads bonded to a paper backed with adhesive. The effect of these glass beads is that incoming light is always reflected back in the same direction as the incoming angle [73], as depicted in Figure 2.6. The use of such tape can drastically increase optical signal levels from an uncooperative surface, however, its use can be undesirable or impractical in some situations; for example, on remote infrastructure where access is limited or light structures, where mass-loading and stiffening can take place changing the vibration characteristics.

---

[3] There exist alternatives, such as Ardrox Reflective Spray although often less effective than retro-refective alternatives.

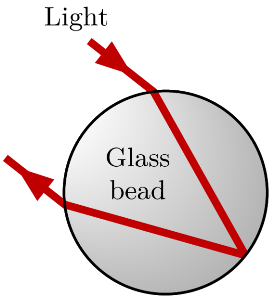


Figure 2.6: Diagram of light being retro-reflected by a glass bead contained in retro-reflective tape.

In scenarios where there is low optical signal level, and surface treatment is not an option, then an infrared LDV may be used. Typical LDVs employ helium-neon laser with a wavelength of 632.8 nm, but infrared LDVs use longer wavelength and more powerful lasers, such as the PSV-500 Xtra Scanning Laser Doppler Vibrometer which is 1550 nm. This difference is due to the increased safety of infrared lasers meaning that the power limit is 10 mW as opposed to 1 mW for visible light lasers [74]. This results in instruments capable of making measurements from uncooperative surfaces and from greater distances.

2.3 Laser Doppler vibrometry and ambient vibration

The fundamental physics of laser Doppler vibrometry dictates that the movement of the sensor head or the target along the beam axis will equally create a Doppler shift in the measurement beam. The result is that the instrument is as sensitive to its own vibration as it is to the target vibration. Therefore, making measurements in environments with high levels of ambient vibrations would negatively affect the data quality. As such, typical implementations isolate the instrument from surrounding vibrations passively using a tripod with compliant feet. However, passive isolation can be inadequate for some applications. For more effective vibration isolation, *active* anti-vibration mounting could be used; however, these solutions can be either too costly or too heavy [25]. Therefore, the development of an LDV insensitive to its own vibration, here referred to collectively as *measurement correction*, has gained increased attention. In general, measurement correction requires independent vibration measurements of any additional velocity contributions to the LDV signal. These additional velocity contributions can be due to a combination of sensor head, scanning head and steering optic vibration; all solutions will be referred to as *LDV measurement correction* herein. This section provides a brief overview of the literature related to measurement correction, with complete reviews contained at the beginning of the relevant subsequent chapters.

The name *measurement correction* refers to the removal of any unwanted velocity contribution of the sensor head from the signal, thereby, fully recovering the target surface vibration velocity. This has been realised in a variety of different ways with varying levels of success. Initial attempts employed a similar approach to that of the differential LDV, but rather than have the secondary beam incident upon an external reference surface, it was incident upon an internal damper [36, 39, 75]. The velocity of this damper, as measured by the second measurement beam, is used to infer the motion of the instrument itself and correct the target measurement. It is important to note that differential LDVs perform this subtraction optically, however, since this damper system acted as an accelerometer, the subtraction was performed electronically following the necessary integration and filtering. Due to the resonances of the internal damper system, a flat frequency response over a sufficiently large range is difficult to obtain, resulting in a correction that is heavily frequency-dependent. Following this, a system was developed which utilised both an accelerometer and an external reference beam to estimate the motion of the instrument on a moving ground vehicle [37]. While this technique significantly reduced the impact of instrument vibration, there were several shortcomings related to the location of the correction measurements and the external reference beam, which limited the performance. Additionally, both the mentioned techniques require complex bespoke optical arrangements making them considerably less accessible for the practising vibration engineer.

Unlike the previous two measurement correction techniques, the third and most recent type of measurement correction utilises a commercial LDV, rather than a bespoke setup, with a single rear-mounted accelerometer [31]; which would be readily available in most vibration laboratories. Despite its relative simplicity, this technique yields a significant reduction of instrument vibration in the measurement signal whilst being considerably more straightforward to implement when compared to the previous attempts. The performance and accessibility of this technique have likely led it to be replicated by other international research groups [40]. This technique has since been expanded to be compatible with scenarios utilising external steering optics, by placing another accelerometer to similarly measure and subtract its vibration contribution [76]; and theoretically shown to be compatible with SLDVs by using three precisely positioned accelerometers whose trigonometric components can act as a measurement colinear and on the beam axis [41]. While the hardware arrangement is well-optimised and is shown to be compatible with a range of application scenarios, there are two drawbacks in the current form. Firstly, the signal processing did not sufficiently handle the accelerometer signal leading to a non-optimal performance. Secondly, the technique is based in the frequency domain, limiting it to stationary signal types. As such, Chapter 3, Chapter 4 and part of Chapter 5 use this hardware arrangement but focus on improving the signal processing.

2.4 Alternative mobile laser Doppler vibrometry techniques

The focus of this thesis is mobile-mounted LDVs, however, there is an alternative proposed technique that is important to discuss, referred to as *flyable mirrors* [32, 33]. This approach focuses on enabling scans of remote surfaces that either have no direct line-of-sight access or would result in a low incidence angle - therefore only measuring a small component of the normal velocity. The approach requires an LDV base station with pitch and yaw control to track a drone-mounted mirror in-flight such that the beam spot is ultimately incident upon the intended target location. An illustration of this can be seen in Figure 2.7.

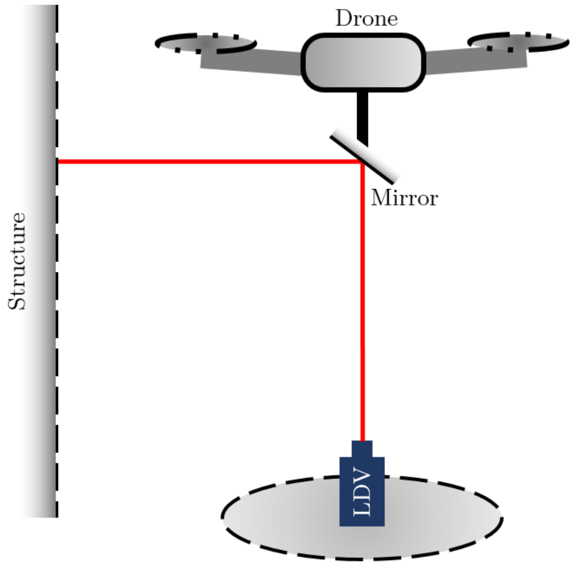


FIGURE 2.7: Illustration of the flyable mirrors concept for beam steering [32].

The advantages of flyable mirrors in comparison to drone-mounted LDVs include lower costs, since smaller drones can be used as the entire instrument is not being lifted; and increased safety, since smaller drones have lower collision risks and potential to cause damage than a larger drone carrying an entire LDV. A laboratory setup was used to simulate a flyable mirror deployment, allowing the efficacy of this technique to be estimated. Here, a shaker was driven with a 5 Hz pure tone - simulating the measurement of interest - while a tethered hovering drone was reflecting the measurement beam towards the shaker spigot. A secondary beam path was also tested which avoided the drone-mounted mirror and recorded the shaker vibration directly, allowing it to be used as a benchmark to compare the effect of the flyable mirror on the measured vibration. It was shown that for this setup, the drone vibration introduced significant signal content in the region of 50 Hz to 70 Hz, which in this scenario was clearly separated from the 5 Hz of the target shaker. However, this clearly defined frequency separation of the target and drone vibration is unlikely to always be the case. As such, it cannot be relied on to distinguish between drone and target vibration.

A follow-up paper shifted the focus onto the removal of the drone's contribution to the measured vibration by employing the aforementioned differential vibrometry approach [33].

In this improved deployment, the measurement beam remains as previously described, being reflected from the ground-based LDV to the structure via a drone-mounted mirror; however, in this scenario the external reference beam is incident upon a retro-reflector attached to the drone. This results in the optical subtraction of the drone's vibration along the reference beam axis from the vibration measurement. An illustration of this can be seen in Figure 2.8.

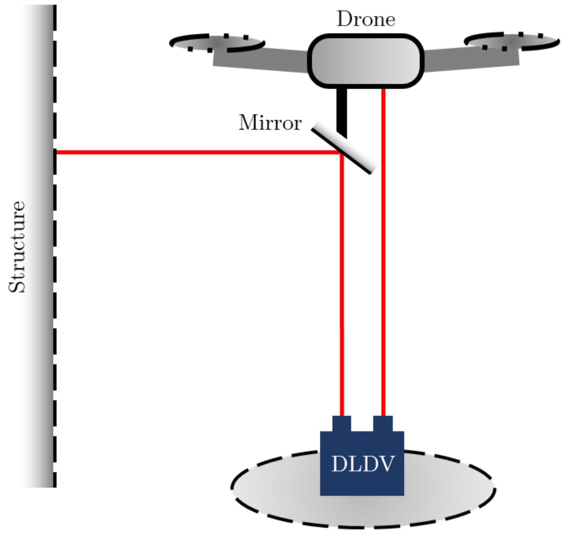


FIGURE 2.8: Illustration of the differential LDV (DLDV) flyable mirrors concept for beam steering *and* measurement correction [33].

It has been previously shown that for optimal subtraction when using external steering optics, the correction measurement should be normal to the mirror surface and as close as possible to the incident beam spot location [76]. This criterion is not satisfied in this configuration - with the subtracted contribution of the drone's vibration occurring along the reference beam axis, not normal to the reflecting steering mirror. While the correction may not be optimal, it may still share enough in common with the normal-mirror axis to lead to an improvement in the signal. A lab-based setup, seen pictured in Figure 2.9, aimed to simulate this deployment scenario, allowing both the LDV and DLDV approaches to be compared to one another in a controlled manner.

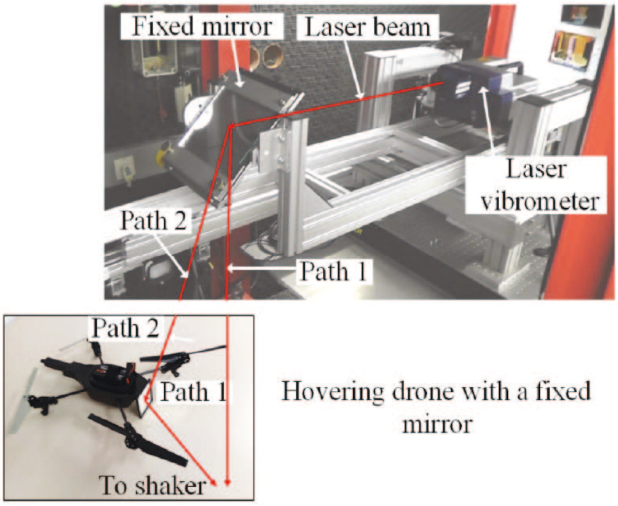


FIGURE 2.9: Flyable mirrors experimental simulation setup aiming to compare the LDV or and DLDV approaches to one another [33]. © 2021 The British Institute of Non-Destructive Testing, reprinted with permission.

Using the setup shown in Figure 2.9, a 10-fold reduction was observed in the overall noise floor of the measurement when compared to the previous single point LDV approach. The results are seen pictured in Figure 2.10, where the 222 Hz vibration of the shaker is clearly visible through either optical path. The technique clearly reduces the introduction of unwanted drone vibration despite its non-optimal arrangement; therefore, positioning this as a viable mobile LDV technique.

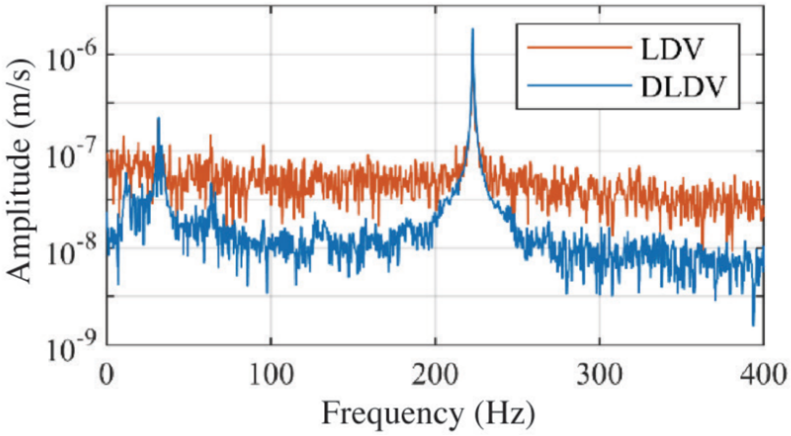


FIGURE 2.10: Flyable mirrors experimental simulation results using either an LDV or an DLDV [33]. © 2021 The British Institute of Non-Destructive Testing, reprinted with permission.

While there is undeniable promise in the unique concept of flyable mirrors, it also presents some technical challenges regarding the simultaneous tracking requirements for the base station and the drone-mounted mirror. While drones have an impressive ability to hold their position, they undergo small amounts of translational motion in flight due to atmospheric effects; as such, modern sensor-laden drones employ an array of sensors to measure their positions in space such that they may be corrected using small adjustments in pitch and roll. These adjustments could make it difficult to maintain a lock on the flyable mirror from the ground-based LDV, as such, a tracking system would likely be required. With a tracking system on the base station, it could be ensured that the beam is incident upon the flyable mirror, however, the drone's motion would result in the beam straying from the intended measurement location on the target. The amplitude of this stray may or may not justify the inclusion of a secondary tracking system on the drone to aim the mirror. These challenges are not dissimilar to those experienced during a drone-mounted mobile deployment, which requires similar measurement correction and aiming capabilities. While drone-mounted mirrors can potentially reduce drone costs and the severity of a crash by offloading some instrumentation to the ground, it similarly ties the drone to that ground station. Restricting movement and similarly increasing the effective standoff distance, while ideal for many applications, most applications discussed herein could benefit from a truly mobile LDV. As such, the remainder of this thesis does not consider this technique.

Chapter 3

Time domain measurement correction

In real-world relevant, field-based vibration measurement scenarios from mobile platforms, it can be reasonably expected that the instrument vibration signals will be *transient* in nature. A signal is considered transient whenever its Fourier expansion requires an infinite number of sinusoids, as opposed to stationary signals, which is expressible as a finite number of sinusoids [77]. Of course, an infinite number of sinusoids is not practically available as this would require an infinitely long measurement duration. As such, when transient vibrations are expected, time domain techniques can be implemented. As discussed in Section 2.3, the optimal hardware arrangement for LDV measurement correction utilises a single-rear mounted accelerometer; however, the compatible signal processing available prior to the work presented herein was based in the frequency domain, therefore, preventing its use in mobile deployments where transient vibrations might be expected. As such, for optimal measurement correction, time domain-based correction techniques are desirable, yet despite this, the available time domain-based techniques were inaccessible to most vibration engineers, requiring the construction of complex bespoke LDVs with unsatisfactory performance [36, 37, 39]. As such, this chapter presents a new time domain-based alternative based on the optimised existing frequency domain hardware arrangement.

First, this chapter will conduct an in-depth review of the time domain-based measurement correction literature. Following this, a new time domain-based technique is described based on the existing single accelerometer frequency domain-based technique hardware arrangement [31]. This is followed by a rigorous performance assessment, contextualising it against the previously established frequency domain-based technique. The work presented in this chapter is orientated around the body of work that was published in the peer-reviewed conference proceedings for the International Conference on Structural Dynamics, EURODYN 2020 [51].

3.1 Review of time domain-based measurement correction techniques

To the author's knowledge, there exists only two time domain-based based LDV measurement correction techniques, both being rather different from one another. This section will outline both techniques and discuss the merits and drawbacks of each based on three main aspects. The first consideration will be the *accuracy*; meaning the technique should yield similar measurements to that of a stationary LDV or contacting transducer - but shown the be the case by means of an empirical and quantifiable metric. The second consideration is *accessibility* - that is, readily available to the practising vibration engineer, and not merely by means of potential eventual commercialisation. And the final consideration is *transient signal compatibility* - since the ambient vibration profiles during mobile deployment are likely to be transient in nature.

3.1.1 Internal damper LDV measurement correction

The earliest mention of LDV measurement correction in any domain was in the year 2000 [36]. The proposed solution featured an internal damper, consisting of a mass suspended between two springs submerged in a viscous fluid. The optical arrangement can be understood as two independent LDVs, one measuring the velocity of the target and one measuring the velocity of the internal damper. Since the velocity of the internal damper is related to the acceleration of the instrument, its velocity measurement is acting as an accelerometer. The demodulated damper signal is then integrated to obtain an instrument velocity estimate, processed via a least means square (LMS) adaptive filter and subtracted from the target vibration measurement, all in the time domain.

In order to test this system, the instrument body was vibrated with two sine waves at 30 Hz and 60 Hz while measuring a target vibrating at 1 kHz. The results are presented in Figure 3.1. While there is a considerable improvement, the velocity trace is still marginally affected by the body vibration. The work does not mention that the damper is unlikely to have a flat frequency response over a sufficiently large frequency range, therefore, the quality of the correction could be heavily frequency dependent. This might have been apparent if the instrument had been vibrated with broadband noise or a chirp excitation over a sufficiently wide frequency range. A rigorous assessment of the system must characterise the correction performance at a wide range of frequencies for it to be a practically viable solution. Similarly, a quantitative performance metric should be introduced to truly gauge the performance of the technique, ideally one utilising a simultaneous reference measurement of the target.

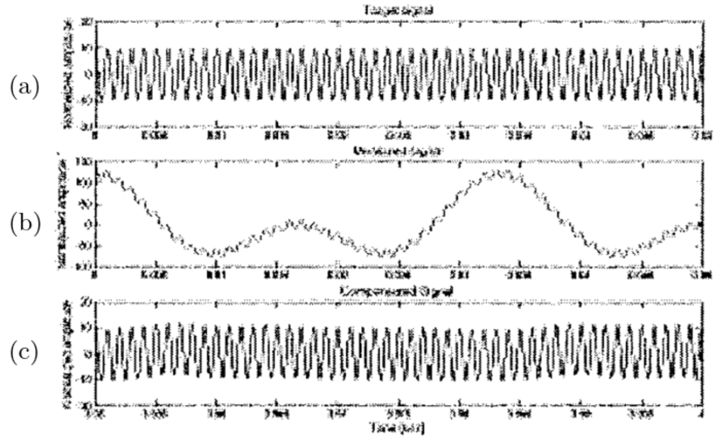


FIGURE 3.1: The time domain signal of (a) the target vibrating at 1 kHz, (b) the measured target vibration, and (c) the corrected measured signal. Unfortunately, a better quality digitisation of the paper does not exist. Here, the vertical axes are velocity and the horizontal axes are time [36]. © 2000 IEEE, reprinted with permission.

In the following year, two follow-up papers covering the same system but in more detail were published [39, 75]. The first was a single-page summary paper that added little new information, however, the results of a second test were described. Here, the LDV body was vibrated in the beam axis with a 1 Hz, 5 cm amplitude sinusoid while the target was vibrating with a 100 Hz 50 µm sinusoid. The results can be seen in Figure 3.2, despite being difficult to interpret, there does seem to be some improvement between the uncorrected measurement in Figure 3.2(c) and the corrected measurement in Figure 3.2(d).

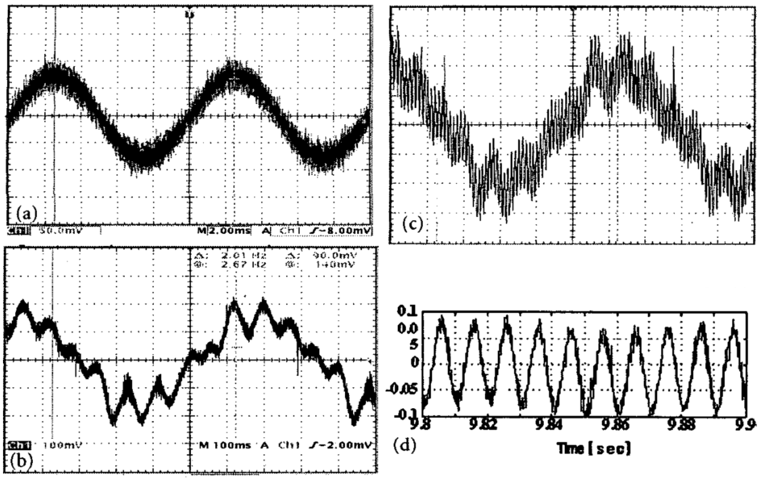


FIGURE 3.2: The time domain signal of (a) the target vibration, (b) instrument vibration, (c) uncorrected LDV signal, and (d) corrected LDV signal [75]. © 2002 IEEE, reprinted with permission.

The second follow-up paper tackled the issues presented by the non-flat frequency response of the damper [39]. The equation of motion for the damper was shown to be:

$$M\ddot{x}_d + C\dot{x}_d + Kx_d = -M\ddot{x}_l \tag{3.1}$$

where M is the mass of the damper, C is the viscosity of the fluid, K is the spring constant, x_d is the displacement of the damper mass with respect to the LDV body and x_l is the displacement of the LDV body to a virtual fixed plane. A damping ratio of $C/M = 0.707$ was selected as this provides a “constant gain” when the instrument vibration frequency is much lower than natural frequency of the damper. The first natural frequency was selected to be at 60 Hz since, the authors claim, that instrument vibrations tend to be much lower than this. However, in reality the expected instrument vibration range is heavily dependant on the application, with vibration profiles during mobile deployment likely to be higher than this.

The effectiveness of the system was tested in a similar fashion to the previous paper but with a single 1.5 Hz sinusoidal instrument vibration. While the team qualitatively claim to have experimentally proven effective removal of the instrument vibration within the range of 1 Hz to 5 Hz, there is no data presented to support this, so the true bandwidth of the measurement correction is unknown. However, taking the stated range to be true, it would still likely be insufficient for mobile deployments. Similarly, it would have been informative to compare the corrected signal with a reference measurement so that some quantitative performance metric could be presented. Despite the poor performance range, this approach is still rather inaccessible as it requires complex bespoke optical arrangement, therefore, would require manufacturing, assembly and optical alignment of the LDV.

3.1.2 External reference beam and accelerometer LDV measurement correction

Later work published in 2011 employed both an accelerometer and an external reference beam to remove the contribution of the instrument vibration [37]. While both the accelerometer and the external reference beam will measure and subsequently subtract some instrument vibration, it's made clear that they have subtly different roles. The accelerometer is used to estimate the vibration of the instrument, while the external reference beam is intended to mitigate the effects of the parasitic coupling introduced by vibrations in the optical fibres used in their setup.

Unfortunately, very little is said in terms of accelerometer signal processing as the paper focuses on bespoke demodulation and Doppler tracking electronics. The accelerometer is described as being subtracted from that of the LDV in a "line-of-sight" measurement. It could be assumed the authors subtracted the *component* of the measurement in the beam direction, however, this is unclear. Similarly, there does not seem to be a signal synchronisation stage prior to the subtraction of the integrated accelerometer signal from the LDV measurement, which is important since the differing signal condition of each transducer type will undoubtedly impact the total obtainable potential performance.

The entirely bespoke system was designed to increase land coverage rates, so it employed a downward-facing five-beam LDV mounted to a moving ground vehicle. However, the nature of this setup prohibited the acquisition of a simultaneous reference measurement of the target vibration as the target location is in motion. Therefore, the system's noise floor is used rather than a reference measurement to assess the measurement correction performance. This experimentally measured noise floor was compared to a theoretically expected noise floor for an instrument insensitive to self-vibration. This was modelled as a combination of shot noise and the aforementioned speckle noise. Shot noise occurs in optical devices since the number of backscattered discrete photons can vary stochastically between samples and is worsened during low-signal conditions with its effects located at the higher acoustic frequencies (>5 kHz). A rigorous analytical derivation for the shot noise is presented in the work which builds on previous works [78, 79], finally yielding the following expression for shot noise:

$$A_{v,sh}(f) = \frac{f\lambda}{\sqrt{\phi_{pe}}} \tag{3.2}$$

where $A_{v,sh}(f)$ is the amplitude spectrum as a function of the frequency f, λ is the wavelength, and ϕ_{pe} is the received photoelectrons per second from the photodetector [37]. Whereas speckle noise mainly contributes to low frequencies (<1 kHz). The paper postulates that the square root of a Lorentzian function can effectively model speckle noise. This postulation was based on several stated factors; firstly, the functional behavior of the Lorentzian function aligned with the observed patterns highlighted by existing work [80]; secondly, the Lorentzian function was chosen for its continuous nature, providing convenience in mathematical modelling; lastly, the Lorentzian function exhibited a meaningful autocorrelation function. Therefore, speckle noise is modelled as the following:

$$A_{v,sp}(f) = \lambda\sqrt{\frac{\pi f_{\text{exc}}^2}{12}}\sqrt{\frac{2\alpha}{\alpha^2 + (2\pi f)^2}} \tag{3.3}$$

where $\alpha = 2\pi f_{\text{exc}}$, with f_{exc} being the exchange rate of the speckle pattern and all other variables are as previously defined [37]. Finally, in order to calculate the total theoretical noise floor due to these effects alone, the speckle noise and shot noise expressions are combined using the following:

$$A_v(f) = \sqrt{[A_{v,sh}(f)]^2 + [A_{v,sp}(f)]^2} \tag{3.4}$$

Figure 3.3 shows the original, uncorrected signal from the LDV at three different cruising speeds along with the theoretically expected noise floor, given by Equation 3.4, Equation 3.3 and Equation 3.4. While for higher frequencies, the experimental measurements match the theory, this is not the case for lower frequencies where the measured noise floor is significantly higher than the theoretical noise floor. This is as would be expected and is caused by the introduction of mechanical vibrations from the moving ground vehicle. There also seem to be a number of resonances being excited, the largest of which is located at roughly 50 Hz and 100 Hz and seem to be somewhat unaffected by the cruising speed.

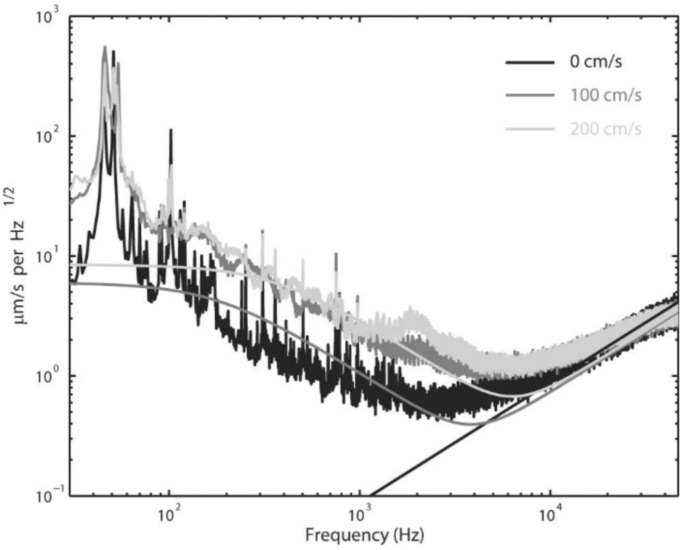


Figure 3.3: Amplitude spectrum of the surface velocity as measured with the laser Doppler vibrometer, uncompensated by accelerometer data or external reference beam. The smooth curves are the theoretically expected noise floors with the remaining curves being the experimental measurements. The colours represent the cruising speeds of the experimental and theoretical pairs [37]. © 2011 The Optical Society, reprinted with permission.

Figure 3.4 shows the data following the subtraction of only the accelerometer signal. This appears to significantly reduce the noise floor at the lower frequencies when compared to the data shown in Figure 3.3. This reduction is particularly apparent on the two resonances located at 50 Hz and 100 Hz which are reduced in amplitude by nearly a factor of 10. However, the measured data does not seem to agree well with the theoretically derived noise floor at each cruising speed.

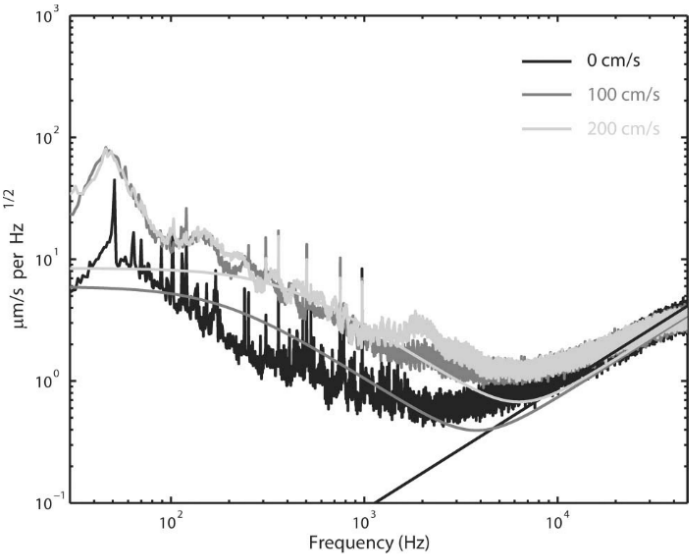


Figure 3.4: Amplitude spectrum of the surface velocity as measured with the laser Doppler vibrometer, corrected with accelerometer data. Solid smooth curves are theoretically expected noise floors due to shot and speckle noise [37]. © 2011 The Optical Society, reprinted with permission.

Finally, Figure 3.5 shows the data following the subtraction of the accelerometer signal *and* the external reference channel. Now the 100 Hz peak is entirely missing, and the 50 Hz peak has undergone a further significant reduction in amplitude. Interestingly, the optical subtraction of the external reference mirror seems to have increased the amplitude of the noise at around 1000 Hz. Despite this, the agreement between the theoretical model and the measurements is still questionable below 1000 Hz, despite the technique leading to a significant reduction in the measured instrument vibration.

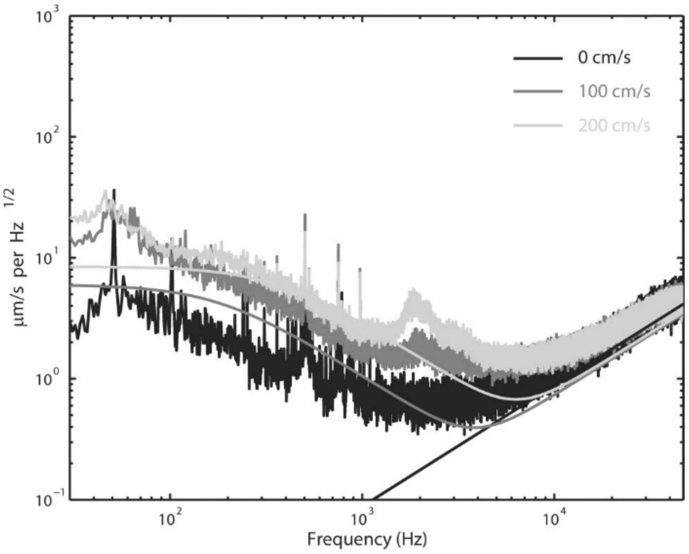


FIGURE 3.5: Amplitude spectrum of the surface velocity as measured with the laser Doppler vibrometer, corrected with both the accelerometer data and the optical reference channel. Solid smooth curves are theoretically expected noise floors due to shot and speckle noise [37]. © 2011 The Optical Society, reprinted with permission.

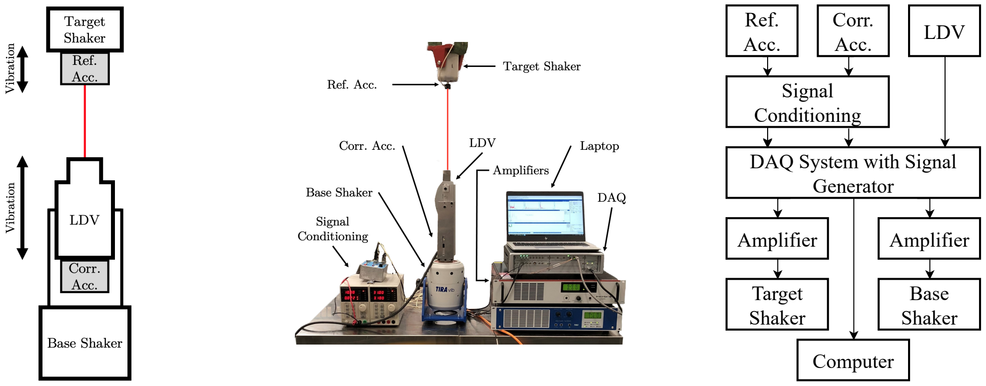
While this technique yields a significant reduction in instrument vibration, there are two main shortcomings. Firstly, the use of an external reference is limited to the quality of the reference surface, if one is available at all. In this scenario, the reference beam is adjacent to the measurement beam on the incoming ground, leading to the optical subtraction of large portions of genuine signal content as the reference location will be undergoing similar motion to the measurement location. The paper acknowledges this by stating that "often, it is more interesting to measure the presence of a vibration rather than quantify its exact amplitude". The second shortcoming of the system is the significant complexity of its implementation. Not only do the optics need to be assembled and aligned, but all the accompanying signal processing electronics would also need to be constructed and configured. Therefore, identifying a major gap within the existing time domain measurement correction works, that is, a technique that can be easily adopted by the practising vibration engineer to make accurate frequency *and* amplitude vibration measurements while being insensitive to self-vibration over a sufficiently wide frequency range.

3.2 A new time domain-based processing technique

Considering that the three main criteria for a correction technique outlined in this chapter are *accuracy*, *accessibility* and *transient signal compatibility*, then it becomes clear that, independently, none of the existing time domain techniques are sufficient. As such, in order to satisfy all three criteria, this section describes the development of a time domain-based technique which utilises the same hardware arrangement as the existing frequency domain-based work described in Section 2.3, but with new time domain-based signal processing; making it compatible with transient signal types and therefore satisfying all three criteria. This frequency domain technique signal processing is described in more detailing the following chapter; however, for the purposes of this chapter, it is only the hardware arrangement that is germane.

3.2.1 Hardware arrangement

It is important for the development of the correction techniques to construct a test setup that simulates real-world LDV vibration scenarios, meaning there should be a vibrating target of interest and a vibrating LDV. Figure 3.6 illustrates the experimental setup used in this work, it allows for the independent control of both the target and the LDV vibration. This setup is similar to the experimental setup employed in existing frequency domain measurement correction work [31]. Here, the target vibration is the measurement of interest, while the base vibration simulates the effects of instrument vibration on the LDV measurement. Both the target and the base vibrations were realised using electrodynamic shakers independently driven using uncorrelated broadband white noise up to 200 Hz, generated by a Siemens Digital Industries Software Simcenter SCADAS Mobile data acquisition system and accordingly amplified. The base shaker was a Tira Vibration exciter S 51120 amplified by a Tira Vibration BAA 500, and the target shaker was a Brüel & Kjær V201 M4-CE amplified by a Brüel & Kjær LDS LPA100. While a flat shaker/amplifier response over the frequency range of interest may be desirable, it is not essential since the correction algorithm should be effective, irrespective of level and phase, across the frequency range.



(a) General schematic of the setup used for this work.

(b) Physical setup with the laser beam highlighted.

(c) Block diagram representation.

FIGURE 3.6: Experimental setup used to simulate a vibration measurement while the instrument is experiencing vibration. The labels "Corr. Acc." and "Ref. Acc." represent the correction and reference accelerometers, respectively.

A custom-made aluminium mounting bracket was used to fix a Polytec NLV-2500-5 Compact Laser Vibrometer to the base shaker so that the laser beam axis was aligned with that of the vibration. An Endevco 770F-10-U-120 (20.4 $\mathrm{mV/ms^{-2}}$ nominal) DC-response accelerometer was mounted to the bracket with its sensitive axis colinear to and on the beam axis. The target shaker was suspended directly above the LDV from an overhead crane, providing isolation from the large base shaker. A second Endevco accelerometer of the same model was mounted to the spigot providing the 'true' vibration measurement. As in earlier work [38], a second, *fixed* LDV could equally be used for the true vibration measurement. However, one or both beams would need to be off-axis to enable optical access, which may require the angular misalignment to be determined and accounted for. Similarly, the vibration produced by the large base shaker may be hard to isolate from the second LDV. Therefore, the use of the reference accelerometer was preferred since all sensitive axes could be aligned without obstruction, and no vibration isolation was required since the target shaker was mounted from an overhead crane.

The usage of accelerometers, while practical, is not without limitations. Primary among these is the flatness of their amplitude and phase response. Unlike LDVs, accelerometer performance is typically limited by the first mechanical resonance of the mass-spring system. However, they are relatively low cost, are readily available and can offer acceptable performance in the context of mobile LDV measurement campaigns, which are the focus of the solutions developed in this body of work. In general, a frequency range from several Hz to several hundred Hz is considered appropriate with vibration levels on the order of several tenths to several tens of $\mathrm{mm\,s^{-1}}$. Over such a relatively narrow frequency range, it is appropriate to compensate for the amplitude and phase response with a straightforward relative calibration (to the LDV), and this will be described in detail subsequently.

To better understand some of the other limitations of accelerometers, it is important to cover their construction and operational principles briefly. As mentioned previously, the core of an accelerometer is a mass-spring system. When the accelerometer is subjected to any form of mechanical shock or vibration, this causes the spring to move or deflect from its equilibrium position. This deflection is then converted into an electrical signal, typically through a transducer which could be capacitive, piezoelectric, or piezoresistive. The transducer transforms the mechanical motion into an electrical signal which can be amplified, conditioned, and eventually measured. The resulting electrical signal is thus a representation of the measured acceleration. However, translating mechanical motion into an electrical signal through the mass-spring system comes with issues, such as noise. This can be inherent in the transducer itself (like thermal noise), or it can be external, like shot noise (caused by the quantum nature of electricity), mains noise (interference from electrical power lines), amplifier noise, or cable sensitivity to vibration. These noise factors can introduce unwanted signals that can interfere with the desired accelerometer signal, thereby reducing the precision and accuracy of the measurements. While these would somewhat affect the final obtained performance, the relative amplitudes of these tend to be orders of magnitude smaller than the measurement of interest, therefore, they are not considered further.

Another limitation of accelerometers is that they typically exhibit a small amount of transverse sensitivity, which might degrade correction performance in the presence of significant off-axis vibration; in this case, it is only 3% [81]. In the experimental setup used here, the inevitable rocking motion of the shaker is minimised by centring the mass distribution on the vibration axis; therefore, this effect is considered negligible. Nevertheless, this and some misalignment between the shaker axes also results in some motion of the LDV beam on the target. While this motion was insufficient to cause the laser beam to deviate substantially from the region of interest on the target, pseudo-vibrations in the LDV signal, which include speckle noise, are associated with such relative motion of the laser beam across the target surface, and these cannot be corrected by the means proposed in this paper. However, combined LDV sensitivity to transverse vibration as a result of both phenomena is on the order of 0.1% [82] and is therefore also considered to be negligible in the context of sensitivity to sensor head vibration.

The Siemens acquisition system was used to record the various time data throughput vibration signals at the maximum sampling frequency of 204.8 kHz for a duration of 8 s. This high oversampling factor assists in the accurate synchronisation of the three signals in the time domain. The acquired data were processed as five separate 1.6 s data lengths. These were similarly processed in the frequency domain for comparison with these acquisition parameters leading to a spectral resolution of 0.625 Hz and a bandwidth of 102.4 kHz. An identical hardware arrangement is also used in Chapter 4; however, this section is not repeated.

3.2.2 Performance metrics

Ultimately, and only possible in the lab-based experimental validation, the reference accelerometer can be used to quantify the efficacy of the correction procedure. This is achieved by comparing the corrected LDV signal with the processed reference accelerometer signal. Signal similarity can be quantified in a number of ways; however, to keep this work within the context of the existing frequency domain-based work, the same metric is adopted - known as the *error reduction*. This metric is essentially the ratio of the error in the signal before and after correction on a logarithmic scale [38]. When taking a mean across multiple measurements as done here, then the formulation for the mean error reduction can be written as:

$$\overline{R} = -10\log_{10}(\overline{r}) \ \text{dB} \tag{3.5}$$

with:

$$\overline{r} = \frac{1}{M}\sum_{m=1}^{M}\frac{\text{MSE}_m^{\text{corr}}}{\text{MSE}_m^{\text{meas}}} \tag{3.6}$$

where $\text{MSE}_m^{\text{corr}}$ and $\text{MSE}_m^{\text{meas}}$ are the mean squared error of the signal before and after correction, respectively, and M is the total number of data lengths ($M = 5$ here). Similarly, the uncertainty in $\overline{R}$ is taken to be the standard error of the mean, SEM, expressed in the form $\overline{R}^{+SEM}_{-SEM}$ throughout this thesis. The SEM is calculated using:

$$\pm SEM = -10\log_{10}\left(\overline{r} \pm \frac{\delta}{\sqrt{M}}\right) \ \text{dB} \tag{3.7}$$

where δ is the standard deviation of the M samples of r. Given by:

$$\delta = \sqrt{\frac{1}{M-1}\sum_{m=1}^{M}(r_m - \overline{r})^2}. \tag{3.8}$$

where r_m is the calculated r for the mth data length. So far, all equations in this section have been kept general such that they apply to both the time and frequency domains. To calculate the error reduction in a specific domain, then $\overline{r}$ is calculated using a domain-specific formulation of the MSE.

