Master Thesis

Dual-beam manually-actuated distortion-corrected imaging (DMDI): 2D scanning using single-axis galvanometer with automated distortion correction

Author:

Madeline Harlow

Supervisors:

Prof. Dr. Markus Rudin (Academic)

Dr. Anthony Lee (Project)

Vancouver & Zurich 2018

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Provincial Health Services Authority

Abstract

High resolution optical imaging modalities, such as optical coherence tomography, confocal, and multiphoton microscopy, continue to show promise for diagnostic imaging. These modalities are commonly employed using 2D scanners, however such scanners can often have limited field-of-view and are susceptible to clinician or patient motion artefacts. A number of techniques have been investigated to overcome these shortcomings, including allowing for manual actuation of handheld probes. Recently our research group at BC Cancer Research Centre demonstrated a new imaging paradigm, called dual-beam manually-actuated distortion-corrected imaging (DMDI), which exploits the predictable path and spatial separation of two beams to calculate and correct the scanning distortion caused by manual actuation of the imaging apparatus and/or the sample. DMDI was first implemented using a rotating dual-beam micromotor catheter which demonstrated the feasibility of DMDI, however the complex scan pattern as a result of the catheter design resulted in tedious and manual distortion correction. Further, alternative implementations of DMDI should be explored to demonstrate the versatility and applicability of this modality.

In this thesis, I demonstrate a galvanometer implementation of DMDI with automated distortion correction. The single-axis galvanometer allows for tracing of two roughly parallel lines onto a sample which can be manually actuated. Distortion caused by the scanning pattern and manual actuation can be corrected by estimating the effective beam path of both beams onto the sample. The effective beam path is determined by characterizing the scanning pattern, achieved by a one-time calibration, and estimating the effective sample displacement, achieved by estimating the velocity profile through identifying common and unique features in the images of the two beams. With a scanning pattern of two roughly parallel lines, identifying common and unique features can be done in an automated fashion. Specifically I developed automated frame correlation to allow for detecting common features with potential displacement in two dimensions, thus sensitive to 2D actuation. Only the information from two distorted images and the calibrated scan pattern are used to produce two corrected images.

Using *en face* OCT as the imaging modality, I demonstrate DMDI with automated distortion correction in imaging of a printed paper phantom, beef tongue, kiwi, dragon fruit, and fingerprint. Distortion correction is possible for manually-actuated motion both perpendicular and parallel to the galvanonmeter-scanned lines. Through biological validations, the correction method estimated constant velocities with an average error of 0.99% and produced favorable results in the correction of 2D manual actuation. Additionally, I developed a simulation of the imaging system and actuation to allow for fast investigation of scanning parameters and arbitrary velocity profiles. For example it was observed that beam separation influences the range of velocity sensitivity. With this work as a foundation, DMDI with automated distortion correction becomes a more attractive and accessible imaging modality with a promising future in clinical applications.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by my reviewers.

Acknowledgments

I would like to acknowledge and sincerely thank Dr. Anthony Lee for his invaluable insight and guidance as my supervisor throughout this thesis and for following through with his brain child so that we could have such an interesting and fulfilling project to work on together. I would also like to thank the entire Integrated Oncology department at BCCRC for their expertise and assistance when it was needed. I would like to thank Alex for all his help in the simulations and thank you to Geoffrey for being my rubber duck. I would like to thank Hanna for dedicating her valuable time to review and critique my report. I would also like to thank Prof. Dr. Markus Rudin for his role as my academic supervisor and allowing me to explore a new place in research as well a new place in this world. Lastly, thank you Thomas for following me out here, taking care of me, and keeping me sane and fed throughout this whole process.

Table of Contents

1 Introduction

1.1 Optical scanners

High resolution optical imaging modalities such as optical coherence tomography (OCT), confocal and multiphoton microscopy continue to show promise for diagnostic imaging. These imaging modalities commonly employ 2D scanning mechanisms that scan the sample in regular, pre-defined patterns. Examples of these scanners, shown in Fig. 1, include rotary motors with mechanical pullback e.g. in catheters [1, 2], MEMs mirrors [3, 4], and galvanometers [5]. A galvanometer, illustrated in Fig. 1c), is a scanning technique consisting of rotating mirrors which redirect impinging beams in linear or raster patterns onto a sample plane. This particular scanner has been used for many applications such as laser etching [6] and in confocal microscopes [7] as it allows for high precision 2D planar scanning. However, the scanners mentioned can be susceptible to artefacts due to patient or clinician motion or often have limited field-of-view (FOV) making it challenging to image large sample regions. To be attractive for clinical applications, a scanner should be insensitive to patient or clinician motion, ideally be actuated and controlled manually by the clinician thus relatively simple and easy to use, and, for many applications, be able to scan large regions of interest with high image resolution and accuracy.

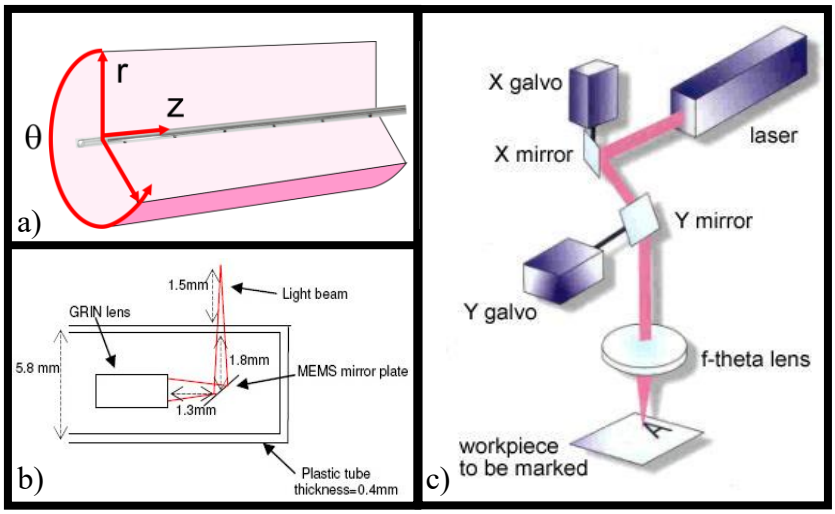

Fig. 1. 2D scanner examples. a) Rotary motors, here applied in a catheter with mechanical pullback [1], can allow for radial 1D scanning of the imaging beam. A motorized pull back allows for scanning along the z-axis, thus 2D scanning. b) 2D MEMs mirrors, here applied in an endoscopic probe [4], allow for 2D beam actuation with very small hardware. c) A galvanometer can direct an impinging beam to scan a sample plane [8].

A number of techniques to increase FOV have been investigated. For example, many imaging modalities increase FOV by capturing multiple overlapping images which are stitched together to form a mosaic [9, 10]. However this requires many scans, which can be susceptible to patient motion, increases procedure time, and it often requires rigorous image processing to form the mosaic. In rotary catheters a motorized pullback [1, 2, 11] can be implemented to capture a longer FOV along the axis of pullback. The length of the FOV is limited by the stroke size of the motor. A motorized pullback removes distortion caused by manual actuation, as a constant and known pullback velocity

is applied in a scan, however clinician or patient motion artefacts may still arise and the additional hardware adds complexity to the imaging setup. Manually scanning of a handheld probe is an attractive technique for increasing and even controlling FOV. For example this method is found in imaging capsule applications [12, 13]. However, non-uniform scan velocities along the axis of pullback result in motion artefacts which distort the image significantly. Further, knowledge of the probe displacement between consecutive scans is not available, making it challenging to locate precisely where a particular scan was acquired. A number of groups have researched techniques to overcome manual motion artefacts and these techniques can be classified into two categories: sensored and sensorless.

Sensored techniques introduce additional hardware, such as magnetic sensors [14, 15] as shown in Fig. 2, IR optical tracking [16], linear position encoders [17], and IR cameras [18], used to track the imaging probe with respect to the sample. These approaches allow for imaging probe motion correction but are not sensitive to sample motion. Furthermore, the required additional hardware adds complexity to the design, can in some cases be costly, and potentially introduces additional challenges. For example, magnetic tracking introduces the challenge of ensuring the procedure environment is free of magnetic field distorting materials to maintain tracking accuracy.

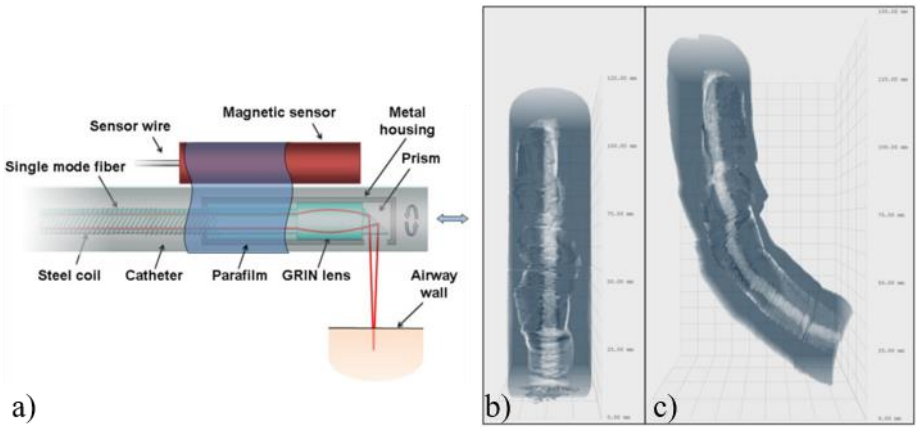

Fig. 2. A magnetically tracked imaging probe is an example of a sensored 2D scanning technique [14]. a) Schematic of the imaging probe. b) 3D reconstruction image of a phantom without sensor assistance. c) 3D reconstruction image of phantom with sensor assistance.

Sensorless techniques exploit information and inherent characteristics from the acquired image to correct for motion. For example, speckle decorrelation is a technique used in motion tracking of ultrasound [19] and OCT probes [20-23] where speckle texture is used to infer dynamics. The cross correlation coefficient of consecutive or slightly displaced A-lines have a functional dependency on the magnitude of lateral displacement and therefore can be exploited to correct for lateral motion distortion. Ahmad et al. were the first group to demonstrate this concept and proved 1D correction of ‘smear’ distortions, caused by low to zero lateral speeds or oversampling, was feasible by identifying and removing consecutively correlating A-lines, shown in Fig. 3. Liu et al. demonstrated that speed of

an imaging probe can be estimated from the correlation information of oversampled data [21]. However extension to 2D motion is not trivial as it is challenging to oversample in 2D.

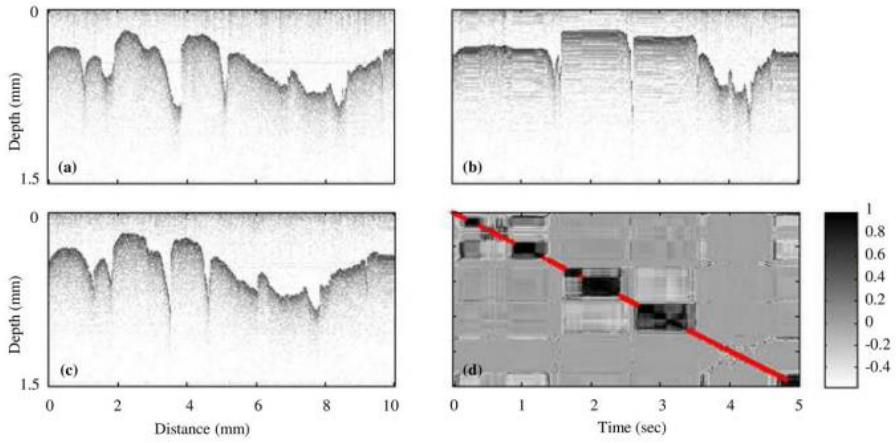

Fig. 3. Ahmad et al. presented a cross-correlation based technique for correcting variable lateral speeds. a) Image of plasticine sample uniformly scanned using a motorized stage. b) Image scanned by hand, i.e. non-uniform. c) Corrected non-uniform scan. d) Cross-correlation matrix with red points showing the A-lines selected for image correction. [20]

A sensorless technique is attractive for simple and cost effective imaging systems yet there is an unmet demand for a large FOV 2D scanning technique with manual actuation distortion correction. As well, the techniques mentioned are susceptible to sample motion causing potential artefacts, which remains a challenge when imaging living, moving patients.

