SELF-WITNESSING COHERENT IMAGING FOR AUTOMATIC NOISE FILTERING AND ARTIFACT REMOVAL IN LASER PROCESS MONITORING

by

TESSA J. H. KRAUSE

A thesis submitted to the
Department of Physics, Engineering Physics, and Astronomy
in conformity with the requirements for
the degree of Master of Applied Science

Queen’s University
Kingston, Ontario, Canada
March 2022

Copyright © Tessa J. H. Krause, 2022
Abstract

High-power lasers are rapidly becoming standard tools in advanced manufacturing, mainly in the form of laser welding, laser cutting, and laser additive manufacturing. Of these applications, laser welding in the electric mobility sector—particularly in the manufacturing of battery packs—presents unique challenges. Weld depth needs to be precisely controlled to ensure joint strength and to ensure the weld does not puncture into the lithium ion cell. These processes often involve highly reflective metals (such as copper), which have material properties that lead to unstable welds; this requires an unprecedented level of control to ensure weld quality and depth. To ensure weld quality, we need in-line monitoring during the process. Inline Coherent Imaging (ICI) is a process monitoring technique that has been demonstrated to measure keyhole depth (down to 15 µm axial resolution) at high camera rates (200 kHz), with unparalleled sensitivity and dynamic range. Unfortunately, it suffers from sources of noise such as speckle and imaging artifacts that can pose a significant challenge to quality assurance and closed-loop control. To mitigate these problems, I have integrated a second, automatically synchronized, imaging channel into a standard ICI system, by exploiting a previously unused part of the imaging window. This “witness” image makes it possible to identify real signal based on correlation and filter out the uncorrelated noise. This has allowed feature-identification intensity threshold
values to be lowered from 12 dB to 8 dB. Using this system, we have demonstrated the complete removal of autocorrelation artifacts with no loss of imaging rate or spatial resolution compared to standard ICI. Between the first and second design for this technique, stability in the system increased by a factor of 12. When applied to imaging laser keyhole welding, I see improvements in the effective imaging rate of up to 79% relative to standard ICI for spot welds performed over 30 ms on 1000 series aluminum at 1100 W.
Acknowledgments

I am extremely fortunate to stand on the shoulders of many talented researchers, including past and current group members, collaborators, and the wider Queen’s community. Over the course of my degree, I have had the privilege of working alongside Ultrafast group members including Troy Allen, Tristan Fleming, James Godfrey, Cameron Romanet, and Heidi Miller. Research in the Ultrafast group is truly a collaborative process, and I am extremely grateful to have benefited from so many brilliant and driven researchers. I would also like to thank Declan Bracken, an undergraduate I hired through the CREATE-MAPS summer researcher program, who contributed to my research far beyond what I expected from a second year engineering student.

During my degree, I had the opportunity to explore microfabrication techniques for my experiments. I would specifically like to thank Dr. Graham Gibson at Nanofab Kingston for his patience and insight which made the fabrication of my final optics possible.

On a personal note, I would like to thank Elliot for his support through this degree. Graduate school can be an isolating experience, but I was fortunate that he ensured I was never alone. I would also like to thank my parents for the love of learning that they instilled in me from an early age; without it my education would have been far less enjoyable.
Finally, I would like to extend my sincerest thanks to my supervisor, Dr. James Fraser. I started my degree in what we now know was just the beginning of the COVID-19 pandemic. Though the pandemic affected many aspects of my degree, the quality of supervision I received was second to none. James’ commitment as an educator and researcher ensured that remote work would not impact the quality of my degree, and for that I am immeasurably grateful.
Contents

Abstract i

Acknowledgments iii

Contents v

List of Tables viii

List of Figures ix

List of Abbreviations xi

Chapter 1: Introduction 1

Chapter 2: Background 5

2.1 Motivation ................................................................. 5
2.2 Laser Welding ............................................................. 6
    2.2.1 Melt Pool Dynamics ............................................. 7
2.3 Existing Laser Welding Process Monitoring Solutions ............. 10
2.4 Inline Coherent Imaging ............................................. 12
    2.4.1 Optical Coherence Tomography ................................. 13
    2.4.2 Mathematical Basis for Low Coherence Interferometry .... 15
2.4.3 Dispersion ......................................................... 18
2.4.4 Sources of noise ............................................... 19

Chapter 3: Pre-Existing Infrastructure 22
3.1 Processing Beamline ........................................... 22
3.1.1 Motion Control ................................................ 23
3.2 Inline Coherent Imaging ....................................... 24
3.2.1 Data Processing ............................................... 25

Chapter 4: Initial self-witnessing inline coherent imaging implementation 31
4.1 Introduction ....................................................... 31
4.2 Methods and Materials ........................................ 31
4.2.1 Self-Witnessing Inline Coherent Imaging ................ 31
4.2.2 Laser Keyhole Welding ..................................... 33
4.3 System Characterization ....................................... 33
4.3.1 Static Surface Imaging ...................................... 33
4.3.2 Assessing the Correlation ................................. 34
4.4 Noise Filtering Using Witness Peak Identification ........... 42
4.5 Results ........................................................... 43
4.5.1 Weld Imaging .................................................. 45
4.6 Conclusion ....................................................... 47

Chapter 5: Stability improvements to self-witnessing coherent imaging 49
5.1 Introduction ....................................................... 49
5.2 Design ................................................. 50
5.3 Fabrication .......................................... 51
  5.3.1 First Production Run .......................... 52
  5.3.2 Second Production Run ....................... 53
5.4 Apparatus ............................................ 55
5.5 Data Processing ..................................... 57
  5.5.1 Homodyne Interpolation ....................... 57
  5.5.2 Intensity Filtering Approach .................. 58
  5.5.3 Phase Filtering Approach ...................... 59
5.6 Characterization .................................... 60
5.7 Exploiting the Improved Correlation ............ 67
5.8 Weld Imaging ....................................... 69
5.9 Conclusion .......................................... 72
5.10 Future Work ....................................... 73

Chapter 6: Conclusion ................................. 76

Bibliography ............................................ 79
List of Tables

5.1 Noise intensity broken down by intensity for different filters . . . . . . 69
# List of Figures

2.1 Laser welding mode diagram .............................................. 8  
2.2 Energy loss flowchart for laser welding ................................. 9  
2.3 Schematic of a Michelson interferometer ............................... 13  
2.4 Pixel intensity histograms from regions where signal (top left) and noise (top right) is expected in image taken while moving over a variable reflective surface (bottom). Region where signal is expected is highlighted by an orange line, and region where noise is expected is within the blue box on the bottom image. ................................... 20  
3.1 Existing inline coherent imaging apparatus ............................ 25  
4.1 Experimental apparatus for first implementation of SW-ICI .......... 32  
4.2 Single and dual arm images of a moving stainless steel sample .... 35  
4.3 2D histogram demonstrating intensity correlation for dual-arm image 36  
4.4 $z$-transformed interferogram for dual-arm peaks .................... 37  
4.5 Variation in path length between reference arm paths .............. 39  
4.6 2D histograms for signal and noise for dual arm images ............ 40  
4.7 $z$-transformed interferogram demonstrating discretization of intensity measurements .................................................. 43
4.8 Histogram of intensity composition for single arm and self-witnessed noise pixels

4.9 Unprocessed and processed self-witnessing keyhole welding data

4.10 Tracking for different signal intensities for keyhole welding

5.1 Process flow for step edge optic manufacturing

5.2 Picture of undiced step edge optics

5.3 Step edge reference arm schematic

5.4 Path length variation between the two paths on the step edge

5.5 2D histogram showing intensity correlation for step edge data

5.6 2D histograms for phase difference of signal and noise measurements

5.7 2D histograms for phase and intensity for welding data

5.8 Noise intensity composition for different filtering approaches

5.9 Unfiltered, intensity filtered and intensity+phase filtered welding data

5.10 Feature tracking for different intensity ranges on welding data
List of Abbreviations

A-line  amplitude-line.
A/D  analog to digital.
CMOS  complementary metal oxide semiconductor.
FFT  fast Fourier transform.
FWHM  full width half maximum.
HAZ  heat affected zone.
ICI  inline coherent imaging.
IDM  Inline Depth Monitor.
LDD  Laser Depth Dynamics.
NIST  National Institute of Standards and Technology.
OCT  optical coherence tomography.
SD-OCT  spectral-domain OCT.
SNR  signal-to-noise ratio.
SRM  standard reference material.
SS-OCT  swept-source OCT.
SW-ICI  self-witnessing ICI.
TD-OCT  time-domain OCT.
Chapter 1

Introduction

Due to the widespread adoption of electric vehicles for both consumer and industrial end-users, more lithium ion batteries are being manufactured than ever before. Lithium ion battery manufacturing relies on joining processes that are effective and safe when applied to electrically conductive materials such as aluminum and copper, while being in close proximity to highly-reactive lithium ion. These properties have made the use of traditional joining approaches challenging, as excess heat input into the lithium ion can have catastrophic consequences. As such, laser keyhole welding is quickly being adopted by the manufacturing industry as its high power density allows welds to be performed quickly with high axial precision. By performing the joining process quickly, heat input into the cell is minimize, enhancing the safety of welding materials containing lithium ion. Despite these benefits, laser welding of materials such as aluminum and copper present significant challenges in maintaining process stability due to the optical and thermal properties of such materials. If the laser penetrates into the cell during processing, the laser interaction with the lithium ion can lead to fire and explosion. In contrast, if the laser penetration is too shallow, the final weld seam may not have the required strength. This can lead to the weld joint
failing once the final product is in the hands of the end user. A partial failure of the weld joint can lead to heat build up on the cell while it is in use, potentially leading to thermal runaway. As such, it is critical to ensure process stability to ensure safety during manufacturing, and the reliability of the final product.

To ensure process stability, we need *in situ* monitoring approaches that are capable of tracking process dynamics with fluctuations at up to 10 kHz. The importance of quality assurance for laser welding is apparent in the number of different solutions that have been developed by both academia and industry to monitor weld quality. Approaches range from indirect sensor based systems that monitor process characteristics, to systems that take direct measurements of the weld while it is being performed. One direct measurement approach makes use of near infrared coherent imaging (originally developed for ophthalmological imaging) to track weld penetration, where a second imaging beam is coaxially aligned with the processing beam. Coherent imaging offers high-spatio temporal resolution, and a large dynamic range, which can offer near-real-time measurement to inform closed-loop control in laser welding processes. As with most imaging technologies, there are significant challenges in distinguishing signal from noise as signal approaches the noise floor. For signal-starved imaging applications, this may mean missing important weld dynamics, such as over or under penetration, because low intensity signal was treated as noise. In the case of lithium-ion battery manufacturing, this may mean losing depth measurement signal just as the laser penetrates through the cell and into the lithium ion. As such, I have developed an alternate metric to traditional thresholding to distinguish signal from noise. This technique, self-witnessing coherent imaging, is similar to matched-detection schemes used in other coherent imaging approaches, however without the
requirement for a specialized light-source, or a second detector.

Before discussing the development of this novel technique, it is important to understand the laser welding process, and existing monitoring techniques. In Chapter 2, I discuss the theory pertaining to laser keyhole welding, and what makes this process so challenging to control. In laser keyhole welding, we are dependant on a balance of energies and pressures to maintain a stable process. These balances are ultimately reliant on the stability of molten metal being irradiated with a high power laser, going through rapid temperature and phase changes, neither of which speak to inherent stability. Should the stability of the process not be maintained, the resulting joint is likely to have safety-critical failures, such as a low breaking force, or poor electrical conductivity in the case of electrical components. This leads to a discussion of existing control approaches in laser keyhole welding, that are used to monitor weld quality. Finally, I discuss a coherent imaging technique that can be used to monitor weld penetration, that is currently the gold standard in industrially applicable \textit{in situ} process control.

In Chapter 3, I discuss the existing infrastructure that I built upon to perform my research. In this chapter, I first discuss the physical infrastructure put in place by past group members, such as the processing beam line. I also describe the integration of the coherent imaging system into the beam line. Finally, I discuss the methods used to process the coherent imaging data, much of which were created by former group members.

In Chapter 4, I discuss my first implementation of a self-witnessing coherent imaging approach. Much of the contents of this chapter have been published in the Journal
of Optics and Lasers in Engineering [1]. The first implementation made use of off-the-shelf components to create a double reference arm in a coherent imaging system, effectively creating a second channel without the requirement of a second spectrometer. Using this apparatus, I demonstrated the value of adding a second path within the reference arm, as I was able to increase the effective imaging rate during welding by as much as 67%. In using off-the-shelf components, I struggled with maintaining stability between the two arms, seeing shifts of up to 300 nm in the difference between the two path lengths. It seems that these shifts were a result of vibrations in the various optical mounts caused by air currents and vibrations in the lab space. Though these shifts did not significantly impact my ability to use intensity as a metric to identify signal, it was impossible to make use of phase due to the lack of stability between the two arms.