To calculate $\overline{r}$ in the time domain, the following formulation of the MSE is used:

$$\text{MSE}_m^{\text{signal}} = \overline{\left(U_m^{\text{signal}}(t) - U_m^{\text{true}}(t)\right)^2} \tag{3.9}$$

where $U_m^{\text{true}}(t)$ is the velocity obtained from the reference transducer, $U_m^{\text{signal}}(t)$ is either the measured or corrected LDV signal and $\overline{(\cdot)}$ signifies the time average. Whereas, to calculate r in the frequency domain, the following formulation of the MSE should be used for the mth spectra of N spectral lines [41]:

$$\begin{aligned}\text{MSE}_m^{\text{signal}} =& (a_{0,m}^{\text{signal}} - a_{0,m}^{\text{true}})^2 + \\ & \frac{1}{2}\sum_{n=1}^{N}(A_{n,m}^{\text{signal}} - A_{n,m}^{\text{true}})^2 + (B_{n,m}^{\text{signal}} - B_{n,m}^{\text{true}})^2\end{aligned} \tag{3.10}$$

where $A_{n,m}^{\text{signal}}$ and $B_{n,m}^{\text{signal}}$ are the real and imaginary parts, respectively, of either the measured or corrected LDV signal at the nth spectral line for the mth spectra The same notation applies to $A_{n,m}^{\text{true}}$ and $B_{n,m}^{\text{true}}$, which are the reference accelerometer equivalents. Similarly, $a_{0,m}^{\text{signal}}$ and $a_{0,m}^{\text{true}}$ are the DC component equivalents.

3.2.3 Accelerometer sensitivity and time delay estimation

Prior to any measurement correction, a one-off relative sensitivity determination and time delay estimation should be carried out to optimise performance. To do this, a vibration measurement is obtained from all transducers with their sensitive axes aligned. Practically achieved using an arrangement with the LDV positioned directly above and focused on an accelerometer stack mounted to the target shaker's spigot. Care was taken to eliminate contamination from ambient vibration by placing the entire arrangement on an anti-vibration base, as shown in Figure 3.7. This setup also assisted in keeping all sensitive axes aligned. Aside from the omission of the base shaker, all other components on the signal conditioning and acquisition side were identical to that described in Section 3.2.1, as were the acquisition parameters.

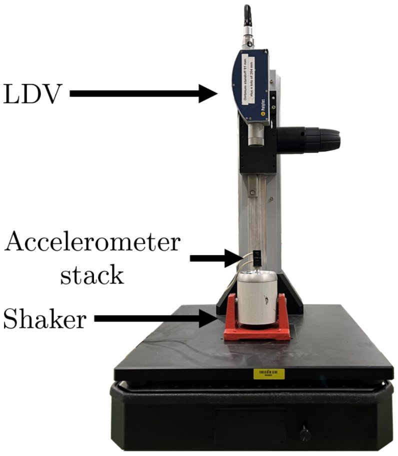


FIGURE 3.7: Annotated photograph of the calibration setup used.

While a secondary setup has been used here for the calibration measurement, it is not strictly required. The setup depicted in Figure 3.6 could be sufficient by enabling only the target shaker to calibrate the reference accelerometer and only the base shaker to calibrate the correction accelerometer. Similarly, both accelerometers could be simultaneously calibrated by placing them in a stack on the target shaker or onto the LDV mount and then driving the appropriate shaker.

Figure 3.8 schematically depicts this time domain-based relative calibration procedure for a single accelerometer, which is applied to both the correction and reference accelerometers. Since accelerometers measure acceleration but LDVs measure velocity, then some differentiation or integration must occur. While it is possible to differentiate the LDV signal and process the signals as accelerations, it is not ideal. Velocity is considered the optimal vibration parameter for its flat representation of all frequencies [83]; therefore, the accelerometers were integrated, as can be seen in Figure 3.8.

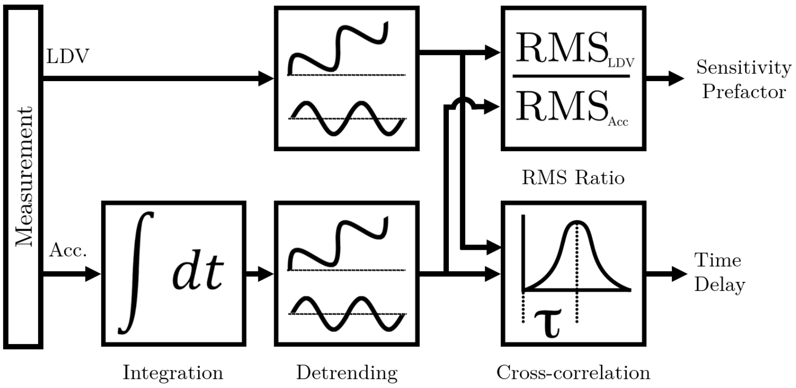


FIGURE 3.8: The time domain-based accelerometer calibration procedure with “Acc.” representing either the correction or the reference accelerometer signal.

Integration is straightforwardly achieved here using the cumulative trapezoidal method. However, the integration of accelerometer signals commonly leads to the introduction of errors such as a DC offset and drift. Detrending is used to remedy this and is achieved by subtracting the least squares fit of a first-order polynomial from the integrated signal. Since this might remove genuine as well as spurious signal content, the LDV signal is subjected to the same for consistency.

With the accelerometer and LDV signals represented as velocities and detrended, a root mean square (RMS) ratio can be used to revise the accelerometer sensitivities. All subsequent measurements are then acquired with this adjusted sensitivity. Meanwhile, the finite time delays between the LDV and accelerometer signals, which occur as a result of differences between the signal conditioning electronics in the measurement chain, are estimated using a cross-correlation function as follows [40]:

$$r_{xy}(\tau) = \frac{1}{T}\int_0^\infty x(t)y(t+\tau)dt \tag{3.11}$$

where τ is the time delay between the signals, x and y are the two signals, t is time and $r_{xy}(\tau)$ the cross-correlation function in which the peak will occur at the time delay between the signals.

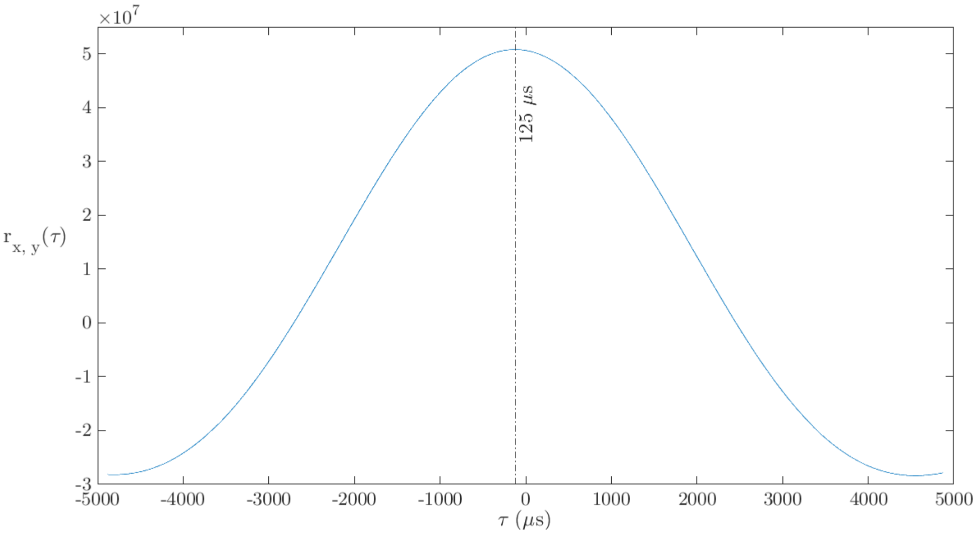


FIGURE 3.9: A cross-correlation plot of a single 1.6 s data length, with the maximum highlighted by the dotted line and the corresponding τ overlaid.

Figure 3.9 shows one of the five cross-correlation functions, $r_{xy}(\tau)$, generated to measure the time delay between each accelerometer and the LDV. Across the five measurements, an average of -125.0 ± 2.4 µs was obtained. Here, the associated uncertainty is half that of the time step, dt. This means that the accelerometer signal leads that of the LDV, which is as expected since obtaining a measurement involves more complex digital signal processing steps, such as demodulation and digital filtering, which can introduce additional latency compared to the simpler and more direct analog-to-digital conversion process used in accelerometers. Moving forward, this is the value used throughout this chapter for both correction techniques. It should also be noted that, due to the signal truncation following time domain synchronisation, the time and frequency domain approaches are not of *exactly* identical signal content. However, the difference is only 26 out of over 300,000 samples, and it is therefore highly unlikely this will significantly affect the results.

3.2.4 Time domain correction algorithm

Figure 3.10 schematically depicts the time domain-based correction technique. The first, integration, and second, detrending, stages are consistent with those of the previous section. The third synchronisation stage accounts for the finite time delay between the accelerometer and LDV signals. This is achieved by time-shifting each accelerometer signal relative to the LDV signal by the amount previously determined using the cross-correlation, given by Equation 3.11. As this is only possible in integer units of the time step, a high sample frequency is required to increase the accuracy of the synchronisation. However, this time-shifting results in regions at the start and at the end of the measurement duration where samples for all three signals are not present. As a result, these regions are truncated.

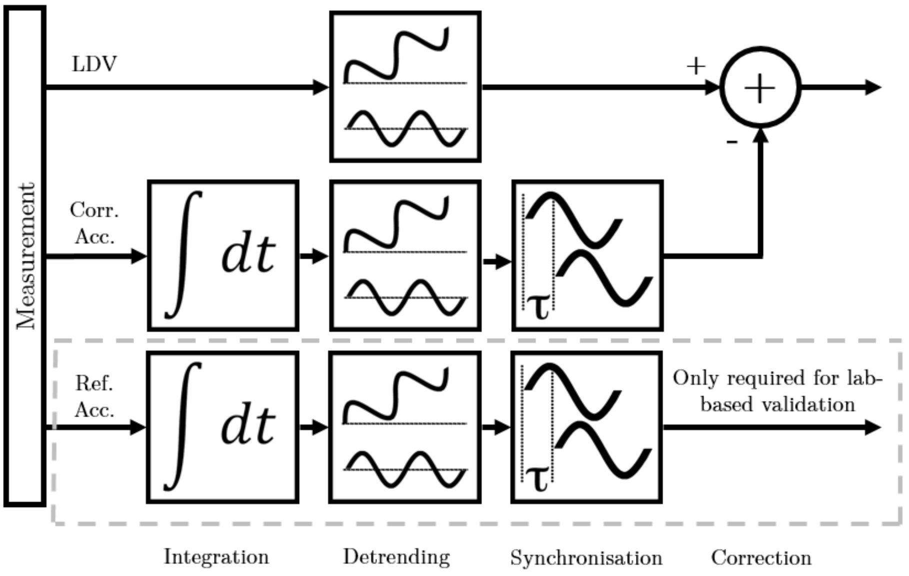


FIGURE 3.10: The time domain-based measurement correction technique with “Corr. Acc.” and “Ref. Acc.” representing the correction and reference accelerometer signals, respectively.

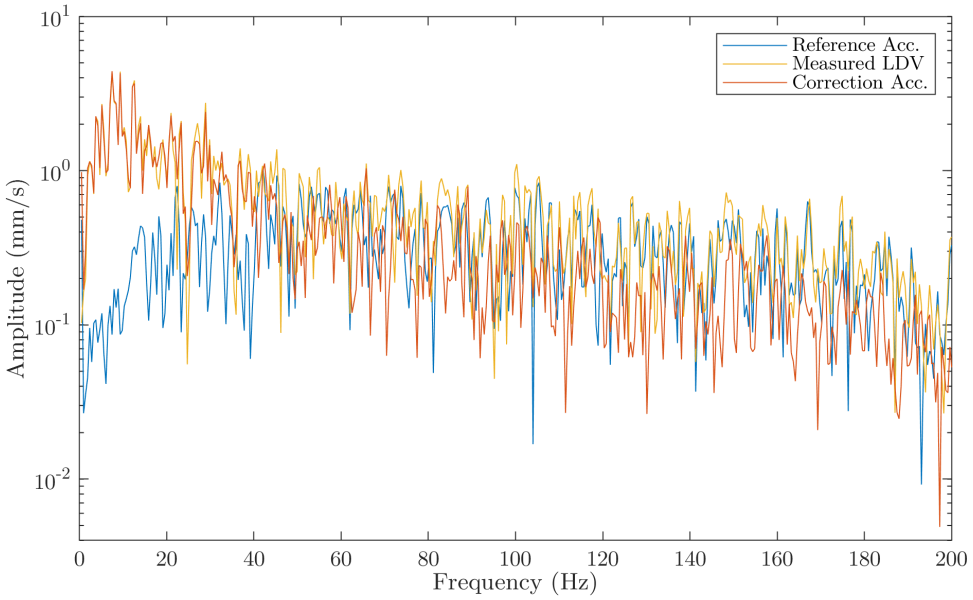
The final stage in Figure 3.10, correction, refers to the removal of the instrument vibration from the signal and is given mathematically by [38]:

$$U_{\mathrm{corr}}(t) = U_{\mathrm{m}}(t) - U_0(t) \tag{3.12}$$

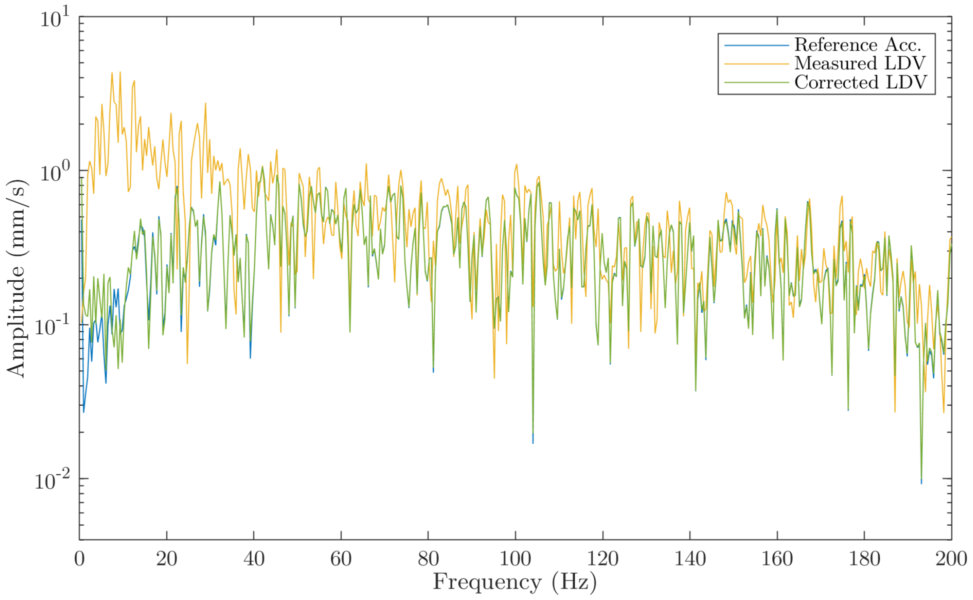
where $U_{\mathrm{m}}(t)$ is the measured LDV signal, $U_0(t)$ is the processed correction accelerometer signal and $U_{\mathrm{corr}}(t)$ is the fully corrected LDV signal.

3.2.5 Established frequency domain and time domain techniques performance comparison

As can be seen qualitatively by comparing Figure 3.11 (a) and Figure 3.11 (b), the frequency domain-based technique yields significant improvements in the corrected versus the uncorrected LDV signal over the range 15 Hz to 100 Hz. However, the performance below 15 Hz is relatively poor, which at the time of publication [51] was thought to be caused by the lower signal level in this range owed to the shaker-amplifier dynamic characteristics. While this may have been partially true, the main cause of this behaviour is incomplete accelerometer signal handling which is the focus of the following chapter.



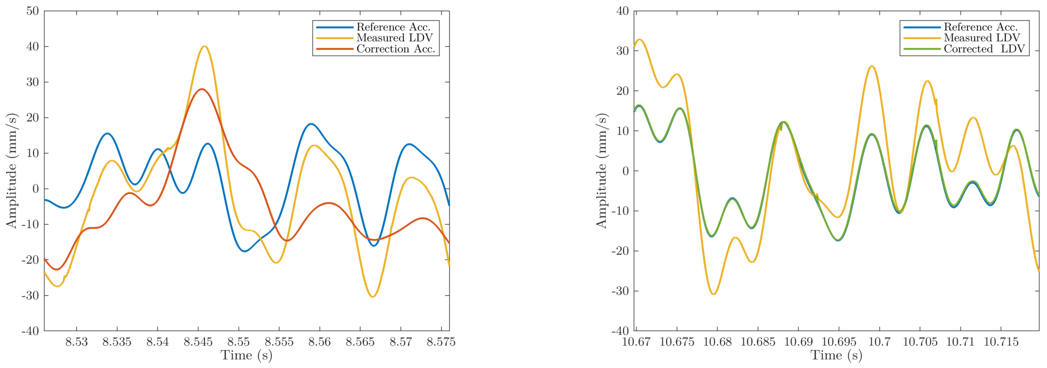
(a)



(b)

FIGURE 3.11: A spectra for a 1.6 s segment (df = 0.625 Hz) in the range 0.625 Hz - 200 Hz: (a) all signals before correction and (b) reference accelerometer, measured and corrected LDV signal.

As can be seen in Figure 3.12 for a 100 ms segment of data, the time domain-based technique developed here offers significant improvement in the corrected LDV signal. However, in the time domain, the effect of speckle noise is apparent, manifested as instantaneous spikes not present in the reference accelerometer signal. In this scenario, the effects of speckle noise may be reduced using a low-pass filter.



(a) (b)

Figure 3.12: A 100 ms segment of data from time domain technique: (a) all signals before correction and (b) reference accelerometer, measured and corrected LDV signal.

The quality of the correction for the two alternative methods can be compared quantitatively using the mean error reduction, which can be seen in Table 3.1. Here it is shown that the time domain-based technique outperforms the established frequency domain-based technique by a factor of eight. Therefore, placing the new time domain-based technique as the likely preferred measurement correction technique for both transient and stationary signal types at the time of publication. It should also be noted that by looking at Figure 3.12 there is little improvement to be made, therefore, future improvements to the technique will likely be minor.

Table 3.1: The mean error reduction for the five 1.6 s segments along with their logarithmic uncertainties calculated as the standard error of the mean.

Technique	$\overline{R}$
Established frequency domain [31]	$25.3^{+1.8}_{-1.3}$ dB
Time domain [51]	$34.5^{+2.1}_{-1.4}$ dB

3.3 Chapter summary and discussion

Recent advances in the use of LDVs for measurement campaigns in which the instrument itself is subject to vibration have led to an increasing number of measurement correction techniques. Practical implementation of these techniques involves the determination of the sensor head vibration and subtraction of this in post-processing, which in the presence of transient leave only the time domain techniques. These existing techniques have been evaluated on three main criteria: accessibility, accuracy and transient signal compatibility. The result of this assessment showed that existing time domain techniques, while being compatible with transient signal types, require complex bespoke optical arrangements and yield questionable performance. However, the existing frequency domain technique hardware arrangement is highly accessible using only a rear-mounted accelerometer and a commercial LDV system. As such, this chapter presented a new signal processing technique based on the existing single-accelerometer hardware arrangement but using only time domain techniques. This results in a technique that is accurate, accessible and compatible with transient signals.

This new time domain-based technique was validated using an experimental arrangement consisting of a vibrating LDV instrumented with a correction accelerometer and a vibrating target similarly instrumented with a reference accelerometer to provide a 'true' vibration measurement. Throughput time data for stationary vibration signals were acquired and processed using both the frequency and new time domain-based techniques. It was shown that, while both approaches lead to improvements in the quality of the corrected LDV measurement, the time domain-based approach presented in this chapter yields a mean error reduction value eight times higher than the frequency domain counterpart, yielding near-perfect measurement correction. This new technique offers a viable alternative to the established frequency domain equivalent for *stationary* vibration signals, provided time data signals can be acquired with a high oversampling factor. Moreover, it now extends the capability to perform complete correction of LDV measurements in the presence of *transient* instrument vibration, which is expected to be present during real-world mobile deployments.

Chapter 4

Frequency domain measurement correction and signal synchronisation

As highlighted in the previous chapter, there is a large performance discrepancy between the established frequency domain and the new time domain-based single-accelerometer techniques. While the new time domain technique presented herein is compatible with stationary and transient signal types, it requires a high oversampling factor to enable accurate signal synchronisation. Similarly, practising vibration engineers working with stationary signal types may prefer a frequency domain-based approach. Therefore, the first portion of this two-part chapter presents a revised frequency domain-based technique. The improved technique's performance is compared against the established frequency domain-based technique along with the time domain-based technique described in the previous chapter for stationary signal types. This improved technique also involves the introduction of an improved relative calibration procedure which is similarly compared to the established frequency domain-based technique. The same experimental arrangement used in the previous chapter is then used to rigorously compare the time, frequency and improved frequency domain-based techniques, showing that the revisions made successfully close the performance discrepancy observed in the previous chapter. The combination of the signal processing presented in this chapter and the previous chapter set a new gold standard for single accelerometer measurement correction signal processing, improving expected performance by up to a factor of eight.

The second portion of this chapter focuses on the effects of signal synchronisation on measurement correction performance. This is important for both domains as the quality of the measurement correction is related to the accuracy of the signal synchronisation since any error therein will adversely affect the quality of the corrected velocity estimate. Therefore, this chapter also describes the development of an analytical model that relates synchronisation error to the velocity estimate error. This is essential for two main reasons. Firstly, the time delay estimate will *always* have an associated uncertainty, however small and in whichever domain. Secondly, the implementation of the synchronisation is constrained to integer multiples of the time step in the time domain; therefore, even an exact delay estimate is unlikely to lead to perfect synchronisation. While interpolation could be used to upsample time domain data to enable sub-time step alignment, this is not always desirable. Therefore, a thorough investigation into the relationship between the time step and the synchronisation error is required such that the errors in the time delay estimate can be propagated and, in the time domain, an optimal sampling frequency can be selected, maximising the performance. This chapter is orientated around the body of work that was published in the Journal of Sound and Vibration [52].

4.1 Review of frequency domain-based measurement correction techniques

The frequency domain-based measurement correction works can be broken down into two categories: a single accelerometer or a dual accelerometer arrangement to estimate the instrument vibration. In either case, these techniques have been shown to effectively remove the instrument vibration from the measurement while being highly accessible to the practising vibration engineer. However, as discussed in the previous chapter, the accuracy of this technique could be improved by implementing proper accelerometer signal handling.

4.1.1 Dual accelerometer measurement correction

Frequency domain-based LDV measurement correction was first realised using a pair of accelerometers in 2017 [38]. They were mounted with equal but opposite coordinates in the two axes orthogonal to the laser beam in order to measure and therefore subtract the instrument vibration from the measured signal, as depicted in Figure 4.1. This arrangement was shown by means of vector calculus to deliver theoretically perfect correction in all six DOF.

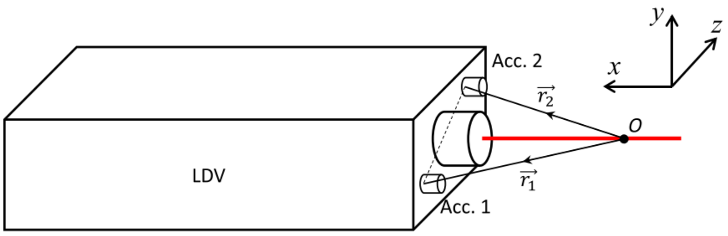


FIGURE 4.1: A schematic of the dual accelerometer LDV measurement correction technique [38]. © 2017 Elsevier, reprinted with permission.

The corrected LDV measurement can be obtained by subtracting the measured instrument vibration from the LDV measurement, as shown by Equation 4.1.

$$U_{\text{corr}} = U_{\text{m}} - U_0 \tag{4.1}$$

where U_{m} is the velocity measured by the instrument, U_0 is the instrument velocity estimate and U_{corr} is the final corrected measurement. Practically, this was implemented using Equation 4.2 and Equation 4.3:

$$\begin{aligned}\Re[FT(U_0)]_{\omega_n} = \frac{1}{2\omega_n}\Bigg(&\cos(\omega_n\tau_1)\Im[FT(a_1)]_{\omega_n} + \cos(\omega_n\tau_2)\Im[FT(a_2)]_{\omega_n}\Bigg)\\ -\frac{1}{2\omega_n}\Bigg(&\sin(\omega_n\tau_1)\Re[FT(a_1)]_{\omega_n} + \sin(\omega_n\tau_2)\Re[FT(a_2)]_{\omega_n}\Bigg)\end{aligned} \tag{4.2}$$

$$\begin{aligned}\Im[FT(U_0)]_{\omega_n} = -\frac{1}{2\omega_n}\Bigg(&\cos(\omega_n\tau_1)\Re[FT(a_1)]_{\omega_n} + \cos(\omega_n\tau_2)\Re[FT(a_2)]_{\omega_n}\Bigg)\\ -\frac{1}{2\omega_n}\Bigg(&\sin(\omega_n\tau_1)\Im[FT(a_1)]_{\omega_n} + \sin(\omega_n\tau_2)\Im[FT(a_2)]_{\omega_n}\Bigg)\end{aligned} \tag{4.3}$$

where $\Re$ and $\Im$ denote the real and imaginary parts, ω_n denotes the nth spectral line of the Fourier transform FT, a_1 and a_2 denote the measured time domain accelerations of the two accelerometers, τ_1 and τ_2 denote the time delays of the two accelerometers relative to the LDV. These time delays were obtained using via a relative calibration procedure using a similar hardware arrangement to that described in Section 3.2.3. This essentially involved the simultaneous acquisition of vibration data using both the LDV and the accelerometers with their sensitive axes aligned. Following this, the phase difference between the integrated accelerometer and LDV signals can be used to calculate the time delay between the two transducer types, as shown in Figure 4.2[1].

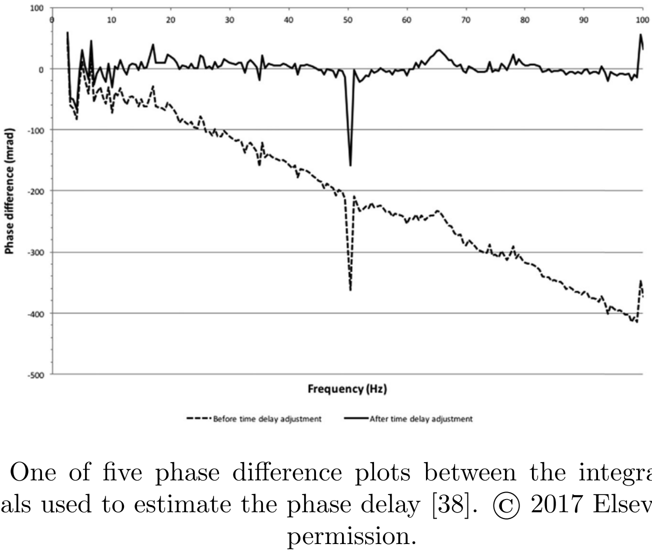


FIGURE 4.2: One of five phase difference plots between the integrated accelerometer and LDV signals used to estimate the phase delay [38]. © 2017 Elsevier, reprinted with permission.

---

[1] This time delay estimation technique is only valid for frequency-independent time delays as is seen here

To verify this correction technique, two different experimental setups were used. First, the setup shown in Figure 4.3 was constructed to simulate instrument vibration in the direction of the beam axis using a large shaker to which the LDV was mounted. The target vibration was simulated using a smaller shaker, and the reference measurement was taken using a stationary LDV mounted above. This experimental arrangement and signal processing yielded a mean error reduction, using Equation 3.5 and Equation 3.10, of 30 dB, representing a significant reduction in the contribution of instrument vibration in the measured signal.

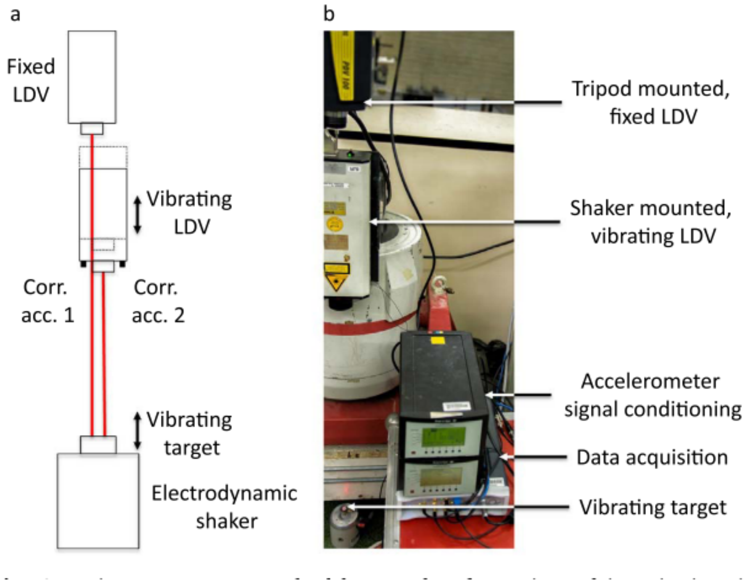


FIGURE 4.3: A diagrammatic representation (a) and a photograph (b) of the experimental setup to verify the dual accelerometer correction technique during one DOF sensor head vibration [38]. © 2017 Elsevier, reprinted with permission.

To further validate this correction technique, a "6-poster" platform that could undergo six DOF vibration was used, with the reference measurement now being taken using an accelerometer rigidly mounted to the target. The setup is pictured in Figure 4.4 with the target now being a mock steering wheel also mounted on the platform. Using this setup, corrections with an error reduction of up to 19 dB were demonstrated. It should be noted that this decrease in the correction performance is likely caused by there being lower levels of instrument vibration on this platform than on the single DOF shaker experiment; hence, there is less room for improvement.

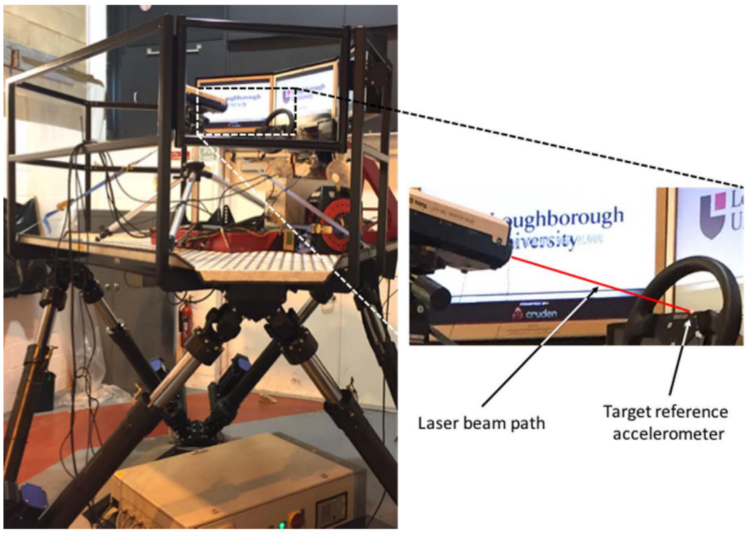


Figure 4.4: A photograph of the experimental setup used to verify the dual accelerometer correction technique during six DOF sensor head vibration [38]. © 2017 Elsevier, reprinted with permission.

4.1.2 Single accelerometer measurement correction

In the year following the initial paper [38], a follow-up paper was published that showed measurement correction to be possible in the presence of six DOF vibration using only one of the two correction accelerometers [31]. Through some additional vector calculus, it was rigorously shown the accelerometer must be mounted colinear to and on the beam axis to avoid any cross-axis sensitivity. Practically, this was achieved by mounting the LDV on a rigid aluminium frame with an accelerometer mounting surface behind the LDV, satisfying these criteria. This frame could then be mounted to the shaker to undergo one DOF vibration while aimed at the vibrating target, mounted from the ceiling for effective isolation from the base shaker. The target was in the form of a speaker box with an accelerometer mounted to the centre to act as the reference measurement. As can be seen in Figure 4.5, this setup also contained similar mounting points for the dual accelerometer configuration discussed in the previous section to allow for a thorough comparison between the two techniques.

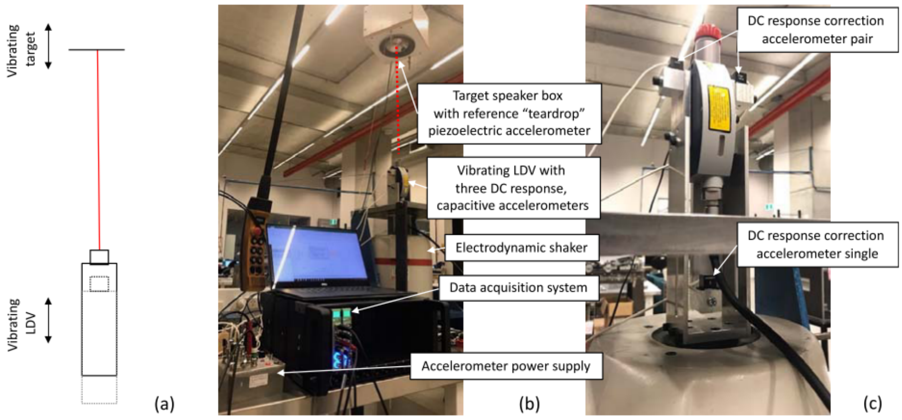


FIGURE 4.5: Experimental arrangement to compare the double and single accelerometer frequency domain-based approaches [31]. CC BY 3.0.

While not presented in the paper, Figure 4.6 was generated to explain the signal processing used for this technique in a style consistent with other such diagrams presented herein to allow for easier comparison. It captures the same concepts as the dual accelerometer technique, except since there is only one accelerometer now, no averaging is required. The procedure involves first windowing and transitioning into the frequency domain using an FFT of all signals. Followed by the integration and synchronisation of the accelerometer signals and, finally, subtracting the processed correction accelerometer signal from the LDV signal.

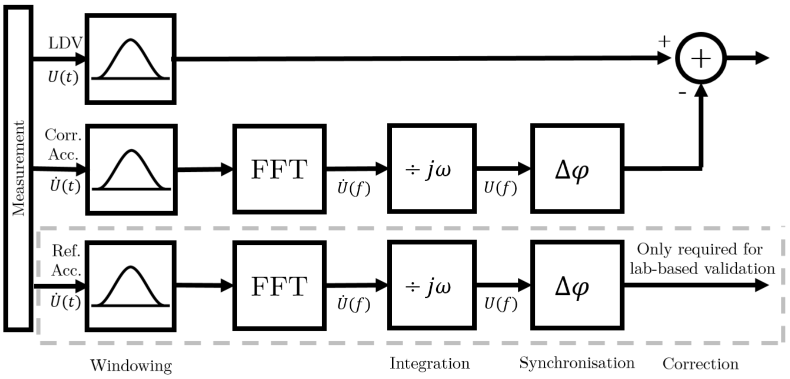


FIGURE 4.6: A functional diagram representing the established frequency domain-based technique developed based on the technique described in the paper (*this diagram was not present in the paper*) [31]. The signals "Corr. Acc." and "Ref. Acc." represent the correction and reference accelerometer signals, respectively.

When both the target and the base shaker were excited using white noise, both the single accelerometer and the dual accelerometer techniques yielded an error reduction of 17.5 dB, showing the single accelerometer technique is as effective as the dual accelerometer technique whilst being considerably easier to implement. Figure 4.7 shows a spectrum before correction and after correction by the single and double accelerometer techniques. Overall, the corrected signals tend to agree with the reference measurement. However, below about 5 Hz, the corrected LDV and the reference measurements deviate from one another.

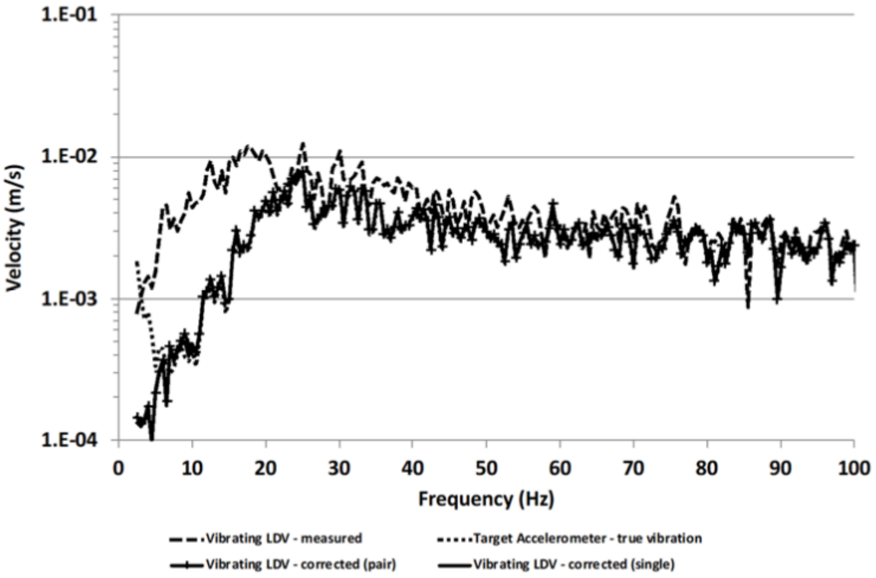


FIGURE 4.7: A spectra of the target vibration compare to the uncorrected LDV and the two corrected LDV techniques [31]. CC BY 3.0.

A similar trend can be observed in the plot of the mean phase difference between the corrected LDV and reference measurement, seen in Figure 4.8. Here, the mean phase difference begins to dramatically increase for frequencies below about 15 Hz. Therefore, it is likely that whatever is causing the performance gap described in Chapter 3 is a phenomenon that largely affects the lower frequencies with a rapid decay as the frequency increases, beginning to hint at the improvements to be made to the existing signal processing.

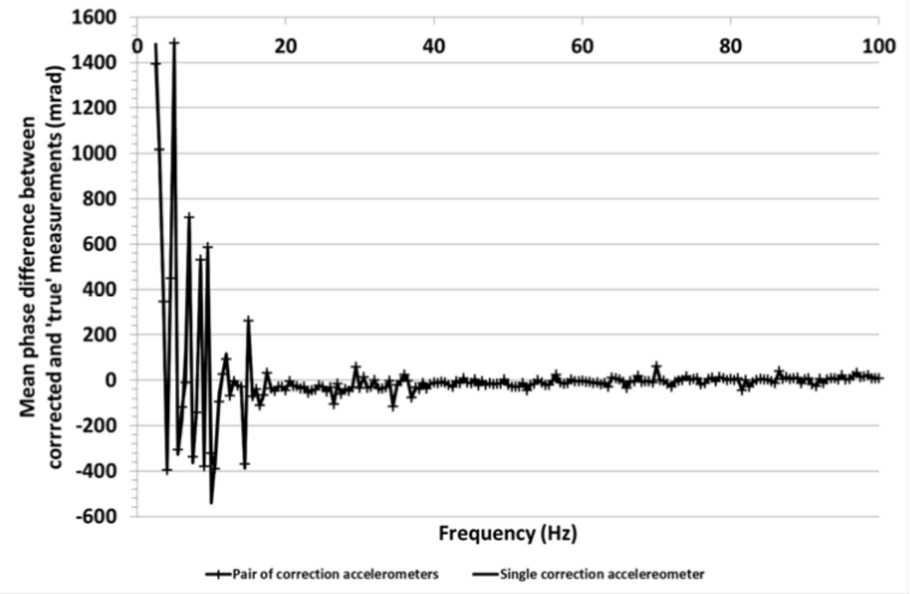


FIGURE 4.8: The mean phase difference between the corrected LDV signal with either a single or double accelerometer, and the reference accelerometer [31]. CC BY 3.0.

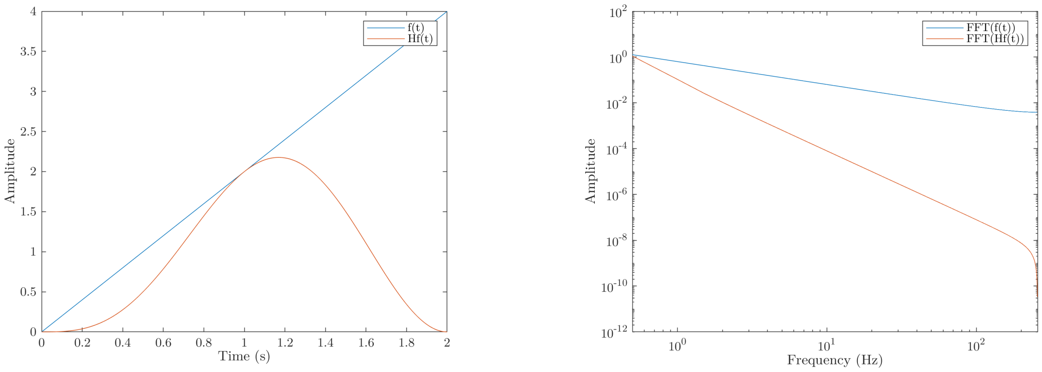
4.2 Improved frequency domain-based processing technique

Similar to the previous chapter, the signals in this work have been handled as velocities since it is the optimal vibration parameter for its flat representation of all frequencies [83]. Therefore, the accelerometer signals must be integrated. However, the integration of discretised signals commonly leads to the introduction of errors which can manifest themselves as drift. Drift is more readily identified and relatively easily removed in the time domain by subtracting a first-order least squares fit. Although possible to remove in the frequency domain, it is less noticeable not as easily removed, which is likely why the established technique neglected to include this. Conversely, integration is readily implemented and shown to be more accurate in the frequency domain [84].