1.2 DMDI with DBMC

The Integrative Oncology Department at BC Cancer Research Centre (BCCRC) recently demonstrated a new imaging paradigm, called dual-beam manually-actuated distortion-corrected imaging (DMDI), which allows for sensorless manual scanning of an imaging probe or sample with correction distortion due to 2D probe and sample motion. DMDI utilizes two imaging beams that scan mechanically along one dimension and are spatially separated by a known distance in the orthogonal dimension, which can be manually scanned. As the sample is manually scanned, the time interval of one beam passing the same sample feature as the other beam can determine the average velocity of the sample or imaging probe within that interval. In the general case, given the known spatial separation, denoted d, of two beams, labeled A and B, and the time at which each beam captures the same sample feature, denoted $t_{A,k}$ and $t_{B,k}$ respectively, the average velocity in that time interval for that particular feature, k, is given by Eq. 1.

$$\bar{v}_{AB,n} = \frac{d}{t_{B,k} - t_{A,k}} \tag{1}$$

This concept is illustrated in one dimension in Fig. 4, where a time series is shown of an imaging probe, e.g. a catheter, with two spatially separated beams manually scanning past two sample features. This example can be expanded to two dimensional scanning by mechanically rotating the beams for radial scanning to, for example, image a tubular structure. The estimated average velocity for many features

throughout the scan can be integrated over time to determine displacement which allows for re-mapping of the distorted images from either beam to produce distortion-corrected images.

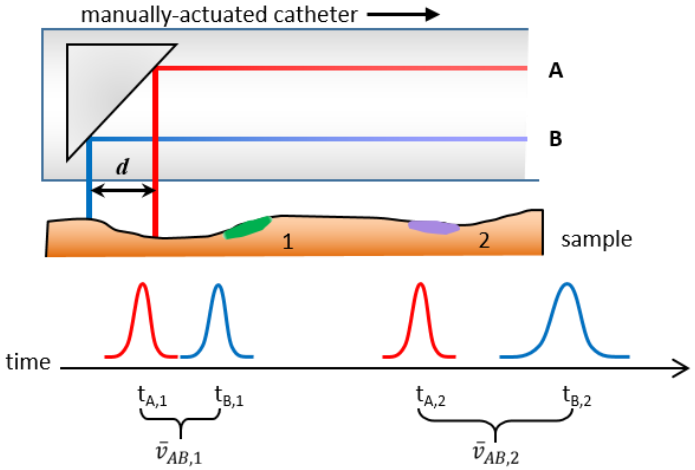

Fig. 4. Illustration of the concept of DMDI adapted from [24] presentation slides. A catheter with spatially separated beams, A and B, manually scans in one dimension across a sample with two features. The time series shown describes the catheter passes the first feature with both beams and the second feature with beam A at the same velocity but beam B passes the second feature at a slower velocity. The slower velocity would cause the second feature in the image produced by beam B to appear stretched compared to the A image. The time interval for each beam to pass the same feature and the known separation of the beams are used to calculate the average velocity within that interval, denoted $\bar{v}_{AB,n}$.Knowledge of the velocity profile allows for correcting the distortion caused by it.

Distortion in DMDI images can be attributed to two sources of motion: the scanning pattern of the imaging apparatus and the motion of the sample and/or imaging apparatus. The scanning pattern distortion is an inherent characteristic of the imaging apparatus and can cause distortion to various degrees of severity depending on the scanning technique. Motion or manual actuation can result in stretching, compressing, shearing, or warping of image features depending on the magnitude and direction of the motion. Distortion correction can be achieved by estimating the effective beam paths of the two beams due to these sources of motion. Estimating the effective beam path can be done in two steps. First, determining the scanning pattern of the imaging apparatus which can be achieved through a one-time calibration. Second, determining the actuation profile which can be achieved through identifying common and unique features in both images used to estimate the effective displacement of the sample due to actuation of the sample and/or imaging apparatus.

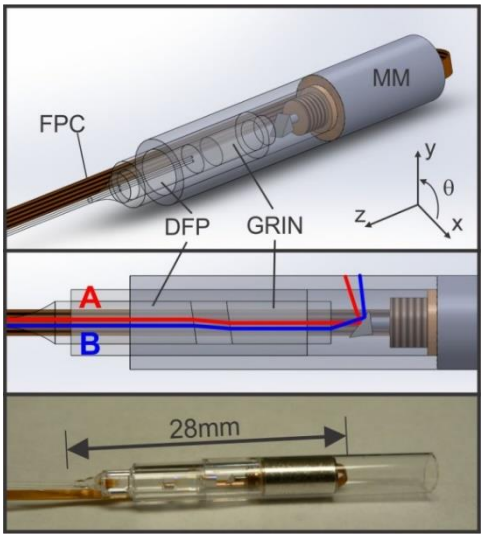

Fig. 5. Dual-beam micromotor catheter. A 3D model of the catheter [top] and photograph [bottom] where FPC = flexible printed circuit, DFP = dual fiber pigtail, GRIN = graded index lens, and MM = micromotor. A rotating prism scans the two impinging beams, A [red] and B [blue] radially, tracing elliptical patterns on the circumference of the catheter, illustrated in [middle]. [24]

DMDI was first demonstrated using a dual beam micromotor catheter (DBMC) [24], shown in Fig. 5. The two beams of the catheter are mechanically scanned by a rotating prism producing two elliptical traces on the circumference of the catheter. This configuration results in a complex scanning pattern which produces an ‘s’-shaped distortion in the images, as can be seen in Fig. 6a). Due to the complicated scan pattern of the DBMC, it is difficult to identify common and unique features in an automated fashion, as sample features can appear very different in both images. Therefore, in the DBMC implementation of DMDI, the common and unique features were identified through manual selection of control points which co-register features in both images, shown in green in Fig. 6a). The control points describe the time interval of each beam passing the same feature, and given the known beam separation as a function of rotation of the micromotor, the velocity of the rotation of the catheter, Fig. 6c), and the axial actuation of the catheter, Fig. 6d), can be estimated. From the velocity estimates, displacement can be determined and used to correct the distorted image, shown in Fig. 6b).

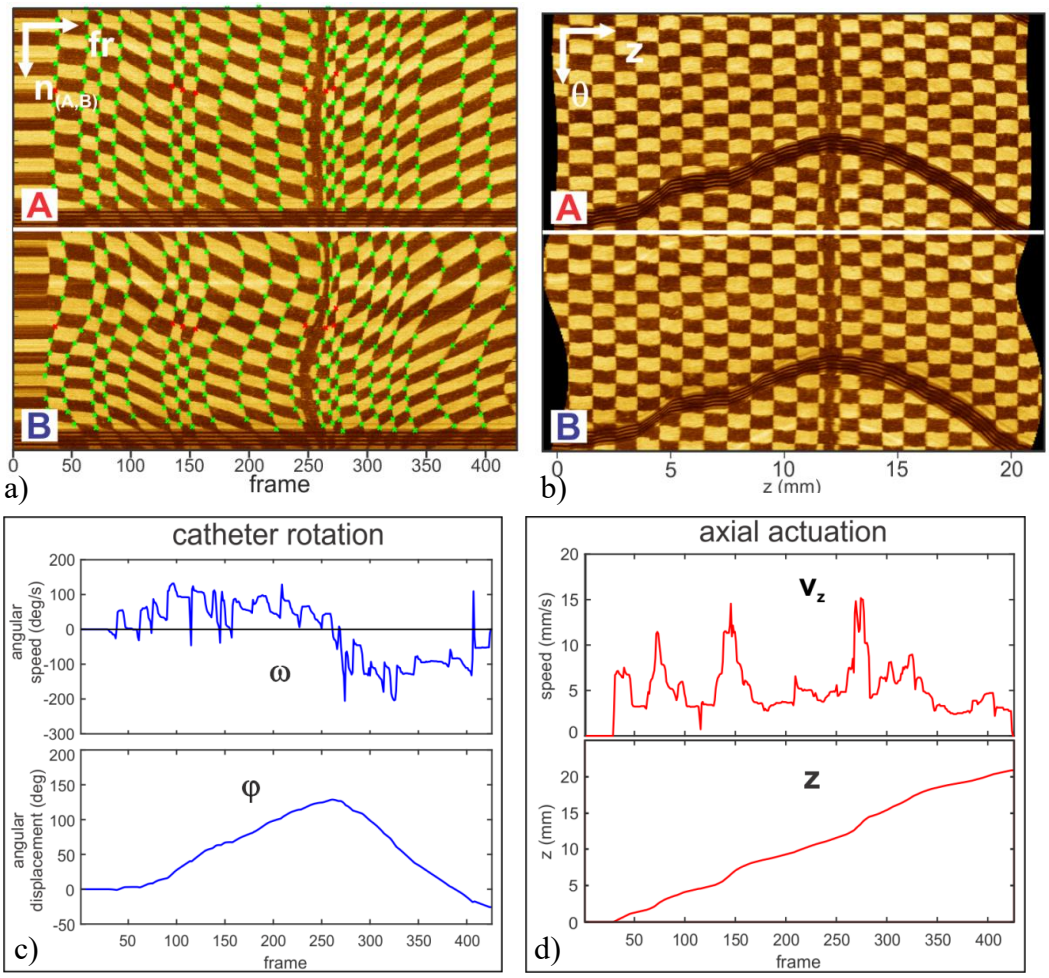

Fig. 6. Images of a checkerboard phantom acquired with manual actuation of the DBMC in the z- and theta-axis. a) The distorted images with manually selected control points [green] for velocity estimation. b) Distortion corrected images. c) Catheter rotation velocity estimate and corresponding displacement used for distortion correction in the image vertical axis. d) Axial actuation velocity estimate and corresponding displacement used for distortion correction in the image horizontal axis.

1.3 Thesis motivation

The previous work demonstrates DMDI is a feasible and promising modality allowing for manual two-dimensional scanning of a probe or sample with distortion corrected images. The DBMC implementation of DMDI could be useful for *in vivo* imaging of internal vessels, air ways, or tubular organs. However the biggest limitation of the technique presented is the laborious procedure of manually selecting numerous co-registration control points which are required for distortion correction. To make DMDI more accessible and attractive for future clinical applications, a demand for automated distortion correction is present. Furthermore, alternative implementations of DMDI should be explored to investigate the versatility and potential applications of the modality.

The scope of this thesis is to investigate a new implementation of DMDI and develop automated distortion correction. A new implementation of DMDI will demonstrate the versatility and applicability of this imaging modality. Specifically a galvanometer implementation could be applied to alternative imaging sites such as the skin or the oral cavity. Furthermore, a galvanometer implementation can

simplify the complicated scanning pattern and allow for automated distortion correction which will make DMDI more accessible and attractive as a potential clinical tool.

1.4 Thesis aims

To explore a new implementation of DMDI and develop automated distortion correction, a number of aims must be met.

First, I will design and construct an optical set up and galvanometer configuration. This will include designing, constructing, and profiling an optical probe. The optical probe must be aligned to the galvanometer which needs to be optimized for scanning in this application. This aim is described in Chapter 2.

Next, I will acquire sets of calibration data. These images may need to be preprocessed, thus I will need to determine the preprocessing procedure. I will determine how to calibrate the system and correct images without any motion distortion. I will calibrate the galvanometer system by characterizing the beam pattern. This aim is described in Chapter 3.

Next, I will develop automated distortion correction. This includes selecting a phantom I will image to acquire a training set, preprocessing said training set, and determining a robust, accurate, and automated method of co-registering common and unique features in each image. These features in turn will be used to discern the manual actuation profile and generate distortion corrected images. This aim is described in Chapter 4.

Finally, I will acquire test data sets of various biological samples to validate and test the galvanometer implementation and the automated distortion correction. This will include any sample preparation required and multiple iterations of image acquisition to test variable scanning conditions. This aim is described in Chapter 5.

In supplement to the various validation tests, there is opportunity to simulate the distortion caused by this imaging system and actuation velocities to allow for investigations of different scanning parameters. I will develop a simulation which constructs distorted images due to the imaging system, which can be parameterized to allow for changing of the scanning parameters, and arbitrary actuation velocities such that the effects of these scanning parameters and actuation velocities on correction accuracy can be investigated. This aim is described in Chapter 6.