In Chapter 5, I make use of a new approach to generate the double reference arm using a custom mirror that I manufactured at Nanofab Kingston. As the second arm is now generated by splitting the beam over a 93 µm step in the mirror, stability is improved by an order of magnitude (in comparison to the 50:50 approach), as both arms are now in the same plane, reliant on a single optic. As a result of this improved stability, I can now exploit the phase locking and intensity correlation between peaks generated by real signal. By exploiting phase and intensity as filtering metrics, I am able to further reduce the prevalence of noise, increasing my effective imaging rate by 79%.

In Chapter 6, I summarize my findings and discuss next steps that could be taken to build upon this research effort.
Chapter 2

Background

2.1 Motivation

The mass production of lithium ion battery packs for e-mobility applications has forged the way for a renaissance of laser beam welding research [2]. Battery manufacturing requires the joining of electrically conductive materials, such as aluminum and copper, that present a variety of challenges for both traditional and laser joining processes [3]. Laser beam welding is well-suited to battery manufacturing processes since the welding beam has a high energy density in comparison to traditional welding methods [4]. High energy density means that the heat affected zone (HAZ) is smaller, and welds can be performed faster. Both of these factors mean that less heat is input into the part.

Modern batteries for e-mobility applications typically exploit a lithium ion chemistry, due to its high energy density, and its ability to recharge quickly. The issue with using a lithium ion chemistry is that it is prone to catastrophic failure when exposed to heat. As such, it is critical to minimize heat input during manufacturing, which is why the e-mobility sector has been an early adopter of laser welding as a standard
2.2. LASER WELDING

manufacturing technique.

The challenge with welding materials such as aluminum and copper are that they have some properties that make achieving a stable laser weld challenging. This means that sudden changes in penetration depth are not uncommon. Should the penetration depth suddenly increase, the laser may puncture into the lithium ion cell. When punctured, the lithium will experience a rapid exothermic oxidation, leading to fire, and potentially explosion. Should the laser penetration decrease, this can lead to weaker joints with variable conductivity, which can lead to heat build up on the cell once it is in use.

As such, it is critical that we make use of inline quality assurance techniques to ensure that a process can be corrected prior to puncturing into the cell, or leaving a finished product with out-of-specification conduction properties. As weld dynamics can change at rates of up to 10 kHz (for welding processes of interest), we must achieve a measurement rate of at least 20 kHz, which can be a challenge given existing quality assurance techniques [5].

2.2 Laser Welding

When considering the monitoring of laser welding processes, it is important to understand the underlying mechanics of the process itself. Most laser welding processes can be divided into one of two categories: conduction mode, or keyhole mode (shown in Figure 2.1). Conduction mode is the most intuitive, where the laser heats the solid material until it transitions to its liquid phase. The challenge with using a laser to weld metal, is that many metals are highly reflective to the infrared processing beam. At 1040 nm, reflectance for metals such as steel, aluminum and copper range from
70 % to 99 % [6]. Due to the speed of the process, the depth into the material that can be melted is severely limited. As such, conduction mode welding is typically the process of choice for joints with low strength requirements, rather than welds that require greater strength.

To achieve greater penetration (and therefore strength), we make use of a phase transition from liquid to vapour within the molten material. When the laser vaporizes the melt, a vapour channel is formed. This vapour channel leads to multiple reflections, allowing for greater absorption than if the beam only has one interaction with the melt (as in conduction welding) [7]. During conduction mode welding, we can anticipate that the coupling energy will correspond to the reflection/absorption ratio for a given material at the wavelength of the laser [5]. Energy absorption in keyhole mode welding is far greater, due to multiple reflections within the keyhole. Allen et al., determined that keyhole welds performed on 316L stainless steel had an absorbance of over 0.8 (whereas conduction welds have an absorbance of 0.3 on the same material) [8].

### 2.2.1 Melt Pool Dynamics

In order to achieve the level of energy absorption required to generate sufficiently deep welds, the geometry of the keyhole (or vapour channel) is of the utmost importance. Keyhole formation is a result of the pressure that the vapour released once the molten material has absorbed sufficient energy from the incident laser beam. At first glance, keyhole formation seems like a relatively simple thermodynamics problem, where heat is applied to a solid, which subsequently transitions to a liquid, and then a gas. However, if we consider that the heat being input into the system is reliant
2.2. LASER WELDING

Figure 2.1: Conduction mode spot weld (left), keyhole mode spot weld (center), and solidified keyhole mode weld after laser has been turned off and melt solidifies (right).

on the absorption of a beam of light, into a material with a potentially non-uniform surface, with absorption properties that depend on its phase and temperature, it is clear that this is a non-trivial problem.

Keyhole geometry is ultimately governed by a balance of energies and a balance of pressures [7]. When considering the energies at play, the input energy is what is absorbed from the incident laser beam, with losses occurring due to phase transitions, conduction into the surrounding material, and convection into its surroundings. These losses for both conduction and keyhole mode welding are shown in Figure 2.2. In addition, due to Marangoni forces and recoil pressure applied on the melt pool, significant energy is lost to the conversion to kinetic energy [9]. Energy losses due to the phase transition from solid to liquid, and conduction into
2.2. LASER WELDING

the surrounding area are dependant on the melt pool geometry. As a relatively small volume of material goes through phase change, little energy is actually lost due to the latent heat of fusion. The majority of loss is a result of heat conduction to the solid material surrounding the melt. Ayool et al. found that for spot sizes less that 1 mm in diameter, the energy loss due to conduction can be over 8 times that lost through phase transition on steel samples [10]. Once the vaporization temperature

![Figure 2.2: Chart detailing the energy losses that occur as a laser weld moves between conduction and keyhole modes. $T$ represents the temperature of the area on the workpiece where the beam is incident, $T_L$ and $T_V$ are the melting and vapourization temperatures respectively specific to the material being processed, $T_S$ is the surface tension of the molten material, and $P_V$ is the vapour pressure. The most significant energy loss is typically due to conduction, as other losses are proportional to the mass of the material the laser is interacting with. Losses due to conduction are proportional to the mass of the entire workpiece.

for the specific material is reached, we can expect the formation of a keyhole due to vapour pressure of the ablated material. For a stable keyhole to form, the vapour pressure must be balanced by the surface tension of the molten material. Should the
keyhole collapse due the surface tension exceeding the vapour pressure in the pressure balance, the angle of incidence will sharply decrease. This will lead to a drop in energy absorption, further leading to a drop in vapour pressure. Conversely, should the vapour pressure exceed the surface tension, melt can be ejected from the keyhole. Melt ejections can lead to metal resolidifying as spatter on the surface of the part. Spatter is undesirable both from a cosmetic perspective, and for electrical parts it must be avoided to ensure the spatter does not bridge a connection and cause a short circuit.

Should the keyhole collapse, this will likely lead to a decrease in weld penetration, decreasing the overall strength of the joint. In addition, keyhole collapse can lead to the formation of pores within the processed material [11]. These pores have negative affects on both the mechanical and electrical properties of the joint. To create high quality joints, process stability is critical, but is difficult to achieve without the use of *in situ* measurement tools.

Keyhole dynamics typically occur at rates up to 10 kHz [5]. Therefore, based on Nyquist theory, measurement tools must make measurements at a rate of at least 20 kHz to fully capture changes in the keyhole.

2.3 Existing Laser Welding Process Monitoring Solutions

Laser welding is used to assemble a number of safety critical parts across industries ranging from medical to aerospace. As a result, there are a number of different approaches that have been used to understand and improve quality assurance on laser welding processes.

These methods can be divided into two categories *ex situ* and *in situ*. The most
common *ex situ* technique is transverse or longitudinal weld seam sectioning. Weld sectioning is generally regarded to be the gold standard in terms of weld quality assurance [12]–[14]. By sectioning the weld, one can determine the depth, while identifying pore formation, and in the case of dissimilar metal welding, identify intermetallics. Though sectioning offers a wealth of information about the weld, there are a number of disadvantages. The most significant disadvantage is that it is destructive, and therefore relies on taking a sample of a particular production run to be representative of all parts in the run. Another issue is that the measurements must be taken post-welding, and therefore cannot be used to correct the process on that particular part; anything learned from sectioning must be applied on subsequent welds. In addition, placing the section in the correct position can be challenging. For longitudinal sectioning (which gives the depth across the entire weld), this means cutting and polishing the part such that the section is exactly down the center line of the weld.

Due to the challenges of sectioning, both industry and academia have been motivated to find *in situ* techniques for taking measurements of laser welding processes. *In situ* monitoring can be further divided into indirect and direct monitoring approaches. Indirect process monitoring involves using a sensor (typically a photodetector), and characterizing what the sensor output is for a “good” weld. During production, sensor outputs are then compared to the original data taken for “good” welds. Should the sensor output vary from the original data, the part is either examined further, or discarded. The advantage of sensor-based monitoring systems is that they offer *in situ* process monitoring at relatively low economic cost. The issue with indirect monitors in general is that they are only capable of showing operator when the process has deviated away from the original process. They are not capable of informing the
2.4. INLINE COHERENT IMAGING

operator how the process has changed, only that it has changed. The benefit of moving towards direct monitoring solutions, is that these techniques have an output that can be correlated directly with a physical property of the weld, such as depth. Recent advances in high-brightness x-ray imaging have yielded beautiful high-speed cross-section video of laser welding and additive manufacturing, revealing previously hidden dynamics, but is limited to highly constrained part geometries and restricted to a few highly specialized research facilities (e.g., synchrotrons) [9], [15]. Due to the expense, constrained part geometries, and safety issues associated with the amount of radiation require to image through metal, x-ray imaging is generally not industrially viable. Coherent imaging has offered a more industrially viable approach to direct process monitoring, by using a secondary laser beam inline with the processing laser to measure the depth of the keyhole. Commercial coherent imaging systems are available from Precitec Gmbh. in the form of their Inline Depth Monitor (IDM) product, and from IPG Photonics Ltd. in their LDD sytems. In industry, this technique is typically referred to as optical coherence tomography (OCT) or as inline coherent imaging (ICI).

2.4 Inline Coherent Imaging

Inline coherent imaging (ICI) is a process monitoring technique similar to spectral-domain OCT (SD-OCT) that measures sample height during laser processing. ICI has shown promise in a number of applications such as laser ablation, drilling, welding and 3D additive manufacturing. Integration into a laser cell provides precise dynamic depth measurement of the sample at one location, or full 2D morphology through raster scanning. This technique has been used to measure process quality on materials
such as metal, silicon, and bone at speeds of up to 312 kHz and to achieve closed-loop control both manually and on-the-fly [16]–[20].

2.4.1 Optical Coherence Tomography

It would be incomplete to discuss the development of ICI without considering its parent technique, optical coherence tomography (OCT). The basis of OCT makes use of a Michelson interferometer where interference between a sample and reference arm for a spread of wavelengths can be measured. Information about the optical path length mismatch between the arms is embedded in the resulting interference spectrum. In OCT, there are a number of different approaches to collect interference spectra, where the majority are based around a Michelson interferometer, as shown in Figure 2.3.