For an arbitrary signal, $s(t)$, the actual measurement in the presence of drift can be understood as the addition of both $s(t)$ and the drift, given by $f(t)$. The effects of this drift on the frequency domain representation of a signal can be visualised by taking the FFT of $f(t)$, where it is best represented as first-order function:

$$f(t) = mt \tag{4.4}$$

where t is time and m is the gradient of the drift. Since the most common window type when dealing with vibrations is a Hann window, H, then its effects should also be considered. However, for completeness, both the FFT of the Hann windowed function, $Hf(t)$, and the original function, $f(t)$, have been plotted, see Figure 4.9. This example demonstrates how the errors being introduced by the drift would manifest themselves at the lower frequencies, agreeing with the observed errors in previous work [31]. Therefore, the integration of an appropriate detrending stage into the established frequency domain-based technique should increase the technique's performance.



(a) (b)

FIGURE 4.9: A 2 s segment of simulated data at 512 Hz for $f(t) = 2t$ in both (a) the time domain and (b) the frequency domain ($df = 0.5$ Hz), with and without a Hann window

4.2.1 Accelerometer sensitivity and time delay estimation

Using the setup described in Section 3.2.3, a dataset is collected with all the sensitive axes aligned and measuring the same vibration. The acquired signals are then processed according to the procedure shown in Figure 4.10 for a single accelerometer channel. In earlier work [31, 38], signals were directly captured as frequency spectra, calculated from Hann-windowed time blocks because the excitation was broadband white noise. Following frequency domain integration of the accelerometer signals, time delay and updated sensitivity values were determined as per the ultimate step in the diagram. In the revised approach, time data are instead acquired with the accelerometer signal immediately converted to frequency domain representation, albeit without prior application of a Hann window to the time data blocks. While perhaps considered unconventional, the lack of windowing is an essential part of the technique as it enables the preservation of the overall time domain waveform, thereby allowing accurate detrending in the subsequent stages.

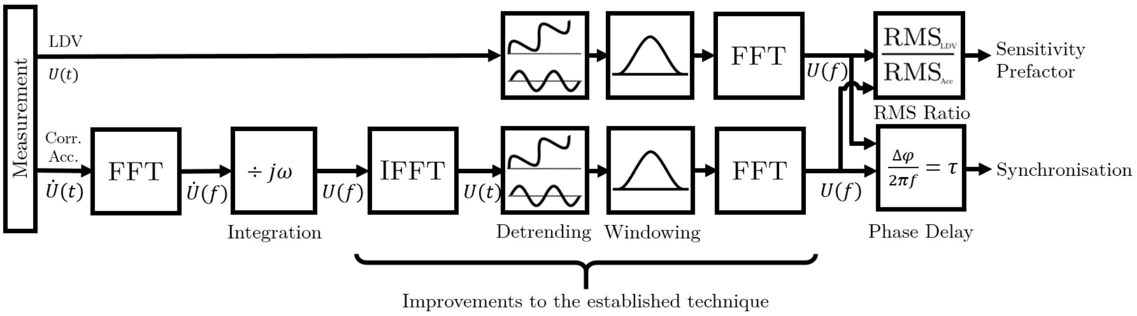


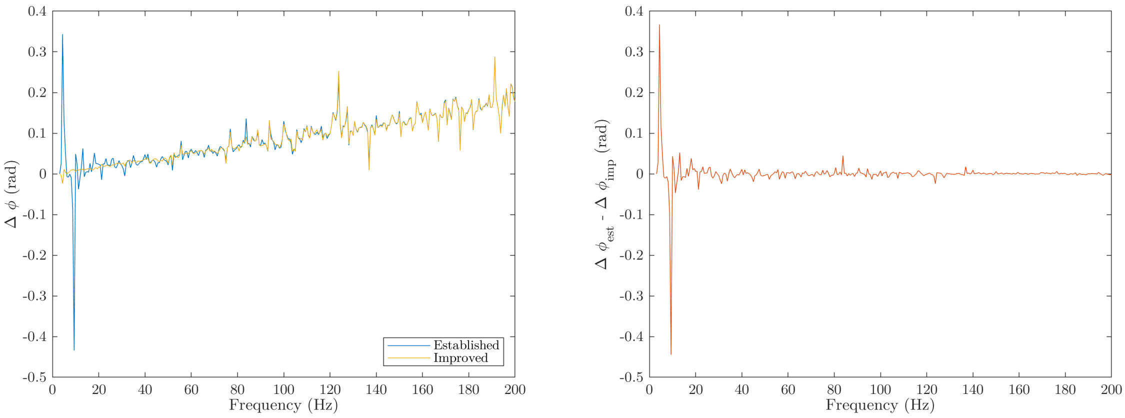
FIGURE 4.10: A schematic of the new frequency domain-based relative calibration procedure. The improvements are highlighted by the curly brackets and include the addition of the IFFT, detrending and FFT stages, along with moving the windowing stage from after the measurement block to just before the second FFT. The signal "Acc." represents that obtained from either the correction or reference accelerometer.

As can be seen in Figure 4.10, a $j\omega$ division is used in the frequency domain to integrate the accelerometer signal. Removal of the resulting drift is achieved by the subtraction of a first-order least-squares fit from the time domain integrated signal. Since this detrending step will act to remove not only the spurious but also some genuine signal content, the same operation must be applied to the measured LDV signal; as done in the previous chapter. Both signals are now converted to the frequency domain in the usual way, and implementing a Hann window on the time data blocks if required. The required sensitivity prefactor and synchronisation parameters are obtained by taking the ratio of the RMS values and the phase difference between the signals, respectively.

Figure 4.11 (a) shows phase difference plots generated from a single time data block using the established and the improved frequency domain-based method. By comparing the two curves, it becomes obvious that the improved technique leads to increased agreement between the two types of transducers. As can be seen in Figure 4.11 (b), the improvement occurs mainly occurs at the lower frequencies. For a system with a constant phase delay, the group delay can be written as:

$$\Delta\phi(f) = -2\pi f \tau_{\text{meas}} \tag{4.5}$$

where $\Delta\phi$ is the phase difference, τ_{meas} is the measured time delay and f is the frequency. Therefore, the least squares fit can be used to extract τ_{meas} from the detrended data set. For this dataset, a value of $\tau_{\text{meas}} = -133.3 \pm 1.8$ µs was obtained using the improved technique. For consistency, all correction techniques in this chapter will use this time delay estimate. Since the required accuracy for the time delay estimate is related to the frequency content of the instrument vibration, more advanced estimate techniques may be required if the instrument vibration increases in frequency. Therefore, the use of more advanced signal processing techniques, such as the smoothed coherence transform [85], may be important for these applications as they yield more accurate estimates which are less affected by measurement noise.



(a) (b)

FIGURE 4.11: Phase differences for a single, 1.6 s data length ($df = 0.625$ Hz), using the established [31], $\Delta\phi_{\text{est}}$, and improved, $\Delta\phi_{\text{imp}}$, frequency domain-based methods; a) phase differences and b) comparison between differences.

4.2.2 Improved frequency domain correction algorithm

Correction of the LDV measurement similarly requires integration of the correction accelerometer signals with detrending; therefore, being essential for optimal performance, as such, a revised post-processing approach is shown in Figure 4.12. As with the previous section, the differences between this improved and the previously established approach are largely captured in the steps between the IFFT and the second FFT. Again, the signals are now captured in the time domain, whereas previously frequency spectra were captured directly. In this case, the integration-related steps are conducted on both the correction and reference accelerometers, these having had their relative sensitivities adjusted and time delays estimated. The latter of the two accelerometers is only intended for use in the laboratory research campaign, which provides a 'true' vibration measurement for correction performance; for subsequent real-world campaigns, there is no such device since, otherwise, there would be no need to develop the LDV capability for this purpose.

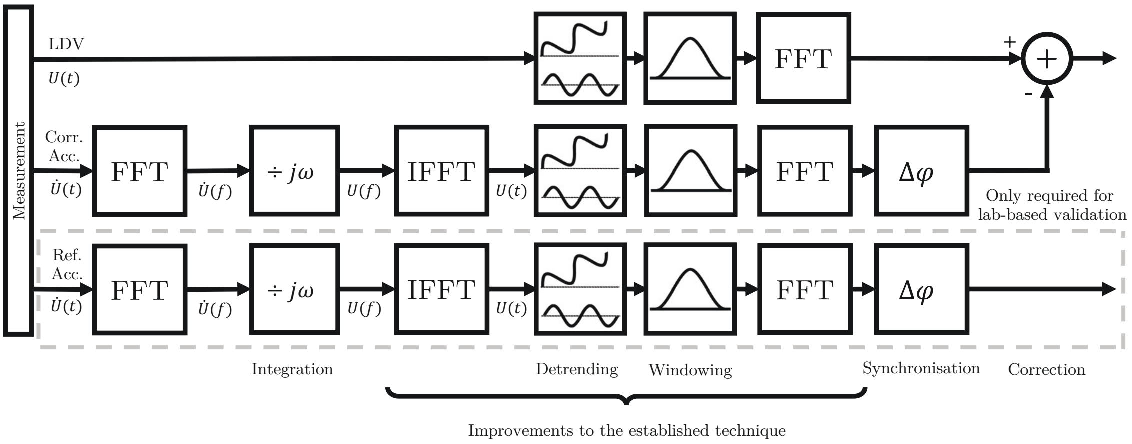


Figure 4.12: A functional diagram representing the improved frequency domain-based technique. The improvements are highlighted by the curly brackets and include the addition of the IFFT, detrending and FFT stages, along with moving the Windowing stage from after the measurement block to just before the second FFT. The signals "Corr. Acc." and "Ref. Acc." represent the correction and reference accelerometer signals, respectively.

It is important also to note here that the LDV measurement must also be subject to the detrending step; otherwise, the corrected LDV signal may contain some signal content that was removed during the accelerometer detrending. Following the second FFT, the correction processing is similar to that previously described Figure 4.6, the exception being that here the accelerometer signals are already in velocity. Incorporating the previously determined time delays, before subtracting the correction accelerometer signal from the LDV (in complex representation), yields the *corrected* LDV signal for direct comparison with the 'true' vibration, given by the reference channel. The correction performance is then quantified using the error reduction, given by Equation 3.5, with the frequency domain formulations of the MSE, given by Equation 3.10. For practical applications, if the user so wishes, an IFFT may be applied to return the signal to the time domain.

4.2.3 Improved frequency domain technique performance assessment

It is useful to visualise the performance difference between the two techniques as a function of the frequency. To do this, the error reduction can be calculated for each spectral line and plotted. The MSE in Equation 3.5 can be substituted for the square error, SE, to preserve the frequency information. This plot is improved if a mean of each spectral line, n, is taken across the multiple spectra, m. Algebraically, this is given by:

$$\begin{aligned} \mathrm{SE}_{\mathrm{signal}}(n) &= \frac{1}{M}\sum_{m=1}^{M}(a_{0,m}^{\mathrm{signal}} - a_{0,m}^{\mathrm{true}})^2 && \text{for } n = 0 \\ &= \frac{1}{2M}\sum_{m=1}^{M}(A_{n,m}^{\mathrm{signal}} - A_{n,m}^{\mathrm{true}})^2 + (B_{n,m}^{\mathrm{signal}} - B_{n,m}^{\mathrm{true}})^2 && \text{for } n > 0 \end{aligned} \tag{4.6}$$

where all symbols are as previously defined. Substituting the SE in place of the MSE in Equation 3.5) would then give:

$$R(n) = -10\log_{10}\left(\frac{\mathrm{SE}_{\mathrm{corr}}}{\mathrm{SE}_{\mathrm{meas}}}\right)\ \mathrm{dB} \tag{4.7}$$

which can be seen plotted in Fig. 4.13 as a function of frequency.

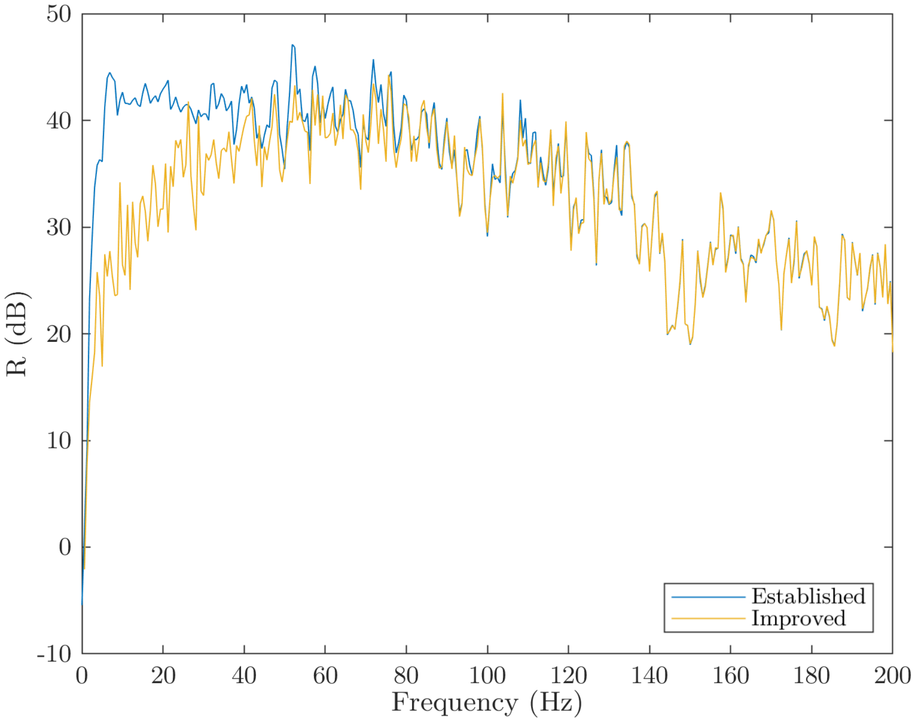


FIGURE 4.13: A plot of $R(f)$ obtained from $R(n)$ for both the established [31] and the improved [52] frequency domain-based techniques using five 1.6 s data lengths ($df = 0.625$ Hz).

To obtain the data shown in Figure 4.13, the reference channel for both techniques were processed identically and according to the technique presented in Figure 4.12. These data show that the improved frequency domain-based technique outperforms the established technique for frequencies below 100 Hz. However, for frequencies above 100 Hz, the difference is less noticeable, which is expected since the errors mitigated by detrending are largely at the lower frequencies. For further proof that this was indeed the cause of the performance gap seen in Chapter 3, the mean error reduction, $\overline{R}$, of these two techniques can be compared to that of the time domain technique.

Table 4.1: Error reduction of the three measurement correction techniques with their associated logarithmic uncertainties.

Technique	$\overline{R}$
Established frequency domain [31]	$25.0^{+1.8}_{-1.3}$ dB
Improved frequency domain [52]	$33.5^{+1.2}_{-0.9}$ dB
Time domain [51]	$34.3^{+0.9}_{-0.7}$ dB

Table 4.1 shows the error reduction for the three techniques, along with their associated logarithmic uncertainties. Firstly, it shows that the improvements made to the frequency domain technique lead to a seven times increase in $\overline{R}$. It is important to note that this is not in contradiction with the previously quoted eight times performance gap [51] since these were different datasets with likely slightly different gains set on the shakers. Comparing the performance of the improved frequency domain-based technique to that of the time domain-based technique described in Chapter 3, it can be seen that while the time domain-based technique slightly outperforms that for the frequency domain technique by 0.8 dB, however, the two are still within the associated uncertainties of one another. This shows that the previously observed performance gap was indeed caused by the lack of proper accelerometer handling in the established frequency domain technique.

4.3 Theoretical generalisation relating synchronisation error to correction performance

With performance inconsistency now resolved, the focus of the remainder of this chapter is an analytical model which relates the synchronisation error, $\Delta\tau$, to measurement correction quality. The main focus will be on time domain-based processing since the quality of the synchronisation is restricted to integer multiples of the time step. Although interpolation could be used on the time data could enable sub-time step alignment, this is not always preferred. Therefore, this model can be used to predict the system requirements necessary to obtain high-quality time domain data based on the sampling frequency and the measured time delay.

4.3.1 Relating the error reduction to the synchronisation error

To relate the synchronisation error, $\Delta\tau$, to the error reduction, R, it is assumed that $\Delta\tau$ is the primary factor which affects the quality of the corrected velocity estimate, $\mathrm{MSE}_{\mathrm{corr}}$. While other factors may also affect $\mathrm{MSE}_{\mathrm{corr}}$, this model is not concerned with them. Therefore, a relationship between $\mathrm{MSE}_{\mathrm{corr}}$ and $\Delta\tau$ is required to relate $\Delta\tau$ to R.

To do so, the corrected LDV signal, $v_{\mathrm{corr}}(t)$, can be written as follows:

$$v_{\mathrm{corr}}(t) = v_{\mathrm{meas}}(t) - v_{\mathrm{acc}}(t) \tag{4.8a}$$

where $v_{\mathrm{meas}}(t)$ is the target velocity measured by the LDV and v_{acc} is the velocity of the LDV instrument itself, measured by the correction accelerometer. Rewriting Equation 4.8a) to encapsulate the synchronisation error expressed as $v'_{\mathrm{corr}}(t)$:

$$v'_{\mathrm{corr}}(t) = v_{\mathrm{meas}}(t) - v_{\mathrm{acc}}(t + \Delta\tau) \tag{4.8b}$$

The velocity error, $\Delta v_{\mathrm{corr}}(t)$, can then be defined as the difference between $v_{\mathrm{corr}}(t)$ and $v'_{\mathrm{corr}}(t)$:

$$\Delta v_{\mathrm{corr}}(t) = v_{\mathrm{acc}}(t) - v_{\mathrm{acc}}(t + \Delta\tau) \tag{4.9}$$

Now a discrete Fourier expansion can be applied and, since $\Delta\tau$ is small, the small angle approximation can also be applied:

$$v_{acc}(t) = \frac{a_0}{2} + \sum_{n=1}^{N} A_n \cos(n\omega_0 t) + B_n \sin(n\omega_0 t) \tag{4.10a}$$

$$\begin{aligned} v_{\mathrm{acc}}(t+\Delta\tau) = \frac{a_0}{2} + \sum_{n=1}^{N} A_n \Bigg(&\cos(n\omega_0 t) - n\omega_0 \Delta\tau \sin(n\omega_0 t) \Bigg) \\ +B_n \Bigg(&\sin(n\omega_0 t) + n\omega_0 \Delta\tau \cos(n\omega_0 t) \Bigg) \end{aligned} \tag{4.10b}$$

where A_n and B_n are constants for each spectral line, a_0 is the DC component and ω_0 is the spectral resolution. Substituting these expansions back into Equation 4.9) and simplifying the expression gives:

$$\Delta v_{\mathrm{corr}}(t) = \Delta\tau \sum_{n=1}^{N} -A_n n\omega_0 \sin(n\omega_0 t) + B_n n\omega_0 \cos(n\omega_0 t) \tag{4.11}$$

Therefore, $\mathrm{MSE}_{\mathrm{corr}}$, or $\overline{\Delta v_{\mathrm{corr}}(t)^2}$, can be expressed as:

$$\mathrm{MSE}_{\mathrm{corr}} = \frac{\Delta\tau^2}{2} \sum_{n=1}^{N} (A_n n\omega_0)^2 + (B_n n\omega_0)^2 \tag{4.12}$$

Inevitable sources of error other than synchronisation error mean that the MSE will never be zero in practice. To account for this, an additional term, c, is introduced:

$$\mathrm{MSE}_{\mathrm{corr}} = c + \frac{\Delta\tau^2}{2} \sum_{n=1}^{N} (A_n n\omega_0)^2 + (B_n n\omega_0)^2 \tag{4.13}$$

where c is a constant representing the lowest practically obtainable MSE with a given setup. In order to relate this to the mean error reduction, Equation 3.5) can be rearranged into the following form:

$$\frac{\mathrm{MSE}_{\mathrm{corr}}}{\mathrm{MSE}_{\mathrm{meas}}} = 10^{-\frac{R}{10}} = r \tag{4.14}$$

where r has been defined for convenience. Combining this with Equation 4.13), the following can be written:

$$r(\Delta\tau) = \frac{c + \frac{\Delta\tau^2}{2} \sum_{n=1}^{N} (A_n n\omega_0)^2 + (B_n n\omega_0)^2}{\mathrm{MSE}_{\mathrm{meas}}} \tag{4.15}$$

The value of $r(\Delta\tau = 0)$ can then be described as "optimal", and denoted by r_{opt}. Therefore, Equation 4.15) can also be written as:

$$r(\Delta\tau) = r_{\text{opt}} + K\Delta\tau^2 \tag{4.16}$$

where r_{opt} and K have now incorporated all remaining constants. Both r_{opt} and K can be experimentally determined for a given setup.

Finally, an expression for $R(\Delta\tau)$ can be written by substituting Equation 4.16) into Equation 3.5):

$$R(\Delta\tau) = -10\log_{10}(r_{\text{opt}} + K\Delta\tau^2) \tag{4.17}$$

However, to make use of this relationship, the synchronisation error must be derived and this differs for the frequency and time domain.

4.3.2 Frequency domain synchronisation error

The synchronisation error in the frequency domain, $\Delta\tau_f$, is simple since it only depends on how accurately the signal delay is known since sub-time step alignment is possible. Mathematically this can be defined as:

$$\Delta\tau_f = \tau_{\text{meas}} - \tau_{\text{true}} \tag{4.18}$$

where τ_{meas} is the measured time delay and τ_{true} is the true time delay. Practically, τ_{true} is the theoretical unknowable exact true time delay between the signals and τ_{meas} is ideally determined using the procedure outlined in Section 4.2.1. Substituting this into Equation 4.17) gives the following:

$$R(\tau_{\text{meas}}) = -10\log_{10}\left(r_{\text{opt}} + K(\tau_{\text{meas}} - \tau_{\text{true}})^2\right) \tag{4.19}$$

Therefore, this frequency domain model predicts that the mean error reduction will have no sampling frequency dependence.

4.3.3 Time domain synchronisation error

The synchronisation error in the time domain, $\Delta\tau_t(dt)$, is not only dependant on the accuracy of the measured time delay, but also on the time step. This is given by:

$$\Delta\tau_t(dt) = \left\lfloor \frac{\tau_{\text{meas}}}{dt} \right\rceil dt - \tau_{\text{true}} \tag{4.20a}$$

where dt is the time step and $\lfloor ... \rceil$ denotes the nearest integer. Equation 4.20a) can now be written in terms of the sampling frequency, f_s, instead of the time step since that is the adjustable acquisition parameter:

$$\Delta\tau_t(f_s) = \left\lfloor \tau_{\text{meas}} f_s \right\rceil \frac{1}{f_s} - \tau_{\text{true}} \tag{4.20b}$$

Since τ_{true} is the theoretically true value, it bares little practical significance. Moving forward, it will be assumed that $\tau_{\text{meas}} \approx \tau_{\text{true}}$ so that the effect of the sampling frequency alone on the error reduction can be thoroughly assessed; both will now be denoted as τ. This assumption also results in both positive and negative synchronisation errors having an equivalent negative effect on the velocity estimate. Combining Equation 4.17) and Equation 4.20b) gives:

$$R(f_s) = -10\log_{10}\left(r_{\text{opt}} + K\left(\left\lfloor \tau f_s \right\rceil \frac{1}{f_s} - \tau\right)^2\right) \tag{4.21}$$

The general form of the time domain model, with the significant features labelled, can be seen in Figure 4.14. As can be seen, there are three distinct regions of behaviour. The first, Inactive Region, displays no sensitivity to the sampling frequency because the time step, dt, is too small for any temporal alignment to take place; as such, no performance change occurs. Temporal alignment becomes possible once the time step is less than $\frac{1}{2\tau}$, representing the beginning of the Transitionary Region. This region is characterised by a sharp performance increase, as the decreasing time step allows for increasingly accurate synchronisation. The third and final region, the Oscillatory Region, is characterised by oscillations in the error reduction which decrease in amplitude as the frequency increases. The peaks of these performance oscillations occur at integer multiples of τ^{-1} Hz, since these locations are where τ becomes divisible by an integer number of time steps, leading to an increase in the accuracy of the temporal alignment. Similarly, the performance troughs occur at half-integer multiples of τ^{-1} Hz, where the synchronisation error is maximised.

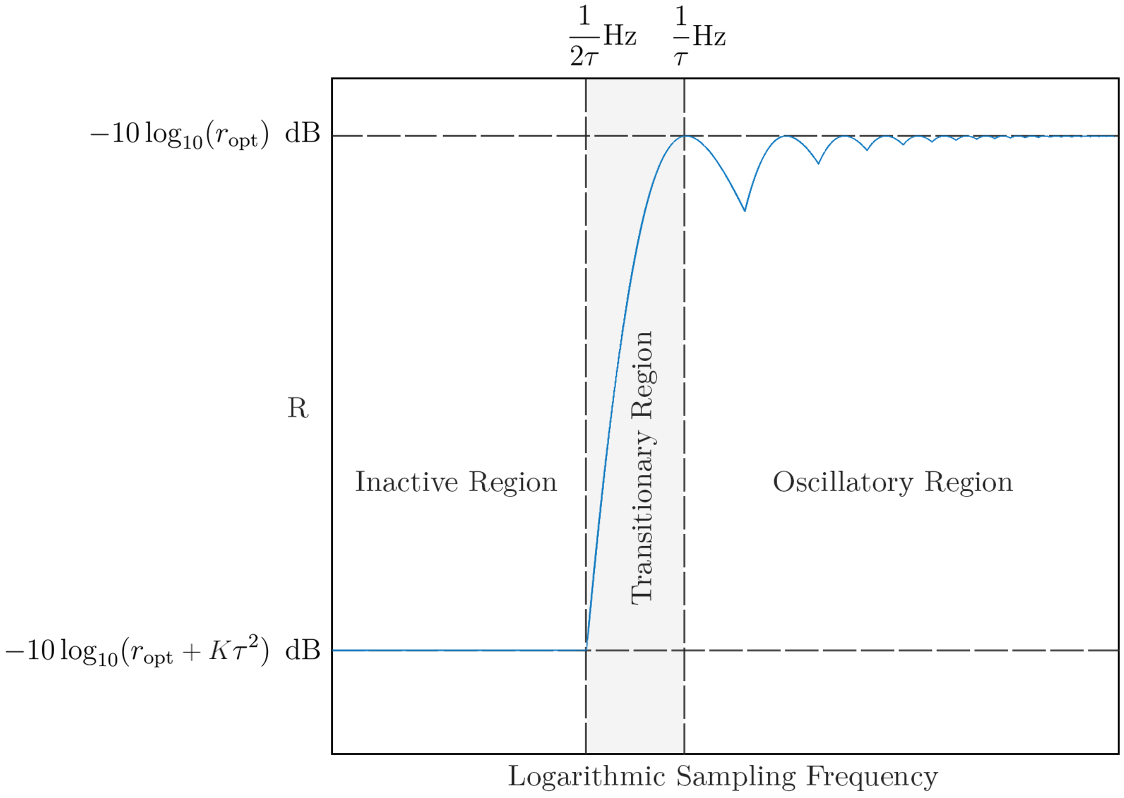


FIGURE 4.14: A general plot of the time domain error reduction model as a function of the sampling frequency with three distinct regions labelled. Higher values on the vertical scale represent better performance.

Since a continuous range of sampling frequencies is rarely available, a more practically relevant example is Equation 4.21) plotted at the sampling frequencies available on the Simcenter SCADAS Mobile data acquisition system, as seen in Figure 4.15. Although the highest sampling frequency, 204.8 kHz, shows a significant reduction in error, a similar performance could have been obtained by selecting a sampling frequency of 8.192 kHz, 16.384 kHz or 40.960 kHz. In fact, the aforementioned frequencies have a error reduction 0.03 dB higher than 204.8 kHz. This shows that the time domain model can still predict the optimal sampling frequency to maximise performance without needing to measure values for the constants K and r_{opt}.

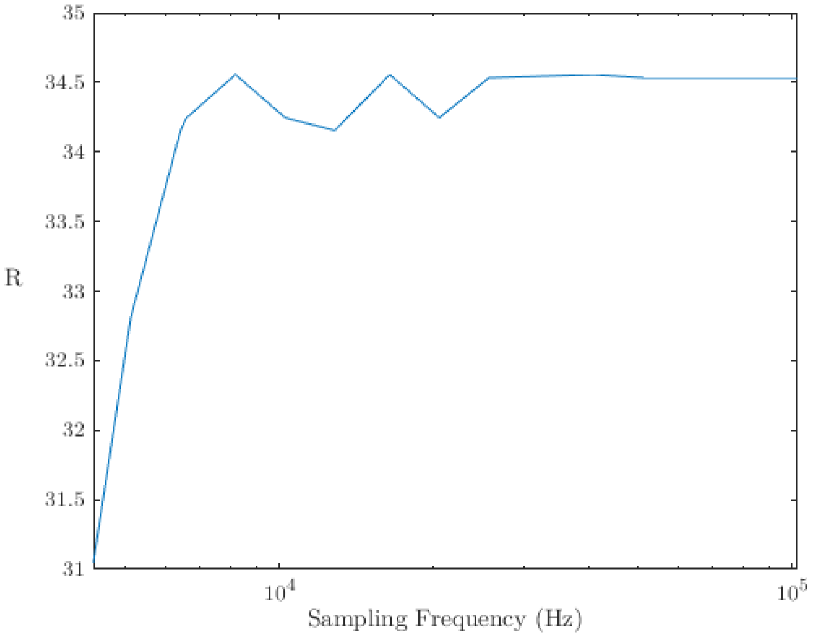


FIGURE 4.15: A plot of error reduction as a function of the sampling frequencies available on the Simcenter SCADAS Mobile data acquisition system. This is plotted with values of $K = 35 \times 10^3\ \text{s}^{-2}$ and $r_{\text{opt}} = 350 \times 10^{-6}$.

4.3.4 Time domain constants determination

Since the main functionality of the time domain model is to enable the informed selection of the sampling frequency, knowledge of the constants is not necessary. However, in order to validate the time domain model it will be compared to an experimentally measured $R(f_s)$, denoted by $R_{\mathrm{ex}}(f_s)$, meaning the constants are required since they affect the model's relative proportions in the vertical axis.

The first constant, r_{opt}, is calculated using:

$$r_{\mathrm{opt}} = 10^{-\frac{\max(R_{\mathrm{ex}})}{10}} \tag{4.22}$$

where $\max(R_{\mathrm{ex}})$ is the maximum value of $R_{ex}(f)$. Similarly, K, can be calculated using:

$$K = \frac{r_{\mathrm{opt}} - 10^{-\frac{R_{\mathrm{ex}}(f_s)}{10}}}{\left(\left\lceil \tau_{\mathrm{meas}} f_s \right\rceil \frac{1}{f_s} - \tau_{\mathrm{meas}}\right)^2} \quad \text{for} \quad f_s > \frac{1}{2\tau_{\mathrm{meas}}} \text{ Hz} \tag{4.23}$$

where all symbols are as previously defined. The sampling frequency here must be larger than $\frac{1}{2\tau_{\mathrm{meas}}}$ Hz as the time domain model does not predict any behaviour in the Inactive Region so scaling using these data will lead to erroneous predictions.

4.4 Experimental validation of the time domain model

This section aims to show the relative performances of both the improved frequency domain and time domain techniques over a range of frequencies and, by doing so, confirm the time domain model for $R(f_s)$, given by Equation 4.21). Thus, providing the user with two instrument vibration correction techniques when faced with stationary signal types. The experimental arrangement and acquisition parameters used in the following was common with that used in Section 3.2.1 and Section 3.2.2.

4.4.1 Data collection and processing

To validate the time domain model presented in Equation 4.23), the error reduction needs to be characterised against the sampling frequency and thus data is acquired at the highest available sampling frequency of 204.8 kHz and iteratively downsampled to simulate acquisition at lower sampling frequencies. The process was implemented in MATLAB and Figure 4.16 illustrates this code for the time domain and improved frequency domain-based techniques.

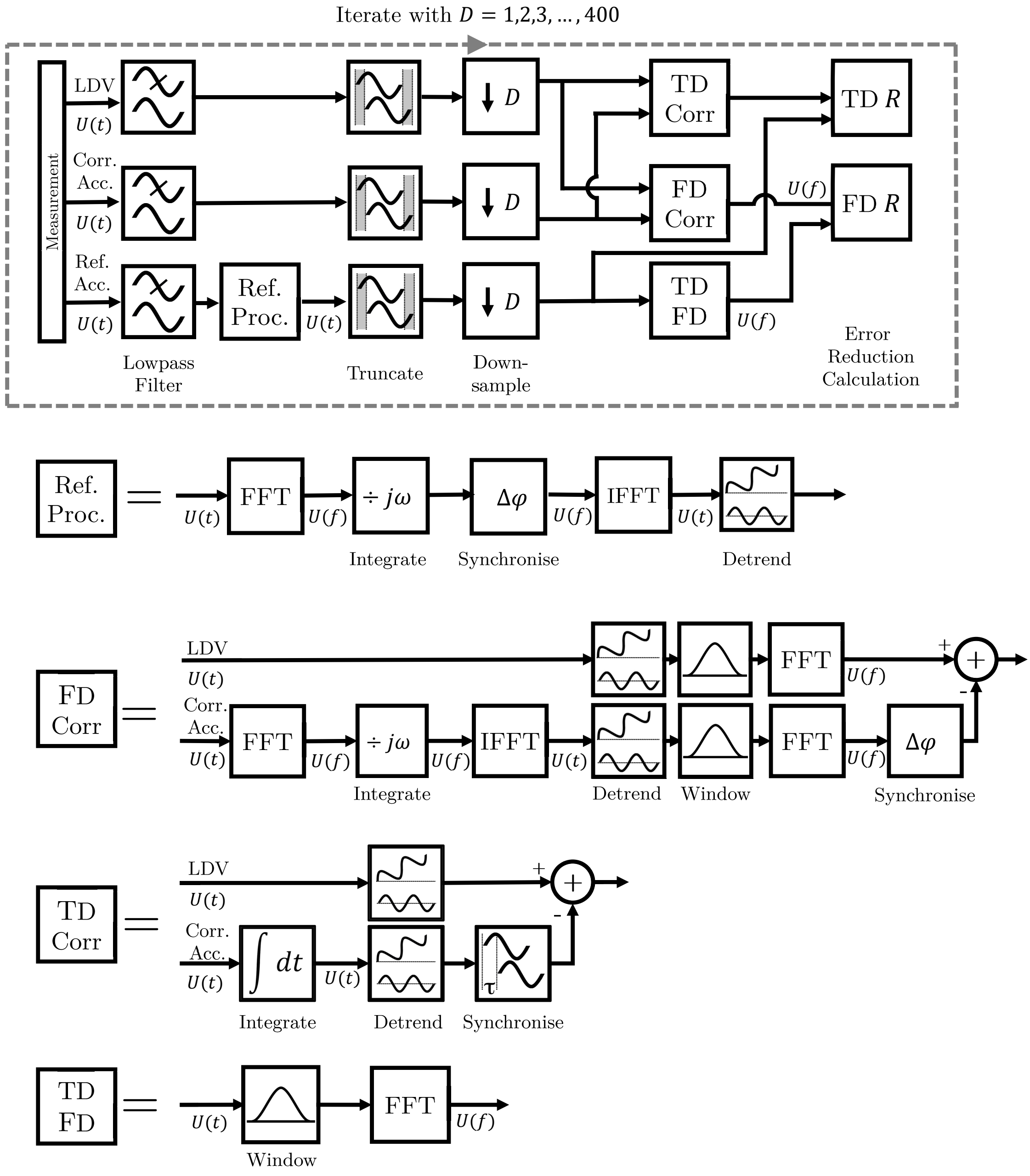


Figure 4.16: A schematic of the code used to characterise the error reduction as a function of sampling frequency. Where D is a downsampling factor and "TD Corr" and "FD Corr" represent both the time domain [51] and improved frequency domain-based correction techniques [52], respectively.

As shown in Figure 4.16, measured data are immediately lowpass filtered using a finite impulse response digital lowpass filter with a 200 Hz cut-off frequency which coincides with the maximum frequency of the vibration. Any spurious higher frequency signal content which might otherwise have been aliased into the frequency range of interest following the downsampling is thereby rejected. Following this, the reference accelerometer signal is then subjected to the same frequency domain detrending and synchronisation steps as previously described. The sole purpose of this special treatment is to make sure that the reference signal is as close to the ‘true’ vibration. This includes both the integration and synchronisation accuracy.

Given that the Fourier transform presumes that a signal is periodic, depending on which signal is leading, the phase shift would have caused a piece of the reference accelerometer signal’s beginning or end to wrap around to the opposite side of the signal. To fix this, any sample in this “wrapped” region is removed from all three signals in the Truncation stage, shown in Figure 4.16. The signals are then downsampled by collecting each Dth sample from the original signals, simulating an acquisition with a lower sampling frequency. The penultimate stages named “TD Corr” and “FD Corr” represent the two correction algorithms. This process was looped in the code with $D = 1, 2, ..., 400$, giving a minimum sampling frequency of 512 Hz and a total of 400 data points for each correction technique. The final output is, therefore, two data sets describing the performance of each correction algorithm as a function of the sampling frequency.

4.4.2 Model validation and sample rate dependent performance assessment

The frequency domain model, given by Equation 4.19), predicts no sampling frequency dependence as sub-time step synchronisation is possible in the frequency domain. However, the time domain model, given by Eq (4.21), predicts a reasonably strong dependence due to this time step synchronisation limitation. Comparing Figure 4.17 and Figure 4.14, the three previously defined regions are clearly identifiable from the experimental data. In particular, the important Transitionary Region is clearly shown; this is where the time domain-based technique’s performance exceeds that of the frequency domain-based approach.

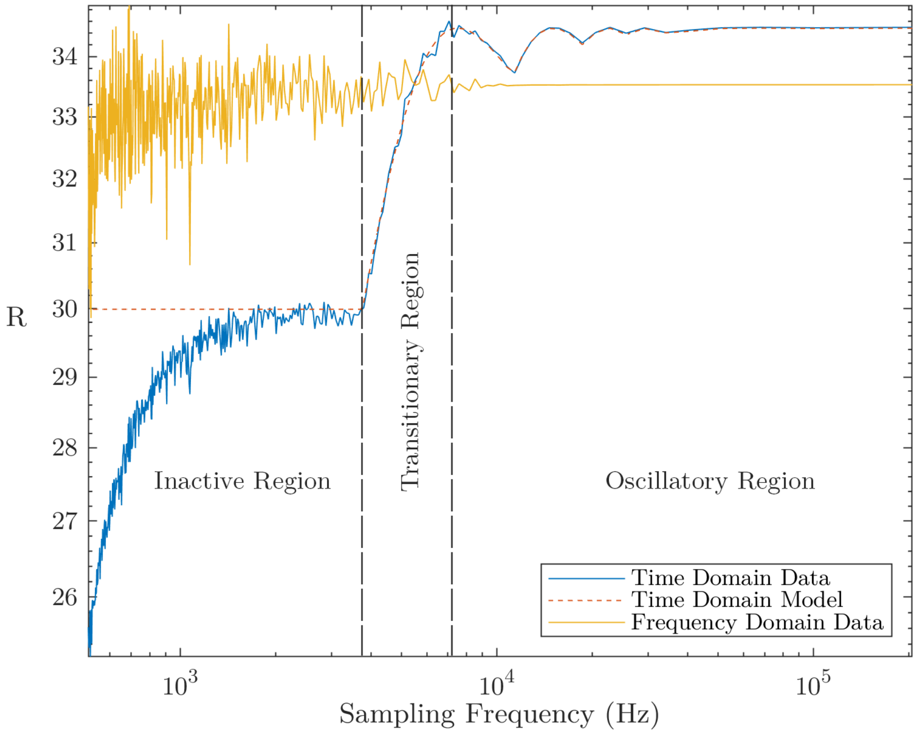


FIGURE 4.17: A plot of the experimentally derived error reduction as a function various sampling frequencies for the improved frequency domain and the time domain-based techniques. The error reduction is calculated using Equation 3.5) and the domain-specific formulations for the MSE. The time domain model, given by Equation 4.21), is also plotted for validation purposes.

In order to validate the time domain model, the time domain data can be compared to the time domain model in Figure 4.17; from this, two shortcomings can be seen. Firstly, the minor noise-like variations that can be seen in the time domain data at frequencies below 10 kHz are not predicted by the time domain model. However, both correction techniques show these variations, which are most likely the result of distinct phenomena since they don't occur at common frequencies. Secondly, the time domain model fails to predict the behaviour in the Inactive Region - where an increase from 25 dB to 30 dB is observed. The three-times increase seen in the time domain experimental data is likely caused by a decrease in the quality of integration at the lower sampling frequencies when using the cumulative trapezoidal method. It is no shortcoming of the time domain model since only the effects of signal synchronisation were considered, a more complex model would likely be required to understand these other phenomena.