2 Imaging System and Acquisition

In this chapter, the imaging system and acquisition process is described. First, OCT and the OCT imaging system are introduced. Then the imaging head requirements and set up for this implementation of DMDI are described. Next, the samples used and their corresponding preparation procedures are described. Lastly, the preprocessing technique used on acquired images is described.

2.1 Introduction to OCT

Due to its availability in the Integrative Oncology lab at BCCRC, *en face* OCT is the imaging modality used in this thesis, however in theory any point scanning modality could be used. OCT allows for high resolution, non-invasive, *in vivo*, cross sections and 3D volumetric imaging of tissue by measuring echoes of backscattered light [25]. Analogous to ultrasound, the magnitude and echo delay time of backscattered light is measured and image contrast is obtained at boundaries of changing optical properties. In 1D, a profile of backscattered light intensity as a function of sample depth is called an A-scan, or A-line, shown in Fig. 7a). Through transverse scanning, a series of A-lines generates a 2D cross-sectional image of the sample, called a B-scan, shown in Fig. 7b). Lastly, a series of B-scans generates a 3D volumetric data set, called 3D-OCT shown in Fig. 7c). An *en face* OCT image, used throughout this thesis, is produced from 3D volumetric data by taking the mean intensity across a range of depths, resulting in a 2D image depicting the surface of the sample. An example to illustrate *en face* imaging acquired using the imaging system constructed in this thesis will be shown in section 2.5 once more details of the imaging system are introduced.

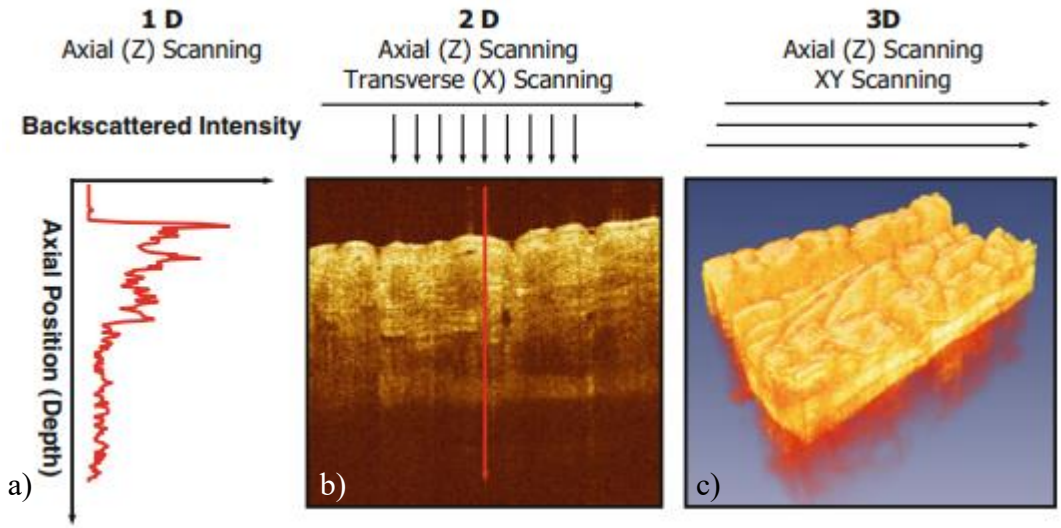

Fig. 7. Types of OCT scans. a) In 1D, an A-scan, or A-line, provides measurements of backscattered light as a function of sample depth. b) Transverse scanning to acquire a series of A-lines generates a cross sectional image of the sample, also called a B-scan. c) A series of B-scans, acquired e.g. through raster scanning, generates a 3D volumetric data-set of the sample, also called 3D-OCT. [25]

As light travels much faster than sound, the same detection methods used in ultrasound cannot be used since a much higher time resolution is required [25]. Instead, interferometry is used. Interferometry utilizes coherent light that is split and travels to either a mirror where it reflects back

or to the sample where it backscatters, the first being called the reference arm and the second the sample arm. The two arms are combined to form an interference pattern. The resulting interference pattern is due to the phase difference of the two arms caused by the presence of the sample. Thus this phase difference describes the optical properties of the sample. By changing the length of the reference arm or by sweeping the source across a range of wavelength, varying depths of the sample can be resolved, i.e. tomography can be achieved.

2.2 OCT imaging system

The OCT imaging system used in this work is schematically shown in Fig. 8. A 23 mW, 100 kHz 1310 nm swept-source laser (AXP50125-6, Axsun Technologies, Billerica, MA) is 50/50 split and fed into two parallel Mach-Zehnder OCT interferometers. The sample arms of each interferometer feed into the imaging head, described in the next sections. OCT signals returning from the imaging head, and sample, are combined with their respective reference arms and fed into a pair of balanced detectors (PDB480-AC, Thorlabs Inc., Newton, NJ) which are simultaneously digitized by a dual-channel, k-clocked acquisition board (ATS9350, Alazar Technologies, Inc., Point Claire, QC). The path length difference between the two interferometers is approximately 2 m, sufficiently different such that there is no OCT cross-talk between them.

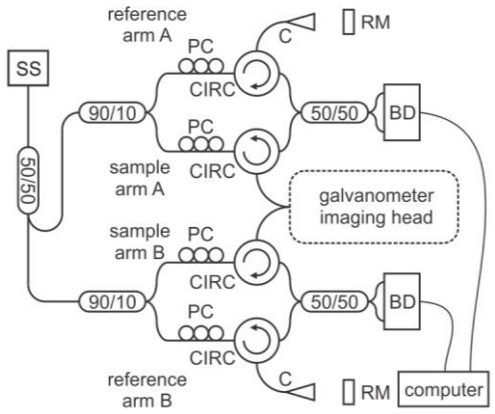

Fig. 8. A schematic diagram of dual-channel OCT imaging system. SS = swept source laser, PC = polarization controller, C = collimator, RM = reference mirror, CIRC = circulator, BD = balanced detector.

2.3 Imaging head

This section describes the imaging head requirements and configuration. The imaging head defines the scanning pattern of the DMDI imaging system as it is responsible for scanning the incoming sample arms of the OCT imaging system onto the sample.

2.3.1 Imaging head requirements

In the DBMC implementation of DMDI, the rotating prism traced the two beams in a complex scanning pattern, which made it difficult to implement an automated feature-matching technique. A potential solution to this problem is an implementation which allows for a simpler scanning pattern, such as two parallel lines. To accomplish this, some requirements of the imaging system should be met:

1. Produce two collimated beams
2. A scanning method to trace beams into two parallel lines, e.g. a single-axis galvo mirror.
3. Both beams should penetrate the sample at the same angle of incidence throughout the scan, such that information from both beams are comparable to each other and ideally beam paths are normal to the sample plane.
4. Beams should be coplanar such that they can impinge on the mirror on its axis of rotation.
5. The angle between the two beams impinging on the mirror should be small, such that separation of the beams at the sample is also small.

2.3.2 Galvanometer imaging head

Beginning with the second requirement, in this implementation a single-axis of a two-axis galvanometer scanner head (GVS002, Thorlabs Inc., Newton, NJ) is used to trace two parallel lines at the sample. A standard component of galvanometer scanners is a telecentric f-theta lens, also called a scanning lens. A scanning lens compensates for the changing beam angle caused by the rotating galvo mirror to allow for scanning of a flat sample plane. As a beam is redirected from a rotating mirror, the angle of the beam would be a function of the mirror rotation angle, producing a semi-circular scanning field at the sample if not compensated for, as illustrated in Fig. 9a). A scanning lens focuses the beam onto a flat scanning field ensuring that the beam is normal to the scanning field at all times, illustrated in Fig. 9b). Use of a scanning lens is critical for ensuring both beams penetrate the sample at the same angle and therefore capture the same sample information when scanning at the same location. The scanning lens available in the lab (LSM02-BB, Thorlabs Inc., Newton, NJ) has a working distance of 7.5 mm and a numerical aperture (NA) of 3.7x3.7 mm^2, indicating an impinging beam with a spot size of 3.7 mm in diameter will achieve optimal focus at the sample located 7.5 mm from the lens.

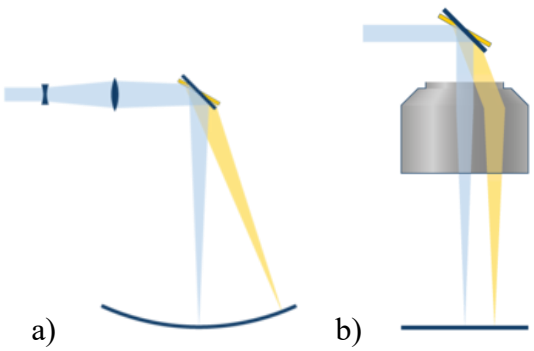

Fig. 9. a) A galvo mirror produces a semi-circular scanning field. b) A telecentric scanning lens focuses the beam to a flat scanning field ensuring the beam is normal to the scanning field at all times. [26]

Addressing the first requirement, two collimated beams can be produced numerous ways. Perhaps the simplest solution to understand would be two single fibers each connected to two single collimators producing two collimated beams. In order to satisfy the requirements 3 through 5, these

two beams must be carefully aligned to the single axis galvo mirror such that both beams are coplanar, impinge the mirror on its axis of rotation, and do so at an angle that produces a small beam separation at the sample. With two free moving probes, there are 5 degrees of freedom to overcome in order to have two aligned and coplanar beams in addition to the degrees of freedom required to be resolved in order to align to the galvanometer. In this configuration, alignment is time consuming and challenging but allows for the advantage of controlling the beam separation at the sample. Another solution for producing two collimated and coplanar beams is to use a dual fiber pigtail (DFP) and graded index lens (GRIN) pairing. A DFP contains two aligned fibers, allowing for two aligned and coplanar beams. The two beams can become collimated when positioned at the working distance of the GRIN. The intersection point of the two beams occur at the face of the GRIN and diverge from each other outward. The DFP and GRIN configuration produces one probe with two aligned, coplanar, and collimated beams which avoids the 5 additional degrees of freedom introduced by the previous configuration thus simplifying the galvanometer alignment. A shortcoming of this configuration is effectively revoking control over the beam separation as beam separation at the face of the DFP is fixed by its design. To achieve focus at the sample, the beams should meet at the parafocal distance of the scan lens, located just in front of the galvo-mirror on the input side. Thus the face of the GRIN must be positioned as close as possible to this parafocal distance. Another shortcoming of this configuration is a loss of resolution at the sample due to a small beam size at the scan lens. The two collimated beams exiting the GRIN will be small, on the order of 300-400 μm, which is an order of magnitude smaller than the optimal NA for the scan lens. Due to the time sensitive nature of this thesis time-line, the optical set up was built for the purpose of investigating the implementation and quickly acquiring data to be used to develop automated distortion correction as opposed to achieving the best resolution or variable beam separation. Therefore the second configuration was selected.

The galvanometer imaging head configuration is illustrated in Fig. 10. A dual SMF-28 fiber pigtail (DFP, KFP-P-2-8D-250S-250-1310-N, Photop Technologies, Inc., Santa Clara, CA) is aligned with a graded index lens (GRIN, GRIN2313A, Thorlabs Inc., Newton, NJ) in a glass ferrule at a separation of approximately the working distance (0.242 mm) to achieve collimation. Detail steps for this probe construction procedure are included in Appendix A. Both the DFP and GRIN lens are polished at 8° and anti-reflection coated for 1310 nm. An 8° polish is used at all component junctions to prevent back reflections which can cause disturbing artefacts in the image. The fibers within the DFP are aligned on the pivot axis of the 8° face with a separation of 250 μm and each equidistant from the center of the face.

The glass ferrule, with the DFP and GRIN, is aligned to a single-axis galvo mirror such that the two beams exiting the face of the GRIN hit the mirror near its rotational axis. The beams exiting the face of the GRIN diverge from each other. This diverging angle as well as the distance traveled to reach the sample will determine the final beam separation at the sample. In order to minimize this separation

as well as achieve optimal focus at the sample, the GRIN is positioned as close as possible to the parafocal distance of the scan lens and the mirror without risk of collision when the mirror rotates. Since DMDI is achieved through imaging the same sample feature with two spatially separated beams, minimal beam separation is desirable as it reduces the distance the sample must travel to be captured by both beams. The galvo mirror, clocked by the laser sweep trigger, is driven by a modified saw tooth waveform which maximizes the forward scanning segment of one scan period while not exceeding the maximum acceleration limit during the minimized retrace scanning segment. To simplify the problem, only the forward scanning segment will be used as the region of interest for processing, therefore maximizing the forward scanning segment allows for higher sampling within the segment and a longer duty cycle. A small strip of reflective metalized mylar placed near one turning point of the galvo scan pattern, such that it occludes the sample in both beams at the same time, is used as an optical frame synchronization marker. Lastly, a lens tube (SM2L1, Thorlabs Inc., Newton, NJ) surrounds the scanning lens and is used to maintain the sample in focus at the working distance of the scanning lens during a scan.