![Figure 2.3: Depiction of a simplified free-space OCT system based around a Michelson interferometer.](image-url)
The first implementation of this technique was demonstrated by Dr. James Fujimoto’s group in 1991 and is often referred to as time-domain OCT (TD-OCT) [21]. In TD-OCT, a beam from a low coherence broadband light source is split using a Michelson interferometer into a reference and sample arm. The sample arm beam is incident on the object being measured and the reference arm path length is scanned. The resulting intensity of the interference spectra is detected by a simple photodetector. The intensity will be at a maximum when the sample and reference arm path lengths are perfectly matched. The challenge with TD-OCT is that image acquisition rate is limited by the speed at which the optical path length in the reference arm can be scanned. This is particularly challenging for ophthalmological imaging, as motion artifacts can be introduced if the subject moves even slightly over the course of the image acquisition. In addition, low sampling rate means that averaging cannot be used to smooth out noise, leading to a lower signal-to-noise ratio (SNR). [22], [23]

Spectral-domain OCT (SD-OCT) eliminates the need to sweep the optical path length in the reference arm by replacing the photodetector with a spectrometer. In comparison to TD-OCT, where interference is being taken as function of reference arm position, interference intensity is now being taken as a function of wavelength. This means that the interference for the entire light source bandwidth can be queried near simultaneously. The resulting interference spectra can then be Fourier transformed to extract the optical path length difference between $N$ interfaces in the sample arm with the reference arm. By eliminating the moving reference arm, the acquisition rate limitation is now based on how quickly interferograms can be collected. [22], [24]

The challenge with SD-OCT is that the speed is ultimately limited by the acquisition rate of the CMOS line camera and A/D conversion rates. In addition, the
spectrometer apparatus used to collect interferograms is costly, which limits the viability of implementing noise reduction techniques such as dual-matched detection. These issues have driven the development of the third and most recent implementation for OCT: swept source. Swept source OCT (SS-OCT) can be thought of as a combination of SD-OCT and TD-OCT, where we still collect an interferogram. However, we sample interference as a function of wavelength by sweeping the light source wavelength as a function of time [25]. The advantage of moving to SS-OCT comes down to acquisition rate, and the ability to move to a lower cost, higher speed photodiode detector (in comparison to a CMOS line camera/spectrometer). As such, in SS-OCT the limiting component for acquisition rate is no longer the detector, it is the lightsource sweep rate, which is controlled by high-speed Fabry-Perot filters. Commercially available swept source light sources are capable of sweep rates up to 3 MHz. Due to the low cost of photodiodes (especially in comparison to a spectrometer), it is economically viable to implement a dual matched detection scheme. This technique helps to eliminate noise in the signal, potentially improving signal quality. Unfortunately, the swept source light source is far more expensive that the combined cost of a broadband light source and spectrometer, as used in SD-OCT. As such, SD-OCT remains the more economic imaging system in comparison to SS-OCT. [26], [27]

2.4.2 Mathematical Basis for Low Coherence Interferometry

Let us start by representing the electric field resulting from our light source as a polychromatic plane wave propagating in time $t$, and space $z$. Due to the superposition principle, we can decompose the expression such that the electric field from a single
2.4. INLINE COHERENT IMAGING

2.4. INLINE COHERENT IMAGING

The frequency is given by

\[ E_{\text{source}}(\omega) = s(w)e^{i(\omega t - k(\omega)z)}. \]  

(2.1)

As we make use of a 50:50 beam splitter, we assume that light is evenly split down the reference arm and the sample arm. This gives

\[ E_{\text{ref}}(\omega) = E_{\text{source}}(\omega)\frac{1}{\sqrt{2}}r_r e^{i2k(\omega)z_r}, \]  

(2.2)

where the \( \sqrt{2} \) term comes from the fact the intensity is split, \( r_r \) is the reflectivity of the surface in the reference arm, and the exponential includes the phase shift from travelling down the reference arm path. To model the sample arm path, we are forced to account for the reflectivities of each interface \( r_{s1}, r_{s2}, \ldots, r_{sN} \) to the electric field, \( E_{\text{source}}(\omega) \).

\[ E_{\text{sample}}(\omega) = \frac{E_{\text{source}}(\omega)}{\sqrt{2}} \sum_{p=1}^{N} r_{sp} e^{i2k(\omega)z_{sp}} \]  

(2.3)

We can now superimpose the electric fields from the reference arm and the sample to model the electric field at the detector,

\[ E_{\text{det}}(\omega) = \frac{E_{\text{source}}(\omega)}{\sqrt{2}} r_r e^{i2k(\omega)z_r} + \frac{E_{\text{source}}(\omega)}{\sqrt{2}} \sum_{p=1}^{N} r_{sp} e^{i2k(\omega)z_{sp}}, \]  

(2.4)

for \( N \) interfaces in the sample arm, and a single interface in the reference arm.

Ultimately, the photocurrent generated by the detector is proportional to the light intensity rather than the electric field. The conversion from electric field to intensity is given by:
\[ I_{\text{det}}(\omega) = \frac{n c \epsilon_0}{2} |E_{\text{det}}(\omega)|^2 \]

\[ = S(\omega) \frac{n c \epsilon_0}{4} (R_r^2 + \sum_{p=1}^{N} R_{sp}^2 + \sum_{p=1}^{N} R_r R_{sp}(e^{i 2k(\omega) \Delta z_{r,sp}} + e^{-i 2k(\omega) \Delta z_{r,sp}}) ) + \sum_{p=1}^{N} \sum_{q=p+1}^{N-1} R_{sp} R_{sq}(e^{i 2k(\omega) \Delta z_{sp,sq}} + e^{-i 2k(\omega) \Delta z_{sp,sq}}) \]  

\[(2.5)\]

where \( n \) is the index of refractive index of the propagation media, \( c \) is the speed of light, \( \epsilon_0 \) is the permittivity of free space, and reflectivities of each interface are now denoted as \( R_{s1}, R_{s2}, ..., R_{sN} \) to denote that the reflectivites are now with regard to intensity (not electric field). By making use of Euler’s rule, and for convenience consider \( I_{\text{det}} \) as a function of \( k \) instead of \( \omega \), we are left with

\[ I_{\text{det}}(k) = S(k) \frac{n c \epsilon_0}{4} (R_r^2 + \sum_{p=1}^{N} R_{sp}^2 + \sum_{p=1}^{N} 2 R_r R_{sp} \cos(2k \Delta z_{r,sp})) + \sum_{p=1}^{N} \sum_{q=p+1}^{N-1} 2 R_{sp} R_{sq} \cos(2k \Delta z_{sp,sq})). \]

\[(2.6)\]

To determine the backscattered intensity as a function of \( \Delta z \) of 2.6, we take the Fourier transform, leaving us with

\[ I_{\text{det}}(z) = \frac{n c \epsilon_0}{4} \gamma(z) \otimes (R_r^2 + \sum_{p=1}^{N} R_{sp}^2 + \sum_{p=1}^{N} R_r R_{sp} \delta(z \pm \Delta z_{r,sp}))+ \sum_{p=1}^{N} \sum_{q=p+1}^{N-1} R_{sp} R_{sq} \delta(z \pm \Delta z_{sp,sq})). \]

\[(2.7)\]
Where $\gamma(z)$ is the Fourier transform of the light source spectrum, $\otimes$ represents the convolution operation, and $\delta$ represents the Dirac-Delta function. Equation 2.7 shows that there are three types of spectral contributions: steady state or DC resulting from the first term, interference between light travelling down the sample arm with light in the reference arm in the second term, and an autocorrelation contribution resulting from interference between light travelling down different optical paths in the third term. In an ideal system, the first term can be completely removed using a background subtraction, and the last term will typically have a negligible contribution due to diffuse scattering in the sample arm.

### 2.4.3 Dispersion

Dispersion mismatch occurs when optical media between the two arms in the interferometer are not exactly the same. Physically, dispersion effects are a result of the index of refraction of optical media, such as fused silica, having a wavelength dependence. This means that different wavelengths of light travel different optical path lengths. In OCT and ICI, this manifests first as a $\lambda$-dependant phase shift in the interferogram. After the interferogram is Fourier transformed, this results in a broadening of the peak, which limits axial resolution.

The most obvious solution to avoid dispersion mismatch is to ensure that the path-lengths in the interferometer contain identical optical media. This is not always practical in ICI, as one arm is integrated into a commercial laser material processing head. As such, approaches have been developed to compensate for dispersion in post-processing of the interferograms. \[28\]
2.4. SOURCES OF NOISE

Coherent imaging has unparalleled sensitivity and dynamic range, as detected signal scales with the electric field (not intensity) backscattered from the sample, and can be amplified in the optical domain by increasing the reference arm power. Unfortunately, coherent imaging also suffers from speckle, autocorrelations, and various sources of noise (such as detector and shot noise) [29]–[32]. Speckle noise is one of the most significant causes of image degradation in coherent imaging techniques. Significant time and energy has been devoted to reducing speckle using techniques ranging from averaging to machine learning. Speckle manifests as a reduction of “true” signal intensity due to destructive interference between the out and back paths in the sample arm [29], [33].

We can see this phenomenon in Figure 2.4(bottom) where I have imaged over a variably reflective surface. In 2.4(top left) I have manually segmented the region where I expect to see signal, and have recorded the pixel intensities in that region. Within the region I expect to see signal, I recorded intensities between 5 and 37 dB. Though it is possible that some of this variability is a result of light not coupling back into the sample arm, it is likely that much of this variability can be attributed to speckle.

In considering the intensity composition in a region I expect to see noise, I see that there are noise pixels with intensities up to 12 dB (shown in Figure 2.4(top right). This means in an imaging application where I do not have a predefined surface, the system will be blind any time the signal intensity goes below 12 dB (since the signal will be indistinguishable from noise). This is particularly problematic for applications that require low-latency feedback, as an otherwise “bright” surface can appear completely dark, rendering the system momentarily blind. In OCT, speckle noise can
Figure 2.4: Pixel intensity histograms from regions where signal (top left) and noise (top right) is expected in image taken while moving over a variably reflective surface (bottom). Region where signal is expected is highlighted by an orange line, and region where noise is expected is within the blue box on the bottom image.

be handled by averaging over acquisitions (due to the static nature of the subject being imaged). However, this is typically not possible when coherent imaging is implemented for monitoring laser processes. Due to the highly dynamic nature of laser welding processes, averaging would likely have the effect of blurring signal into the noise. Therefore, in coherent imaging for laser process imaging, we should anticipate some signal loss to speckle.

Industrial deployment of coherent imaging metrology often involves integration into laser cells with coatings that are either not optimised for the imaging light, or that degrade during processing. Even weak multiple reflections within one arm leads
to autocorrelations that can be misinterpreted as real interfaces [31]. Removal of these false interfaces can be done through image processing, but only if they remain static during laser irradiation. A better approach to removal is balanced detection, such as used in SS-OCT, where autocorrelation terms are automatically removed on-the-fly. Unfortunately, this technique requires a highly specialized (and expensive) laser for suitable line imaging rates (100 kHz or higher) and optical and detection components that are extremely well matched over 10s of nm bandwidth[34], [35]. Altal et al. implemented a balanced-detection approach in spectrometer-domain ICI with modest signal to noise improvements (2.13 dB) that avoided the need of a specialized light source, but required synchronized and balanced spectral measurements [36].

In addition to removing autocorrelations, matched detection is also useful for reducing noise related to the detector. This could include thermal noise from the silicon detector, or noise introduced during the analog to digital conversion. As only the signal will be constant between the two detectors, random noise from the detector can be filtered out.
Chapter 3

Pre-Existing Infrastructure

Prior to pursuing my own research, it was critical to understand the pre-existing infrastructure within the Fraser group lab. Our group has a laser material processing cell built and developed by previous group members such as Dr. Paul Webster, Dr. Cole Van Vlack, Chris Galbraith and others. This cell is capable of welding, cutting, and drilling among other processes. In addition, this cell is equipped with an inline coherent imaging system that is the precursor to the commercial LDD system produced by IPG Photonics Canada Ltd. While considering the existing physical infrastructure, it was also important to understand the software used to control the cell and imaging system (much of which was generated by past group members).

3.1 Processing Beamline

The processing laser is a multimode diode pumped ytterbium doped fibre laser made by IPG Photonics (IPG Photonics YLS-1000-IC), with a nominal maximum output power of 1100 W, a central wavelength of 1070 nm ± 1.5 nm. The output of the gain stage in the fibre laser is coupled into a 5 m long feeding fibre with a 50 µm core diameter. This fibre is then coupled to a 5 m delivery fibre with a 100 µm fibre core.
3.1. PROCESSING BEAMLINE

The delivery fibre is then coupled into the processing head (Laser Mech Accufibre PLYDH0209) through a 60 mm focal length watercooled collimator. At the output of the collimator, the beam travels in free space towards a dichroic mirror at a 45° angle. The dichroic is coated to reflect the 1070 nm process beam and to transmit the 840 nm imaging beam. The beam is reflected off the dichroic mirror down towards the focusing optic with a 150 mm focal length. Below this lens is a sacrificial cover glass to ensure that the focusing optic does not become contaminated. To further protect the cover glass, there is the option to blow compressed gas just below the cover glass, perpendicular to the optical axis. This creates another barrier to prevent smoke or spatter from contaminating the cover glass. We also have access to a second gas nozzle that can be used to create a gas flow over the melt. By blowing inert gas over the melt, we prevent oxidation when the melt is solidifying.

When the system was first commissioned in 2012, a beam profile was taken of the processing beam line using a Primes beam diagnostic system. At time of commissioning, the beam had a waist of 210 µm and a Rayleigh length of 3.7 mm. Since commissioning, Troy Allen has verified this beam waist, and found the waist to be 238 µm.