However, as can be seen in Figure 4.17, the time domain model almost exactly describes the experimental data in the Transitionary and Oscillatory Regions, which are the regions of interest here. With the time domain model validated, its main utility is its ability to determine when the time domain technique will perform optimally, *a priori* based on the time delay estimate and the sampling frequency. That is, significant improvements in the correction will occur if a data acquisition system is used with a sampling frequency larger than $\frac{1}{2\tau}$ Hz. Similarly, any sampling frequency larger than τ^{-1} Hz will not yield any substantial increase in performance and is, therefore, unnecessary. Also, if possible, a sampling frequency should be selected close to a performance peak located at $\frac{n}{2\tau}$ Hz with $n = 1, 2, 3, ...$

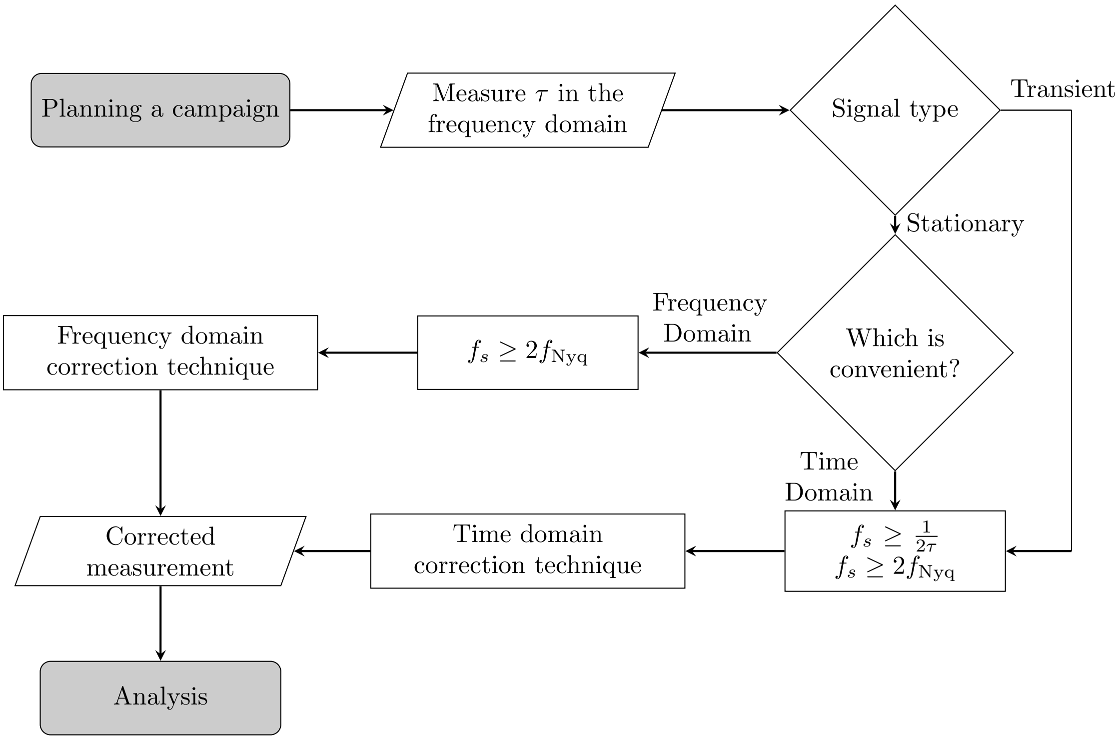


FIGURE 4.18: A flow chart describing when to use either the time domain or the improved frequency domain-based technique, both described herein. The outcome is based on the sampling frequencies available to the user and τ.

Comparing the performances of the two techniques in the region $f_s > \frac{1}{2\tau}$ Hz, as shown in Figure 4.17, there is about a decibel of performance gain to be made by using the time domain-based technique. The exact reason for the difference in performance is yet to be determined. However, since this is small, either technique could be used with little difference in overall performance. Similarly, the user could select a sampling frequency close to a performance peak, τ^{-1} Hz, in order to optimise the time domain-based technique performance. However, the benefit of this marginal with the performance change between a performance peak and trough in the Oscillatory Region to be just over a decibel. Figure 4.18 summarises these generalised findings based on the vibration signal type and the sampling frequency, advising the user which is the most appropriate technique to use for a given measurement campaign. When the vibration is stationary in nature, the user can select either technique. When the vibration is transient in nature, the user must use the time domain-based technique with an an appropriate sampling frequency to optimise performance.

4.5 Chapter summary and discussion

For stationary signal types, both time and frequency domain-based techniques should be able to handle the measurement correction equivalently. However, the previous chapter showed that there exists an eight times performance gap between the established frequency domain-based technique and the time domain-based technique. As such, this chapter aimed to close the observed performance gap, with an improved *frequency domain*-based technique being developed. By applying a modified frequency domain-based technique, which included a detrending step prior to implementation of the correction, a seven times performance increase was obtained. To make this detrending possible, it is necessary that no window is applied to the sampled data until after detrending. Particular improvement was shown to be found in frequencies below 100 Hz, which is arguably a major benefit since many applications of interest in these techniques are expected to be focused on mechanical vibrations experienced during mobile deployments. In all likelihood, the improved technique presented herein should become the new gold-standard frequency domain-based technique, with the established technique becoming obsolete. It should be noted that while previous work showed an eight-times difference in the performance, the seven-times difference in performance is not contradictory as the error reduction is dependant on the relative levels of the target and instrument vibration, therefore, is not necessarily consistent across campaigns.

Since correction measurements are typically obtained using accelerometers, in addition to the requisite integration, signal synchronisation is necessary due to differences in signal conditioning. It is commonly known that the effectiveness of the correction will depend on how well the signals are synchronised. Therefore, when working in the *time domain* and when interpolation is not desirable, the sampling frequency would then contribute to the synchronisation error as time shifts are only possible in integer units of the time step. As such, an analytic model describing the relationship between the synchronisation error and the performance was derived and formulated in terms of the sampling frequency and the error reduction. The derived model was validated using significantly oversampled experimental data, downsampled to simulate acquisition over a wide range of sampling frequencies with excellent agreement found. This method of experimentally obtaining the sampling frequency dependence also enabled the comparison of the time and frequency domain-based techniques across an extended range.

Comparing the relative performances over this extended sampling frequency range leads to the definition of distinct regions, within each of which the measurement correction can be optimised by selecting the appropriate technique. These findings were then generalised, based on two parameters, to determine the minimum sampling frequency necessary for the time domain-based technique to outperform the improved frequency domain technique. Given the now two viable and equally effective correction techniques developed herein, each with their own set of requirements and the specifics of the vibration measurement of interest, a framework was developed to allow the user to conveniently select the appropriate correction technique *and* sampling frequency *a priori*. This enables the definition of the optimal hardware characteristics required for a given measurement campaign, which is important for the efficient and practical integration of such sensor solutions.

Chapter 5

Multi-Axial Simulation Table Experimentation

The focus of this thesis up until this chapter has been on the translational motion in the beam axis. However, the introduction of arbitrary non-beam DOF motion, as would be expected during an in-field mobile deployment, introduces *speckle noise* and *aiming errors*. Speckle noise is discussed in Chapter 2 and arises due to relative motion between the LDV beam and measurement surface, whereas aiming errors will cause the measurement beam to record the vibration of an area different to the intended measurement location. Both of these effects can be mitigated by reducing the relative motion between the incident beam spot and the measurement surface, which is the initial focus of this two-part chapter.

Techniques that can maintain an LDV fix on an arbitrarily moving target or throughout arbitrary instrument motion, are known as arbitrary path tracking LDV [4]. While few techniques exist, the most promising of the existing work utilises a vision-based system to locate the measurement location in combination with an SLDV setup to steer the beam in the appropriate direction [86, 42]. This technique will likely be suitable for many terrestrial applications, however, due to the necessary onboard optical and computational resources, the additional weight can make it unsuitable for drone-based applications. As such, this chapter first describes a novel type of arbitrary path tracking tailored specifically for drone-based deployments, similarly utilising an SLDV setup to steer the beam and maintain a fix on a stationary target as the drone undergoes motion. In order to assess the performance of the system, the Multi-Axial Shake Table (MAST) is used to simulate the motion of a hovering drone subject to atmospheric turbulence.

Previous work has shown that it is vital for the correction measurement to be colinear to *and* on the beam axis, otherwise, the corrected measurement could have some residual sensitivity to vibration in the non-beam DOFs [31]. While these criteria are straightforwardly achieved for single-point LDVs by mounting the correction transducer to the rear of the instrument, it is difficult to satisfy during a scanning scenario which would be required during the proposed tracking scenario. Previous literature has considered this situation and theoretically shown through a rigorous vector analysis the required transducer arrangement to enable correction measurements on an SLDV [41]. If three independent transducers are positioned precisely relative to the final steering mirror in the optical path, a triaxial measurement can be taken, whose components can be resolved into a single velocity measurement colinear *and* on the moving beam axis. As such, the second part of this chapter describes the incorporation of the improved signal processing described in Chapter 3 and Chapter 4 into the existing SLDV measurement correction framework, followed by the first experimental validation of this during six DOF vibration using the MAST to simulate the vibration of a vehicle during deployment.

5.1 The galvanometer gimbal

One potential mobile deployment scenario is on an Unmanned Aerial Vehicle (UAV), specifically a drone. With the increase in popularity of drones and the advances in measurement correction, their potential use for vibration measurements has already been recognised for both intelligence gathering [87] and remote structural vibration measurements [32, 33]. However, the applications are by no means limited to these. One promising example of a field that stands to benefit is LDV-based mobile buried landmine detection. Currently, this is conducted from a terrestrial vehicle with a mounted speaker to insonfiy the ground [22, 23, 24]. The sound waves generate seismic waves within the soil, which cause the mine to vibrate and resonate, producing a measurable increase in the vibration directly above the mine at the ground surface. If this work could be replicated from a drone, as it has been done with ground penetrating radar [44], it would enable the identification of buried landmines to be conducted more efficiently with reduced risk of accidental detonation, preventing the loss of life and assets. While the measurement correction techniques discussed herein remove the effects of motion along the beam axis, any translational motion in the two orthogonal axes and any rotational motion would act to divert the beam from the intended measurement location. Image-based tracking solutions could likely handle these tracking requirements [42, 86]; however, these systems tend to be heavy due to the required high-speed camera, additional optics and an onboard computer sufficiently powerful to perform low-latency tracking. While these vision-based systems are viable in scenarios where either or both the LDV and the target are mobile, this functionality is often excessive. For most applications utilising a mobile LDV discussed herein, the target is likely to be stationary; therefore, it is only the effects of the vehicle motion itself that must be mitigated. As such, this section will explore a novel alternative tracking approach tailored specifically to the requirements surrounding drone-based deployment.

The problem of tracking from a drone can be simplified by only considering a hovering scenario. Modern, sensor-laden drones have an impressive ability to hold their translational position in space by constantly making small adjustments in pitch and roll. While some translational motion will inevitably occur, the magnitude depends on the type of drone and the weather conditions. For this section, it will be assumed that this motion is negligible. Therefore, only the pitch and roll need to be corrected in this hovering scenario. Devices that compensate for rotational motion are known as gimbals and are commonly used within the film industry [46]. However, these compensate for whole-body rotation by counter-rotating the entire device. This can cause them to be both heavy, due to the powerful actuators required; and sluggish, due to the inertia of the camera (or LDV). Therefore, this section describes the development of a novel type of gimbal that utilises the galvanometer scanning mirrors typically found in SLDV systems to counter-rotate the beam rather than the entire instrument.

5.1.1 Existing arbitrary path tracking solutions

Tracking LDV (TLDV) systems come in various forms, with most published material assuming a stationary instrument and a mobile target. The simplest type of tracking LDV system is referred to as *self-tracking*. In this scenario, a mechanically coupled optical arrangement is utilised to redirect the beam onto a moving measurement surface, usually a rotating one [88, 89]. A similar effect can also be achieved using an SLDV system, with the addition of positional feedback, typically a rotary encoder. This is known as *tracking* LDV and has been similarly applied to rotating machinery [90, 91, 92]. If the rotation angle is known, then the scan location can be calculated using a kinematic model, allowing the required SLDV pitch and yaw mirrors to be straightforwardly calculated. However, in a mobile deployment scenario, none of these are suitable tracking techniques since the vehicle motion is arbitrary, meaning there exists no *a priori* knowledge of the expected trajectory.

Arbitrary tracking has been achieved on windscreen wipers [91], whereby two string potentiometers are attached, then using the extended string length, the two-dimensional location of the windscreen wiper can be solved for if the two potentiometer locations are known. A different yet similar approach employed a commercial Vicon MX videogrammetry system which uses multiple infrared cameras to track retro-reflective markers in 3D [93]. These markers act as distinct tracking features when illuminated by the infrared ring lights surrounding each camera lens. Therefore, the locations of the markers can be determined in a 3D virtual coordinate system by triangulating the feature coordinates based on the known relative camera positions. As such, when these markers are mounted to an arbitrarily moving target, the motion can be tracked. Finally, using the location of the SLDV within this virtual coordinate system, the vector direction from the LDV location to the target location can be calculated and used to drive the steering mirrors. Despite the reported success of both these arbitrary path-tracking solutions, they both are impractical for mobile deployments since they involve instrumenting the target.

Image-based tracking laser Doppler vibrometry (iTLDV) requires no target instrumentation by utilising a single in-line camera to obtain the relative angular location of the target [42, 86]. Much like TLDV, iTLDV systems utilise a bespoke scanning head that accommodates an in-line high-speed camera that shares the beam axis by mixing the two optical paths using a semi-transparent mirror. If correctly aligned, this acts to remove any parallax error completely. The frames from the high-speed camera are received by a computer running a tracking algorithm that searches for the target to calculate the angle for each respective mirror in order to recenter the measurement location. To avoid scenarios where the system might track the beam spot itself rather than the target, an optical filter is also attached to the camera, which blocks the laser wavelength. Initially, this tracking algorithm searched for a marker placed onto the target [42]; however, more recent work utilising more advanced image processing techniques did not require any markers [86]. The measurements taken using this system included active windshield wipers [42], and a table tennis ball while in-play [86].

A benefit of iTLDV is that it functions irrespective of target or vehicle motion, yet despite the performance and flexibility of this technique, there are a number of drawbacks. Firstly, in order to accommodate an inline camera, modifications should be made to existing SLDV setups, likely voiding any manufacturer warranties, otherwise requiring a bespoke system to be built, which can be costly. Secondly, the required onboard computational and power resources will be heavy. As such, when considering a hovering drone undergoing small adjustments in pitch and yaw, then iTLDV solutions can be unsuitable and excessive. A more practical solution would be the galvanometer gimbal proposed herein.

5.1.2 Limitations of gimbal-based solutions

The galvanometer gimbal is suited for any mounting angle on the drone, however, this section specifically focuses on downward-facing applications, such as unexploded buried landmine detection. In this downward-facing configuration, the yaw and pitch mirrors of the LDV become aligned with the roll and pitch rotations of a hovering drone allowing each mirror to correct a single rotational DOF; whereas in the forward-facing scenario, the roll motion of the drone would require the coordination of both mirrors determined by the rotation matrix. In either case, some misalignment between the drone's centre of rotation in the vertical plane is probable since the LDV system must be mounted on the underside. The result of this misalignment between the galvanometer gimbal mirrors and the centre of rotation will result in a small amount of translational offset as the drone makes adjustments even for a gimbal with hypothetical perfect performance. This can be seen illustrated in Figure 5.1.

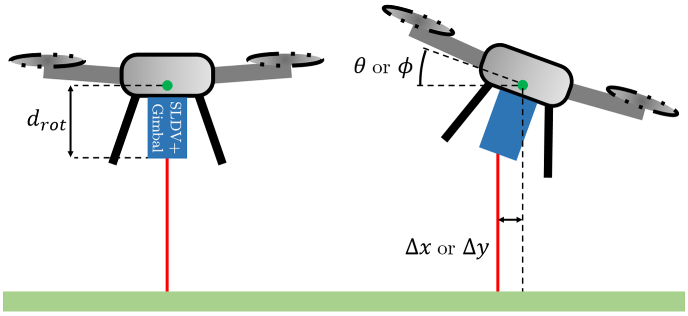


FIGURE 5.1: A depiction of the translation error introduced when the drone centre of rotation, given by the green circle, and the mirror axes are misaligned. Here, the size of the misalignment is given by d and the transitional error is given by Δ.

The pitch and yaw SLDV mirrors are independent and only separated by a couple of centimetres, therefore, it will be assumed that the optical path length from either mirror to the target is equal. If care was taken to mount the SLDV beneath the centre of the drone, which is necessary for in-flight stability, then the relationship between the pitch, θ_{in}, and the roll ϕ_{in}, of the drone and the translational offset introduced by the z offset, d_{rot} is given by:

$$\Delta x = d_{rot} \sin(\theta_{in}) \tag{5.1}$$

and

$$\Delta y = d_{rot} \sin(\phi_{in}) \tag{5.2}$$

Therefore, the total Euclidean error introduced by the z misalignment for a given pitch and roll angle of the drone is given by:

$$\Delta = d_{rot}\sqrt{\sin^2(\theta_{in}) + \sin^2(\phi_{in})} \tag{5.3}$$

5.1.3 Galvanometer gimbal control box

A control box for this galvanometer gimbal was developed using a BNO080 inertial measurement unit, an ESP-32 microcontroller, two MCP4725 12-bit Digital-to-Analogue Converters (DACs), and an operational amplifier circuit. This control box would then interface with an SLDV system to become the galvanometer gimbal by appropriately driving the mirrors, a block diagram of this integrated system can be seen in Figure 5.2. The ESP-32 was selected since it has a higher clock speed than most consumer microcontrollers whilst also having two cores - allowing for both the steering mirror positions to be updated simultaneously. The 12-bit DACs were used since they provide 16 times the resolution of the internal 8-bit DACs in the ESP-32. An operational amplifier circuit was required to properly scale the 0 V to +3.3 V DAC output to the required voltage range for the SLDV steering mirrors used - in this case, ± 3.3 V. Finally, a BNO080 9-axis IMU is used to measure the rotation of the system. This was specifically selected over more popular and cheaper IMUs, such as the MPU6050, for its virtual reality head-tracking applications. This allowed for low-latency and pre-processed roll, pitch and yaw outputs from the Kalman filtering-based sensor fusion occurring on its onboard ARM Cortex M0+. Since the position estimate calculations are running on a separate microcontroller, the processing delays are further reduced by offloading this from the main ESP-32 microcontroller. A schematic of the system can be seen in Figure 5.2.

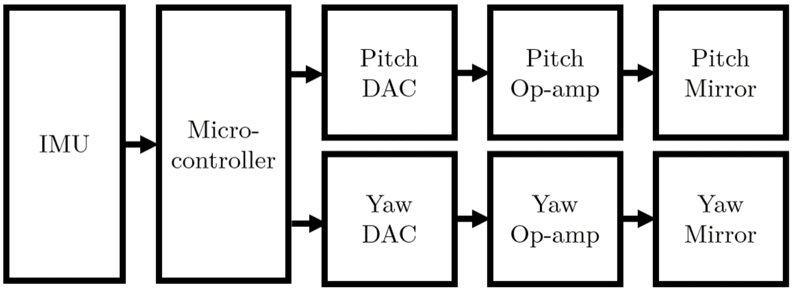


Figure 5.2: Block diagram schematic of the galvanometer gimbal control box components and the galvanometer mirrors of the interfacing SLDV system to achieve pitch and yaw tracking.

The electronics of the system sit inside a 3D printed housing with external BNC jacks such that it can be quickly set up by simply connecting the two mirrors to their appropriate jacks and a third BNC cable supplying 6.6 V for the operational amplifier. The enclosure has an external button on the side which will trigger the microcontroller to establish a new lock on a specific angle, re-centring the beam. The enclosure is then mounted to the scanning head using any non-permanent solution; in this case, large zip ties were used.

5.1.4 Experimental arrangement

The galvanometer gimbal system, which includes both the control box and SLDV, was mounted to the centre of the MAST in order to simulate a hovering drone and characterise the frequency response. To do so required simultaneous measurements of both the MAST angle, referred to as the *input*, and steering mirror angles, referred to as the *output*. While both the MAST and the galvanometer mirrors have position feedback, due to the differing signal conditioning, these would both experience some finite time delay, especially from the MAST, which performs kinematic model calculations in order to estimate its position in space. These finite time delays would introduce errors in the measured phase response. Whilst these could be measured and subsequently removed, an alternative approach was employed. This involved the use of a projection screen orthogonal to the MAST where the LDV laser is incident upon alongside an additional laser of a different wavelength fixed to the galvanometer gimbal system. Therefore, both the pitch and yaw of the MAST and of the galvanometer gimbal system can be known in perfect synchronisation. These locations could then be recorded using a high-speed camera, converting the beam spot locations into their respective angles once the video is analysed in post.

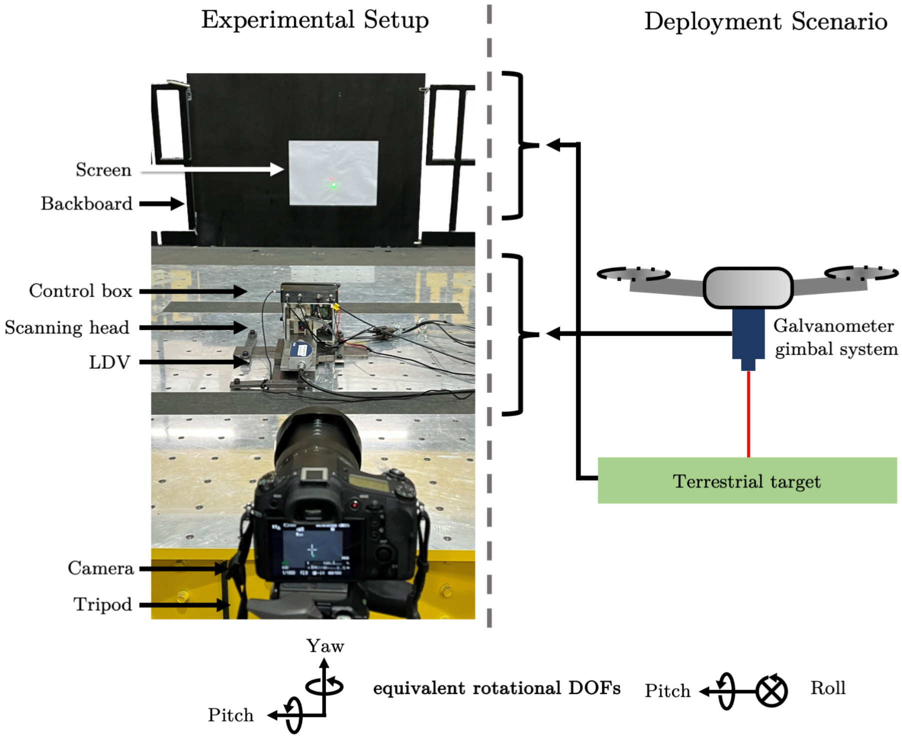


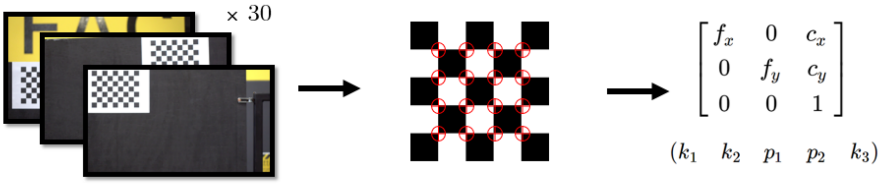
FIGURE 5.3: Annotated photograph of the experimental arrangement used to characterise the frequency response of the galvanometer gimbal system alongside a diagram representing the downward-facing drone deployment. The coordinate system of the MAST evaluation is shown alongside that of the intended drone deployment - illustrating how the MAST yaw becomes the drone roll when downward-facing.

The positions of these beams were recorded using a Sony RX10 II, at 120 fps while the MAST underwent rotational motion in pitch and yaw with the lasers aimed at a white screen 2.55 m from the galvanometer gimbal system. The videos were analysed, giving six 10 s data lengths resulting in frequency domain data with a bandwidth of 60 Hz and a spectral resolution of 0.1 Hz. While the oversampling factor may seem unnecessarily high, it is essential since this increases the shutter speed such that each beam may be resolved as a circular spot rather than a streak - which was the case with the typical frame rate of 30 fps. The SLDV system used here employed the same pair of galvanometer scanning mirrors (GSI Lumonics 000-3008561), associated electronics and optical layout as are in the Polytec Scanning Vibrometer PSV 300; however, in this application used in conjunction with a Polytec NLV-2500-5 single-point vibrometer - despite no actual LDV measurements being taken.

The rotation of the MAST was displacement-driven in closed-loop control white noise with an RMS amplitude of 1° in the range of 0.5 Hz - 5 Hz. This led to a maximum observed angle of about 5°, which was near the limit of the MAST range; however, the galvanometer gimbal system in its current form is capable of tracking in the range of ± 13°. These parameters were selected to simulate extreme conditions for hovering drones. During deployment, the system is intended to be downward facing rather than forward facing; however, due to the practical limitations of implementation, this couldn't be replicated. To account for the different mounting orientations, the appropriate rotational degrees of freedom were driven such that it simulates a downward-facing deployment, illustrated in Figure 5.3. It is important to note that in this setup, the SLDV was aligned with the centre of the table, however, there was an inevitable vertical offset from the SLDV to the MAST surface by about 15 cm, which will manifest itself as some Euclidean misalignment error in the vertical axis of the screen (pitch of SLDV/MAST). During a typical drone deployment the Euclidean misalignment error would affect both axes, as discussed in Section 5.1.2 with an amplitude given by Equation 5.3. To estimate the amplitude of the translational offset introduced in this simulated deployment, the single-axis formulation of the translational offset error can be used, given by Equation 5.2. Taking a maximum rotation of 5°, then the *maximum* error introduced by this misalignment is only 1 cm.

5.1.5 High-speed vision tracking and data analysis

Since vision-based systems will suffer from some degree of lens distortion, the first step is to derive the distortion coefficients and the camera matrix required to undistort the later frames. To do so requires a series of images of a printed checkerboard pattern at various locations within the camera's field of view. These images are then processed using the OpenCV library in order to calculate the aforementioned parameters, thus enabling the undistortion of the experimental frames used in the data collection. This process is



Calibration frames Find 2D coordinates Parameter calculation

FIGURE 5.4: The typical camera calibration procedure in OpenCV which implements previously published work [94].

outlined in Figure 5.4, and begins by importing 30 images; these many images are not necessarily required, but the process is generally improved by increasing the number of images. The next stage is to determine the 2D locations of the intersection points on the checkerboard pattern in pixels using the `findChessboardCorners` function and further refine it into a subpixel value by utilising the `cornerSubPix` function. The latter is able to make subpixel measurements by analysing the gradients of the colours to estimate the true location of the corners. Finally, these coordinates are fed into the `calibrateCamera` function, outputting the camera calibration parameters, f_x, f_y, c_x, c_y, k_1, k_2, k_3 and p_1. This procedure is a standard OpenCV implementation of previously published work [94].

A separate Python script was written in order to efficiently and accurately analyse the experimental frames, extracting the pitch and yaw for each of the two beams - equivalent to the pitch and the roll of a drone during downward-facing deployment. This process is outlined in Figure 5.5. The first stage in this process is to apply the previously obtained distortion parameters in order to undistort the frame. Following this, a small adjustment to the skew of the image is made since the camera is slightly elevated from the screen in order to see over the MAST and the mounted setup. In order to isolate the beams within the image, a background subtraction algorithm is used; specifically, a Gaussian mixture-based background/foreground segmentation algorithm [95, 96]. This feature is conveniently integrated into OpenCV and is fairly straightforwardly implemented, resulting in a masked image containing only the regions which have moved since the previous frame. This algorithm performs best when there are a few frames at the beginning of a video segment only containing the background, therefore, each recorded video here starts with the lasers switched off.

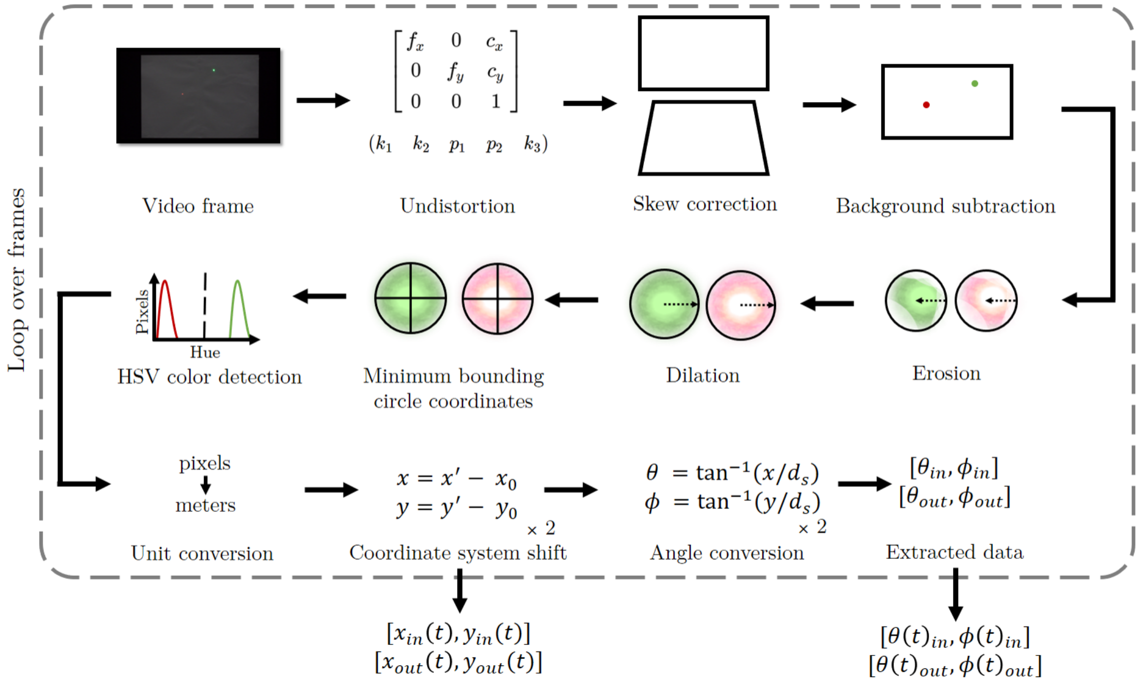


FIGURE 5.5: A depiction of the processing algorithm for a single frame, subsequently repeated for all frames. Where x', y' are the Cartesian coordinates for each beam relative to the frame origin, whereas x and y are relative to the zero position of each beam.

This background subtraction algorithm leaves two clusters of pixels remaining, representing the two incident beam spots; however, their boundaries tend to be non-circular due to some tolerance in the background subtraction algorithm. As such, the next step in Figure 5.5 is a pair of common morphological operations known as erosion and dilation, resulting in a mask boundary that accurately represents the circular boundaries of the two beams. In order to extract the pixel coordinates of these two beam spots, a minimum bounding circle is then applied to each of the two pixel clusters, with the centre location taken to be the beam location. These two pixel clusters then need to be classified based on their colour, as such, the frame is converted from the standard Red Green Blue (RGB) colour space into Hue Saturation Value colour space. Now, the two pixel clusters will have distinctly different mean hue values corresponding to green and red, therefore, allowing them to be easily distinguished from one another.

With all the image processing stages now complete in Figure 5.5, the pixel coordinates of the two beams are then converted into metres using the known size of the screen. However, these are in the reference frame of the image, therefore, this is transformed such that the coordinates are relative to each beam's respective zero position when the MAST is stationary. Now, the pitch and yaw can then be calculated for the MAST and the galvanometer gimbal system using some basic trigonometry and the measured distance to the screen. The repetition of this process across all video frames results in four time domain data types, x, y, Θ and Φ for both the input and output, which is analysed using both time and frequency domain techniques to assess the performance of the system in the next section.

5.1.6 Performance assessment

The performance of the system will be first assessed by observing the deviation of the two beams, representing the scenarios with and without tracking. Figure 5.6 shows the Euclidean coordinates across the screen relative to their respective zero locations for a single 10 s segment. It can be seen that the tracking beam moves considerably less than the fixed beam - which represents the motion of an LDV beam without tracking. Quantifying this decrease in beam movement can be done by calculating the mean absolute deviation across the six 10 s data lengths - giving 8.2 cm for the fixed beam compared to 2.6 cm for the tracking beam, representing a 68% reduction in beam movement.

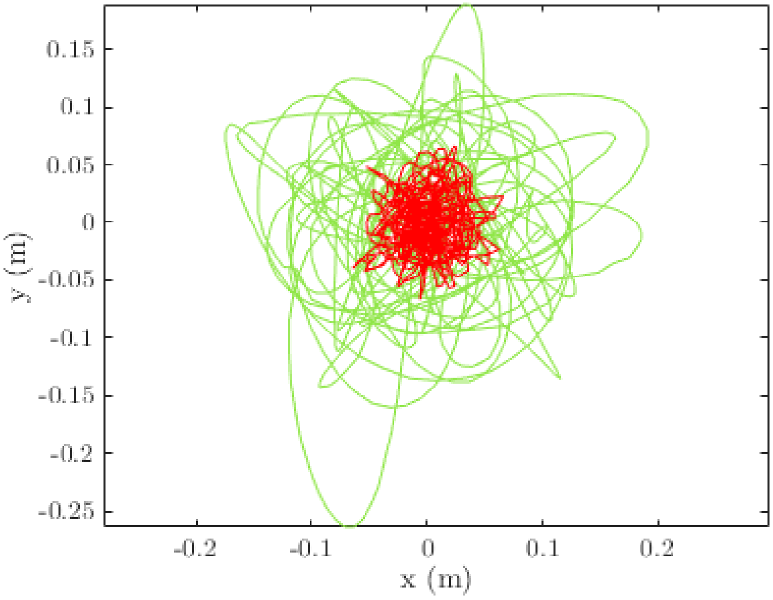


FIGURE 5.6: A plot of the Euclidean movement of the fixed beam, in green, and the tracking beam, in red, across the surface of the screen from their respective zero locations for one of the six 10 s data lengths, with each of the traces comprised of 1200 position measurements.

The second important aspect to consider is the frequency response of the system. This is straightforwardly calculated in the frequency domain following an FFT for each of the six data lengths. A mean Bode plot of all six data lengths can be seen in Figure 5.7.

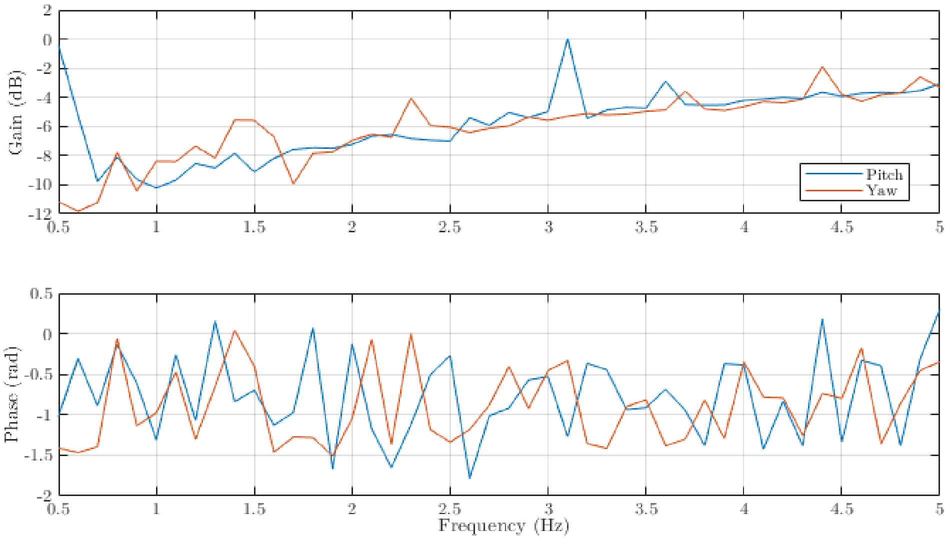


FIGURE 5.7: The mean frequency response of the pitch and yaw axes of the galvanometer gimbal system, with the motion of the MAST as the input and beam angle as the output (df = 0.2 Hz).

Figure 5.7 shows that the galvanometer gimbal system is more capable of correcting for lower frequency motion than higher frequency motion, with its performance steadily declining from about -10 dB at 0.75 Hz to about -3 dB at 5 Hz. However, there exists a large discrepancy between the two axes below 0.75 dB with the yaw axis at -11.2 dB while the pitch axis is only at -0.4 dB. The mean gain of the pitch axis is -5.1 dB while the mean gain of the yaw axis is -5.4 dB. The slight performance discrepancy could be caused by the translational offset error, discussed in Section 5.1.2, which only affects the pitch axes in this configuration. However, this seems unlikely as most of this occurs mostly below 0.75 Hz, whereas the translational offset error would manifest equally for all frequencies. An alternative explanation for this discrepancy could be the behaviour of the IMU. For example, rotations in yaw would rely on the internal gyroscope and magnetometer, whereas rotations in pitch would rely on the internal gyroscope, magnetometer *and* the accelerometer - as the change in the orientation of gravity would be detectable. The IMU employs a Kalman filter to internally fuse the various sensor data; therefore, the re-calibration of the filter to increase the sensitivity in pitch to reduce this dead zone could help.

5.2 Scanning laser Doppler vibrometry measurement correction

It is vital that the correction measurement is colinear to *and* on the beam axis, otherwise, the corrected LDV measurement could have some residual sensitivity to vibration in the other axes [31]. For SLDV systems, obtaining a measurement that adheres to these strict measurement criteria is difficult since the beam angle relative to the LDV body is not fixed; therefore, the accelerometer cannot simply be mounted to the rear. However, as the previous section discussed, the successful integration of measurement correction with the SLDV system is vital for arbitrary path-tracking systems. A theoretical framework previously developed described an arrangement that enables measurement correction on SLDV systems using three independent accelerometers [41]. However, the described arrangement has yet to be experimentally validated and employed the previously established frequency domain signal processing which has been shown to perform relatively poorly in the previous chapters. As such, this section describes the extension of this theoretical framework to incorporate the updated signal processing techniques described in Chapter 3 and Chapter 4. This updated SLDV measurement correction framework is then validated experimentally under scanning scenarios and in the presence of six DOF vibration for the first time using the MAST.

5.2.1 Review of existing work

Both iTLDV and the galvanometer gimbal proposed herein are promising tracking techniques for mobile deployments, both of which utilise SLDV systems. Recent work explored potential SLDV measurement correction possibilities, with a rigorous vector analysis concluding on the optimal accelerometer arrangement for correction in the presence of full six DOF vibration [41]. The previously shown correction measurement criteria of colinear *and* on the beam axis still applying in this scenario [31, 38, 76], the exact transducer arrangement is difficult to implement practically. Through a rigorous vector analysis, it was shown that the optimal correction measurement locations are three independent orthogonal transducers, rigidly mounted such that their sensitive axes intersect at the centre of the final mirror surface.

By taking the appropriate components of each orthogonal transducer, it is possible to make a *virtual* measurement that is colinear to the beam, regardless of the scanning angle. Similarly, by having their sensitive axes intersect at the centre of the final steering mirror, then the measurement is approximately on the beam axis. This cannot be exact since the beam deflection incurred following the reflection from the first mirror to the point of incidence on the secondary mirror. This was taken into consideration during the derivation and shown to have a negligible effect since the point of incidence moves on the order of a few millimetres, depending on the scan angle and the distance between the two orthogonal steering mirrors. The considerable benefit of this proposed technique lies in its ability to similarly be applied to any commercial scanning LDV without any modification to the system, thus keeping the system accessible to the practising vibration engineer for real-world industrial vibration measurement challenges.

Mathematically, the magnitude of the compensation measurement, U_{corr} is derived using the pitch and yaw mirror angles. The mirror angles are half of the optical beam angles, as such, the following has been given in terms of the beam angles to remain consistent with the previous section:

$$U_{corr} = -U_x \cos(\phi_{out}) \sin(\theta_{out}) + U_y \sin(\phi_{out}) - U_z \cos(\phi) \cos(\theta) \tag{5.4}$$

where U_x, U_y and U_z are the x, y and z velocities corresponding to the three specific transducer locations. In this case, accelerometer measurements were subsequently integrated to obtain the velocities. The signal processing employed was common to that of previous work [31, 38, 76] and as outlined in Figure 4.6. As was shown in Chapter 4, the signal processing in these existing works does not adequately handle the accelerometer signals and is limited to stationary signal types. As such, the following section will integrate both the time domain technique, described in Chapter 3; and the improved frequency domain technique, described in Chapter 4, into this SLDV measurement correction scheme.

5.2.2 Improved SLDV measurement correction signal processing

Here, the previously proposed SLDV measurement correction approach [41] is extended to be compatible with the time domain approach discussed in Chapter 3 and the improved frequency domain approach described in Chapter 4. These approaches are largely the same as those previously described, with the necessary accelerometer integration, detrending, and synchronisation duplicated for all three orthogonal correction accelerometers, corresponding to the x, y and z axes. However, both stages contain an additional processing stage where the three orthogonal velocity components are used in conjunction with the current scanning mirror angles in order to calculate the required velocity component which is both colinear *and* on the beam axis. The remaining LDV and reference accelerometer channel handling is no different to the single accelerometer techniques in the previous chapters. The updated SLDV signal diagrams for the improved frequency domain and the time domain can be seen in Figure 5.8 and Figure 5.9, respectively.