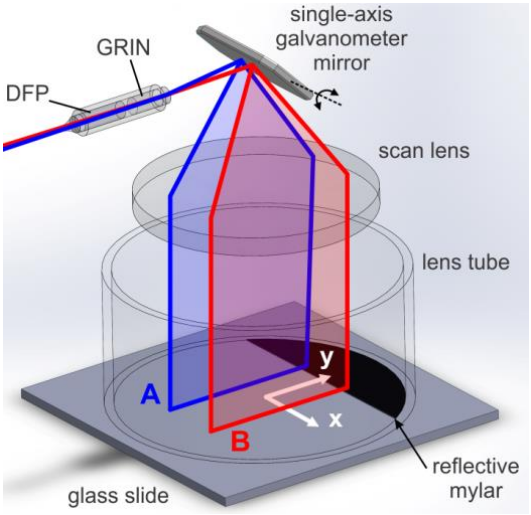

Fig. 10. Galvanometer imaging head configuration. An illustration of galvanometer configuration. DFP = dual fiber pigtail, GRIN = graded index lens. Approximate paths of beams A and B are shown in blue and red respectively.

The approximate light path of this implementation is also illustrated in Fig. 10 as blue and red lines for beams A and B respectively. Due to the 8° polish of the DFP and GRIN, the beams exit the GRIN slightly downward of the horizontal plane, which is compensated for in the alignment. As the mirror rotates, both beams, which are diverging from each other, reflect off the mirror, pass through the scan lens and trace two roughly parallel lines on the sample located at the working distance of the scanning lens. Due to the small size of the collimated beam impinging upon the scan lens, the resultant focused spot size of each beam at the sample is roughly 80x80 µm FWHM measured with a beam profiler (BeamMap2, DataRay Inc., Redding, CA). This is a relatively low resolution compared to typical OCT resolutions on the order of 10 µm.

Lastly, for notation purposes, a coordinate system is established with the x-axis defined perpendicular to the scan lines, originating at the midpoint between the scan lines, and the y-axis defined parallel to the scan lines at the midpoint of scan line B, as depicted in white in Fig. 10.

2.4 Sample preparation and imaging

Various samples are used to calibrate, train, and test in this work. For calibration and training, black phantoms, of lines and a QR code respectively, printed on standard white printer paper were used. For testing, biological samples were used, including beef tongue, kiwi, dragon fruit, and a fingertip. All samples are imaged with a standard transparent glass slide between the sample and the lens tube of the imaging head to ensure a consistent and flat surface slides against the lens tube during a scan. This is particularly important when scanning samples with irregular surfaces such as the biological samples. For biological samples, the interface between sample and glass was filled with water.

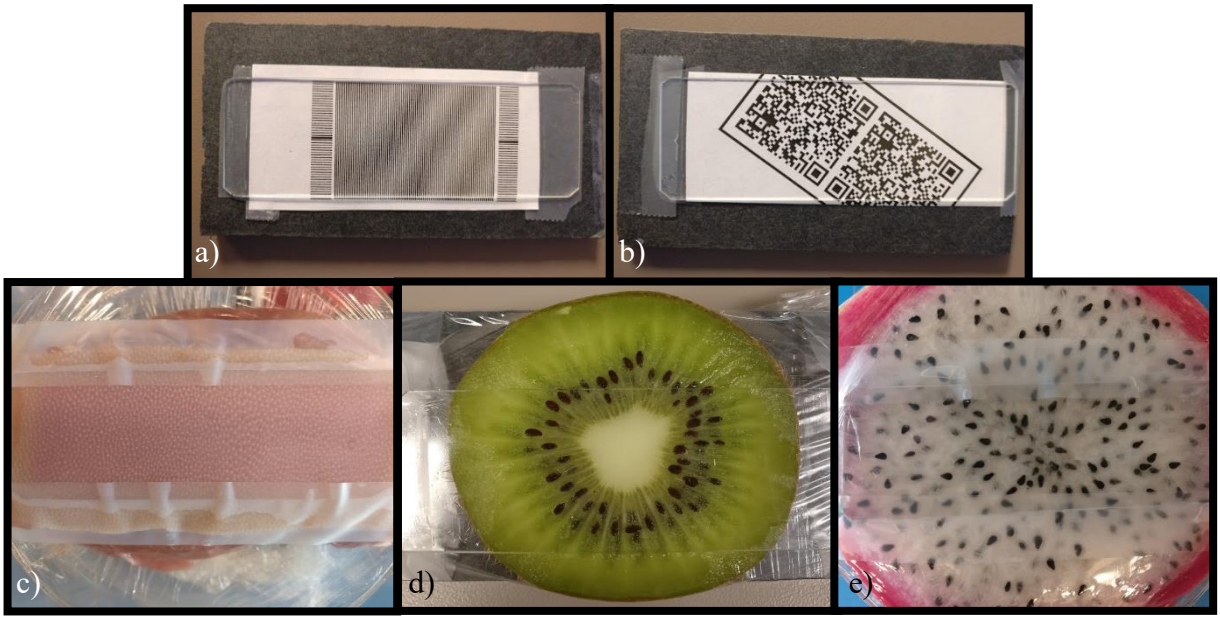

Fig. 11. Paper [top] and fruit [bottom] samples were used to train and test. All samples are covered with a standard transparent glass slide. a) The calibration pattern with vertical lines equally spaced by 0.339 mm and horizontal lines equally spaced by 0.698 mm with the exception of one pair with the separation filled in. b) The QR code training phantom with smallest feature ~0.53x0.53 mm. c) A beef tongue surface sample. d) A kiwi cross section. e) A dragon fruit cross section.

To prepare the paper samples, the printed calibration pattern, Fig. 11a), and testing phantom, Fig. 11b), were taped face up and flat onto a flat platform with a glass slide taped flush above it. The beef tongue, kiwi, and dragon fruit samples, shown in Fig. 11c-e), were prepared by taking a surface sample for the tongue and a cross sectional slice for the fruits and placing the sample in a Petri dish or on a flat dish with a glass slide placed on top and water, in between, filling any air gaps. The slide was firmly held in place with plastic wrap and tape with a window cut into the plastic wrap to ensure the scanning region was not occluded. In this configuration, the prepared sample can be pressed to the lens tube with consistent pressure and contact for scanning. Similarly for the fingertip, a glass slide was pressed

to the lens tube with a wetted index finger, shown in Fig. 12b), while ensuring to maintain consistent pressure and contact.

To acquire an image, as demonstrated in Fig. 12, the sample is pushed in the plane normal to the opening of a lens tube. The lens tube surrounds the scanning space and keeps the sample in focus throughout the scan. For calibration and validation, the samples were pushed using a linear stepper motor while the other samples were manually pushed by hand to allow for motion in both x and y axis. Care was taken to avoid rotation of the sample during actuation.

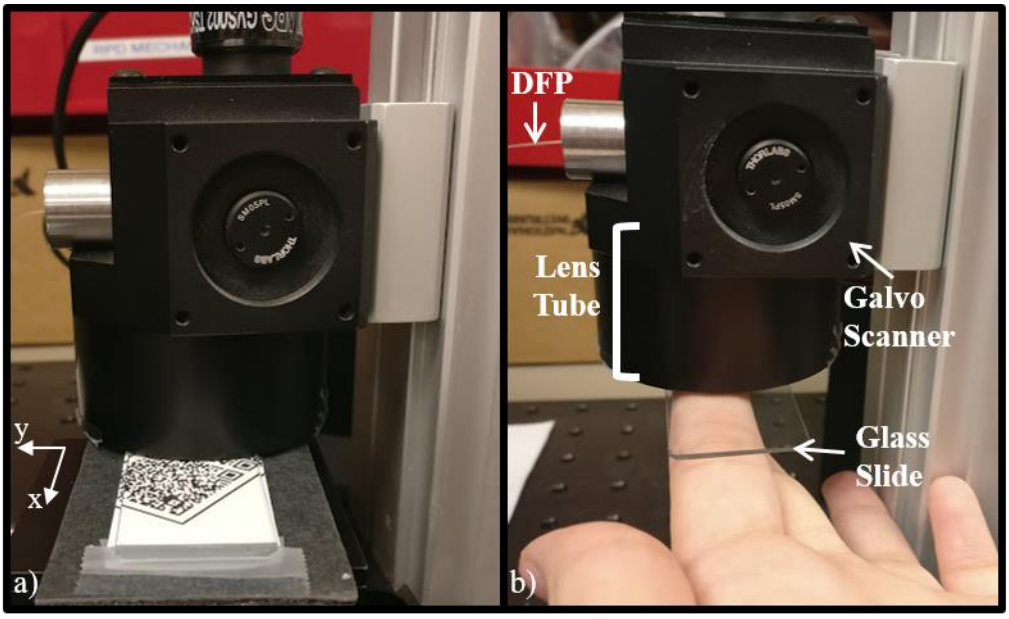

Fig. 12. Demonstration of imaging a sample. Either the imaging head is pushed down flush onto a flat sample, as in a), or the sample is pushed to the lens tube by hand, as in b), ensuring to maintain consistent pressure and contact. DFP = dual fiber pigtail.

2.5 Image acquisition and preprocessing

Two imaging channels, A and B, are acquired as continuous streams of equally time-spaced A-lines with index N. Recall a series of A-lines across one dimension, such as along the y-axis of the scan lines, constructs a B-scan. An example of a B-scan acquired from beam A along one galvo cycle and therefore one y-axis scan line of a fingerprint is shown in Fig. 13a). Features unique to a fingerprint OCT image are labeled as well as features unique to this imaging system, such as the forward and retrace scanning regions and the appearance of the glass slide, which rolls over the top and wraps around the OCT image. *En face* OCT images are formed by taking the mean intensity projection over the OCT depth dimension, labeled in yellow in Fig. 13a), of each A-line and reshaping the continuous stream such that each column, called a frame, contains 1000 mean intensity A-lines equivalent to the time interval of one galvo scan period. The A-line index within a given frame is denoted n, ranging from 1 to 1000. Given a laser frequency of 100 kHz and one frame consists of 1000 A-lines, the corresponding frame rate is 100 frames/s. The *en face* image A of the fingerprint is shown in Fig. 13b) with the black line indicating where the B-scan in Fig. 13a) is located.