3.1.1 Motion Control

To move the workpiece relative to the processing beam, we make use of three Aerotech ball screw bearing linear motion stages. The maximum travel in the x-axis is 400 mm, whereas the maximum travel in the Y and Z axes is 300 mm. The maximum velocity for all three stages is 300 mm/s. Stages are driven by brushless motors, also supplied
3.2. INLINE COHERENT IMAGING

by Aerotech. The motion system is controlled by a computer using the A3200 software suite.

3.2 Inline Coherent Imaging

An inline coherent imaging system is integrated into the macro processing cell to allow for the measurement of laser welding processes such as welding, cutting, and drilling, as shown in Figure 3.1. This system is a spectrally-resolved interferometer. We can divide the system into four sections corresponding to each arm of the Michelson interferometer system is based around. The imaging beam is generated by a super luminescent diode (SLD), central wavelength of 840 nm and a 25 nm bandwidth. To prevent light from going back into the light source, the output is passed through an optical circulator. The output of the circulator then goes into a 50-50 fibre coupler, which splits the light into a sample arm and a reference arm.

In both the sample and reference arm, immediately after exiting the 50-50 beam splitter, the light is passed into a polarization controller. To ensure dispersion matching, light is then passed into one of two matched single-mode fibres. In the sample arm, the light is then coupled into a collimator attached to a 2-axis galvanometric scanner. This scanner is mounted to the welding head. The galvo mirrors are set such that the imaging beam is coaxially aligned with the processing beam. In the reference arm, the fibre is also coupled into a collimator. The output of this collimator is then directed down a free space path roughly equal to the optical path length in the sample arm. The light returning from the sample and reference arms is then recombined in the 50-50 fibre coupler. The combined light then passes out the fourth path from the beam splitter through a collimating lens and is directed to a diffraction grating. The
3.2. INLINE COHERENT IMAGING

Figure 3.1: Apparatus for inline coherent imaging integrated into a laser material processing cell. The system is split into four sections corresponding to the four arms in the Michelson interferometer created by the 50/50 fibre splitter. The spectrometer, beam delivery and reference arm are all external to the cell, whereas the sample arm shares a beam line with the processing laser beam.

diffracted light is then collected by a high-speed CMOS line camera. It is important to note that all fibre components are single-mode.

3.2.1 Data Processing

Spectrometer Calibration

When we collect an image from the line camera, we have intensity in arbitrary units as a function of pixel number. As the OCT equation requires intensity as a function
of $k$, we need a way of mapping pixel number to $\lambda$. This can be achieved by using the spectrometer to acquire a spectrum from a light-source with a well characterized spectrum. We typically choose to make use of an argon lamp to provide a calibration spectrum.

By matching a known argon spectrum, generated by the National Institute of Standards and Technology (NIST), to the collected spectrum we now have several pixel-wavelength pairs [37]. These pairs are then fit to a 4th order polynomial; we now have the mapping function between pixel and wavelength:

$$\lambda(p) = c_4 p^4 + c_3 p^3 + c_2 p^2 + c_1 p + c_0$$

(3.1)

where $p$ is the pixel number and $c_n$ are the fitting coefficients to a 4th order polynomial.

$\lambda$ to $z$ transform

In the mathematical discussion of OCT signals in Chapter 2, I refer to using a Fourier transform to convert the interferogram from $k$-space to $z$-space. Instead of using a fast Fourier transform (FFT), to achieve the mapping to $z$-space, we make use of a matrix multiplication technique, which will be referred to as a homodyne transform henceforth. This technique operates on a similar principle to that of a lock-in amplifier. [38]

The homodyne transform achieves this mapping using a pre-generated homodyne look-up table. This table is an $nxm$ matrix with dimensions corresponding to the number of wavelength samples, $n$, and the number of $z$ bins, $m$ (typically half the number of wavelength samples). To determine the $z$-range, we first must calculate
the field of view from the wavelength bin size, $\Delta \lambda$. The single-sided field of view is given by

$$FOV = \frac{\lambda_0^2}{4\Delta \lambda}$$

(3.2)

therefore the z range will be from 0 to the field of view value over $m$ samples.

To compensate for dispersion resulting from mismatched optical media between that sample and reference arm, we apply a $\lambda$-dependent phase correction to the interferogram. This is achieved by experimentally determining the ideal second and third dispersion coefficients, $d_2$ and $d_3$, respectively, and using those to dispersion correct the collected interferogram.

Since the interferogram is not symmetric about $k = 0$, we can expect the $z$-transformed spectrum to be complex. To account for the real and imaginary components, we generate one matrix using cos terms, and one using sin.

$$A_{n,m} = \cos\left(\frac{4\pi z_n}{\lambda_{d,m}}\right)$$

(3.3)

$$B_{n,m} = \sin\left(\frac{4\pi z_n}{\lambda_{d,m}}\right)$$

(3.4)

We then use matrix multiplication to multiply the collected (and background subtracted, if necessary) interferogram, $C$, which yields the complex vector $D$

$$D = A \cdot C + iB \cdot C$$

(3.5)

where by taking the magnitude of $D$, we will have intensity as a function of $z$ for all $z$ inputs into the homodyne matrix. Alternatively, we could chose to take the angle
3.2. INLINE COHERENT IMAGING

to find phase as a function of $z$.

The reason we choose to use matrix multiplication instead of an FFT is primarily because of speed. Though it may seem to be an outlandish claim that matrix multiplication would ever be faster than taking a FFT, for this application it is. If we choose to use a FFT, we are forced to interpolate our interferogram (which is linear in $\lambda$) such that it is linearly sampled in $k$ (where $k$ is inversely proportional to $\lambda$). By using the homodyne approach, we are able to avoid interpolating. On my personal computer, this leads to a speed improvement of 41% when comparing the homodyne approach with the interpolate then FFT approach (for an 896x448 homodyne matrix).

**Background Subtraction**

A background is taken by blocking the sample arm, and taking a number of acquisitions with the same integration time as will be used when data is acquired later. These acquisitions are then averaged to create an average background data set. This averaged data set is then subtracted off of all acquisitions prior to transform to $z$-space. Mathematically, the background corresponds to the first term in

$$I_{det}(k) = S(\omega)\frac{n\epsilon_0}{4}(R_r^2 + \sum_{p=1}^{n} R_{sp}^2 + \sum_{p=1}^{n} 2R_rR_{sp}\cos(2k\Delta z_{r,sp}))$$

$$+ \sum_{n=1}^{N} \sum_{m=n+1}^{N-1} 2R_{sn}R_{sm}\cos(2k\Delta z_{sn,sm})).$$

(3.6)

**Noise Floor Equalization**

After transforming the interferogram from $\lambda$-space to $z$-space, we are left with intensity in arbitrary units as a function of $z$ in units of distance. Prior to converting intensity to units of dB, we equalize to the noise floor. The reason we equalize to
the noise floor is to minimize the effect of signal roll-off as we move to a longer relative path length mismatch. As path length mismatch increases, the fringe spacing decreases, making fringes harder to detect due to limited spectral resolution. By assuming the noise floor is flat, and normalizing to the noise floor at a particular depth, we can somewhat reduce the negative effects of roll off. This is achieved by first subtracting the averaged background from \( N \) background acquisitions, and transforming the backgrounds to \( z \)-space. We then take the root mean square of the \( z \)-space transformed backgrounds. This yields the noise floor as a function of \( z \). To convert to dB, we take

\[
I_{dB}(z) = 20 \log \left( \frac{I_0(z)}{I_n(z)} \right) \tag{3.7}
\]

where \( I_{dB} \) is the intensity of the acquisition in dB, \( I_0 \) is the intensity of the acquisition in arbitrary units, \( I_n \) is the noise floor intensity at a given \( z \).

**Weighted average for sub-micron \( \Delta z \) identification**

In order to extract \( \Delta z \) from a point spread functions where we have intensity as a function of \( z \), the standard approach is relatively simple. We set a threshold (in dB), which is sufficiently high as to not misidentify noise as signal, and we take the \( z \) bin of the pixel that is the brightest in the acquisition (as we can typically assume the brightest pixel corresponds to the path length mismatch between the sample and reference arms). This technique yields a measurement of \( \Delta z \) to within half of a depth bin, which is determined by

\[
l_z = \frac{\lambda_0^2}{2 \Delta \lambda N_p} \tag{3.8}
\]

where \( N_p \) is the number of detector pixels, and \( l_z \) is the bin size. This equation can be derived by dividing the expression for the field of view given in Equation 3.2, by the
number of detector pixels over 2. For our system, this means our precision is limited to 5 µm when using the brightest pixel above threshold approach.

For certain applications this is not sufficient. One way to improve resolution is to increase ∆λ, but this is at the cost of a decreased field of view, and requires modification to the spectrometer. As sacrificing field of view for resolution is typically not desirable, Kanko (a former graduate student in the group) developed a weighted averaging technique to make sub-micron precision possible, without making changes to the spectrometer [39]. This technique approximates Gaussian fitting (a computationally intensive operation) on the pixels that make up the peak around the brightest pixel, yielding to an increase in precision of up to 600% (in comparison to 800% for Gaussian fitting). This represents a decrease in uncertainty from ±5 µm to ±0.8 µm. The weighted average value for ∆z, ∆z_{wa} is given by,

$$\Delta z_{wa} = \sum_{n=p_s}^{p_e} z(n)I(n) / \sum_{n=p_s}^{p_e} I(n)$$  \hspace{1cm} (3.9)

where \(p_s\) is the starting pixel for the weighted average, \(p_e\) is the final pixel, and \(z(n)\) and \(I(n)\) represent ∆z and intensity respectively for a given pixel value \(n\). [39]

Ultimately this technique is dependent on the number of pixels that contribute to the peak. As such, we see a decreased benefit for low intensity peaks, as noise begins to contribute to the weighted average more substantially than for bright peaks.
Chapter 4

Initial self-witnessing inline coherent imaging implementation

4.1 Introduction

To determine whether self-witnessing ICI SW-ICI could be used to filter imaging noise, I designed an apparatus that could be used in combination with an existing ICI apparatus. For the first prototype of a dual-path reference arm, I chose to use a free-space 50/50 splitter terminated with gold mirrors. I made use of off-the-shelf components so that the prototype could be built and tested quickly and at a low cost as a proof-of-concept.

4.2 Methods and Materials

4.2.1 Self-Witnessing Inline Coherent Imaging

In this experiment, I used two different reference arm configurations, both of which are shown in Figure 4.1(b). The single path reference arm is used to generate a baseline of typical ICI data. The dual path reference arm uses a free-space 50/50
4.2. METHODS AND MATERIALS

Figure 4.1: (a) Experimental apparatus for inline coherent imaging of laser welding. (b) Typical reference arm configuration using a single path (bottom) and modified reference arm configuration to capture correlated images separated by pre-set depth (top).

beam splitter to re-direct half of the light to a second gold mirror. The two reference arm paths in the dual path configuration have a known path length mismatch, given by $\Delta r$. Given the application, 90 $\mu$m was a convenient choice since it allowed the two peaks to be easily resolved. I have observed that for highly reflective surfaces (i.e. a polished gold mirror), I can expect a full width half maximum (FWHM) of as low as 13 $\mu$m, whereas for diffuse scattering workpiece (such as a rough stainless steel sample), a FWHM of up to 30 $\mu$m is not uncommon. By selecting 90 $\mu$m, for even a rough surface, this leaves at least 30 $\mu$m between the peaks. Choosing a mismatch that is an integer multiple of the depth bin size (10 $\mu$m) facilitated image processing that exploits correlation between pixel values and subsequently excludes uncorrelated noise and artifacts.
4.2.2 Laser Keyhole Welding

Spot welds were performed at an average power of 360 W over 10 ms on 316L stainless steel samples prepared from NIST standard reference material (SRM) 1155a. This material was selected as its thermophysical properties have been well-measured [40]. The beamline used to perform these welds is further detailed in Chapter 3.

4.3 System Characterization

4.3.1 Static Surface Imaging

To demonstrate self-witnessing ICI, I first imaged a stainless steel sample moving at a vertical speed of 3 mm/s. This was an intermediary step to show imaging over a large range of path length differences while ensuring no morphology changes in the sample. Figures 4.2(a) and 4.2(b) show ICI measurements of a moving interface for single arm ICI and self-witnessing ICI, respectively (with imaging light incident on the sample from above). One pixel represents the measured intensity of the light backscattered from one “depth bin” (each pixel is 9.98 µm, as dictated by the spectrometer specifications). For multiple measurements made sequentially, the x-axes represent time and the backscattered intensity is encoded as brightness (on a logarithmic scale). Zero on the y-axes correspond to the initial measured position of the interface. Due to the high temporal resolution of the data (2 µs), the 50,000 vertical lines have been downsampled to 5,000 using the method described in Chapter 3.