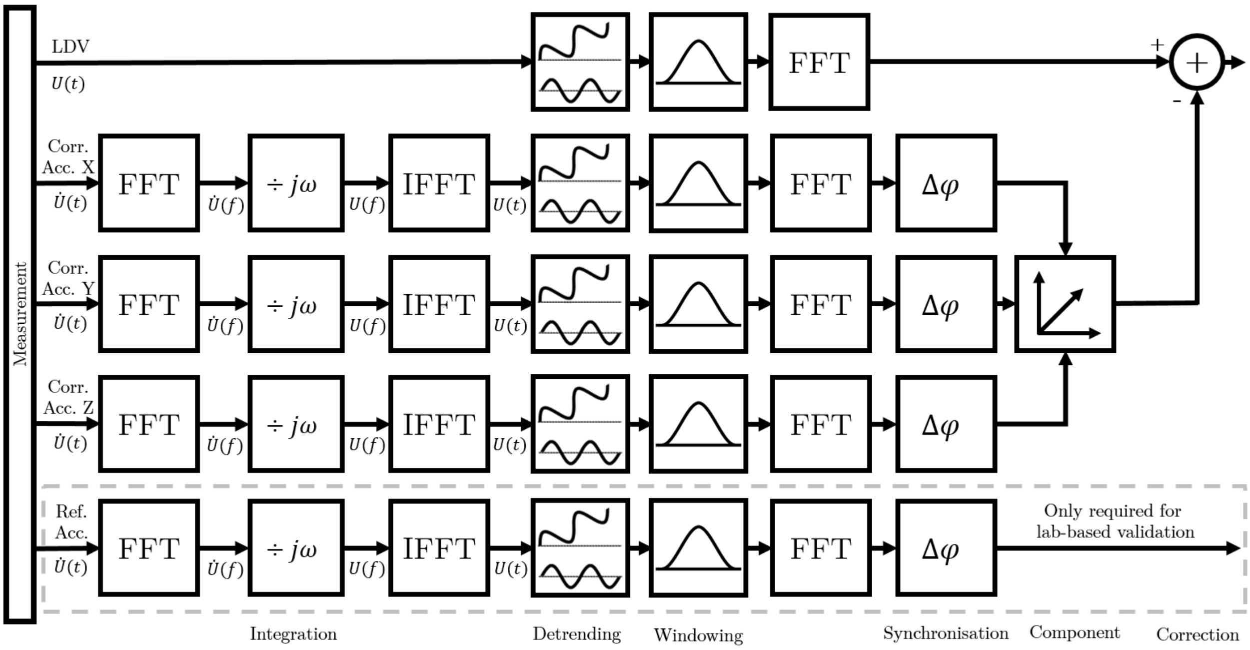


FIGURE 5.8: Signal diagram for SLDV measurement correction in the frequency domain using the improved technique. With the Component block calculated using Equation 5.4.

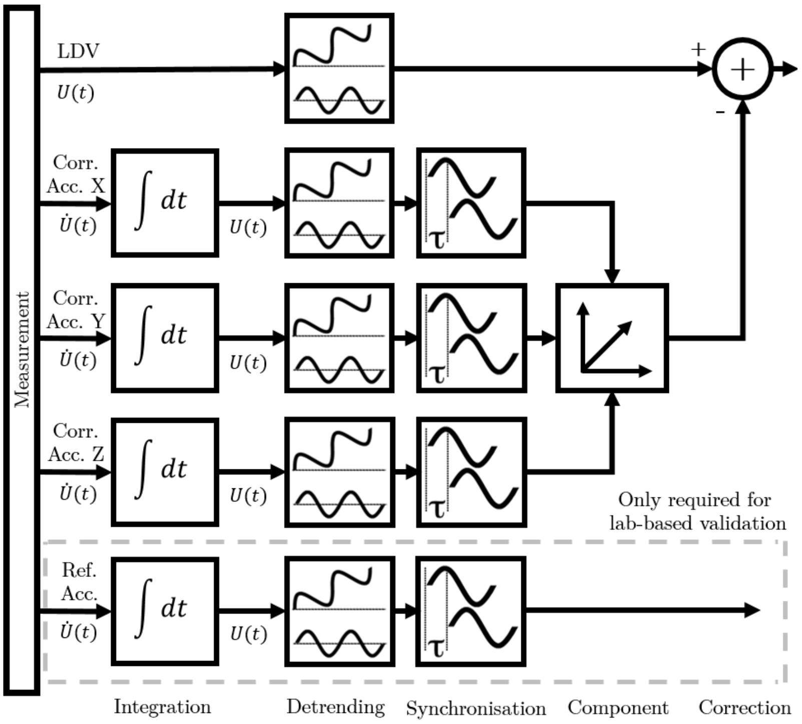


FIGURE 5.9: Signal diagram for SLDV measurement correction in the time domain. With the Component block calculated using Equation 5.4.

5.2.3 Experimental arrangement

The setup employed the same pair of scanning mirrors, associated electronics and optical layout as are in the Polytec Scanning Vibrometer PSV 300; however, in this scenario used in conjunction with a Polytec NLV-2500-5 single-point vibrometer. This was mounted to a stiff, steel base which, by using adjustable straps, could be positioned such that the centre of rotation of the MAST was aligned in the horizontal axes with the axis of the final mirror of the scanning head. The positioning is important to minimise the translational offset experienced by the SLDV sensor head when the MAST is directed to a pitch or yaw angle. Accompanying bespoke steel mounts were constructed such that the x and z correction accelerometers could be mounted to the table and similarly adjusted to be aligned with the centre of the final steering mirror reflecting surface. The final y correction accelerometer was simply mounted to the surface of the MAST in alignment with the centre of rotation and the axis of the final steering mirror. These three accelerometer mounting locations are as outlined in previous theoretical works [41]; a diagram depicting the arrangement is shown in Figure 5.10.

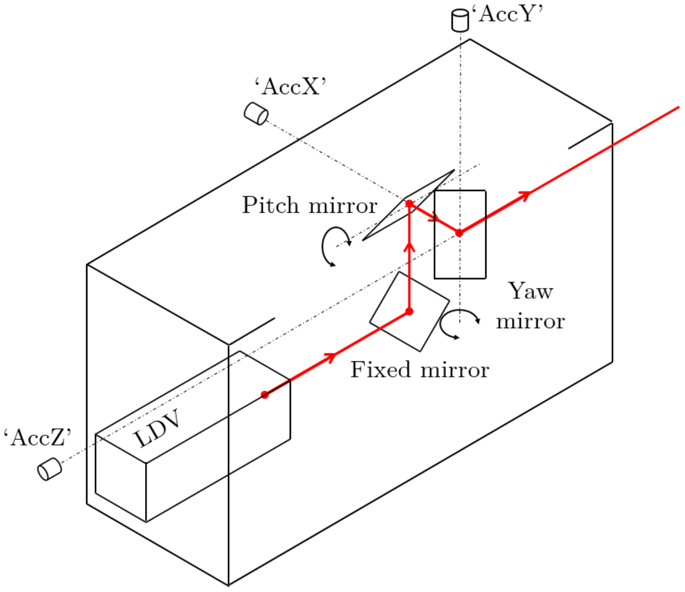


FIGURE 5.10: Illustration of the three orthogonal correction accelerometers positioning, 'AccX', 'AccY' and 'AccZ'.

The reference accelerometer was mounted to the spigot of the target shaker using synthetic beeswax with the SLDV measurement beam incident upon its top surface. Retro-reflective tape was also adhered to the reference accelerometer top surface to ensure optimal SLDV signal strength. All four accelerometers used were Endevco 770F-10-U-120 (20.4 mV/ms^{-2} nominal), as in previous chapters. The target shaker was a Modal Shop SmartShaker K2004E01 with its vibration axis aligned with that of the LDV beam using an adjustable tripod at a 2.2 m distance from the rotation centre of the MAST. During this initial setup, the MAST and SLDV mirrors were in their zero positions, as can be seen in the annotated diagram Figure 5.11. Following the initial alignment, the MAST was directed to a 4° yaw angle and a -4° pitch angle with the opposite angles applied to the SLDV system such that the measurement beam remained incident upon reference accelerometer.

The signals of all three correction accelerometers, the reference accelerometer and the LDV were acquired using a Siemens Digital Industries Software Simcenter SCADAS Mobile data acquisition system. The data were collected at a sampling frequency of 8192 Hz for a duration of 10 s. The sampling frequency was selected using Equation 4.20b derived in Chapter 4; this can be used to select the optimal sampling frequency to minimise the synchronisation error incurred due to the integer time-step limitation imposed during time domain synchronisation. The MAST was driven using closed-loop control in velocity with 0.5 Hz - 50 Hz white noise in all six DOFs, with a translational RMS amplitude of 10 $\mathrm{mm\,s^{-1}}$ and a rotational RMS amplitude of $0.25^\circ\,\mathrm{s}^{-1}$. The resulting incident beam spot motion from the combined six DOF vibration was at maximum 0.5 cm in diameter, which was comfortably contained on the 1.5 cm by 1.5 cm top surface of the accelerometer. The target shaker was driven in velocity with 0.5 Hz - 50 Hz uncorrelated white noise, with the shaker vibration occurring only along the beam axis and the MAST vibration in all six DOFs. The acquired data were processed as five separate 2 s data lengths, similar to before. In the frequency domain, these acquisition parameters lead to a spectral resolution of 0.5 Hz and a bandwidth of 4096 Hz.

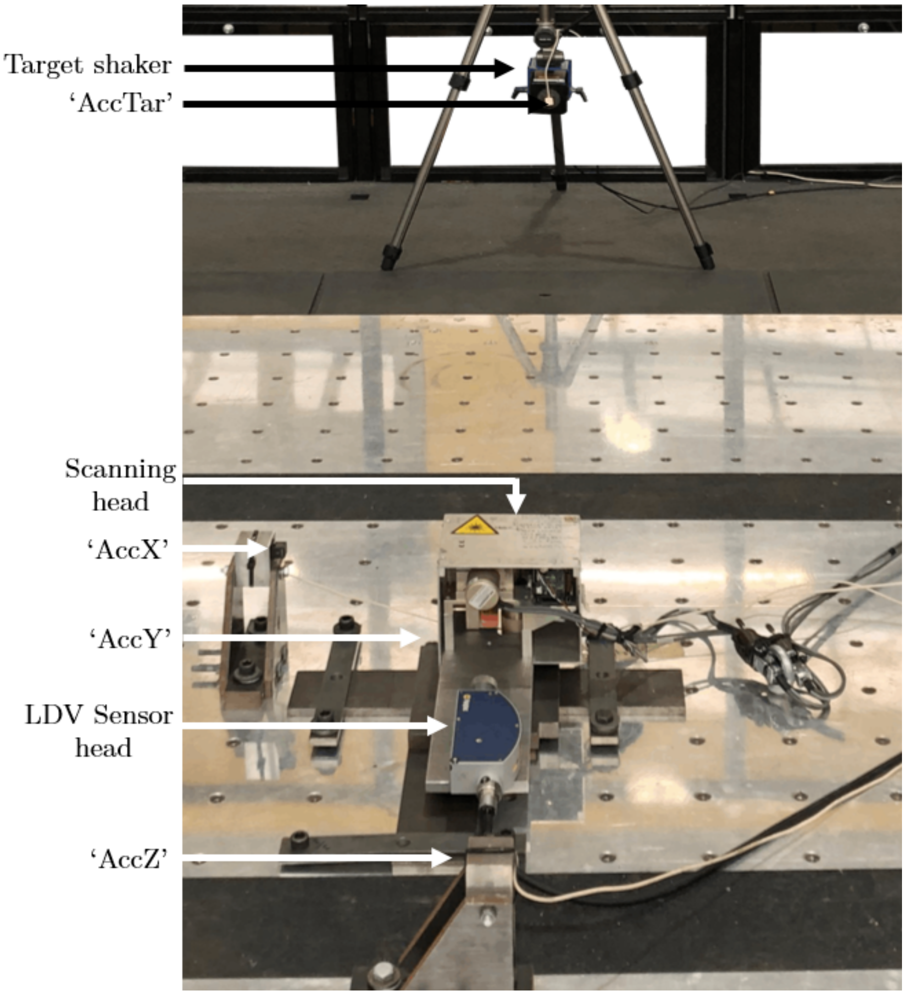


FIGURE 5.11: An annotated photograph of the SLDV mounted to the MAST with ‘AccX’, ‘AccY’ and ‘AccZ’ being the three orthogonal correction accelerometers and ‘AccTar’ being the reference accelerometer mounted to the target shaker spigot.

5.2.4 Performance assessment

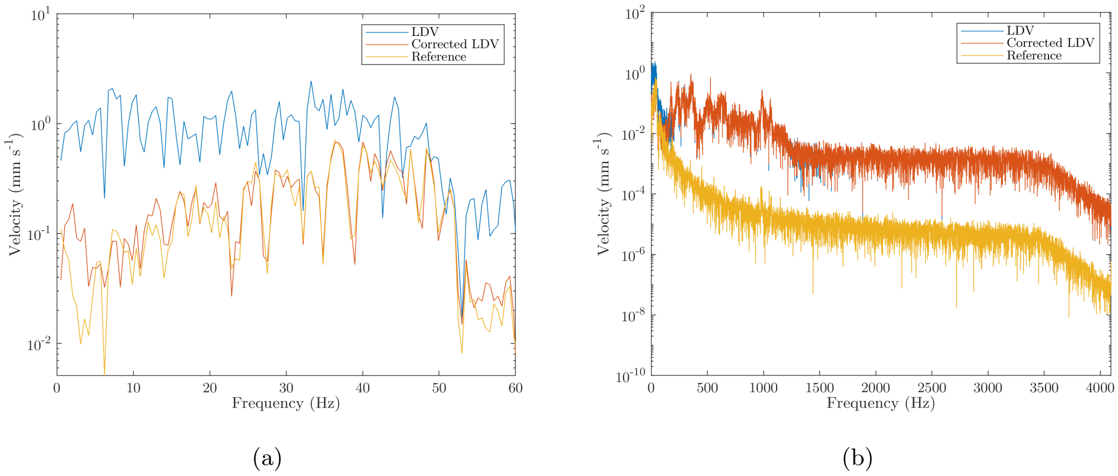
In order to gauge the performance increase gained from the triple accelerometer arrangement [41] as opposed to the single accelerometer arrangement [97] under scanning scenarios, the correction can be carried out using either just the rear-mounted correction accelerometer or using all three and equation Equation 5.4 to calculate the relevant component in the beam direction. As such, Table 5.1 shows the results of the time and improved frequency domain techniques presented herein in either the single accelerometer or triple accelerometer correction. For a fair comparison, the mean error reduction in the frequency domain is calculated over the entire measurement bandwidth rather than only the region of interest.

This is essential for a fair comparison between the signal processing based either in the time or frequency domain.

Table 5.1: Mean error reduction of the two measurement correction techniques developed herein [51, 52]combined with either the single accelerometer [31] or the triple accelerometer [41] hardware arrangement. These are calculated using the domain-specific formulations over the entire frequency range, presented with the associated logarithmic uncertainties.

Transducer arrangement	Time domain signal processing [51]	Frequency domain signal processing [52]
Single accelerometer [31]	$7.5^{+0.2}_{-0.1}$ dB	$7.6^{+0.2}_{-0.2}$ dB
Three accelerometers [41]	$7.7^{+0.2}_{-0.1}$ dB	$7.8^{+0.3}_{-0.3}$ dB

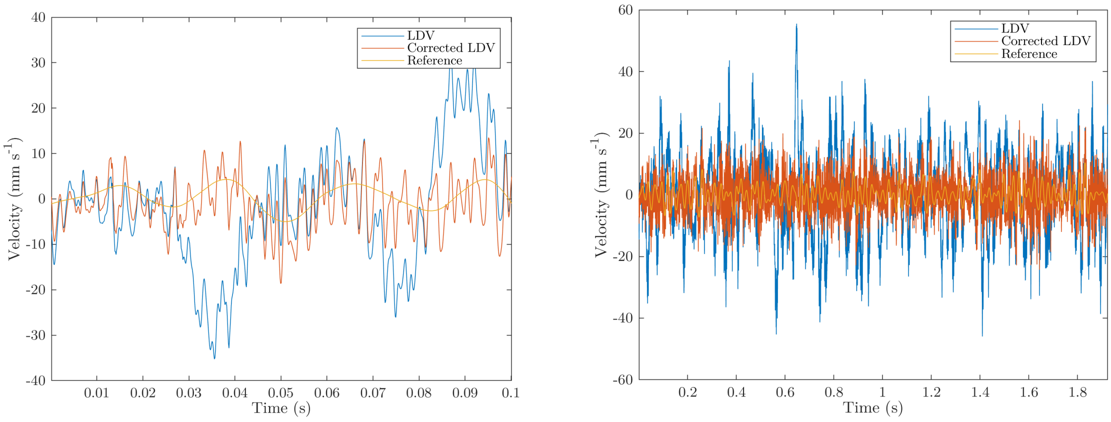
While Table 5.1 does, in fact, show that there is an increase in the mean error reduction when using the triple accelerometer arrangement, it is minimal, with the associated uncertainties of all four combinations of techniques and hardware arrangements containing large overlap. The mean error reduction is dependent on the relative levels of target vibration and instrument/sensor head vibration, therefore, it cannot be expected that these values are to match previous chapters; however, this drastic drop in performance is likely caused by additional factors. In order to understand the source of this poor performance, a spectrum of the improved frequency domain technique can be seen in Figure 5.12.



(a) (b)

FIGURE 5.12: Frequency domain-based correction technique spectra ($df = 0.625$ Hz) plotted over the frequency range of interest (a) and entire frequency range (b).

Figure 5.12 shows the spectra of the LDV signal before and after correction, along with the reference measurement. While Figure 5.12 (a) shows that the correction was successful in considerably reducing unwanted signal content, Figure 5.12 (b) shows that the LDV noise levels were higher across the entire frequency band. This is similarly apparent in the time domain plots, shown in Figure 5.13. The entire data length in Figure 5.13 (b) shows a clear reduction in signal amplitude, resulting in a signal closer to that of the reference measurement. Similarly, Figure 5.13 (a) shows 1 ms of data and the same can be seen, however, here it is clear that the LDV signal is contaminated by higher frequency noise, as is apparent in the spectra shown in Figure 5.12.



(a) (b)

FIGURE 5.13: Time domain-based correction technique time trace over 1 ms (a) and the entire 2 s data length (b).

The mean error reduction is weighted across each sample in the time domain, and similarly, each spectral line in the frequency domain. Therefore, a higher noise floor across the entire bandwidth is sufficient to skew the results in either domain. One obvious solution is not to oversample when not required, allowing the anti-aliasing filters to remove this signal content; however, in this scenario, the high oversampling factor is required for the synchronisation stage in the time domain-based technique, and therefore, to keep it a fair comparison, also the improved frequency domain based-technique. Calculating the mean error reduction over the frequency range of interest is readily achieved in the frequency domain by only including those spectral lines of interest. Similarly, filtering could be applied to the time domain to reject any signal content outside of the measurement range of interest whilst still maintaining the required high-sampling rate for synchronisation. However, in this section, to keep the two techniques comparable, the time domain correction signals are converted into spectra such that both may receive identical treatment during the calculation of the mean error reduction over only the range of interest (0.5 Hz - 50 Hz) in the frequency domain using Equation 3.7.

Table 5.2: Mean error reduction of the two measurement correction techniques developed herein [51, 52] combined with either the single accelerometer [31] or the triple accelerometer [41] hardware arrangement. Calculated using the frequency domain formulation over the frequency range 0.5 Hz - 50 Hz, presented with the associated logarithmic uncertainties.

Transducer arrangement	Time domain signal processing [51]	Frequency domain signal processing [52]
Single accelerometer [31]	$21.0^{+0.8}_{-0.8}$ dB	$20.9^{+0.9}_{-0.7}$ dB
Three accelerometers [41]	$29.3^{+5.9}_{-3.8}$ dB	$28.0^{+7.3}_{-2.6}$ dB

Table 5.2 shows the same data as the previous table, but with the mean error reduction calculated in the frequency domain over the frequency range of interest (0.5 Hz - 50 Hz). This shows a dramatic increase in the reported performance between the single accelerometer and the triaxial arrangement. If evaluated on a linear scale instead of a logarithmic one, this would correspond to an increase of 6.8 times in the time domain technique and 5.1 times in the improved frequency domain technique, respectively. This falls in line with the qualitative gains observed in Figure 5.12 (a). This demonstrates unequivocally the value of the previously proposed hardware arrangement, in combination with the improved signal processing techniques presented herein. It is possible that the increased noise in the signal between 100 Hz and 1.2 kHz is caused by the vibration of scanning mirrors; however, this is the subject of further study.

5.2.5 Iterative reference frame alignment

For proper correction, it is essential that the axes of the accelerometers be aligned with the coordinate axes of the SLDV. While great care was taken during the MAST setup, some misalignment is inevitable since the galvanometer mirrors in the scanning head are not guaranteed to be similarly aligned with the body of the scanning setup. This is particularly present since this semi-bespoke SLDV setup is difficult to precisely to align as the tension screws which mount the galvanometer mirrors introduce some twist upon tightening. Best efforts were made to ensure alignment but some angular error is likely inevitable. As such, this section describes a post-processing-based iterative alignment approach intended to fine-tune the mirror angles relative to that of the LDV body, and hence, accelerometers. This could similarly be applied to factory-aligned SLDV systems to ensure that the body of the LDV and the internal galvanometer mirrors are aligned.

This reference frame alignment technique is based on the premise that the correction performance will peak when the *true* pitch and yaw values are used to calculate the appropriate component of the three orthogonal correction accelerometers. Therefore, in order to find the true angles, the correction algorithms can be iteratively run to search for the combination of pitch and yaw values which leads to a peak in performance. Since this process does not take too long to compute, every combination of the pitch and yaw values are tried within an expected range, $\pm\sigma$, at a predefined resolution, δ. Other more sophisticated search algorithms could be deployed here to reduce computation time rather than computing every combination of pitch and yaw; however, there is little advantage to this as it only took a minute or so to complete. The angles for which the correction should be computed can be expressed as:

$$\Theta = (\Theta_E - \sigma), \quad (\Theta_E - \sigma) + \delta, \quad (\Theta_E - \sigma) + 2\delta, ..., (\Theta_E + \sigma) \tag{5.5}$$

and

$$\Phi = (\Phi_E - \sigma), \quad (\Phi_E - \sigma) + \delta, \quad (\Phi_E - \sigma) + 2\delta, ..., (\Phi_E + \sigma) \tag{5.6}$$

for the yaw and pitch beam angles, respectively. Where Θ_E and Φ_E are the expected yaw and pitch values respectively; in this case since the angle of the MAST. The result of these computations will result in a $(2\sigma/\delta) \times (2\sigma/\delta)$ grid of mean error reductions, whose maximum will correspond to the *true* yaw and pitch beam angles. Both the time and improved frequency domain techniques have been tested here, with this described procedure seen depicted in Figure 5.14.

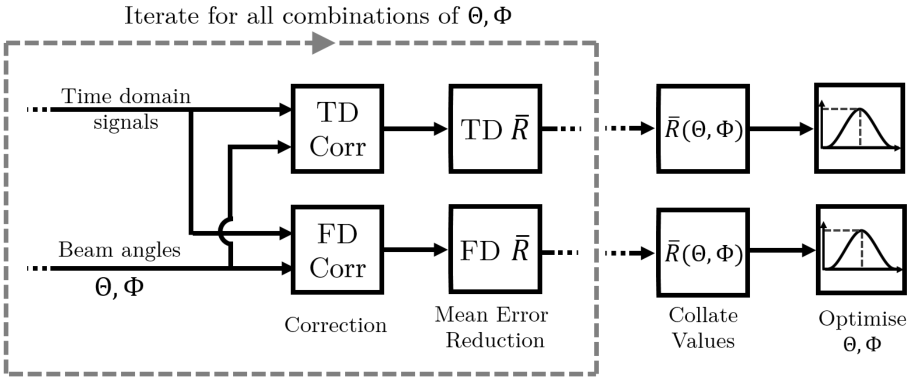
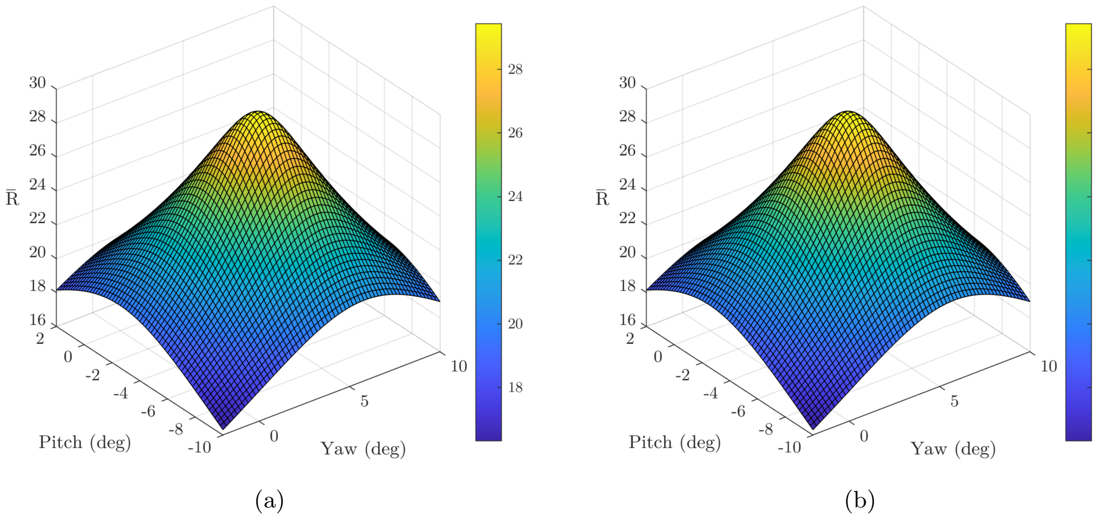


FIGURE 5.14: Illustration of the iterative reference frame alignment procedure for both the time and improved frequency domain techniques. Where the beam angles are used to calculate the component of the instrument velocity in the direction of the measurement beam using Equation 5.4.

For this work, the expected yaw and pitch values were $\Theta_E = 4°$ and $\Phi_E =$ -4°, corresponding to the MAST angle. Similarly, the selected angular search resolution was $\sigma = 0.2°$. Finally, the search range was selected to be far larger than the expected misalignment at $\sigma = 6°$ to better visualise the performance drop-off at larger misalignments for the purposes of this section. Practically, a value reflecting the maximum estimated misalignment should be used. These were then used to calculate the mean error reduction for every combination of angles within this range with each correction algorithm in Figure 5.14 representing the same procedure used in the previous section, that is a mean error reduction of the same five 2 s data lengths previously described.



(a)

(b)

FIGURE 5.15: Surface plot showing the relationship between measurement correction beam angle and the obtained performance with either the (a) time domain technique and (b) improved frequency domain technique.

Figure 5.15 shows the surface plots for both the time and improved frequency domain techniques. Both techniques show a true beam angle of -4.4° in pitch and 4.4°; therefore, the angular offsets are -0.4° in pitch and 0.4° in yaw. Since this is the beam angle, the mechanical misalignment of each steering mirror is half of that - 0.2°. Recalculating the data shown in Table 5.2 using the optimised beam angles, the three accelerometer technique performance increased to $29.4^{+0.8}_{-0.8}$ dB and $28.2^{+6.3}_{-3.9}$ dB for the time and frequency domain, respectively. These improvements are within the associated uncertainties and potentially do not justify the application of this procedure, however, it does give the user confidence that the system is fairly well aligned. SLDV systems with larger misalignments may yield more significant performance gains from this fine-tuning stage.

5.3 Chapter summary and discussion

Up to now, existing arbitrary motion LDV tracking solutions have been vision-based systems. While these systems are capable of tracking in the scenario where both the LDV and the target are mobile, they utilise an inline high-speed camera, meaning either the modification of a commercial SLDV system or a bespoke scanning setup. Similarly, these vision-based systems must be coupled with the necessary power and computational resources to function. The combination of these factors results in a system that is costly and heavy - which is not perfectly suited for drone-based LDV deployment. For scenarios involving a stationary target and a hovering drone, the requirements can be reduced such that only rotational motion in the pitch and roll axes require correction. While traditional gimbals could be used, they can similarly be heavy as they rotate the entire instrument, which requires powerful actuators. Therefore, this chapter proposed a novel galvanometer mirror gimbal intended specifically for downward-facing drone-based applications. This type of system is considerably cheaper and lighter than the vision-based alternatives. Similarly, it can simply be mounted to the body of an existing commercial SLDV system requiring no modifications.

The assessment of the system performance was carried out on the MAST in an arrangement that would simulate a hovering drone. The data acquisition was carried out using using an optical arrangement whereby both the motion of the MAST and the response of the LDV could be measured in perfect synchronisation when recorded with a high-speed camera. This data was then analysed using a combination of time and frequency domain techniques. The results of this assessment showed that the LDV beam motion could be reduced by 69%, with a gain of -5.1 dB and -5.4 dB for the pitch and yaw axes, respectively. Therefore, showing that this type of system represents a viable tracking system with benefits uniquely suited to drone-based LDV measurement campaigns, complementing the steadily expanding body of work pushing the LDV application envelope into new and promising areas.

While measurement correction has been rigorously investigated and improved prior to this chapter, the focus was on LDV systems, rather than SLDV systems. While measurement correction is straightforwardly achieved on an LDV, the situation becomes more complex on an SLDV since the correction measurement must be colinear to *and* on the beam axis, despite the variable beam angle. Innovative work prior to this thesis showed by the means of rigorous vector analysis that these criteria can be met using three independent precisely positioned contacting transducers [41]. By mounting them such that their orthogonal sensitive axes intercept in the centre of the final mirror surface, their components can be resolved such that a velocity measurement that is colinear and on the beam axis is obtained.

While this aforementioned SLDV measurement correction hardware arrangement holds great promise, the work employed the previous signal processing approach, which has since been superseded by the techniques described in Chapter 3 and Chapter 4. As such, the second part of this chapter focused on the expansion of this triaxial accelerometer arrangement to incorporate the signal processing techniques proposed herein. This updated SLDV measurement correction technique was then validated on the MAST undergoing six DOF white noise vibration during a four degree pitch and yaw scan angle. When compared to the single accelerometer technique, results show that the triaxial arrangement leads to a performance increase of 6.8 times and 5.1 times for the time and frequency domain techniques, respectively. This demonstrates unequivocally that serious performance gains can be met by implementing the triaxial arrangement in scanning scenarios.

Following this, a new iterative reference frame alignment approach is proposed. Using the premise that measurement correction performance is optimised at the 'true' angle, the correction can be iteratively run over a range of marginally adjusted angles, creating a surface plot of performance against measurement correction scan angles. An optical mirror misalignment of 0.2° in both mirror axes was determined, resulting in a slight performance increase. A mirror misalignment this small is difficult to remove by hand, as such, this system can be considered to be aligned, with the adjusted pitch and yaw values being optional. While the benefits here were minimal, the true value in this iterative fine-tuning approach would be for either determining the misalignment in a system known to be out of alignment such that it may be accounted for, or for validating the alignment of a system, as was achieved here.

The combination of these two sections results in a practically viable tracking *and* measurement correction technique which when used in tandem can mitigate the effects of speckle noise, aiming errors and vehicle vibration. In order to confirm the true efficacy of this combined technique, the two sub-components should be integrated into a real deployment scenario, where both correction and tracking are taking place. This would require the angles of the two orthogonal mirrors to be simultaneously logged as they are tracking, enabling the appropriate measurement correction components to be calculated for each data point in the time series; naturally lending itself to the time domain technique. This test should be experimentally validated on the MAST and include a reference measurement, as has been shown in previous chapters, such that the performance can be quantified with and without the combined motion tracking and triaxial measurement correction. This should experimentally prove that there are reductions in speckle noise, aiming error and vehicle vibration. Prior to this test, the galvanometer gimbal system could be improved in several ways. Firstly, integrating an IMU with a higher output rate would increase the system's responsiveness, reducing the introduced error at any instant. Secondly, while this system is intended for downward-facing applications, such as buried landmine detection, it should be extended to work with any mounting angle. The mirrors on an SLDV correspond to pitch and yaw which are straightforwardly tracked during a downward-facing application as these translate readily to the pitch and roll of the drone. However, if the system were forward facing, the two orthogonal mirror angles would need to work in tandem to recenter the target location; these angles are readily calculated using a rotation matrix.

Chapter 6

Non-contact vibro-acoustic object recognition

The work described in previous chapters outlines a comprehensive framework for both measurement correction and arbitrary path tracking, the amalgamation of which enables the deployment of LDVs on mobile vehicles and unlocks countless potential applications. Some examples already receiving some interest include: buried landmine detection [98, 99, 100], terrestrial seismology [25], orbital seismology [27, 28, 29], structural health monitoring and intelligence gathering from drones [30, 31, 32]. As such, it is straightforward to imagine the extension of these various contemporary vibro-acoustic detection, localisation and classification techniques into solutions that can be deployed on mobile *autonomous* vehicles, especially since this immediately extends their application envelope to scenarios in hostile environments, not suitable for human presence.

In that context, the work presented in this chapter adds a new addition to the ever-expanding body of LDV applications: non-contact vibro-acoustic object recognition. Prior to the work presented herein, acoustic object recognition approaches described in the literature involved *contact* excitation techniques. These excited an audible response in the object using either a simple actuator [47, 48], or a multiple degrees-of-freedom robotic arm

[49, 50], such that the recorded sounds could be classified using a variety of signal processing and machine learning techniques. The simple actuators excited the objects using a shake motion [47, 48], whereas robotic arms employed various exploratory behaviours such as *lift, shake, drop, crush, and push* motions to be compatible with a broader range of objects [49, 50]. All these techniques have focused on the acoustic classification of household objects and have obtained accuracies as high as 96.7%. Despite this success, exciting an object via touch has some drawbacks. Firstly, contact necessitates physical access to the object, requiring the robot to move towards the object, slowing down and complicating the task. Secondly, the object must be excited with enough force to produce an audible response; therefore, some fragile objects could be damaged during the excitation - for example, *dropping* a glass object. Therefore, the new vibro-acoustic object recognition technique presented in this chapter addresses these shortcomings with an entirely non-contact technique.

To achieve an entirely *non-contact* vibro-acoustic object recognition technique, the contacting actuator of the robot was substituted with an acoustic excitation generated by a loudspeaker. However, the response due to this acoustic excitation will be orders of magnitude lower in amplitude than one excited by touch, therefore, generating little to no detectable sound. As a result, an LDV is used in place of the conventional microphone used in earlier studies to measure the low-amplitude vibrational response of the object itself rather than measuring the sound the object produces as a result of the excitation. The measured responses are then processed using various spectrogram-based techniques and used to train a convolutional neural network (CNN) via transfer learning (TL). The training set size and pre-processing techniques are then optimised using a rigorous five-fold cross-validation, resulting in a CNN which can obtain an accuracy of 99.7%. As practical applications will require the technique to classify objects belonging to broader object classes rather than a specific instance of an object, the technique is then shown to successfully recognise never-before-seen instances of an object class with similar near-perfect accuracy. This performance exceeds that of existing *contact* techniques whilst simultaneously introducing *non-contact* functionality, positioning it as a viable object recognition technology for the potential deployment on autonomous systems and other machine automation tasks. This chapter is orientated around the body of work that was published in the Sensors special issue named Artificial Intelligence-Based Audio Signal Processing [53].

6.1 Review of related work acoustic object recognition work

As mentioned in the previous section, there are no non-contact acoustic object recognition techniques described in the literature; however, this section will cover the relevant aspects of the existing contact-based techniques to allow the work presented in this chapter to be contextualised. The following discussion will be divided into two main categories, simple shake actuators and multiple DOF robotic arms. Due to the nature of the shake excitation, the test objects are largely limited to those which emit a sound when shaken; however, multiple DOF robotic arms explore a wider range of objects and excitations. All the discussed techniques employed traditional machine learning algorithms, as opposed to the NN-based classification that will be used in the new method proposed in this chapter.

6.1.1 Simple shake actuators

Simple shake actuators consist of a mechanism able to undergo a reciprocating motion and a mount to hold the object being excited. The simplest of these consisted of a slider-crank mechanism, leading to a linear reciprocating motion with an amplitude of 10 cm and a frequency of about 2 Hz that was transferred to a container via a clamp at the end of the arm - intending to mimic how a human might shake a container [48]. The objects studied can be seen in Figure 6.1 and represent an assortment of 12 items, such as nuts, bolts and even chocolate. The unusual variety of objects here is partly due to the focus of the work being specifically on not only identifying the objects within the container but also estimating the number of the specific object within the container.

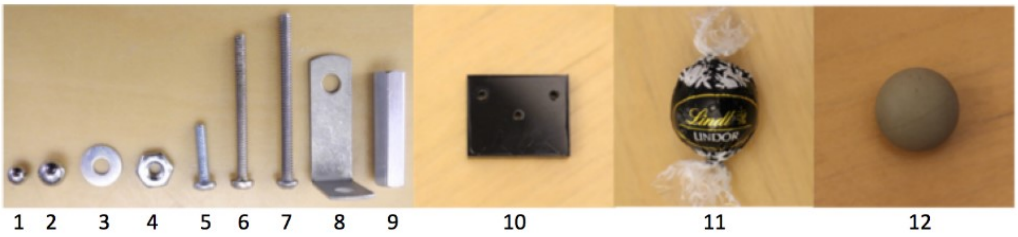


Figure 6.1: Photograph of the 12 object types used for acoustic object recognition with a simple DC motor shake actuator [47]. © 2017 IEEE, reprinted with permission.

The work focused on the use of a contact microphone, also known as a piezo microphone, to measure the responses on the exterior of the container while it is being excited using the shake actuator - directly measuring the vibrations of the container in a similar fashion to the LDV-based technique of this chapter. A support vector machine was used to classify 1 s measurements of 28 g batches of all 12 items seen in Figure 6.1. The mean accuracy when classifying the responses was 93.8%. Following this, and less related to the work proposed in this chapter, the paper goes on to assess the ability of the technique to estimate the number of objects seen in Figure 6.1 within the container, showing that an accuracy as high as 99% is obtainable for some objects when the acceptable tolerance of the estimation is ±2 objects.

The second paper utilised a pneumatic actuator to shake objects within a small arc, similar to that of a human shake [47], as can be seen in Figure 6.2. A feedback potentiometer was integrated into the arm to enable closed-loop control, driving the arm with a sinusoidal signal. The frequency of this shaking was not stated but lasted for a duration of 10 s, ten times longer than that of the previously described work. The end of this arm contained a simple G-clamp to allow it to grasp objects of various shapes and sizes.

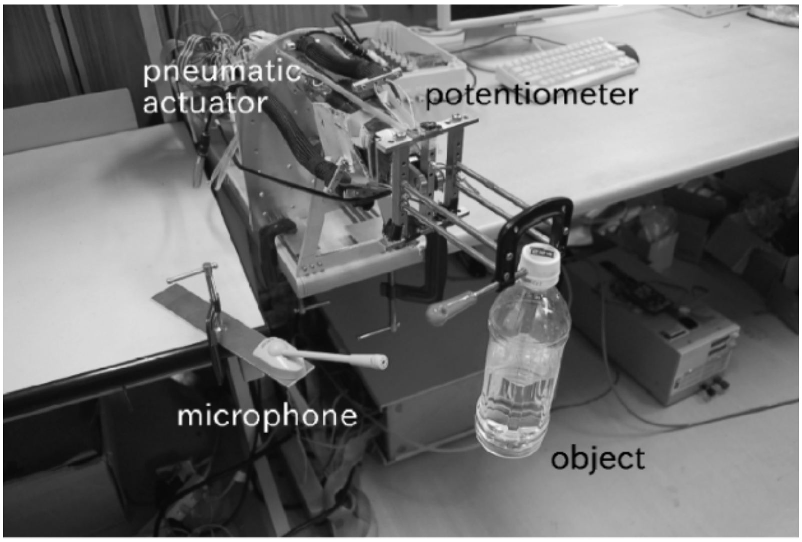


Figure 6.2: Photograph of the pneumatic shake actuator used for acoustic object recognition [47]. © 2022 Informa UK Limited, reprinted with permission.

The work focused on 3 broad categories of objects, seen pictured in Figure 6.3, consisting of rigid objects, such as a spanner; paper materials, such as a newspaper; and water bottles of various shapes and sizes. This represents a limited set of objects, especially when considering that the rigid objects will not emit a sound when shaken and, therefore, would be easily distinguishable. However, unlike the other works described in this literature review, this work investigated *broader object classes*, which requires the classifier to identify features in the recorded sound common to all instances of the objects; for example, understanding the general sound of paper rather than the sound of a specific piece of paper, therefore, identifying a never-before-seen instance of a broad class. This approach is more in line with the form of acoustic object recognition that people are capable of; for example, anyone could distinguish the sound of rustling paper from the sound of a shaken water bottle - regardless of the specific paper or water bottle used.

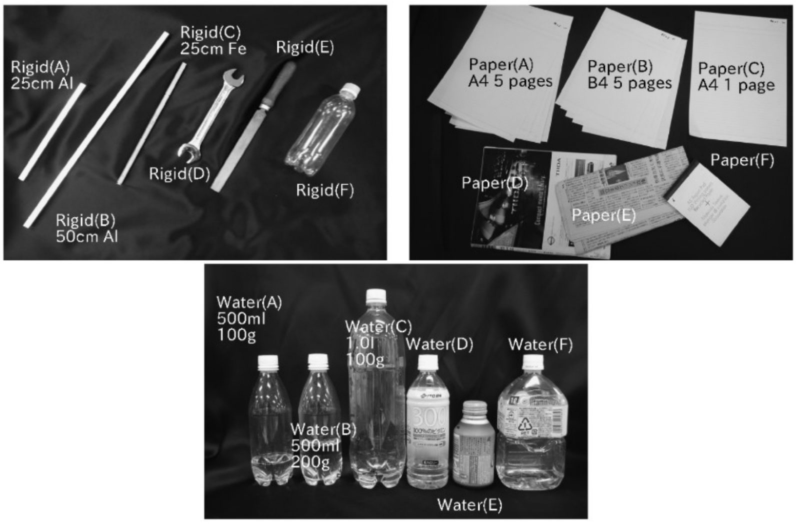


Figure 6.3: Photograph of the 3 broad objects categories used for acoustic object recognition with a simple shake actuator [47]. © 2022 Informa UK Limited, reprinted with permission.