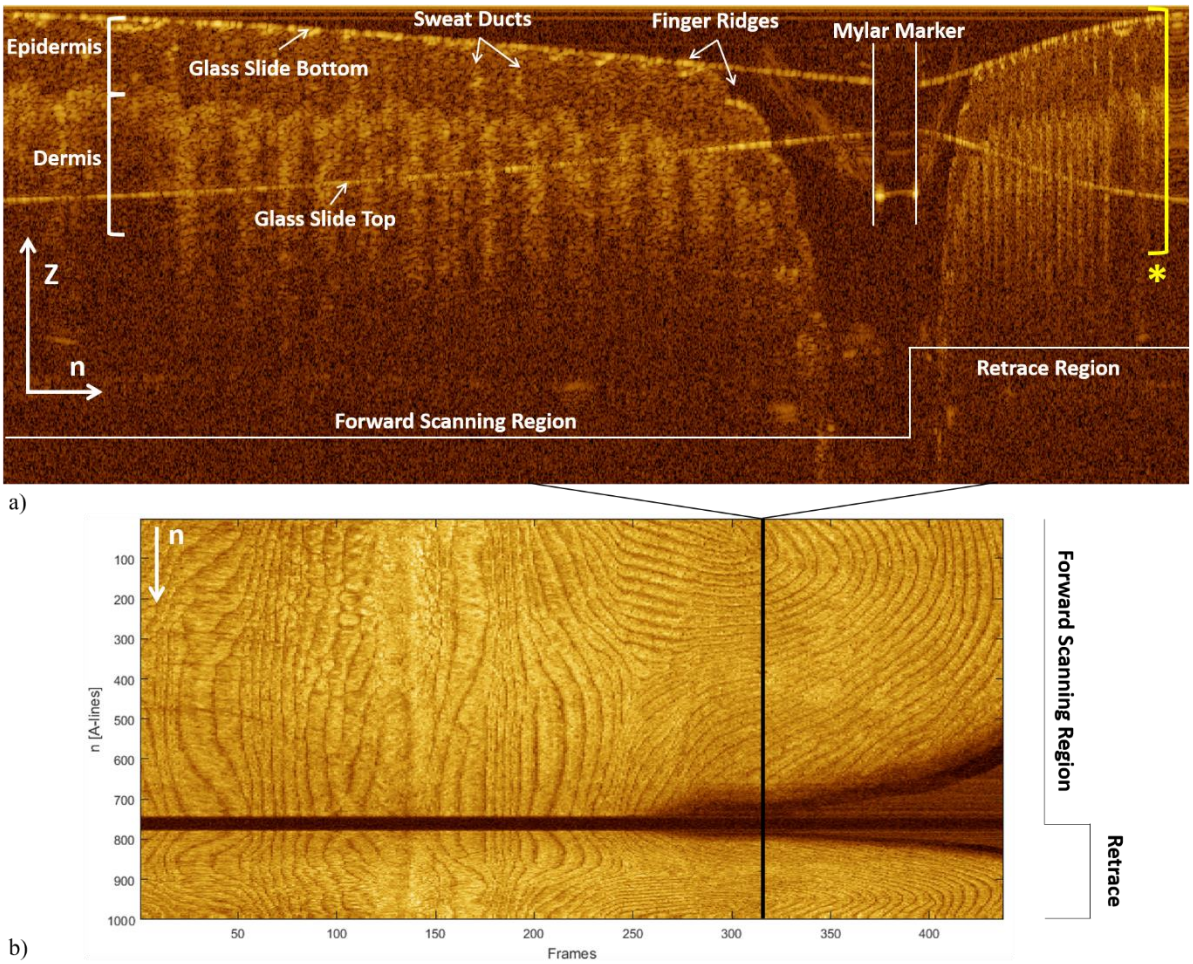

Fig. 13. An example a) B-scan of a 3D-OCT data set acquired of a manually actuated fingerprint and corresponding b) *en face* OCT image. The black bar in b) indicates where a) is located in the scan. Finger specific OCT features such as the epidermis, dermis, finger ridges, and sweat ducts are labeled as well as implementation specific features such as the mylar marker and the forward and retrace scanning regions. The top face of the glass slide rolls over the top and wraps around the image. The yellow bar [*] indicates the depth at which the mean intensity is calculated to form the *en face* image.

The image acquisition can begin at any point in the galvo scan period; however the frame synchronization marker will always occlude the sample at the same point in a scan period. Therefore the marker is used to align images such that the beginning of each frame corresponds to the start of the forward scanning segment. This is achieved by cropping the beginning of the continuous streams until the start of the first full scan period, indicated at n_{roi} = 743 A-lines before the beginning of the marker, is found before reshaping the streams. Once aligned, shown in Fig. 14, the forward scanning region up to the marker is isolated as the region of interest for correction, thus discarding the marker and the data acquired during the retrace segment. This results in two images with dimensions 743 A-Lines by number of frames, depending on the total scan time.

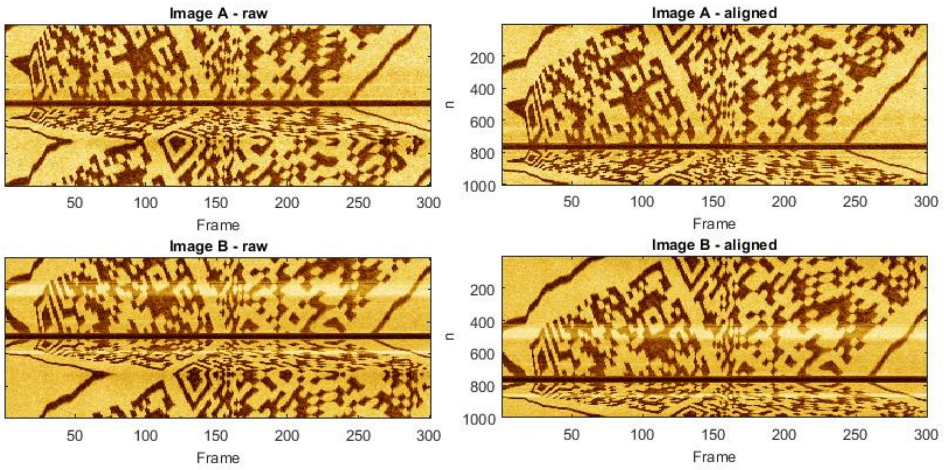

Fig. 14. Raw [left] and aligned [right] images from beams A [top] and B [bottom] of the QR code testing phantom which was scanned with manual actuation.

As can be seen in Fig. 14, strong reflection artefacts are present in the B beam, due to reflections caused by the galvanometer alignment. To mitigate the appearance of reflection artefacts, the aligned and cropped images are flat-field corrected by using a pseudo flat-field obtained from imaging without a sample, illustrated in Fig. 15. Flat field corrected image is calculated using Eq. 2

$$im_{ffc} = \frac{im}{im_{ff}} * \mu_{im} \tag{2}$$

where im_{ff} is the flat field image and μ_{im} is the mean intensity of the image. The image without a sample was averaged across all frames to produce a single frame average which can be used to correct single frames of a sample image. In other words, a pseudo flat field image can be produced by repeating the single frame average to match the number of frames of the image to be corrected, which is shown in Fig. 15.

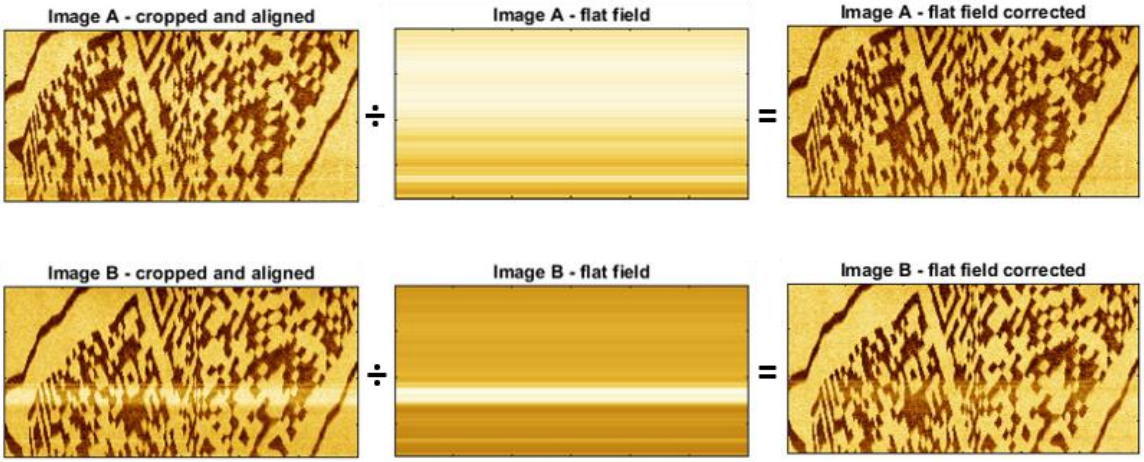

Fig. 15. Flat field correction [right] of cropped and aligned images [left] from beams A [top] and B [bottom] using a pseudo flat field [middle] produced by repeating in each column the total frame average of an image without a sample.

3 Calibration

Calibration is required in DMDI to characterize the scanning pattern to allow for distortion correction. Characterizing the scanning pattern involves defining the x and y position of each beam as a function of the galvo mirror rotation in time. Once characterized, one can correct for distortion caused by the scanning pattern and map raw images from their inputs of A-lines and time into dimension units. Calibration is achieved by imaging a calibration pattern which provides a ground truth that can be exploited to determine the conversion to dimension units. In this work, the calibration pattern shown in Fig. 11a) was used. The calibration pattern consists of equally spaced black lines, with 0.339 mm separation from first edge to next first edge for the vertical lines and 0.698 mm separation for the horizontal lines, and is printed on standard white printer paper. The separation between one pair of horizontal lines was filled to provide a discernible feature, later used to vertically register the beams

Two steps are required for calibration resulting in two sets of images to be acquired. First step is to calibrate n [A-lines] to y [mm], requiring a stationary image of the calibration pattern with the vertical lines aligned perpendicular to the scan lines. The second step is to calibrate the scanning pattern, requiring images of the calibration pattern pushed at a known and constant velocity with the vertical lines aligned parallel to the scan lines.

3.1 Calibrate n to y

The first step of calibration is to determine the y positions of the beams as a function of n which is possible from the stationary image of lines aligned perpendicular to the scan lines. The resulting images from each beam are a series of horizontal lines, shown in Fig. 16, which represent the y-axis extent of each scan line, from which the scan line lengths can be determined. To do so, the images are thresholded and the negative-going edges of the lines are extracted, empirically fitted, and optimized using least-squares minimization. The resulting scan line lengths are 10.209 mm and 10.285 mm for beams A and B respectively. With n_{roi} = 743 A-lines, the corresponding average scan resolution, denoted q, is 0.0138 mm/A-line.

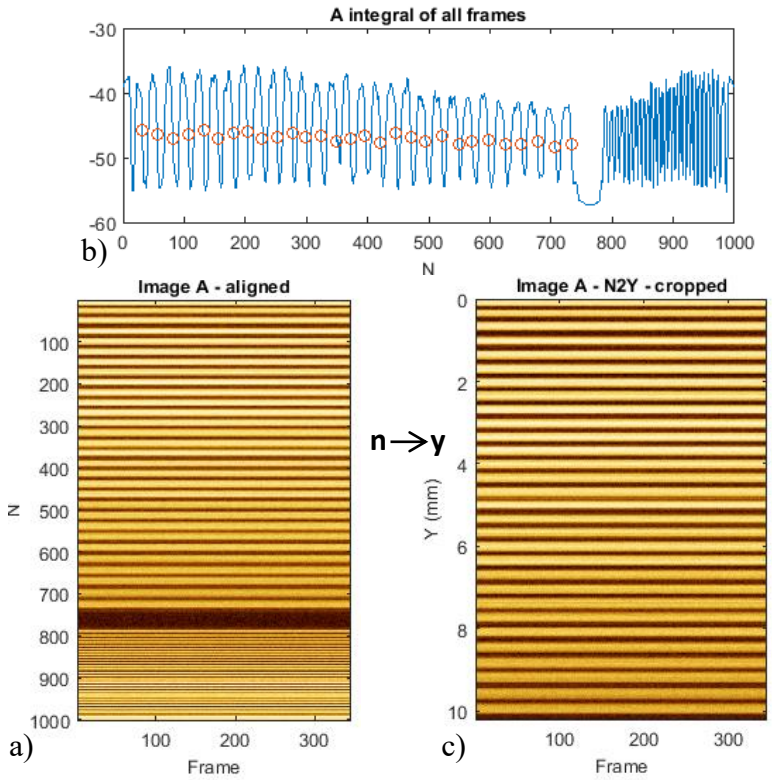

Fig. 16. n-to-y calibration for each beam is determined from stationary images of lines, with known separation, aligned perpendicular to the scan lines a). The integral is taken across all frames b) in blue and the negative going edges are identified b) in red. The negative going edges are empirically fitted to determine the mm displacement of the beam at any point in one frame. This provides the scan line length and the conversion factor to convert the active region of the scan from A-lines to mm c).

3.2 Calibrate scanning pattern

The second step of calibration is to determine the scanning pattern and spatial relationship of both beams and is possible from the constant velocity images of lines parallel to the scan lines. This was achieved by imaging the calibration phantom aligned parallel to the scan lines at v_x = 10 mm/s using a linear stepper motor. The calibration images are preprocessed, shown in Fig. 17a), then cropped and thresholded to isolate and binarize the vertical lines. From the isolated, binary vertical lines, all edges are extracted and used to determine the frame separation, or time delay, of each beam capturing the same vertical line in each row of the image. In other words, the frame index of each edge in each row for both beams is used to determine when each beam intersects the same edge in the scan, thereby defining the scanning pattern as a function of time.