The most prominent feature in Figure 4.2(a) is the high-intensity peak moving in height (Δz) over time. This bright peak (more than 15 dB above the noise floor) corresponds to the position of the steel sample. Note that since the imaging light cannot penetrate into the metal interface, any signal from below this surface is due
to noise. Note that at this stage, no thresholding has been applied. This means that an intensity of 0 dB corresponds to the average intensity at a given depth bin in the background acquisition (more details on the background subtraction approach can be found in Chapter 3). With the dual arm configuration (Figure 4.2(b)), a second image overlays the first but offset by $\Delta r$ (90 $\mu$m) below the original interface. Any interface that arises from interference between the sample and reference arm will manifest as a “primary” and a “witness” peak. In both images, there is also a faint feature at 300 $\mu$m with an intensity fluctuating between 8 dB and 12 dB. This is caused by an autocorrelation in the sample arm from an imperfectly coated optic and has the potential to interfere with tracking interface positions. In standard ICI, these artifacts can be avoided by limiting the field of view (i.e., ensuring your sample never enters the $\Delta z$ range near the artifact) or by setting a high intensity threshold that will eliminate the artifact completely. However, limiting field of view is not possible for all applications and raising the intensity threshold will necessarily remove real signal.

4.3.2 Assessing the Correlation

The image shown in 4.2(b) demonstrates an image of relatively constant intensity moving through the field of view. To assess how well correlated the peaks are across measurements of varying intensity, I chose to image over a variably reflective surface, shown in 4.3(left).

By comparing the intensity of the nth and the n+9th pixel across an entire intensity plot, one immediately notices that there is a high-density clustering of points along the $y = x$ line, as expected. However, there are other clusters of points that
Figure 4.2: Standard (a) and unprocessed self-witnessing (b) ICI depth measurements of a static reflective stainless steel sample for a relative vertical scan speed of 3 mm/s. A horizontal streak is apparent at 300 µm originating from an autocorrelation in the sample arm due to imperfect optical coatings. The mismatch between the dual arm interfaces is set to approximately 90 µm (9 pixels). All pixels below 5 dB are set to black.
Figure 4.3: 2D histogram (right) showing the relationship between the intensity of the $N^{th}$ pixel, and the $N + 9^{th}$ across each A-line in the image on the left. This image was taken while moving over a variably reflective metallic surface, which led to the variation in intensity in the image. Note that pixel pairs in region b on the histogram correspond to primary/witness pairs, whereas pairs in section d correspond to noise. Pairs in section a and c correspond to either a primary, or witness pixel paired with noise.

do not follow the expected trend. To determine why this was, I first examined the $z$-transformed interference from a single A-line.

When I examined the two peaks in Figure 4.4, it is clear that the witness (right peak) is wider than the primary peak (left peak). Though this does not impact the brightest pixel in each of the primary and witness peaks, it will have a significant impact on the ability to find a pair of pixels of similar intensity separated by 9 pixels in the shoulders (as shown in Figure 4.4). To determine why the peak geometries are not matched, I must consider the optical paths in the dual-arm reference arm.

When comparing the amount of glass in each of the paths, the path that travels through the 50/50 clearly travels through more than the reflected path. To compensate for this, I added a fused-silica window, made of an identical material to the 50/50 splitter in the reflected path. This was done in an attempt to dispersion match the two paths, but the fused silica window did not have an anti-reflection coating.
Figure 4.4: Peak geometries for a primary and witness pair. Intensities are well matched at the peak, whereas pixels in the shoulders are less well matched despite both the peaks and shoulder pixels being compared having a separation of 9 pixels.

When examining the detected spectra of just the arm with the fused silica window, it seemed that the spectra was modulated as if an autocorrelation term was present. This resulted in a widening of the peak resulting from interference between the reflected path in the reference arm and the sample arm. The consequence of this was that the primary peak was narrower than the witness peak, which made matching pixels, particularly in the shoulders of each peak, far more challenging than if the peaks had a similar width.

Another issue I noticed early on during data collection was that the distance between the primary and witness peak varied as a function of time. By using the weighted average technique (detailed in Chapter 3) to determine a sub-depth bin precise location of the peak resulting from the interference between the two reference arm paths
on frames taken over 100 ms, I see a variation of $69 \pm 2$ hz with a standard deviation of about 312 nm as shown in Figure 4.5. The position was calculated by taking the weighted average about the maximum pixel value in the background (no sample arm data). I then took the average of all calculated positions and subtracted that from each individual position to calculate the variation from the mean for each point. There is clearly a slower drift in the $\Delta r$ variation that must be accounted for. The fast and slow variations in $\Delta r$ suggests that there is a degree of instability in the reference arm apparatus, likely due to vibrations on the optical table introduced by various electronics, in addition to air currents through the room. As the two mirrors are mounted orthogonally, it is unsurprising that the path length between the two is shifting. Though this is not ideal, 312 nm only represents 3 % of a depth bin (9.98 $\mu$m), therefore it should not significantly affect the ability to identify a witness peak.

Though the mismatched peak widths and path length instability are not ideal, the strong correlation along $y = x$ in Figure 4.6(b) is extremely promising in moving towards creating a novel way of distinguish signal from noise, without relying exclusively on thresholding well above the noise floor. To aid in our ability to separate signal from noise algorithmically, we first made a correlation plot for just a region without signal (Figure 4.6(a)), to help us better understand the noise. Traditional ICI noise suppression relies on excluding data below a specific threshold. This corresponds to a region to the left of a vertical line on Figure 4.6(a). Since intensity between primary and witness peaks are highly correlated, SW-ICI allows us to consider the average of the nth and n+9th pixel before thresholding. The graphical representation of this cutoff is a diagonal line as shown on Figure 4.6(a) for various thresholds. The associated pixel error rates and estimated line error rates are included in the inset. This
Figure 4.5: Variation in the position of the peak resulting from interference between the two arms in the reference arm taken over 0.15 s with a camera integration time of 2 $\mu$s. There is a standard deviation of $\pm 312$ nm on the $\Delta r$ variation. Frames are taken at 200 kHz with an integration time of 2 $\mu$s.

data suggests that I can set the threshold for SW-ICI to approximately 10 dB and achieve a similar estimated line error as that given by a 12 dB threshold in single arm ICI. Line error is further decreased by rejecting points outside of the correlation acceptance window.

The specifics of the filtering process required consideration of both the signal and noise, particularly in the intensity region where they overlap. As such, pixel pairs from Figure 4.3 where pixel pairs were segmented into noise (Figure 4.6(a)) and brightest pair in each line (Figure 4.6(b)). By taking the brightest pair in each line, I was able to exclude the pairs that are a result of the mismatched shoulders. Figure 4.6(a) shows pixel intensities for a region of a SW-ICI image that contains no interfaces. As expected, the pixels show no correlation. Many pixels that normally
4.3. SYSTEM CHARACTERIZATION

Figure 4.6: (a) Correlation between the nth and n+9th pixel for a region of noise taken from a SW-ICI image. Pixel error refers to the percentage of noise pixels above a given threshold (T). Line error refers to the calculated probability of camera lines having at least one noise pixel above T. (b) Correlation between primary peak intensity and witness peak intensity from a SW-ICI image taken from scanning a static steel sample. Data is only included for peaks believed to be caused by real signal (they are 9 pixels apart and exist on the expected position of the interface). (c) Difference between the peak intensities as a function of average peak intensity. The success rate represents the percentage of camera lines with a measurement that satisfies the window condition. Figure adapted from [1], using data processing scripts generated by Troy Allen.

would survive a threshold cutoff can be easily discarded since they do not correlate with another high intensity pixel. In effect, the chance of getting two bright noise pixels (from the tail of the distribution of Figure 4.8) exactly 9 pixels apart is unlikely. This uncorrelated behaviour is in sharp contrast to signal from a real interface (steel),
as captured in Figure 4.6(b) (for a horizontal scanning speed of 10 mm/s). In this idealized situation, every pixel corresponds to a “true” interface since I have included only maximum pixels from the two brightest peaks in every image line and excluded any peaks that are not 9 pixels apart or not in the location of the interface. Even though measured backscattered intensities from the steel interface varied by orders of magnitude, the primary and witness peak are highly correlated. To apply this technique to weld depth monitoring applications, where the location of the true signal among the noise is not known \textit{a priori}, I must determine an acceptance window that only considers potential primary and witness pixels that are close enough in intensity. The correlation acceptance window can be informed by the spread in peak intensity difference taken from this ideal case, re-examined in Figure 4.6(c). A large acceptance window will preserve all the pixels that correspond to interfaces (i.e., 100% “success rate”) but will be of little value in excluding noise. For the imaging illustrated in Figure 4.6(c) (which is a flattened interpretation of the data shown in Figure 4.6(b)), a ± 1 dB acceptance window preserves 70% of the interfaces. If the primary source of noise is multiplicative, then the acceptance window can be set to a constant dB value, determined by the magnitude of this multiplicative factor. Multiplicative noise would manifest as a flat distribution in Figure 4.6(c) which is clearly not the case for these results. I observed that the spread of the differences in peak intensities decrease for higher peak intensities. The observed shape corresponds to a noise model that follows the square root of peak intensity, suggesting that some applications might benefit from an acceptance window that adjusts based on the peak intensity level. A key consideration in selecting the acceptance window range is the sensitivity of the particular application to line errors (i.e., false detections) and lost interfaces. For
example, a low-latency closed-loop control application that uses sample height as a feedback parameter might require a narrow acceptance window since false depths may make the feedback process unstable.

4.4 Noise Filtering Using Witness Peak Identification

In an ideal scenario, where the peaks are perfectly matched, and separated by exactly an integer multiple of the depth bin size, it would be sufficient to compare the intensities at the nth and n+9th pixel. Unfortunately, due to the difference in peak geometries, and that the separation between the two peaks being slightly smaller than 9 depth bins, an alternate approach was derived to preserve as much of the shoulder as possible. One approach to preserve the shoulders would be to set a larger acceptance window. Unfortunately, this will result in preserving more noise. As such, I instead chose to take an approach where I smooth the peaks so that the primary and witness pixel in question fall within a smaller acceptance window.

To ensure the maximum intensity value present is preserved, I average the nth and the n-1th peak in the witness, and the n+9th and n+8th in the primary peak (as labeled in 4.7). If the average on the witness peak is within an acceptance window of the average on the primary peak, the measurement is preserved. This is achieved by saving the maximum intensity between the four highlighted pixels into the n+9th pixel. If the average does not fall within the acceptance window, the n+9th pixel will be set to zero. As all pixels in the witness peak will ultimately be discarded (as the witness does not have a witness), this ensures that the maximum intensity, whether it is present in the witness or the primary will be preserved. In the case of the primary and witness peak shown in 4.7, the pixel with the greatest intensity is n+9th pixel.
Figure 4.7: Intensity as a function of depth bin (*) with a Gaussian fit of the peaks overlaid as a dashed line. 4 pixels are highlighted to demonstrate the difference in pixel intensities in the primary and witness peaks.

and the nth pixel is within 1 dB, therefore the original approach of comparing the nth and n+9th approach would be effective. Unfortunately, this is not always the case due to the instabilities (as shown in Figure 4.5) in the reference arm. As such this smoothing approach helps to ensure I preserve real measurements while decreasing the prevalence of noise pixels in measurements with less well matched primary and witness peaks.

4.5 Results

Using the approach detailed in 4.4 on Figure 4.2(b), yields Figure 4.8(bottom left). The most obvious result is that the “double image” is reduced to a single one, since the lower witness peak has no witness of its own. The autocorrelation artifact at 300 µm has also been completed removed, as well as much of the noise. This is made
more clear in Figure 4.8 (right) where I can see that I have eliminated a significant amount of noise, particularly in the tail region. This is well supported by the results shown in Figure 4.6(a), which shows that, for example, only 0.0004% of correlated pixel pairs will average above 9 dB (in comparison to 0.04% of pixels in a standard ICI configuration). While these values are small and the improvement may seem inconsequential, the resulting line error probability improves from 16% to 0.18%. As expected, this method has the greatest impact on eliminating noise in the tail region. Since the noise tail is what determines where I must set the threshold, this is a promising result.