This work used a traditional microphone in order to record the sounds of the objects being shaken, which were converted into amplitude spectra. These spectra were then classified using self-organising maps to identify a previously unseen instance of a broader object class. Results showed that water bottles could be identified with an accuracy of 92.5%, paper with an accuracy of 95% and rigid objects with an accuracy of 100% - giving a mean accuracy across all three classes of 95.8%. It is unsurprising that rigid objects obtained a 100% classification accuracy since they would simply emit no sound when shaken. It is important to note that with only three potential classes, a faulty classifier would correctly identify the object with an accuracy of 33% of the time, meaning these accuracies would likely decrease as more object classes were introduced.

6.1.2 Robotic arm actuators

While robotic arms are orders of magnitude more complex than the simple actuators described, they are able to enact a range of exploratory behaviours. The simplest implementation of these utilised a Universal Robotics UR5 six DOF robotic arm in order to excite small containers with three different rotations and one linear shake motion. Using an AG-95 end effector, the arm picked up each container to then excite an audible response, which is then measured using a nearby traditional microphone, as can be seen pictured in Figure 6.4.

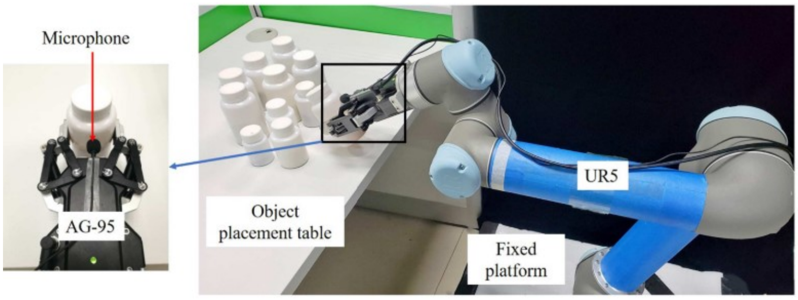


Figure 6.4: Photograph of the UR5 robotic arm used to excite an audible response from various filled containers [50]. CC BY 3.0.

There was a total of 10 containers filled with a wide range of contents, largely pantry items and medicine, as can be seen in Figure 6.5. These represent a diverse and challenging set of contents, for example, including two types of rice which will likely be difficult to distinguish from one another. Similarly, with 20 types of container contents, the chance of a faulty classifier correctly identifying the item is reduced to only 5%.

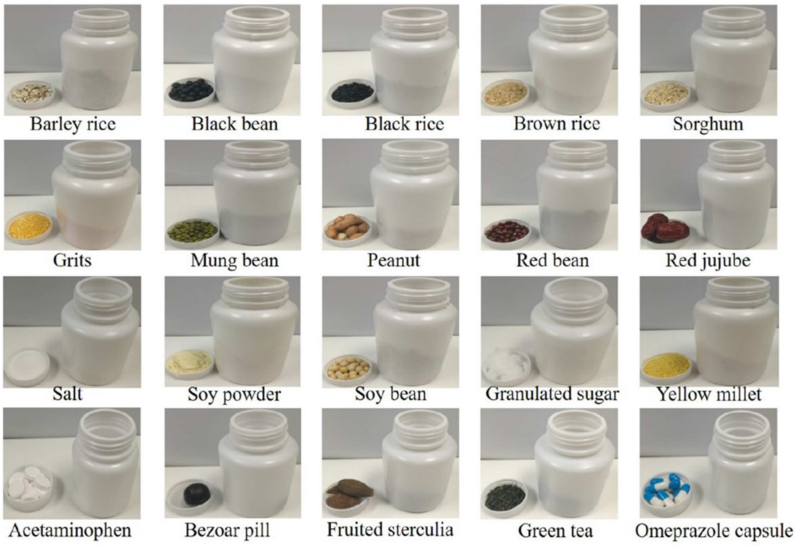


Figure 6.5: Photograph of the contents of the 20 filled containers used in the acoustic object recognition work [50]. CC BY 3.0.

The sound data were measured for a total of 6 s, starting prior to the beginning of each exploratory behaviour. These measured sounds were then converted into Mel-Frequency Cepstral Coefficients, which are commonly used in audio analysis. This work compares four traditional machine learning algorithms (k-nearest neighbour, support vector machines, sparse representation classification and Kernal-k-nearest neighbour) to find which could best identify the contents of each container. Results showed that the data obtained by rotating the container through 180° yielded the best results, with accuracies for the four machine learning techniques ranging between 75% and 86%.

The final paper utilised the most biomimetic approach of all the robots, using an entire upper-torso humanoid robot, with two seven DOF Barrett WAM arms and two three-finger Barrett hands as end effectors [49]. The robot's head was equipped with a microphone, keeping the experimental setup in line with how a person would both interact and hear objects. The robot employed the widest range of exploratory behaviours, including *lift, shake, drop, crush, and push* motions. The experimental setup can be seen in Figure 6.6.

FIGURE 6.6: Photograph of the humanoid robot used in existing acoustic object recognition work [49]. © 2011 SAGE Publications, reprinted with permission.

This work employed the largest set of items than any of the previous works, generating a library of 50 items, as can be seen in Figure 6.7. These items consisted of a range of household items made from materials such as plastic, fabric and more. This work recorded both the sounds the items emitted when excited and the proprioceptive data in the form of torque time series from the robot's actuators, further mimicking the human experience.

FIGURE 6.7: Photographs of the 50 household object selection used in existing acoustic object recognition work [49]. © 2011 SAGE Publications, reprinted with permission.

Each object was then classified by using the actuator torque data and the discrete Fourier transform of the audio data using self-organising maps for each exploratory behaviour. The probabilistic prediction of each exploratory behaviour was then used together with a weighted combination to predict a single class. A similar weighted combination was also taken using the proprioceptive data and the audio data. The results show that using only audio data or proprioceptive data obtained an accuracy of 93%, however, when these two were similarly combined using a weighted output, the accuracy increased to 98.2%.

6.2 Non-contact data collection and data pre-processing

With the existing contact acoustic recognition techniques summarised, this section will now describe the proposed non-contact technique hardware arrangement and data handling.

6.2.1 Object selection

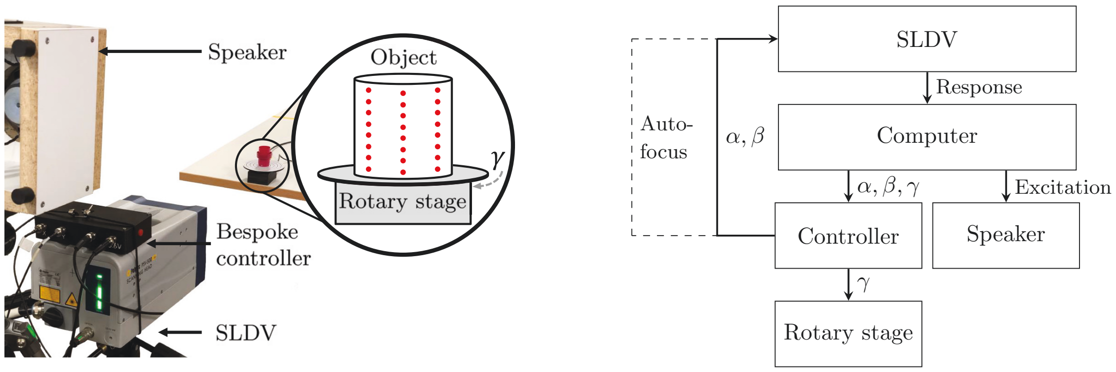
The objects used in this new non-contact vibro-acoustic object recognition technique are shown in Figure 6.8. The objects were selected to satisfy the following three main criteria. Firstly, objects were selected to enable this research to remain within the scope of the previous work presented in Section 6.1. These objects are made from a range of common materials, including glass, ceramic, metal, plastic, and wood. Secondly, triplets of similar objects were included to assess the sensitivity of this technique to distinguish similar objects *and* the ability to generalise across broader object classes, similar to previous work [47]. As such, there are four main subgroups composed of three similar objects: i) tennis balls, ii) table tennis balls, and iii) full or iv) empty soda cans. For example, the system could either distinguish the three tennis balls from one another *or* it could identify that a never-before-seen instance of tennis ball belongs to the broader object class of “tennis balls”. Finally, since a rotary stage is used to automatically collect vibro-acoustic data, the objects all possess some degree of axial symmetry and are small and light enough to fit on and be manipulated by the rotary stage.

(a) (b) (c) (d) (e)
(f) (g) (h) (i) (j)
(k) (l) (m) (n) (o)
(p) (q) (r) (s) (t)
(u) (v) (w)

FIGURE 6.8: The 23 household objects characterised in this paper: (a) bottle cap, (b) AA battery, (c) empty jar, (d) empty container, (e) small wooden block, (f) porcelain mug, (g) metal cup, (h) plastic cup 1, (i) plastic cup 2, (j) plastic cup 3, (k) table tennis ball 1, (l) duct tape, (m) tennis ball 1, (n) tennis ball 2, (o) tennis ball 3, (p) table tennis ball 2, (q) table tennis ball 3, (r) empty 375 mL soda can, (s) empty 250 mL soda can, (t) empty 500 mL soda can, (u) sealed 375 mL soda can, (v) sealed 250 mL soda can, (w) sealed 500 mL soda can. Here, the empty soda cans have been drained of their containing liquid, and the sealed soda cans were full of soda.

6.2.2 Automated data acquisition system

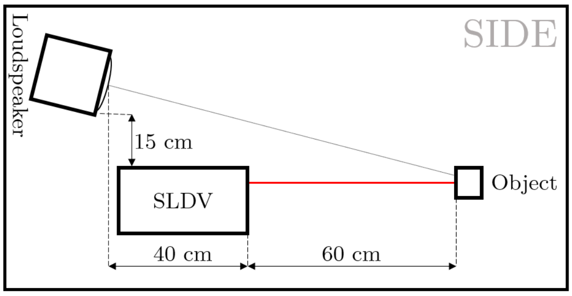
It is challenging to predict in advance how much data will be needed to produce a specific performance outcome when applying deep learning to a new problem. As a result, a large number of responses were collected for each object, enabling the determination of the ideal data set size by taking subsets of the entire obtained data to simulate the smaller data sets. A bespoke automated data collection system was built to enable the rapid collection of hundreds of measurements per object, as seen in Figure 6.9. The rotary stage was custom-built and consisted of an Arduino Nano and a 28BYJ-48 stepper motor, contained in a 3D printed housing with a similarly 3D printed turntable attached to the motor shaft, allowing the objects to be rotated about a vertical axis in front of the speaker and SLDV. In order to control the SLDV, the bespoke galvanometer control box described in Chapter 5 was used, receiving mirror angle commands from the laptop over serial communication, then applying the appropriate voltages to the scanning mirrors in the SLDV; with the IMU playing no role here. Two ±3V analogue outputs - one per scanning mirror - and a single digital output to the rotary stage enabled control of the object rotation and SLDV beam orientation.



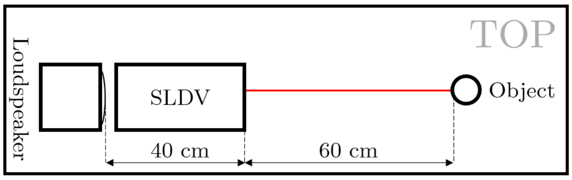
(a) (b)

FIGURE 6.9: Experimental arrangement; (a) showing the physical setup with the SLDV measurement grid annotated, (b) block diagram of the components where solid and dashed lines are wired, and wireless connections, respectively. α and β are the scanning mirror angle command signals while γ is the rotary stage angle command signal.

Since LDVs measure the surface velocity of a target in the direction of the laser beam at the incident beam spot, any change in the beam angle and measurement location will change the nature of the measurement, thereby adversely affecting the subsequent object classification performance. While the object location is fixed in this work, the eventual practical application of this technique would need to be insensitive to the relative object-robot positioning. As such, care was taken to sufficiently diversify the training data to better represent a real application scenario. Practically this was achieved by moving both SLDV steering mirrors and the rotary stage to collect a total of 864 responses at various locations and angles of incidence. A schematic with the relevant dimensions labelled as can be seen in Figure 6.10; it can be noted that a portion of the measurement content will be the vibration of the SLDV steering mirrors themselves due to the proximity to the loudspeaker, however, it is unlikely to affect the classification accuracy as this is the same for each object. The responses were collected over an 8 × 3 grid for every 10° increment of the rotary stage. The height and width of the measurement grid for each object were chosen to be 80% of the total height and width of the particular object. Once responses at each of the 24 grid points are complete, the rotary stage rotates by 10 degrees, and another 8x3 scan commences. This procedure was repeated until scans were collected around the entirety of the object, giving 864 responses and taking approximately 20 minutes per object.



(a)



(b)

FIGURE 6.10: A side view and top view schematic of the test apparatus with the relevant approximate dimensions annotated.

The entire operation was orchestrated by a laptop running a Python script that sent angle commands to the galvanometer gimbal control box via a USB serial link; the controller then applied the corresponding voltages to the External Scanner Control of the Polytec PSV-500 Xtra Scanning Laser Vibrometer to adjust the mirror positions[1] and then relayed the required object orientation angle to the rotary stage. The acoustic excitation imparted on the target was a 1 s long, 1 Hz to 20 kHz linear chirp played during each

---

[1] External Scanner Control (EXT) is an optional extra on some Polytec scanning laser Doppler vibrometers.

measurement through a 125 mm 5 Ω AS3007 loudspeaker, amplified by a Kemo Electronic 12W M032N amplifier module. The selected frequency range was chosen to be as broad as possible while utilising a standard audible frequency loudspeaker. The maximum A-weighted sound pressure level recorded at the object position was 67 dB re 20×10^{-6} Pa, measured with a Digitech QM1592 sound level meter. Both the excitation and response signals were played and recorded through the headphone and microphone jacks on the laptop for convenience, using a sampling frequency of 44.1 kHz. The room itself was a concrete-walled laboratory of about 8 m by 8 m with little ambient sound.

6.2.3 Management of measurement challenges

As discussed in Section 2.2, successful LDV measurements usually rely on the target surface being optically rough such that sufficient light may be collected in direct backscatter by the instrument optics. Importantly, this enables vibration measurements of surfaces that are not normal to the beam. However, optically rough surfaces de-phase monochromatic coherent light, leading to the formation of a speckle pattern [64]. For a moving target or laser beam, the speckle pattern will change in sympathy to the movement which can present a number of measurement challenges [4]. While the PSV-500 Xtra includes some automatic features to overcome some speckle challenges, such features were not available as part of this work since externally generated signals controlled the scanning mirrors. Laser beam auto-focusing is another increasingly important feature of commercially available scanning LDVs, as this also contributes to optimising the chances of obtaining high-quality automated measurements across the range of points on the object surface. As such, a custom solution was developed to handle these two aspects and maximise the collection of high-quality object responses in this campaign.

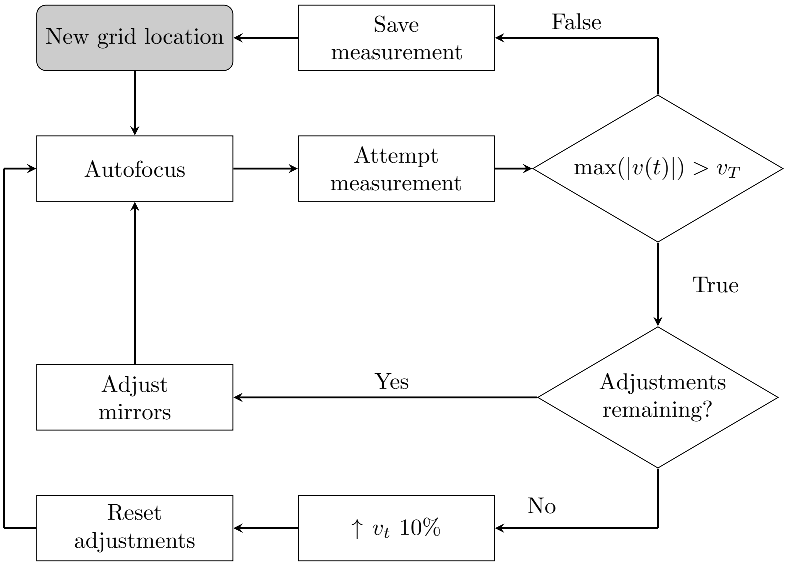
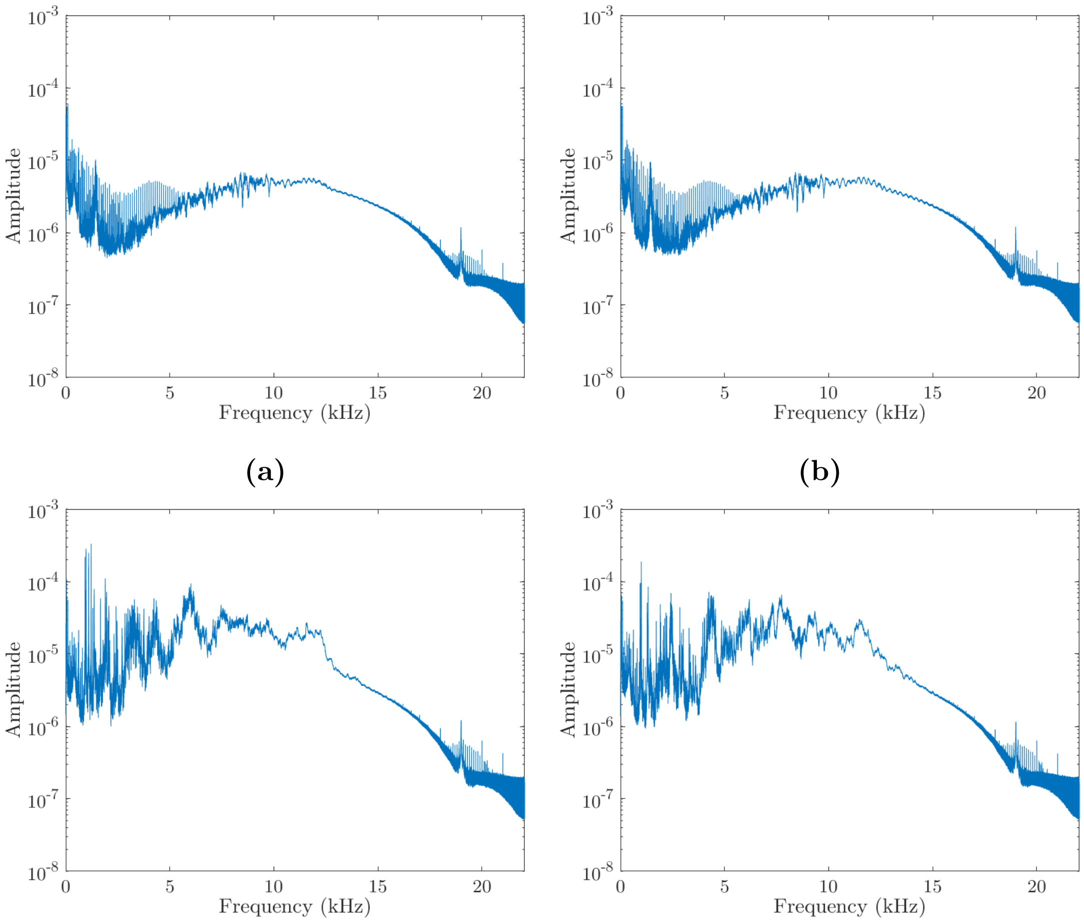


FIGURE 6.11: The measurement procedure where t is time, $v(t)$ is the LDV measurement and v_T is the threshold velocity, initially set to twice the root mean square of the first object measurement.

Figure 6.11 illustrates the measurement procedure from the moment the measurement beam has been directed to a new grid location on the target. First, the galvanometer gimbal control box emulates a Bluetooth keyboard connected to the integrated PSV-500 Xtra Data Management System and issues a `Ctrl+E` keystroke shortcut to have the scanning head complete “fast auto-focus” operation. Now that the laser beam is focused at the measurement location, the loudspeaker signal is outputted via the headphone jack, and a simultaneous recording is made from the SLDV via the microphone jack. The amplitude of this velocity signal, $v(t)$, is interrogated and should it exceed a predefined threshold, v_T, it is rejected on the basis that a high-amplitude inducing laser speckle noise “drop-out” likely occurred; otherwise, the measurement is saved. This threshold value is initially equal to twice the RMS of the first object measurement. In the event of a rejection, the laser beam position is adjusted by a 5 mrad scan angle, sequentially in $\pm x$ or $\pm y$, prior to the process repeating from the autofocus step until a satisfactory measurement is acquired. If the original measurement location and the four adjusted measurement locations fail to obtain a satisfactory measurement, it is reasonably concluded that the algorithm is incorrectly identifying measurements as containing speckle drop-outs. As such, the threshold is raised by 10% to prevent the algorithm from becoming stuck in an infinite loop, and then the process is repeated. Finally, once a satisfactory measurement is obtained for that measurement location, the measurement beam is directed to the grid location or the object is rotated to the next rotary stage angle, as shown in Figure 6.9 (a). It should be noted that this revised threshold value remains for all subsequent grid locations until another false speckle drop-out is detected, at which point it will be further increased. This allows the system to determine an appropriate threshold value for each object autonomously.

6.2.4 Data pre-processing

The 864 responses for each object were stored in the form of audio files (*.wav) for later processing. Figure 6.12 shows the mean spectra of all 864 responses for two of the tennis ball and two of the soda cans to aid in visualisation. It can be seen that most of the resonances occur below 15 kHz, and that while some objects like the tennis balls share many features, objects such as the soda cans do not as they are geometrically rather different to one another. It should also be noted that while the mean spectra are presented here, the measured responses for a given object varied depending on the exact nature of the measurement location; for example, a measurement on a modal node would vary from that to a measurement on a modal anti-node, similarly, a measurement normal to the vibrating surface would vary from that measured at a sharp angle of incidence. In order to classify these responses the audio files were processed into images, enabling the application of image processing CNN; the raw measurements have been made publicly available [101]. Since this pre-processing will affect classification accuracy, four similar but alternative spectrogram-based pre-processing techniques were applied so that they could be compared to one another. For each measurement, the first two data types were the spectrogram, showing the frequency content plotted against time; and the mel-spectrogram, instead showing the frequency content on a non-linear scale (the mel scale). The mel-spectrogram uses a non-linear frequency scale based on human auditory perception and is a common pre-processing technique for audio classification tasks [102, 103]. Since these types of spectrograms devote more image area to the lower frequencies, they should allow the NN to more effectively generalise as most of the object resonances are located in these lower frequencies, as shown in Figure 6.12. With a focus on reducing the impact of measurement noise on classification accuracy, an additional two data types were included where the spectrograms and mel-spectrograms were modified such that any signal content above the excited frequency at that instant is removed; this will be referred to as "cropping" throughout this paper. This results in a total of four data types: spectrograms, mel-spectrograms, cropped spectrograms and cropped mel-spectrograms, as can be seen in Figure 6.13. Each of these will be used to train a CNN so that the efficacy of these data pre-processing techniques can be compared.



(a) (b)

(c) (d)

FIGURE 6.12: Mean spectra (df = 1 Hz) of all 864 Hann windowed responses with subfigures (a) and (b) being those of the tennis balls (m) and (s). For comparison, subfigures (c) and (d) show those of the soda cans (r) and (s). Since there were recorded via the laptop microphone port, the amplitude has no units. The mean spectra of all 23 objects can be found in the publication [53].

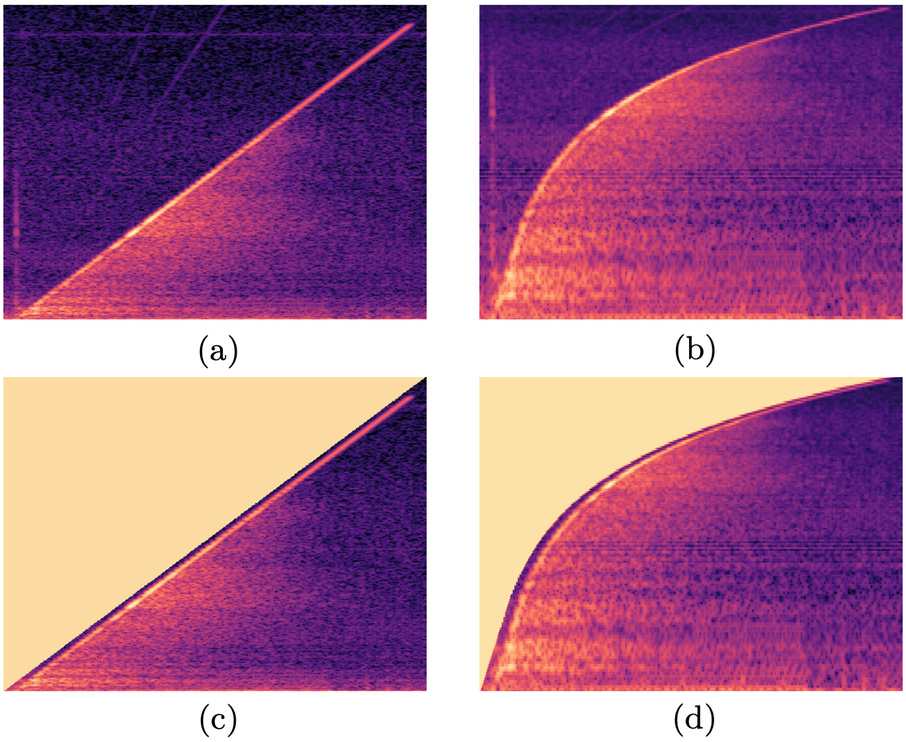


FIGURE 6.13: A response of object (j) shown as the four data types used to train the various CNNs; (a) spectrogram, (b) mel-spectrogram, (c) cropped spectrogram, and (d) cropped mel-spectrogram. The axes were not presented to the CNN, similarly, the colours represent normalised amplitude.

The pre-processing was implemented in Python using the Librosa library [104]. Each response was first used to generate a normalised spectrogram with a window size of 1024 samples, or 23.22 ms, which gives a spectral resolution of 0.043 Hz. The windows had an overlap of 6.25% and a maximum frequency bin of 22.05 kHz. The spectrograms were then "cropped" between opposing corners, shown in Figure 6.13, removing any spectral information above the excitation frequency by setting these areas of the spectrogram to zero amplitude. Finally, the mel-spectrograms and the cropped mel-spectrograms were generated using their spectrogram counterparts with 150 bin frequencies. Convolutional neural networks (CNNs) were then trained and assessed on each of these four data types, which were in the form of 496 x 369 pixel images (*.png) as exported by Librosa.

6.3 Convolutional neural network training and regularisation

6.3.1 Summary of fundamental neural network concepts

CNNs are a variety of neural network (NN) specialised in processing data with grid-like topologies, therefore, excelling at image recognition tasks [105]. Since audio data can be readily represented as an image using spectrogram techniques, which maintain this grid-like topology, CNNs have been commonly employed for audio classification tasks [103]. The name and fundamental structure of the NN is inspired by that of a biological brain, as it is intended to mimic neurons in a brain signalling to one another [106]. A general schematic of basic NN can be seen in Figure 6.14. The input nodes are those which receive information, such as the brightness of a pixel[2], and pass this information on via a connection to a node in the following layer. Each hidden layer node uses an internal bias and the assigned weights of each connection to calculate an output value[3] [107]. This process continues across each layer of the NN until the output layer, where the result is acquired; for example, the classification of the image.

---

[2] For a monochromatic image; otherwise, it is the brightness of a channel within a pixel.

[3] Activation functions are not mentioned here for simplicity.

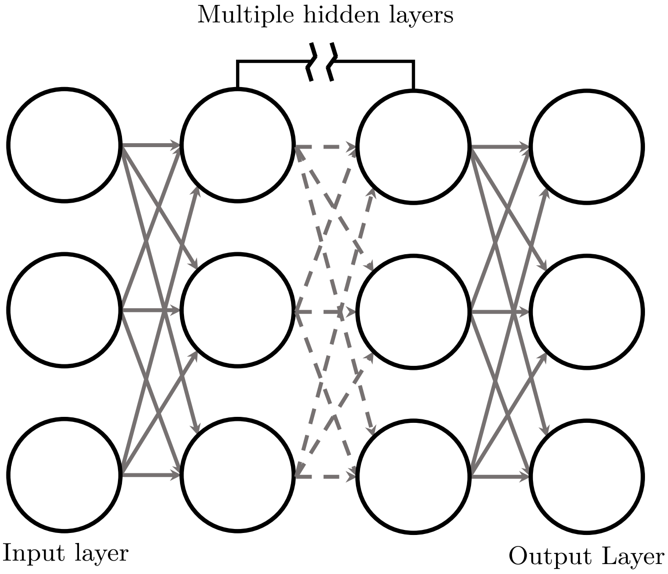


Figure 6.14: A general schematic of a basic NN, with circles representing nodes and arrows representing how information is passed between them.

It is the weights and biases that determine the specific output of a NN, given a specific input, and they are assigned through a process known as *training*. In the case of supervised learning used herein, training is performed using a labelled data set, known as a *training set*. The training set is used during training to tune the weights and biases such that the NN outputs best agrees with the labels of the training set. The tuning of the weights and biases is regulated by various parameters known as hyperparameters [107]. The correct selection of the hyperparameters is a vital part of training any NN, as these influence the ultimate performance. One such hyperparameter is the number of epochs a NN is trained for, with an epoch being one iteration of the training data by the NN - since training involves making repeated iterations over the training data to arrive at the optimal weights and biases. A common obstacle when training a NN is obtaining high performance on the training data but lower performance on a new unseen data set; this is known as overfitting and should be avoided.

6.3.2 Regularisation to prevent overfitting

There are a number of techniques intended to prevent overfitting and are known as *regularisation* techniques, but the most common of them is known as *early stopping* [108]. Early stopping uses a second data set, known as the *validation set*; the NN is not trained on this data set, but it does make predictions on it. These predictions are used to determine if the NN has successfully generalised such that the predictions also apply to the validation set, or if the NN is can only make accurate predictions on the training set. Practically, this means training for a number of epochs which minimise *both* training loss and validation loss, where loss is a metric that captures the performance of the NN when making predictions on either data set.

Typically, a dataset will be split into 80% training set and 20% validation set. However, sometimes split into 70% training set, 20% validation set and 10% *test set* when particular rigour is required [107]. Despite the validation set only being used to calculate the performance, not tune the weights and biases, sometimes, through the selection of specific hyperparameters, slight overfitting can still occur. Therefore, to obtain the most representative real-world estimate of the performance, the test set is set aside in the early stages and is not used at all for training but at the end of the work to gauge the performance.

6.3.3 Training methodology for response classification

Rather than training the CNN from scratch, transfer learning (TL) was applied. TL takes a NN which has already been trained for a related task and retrains it for the task at hand, thus "re-purposing" the NN. In this approach, only the weights and biases near the output layer are tuned, with the output layer being modified for the new classes; therefore, the network retains knowledge learned from the initial data set. TL has three main general advantages when compared to training a new network [109]. Firstly, the accuracy after one epoch of training will be higher. Secondly, the rate of improvement with increasing epochs will be steeper. Finally, the performance plateau at the higher epochs will still remain higher than for traditional techniques. These advantages are due to the increased contextual knowledge available to the NN from the pre-training data set, which is being applied to the new data set.

The pre-trained CNN used in this work is ResNet-18 [110] which has already been trained on the ImageNet database [111]. It is a relatively lean CNN at only 18 layers deep, meaning the training and subsequent inference will require less time and processing power than its larger counterparts. This was performed on the Google Colab cloud platform in Python using the Fastai library [112]. A Google Colab Pro subscription allowed access to either NVIDIA V100 or A100 graphical processor units, depending on availability. All of the CNNs described in this chapter were trained using a batch size of 16 for 14 epochs using the `fit.one_cycle` function. This training methodology is an implementation of cyclical learning rate [113, 114] and super-convergence [115] principles published in the literature. The practical implications of this technique allow for the hyperparameters of the learning rate, momentum, and weight decay to be automatically determined, resulting in CNNs which can outperform those trained with traditional hyperparameter tuning techniques for some applications [115]. In order to measure the performance of the various CNNs the classification accuracy is calculated on the test set, while in many applications this would not be appropriate, here it is valid since the data is *balanced*, meaning each class contains the same number of responses [116]. During the analysis, the data can be split in such a way that some classes contain one more or less response than other classes. For example, the 10% test set split of all 19,872 responses results in 1987.2, as such, the resulting test set contains 1987 responses meaning one class has one less response. This is a common occurrence that introduces a negligible bias into this accuracy-based performance assessment, whereas, for larger class imbalances - roughly ranging up from a two times difference - other performance metrics become more reliable, such as the F1-score [117].

6.4 Pre-processing and dataset size optimisation

This section explores the impact of data pre-processing and total training set size on the performance of the various CNNs. Here, the 864 audio files for each of the 23 objects in Figure 6.8 were used to generate 864 of each of the four data types shown in Figure 6.13. This results in a total of 19,872 images for each of the four pre-processing techniques. Of these, 10% is set aside as the test set used to determine the accuracy of each CNN in the final stages. This test set is comprised of the same 1988 responses for each technique, only with a different pre-processing technique applied. As discussed in the previous section, 10% of 19,872 results in there being either 86 or 87 responses for each object class in the test set. The remaining 17,884 images are taken to be the training set. Finally, to simulate smaller data sets, the number of responses is reduced by a factor of $m = 1, 2, 4, 6, 8, 10$. This results in six training set sizes and the four pre-processing techniques, giving a total of 24 data sets used for this section.

The test set was comprised of a randomly selected stratified sample and was set aside for each pre-processing technique to evaluate the performance of the CNNs later. Five-fold cross-validation was then used to rigorously compare the effects of the four pre-processing techniques and data set size on the CNN performance. This divides the remaining data into five “folds”; four of these are used as the training set, with the remaining fold used as the validation set. This process repeats five times, resulting in five separate CNNs for each iteration, meaning that all the data not in the test set is used for both training and validation, generating a total of 120 CNNs for the 24 data sets. For a fair comparison, all CNNs for all values of m were compared against the identical test sets for each pre-processing type. Meaning all the metrics presented in this section were generated using the same test set just with differing pre-processing applied for the four techniques. This procedure is depicted in Figure 6.15.

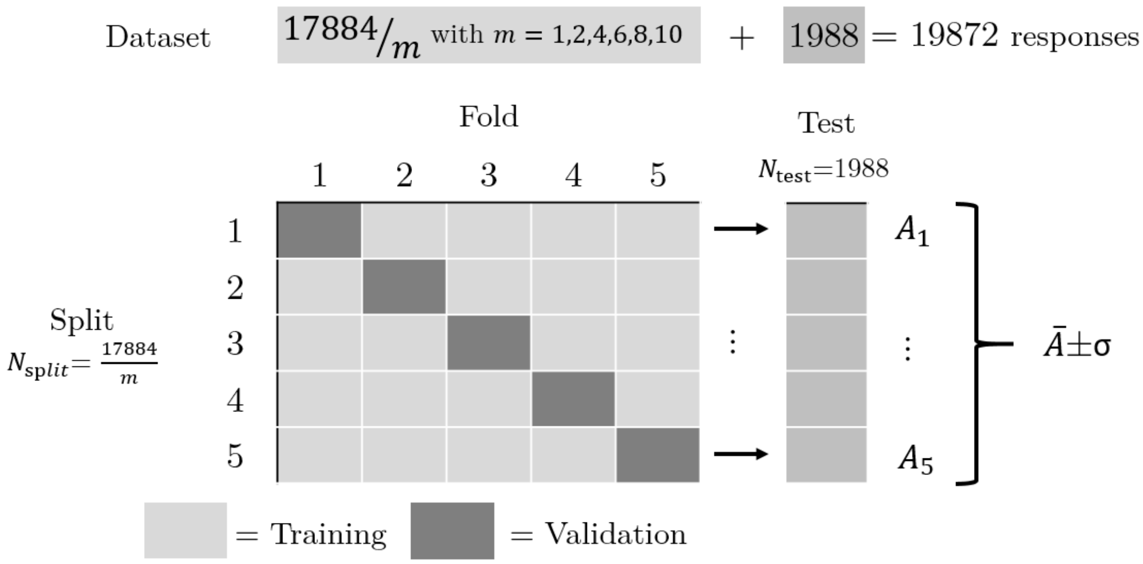


Figure 6.15: Illustration of the five-fold cross-validation procedure used to compare the four pre-processing techniques at six data set sizes. Where m is the fraction of the data set used in the five-fold cross-validation, A is the accuracy obtained for each split, $\bar{A}$ is the mean accuracy of all five splits, σ is the associated standard deviation of the mean accuracy, N_{split} is the total size of that split of the training data set, and N_{test} is the test set size.

The results of the five-fold cross-validation procedure for the 24 data sets are shown in Figure 6.16. Overall, all four pre-processing techniques yield higher accuracies for larger data set sizes, with all four techniques tending towards 100%. However, for smaller data set sizes, the performances of the four techniques begin to diverge. The worst-performing CNNs used spectrograms, reaching as low as 72.16% for the data set containing 78 responses used per object. The second worst technique was the cropped spectrogram, however, its performance was fairly similar to mel-spectrograms, exhibiting a maximum performance difference across all data set sizes of about 2%. Finally, the CNNs which utilised cropped mel-spectrograms performed the best of the four techniques and exhibited considerably lower sensitivity to data set size than the other techniques; while also being the most accurate when the entire data set was available. The cropped mel-spectrograms obtained an accuracy of 87.74% while only being trained on 78 responses per object, which is potentially sufficiently small to make manual data acquisition a viable alternative to the bespoke automated arrangement used here.

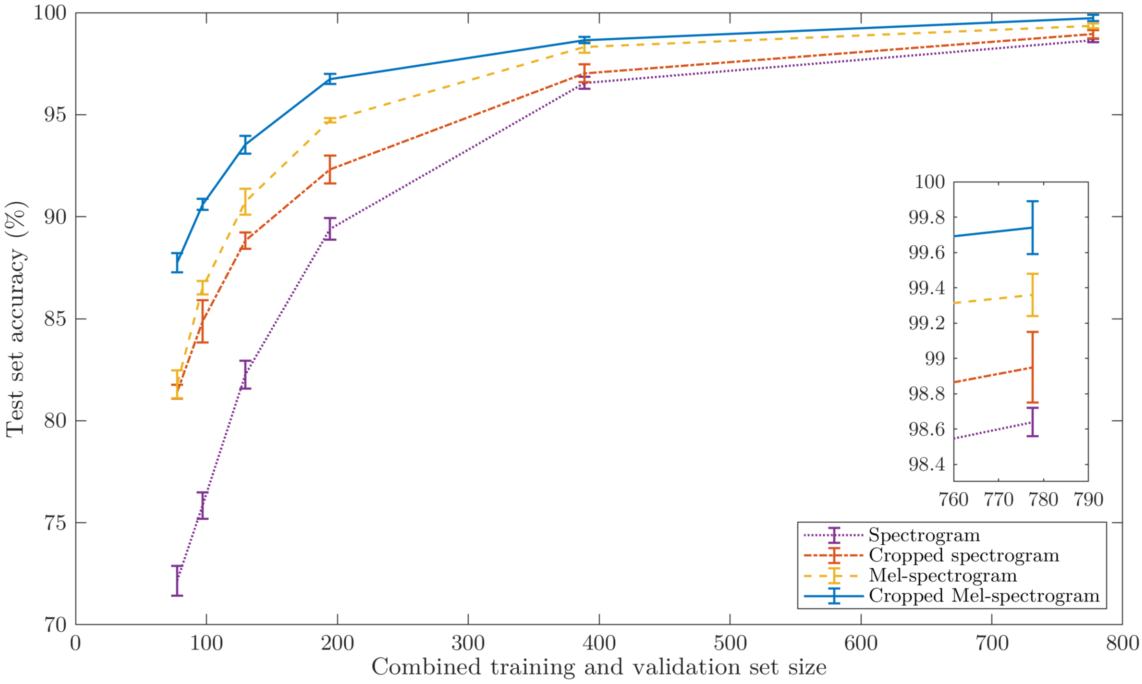
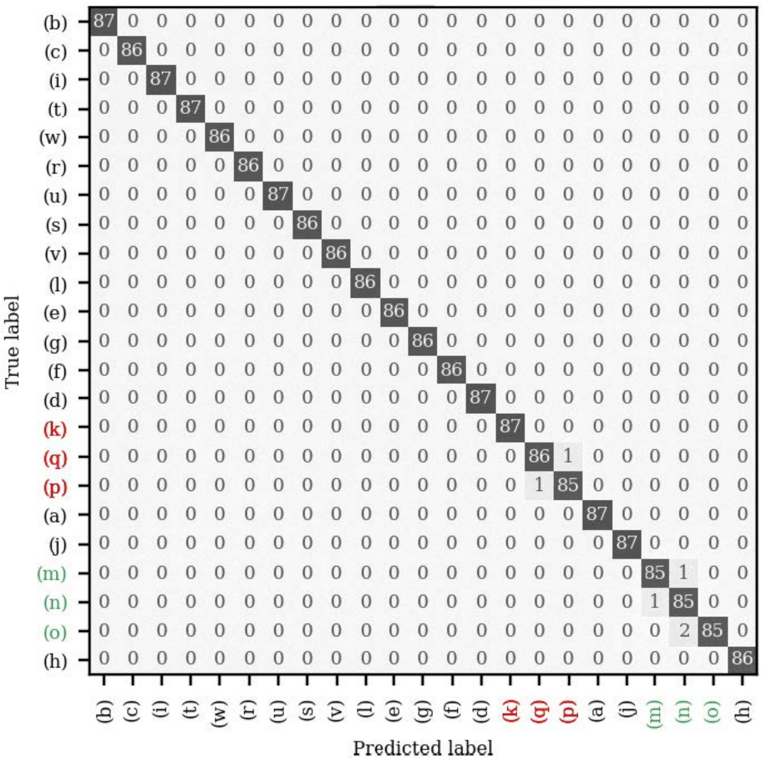


FIGURE 6.16: The accuracies and the associated standard deviations for each set of five CNNs in the five-fold cross-validation on the 24 data sets - four preprocessing types at 6 data set sizes. Overlaid on the right is an expanded plot representing four data points at the full data set size.