In previous work on DMDI [24], the functions scan pattern, S_{fr}, and fixed pattern, F_{fr}, given in Eq. 3 and 4 respectively, in units of frames were defined. The scan pattern describes the distance between each beam at a given row index, or the y-axis in this implementation, in a single frame. The fixed pattern describes the deviation of each scan pattern from the perfect geometry, or a straight line in this implementation, in a single frame.

$$S_{fr}(y) = \frac{1}{2}[b_A(y) - b_B(y)] \tag{3}$$

$$F_{fr}(y) = \frac{1}{2}[b_A(y) + b_B(y)] \tag{4}$$

Similar to the first calibration step, S and F are empirically fitted and optimized using least-squares minimization from simple linear regressions of each row. The linear regressions of each row provides a slope (m) indicating the spacing between the vertical lines, which should be a constant, and intercept (b_A and b_B) indicating the average shape of the vertical lines as a function of y. The difference of these intercepts, plotted in Fig. 17, indicates the difference of frames of each vertical line between both images in each row, or in other words the scan line separation in frames as a function of y. The mean scan line separation in frames, denoted Δfr_{AB} and plotted in red in Fig. 17, is 23.2211 frames. Given v_x = 10 mm/s and a frame rate of 100 frames/s, the scan line separation, denoted s_o, is 2.32 mm using Eq. 5.

$$s_o = \frac{v_x}{f_{frame}} \cdot \Delta fr_{AB} \tag{5}$$

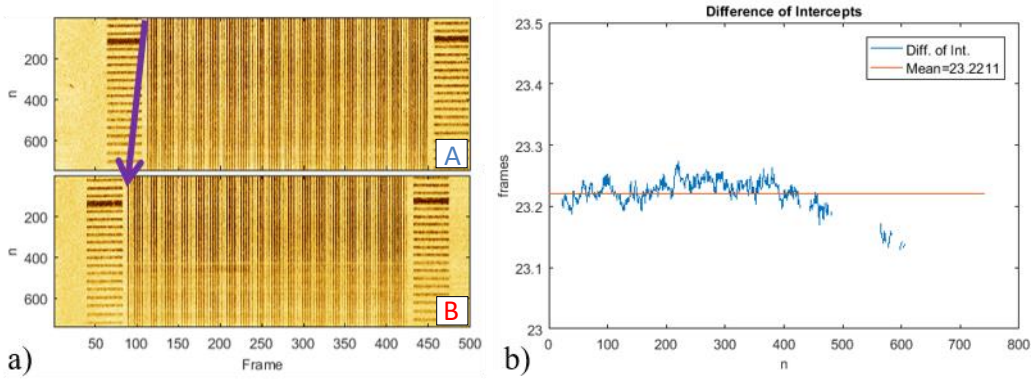

Fig. 17. Scan line length calibration. The scan line separation, is determined from the average frame separation of the vertical lines in both images, shown in a), for each row (shown by purple arrow), which is plotted in blue along the x-axis of the b) plot. The red line indicates the average frame separation for each row, equivalent to 23.2211 frames.

The largest deviation from the mean observed is <0.1 frames, which is equivalent to 10 µm. Given that the resolution of this optical set up is 80 µm, the assumption that the scan lines are linear and parallel will be used. This assumption simplifies Eq. 3 and 4 to constants, as shown by Eq. 6 and 7 of S and F in units of distance.

$$S = \frac{1}{2}s_o \tag{6}$$

$$F = f_o \tag{7}$$

where f_o is a constant representing an arbitrary horizontal offset of both images.

Furthermore, as indicated by a green arrow in Fig. 18, there exists a vertical offset between images A and B. The vertical offset, denoted y_{off}, was determined from the A-line difference of the negative going edges produced by the thicker line feature in both images, made possible by observing the frame integral of a region of interest for each image. The offset is 21 A-lines which is equivalent to 0.290 mm.

This indicates that scan line B is shifted 0.290 mm up from scan line A and captures slightly different regions of the sample. This offset is due to the imperfect manual alignment of the imaging head described in Chapter 2.

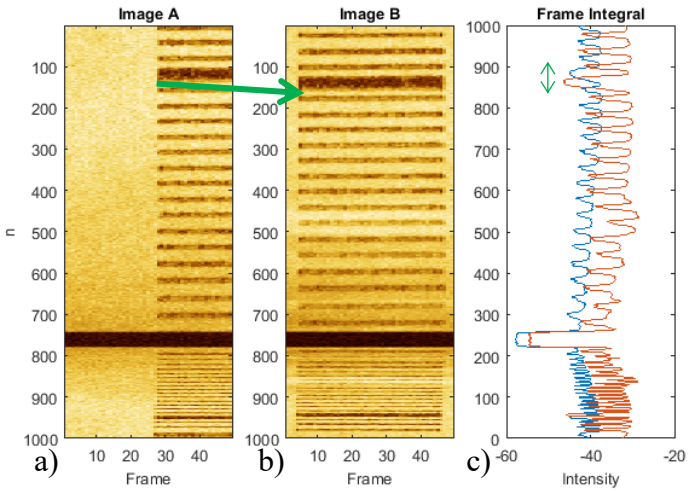

Fig. 18. A vertical offset of 21 A-lines or 0.290mm is observed. An intensity integral, c), across all frames within regions of interest of images a) A and b) B was calculated to observe the A-line difference of the two images capturing the same thicker horizontal line feature.

The resulting calibration information, shown in a), characterizes the scanning pattern. If the sample is pushed at a constant v_x with $v_y = 0$, i.e. aligned perpendicular to the scan lines, then the corrected images can be mapped from the input images with x-axis linearly scaling by displacement per frame resulting from v_x and y-axis linearly scaling by q from the first step of calibration. Furthermore, the corrected images are registered by horizontal and vertical translations equivalent to S and y_{off} respectively. For example, the calibration images are corrected using the calibration results and v_x = 10 mm/s, shown in Fig. 19b). However, if the velocity profiles are unknown or not constant, they must be estimated and utilized to produce distortion corrected images, as described in the next chapter.

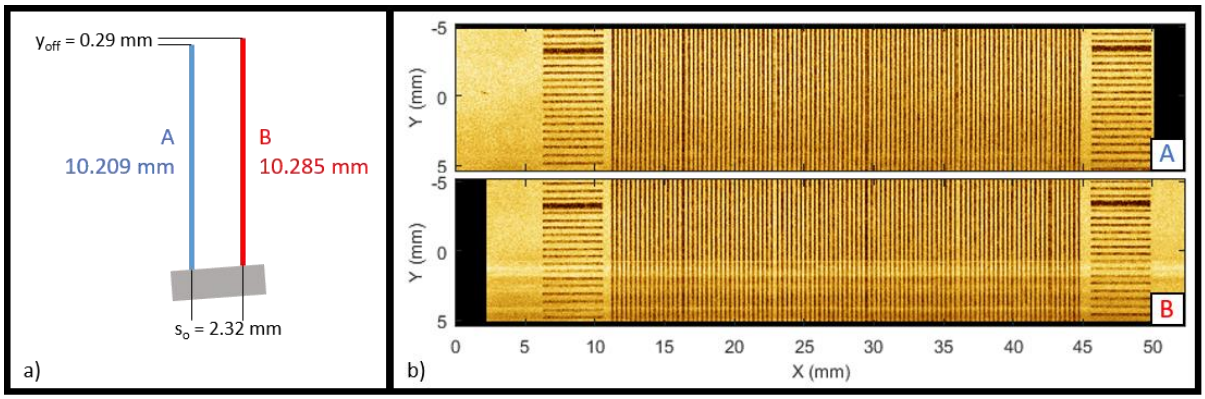

Fig. 19. a) Illustration of the calibrated scanning pattern. b) Corrected images of the calibration phantom imaged at v_x = 10 mm/s.

4 Automated Distortion Correction

In DMDI, distortion correction is achieved by estimating the effective beam path of the imaging beams onto the sample, achieved through combining the beam displacement due to mechanical scanning and the displacement due to actuation of either the sample or the imaging head. The actuation displacement can be estimated by identifying unique and common features in images from both beams. Given the time interval of imaging a sample feature by both beams and the known spatial relationship between the beams, an average velocity can be calculated for that time interval. With many common and unique features, a velocity profile, and therefore displacement, can be estimated.

A challenge for identifying these common and unique features is the fact that scan patterns and velocity profiles can cause the same sample feature to appear very different in each image. As sample features can appear very different, features cannot be easily extracted using automated feature detectors, which are commonly based on feature appearance or geometric transformations such as scaling or affine. Thus, in the first demonstration of DMDI [24] feature matching was done manually by selecting control points to co-register the two images, resulting in a processing time on the range of hours. However, due to the scanning configuration of this implementation, the scan pattern is simplified to two roughly parallel lines, meaning that both beams will capture similar sample regions in a given single frame. With parallel scan lines, frames or portions of frames in each image can be directly compared to identify when common and unique sample features were captured, which is not possible with more complicated scan patterns.

For this implementation of DMDI, distortion correction is achieved by estimating the x and y velocities as a function of time, $v_x(t)$ and $v_y(t)$ respectively, in order to calculate the sample displacement and interpolate the corrected images. The velocity functions can be estimated from feature-matching frame pairs which are found using automated frame correlation.

This chapter describes the methods for extracting feature-matching pairs, estimating the velocity profiles, and producing the distortion corrected images with each step illustrated by an example with manually actuated images of a black QR code phantom printed on standard white printer paper. The sample was fixed onto a 1D translating stage attached perpendicular to a sliding rail which was aligned perpendicular to the scan lines. The translating stage allows the sample to be pushed back and forth in the y-axis while pushing the stage along the sliding rail allows the sample to be pushed in the x-axis. This configuration allows for manual motion in the x and y axes without allowing for the sample to rotate. The distorted preprocessed images are shown in Fig. 20. In this example, the sample moves in the positive x-axis such that scan line A captures sample features before scan line B. In this scenario, image A is analogous to lead image and image B is analogous to lag image.

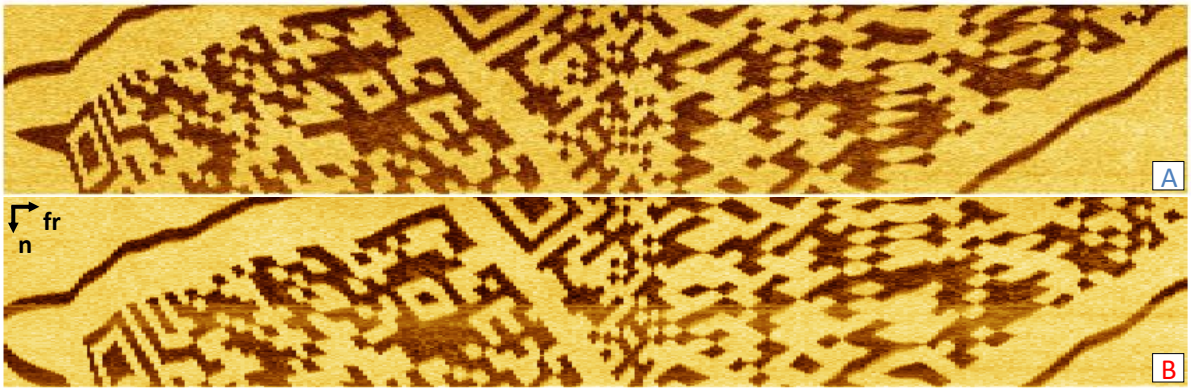

Fig. 20. Preprocessed distorted A and B images of a manually actuated black QR code printed on standard white printer paper. The smallest features are 0.53x0.53 mm^2.

4.1 Automated frame correlation

Due to the scanning configuration of this implementation, scan lines are roughly parallel and therefore individual frames, or segments of frames, from each image which pass the same feature are similar, even with varying velocity patterns. Thus, frames of each image can be compared to find common and unique features. The similarity measure used to compare frames is normalized cross correlation which allows for sensitivity of image structures without sensitivity for intensity differences, since images A and B may have different intensity profiles.

The frames of images A and B which capture the same feature, called feature-matching pairs, are determined by extracting points from a 3D correlation coefficient matrix constructed to indicate how strongly a segment of a given lead image frame is correlated to a full lag image frame at a particular vertical shift, or Δy_{AB}, to one another. If $v_y = 0$, there will be no vertical shift and Δy_{AB} will equal a constant value, Δy_o. The matrix, visualized in Fig. 21, is structured such that dimension 1 and 2 are A and B image frames while dimension 3, or the depth dimension is Δy_{AB}, and the value saved at a given index is the normalized cross correlation coefficient of that combination, calculated using MATLAB's normxcorr2 [27]. A segment of a given lead image frame is defined as a frame with 100 A-lines cropped from the top and bottom. Effectively, the lead image frame segment is being scanned along each y index of the lag image frame to calculate a correlation coefficient of the overlapping information. Using only a segment of the lead image frame reduces the calculation time of the cross correlation at the cost of limiting the maximum measurable v_y.