![Figure 4.8: Histogram comparing noise intensities for image taken using standard ICI (top left) and image taken using the dual arm reference arm after processing using the self-witnessing algorithm with a ±2 dB acceptance window (bottom left). Noise histogram corresponds to a 200000 by 200 pixel window chosen to exclude the moving interface in both images.](image)

The interface in the processed image also appears to be higher resolution than the original. Though the peak heights of the primary and the witness are highly correlated, the variations of the shoulders of the two point-spread functions were found to be larger. As a result, the SW-ICI approach tends to suppress shoulders. This should not be interpreted as an improvement in axial resolution (how close
two interfaces can be and still be distinguished). SW-ICI is appropriate for “sparse” images where backscattered light comes primarily from one depth, such as for a metal or other opaque material. Semi-transparent materials could lead to interfaces that are $\Delta r$ apart causing primary and witness peaks to overlap, making processing difficult. Increasing $\Delta r$ would avoid this problem, dependent on the material being imaged.

4.5.1 Weld Imaging

To test the performance of SW-ICI on a highly dynamic application, I used the technique to image spot welds on 316L stainless steel. Figure 4.9(a) shows SW-ICI data collected from a 10.4 ms spot weld with average power of 360 W (temporal resolution $10 \mu s$). After the laser is turned on at 0 ms, I can see that the correlated interfaces track downwards relatively smoothly, indicating that a keyhole is forming. After about 1 ms, the smooth growth of the keyhole transitions into slower, fluctuating growth. This behavior is consistent with previous observations of the initial stages of keyhole formation [8], [15], [17]. After the initial formation, gaps in the image are observed likely arising from speckle and tilted keyhole base, impacted by keyhole instabilities such as keyhole collapse or sidewall protrusions. These gaps would be even more prominent if the threshold level was set to avoid noise (e.g., 12 dB as used in typical ICI image processing). The amount of noise as shown in Figure 4.9(a) (with 5 dB thresholding) is unacceptable for interface tracking. The SV-ICI algorithm effectively removes most of this noise without needing to resort to harsh thresholding (Figure 4.9(b), acceptance window of $\pm 3$ dB). It is apparent that I have not eliminated all noise using this method, as there are still pixels remaining that are far below the expected position of the interface. However, most of the higher
intensity noise pixels have been removed (similar to Figure 4.8 inset), allowing a lower threshold to be used. The particular threshold value can be chosen based on the maximum line error rate permitted by the application.

Figure 4.9: ICI depth measurements for a single spot weld performed on 316L stainless steel samples using a laser power of 360 W. Unprocessed SW-ICI data is shown in (a), and data processed using an acceptance window of ±3 dB is shown in (b).

I have shown that SW-ICI has the ability to suppress high-intensity noise without sacrificing low-intensity signal, suggesting that signal can be salvaged from a previously inaccessible intensity range. To quantify this, I compare the number of lines in Figure 4.9(b) that have a brightest pixel above 12 dB (orange), to those with intensities between 9 and 12 dB (blue) (Figure 4.10). A typical threshold in standard ICI is
4.6 Conclusion

Though the results of this first attempt at SW-ICI are promising, significant improvements can be made by improving both how stable, and how well-matched the peaks are. The instability meant that it was challenging to predict the location of witness
peak with greater precision than a depth bin. This means that the acceptance windows need to be set larger, limiting the amount of noise I can expect to exclude. Due to the poorly matched peak geometries, pixels in the shoulder of the peak are nearly always discarded. Though I do not necessarily need the shoulder pixels for applications like closed loop control, the resulting image looks more sparse as I am throwing away pixels that are a result of interference between the reference and sample arm.
Chapter 5

Stability improvements to self-witnessing coherent imaging

5.1 Introduction

The first implementation of a self-witnessing coherent imaging described in Chapter 4 showed promise as an effective approach for distinguishing between signal and noise, apart from traditional thresholding. That being said, the first implementation of this approach suffered from challenges such as difficult optical alignment, and more significantly, a lack of stability in the path lengths of the two arms as discussed in Chapter 4. To simplify alignment and improve stability, I designed a mirror with a 100 µm step. The step size was chosen to be similar to that used in the previous implementation of this technique (detailed in Chapter 4). By splitting the beam over the step, I have created a two arm reference arm on a single optic, without the need for dispersion compensation between the arms, and reducing the number of optics to be aligned by half. In addition, since both reference arms are on the same optic, any vibrations in one path in the reference will apply to both, which was not true of the previous implementation.
5.2 Design

To create a step edge optic, two different manufacturing approaches were considered: subtractive manufacturing and additive manufacturing. In choosing a manufacturing technique it was important that the final surface be optical quality, such that the light reflected from the optic could couple back into the single mode fiber. In addition the final surface should be highly reflective to minimize losses in the reference arm. Finally, the sidewall on the step needed to be as close to perpendicular to the substrate as possible to avoid coupling light from the edge of the step back into the single mode fiber.

The subtractive approach would require the removal of about 100 µm of silicon to create a step in the silicon, and then depositing a gold surface over the step. The advantage of using the subtractive approach is that the material on either side of the step will have the same material composition. The challenge is that any technique we could use to remove 100 µm of material (laser ablation, deep reactive ion etching, etc...) would leave the processed surface rougher than the unprocessed surface. This is an issue, as this application requires both sides of the step to have an optical grade surface. Due to the depth of the step, any material process would likely be prohibitively slow, and therefore costly, so I looked to an additive approach.

The additive approach required creating an approximately 100 µm step on a silicon wafer, and then depositing a gold surface over the edge. Due to the relatively large height required, neither chemical or physical vapour deposition would be practical. This left the option of bonding a second substrate that was 100 µm thick to the silicon wafer (a challenging process) or coating the substrate in photoresist, and exposing the areas that needed to be raised. Finding a substrate of the correct thickness
was challenging, so I ultimately chose to use SU-8 photoresist to create my step. Coating the wafer in photoresist is achieved by spincoating, where the final thickness is determined by a number of factors including the speed the wafer is spun at, and how the SU-8 is initially deposited on the wafer. The latter is not strictly controlled, which means there is a large degree of uncertainty in the final thickness of the photoresist. A study run by Dr. Graham Gibson at NFK on spincoating SU-8 estimated that for a 100 $\mu$m thickness, I could expect an error of $\pm$ 20 $\mu$m.

5.3 Fabrication

In order to create step edge mirrors, the basic process shown in Figure 5.1 was used.

![Process flow diagram](image_url)

Figure 5.1: Process flow used to generate the first production run of step edge mirrors.
5.3.1 First Production Run

Dr. Graham Gibson at Nanofab Kingston was instrumental in developing the process for manufacturing the step edge optics. The following process was tested by his team prior to my use of the process to create my optics.

Starting with a 4” polished silicon wafer (intrinsically doped in the 100 lattice orientation), I first cleaned the wafer with acetone, then isopropanol, then de-ionized water. I then plasma cleaned the wafer at 700 W for 60 seconds to remove any remaining organic material. After plasma cleaning, the wafer was baked at 200°C for 5 minutes to ensure the wafer was entirely dehydrated.

To create a substrate for the raised step, I then applied a layer of SU-8 to the silicon wafer. This was achieved by spin coating. After applying the liquid SU-8, the SU-8 was spread at 500 rpm for 5 s with 100 rpm/s acceleration, and then spun at 2200 rpm for 30 s with 300 rpm/s acceleration. After spin coating, the wafer was baked at 65°C for 5 minutes and then 95°C for 15 minutes.

After baking, I used maskless photolithography (Intelligent Micro-Patterning SF-100 XPress) to expose the step edge design. On the 4” wafer, there was enough room to expose the step in four different locations to create four optics after cleaving the wafer. The design was exposed using 365 nm light, with a 2.5 second exposure time. After exposure, the wafer was baked at 65°C for 3 minutes, and then at 95°C for 9 minutes. After 9 minutes, the hotplate was turned off, and the wafer was left on the hot plate to cool for 30 minutes to avoid warping or cracking in the SU-8.

To develop the sample, the wafer was placed in SU-8 developer for 9 minutes, while agitating continuously. After developing, the substrate was hard baked at 150°C for
10 minutes, and then allowed to cool on the hotplate for 40 minutes to prevent cracking or distortion.

At this point in the fabrication process, four raised 1” by 1/2” SU-8 rectangles with a height of 93 $\mu$m (as determined by profilometry) were present on the silicon wafer. I then used electron beam evaporation to deposit a 5 nm chromium adhesion layer to help the final gold layer adhere to the silicon (neither chromium or gold adhere well to SU-8, so this step was purely to improve the adhesion of the gold to the silicon). Finally, electron beam evaporation was used to deposit 100 nm of gold.

This process yielded four identical step edge mirrors. One concern with depositing the metal layer on SU-8 was adhesion between the two materials. SU-8 is typically used as a photoresist, and therefore the ability to adhere metal to the surface is not a design target of this material. Immediately after manufacturing, there were no signs of delamination. However, after transporting the parts from NFK to Queen’s, the corners of the mirrors were clearly starting to peel. Fortunately, as the mirrors were quite large, the beam could be aimed at regions on the mirror that were not experiencing delamination.

5.3.2 Second Production Run

After taking preliminary measurements with the mirrors from the first production run, I had the opportunity to send mirrors to be used to image additive manufacturing at the Diamond Research complex at Harwell (working under Dr. Peter Lee, University College of London). This required more mirrors to be manufactured, so I took this opportunity to try to resolve the delamination between the SU-8 and the metallic layers. Materese et. al attempted to solve exactly this problem when attempting to
use SU-8 instead of nickel as an adhesion layer due to the biocompatibility of SU-8 [41]. They found that performing the post-exposure bake after the gold had been deposited yielded far better adhesion to the SU-8 than performing the post-exposure bake before gold deposition as is typical.

For this second run, I also sought to lower the manufacturing cost for these optics. Firstly, I found that the 1”x1” mirrors were far larger than necessary. When fabricating the 1” mirrors, I was only able to fit 4 optics on the 100 mm wafer. For the second run, I made 10 mm by 10 mm mirrors (with a 10 mm by 5 mm raised step), which allowed me to make 10 optics on the same 100 mm wafer.

In addition to decreasing the size of the optics, I also created a photomask, significantly reducing the time and cost of exposing the wafer using maskless photolithography. The mask was fabricated by cutting the step geometries into an opaque cling film using a UV laser cutter. The cut film was then mounted on a glass slide to be used on the masked photolithography system.

In an effort to replicate the process used by Materese et. al, I deposited the gold layer immediately after removing the wafer from the SU-8 developer, without a chromium adhesion layer. After depositing 100 nm of gold, I transferred the wafer to a hotplate set at 100 °C for 15 minutes, followed by a 40 minute cool down on the powered-off hotplate to allow the SU-8 to fully cure. The finished wafer is shown in Figure 5.2. I immediately noticed that the adhesion of the gold to the silicon wafer was far worse than when the chromium adhesion layer is present. Unfortunately, this approach actually made adhesion to the SU-8 layer worse (than with the chromium adhesion layer), leading to complete delamination of the gold layer before the new mirrors could be used. As such, all imaging detailed in this chapter was performed using the
5.4. APPARATUS

first run of optics.

Figure 5.2: Silicon wafer with 10 10mm by 5mm raised SU-8 rectangles with a height of 84 $\mu$m (determined using profilometry) deposited with 100 nm of gold to create an optical grade surface. After dicing, this wafer yielded 12 step edge optics.

5.4 Apparatus

In the first apparatus incorporating the step edge in the reference arm, I used a plano-convex lens with a relatively short focal length (50 mm). We typically use a focusing lens in single-arm ICI to ensure that light reflected from the terminating mirror is coupled into the single-mode fibre. Without this lens, the system would be less robust to rotational motion is the terminating mirror (as long as the mirror is at the beam waist).

During alignment of the reference, we typically block the sample arm to ensure the only light incident on the detector is from the reference arm. After aligning one side of the step, and confirming that I could saturate the detector, I used a stage to move the mirror such that the focused beam was split on the step. As the beam moved
5.4. APPARATUS

over the step, I noticed a significant drop in intensity across all of the pixels, with a minimum reached when the beam was exactly split over the edge. Upon crossing the edge to the other side of the step, the detector was once again saturating. This drop in intensity was troubling, as we make use of the intensity in the reference arm to amplify the (often weak) back-scattered intensity in the sample arm, as shown in Equation 2.7.

A possible reason for this loss of intensity when the beam is on the step is that the step has about a 1° sidewall from the photoexposure. When making use of a small focal length, a larger proportion of the spot is incident on the step, not on the material on either side of the step. It is also possible that this loss has to do with diffraction of the light backscattered from the step-edge optic. An undergraduate thesis student, Alexander Wainwright, is currently in the process of modelling diffraction due to the step to better understand the optical effects of this type of optic.