The results at the largest sample size can be seen more closely in the expanded plot in Figure 6.16. This shows that the accuracy of all four techniques was above 98.64% with non-overlapping associate uncertainties. At this largest data set size, the spectrogram CNNs performed the worst, with cropping leading to a performance boost of over a percent. A similar effect can be observed with mel-spectrogram CNNs, where cropping leads to a performance increase of about half a percent. Overall, cropping seems to have the expected outcome on the performance as the CNNs are able to generalise more effectively if the image area known to be devoid of any features is removed. Similarly, mel-spectrograms had a similarly anticipated advantage over spectrograms, likely because where mel-spectrograms devote more image area to the frequencies with higher densities of resonances, therefore, allowing the CNNs to generalise more effectively. Consequently, the remainder of this chapter will focus on the use of cropped mel-spectrograms as the primary data type in combination with the entire data set, as the classification accuracy using these exceeds that of the others tested here.

Figure 6.17 shows a confusion matrix for one of the five cropped mel-spectrogram CNNs using the entire data set. For context, a confusion matrix of a classifier with 100% accuracy would only have entries on the diagonal from the top left to the bottom right-hand corners of the figure since these locations correspond to the predicted label being the same as the true label. It can be seen that the CNN correctly classified the vast majority of the responses, however, there were to two main regions where errors were made. The CNN misclassified the tennis balls (m), (n) and (o) for one another four times and the table tennis balls (p) and (q) for one another once. These misclassifications might be resolved by using a larger data set, by extrapolating the observed trend in Figure 6.16; using a deeper CNN which is sensitive to more complex features, such as ResNet-34; and by adjusting the parameters in the spectrograms to resolve more spectral features.

Despite these potential improvements to the technique, the performance of this technique rivals that of the existing *contact* object recognition techniques discussed in Section 6.1. While the direct comparison of existing work to this work is difficult due to differences in the number of objects (ranging from 12 to 50), with some using filled containers [48, 50] rather than solid objects [49], or both [47]. However, these techniques can be summarised as performing within the range of 85.5% [50] up to 98.2% [49] - with the best performing work utilising similar objects to those used in this work. However, as was discussed, the best performing of the existing work utilised a combination of actuator torques and the sounds emitted by a combination of five different exploratory behaviours, whereas the technique presented in this chapter can identify an object within seconds using only an acoustic excitation. This contextualises the technique presented in this chapter as a viable alternative since it not only increases recognition accuracy but also introduces rapid non-contact functionality.



	(b)	(c)	(i)	(t)	(w)	(r)	(u)	(s)	(v)	(l)	(e)	(g)	(f)	(d)	(k)	(q)	(p)	(a)	(j)	(m)	(n)	(o)	(h)
(b)	87	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(c)	0	86	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(i)	0	0	87	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(t)	0	0	0	87	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(w)	0	0	0	0	86	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(r)	0	0	0	0	0	86	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(u)	0	0	0	0	0	0	87	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(s)	0	0	0	0	0	0	0	86	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(v)	0	0	0	0	0	0	0	0	86	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(l)	0	0	0	0	0	0	0	0	0	86	0	0	0	0	0	0	0	0	0	0	0	0	0
(e)	0	0	0	0	0	0	0	0	0	0	86	0	0	0	0	0	0	0	0	0	0	0	0
(g)	0	0	0	0	0	0	0	0	0	0	0	86	0	0	0	0	0	0	0	0	0	0	0
(f)	0	0	0	0	0	0	0	0	0	0	0	0	86	0	0	0	0	0	0	0	0	0	0
(d)	0	0	0	0	0	0	0	0	0	0	0	0	0	87	0	0	0	0	0	0	0	0	0
(k)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	87	0	0	0	0	0	0	0	0
(q)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	86	1	0	0	0	0	0	0
(p)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	85	0	0	0	0	0	0
(a)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	87	0	0	0	0	0
(j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	87	0	0	0	0
(m)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	85	1	0	0
(n)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	85	0	0
(o)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	85	0
(h)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	86

Figure 6.17: The confusion matrix of test set inferences of one of the five cropped mel-spectrogram CNNs. Highlighted for convenience are the two groups of objects where misclassification occurred: in green - (m), (n) and (o) - are the tennis balls, and in red - (k), (p) and (q) - are the table tennis balls.

6.5 Broader object class performance

While it was demonstrated in the previous section that it is possible to identify objects solely by their vibrational response to an acoustic excitation, it has not yet been determined whether this technique can be used to recognise broader classes of similar objects; for example, recognise any soda can rather than a specific instance of a soda can. The few misclassifications seen in Figure 6.17 are a subtle indicator that vibrational characteristics of these groups of similar objects may share some features. As such, this section will focus on the system's ability to classify objects into broader object classes.

For this to be possible, the CNN must generalise features related to all objects within a broader class and be able to identify those features in the instance of the object not contained within the training data. As such, this section utilises the grouped objects: tennis balls - objects (m), (n) and (o); table tennis balls - objects (k), (p) and (q); the sealed soda cans - objects (u), (v), (w); and empty soda cans - objects (r), (s), (t). The CNN was trained on two of the three objects within each group, but labelled as the same class, along with all of the previous objects, thereby minimising the likelihood that the CNN will correctly predict the class due to chance. This means that the accuracy of the CNN can then be obtained using the third, unseen objects within each group. Therefore showing that the CNN successfully generalised features common to all three objects within each group, despite having seen only two of them.

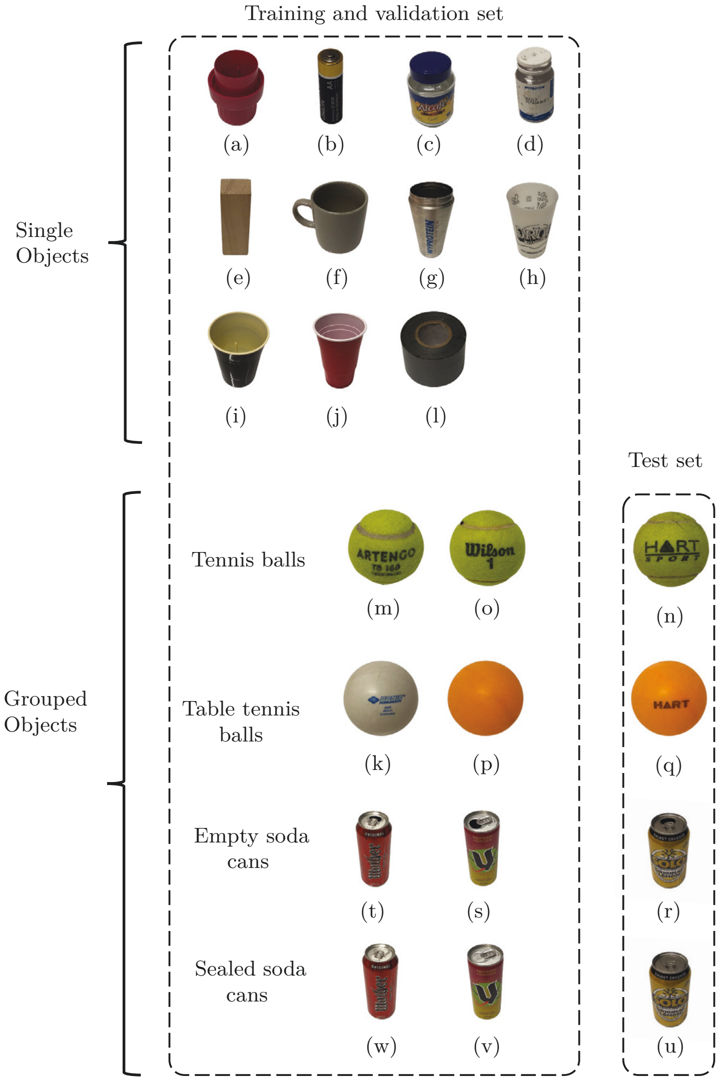
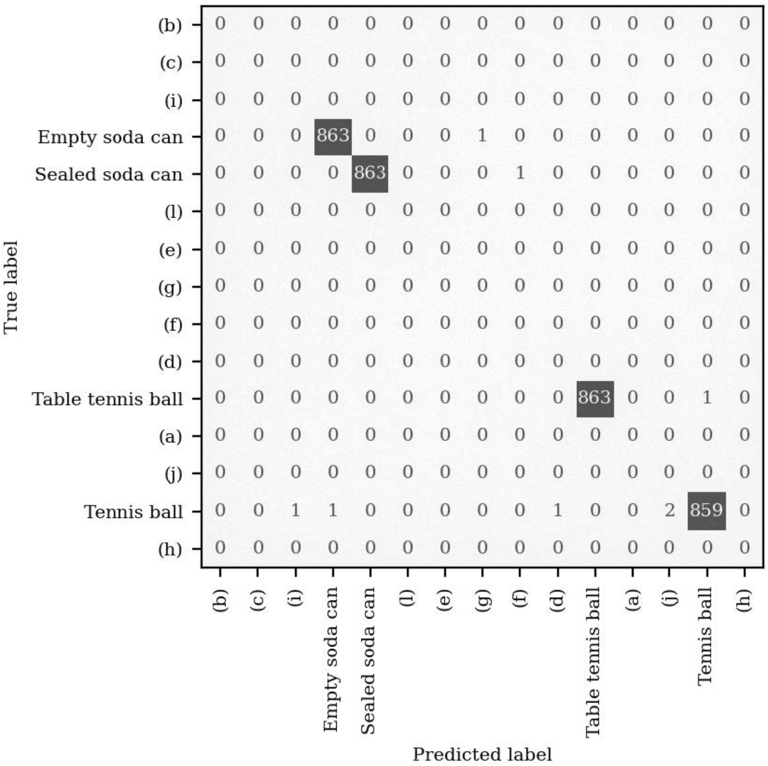


Figure 6.18: The labelling of the 23 objects organised as single objects or grouped objects - table tennis balls, tennis balls, and either full or empty soda cans. Each object in this represents figure represents the full 864 mel-spectrograms.

The data for this section were labelled according to Figure 6.18, implemented by simply modifying the image file names from which the classes were set to be assigned from. Here, ResNet-18 was again trained in line with the procedure previously described using 864 cropped mel-spectrograms for each unique object, along with a further 1728 mel-spectrograms for each pair belonging to each group of objects (since one object is being held out). The data contained within the dashed box in the figure titled “training and validation set” was split into the standard 80/20 validation set test set split for the purposes of training. However, since the performance metric of interest here is how accurately the CNN can classify the never-before-seen objects belonging to these four groups, the test set consisted only of the four hold-out objects, with 864 cropped mel-spectrograms each.

Figure 6.19 shows the confusion matrix for the hold-out objects with near-perfect performance. This represents an accuracy of 99.83% at predicting the class of a never-before-seen object out of a potential 15 options; leaving the likelihood that chance is a substantial contributing factor very low. The number of these misclassifications would likely decrease if the training set contained more than just two instances of an object. As discussed in Section 6.1, one of the cited works similarly looks at the classification of objects belonging to broader object classes, rather than classifying individual objects [47]. That work included three object classes - filled bottles of water, pieces of paper and rigid objects - which were excited using a simple shake actuator. Their results showed that a never-before-seen object belonging to the broader object classes could be recognised with an accuracy of up to 95.8%. However, with such a small number of potential output classes in the data, there is a 33.3% chance that the classifier can obtain the correct output due to chance alone. Similarly, the chosen objects do not represent the most challenging selection, with rigid objects simply emitting no sound when shaken. Similarly, it can be noted that all the aforementioned contact techniques utilise high amplitude induced vibrations within the object by contacting the object, leading to audible non-linearities, like sloshing. In contrast, this technique excites the response using only an external acoustic field, yet, this alternative approach presented in this chapter outperforms the existing technique, obtaining near-perfect classification despite a total of 15 potential class outputs whilst also introducing non-contact functionality.



	(b)	(c)	(i)	Empty soda can	Sealed soda can	(l)	(e)	(g)	(f)	(d)	Table tennis ball	(a)	(j)	Tennis ball	(h)
(b)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(c)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Empty soda can	0	0	0	863	0	0	0	1	0	0	0	0	0	0	0
Sealed soda can	0	0	0	0	863	0	0	0	1	0	0	0	0	0	0
(l)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(e)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(g)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(f)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(d)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Table tennis ball	0	0	0	0	0	0	0	0	0	0	863	0	0	1	0
(a)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(j)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tennis ball	0	0	1	1	0	0	0	0	0	1	0	0	2	859	0
(h)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

FIGURE 6.19: The confusion matrix of the test set used in assessing the ability of the technique to detect broader classes of objects.

6.6 Chapter summary and discussion

This chapter introduced a new non-contact vibro-acoustic object recognition technique that can rapidly recognise previously characterised objects within seconds. This is achieved by exciting a small vibrational response using a loudspeaker which is measured using an LDV and classified using a convolutional neural network trained via transfer learning. The technique was developed and verified using 23 household objects, however, it is not only limited to household objects. A bespoke, automated vibro-acoustic response measurement system was developed specifically to enable the rapid collection of quality raw time data. These were then pre-processed into images for use with the CNN as four data types: spectrograms and mel-spectrograms, with and without cropping the image above the excitation frequency.

The chapter first looked at how the pre-processing approach and training set size may affect the accuracy of the object classification predictions. From this, three main observations can be made. Firstly, all the CNN's accuracies increased for larger training set sizes; which is a fairly typical outcome. Secondly, the CNNs utilising mel-spectrogram inputs performed better than those utilising spectrograms. This is likely due to the higher density of resonances at the lower frequencies, which the mel-scale devotes more space to, therefore enabling more effective classification and generalisation. Lastly, removing any spectral information in the spectrogram or mel-spectrogram above the instantaneous excitation frequency (cropping) increased the accuracy. This is likely caused by cropped regions containing mostly noise, therefore inhibiting the CNN's ability to generalise effectively. While all the pre-processing techniques lead to a sufficiently high accuracy when the training set size was at its largest, the use of cropped mel-spectrograms outperformed the others with an accuracy of 99.74% ± 0.15% on the test set. As such, only the full mel-spectrogram data set was used to train CNNs in the subsequent section.

While the aforementioned performance is near-perfect, a practically viable object recognition technique must be able to detect broader classes of objects, not just specific ones; for example, all soda cans, not just one specific soda can. To test this, four triplets of similar objects were grouped together so that the CNN could generalise across them. When the CNNs predicted the class of the never-before-seen third item in the groups, an accuracy of 99.83% was obtained, confirming that this approach constitutes a practically viable object detection technique. It is important to note that, when generalising to unseen objects, the broader object class must contain some vibro-acoustic similarities, however, it can be reasonably expected that a sufficiently large and diverse training set should still yield performance comparable to that described here.

While the paper has focused on object recognition, it has been shown that the solution is highly sensitive to slight differences in the objects while also capable of learning broader object classes. These features open up many potential applications within various other fields. For example, this technique could be used for defect detection in production lines, where physical changes in an item will change its acoustic fingerprint. This is useful when all that is of concern is that a defect is present in an item, not what exactly the defect is, similar to the roles checksums play in computing. Another potential application is the identification of loose rock in underground mines. Current techniques depend on a worker tapping the roof with a bar and listening to the sound to determine if it is loose, known as "roof sounding" [118]. A more high-tech solution employs the use of a vibration sensor which is similarly tapped on the roof to identify loose rock [119]. However, both require workers enter the newly created area. Therefore, this technique could scan the roof in an entirely non-contact manner prior to entering the area, reducing the risk to human life.

Despite the technique's promising performance, there remain some aspects of the system which would require further investigation and refinement. Firstly, the loud audible chirp used to elicit a response from the object makes the system unpleasant to nearby people. Therefore, it is important to use non-audible frequencies to excite the target. Secondly, the energy of the speaker should be directed in the approximate direction of the target rather than being lost to the surroundings to increase the functional range of this technique. Therefore, integrating an ultrasonic parametric speaker to perform the excitation is an ideal next step for the system. Following this, an in-depth investigation into the effects of object distance can be conducted. Thirdly, the automatic acquisition procedure could be modified to accommodate larger objects with more complex geometries which may not fit onto a rotary stage.

For practical deployment onto autonomous systems, this system would need to be used in conjunction with a complementary sensor able to obtain the relative location of the object of interest. Light Distance and Ranging (LiDAR) sensors are almost ubiquitous on autonomous robots as they can generate dense 3D maps of the surrounding environment for navigation and mapping. As such, this technique could naturally be extended to become compatible with this data type. For example, an unknown object-of-interest may be identified within the LiDAR scan, following which the SLDV can be directed towards it for further vibro-acoustic interrogation utilising the techniques described in this chapter. A proof-of-concept system that fuses point cloud and LDV data is described in Appendix A.

Chapter 7

Conclusions and future work

The laser Doppler vibrometer has become an indispensable tool in vibration engineering, boasting high bandwidths and spatial resolutions unrivalled by traditional contacting accelerometers - all whilst doing so via non-contact means. Since their widespread adoption, their application envelope as vibration transducers has been steadily expanding, encompassing increasingly impactful areas. However, their deployment on mobile platforms has been largely hampered by the motion the instrument might experience in this scenario. Despite the associated challenges, the allure of mobile LDV deployment remains since it can substantially increase land coverage rates whilst simultaneously enabling access to hazardous or remote areas. This field has already received some interest, with research into mobile buried landmine detection, intelligence gathering from drones and structural health monitoring from drones all having received attention. Therefore, this thesis first aimed to develop the necessary infrastructure needed to enable the mobile deployment of LDVs and then explored the potential for their integration within autonomous systems.

In the reference frame of the LDV sensor head, the vehicle motion can be divided into two broad categories: translational motion along the beam axis *or* translational motion in the two orthogonal axes, plus the three rotational degrees of freedom. Since the underlying physics of the LDV dictate that the sensor head is as sensitive to self-vibration in the beam axis as it is to the target vibration in the beam axis, sensor head vibration could compromise the measurement. Whereas motion in the five non-beam DOFs would cause the measurement beam to stray from the intended measurement location and introduce speckle noise. Two separate approaches can mitigate the effects of these two distinct phenomena, referred to herein as *measurement correction* and *arbitrary tracking.*

The best of the pre-existing measurement correction techniques conveniently utilises a single contacting accelerometer in order to measure and subsequently subtract any velocity contribution due to instrument motion. However, improper accelerometer handling suggested that performance gains could be made and the frequency domain-based processing is limited to stationary signal types. As such, Chapter 3 pioneered the development of a new time domain signal processing technique - thereby introducing important non-stationary or transient vibration signal capability - while Chapter 4 delivered a significantly improved frequency domain-based technique which outperformed the original technique by up to eight times. An analytical model was then developed, enabling a rigorous sensitivity analysis of the effects of signal synchronisation. These two techniques were then extended in Chapter 5 to encompass similar existing work which showed that, by using three precisely positioned orthogonal accelerometers, similar measurement correction performance can be achieved under scanning conditions, but with improved signal processing. In all likelihood, the improved signal processing techniques presented herein, based on the existing hardware arrangement, should become the new gold standard for measurement correction signal processing.

With the effects of beam axis motion substantially mitigated and compatible with scanning setups, Chapter 5 also presents a novel instrument motion-tracking solution, intended for downward-facing drone applications. With existing image-based tracking solutions requiring additional onboard computational resources, the resulting weight can make the technique unsuitable for drone deployments. As such, the proposed solution utilises an existing SLDV system, and unlike the image-based alternative, this does not require any modification to the system. This lightweight system functions much in the same way as a gimbal, but counter-rotates the beam orientation rather than the entire instrument, using information obtained from an onboard inertial measurement unit. Testing this on a

Multi-Axial Simulation Table showed the system achieves a 68% reduction in beam motion during what was equivalent to extreme flight conditions.

Finally, with significant steps forward achieved in both the development of an LDV insensitive to self-vibration, and the introduction of a new tracking technique, the focus of the thesis shifts towards potential in-field mobile applications. Chapter 6 explores the common autonomous vehicle task of object recognition, developing a new non-contact vibro-acoustic object recognition technique. This technique involves the remote acoustic excitation of an object such that it can be detected by an LDV. The measured responses were then processed using a range of spectrogram-based techniques and classified using a convolutional neural network trained via transfer learning. Results show that objects may be identified with an accuracy of over 99.8%. Potential further applications include loose rock detection in mines, rapid defect detection on production lines and enhanced clutter rejection for LDV-based buried landmine detection.

7.1 Future work

The galvanometer gimbal system could be improved in a number of ways. Firstly, the integration of an inertial measurement unit with a higher output rate would increase the responsiveness of the system, reducing the introduced error at any instant. Secondly, the code should be extended such that it may be compatible with any mounting angle. The mirrors on an SLDV correspond to pitch and yaw, which is straightforwardly tracked during a downward-facing application as these translate readily to the pitch and roll of the drone. However, if the system was forward facing, the two orthogonal mirror angles would need to work in tandem to recenter the target location; these angles can be readily calculated using a rotation matrix.

The SLDV measurement correction work should be combined with the galvanometer gimbal work to show for the first time a viable tracking *and* measurement correction technique which when used in tandem can mitigate the effects of speckle noise, aiming errors and vehicle vibration. In order to confirm the true efficacy of this combined technique the two sub-components should be integrated such that the angles of the two orthogonal mirrors are simultaneously logged, enabling the appropriate measurement correction components to be calculated for each time step. This naturally lends itself towards the time domain technique, and since it has both stationary and transient signal compatibility, it will be applied. This combined system should be experimentally validated on the Multi-Axial Simulation Table and include a reference measurement as has been done in Chapter 3 and Chapter 4 such that the performance can be quantified with and without the combined motion tracking and triaxial measurement correction. If shown to be successful, a real-world measurement campaign could be conducted which integrates all the measurement correction and tracking knowledge available herein. For example, drone-based structural health monitoring or buried landmine detection.

Finally, Appendix A lays the groundwork by describing a proof-of-concept system that fuses LDV and point cloud data to generate enhanced vibro-acoustic spatially associated maps for autonomous systems. Here, the LDV data is fused with that of a LiDAR unit to present a series of interesting scenes. This capability could eventually be combined with that described in Chapter 6 to allow the autonomous system to glean additional information about objects within a scene based on their vibration profiles while spatially associating this to their specific locations within a scene. While not realised within this thesis, the information gleaned from this when combined with that of the aforementioned object recognition work could allow autonomous systems to carry out enhanced interrogations of specific objects if, for example, the object is not recognised visually. Similarly, the vibrational properties of the object in question could be used in industrial settings to infer the operational state or health of machinery, being regularly autonomously monitored.

Appendix A

Towards enhanced perception for autonomous systems

This Appendix presents the practical extension of the work described in Chapter 6. Here, an additional mapping modality is used in conjunction with the SLDV, allowing an autonomous system to locate objects of interest for future vibro-acoustic interrogation using an LDV.

A.1 Introduction

Chapter 6 describes a non-contact vibro-acoustic object recognition technique, however, practical deployment of the system in its current form is limited as it requires the object of interest to be in the same location each time. An additional sensor type can be used in conjunction with the SLDV to locate objects of interest within the environment so that they may be vibro-acoustically interrogated. Light Distance and Ranging (LiDAR) is an active range-resolving optical remote measurement technique. It has become almost ubiquitous for autonomous systems as it can rapidly generate accurate spatial maps of the surrounding environment for robot perception and navigation [120]. As such, this appendix describes the groundwork required to integrate the non-contact vibro-acoustic object recognition technique presented in Chapter 6 to a practical object recognition technique for autonomous systems.

A.2 Point clouds and LiDAR fundamentals

A point cloud can be understood as a spatial data type consisting of coordinates, most commonly polar or Cartesian coordinates. These coordinates generally represent positions of ‘hits’ in an environment generated by some directional distance sensor. The means by which a point cloud can be generated varies widely. Some mainstream examples include structured light scanners [121], stereoscopic cameras [122], and LiDAR scanners [123]. While all these techniques have their advantages and disadvantages, the most common of them in the context of autonomous vehicles is arguably the LiDAR scanner, as such, this work will focus on its integration.

LiDAR scanners can take various forms, but at their heart is a time-of-flight (TOF) sensor, also known as a single point or 1D LiDAR. These TOF sensors consist of a transmitter and a receiver. The transmitter is typically a laser and emits a short pulse on the order of a few hundred nanoseconds [124]. This light is subsequently backscattered from an incident surface in a similar fashion to the light from an LDV. As the sensor collects this returning light using the receiver optics, the time between emission and detection is used to calculate the distance to the incident surface, d, given by:

$$d = \frac{c\Delta t}{2} \tag{A.1}$$

where c is the speed of light and Δt is the round-trip journey time of returning pulse [125]. Therefore, by scanning this TOF sensor, or multiple TOF sensors, across an environment, the surrounding geometry can be mapped. This scanning is most frequently implemented using an opto-mechanical system, where an actuated mirror surface can direct the measurement beam [126]. This has been previously implemented by using the same orthogonal galvanometric scanning mirror setup employed on SLDVs [127], however, the most common types of scanners are those with continuously rotating reflectors driven by motors and positional feedback [126]. In these setups, the scanner simultaneously measures the distance to the incident surface and the current angular position of the TOF sensor using feedback sensors [128].

Tangential to the LiDAR scanner is the TOF camera, it is a similar device that removes the need to “scan” in order to generate 3D spatial data. This technique exposes the entire scene to an infrared light pulse, then collects the returning light using a camera sensitive to that wavelength [129]. By timing the time taken for the light to return to each individual pixel on the imaging sensor, a depth map of the scene can be constructed simultaneously for the entire FOV [130]. However, these cameras generally have a lower range, spatial resolution and adverse random noise behaviour when compared to LiDAR scanners [131].

A.3 Existing time-of-flight sensor usage with laser Doppler vibrometry

Polytec first integrated a TOF sensor with an LDV on the PSV-400 SLDV as an optional addition [132], which has since become an integrated feature on the PSV-500 SLDV [74]. Referred to as the Geometry Scan Unit, the inline TOF sensor enables a sample geometry measurement by utilising the steering mirrors to scan across the sample surface. Therefore, providing a static baseline for observing the dynamic vibration of the sample. Similarly, a TOF camera has been combined with an SLDV to automate the selection of measurement points on a sample which would otherwise be a lengthy manual procedure [133]. This technique relies on an existing CAD model of the test sample, then using Viewpoint Feature Histograms and Iterative Closest Point algorithms, the sample geometry and pose is automatically determined from the TOF camera data. This removes the lengthy manual point selection phase in existing approaches and automatically obtains measurements at pre-defined measurement locations on the sample surface.

While these existing TOF-LDV techniques are important within the field of vibration engineering, they bear less significance within robotics. Therefore, this section aims to build on the work described in Chapter 6 by describing how the LDV can be integrated into the autonomous system workflow rather than integrating LiDAR into the vibration engineering workflow. As such, the following sections will discuss how vibro-acoustic interrogation can be used within the context of enhanced perception for autonomous systems.

A.4 Hardware arrangement

A bespoke 3D printed adapter plate was mounted to the face of the SLDV using some existing mounting holes; this allowed the LiDAR scanner to be mounted above the SLDV sensor head, as can be seen pictured in Figure A.1. The SLDV used for this work is an infrared-based Polytech PSV-500 Xtra, as it yields a higher return signal from untreated surfaces than the more common visible light variants. Whereas the point cloud was obtained using a LiVOX Mid-40 LiDAR scanner as it can generate dense point clouds over a 40° cone using a non-repeating scan pattern; therefore, the longer the scan duration, the denser the point cloud. Both of these instruments are capable of obtaining dense scans over a similar FOV from under a meter to over a hundred meters making them appropriate for a wide range of potential applications.

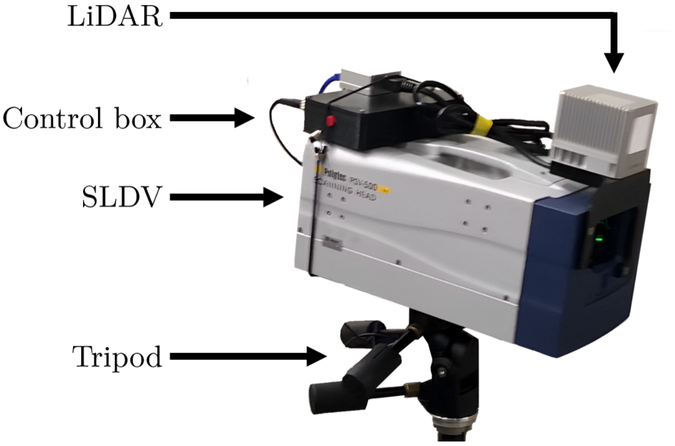


Figure A.1: Annotated photograph of the SLDV LiDAR sensor system.

A laptop orchestrated the entire scan procedure and conveniently collected the data, as can be seen in Figure A.2. This was straightforwardly achieved for the LiDAR scanner as it is designed to interface via an Ethernet port, both receiving commands and returning the point cloud data. In order to control the SLDV, the galvanometer control box described in Chapter 5 was used in the same way as it was in Chapter 6, receiving mirror angle commands from the laptop over serial communication, then applying the appropriate voltages to the scanning mirrors in the SLDV. The measurements from the SLDV were collected by the computer over the microphone jack, again in the same way as Chapter 6. Each measurement was taken using a sampling frequency of 44.1 kHz for a duration of 1 s, providing a spectral resolution of 1 Hz and a bandwidth of 22.05 kHz. These LDV measurements were taken over FOV of ±13° in both pitch and yaw mirrors at intervals 1.5°.

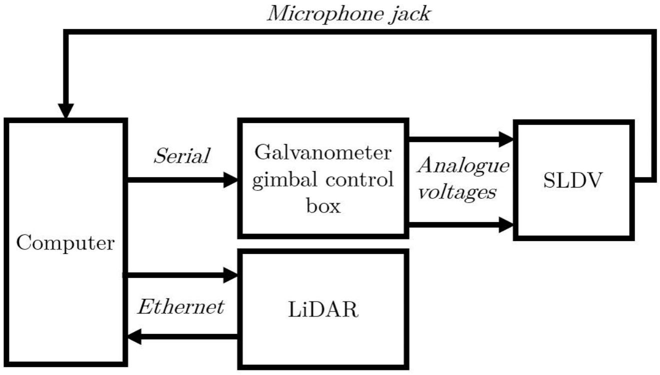


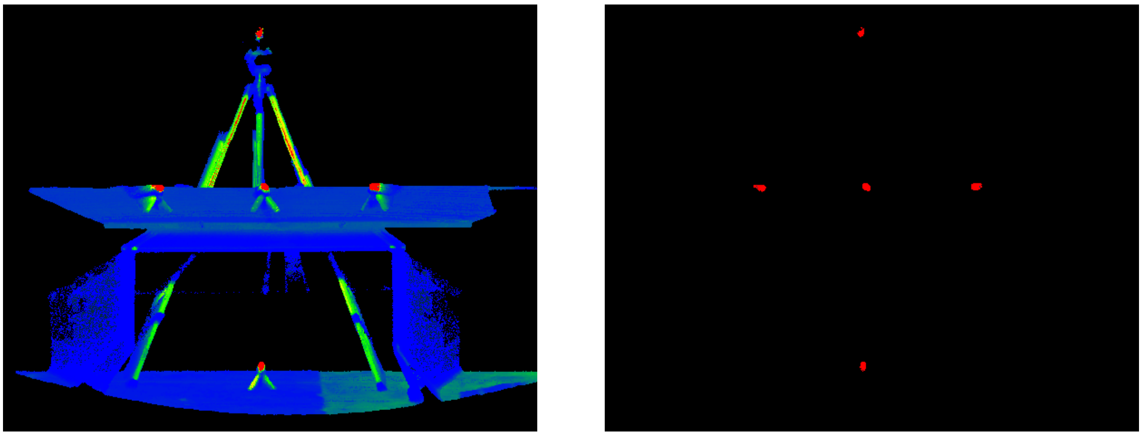
FIGURE A.2: Block diagram of the SLDV-LiDAR measurement system. Here, a laptop orchestrates the acquisition of both sensors, straightforwardly achieved with the LiDAR, and made possible for the SLDV using the galvanometer control box.

A.5 Reference frame alignment

To accurately associate the LDV measurements to the LiDAR data, the relative positions and angles of the two sensors must be known. This can either be precisely measured or experimentally derived via a calibration procedure, as done here. The bespoke 3D-printed mount was designed using 3D files of PSV-500 Xtra provided by the manufacturer such that there would be no angular misalignment between the two sensors. However, in reality, there will be a small amount of angular misalignment due to the limited accuracy of fused deposition modelling 3D printers. For short-range scans, this will have little effect on the reference frame alignment, however, for longer-range scans the angular misalignment is amplified. Therefore, it is best to ensure the reference frames are aligned in rotation.

FIGURE A.3: Photograph of the retro-reflective calibration target used during the reference frame alignment procedure.

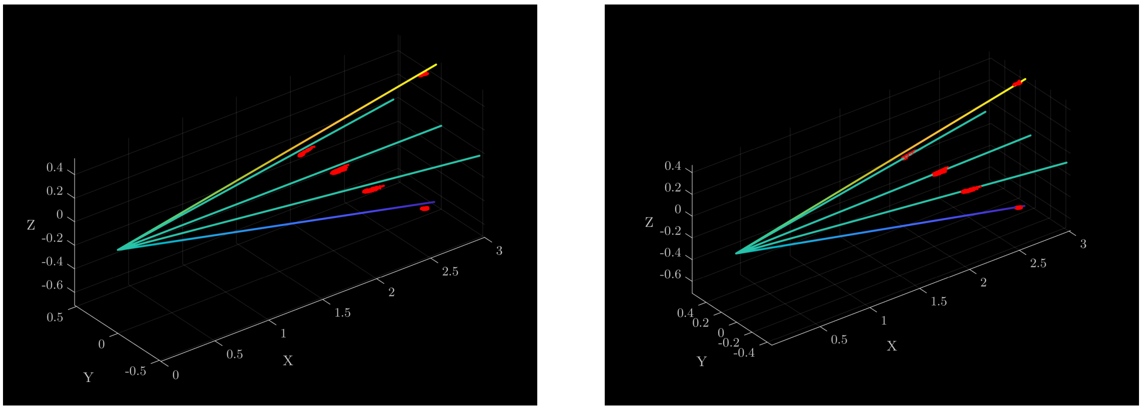
The reference frame alignment procedure requires retro-reflective targets, shown in Figure A.3, as their locations can be easily identified within the reflectivity point cloud. The target is placed within the scene and the LDV laser is directed to its centre, the angles of the steering mirrors are noted, the laser is switched off and a LiDAR scan is obtained. This procedure is repeated with the target at various locations within the scene. As long as the system was not moved between the various scans, the point clouds can be merged, as seen in Figure A.4 (a). The target can be clearly identified within the scene due to its high reflectivity at the five chosen locations. This point cloud can then be filtered as to only include the most reflective points, as seen in Figure A.4 (b).



(a) (b)

FIGURE A.4: The five merged calibration point clouds coloured by reflectivity. (a) before and (b) after filtering by reflectivity.

Since each of the calibration target clusters seen in Figure A.4 (b) has an associated pitch and yaw SLDV beam angle, they can be used to align the two reference frames. The first stage is the generation of point clouds representing the LDV beam vectors with a predefined resolution; here, lines were generated at the appropriate angles with a resolution of 10 points per mm. These LDV beam vector point clouds and the filtered calibration point cloud can be seen in Figure A.5 (a). Since the reference frames of the SLDV and the LiDAR scanner are not aligned, the LDV beam vectors do not coincide with the calibration targets as they did during the experiment, as they do not really originate from the LiDAR scanner origin. In order to align the two reference frames, translational offsets are required, and ideally small rotational offsets. Practically, these can be estimated using the Iterative Closest Point (ICP) algorithm. In this case, applied using the `pcregistericp` function; which is an implementation of published work [134, 135]. This algorithm adjusts the LDV beam vectors' translation and rotation until they best coincide with the calibration target clusters, as they did in reality. The outputted transform between the two point clouds was shown to be (x, y, z) = (1.5 cm, 0.8 cm, -8.0 cm) with Euler angles (α, β, γ) = (0.21°, 0.15°, -1.02°). Therefore, by using this information to apply a coordinate transform to the LDV beam vector point cloud, the two point clouds can be correctly aligned, as shown in Figure A.5 (b). This transform is therefore applied to *all* subsequent scans.



(a) (b)

FIGURE A.5: The generated LDV beam vector point clouds with the merged and filtered calibration point cloud. Shown in (a) before and (b) after reference frame alignment.

By applying the inverse reference frame transform in both translation and rotation to the measured point cloud, both the location of the origin and the direction of the axes are shifted to correspond to that of the LDV; therefore, the non-scanning SLDV beam axis would the same as the z axis in the point cloud data. Therefore, to find the angle of a point in the point cloud relative to the SLDV simple trigonometry can be used, with the pitch being given by:

$$\phi = \tan\left(\frac{z}{x}\right) \tag{A.2}$$

and the yaw being given by:

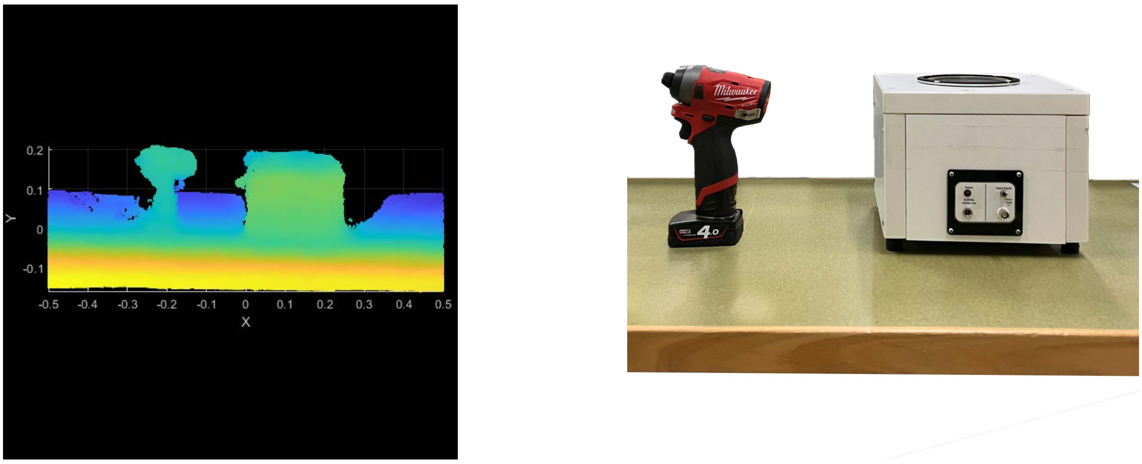
$$\theta = \tan\left(\frac{y}{x}\right) \tag{A.3}$$

where x, y, z are the point cloud Cartesian coordinates. Practically, this means that the SLDV beam angles θ and ϕ can be calculated for each point in the point cloud such that they may be correctly associated. However, since there will not be a unique LDV measurement for every point in the point cloud, they are grouped according to the angular resolution of the SLDV scans. In this work, the LDV scans were acquired with an angular resolution of 1.5°, therefore, any points within the angular range of $\theta \pm 0.75°$ and $\phi \pm 0.75°$ to a scan angle would be associated with that particular LDV measurement.

A.6 Enhanced perception example application

This section explores one potential application of LDV-LiDAR data fusion to enhance robotic perception. This proof of concept scenario focuses on the ability of a robot to take a LiDAR-LDV fused scan, identify and segment objects within the point cloud scene and analyse each object's vibration profile. Some potential applications include enhanced object recognition if combined with the work in Chapter 6; or even autonomous fault detection in industrial spaces, where regular autonomous inspections could check for any significant deviation in the vibration profiles of machinery.

In order to showcase the potential, a fused scan was taken of a scene containing an operational drill and a speaker box outputting a 2 kHz pure tone. The drill was selected as it represents a fairly reasonable analog of an operating machine, whereas the speaker box was included for debugging purposes since it would have a clearly identifiable peak. Figure A.6 shows a 3D point cloud alongside a photograph of this scene. The scene here has already been partially processed by removing any points outside of the region of interest, in this case, an area on a table. All point cloud processing was performed in MATLAB using the Computer Vision Toolbox.



(a) (b)

FIGURE A.6: The test setup used in the preliminary testing of the system. (a) point cloud and (b) photograph of the scene.

Figure A.7 describes the processing procedure following the acquisition of both LiDAR and LDV scans. First, the point cloud is filtered using the `pcdenoise` function in MATLAB; this function is an implementation of existing work [136]. In order to decrease processing times, the scans were also downsampled using the `pcdownsample` function, which utilises techniques found in existing work [137]. This downsampling was implemented as a box grid filter with a spatial resolution of 5 mm. In order to isolate objects within the resulting point cloud, a plane is fit to the data using `pcfitplane`; which is an implementation of the M-estimator Sample Consensus technique [138]. To reduce errors, the fit is limited to upward-facing planes with a tolerance of 10°. This prevents surfaces, such as the speaker's front face, from being mistaken for the relevant table plane. Similarly, the top surface of the speaker is not mistaken for the table surface since it can be safely assumed that the largest plane is that of the table. While in this application, the table surface is identified, this would similarly work for identifying the floor surface. In these less controlled environments, additional criteria could be included to reject non-floor surfaces, such as the relative heights of the planes. Finally, any points within a user-defined tolerance of the plane, specified here to be 1.5 cm, are subsequently removed, leaving a point cloud containing just the objects.