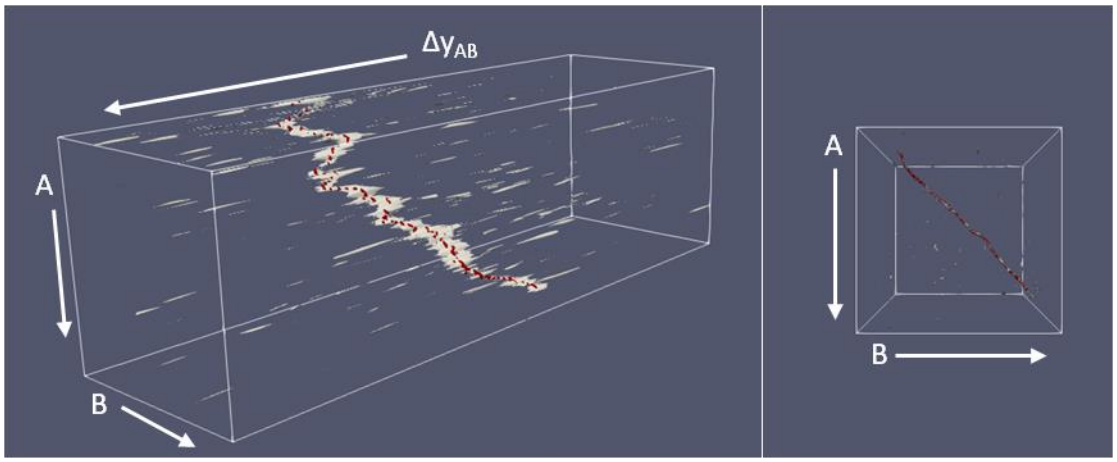

Fig. 21. 3D visualization of correlation coefficient matrix illustrated using Paraview. Grey contour constructed from a correlation coefficient threshold of 0.7. Red indicates the pathway extracted.

The feature-matching pairs are extracted from the matrix by determining the path of high correlation values across the volume, shown in red in Fig. 21. The pathway corresponds to combinations of A frame, B frame, and Δy_{AB} which correlate highly. This indicates matching image structures, or matching sample features, in each image at those instances during the scan. The pathway is found in pseudo 3D, shown in Fig. 22, by first taking the maximum intensity projections of two planes, the Δy_{AB}-lead plane shown in Fig. 22a(i), and the lead-lag plane shown in Fig. 22b(i). The projections are further reduced by finding one maximum per lead frame, shown in Fig. 22a(ii) and Fig. 22b(ii) respectively. Reducing the data to one maximum per lead frame ensures that the points that are kept in the final path are those of higher correlation and therefore higher confidence in matching. Note that individual data points, shown as black points in Fig. 22, have been thickened to be more visible.

First the reduced Δy_{AB}-lead projection is processed. A quasi-Euclidean distance map, shown in Fig. 22a(iii), is calculated and thresholded to create blobs, indicating regions of point(s) in the map that have a neighbor within 3 pixels, which are labeled, shown in Fig. 22a(iv). Then, starting at the first lead frame and $\Delta y_{AB} = \Delta y_o$, the next Euclidean-closest blob is selected, from which the end of that blob is found and the next closest blob is determined. This is repeated until an end condition, determined by estimating the last frame in the lead image which captures a common feature in the lag image, is reached. The resulting pathway of blobs, shown in Fig. 22a(v), is used as a mask on the reduced projection to provide a filtered pathway of high correlation values across the Δy_{AB}-lead plane, shown in Fig. 22a(vi). Practically this means that individual points with a Δy_{AB} value too far from the pathway are filtered out. The points filtered out are used to further filter the lead-lag reduced projection, resulting in Fig. 22b(iii) and the same process is repeated, shown in Fig. 22b(iv)-(vii). The information encoded in the final pathway is the feature-matching pair which indicates the time interval of both imaging channels capturing the same feature and the corresponding Δy_{AB} which describes displacement in the y-axis.

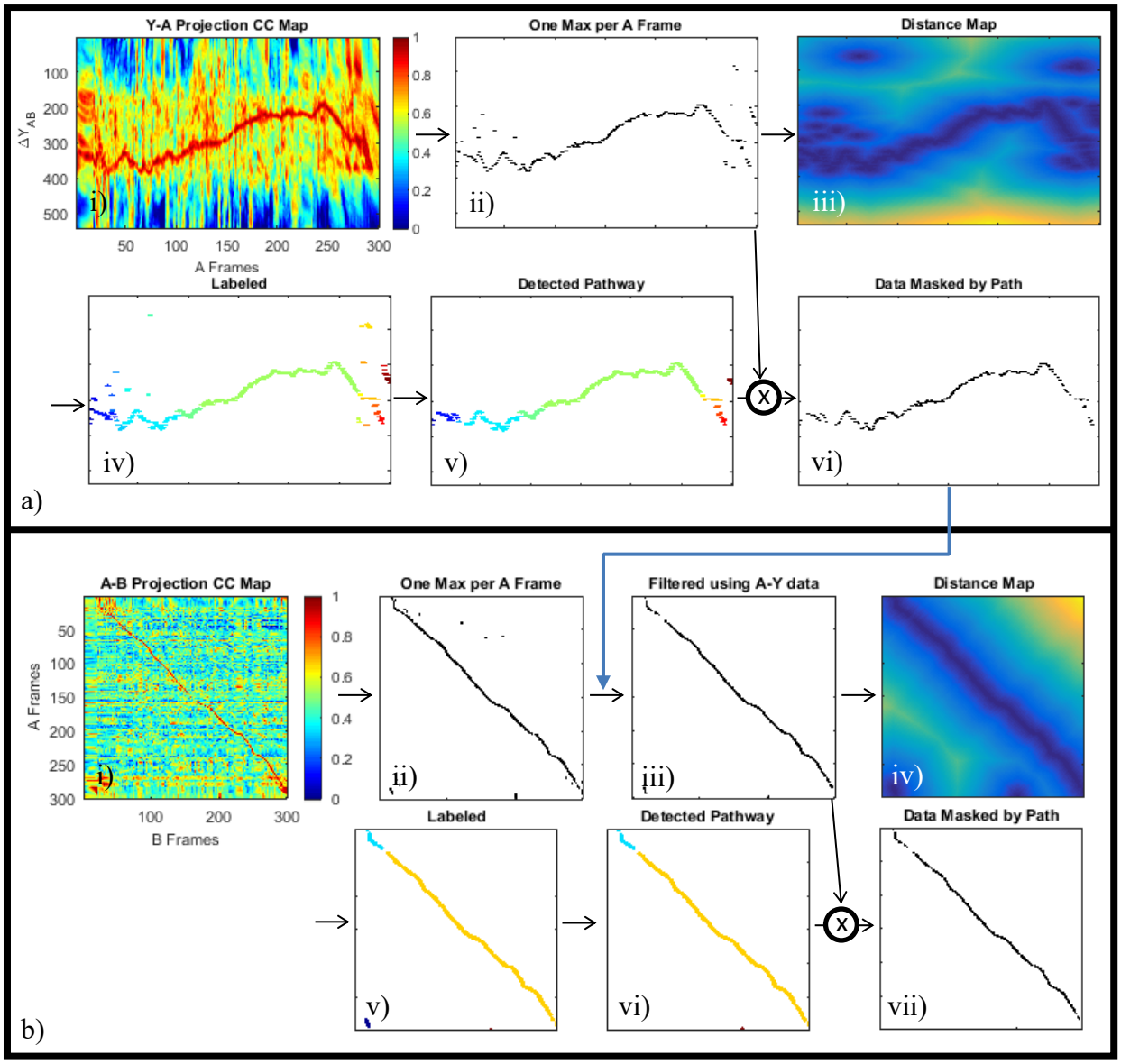

Fig. 22. Automated frame correlation results for manually actuated images of a QR code phantom with smallest features having a width of 0.53 mm. a) Pathway extraction from ΔyAB-A maximum intensity projection a(i). First the projection is reduced to one maximum per A frame (ii), then a quasi-Euclidean distance map is calculated a(iii), thresholded and labeled a(iv) such that pixels with neighbours within 3 pixels are maintained, discarding single lone pixels which have a high probability of being noise. From a(iv) a pathway is found a(v) across the blobs, by minimizing the distance between blobs that are kept, and the pathway is used as a mask on a(ii) to produce the final pathway of highly correlating points a(vi). b) The same process is repeated for the A-B maximum intensity projection with the exception of first filtering out the points removed in a(vi) from the reduced A-B projection b(ii) resulting in b(iii). The final extracted pathway b(vii) encodes the information of A and B frames which highly correlate and therefore are likely to be matching features.

4.2 Velocity estimate

Within the time interval of feature-matching pairs, an average x velocity is determined by the scan line separation while an average y velocity is determined by the vertical shift required to best match the frames in the feature-matching pair, or Δy_{AB}. The average v_y of a feature-matching pair is zero if there is no vertical shift, in which case Δy_{AB} is equal to Δy_o. The time interval of the feature-matching

pair is determined from the difference in time for a sample feature k to appear in both images. This time occurrence $t_{A,k}$ and $t_{B,k}$ are given by

$$t_{AB,k} = \frac{(fr_{B,k} - fr_{A,k})}{f_{fr}} \tag{8}$$

where $fr_{A,k}$ and $fr_{B,k}$ are the frames of a matching feature k in images A and B and f_{fr} is the frame rate equal to 100 frames/s. Given the time each beam captures feature k, the calibrated scan line separation s_o, and Δy_{AB}, the average velocity within k time interval, $\bar{v}_{x,k}$ and $\bar{v}_{y,k}$, can be calculated by Eq. 9 and 10 respectively.

$$\bar{v}_{x,k} = \frac{s_o}{t_{B,k} - t_{A,k}} \tag{9}$$

$$\bar{v}_{y,k} = \frac{q(\Delta y_{AB} - \Delta y_o)}{t_{B,k} - t_{A,k}} \tag{10}$$

With overlapping time intervals $(t_B,t_A)_k$, the instantaneous velocities $v_x(t)$ and $v_y(t)$ can be estimated as

$$v_{(x,y)}(t) \approx \frac{\sum_k \left(\bar{v}_{(x,y),AB,k} \cdot w_k(t)\right)}{\sum_k w_k(t)} \tag{11}$$

where w is the normalized $\left(\sum_{t=t_{A,k}}^{t_{B,k}} w_k(t) = 1\right)$ weighting function defined as

$$w_k(t) = \begin{cases} \dfrac{1}{t_{B,k} - t_{A,k} + 1}, & t \in [t_{A,k}, t_{B,k}] \\ 0, & elsewhere \end{cases} \tag{12}$$

The weighting ensures each feature-matching pair equally contributes when integrated over all time to determine the corresponding displacements, $d_x(t)$ and $d_y(t)$. Any gaps in the velocity functions, caused by little or no feature-matching data, are filled by cubic interpolation before integration. The estimated velocities and displacements of this phantom example are shown in Fig. 23.

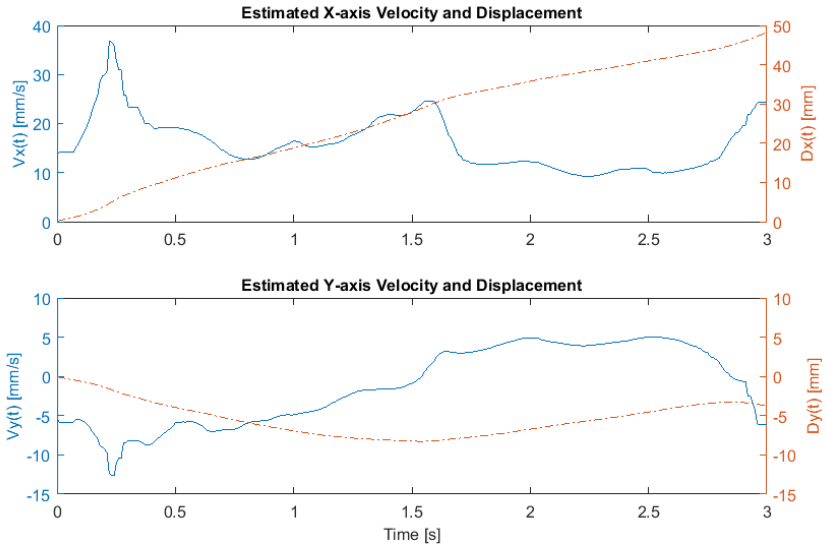

Fig. 23. Velocity [left, blue axis] and displacement [right, red axis] estimates in the x-axis [top] and y-axis [bottom] for the manually actuated QR code example.