In an effort to improve signal, I replaced the 50 mm focal length lens with a 750 mm lens. In both cases, the lenses were placed such that the terminating mirror was at the focal length of the lens. By moving to a longer focal length lens, the focused spot on the step is larger, and therefore the proportion of the beam on the step in comparison to the beam hitting either side of the step is much smaller. This improved coupling back into the single mode fiber to the point I could once again saturate the detector, even when the beam is evenly split over the edge.
5.5 Data Processing

5.5.1 Homodyne Interpolation

Due to the discrete nature of the point spread function, I needed to develop a technique to characterize the actual peak associated with interference, not just the highest pixel value. To determine a more precise centre location of the peak generated by interference between the two reference arm paths (separated by $\Delta r$), I used the weighted averaging technique detailed in Chapter 3. Though this technique yields the centre position, it does not provide the intensity (or phase) at the centre of the peak. To determine the intensity and phase, I generated a synthetic interferogram for the $\Delta z$ value determined using the weighted average. By matrix multiplying the synthetic interferogram by the real interferogram, I can determine the intensity and phase at
Δz. The intensity corresponds to the absolute value of the product, and the phase comes from the angle between the real and imaginary components.

The synthetic interferogram, $I_s(Δz)$ is given by:

$$I_s(Δz) = \exp(ikΔz),$$

where $k$ is the wave number corresponding to the wavelength at each pixel on the detector, $i$ is $\sqrt{-1}$, the subscript $s$ denotes that the interferogram is synthetic, and $Δz$ is now a single value of path length mismatch that I am choosing to query. After matrix multiplication, the absolute value can be taken to yield the intensity, and the angle can be taken to find the phase for a specific $Δz$.

Using this technique will allow for better identification of primary/witness pairs by precisely determining the intensity and phase for a specific $Δz$.

### 5.5.2 Intensity Filtering Approach

The intensity approach is the most similar to that used in the previous chapter for the first implementation of SW-ICI. In this approach, I first must determine the position of the peak resulting from interference between the two paths of the reference arm. To achieve this, I first used the Matlab peak finding algorithm to identify the rough position (precise to the depth bin size of 9.98 µm). Once the approximate position was determined, I made use of a weighted averaging technique (detailed in Chapter 3) to find the center of the peak with a sub-depth bin level of precision.

Once the position of this peak is identified, this number is recorded as the expected distance between two measurements of similar intensity if the intensity in the point spread function for a given $Δz$ is a result of interference between the reference and
sample arm. To identify intensity pairs, I also need to identify how closely the intensities between the primary and witness pixel should be expected to match.

In an ideal apparatus, the intensity would always be split perfectly over either side of the step. However, the stage on which the step edge was mounted did not have sufficient precision. Instead of changing the stage, which would have involved handling the very delicate mirror, I chose to handle this issue during post-processing. To determine the intensity offset between a primary and witness measurement, I used data from a bright surface, just prior to either the laser turning on (in the case of weld imaging), or just before stage motion is initiated (for surface imaging). I then compare the intensity of a primary peak (resulting from interference between the sample and reference arm), and the intensity at the witness located $\Delta r$ below, I now have an intensity offset between the pair of measurements. By adding this offset to the lower intensity of the two measurements, the intensity between the primary and witness measurement should be matched, and we can now filter for matched intensity measurements separated by $\Delta r$ when distinguishing between signal and noise.

### 5.5.3 Phase Filtering Approach

In coherent imaging, we typically use intensity as a function of $\Delta z$ to create our images. By relying on only the intensity (given by the absolute value of the $k$-$z$ transformed interferogram), we are completely ignoring the complex nature of our signal. By taking the phase instead of the absolute value, we have another metric by which we can differentiate signal from noise.

There are particular challenges when it comes to exploiting phase instead of intensity. Phase (even on a peak resulting from real signal) varies greatly depending on exactly
where in $\Delta z$ I choose to sample. If we consider that our detector has a central wavelength of 840 nm, this corresponds to a wave number of $7.48 \, \mu m^{-1}$. The interference spectra from a single interface with a path length mismatch of $\Delta z$ is given by

$$\cos(2k\Delta z),$$

(5.2)

where if we consider only the wavenumber corresponding to the central wavelength, this gives a period of 0.42 $\mu m$. When considering two value of $\Delta z$ separated by even 1 $\mu m$, this is equivalent to a phase shift of nearly 15 rads within the cosine function. As a result of this phenomenon, it becomes impractical to take the phase at $z(n)$, where $n$ corresponds to an integer valued depth bin, and $z(n) + \Delta r$ as was done when considering intensity. Instead, I choose to take a weighted average (detailed in Chapter 3) about $z(n)$ to yield the centre of the peak (if one exists). I then used homodyne interpolation to determine the phase at this location. This process is repeated at the location $\Delta r$ above the centre of the first peak.

### 5.6 Characterization

In an effort to characterize the performance of the new reference arm incorporating the step edge optic, I first wanted to determine how the interference spectra resulting from the two paths in the reference arm changed as a function of time. This was achieved by taking interference spectra in $k$–space and transforming into $z$–space and analysing the peak resulting from the interference between the two paths.

To ensure that the path length distance between the two sides of the step was stable, I took an acquisition over 0.45 seconds at 200 kHz with an integration time of 2 $\mu$s.
5.6. CHARACTERIZATION

Figure 5.4: Variation in the position of the peak resulting from interference between the two arms in the reference arm taken over 0.45 s. There is a standard deviation of ±26 nm on the Δr variation.

This was performed with the sample arm blocked, so that the highest peak in the \( k-z \) transformed image would correspond to the path length mismatch between the two sides of the step. By using a weighted average about the brightest pixel in each A-line, I could identify path length mismatch down to sub-micron precision. I then plotted the variation from the mean path length value, as shown in 5.4. The standard deviation for the path length mismatch is 26 nm (a significant improvement from the previous apparatus with a standard deviation of 312 nm). If we treat the standard deviation of the path length mismatch as the uncertainty, we then have a path length mismatch of 93.29 ± 0.03 \( \mu \)m.

After determining that the measured optical path length difference between the two reference arm paths (\( \Delta r \)) was relatively stable, I then examined the resulting
sample/reference arm interference. I initially imaged over a variably reflective steel surface at a constant velocity (5.5(a)). Using the resulting image, I generated a plot where the intensity of the $z(n)$ pixel is plotted on the x-axis, and the $z(n) + \Delta r$ (calculated using homodyne interpolation) is plotted on the y-axis. The resulting 2D histogram is shown in 5.5(b). In comparing this correlation plot to that previously attempted using the 50/50 approach we have made significant improvements. Pixels above 20 dB have an average correlation factor of 0.9925 when comparing the intensity at $z(n)$, and the intensity at $z(n) + \Delta r$.

Due to the improved stability between the paths in the reference arm, I also chose to characterize the phase difference between pixels separated by $\Delta r$ for both signal and noise. To perform this characterization, I opted to image a gold mirror moving vertically at a constant velocity (3 mm/s) relative to the beam delivery optics (shown in Figure 5.6(a)). This would provide a pair of constant-intensity peaks moving for $\Delta z$ values varying over 750 $\mu$m. I then isolated the peaks corresponding to the top and bottom line in each A-line. As the phase is extremely sensitive to small changes in $\Delta z$, I opted to use a weighted average to find a more precise centre for the location of the upper peak. Next, I used homodyne interpolation to determine the phase at the center of the upper peak. I then subtracted $\Delta r$ from the location of the upper peak, and used homodyne interpolation at this new location to determine the phase of the lower peak. To deal with the phase wrapping, I took the cosine, and then inverse cosine of the difference between the two phases. I then plotted a 2D histogram of the difference in phase as a function of $\Delta z$ (shown in Figure 5.6(b)), and found that the peaks were effectively phase locked with a value of $2.01 \pm 0.01$ radians. The small variation away from 2.01 radians implies that it should be possible to exploit phase
Figure 5.5: 2D histogram (b) of intensities at \( z(n) \) and \( z(n) + \Delta r \) where \( n \) is an integer that corresponds to depth bins in the image of a variably reflective metallic surface shown in (a).
Figure 5.6: 2D histogram (b) of phase difference between the top and bottom peak in (a) as a function of $\Delta z$. A 2D histogram of the phase difference for noise pixels created using the same process as (b) is shown in (c). correlation in addition to the intensity correlation to further reduce noise. When this process was repeated on noise pixels separated by $\Delta r$, it is clear that the phase difference is completely random (shown in Figure 5.6(c)).

To ensure that intensity and phase will remain well correlated for keyhole welding data, I repeated the above analysis on an image taken during a spot weld on 1000 series aluminum using 1100 W over 40 ms. Figure 5.7(a) show 30 ms of this welding process. The laser turned on 5 ms before 0 ms in Figure 5.7(a), and the laser turned
off 5 ms after 30 ms in Figure 5.7(a). In this image, I repeated the phase analysis performed on Figure 5.6(a), for peak pairs in Figure 5.7(a) over 12 dB. The resulting 2D histogram is shown in Figure 5.7(b).

I also repeated this analysis on low intensity peaks resulting from noise, shown in 5.7(c), and found that the phase difference between the pairs did not follow an observable trend. Finally, I repeated the process for generating 2D Histograms comparing peak intensities separated by $\Delta r$, as done in Figure 5.5(b), for the welding data. The resulting image is shown in 5.7(d). It is clear that when comparing the phase histograms for signal from a moving mirror in 5.6(b) with that from welding data in 5.7(b), that the grouping is far tighter for the former than the latter. In Figure 5.7(b), the phase difference between peaks, after handling phase wrapping, was $-0.4 \pm 0.4$ rads. The large variation (especially in comparison to the variation from Figure 5.6(b)) is likely because the welding signal is of far lower intensity than that of the gold mirror. As such, the phase of the noise will have a greater contribution on the calculated phase of the signal. To avoid discarding real signal during phase filtering (without setting the phase acceptance window so large to negate the effect of filtering), I chose to preserve any pixel that survives intensity filtering with an intensity above 12 dB, regardless of its phase relationship with its witness. By only applying phase filtering to lower intensity pixels, any signal I misidentify as noise would have been discarded anyway using traditional thresholding approaches. Any low intensity signal that I salvage then is only adding to the effective imaging rate, while preserving higher intensity signal that might not survive phase filtering.

For pixels above 12 dB (which can be assumed to be signal), the standard deviation
5.6. CHARACTERIZATION

Figure 5.7: (a) shows keyhole data from a spot weld performed at 1100 W on 1000 series aluminum where the laser turned on 5 ms prior to 0 ms on the x-axis. 2D histograms for intensity and phase of welding data 2D histogram in (b) take the difference in phase between signal above 12 dB at $\Delta z$ and $\Delta z + \Delta r$ as a function of $\Delta z$. The same plot is shown for a section of (a) where only noise is present is shown in (c). A 2D histogram (d) comparing intensity in (a) at $z(n)$ pixel, and the intensity at $z(n) + \Delta r$ Data in (a).

on the intensity difference between the peaks is 1.5 dB, and 0.4 dB in the phase difference. This is calculated using pixels above 12 dB in Figure 5.7(d) for the intensity value, and using all pixels in Figure 5.7(b) for the phase value. As such, these values will be used as the acceptance window for filtering welding data moving forward. It is also worth noting that the cluster of points calculated by taking the phase difference between a primary and witness is different between 5.6(b) and 5.7(b) as the welding data was taken days after the mirror imaging. Between data collections, I had to realign the step edge (due to multiple users on the reference arm beam line), and
found that the height of the step increased by $6.45 \mu m$ after realignment. This was likely because the reference arm beam was now on a location seeing more delamination than the location it was on previously. $6.45 \mu m$ corresponds to a shift of about 2.5 radians.

The overall improvement in reference arm stability means that the relationship between a primary and witness pixel can now be characterized by both intensity and phase, even for welding data. This should increase my ability to eliminate noise pixels, hopefully allowing for the tracking of lower intensity signal measurements.