To segment the point cloud into multiple point clouds each containing an object, the spatial separation of the objects is utilised. Practically, this involves clustering the data using the `pcsegdist` function; this utilises the minimum euclidean distance between points from different clusters in order to separate them. The results of this can be seen in Figure A.8 for both the drill and the speaker box. It is important to note that this technique breaks down if the two objects are placed too close to one another, however, later acoustic interrogation could be used to distinguish them.

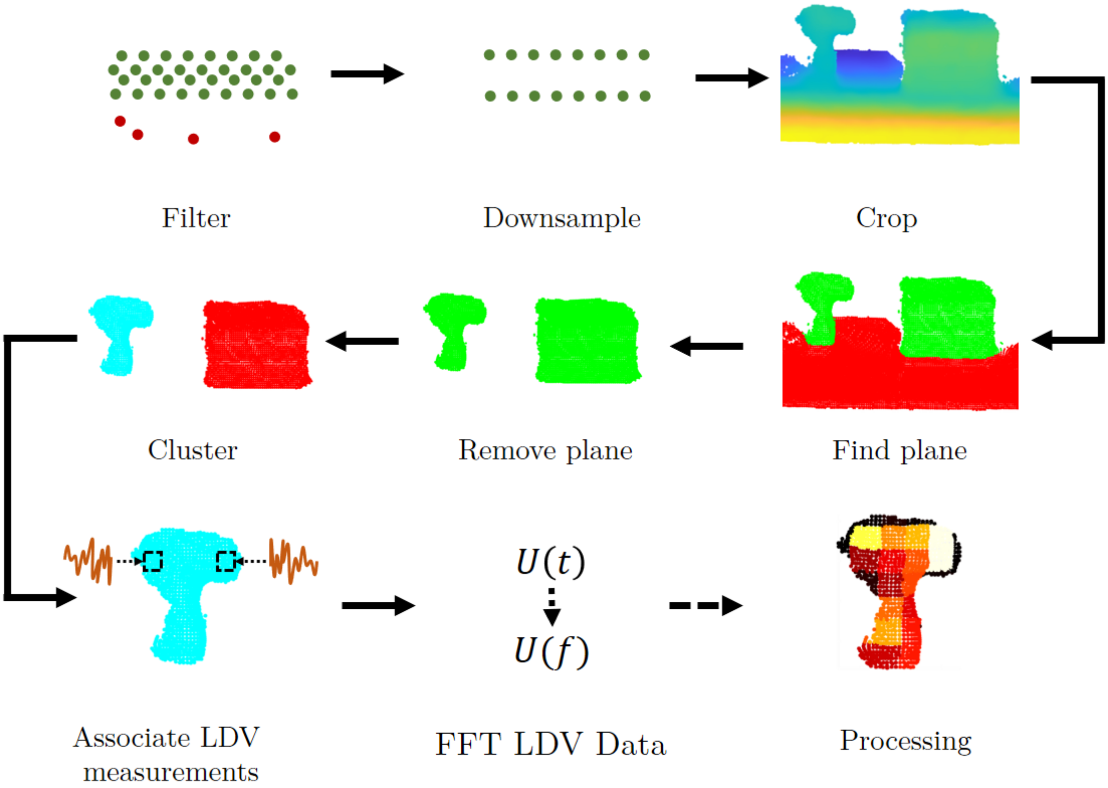
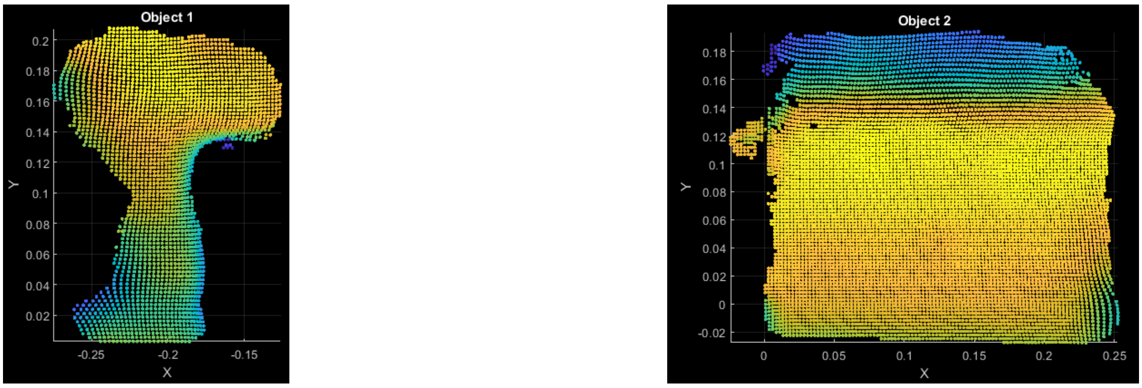


Figure A.7: Processing applied after acquiring vibration measurements and a point cloud scan using the LDV and LiDAR, respectively. Here, the processing stage simply generates spectral-coloured point clouds, however, this could consist of any processing technique.

Now that each object's point cloud has been identified, the LDV scans can be associated with the relevant points; as can be seen in Figure A.7. This is described in Section A.5 but involves calculating the angle of each point in the point cloud relative to the SLDV origin, therefore, allowing the points to be grouped accordingly. Since these data were collected as time data, the next stage is to apply an FFT to convert them into spectra such that they may finally be processed. Ultimately, this work is intended to be used in conjunction with the work presented in Chapter 6 to enhance object recognition, where objects with similar or challenging geometries can be interrogated using an LDV measurement. In that specific scenario, the object coordinates would first be identified and the SLDV would subsequently be directed in that direction, rather than here, where scans have already been acquired over the entire SLDV FOV for a post processing-based approach.



(a)

(b)

FIGURE A.8: The point clouds of the objects identified within the scene, with (a) the drill and (b) the speaker box.

In order to view the highly dimensional data associated with the point cloud, a variety of data visualisation techniques can be applied. Figure A.9 shows one potential approach where the points in the drill point cloud are coloured according to the RMS amplitude of the vibration in each octave. In the lowest frequency range of 11 Hz to 22 Hz, the rocking motion seems to be highlighted. Similarly, the frequency range of 355 Hz to 710 Hz, the area containing the motor is vibrating with a high amplitude, potentially being excited by the frequency of the motor rotation. This imaging modality now allows the autonomous system to infer the operational state of mechanical objects within the environment. Additionally, by merging this work with that described in Chapter 6, enhanced object recognition could be achieved.

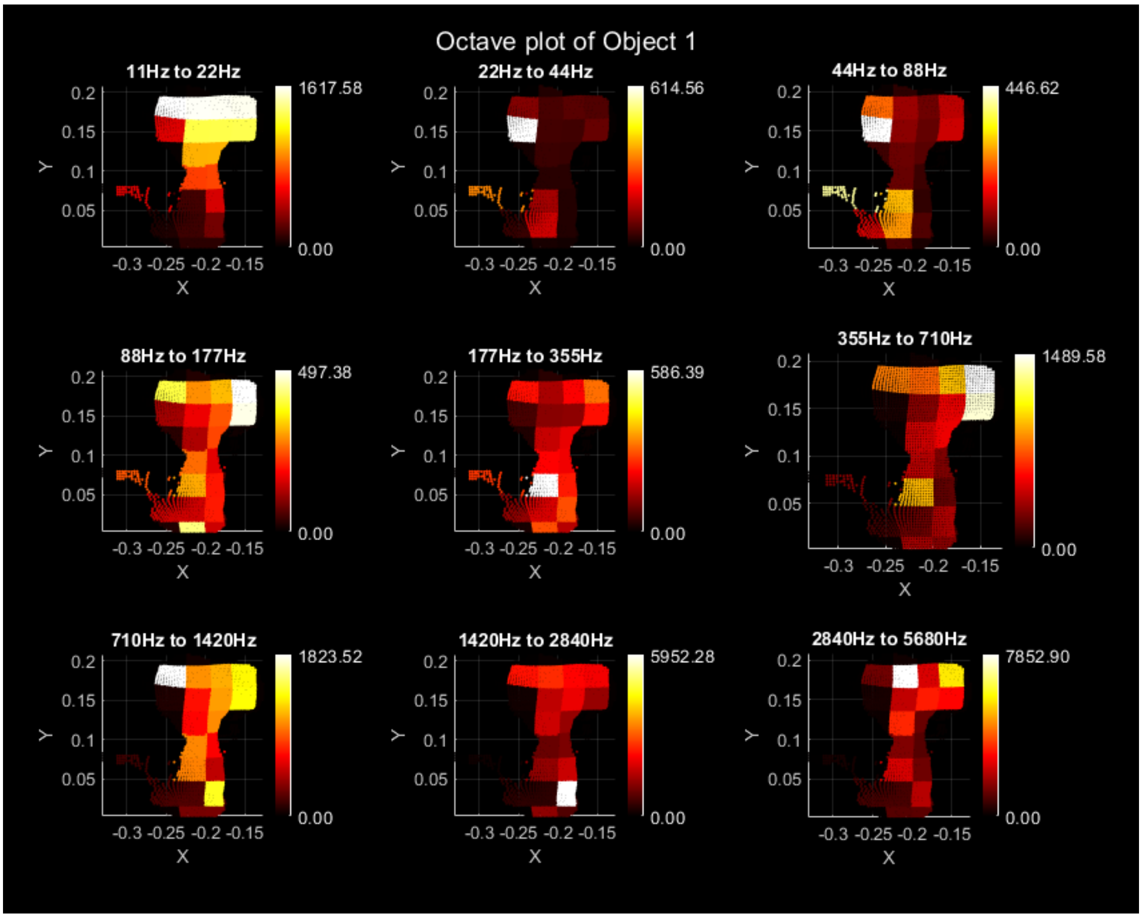


Figure A.9: The point cloud of the first object identified within a scene coloured by the vibration amplitude in the region over various octaves. The units here are arbitrary as the analogue-to-digital converter in the mic port is uncalibrated.

A.7 Summary and discussion

This appendix aimed to establish the groundwork to enable the fusion of LiDAR and LDV sensor modalities for autonomous robots. It extended the work presented in Chapter 6 into a more practically viable object recognition technique by integrating an SLDV into a LiDAR-based autonomous vehicle application. Although recognition itself was not demonstrated, the infrastructure behind the merging of the two sensor modalities was developed along with the automatic localisation of objects of interest within a scene so subsequent vibro-acoustic interrogation. This is a promising concept as it would enable autonomous systems to perceive the vibrations of the surrounding world with a high spatial resolution and sensitivity; unlocking a whole host of new applications. A few examples include autonomous fault detection in machinery - where a robot can could roam a warehouse and autonomously assess the health of the various machines; enhanced object recognition for autonomous systems, by combining this with the work from the previous chapter, objects otherwise confused for one another in a LiDAR scan could be distinguished using their acoustic fingerprint; and enhanced Simultaneous Localisation and Mapping (SLAM), where the vibrations of the surrounding environment could be used as additional landmarks. There are likely many more unforeseen applications that are yet to be conceived.

Currently, the system scans the SLDV over the entire field of view, which is time-consuming and unnecessary, as many of the measurements are not processed. This was convenient for the post-processing approach applied here but likely not part of the final technique. Ideally, localising objects of interest within the scene would directly determine the SLDV mirror angles for further vibro-acoustic interrogation. Secondly, this work can be integrated with that shown in Chapter 6, where similar objects or objects with difficult geometries can be better identified using a combined SLDV-LiDAR approach rather than just the LiDAR data. Finally, this work should be extended into some of the aforementioned applications beyond object detection.

References

[1] A. A. Michelson and E. W. Morley. “On the relative motion of the Earth and the luminiferous ether”. *American Journal of Science* 33-34.203 (1887), pp. 333–345.

[2] T. H. Maiman. “Stimulated Optical Radiation in Ruby”. *Nature* 187 (1960), pp. 493–494.

[3] J. Hecht. “A Short history of laser development of Photo-Optical Instrumentation”. *Optical Engineering* 49.49 (2010), pp. 99–122.

[4] S. J. Rothberg, M. S. Allen, P. Castellini, D. Di Maio, J. J. Dirckx, D. J. Ewins, B. J. Halkon, P. Muyshondt, N. Paone, T. Ryan, H. Steger, E. P. Tomasini, S. Vanlanduit, and J. F. Vignola. “An international review of laser Doppler vibrometry: Making light work of vibration measurement”. *Optics and Lasers in Engineering* 99.July (2017), pp. 11–22.

[5] R. A. Lomenzo, A. J. Barker, and A. L. Wicks. “Laser vibrometry system for rotating bladed disks”. *Proceedings of the 17th International Modal Analysis Conference.* 1999.

[6] I. A. Sever, A. B. Stanbridge, and D. J. Ewins. “Turbomachinery blade vibration measurements with tracking LDV under rotation”. *Seventh International Conference on Vibration Measurements by Laser Techniques: Advances and Applications.* Ed. by E. P. Tomasini. Vol. 6345. International Society for Optics and Photonics. SPIE, 2006, p. 63450L.

[7] B. J. Halkon and S. J. Rothberg. “Vibration measurements using continuous scanning laser vibrometry: Advanced aspects in rotor applications”. *Mechanical Systems and Signal Processing* 20.6 (2006), pp. 1286–1299.

[8] B. J. Halkon and S. J. Rothberg. “Vibration measurements using continuous scanning laser Doppler vibrometry: Theoretical velocity sensitivity analysis with applications”. *Measurement Science and Technology* 14.3 (2003), pp. 382–393.

[9] S. J. Bell J. R.; Rothberg. “Rotational vibration measurements using laser Doppler vibrometry : Comprehensive theory and practical application”. *Journal of Sound and Vibration* 238.4 (2000), pp. 673–690.

[10] C. Rembe, L. Muller, R. S. Muller, and R. T. Howe. “Full three-dimensional motion characterization of a gimballed electrostatic microactuator”. *2001 IEEE International Reliability Physics Symposium Proceedings. 39th Annual.* 2001, pp. 91–98.

[11] S. C. Ko, Y. C. Kim, S. S. Lee, S. H Choi, and S. R. Kim. “Micromachined piezoelectric membrane acoustic device”. *Sensors and Actuators A: Physical* 103.1 (2003), pp. 130–134.

[12] E. M. Lawrence, K. Speller, and D. Yu. “Laser Doppler vibrometry for optical MEMS”. *Fifth International Conference on Vibration Measurements by Laser Techniques: Advances and Applications.* Ed. by E. P. Tomasini. Vol. 4827. International Society for Optics and Photonics. SPIE, 2002, pp. 80–87.

[13] K. Speller, H. Goldberg, J. Gannon, and E. M. Lawrence. “Unique MEMS characterization solutions enabled by laser Doppler vibrometer measurements”. *Fifth International Conference on Vibration Measurements by Laser Techniques: Advances and Applications.* Ed. by E. P. Tomasini. Vol. 4827. International Society for Optics and Photonics. SPIE, 2002, pp. 478–485.

[14] A. Z. Khan, A. B. Stanbridge, and D. J. Ewins. “Detecting damage in vibrating structures with a scanning LDV”. *Optics and Lasers in Engineering* 32.6 (1999), pp. 583–592.

[15] K. Waldron, A. Ghoshal, M. J. Schulz, M. Sundaresan, F. Ferguson, P. Pai, and J. H. Chung. “Damage detection using finite element and laser operational deflection shapes”. *Finite Elements in Analysis and Design* 38 (2002), pp. 193–226.

[16] W. J. Staszewski, B. C. Lee, L. Mallet, and F. Scarpa. “Structural health monitoring using scanning laser vibrometry: I. Lamb wave sensing”. *Smart Materials and Structures* 13.2 (2004), pp. 251–260.

[17] A. C. Okl. “Vibratory signal transmission in plants as measured by laser vibrometry”. *Periodicum Biologorum* 90.2 (1988), pp. 193–196.

[18] S. Oberst, E. Nava-Baro, J. C. L. Lai, and T. A. Evans. “An innovative signal processing method to extract ants walking signals”. *Acoustics Australia* 43 (2015), pp. 87–96.

[19] B. Aicher, H. Markl, W. M Masters, and H. L Kirschenlohr. “Vibration transmission through the walking legs of the fiddler crab, Uca pugilator (Brachyura, Ocypodidae) as measured by laser Doppler vibrometry”. *Journal of Comparative Physiology* 150.4 (1983), pp. 483–491.

[20] Y. Li, P. Segers, J. Dirckx, and R. Baets. “On-chip laser Doppler vibrometer for arterial pulse wave velocity measurement”. *Biomedical Optics Express* 4.7 (2013), pp. 1229–1235.

[21] P Marchionni, L Scalise, I Ercoli, and E. P. Tomasini. “An optical measurement method for the simultaneous assessment of respiration and heart rates in preterm infants”. *Review of Scientific Instruments* 84.12 (2013), p. 121705.

[22] B. Libbey, D. Fenneman, and B. Burns. “Mobile platform for acoustic mine detection applications”. *Detection and Remediation Technologies for Mines and Minelike Targets X* 5794 (2005), p. 683.

[23] T. V. Writer, J. M. Sabatier, M. A. Miller, and K. D. Sherbondy. “Mine detection with a forward moving portable laser Doppler vibrometer”. *Proceedings of SPIE - The International Society for Optical Engineering* 4742 (2002), pp. 649–653.

[24] R. D. Burgett, M. R. Bradley, M. Duncan, J. Melton, A. K. Lal, V. Aranchuk, C. F. Hess, J. M. Sabatier, and N. Xiang. “Mobile mounted laser Doppler vibrometer array for acoustic landmine detection”. *Detection and Remediation Technologies for Mines and Minelike Targets VIII* 5089 (2003), p. 665.

[25] A. Dräbenstedt, X. Cao, U. Polom, F. Pätzold, T. Zeller, P. Hecker, V. Seyfried, and C. Rembe. “Mobile seismic exploration”. *AIP Conference Proceedings* 1740 (2016).

[26] P. Garg, R. Nasimi, A. Ozdagli, S. Zhang, D. D. L. Mascarenas, M. R. Taha, and F. Moreu. “Measuring transverse displacements using unmanned aerial systems laser doppler vibrometer (UAS-LDV): Development and field validation”. *Sensors (Switzerland)* 20.21 (2020), pp. 1–16.

[27] S. W. Courville and P. C. Sava. “Speckle noise attenuation in orbital laser vibrometer seismology”. *Acta Astronaut.* 172 (2020), pp. 16–32.

[28] S. W. Courville and P. C. Sava. “Speckle noise in orbital laser doppler vibrometry”. *Lunar and Planetary Science Conference* 39.8 (2019), pp. 697–699.

[29] P. Sava and E. Asphaug. “Seismology on small planetary bodies by orbital laser Doppler vibrometry”. *Advances in Space Research* 64.2 (2019), pp. 527–544.

[30] J. L. Richmond and B. J. Halkon. “Speaker diarisation of vibroacoustic intelligence from drone mounted laser doppler vibrometers”. *Journal of Physics: Conference Series* 2041.1 (2021).

[31] B. Halkon and C. Chapman. “On the development and characterisation of a synchronised-scanning laser doppler vibrometer system”. *25th International Congress on Sound and Vibration 2018, ICSV 2018: Hiroshima Calling.* Vol. 5. 2018, pp. 2876–2883.

[32] M. A. A. Ismail, A. Bierig, S. R. Hassan, R. Kumme, and P.-T. Bundesanstalt. “Flyable Mirrors : Laser Scanning Vibrometry Method for Monitoring Large Engineering Structures Using Drones”. *Proc. SPIE* 11142 (2019), pp. 5–8.

[33] M Schewe, M. A. A. Ismail, and C Rembe. “Towards airborne laser Doppler vibrometry for structural health monitoring of large and curved structures”. *Insight - Non-Destructive Test. Cond. Monit.* 63.5 (2021), pp. 280–282.

[34] A. I. N. Alshbatat. “Heuristic Optimization Technique to Locate and Avoid Buried Landmines: Drone-based Approach”. *International Journal of Information Technology and Computer Science* 10.11 (2018), pp. 49–57.

[35] M. Schewe, C. Rembe, H. Fritze, H. Wulfmeier, and D. Kohlmann. “Differential Laser Doppler Vibrometry for Displacement Measurements Down to 1 MHz with Amplitude Resolution Below 1 Nm”. *SSRN Electronic Journal* (2022), pp. 1–15.

[36] Tae-Gyu Chang, Yong-Gi Son, Jae-Hwa Kim, Ho-Seung Kim, and Min-Shik Kang. “A laser Doppler vibrometer featured with the in-housed mechanism for adaptive compensation of body vibration”. Institute of Electrical and Electronics Engineers (IEEE), 2002, pp. 535–536.

[37] L. A. Jiang, M. A. Albota, R. W. Haupt, J. G. Chen, and R. M. Marino. “Laser vibrometry from a moving ground vehicle”. *Applied Optics* 50.15 (2011), p. 2263.

[38] B. J. Halkon and S. J. Rothberg. “Taking laser Doppler vibrometry off the tripod: correction of measurements affected by instrument vibration”. *Optics and Lasers in Engineering* 99 (2017). Laser Doppler vibrometry, pp. 3–10.

[39] H. Kim, Y. Lee, C. Kim, T.-g. Chang, and M.-S. Kang. “Laser Doppler vibrometer with body vibration compensation”. *Opt. Eng.* 42.8 (2003), p. 2291.

[40] M. Klun, D. Zupan, J. Lopatič, and A. Kryžanowski. “On the application of laser vibrometry to perform structural health monitoring in non-stationary conditions of a hydropower dam”. *Sensors (Switzerland)* 19.17 (2019), p. 3811.

[41] B. J. Halkon and S. J. Rothberg. “Establishing correction solutions for scanning laser Doppler vibrometer measurements affected by sensor head vibration”. *Mechanical Systems and Signal Processing* 150 (2021), p. 107255.

[42] P. Castellini and E. P. Tomasini. “Image-based tracking laser Doppler vibrometer”. *Review of Scientific Instruments* 75.1 (2004), pp. 222–232.

[43] K. Okumura, H. Oku, and M. Ishikawa. “High-speed gaze controller for millisecond-order pan/tilt camera”. *Proceedings - IEEE International Conference on Robotics and Automation.* 2011, pp. 6186–6191.

[44] R. Burr, M. Schartel, P. Schmidt, W. Mayer, T. Walter, and C. Waldschmidt. “Design and Implementation of a FMCW GPR for UAV-based Mine Detection”. *2018 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility, ICMIM 2018.* Institute of Electrical and Electronics Engineers Inc., 2018.

[45] E. M. Rosen and K. D. Sherbondy. “Performance assessment of mine detection systems”. *Detection and Remediation Technologies for Mines and Minelike Targets V* 4038 (2000), p. 1225.

[46] T. Diaz. “Lights, drone... action”. *IEEE Spectrum* 52.7 (2015), pp. 36–41.

[47] S. Takamuku, K. Hosoda, and M. Asada. “Object category acquisition by dynamic touch”. *Advanced Robotics* 22.10 (2008), pp. 1143–1154.

[48] C. L. Chen, J. O. Snyder, and P. J. Ramadge. “Learning to identify container contents through tactile vibration signatures”. *2016 IEEE International Conference on Simulation, Modeling, and Programming for Autonomous Robots, SIMPAR 2016.* IEEE, 2017, pp. 43–48.

[49] J. Sinapov, T. Bergquist, C. Schenck, U. Ohiri, S. Griffith, and A. Stoytchev. “Interactive object recognition using proprioceptive and auditory feedback”. *International Journal of Robotics Research* 30.10 (2011), pp. 1250–1262.

[50] S. Jin, H. Liu, B. Wang, and F. Sun. “Open-environment robotic acoustic perception for object recognition”. *Frontiers in Neurorobotics* 13 (2019), pp. 1–15.

[51] A. Darwish, B. Halkon, S. Oberst, R. Fitch, and S. Rothberg. “Correction of laser Doppler vibrometer measurements affected by sensor head vibration using time domain techniques”. *Proceedings of the International Conference on Structural Dynamics, EURODYN* 2 (2020), pp. 4842–4850.

[52] A. Darwish, B. Halkon, S. Rothberg, S. Oberst, and R. Fitch. “A comparison of time and frequency domain-based approaches to laser Doppler vibrometer instrument vibration correction”. *Journal of Sound and Vibration* 520 (2022), p. 116607.

[53] A. Darwish, B. Halkon, and S. Oberst. “Non-Contact Vibro-Acoustic Object Recognition Using Laser Doppler Vibrometry and Convolutional Neural Networks”. *Sensors* 22.23 (2022), p. 9360.

[54] R. Holmewood, B. Halkon, and A. Darwish. “Towards real-time vibro-acoustic classification, verification and tracking of in-flight UAVs”. *International Conference on Noise and Vibration Engineering.* 2022.

[55] Y. Yeh and H. Z. Cummins. “Localized fluid flow measurements with an HeNe laser spectrometer”. *Applied Physics Letters* 4.10 (1964), pp. 176–178.

[56] J. Oldengarm, A. H. van Krieken, and H. J. Raterink. “Laser Doppler velocimeter with optical frequency shifting”. *Optics and Laser Technology* 5.6 (1973), pp. 249–252.

[57] N. A. Halliwell. “Laser-Doppler measurement of vibrating surfaces: A portable instrument”. *Journal of Sound and Vibration* 62.2 (1979), pp. 312–315.

[58] A. C. Lewin, A. D. Kersey, and D. A. Jackson. “Non-contact surface vibration analysis using a monomode fibre optic interferometer incorporating an open air path”. *Journal of Physics E: Scientific Instruments* 18.7 (1985), pp. 604–608.

[59] P. Buchhave. “Laser Doppler velocimeter with variable optical frequency shift”. *Optics and Laser Technology* 7.1 (1975), pp. 11–16.

[60] S. Rothberg. “Numerical simulation of speckle noise in laser vibrometry”. *Applied Optics* 45.19 (2006), pp. 4523–4533.

[61] M. Johansmann, G Siegmund, and M Pineda. “Targeting the limits of laser Doppler vibrometry”. *Proceedings of the International Disk Drive Equipment and Materials Association Conference* October 2014 (2005), pp. 1–12.

[62] K. Kiyono and Y. Tsujimoto. “Time and frequency domain characteristics of detrending-operation-based scaling analysis: Exact DFA and DMA frequency responses”. *Physical Review E* 94.1 (2016), pp. 1–3.

[63] E. Tomasini, G. Revel, and P. Castellini. “Laser Based Measurements”. *Encyclopedia of Vibration* (2001), pp. 699–710.

[64] P. Castellini, M. Martarelli, and E. P. Tomasini. “Laser Doppler Vibrometry: Development of advanced solutions answering to technology’s needs”. *Mechanical Systems and Signal Processing* 20.6 (2006), pp. 1265–1285.

[65] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis. “Geometrical Considerations and Nomenclature for Reflectance.” *National Bureau of Standards (US) Monograph* 160 (1977), pp. 1–52.

[66] H. W. Jensen, S. R. Marschner, M. Levoy, and P. Hanrahan. “A practical model for subsurface light transport”. *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics* (2001), pp. 511–518.

[67] J. Vass, R. Šmíd, R. B. Randall, P. Sovka, C. Cristalli, and B. Torcianti. “Avoidance of speckle noise in laser vibrometry by the use of kurtosis ratio: Application to mechanical fault diagnostics”. *Mechanical Systems and Signal Processing* 22.3 (2008), pp. 647–671.

[68] S. J. Rothberg and B. J. Halkon. “Laser vibrometry meets laser speckle”. *Proceedings of SPIE - The International Society for Optical Engineering* 5503 (2004), p. 280.

[69] S. J. Rothberg, B. J. Halkon, M. Tirabassi, and C. Pusey. “Radial vibration measurements directly from rotors using laser vibrometry: The effects of surface roughness, instrument misalignments and pseudo-vibration”. *Mechanical Systems and Signal Processing* 33 (2012), pp. 109–131.

[70] B. J. Halkon and S. J. Rothberg. “Angular (pitch and yaw) vibration measurements directly from rotors using laser vibrometry”. *Mechanical Systems and Signal Processing* 46.2 (2014), pp. 344–360.

[71] S. Rothberg, J. Baker, and N. Halliwell. “Laser vibrometry: Pseudo-vibrations”. *Journal of Sound and Vibration* 135.3 (1989), pp. 516–522.

[72] J. D. Parsons and J. G. Gardiner. “Diversity reception”. *Mobile Communication Systems.* Boston, MA: Springer US, 1989, pp. 189–243.

[73] D. Kiesewetter. “Numerical simulation of the retroreflection by glass beads”. *Seventh International Workshop on Nondestructive Testing and Computer Simulations in Science and Engineering.* Vol. 5400. International Society for Optics and Photonics. SPIE, 2004, pp. 185 –188.

[74] Polytec. “Polytec Scanning Vibrometer 500 Operating Manual” (2012).

[75] H. Kim, Y. Lee, T. Chang, and M. Kang. “Dual beam Laser Doppler vibrometer with body vibration compensation”. *Summaries of Papers Presented at the Lasers and Electro-Optics. CLEO ’02.* The Journal of the Optical Society of America, 2002, pp. 65–66.

[76] B. J. Halkon and S. J. Rothberg. “Restoring high accuracy to laser Doppler vibrometry measurements affected by vibration of beam steering optics”. *Journal of Sound and Vibration* 405 (2017), pp. 144–157.

[77] J. O. Smith. *Introduction to Digital Filters: With Audio Applications.* 2nd ed. W3K Pubishing, 2006, p. 458.

[78] J. H. Shapiro. “Correlation scales of laser speckle in heterodyne detection”. *Applied Optics* 24.12 (1985), p. 1883.

[79] C. A. Hill, M. Harris, K. D. Ridley, E. Jakeman, and P. Lutzmann. “Lidar frequency modulation vibrometry in the presence of speckle”. *Applied Optics* 42.6 (2003), p. 1091.

[80] A. Dräbenstedt. “Quantification of displacement and velocity noise in vibrometer measurements on transversely moving or rotating surfaces”. *Optical Measurement Systems for Industrial Inspection V*. Vol. 6616. International Society for Optics and Photonics. SPIE, 2007, p. 661632.

[81] Endevco. “Variable capacitance accelerometer datasheet - Model 770A-770F” (2022).

[82] P. Martin and S. J. Rothberg. “Pseudo-vibration sensitivities for commercial laser vibrometers”. *Mechanical Systems and Signal Processing* 25.7 (2011), pp. 2753–2765.

[83] F. Herz and R. Nordmann. “Vibration Limits”. *Vibrations of Power Plant Machines: A Guide for Recognition of Problems and Troubleshooting*. Cham: Springer International Publishing, 2020, pp. 245–248.

[84] A. Brandt and R. Brincker. “Integrating time signals in frequency domain - Comparison with time domain integration”. *Measurement: Journal of the International Measurement Confederation* 58 (2014), pp. 511–519.

[85] G. Carter, A. Nuttall, and P. Cable. “The smoothed coherence transform”. *Proceedings of the IEEE* 61.10 (1973), pp. 1497–1498.

[86] L. Miyashita, Y. Zou, and M. Ishikawa. “Vibrotracker: A vibrotactile sensor tracking objects”. *ACM SIGGRAPH 2013 Emerging Technologies, SIGGRAPH 2013* (2013).

[87] J. L. Richmond and B. J. Halkon. “Speaker diarisation of vibroacoustic intelligence from drone mounted laser doppler vibrometers”. *Journal of Physics: Conference Series* 2041.1 (2021).

[88] I. A. Sever, A. B. Stanbridge, and D. J. Ewins. “Turbomachinery blade vibration measurements with tracking LDV under rotation”. *Seventh International Conference on Vibration Measurements by Laser Techniques: Advances and Applications* 6345 (2006), p. 63450.

[89] M. R. Ashory, F. Talebi, H. R. Ghadikolaei, and M. Karimpour. “Investigating the vibrational behaviour of a rotating two-blade propeller by using a self-tracking method”. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 233.3 (2019), pp. 835–847.

[90] P. Castellini and C. Santolini. “Vibration measurements on blades of a naval propeller rotating in water with tracking laser vibrometer”. *Measurement: Journal of the International Measurement Confederation* 24.1 (1998), pp. 43–54.

[91] P. Castellini and R. Montanini. “Automotive components vibration measurements by tracking laser Doppler vibrometry: Advances in signal processing”. *Measurement Science and Technology* 13.8 (2002), pp. 1266–1279.

[92] B. Halkon and C. Chapman. “On the development and characterisation of a synchronised-scanning laser doppler vibrometer system”. *25th International Congress on Sound and Vibration 2018, ICSV 2018: Hiroshima Calling*, pp. 2876–2883.

[93] A. Jennings, J. Black, C. Allen, J. Simpkins, and R. Sollars. “Vibrometer Steering System for Dynamic In-flight Tracking and Measurement”. *Experimental Mechanics* 51.1 (2011), pp. 71–84.

[94] Z. Zhang. “A flexible new technique for camera calibration”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.11 (2000), pp. 1330–1334.

[95] Z. Zivkovic. “Improved adaptive Gaussian mixture model for background subtraction”. *Proceedings of the 17th International Conference on Pattern Recognition* 2.2 (2004), pp. 28–31.

[96] Z. Zivkovic and F. Van Der Heijden. “Efficient adaptive density estimation per image pixel for the task of background subtraction”. *Pattern Recognition Letters* 27.7 (2006), pp. 773–780.

[97] B. J. Halkon and S. J. Rothberg. “Towards laser Doppler vibrometry from unmanned aerial vehicles”. *Journal of Physics: Conference Series* 1149.1 (2018).

[98] J. M. Sabatier and N. Xiang. “Laser-Doppler-based acoustic-to-seismic detection of buried mines”. *Detection and Remediation Technologies for Mines and Minelike Targets IV* 3710.April (1999), p. 215.

[99] R. M. Fikry and H. Kasban. “Deep Neural Networks for Landmines Images Classification”. *Advances in Intelligent Systems and Computing.* Vol. 1261 AISC. Springer International Publishing, 2021, pp. 126–136.

[100] L. A. Jiang, M. A. Albota, R. W. Haupt, J. G. Chen, and R. M. Marino. “Laser vibrometry from a moving ground vehicle”. *Applied Optics* 50.15 (2011), p. 2263.

[101] A. Darwish, B. Halkon, and S. Oberst. “Non-contact vibro-acoustic object recognition using laser Doppler vibrometry and convolutional neural networks dataset”. *Mendeley Data* 405 (2017).

[102] J. Abeßer. “A review of deep learning based methods for acoustic scene classification”. *Applied Sciences* 10.6 (2020).

[103] G. Ciaburro and G. Iannace. “Improving smart cities safety using sound events detection based on deep neural network algorithms”. *Informatics* 7.3 (2020).

[104] B. McFee, C. Raffel, D. Liang, D. P. Ellis, M. McVicar, E. Battenberg, and O. Nieto. “librosa: Audio and music signal analysis in python”. *Proceedings of the 14th Python in Science Conference.* Vol. 8. 2015.

[105] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. “ImageNet Large Scale Visual Recognition Challenge”. *International Journal of Computer Vision* 115.3 (2015), pp. 211–252.

[106] W. S. McCulloch and W. Pitts. “A logical calculus of the ideas immanent in nervous activity”. *The Bulletin of Mathematical Biophysics* 5.4 (1943), pp. 115–133.

[107] J. Howard and S. Gugger. *Deep Learning for Coders with Fastai and PyTorch: AI Applications Without a PhD.* O’Reilly Media, Incorporated, 2020.

[108] L. Prechelt. “Early stopping-but when?” *Neural Networks: Tricks of the trade.* Springer, 1998, pp. 55–69.

[109] E. S. Olivas, J. D. M. Guerrero, M. M. Sober, J. R. M. Benedito, and A. J. S. Lopez. *Handbook Of Research On Machine Learning Applications and Trends: Algorithms, Methods and Techniques - 2 Volumes.* Hershey, PA: Information Science Reference - Imprint of: IGI Publishing, 2009.

[110] K. He, X. Zhang, S. Ren, and J. Sun. “Deep Residual Learning for Image Recognition”. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* 2016.

[111] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. “Imagenet: A large-scale hierarchical image database”. *2009 IEEE conference on computer vision and pattern recognition.* IEEE. 2009, pp. 248–255.

[112] J. Howard et al. *fastai.* `https://github.com/fastai/fastai`. 2018.

[113] L. N. Smith. “Cyclical learning rates for training neural networks”. *Proceedings - 2017 IEEE Winter Conference on Applications of Computer Vision, WACV 2017.* April. 2017, pp. 464–472.

[114] L. N. Smith. “A disciplined approach to neural network hyper-parameters: Part 1 - learning rate, batch size, momentum, and weight decay”. *arXiv CoRR* abs/1803.0 (2018), pp. 1–21.

[115] L. N. Smith and N. Topin. “Super-convergence: very fast training of neural networks using large learning rates”. *arXiv CoRR* abs/1708.0 (2019), pp. 1–36.

[116] M. Kuhn and K. Johnson. *Applied Predictive Modeling.* New York, NY: Springer New York, 2013.

[117] B. Krawczyk. “Learning from imbalanced data: open challenges and future directions”. *Progress in Artificial Intelligence* 5.4 (2016), pp. 221–232.

[118] P. Swanson. “Feasibility of using laser-based vibration measurements to detect roof fall hazards in underground mines”. *Fifth International Conference on Vibration Measurements by Laser Techniques: Advances and Applications.* Vol. 4827. 2002. 2002, pp. 541–552.

[119] A. Brink and M. C. Roberts. “Early Warning and/or Continous Risk Assessment of Rockfalls in Deep South African Mines”. *Proceedings of the Fourth International Seminar on Deep and High Stress Mining.* Australian Centre for Geomechanics, 2007, pp. 437–450.

[120] Y. K. Tee and Y. C. Han. “Lidar-Based 2D SLAM for Mobile Robot in an Indoor Environment: A Review”. *International Conference on Green Energy, Computing and Sustainable Technology (GECOST).* IEEE, 2021, pp. 1–7.

[121] B. Li, Y. An, D. Cappelleri, J. Xu, and S. Zhang. "High-accuracy, high-speed 3D structured light imaging techniques and potential applications to intelligent robotics". *International Journal of Intelligent Robotics and Applications* 1.1 (2017), pp. 86–103.

[122] U. Dhond and J. Aggarwal. "Structure from stereo-a review". *IEEE Transactions on Systems, Man, and Cybernetics* 19.6 (1989), pp. 1489–1510.

[123] S. Royo and M. Ballesta-Garcia. "An overview of lidar imaging systems for autonomous vehicles". *Applied Sciences* 9.19 (2019).

[124] H. K. V. Lotsch, W. T. Rhodes, E. B. A. Adibi, T Asakura, T. W. Hänsch, T Kamiya, F Krausz, B Monemar, H Venghaus, H Weber, H Weinfurter, and W. T. Rhodes. *Lidar*. Ed. by C. Weitkamp. Vol. 102. New York: Springer-Verlag, 2005.

[125] P. Chazette, J. Totems, L. Hespel, and J. S. Bailly. *Optical Remote Sensing of Land Surface: Techniques and Methods.* Elsevier Science, pp. 201–247.

[126] T. Raj, F. H. Hashim, A. B. Huddin, M. F. Ibrahim, and A. Hussain. "A survey on LiDAR scanning mechanisms". *Electronics* 9.5 (2020).

[127] T Hegna, H Pettersson, and K Grujic. "Inexpensive 3-D Laser Scanner System Based on a Galvanometer Scan Head". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 38 (2010), pp. 277–281.

[128] D. Wang, C. Watkins, and H. Xie. "MEMS mirrors for LiDAR: A review". *Micromachines* 11.5 (2020).

[129] M. Hansard, S. Lee, O. Choi, and R. Horaud. *Time-of-Flight Cameras: Principles, Methods and Applications.* Springer London, 2012.

[130] Y. Cui, S. Schuon, D. Chan, S. Thrun, and C. Theobalt. "3D shape scanning with a time-of-flight camera". *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2010), pp. 1173–1180.

[131] D. Anderson, H. Herman, and A. Kelly. "Experimental Characterization of Commercial Flash Ladar Devices". *Proceedings of International Conference on Sensing and Technology (ICST '05).* 2005.

[132] Polytec. “Polytec Scanning Vibrometer 400 Operating Manual” (2004).

[133] S. Sels, B. Ribbens, B. Bogaerts, J. Peeters, and S. Vanlanduit. “3D model assisted fully automated scanning laser Doppler vibrometer measurements”. *Optics and Lasers in Engineering* 99 (2017), pp. 23–30.

[134] P. Besl and N. D. McKay. “A method for registration of 3-D shapes”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14.2 (1992), pp. 239–256.

[135] A. Segal, D. Haehnel, and S. Thrun. “Generalized-ICP”. *Robotics: Science and Systems V.* Robotics: Science and Systems Foundation, 2009.

[136] R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha, and M. Beetz. “Towards 3D Point cloud based object maps for household environments”. *Robotics and Autonomous Systems* 56.11 (2008), pp. 927–941.

[137] F. Pomerleau, F. Colas, R. Siegwart, and S. Magnenat. “Comparing ICP variants on real-world data sets: Open-source library and experimental protocol”. *Autonomous Robots* 34.3 (2013), pp. 133–148.

[138] P. H. Torr and A. Zisserman. “MLESAC: A new robust estimator with application to estimating image geometry”. *Computer Vision and Image Understanding* 78.1 (2000), pp. 138–156.