5.7 Exploiting the Improved Correlation

When comparing the two filtering approaches, intensity and phase, they tend to filter out different levels of noise to varying degrees. Intensity filtering seems to be the most effective for filtering out high intensity noise (10-12 dB), which is unsurprising, as the preservation of a high intensity noise pixel requires the presence of two of the pixels, separated by $\Delta r$. This is ultimately an unlikely occurrence. As such, intensity filtering struggles to minimize lower intensity noise (0-10 dB), due to the statistical probability of two noise pixels separated by $\Delta r$ having a similar intensity. The phase filtering approach does not favour high intensity pixels in the way that intensity filtering does, it filters relatively evenly across all noise intensities. As phase is ultimately random (and independent of intensity), it is just as likely that a low intensity pixel has a phase pair, as it is for a higher intensity pixel.
5.7. EXPLOITING THE IMPROVED CORRELATION

Figure 5.8: Histogram showing noise composition by intensity for data processed just using the intensity filter (with acceptance window of 1.5) in light blue, just the phase filter (with an acceptance window of 0.05 radians) in orange, and with both filters applied in dark blue. The call out demonstrates the noise composition at the tail. This histogram is representative of 2.5 million noise pixels. Note that the sum of all of the counts will not add to 2.5 million due to pixels discarded through filtering.

Due to the differences in the performance of the two filters, it is advantageous to apply both. It is far less likely that a pair of noise pixels will satisfy both the intensity and the phase requirement, significantly reducing the prevalence of noise.

To demonstrate the effect of different filters on pixels that are likely resultant from noise, I passed these segments through each filter, and plotted a histogram of intensities for the resulting image (shown in Figure 5.8). The image segment I chose is comprised of 2.5 million pixels. The representation of each noise intensity is shown in Table 5.1, compared to the representation of each noise intensity in unfiltered noise.

By applying both filters, I decrease the prevalence of noise pixels in the 9-12 dB range by nearly 50 times, and the prevalence of noise pixels in the 6-9 dB range by over 70 times, than when no filter is used. In comparing the results of using both filters in
Table 5.1: Table showing the percent composition for noise plotted in the histogram binned into groups of three ranging from 0 to 12 dB. Composition is given for unfiltered data (not shown in the histogram), intensity filtered data (light blue in the histogram), phase filtered data (orange in the histogram), and data filtered using both metrics (light blue in the histogram). This composition was taken from 2.5 million noise pixels.

<table>
<thead>
<tr>
<th>Intensity (dB)</th>
<th>No Filter</th>
<th>Intensity Filter</th>
<th>Phase Filter</th>
<th>Both Filters</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤0</td>
<td>58.088 %</td>
<td>93.832 %</td>
<td>98.968 %</td>
<td>99.672 %</td>
</tr>
<tr>
<td>0-3</td>
<td>24.272 %</td>
<td>3.376 %</td>
<td>0.533 %</td>
<td>0.159 %</td>
</tr>
<tr>
<td>3-6</td>
<td>14.509 %</td>
<td>2.242 %</td>
<td>0.372 %</td>
<td>0.125 %</td>
</tr>
<tr>
<td>6-9</td>
<td>3.034 %</td>
<td>0.534 %</td>
<td>0.121 %</td>
<td>0.042 %</td>
</tr>
<tr>
<td>9-12</td>
<td>0.098 %</td>
<td>0.016 %</td>
<td>0.006 %</td>
<td>0.002 %</td>
</tr>
</tbody>
</table>

comparison to just intensity filtering (as was done in Chapter 4), I see 8 times less pixels in the 9-12 dB range, and nearly 13 times less pixels in the 6-9 dB range.

This significant reduction in the prevalence of high intensity noise pixels should make it possible to drop the feature identification threshold even lower than was possible when just using intensity filtering.

5.8 Weld Imaging

To assess the performance of both filters, I imaged 10 ms spot welds on 1000 series aluminum performed on the macro-machining cell beam line at 1100 W. During these welds, imaging was performed using the step edge optic. When comparing the unprocessed image in Figure 5.9(a), and the image that is only intensity filtered in Figure 5.9(b), there is an obvious reduction in the noise intensity throughout the image. There is a further reduction in noise intensity when comparing the image filtered using both phase and intensity, shown in Figure 5.9(c) and that with just intensity filtering in Figure 5.9(b). This further decrease made it possible to decrease the image
Figure 5.9: ICI depth measurements for a single spot weld performed on 1000 series aluminum samples using a laser power of 1100 W. Laser turned on at 0 ms with the keyhole forming about 0.5 ms later. Unprocessed SW-ICI data is shown in (a), data processed using just intensity filtering with an acceptance window of ±1.5 dB (b), and data processed using both filters with an intensity acceptance window of ±1.5 dB, and a phase difference acceptance window of 0.4 radians is shown in (c).
5.8. WELD IMAGING

Figure 5.10: Tracked keyhole depth data filtered using intensity and phase filtering with blue circles highlighting tracked pixels above 12 dB (24% of tracked pixels), and between 8 and 12 dB (19% of tracked pixels). A pixel between 8 and 12 dB is only tracked if a bright pixel above 12 dB does not exist in a given line. Full 10 ms of data is shown in (a), whereas a 2 ms segment from 8 to 10 ms is shown to highlight areas where there is little signal from pixels above 12 dB (i.e. the region from 9.4 to 9.7 ms) (b)

tracking threshold a further 1 dB from the 9 dB that was previously required with just intensity filtering. Prior to making use of both intensity and phase filtering, it was not possible to drop the threshold to below 9 dB without misidentifying a significant amount of noise as signal. By lowering the tracking threshold to 8 dB (which is only possible with both intensity and phase filtering), I have gained another 6% of signal
that would have been discarded otherwise. For this weld, this brings the effective imaging rate from 48 kHz (with tracking at 12 dB), to 86 kHz (with tracking above 8 dB). The welds performed are ultimately fairly shallow, and as such we expect to see high effective imaging rates even without dropping the tracking threshold to 8 dB. As welds become deeper, we expect to see a significant decrease in the percentage of bright A-lines. This decrease occurs for the same reason that laser keyhole welding allows greater absorption of the processing beam. I initially planned to take deeper welds in an effort to demonstrate the benefits of this approach for imaging with a lower percentage of A-lines with a pixel above 12 dB. Unfortunately, the mirror completely delaminated shortly after taking the data shown in Figure 5.10(a).

Even with shallow welding data, the value of this approach is demonstrated when we examine the final 2 ms of the weld. In examining the region between 9.4 and 9.7 ms in Figure 5.10(b), we notice that there are no bright pixels above 12 dB within this section. Keyhole dynamics fluctuate at up to 10 kHz [5]. As a result, it is entirely possible that there is a change in the keyhole over the 0.3 ms that we are missing when we must threshold at 12 dB. By dropping the threshold to 8 dB (which is only possible due to the intensity and phase filtering), we have regained visibility on that portion of the welding process. In this case, we see that the keyhole becomes slightly deeper that the last 12 dB pixel at 9.4 ms, to swoop back up at 9.7 ms; this is a feature we would have missed without the introduction of a witness peak for filtering.

5.9 Conclusion

Moving from the 50:50 splitter to the step edge mirror as a method of creating a witness for feature identification has had many benefits. First of all, alignment is far
5.10 FUTURE WORK

The largest challenge in making use of a step edge optic to increase stability within the arms, was actually the stability of the optic itself. The process of creating an approximately 100 µm step on a silicon wafer falls within a rather awkward manufacturing regime. It is on the upper limit of what is typically done in micro-fabrication, while being on the lower limit of what can be achieved using traditional macro-machining methods. This led to a couple of issues, one apparent in the short term, and one
apparent in the long term.

The first issue I encountered was with the process to create the step edge. I was forced to use a photoresist due to the large (by microfabrication standards) step. At the fabrication facility we make use of, this meant spin coating the wafer in an SU-8 photoresist, and relying on the spin curves to achieve the desired thickness. In an ideal world, I would have made the step such that it was exactly an integer multiple of my depth bin size, as this would have eliminated the need for interpolation. Leading up to this fabrication, Dr. Graham Gibson, Director of NFK Kingston, attempted multiple spin coats using SU-8, and found the thickness was only accurate to about +/- 10 µm. This meant that I had to aim for a specific thickness, and account for a likely non-integer multiple of my depth bin in the software.

The second issue that I encountered had to do with the deposition of the gold reflective surface onto the silicon and the SU-8 step. Immediately after manufacturing the mirrors appeared as two flat mirror surfaces separated by a step. Over time, even with a nickel adhesion layer, the metal started to delaminate from the SU-8 step, creating bubbles in the mirror surfaces. Fortunately, I made four mirrors, and on the one I chose to use in my experiment, most of the delamination occurred away from the step edge (at least initially). It took about 5 months (from date of manufacture) of being mounted in the apparatus for the delamination to cause the optic to become unusable.

The seemingly obvious solution to the second issue is to simply use a photoresist to which metal adhere to better. This unfortunately is a non-trivial solution. Photoresists are typically designed such that metals do not adhere well, as this is a desirable property for most applications. As such, it would likely be better to avoid using a
photoresist completely. This could be achieved by using a subtractive approach to create a 100 $\mu$m depression in the silicon wafer, instead of additively creating a 100 $\mu$m step. By moving to a subtractive approach, it may also be possible to better control the depth of the depression, making an integer multiple step height possible. The issue with going to a subtractive approach, is that the depth would likely require deep reactive ion etching, to ensure straight sidewalls, which would be time consuming and expensive.
Chapter 6

Conclusion

I have demonstrated a novel technique for distinguishing signal from noise that does not rely exclusively on intensity thresholding. By introducing a second path length in the reference arm, I have effectively created a dual-channel coherent imaging system without the requirement of a secondary detector. By making use of this secondary measurement, I have used the relative intensities and phase between the two peaks to filter out noise, while preserving signal. In addition to removing random noise, this technique has also proven to be highly effective at removing imaging artifacts, such as those introduced by coatings on the optics in the sample arm. This work has included conference presentations (Canadian Association of Physicists Symposium 2021, and the Optica Laser Applications Conference 2021), as well as a publication on my findings from the first implementation of this technique in the Journal for Optics and Lasers in Engineering [1].

By first implementing this technique in an apparatus that could easily be implemented in the existing infrastructure (detailed in Chapter 3) using off-the-shelf components, I was able to first demonstrate the value of adding a second reference arm quickly and cheaply. By filtering pixels based on whether they had a witness pixel with a similar
intensity, I was able to significantly reduce the prevalence of noise, especially in the ever-problematic tail region between 9-12 dB. By limiting the prevalence of noise in the tail region, I was able to drop the depth tracking threshold from 12 dB (typically used in conventional ICI) down to 9 dB. This led to an effective imaging rate increase of nearly 67% when this technique was used to image a stainless steel keyhole during a weld. Though this apparatus showed promising first results, there were issues in stability that prevented the exploitation of phase as a filtering metric in addition to intensity. As such, I needed to design a more stable reference arm.

To improve stability, I manufactured a custom step edge optic with a separation of 94 \( \mu \text{m} \) between the surfaces. By centering the reference arm beam on the step, I created a two path reference arm on a single optic. This decreased the variation of the path length difference between the two arms from 312 nm to 26 nm. This increased stability meant that I could now exploit phase in addition to intensity for filtering. This made it possible to drop the tracking threshold by an additional 1 dB, increasing the effective imaging rate by 79% in comparison to what would be possible with standard coherent imaging. The step edge offered significant improvements over the technique used in the first implementation of self-witnessing coherent imaging, however there are issues with step edge optic itself that should be resolved prior to further experimentation.

The 100 \( \mu \text{m} \) step size falls into an awkward area between micro and macro machining, where it is on the upper limit of what is achievable using micro fabrication, and the lower of macro machining. In addition, creating optical grade surfaces is a non-trivial endeavour. The step edge mirror I created was sufficient for short term experimentation, however due to adhesion issues, the top surface did eventually completely
delaminate, rendering the optic unusable. This was even when keeping the optic in a temperature and humidity controlled environment, which is not the reality of most spaces where commercial coherent imaging systems are deployed. If the step-edge can be made to withstand environmental factors long term, it does show promise as a potential retrofit upgrade to existing ICI systems, as its implementation requires the change of only one optic in the reference arm.

I have developed a low-cost, easy to implement technique that adds a second channel to existing coherent imaging systems without the use of additional detectors. This technique has significantly increased the effective imaging rate. By increasing the effective imaging rate, closed loop control may now be possible, even for signal-starved applications. In addition, this technique may help to fully characterize keyhole behaviour in processes where achieving the 20 kHz imaging rate (Nyquist rate of keyhole fluctuations up to 10 kHz) was previously a challenge. This is an advance that benefits stakeholders ranging from industry-users attempting to implement closed loop control, to academics attempting to validate keyhole models on signal starved imaging applications. As coherent imaging for laser process monitoring becomes more common across industry and academia, I am hopeful that this technique could be used to increase our understanding of laser material processing.