4.3 Generate distortion corrected images

To generate distortion corrected images, the images must be adjusted to account for the scanning pattern and overall displacement of the sample, effectively converting from A-lines and frames to mm. The coordinate of any point in images A and B is given by

$$x_A(t) = d_x(t) + S - F \tag{13}$$

$$y_A(t) = q \cdot n_A(t) + d_y(t) + y_o \tag{14}$$

$$x_B(t) = d_x(t) - S - F \tag{15}$$

$$y_B(t) = q \cdot n_B(t) + d_y(t) \tag{16}$$

where $q \cdot n_A(t)$ and $q \cdot n_B(t)$ describe the y-axis location in mm of beams A and B respectively at a given time. As F is a constant offset subtracted from both A and B, it can be ignored. Once the corrected image positions are determined from Eq. 13-16, the corrected image pixel intensities are linearly interpolated and plotted, producing scan pattern and sample or imaging head distortion corrected images. The distortion correct images of this phantom example are shown in Fig. 24.

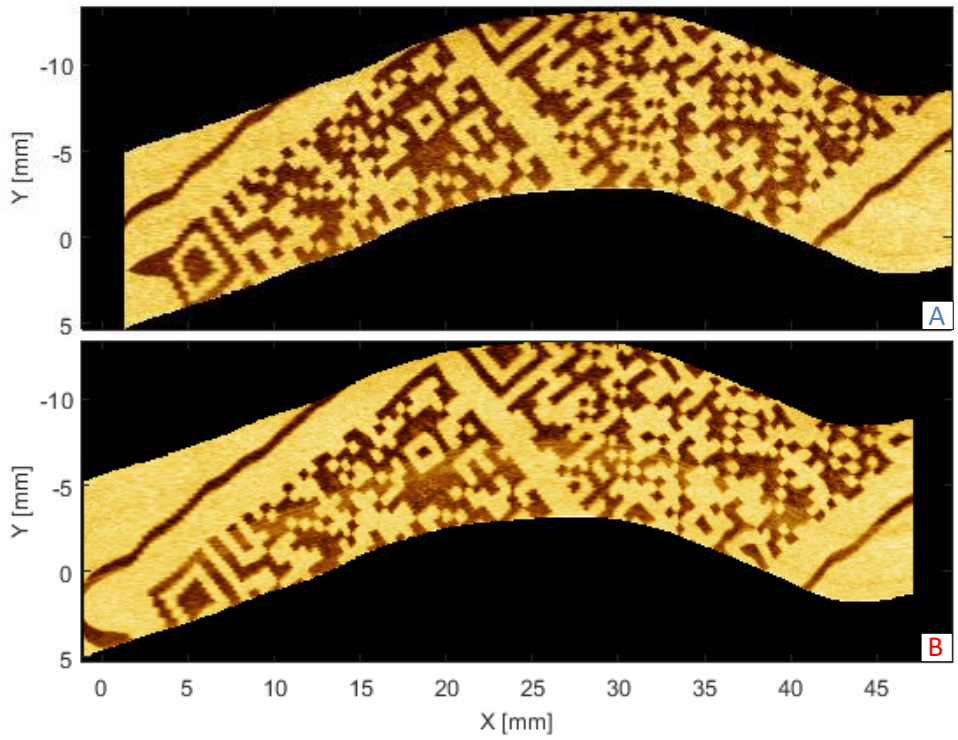

Fig. 24. Distortion corrected images of the manually actuated QR code phantom.

The distortion corrected images can be overlaid on each other, shown in Fig. 25b), to infer the correction accuracy qualitatively. To produce this image, the distortion corrected images A and B are contrast adjusted and normalized in order to produce two images with comparable intensity profiles. Then the lead image is saved into the red channel of a RGB image structure and the lag image is saved into the blue and green channels. In the resulting RGB overlay image, cyan intensity indicates lag image without overlapping lead image, red intensity indicates lead image without overlapping lag image, white indicates both images are high intensity together, and black indicates both images are low

intensity together. In other words, grayscale in the overlapping regions indicates where both images are similar and colours indicate where both images differ. Distortion correction can be further assessed qualitatively by comparing the corrected images to the original sample, shown in Fig. 25a). The size and shape of the features in the corrected images are similar to that of the original sample.

The distortion correction processing time for this example was 86 s.

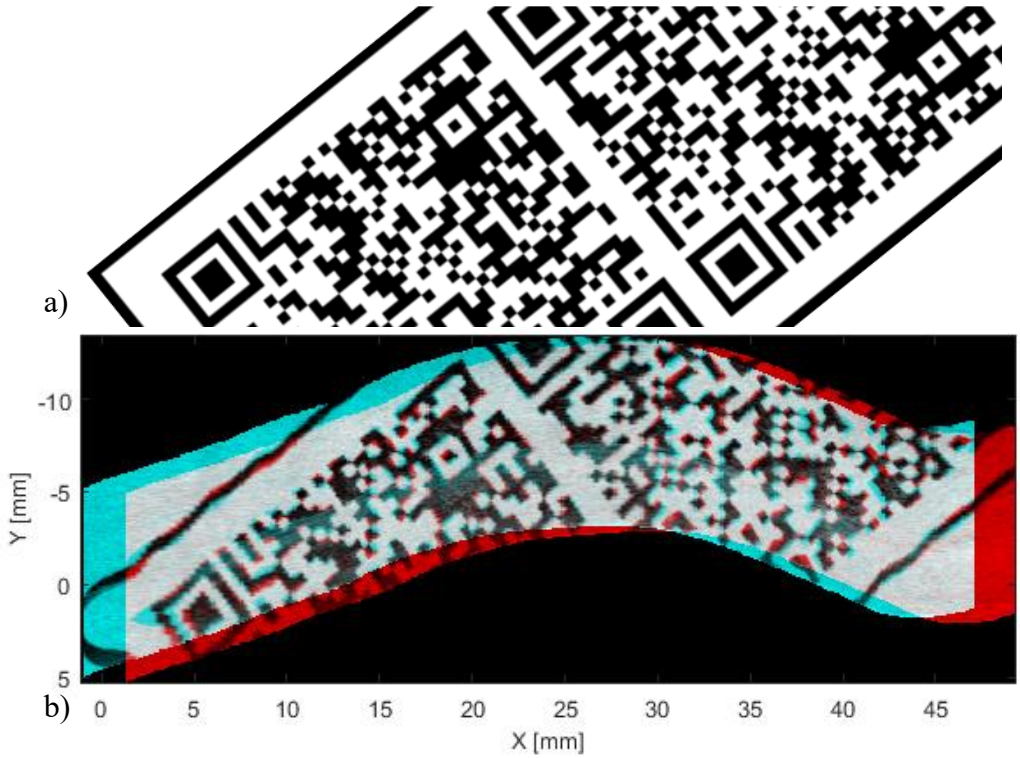

Fig. 25. Qualitatively assess distortion correction results. a) The original digital QR code phantom. b) Overlay image of phantom distortion corrected images. Grayscale in the overlapping regions indicate where both images are similar whereas colors indicate where the images differ.

4.4 Discussion

In this chapter, the method for automated distortion correction was presented. As this method is automated, the processing of images is significantly faster, 86 s compared to on the range of hours, and far less tedious than in the previous work [24]. Through automation, DMDI becomes a more accessible and attractive modality for potential clinical applications.

The QR code phantom allowed for a sample with high contrast and control over the density and uniqueness of features. That is, with the QR code rotated, columns of the sample are unique and rather dense in information. Therefore matching frames of the two images will correlate highly with each other compared to other non-matching frames, producing a strong and clear pathway through the correlation matrix.

The automated distortion algorithm performs well at extracting the manually actuated velocity profiles and correcting the distorted images, as can be seen by the corrected images in Fig. 24 and

corresponding overlay in Fig. 25. Although some non-matching regions in the overlay image demonstrate that the correction is not perfect. Near the beginning of the scan, the first square locator of the QR code is distorted in the corrected A image. This can be attributed to the lack of features in this region of the scan, resulting in a poor velocity estimate. Additionally, poor velocity estimation at the beginning and end of the scan can be attributed to less overlapping feature-matching information.

The v_x limits of this implementation of DMDI are governed by the scan line separation, the frame rate, and the total scan time. The maximum and minimum measurable v_x are given by

$$v_{x,max} = s_o \cdot f_{fr} \tag{17}$$

$$v_{x,min} = \frac{2 \cdot s_o}{t_{scan}} \tag{18}$$

With a scan line separation of 2.32 mm, $v_{x,max}$ is 232 mm/s, however the effective maximum would be smaller, depending on the sample feature size and the desired scanning resolution. At low velocities, a long scan time and therefore a large amount of data would be required to capture enough features in both images to make a velocity estimate. Thus smaller scan line separations may be more desirable for lower velocities. With large scan line separation, the velocity estimate is averaged over a larger interval compared to that of smaller separations. This results in lower velocity estimate accuracy, especially at lower speeds. Depending on a given application, a trade off decision must be made: choose larger separations to be sensitive to larger velocities or choose smaller separations to have higher accuracy at small speeds. In this optical set up the scan line separation is fixed, therefore the imaging system and velocity distortion was simulated with variable beam separations, found in Chapter 6, to investigate the influence of scan line separation of correction accuracy.

Since velocities are estimated from the time interval of a sample feature passing both beams, v_y is inherently dependent on v_x. Due to the nature of how v_y is calculated, zero velocity is possible and therefore there is no $v_{y,min}$. The maximum measurable v_y is given by

$$v_{y,max} = \pm \frac{q \cdot (n_{roi} - n_{win})}{s_o} \cdot v_{x,max} \tag{19}$$

where n_{roi} is the number of forward scanning A-lines equal to 743 A-lines and n_{win} is the window size selected to reduce the size of the searching field and therefore reduce the correlation coefficient calculation time. For all correction cases presented in the work, n_{win} is equal to 100 A-lines. Decreasing n_{win} increases the overlapping information to be compared in A and B and allows for higher sensitivity to larger Δy_{AB} and thus larger y velocities. Since n_{win} is cropped from both the top and bottom of a given lead frame, the same velocity sensitivity is achieved in both positive and negative directions. At increasing v_y, it is expected that the overlapping sample information in matching frames will reduce and correction accuracy will also reduce. Thus, the effects of increasing v_y on correction accuracy are investigated in Chapter 6.

Analyzing the correlation matrix and accompanying maximum intensity projections can shed insight on the scan. Position and shape of the pathway correspond to the velocity profile. For example, in the lead-lag projection, a pathway points located near the diagonal corresponds to higher velocities compared to points located further away from the diagonal. Curved pathways indicate variable velocity profiles while linear pathways indicate constant velocities. Additionally, if v_x is monotonic, then the resulting pathway will only appear in one of the triangular matrices of the lead-lag maximum intensity projection. Note that in this work, samples are restricted to monotonic v_x and the correlation matrix is constructed such that the pathway is always in the upper triangular matrix. This was to allow for omitting the lower triangular matrix in the correlation calculation in order to reduce calculation time by a factor of two and speed up the procedure whilst developing the method. However, in principle, it is possible to adapt the method to allow for and detect v_x in both directions. Non-monotonic v_x would appear as a lead-lag pathway which exists on both sides of the diagonal. *A priori* knowledge of the velocity profile expected in a given application can be exploited to reduce the size and therefore calculation time of the correlation matrix, such as excluding the lower triangular matrix of a monotonic v_x.

Continuing on the concept of calculation time, it should be possible to optimize this method for calculation time such that it can be implemented in real-time. As already mentioned, calculating the correlation matrix, which is the most costly aspect of the technique, can be reduced and optimized to run faster. In addition, developing this technique outside of MATLAB can drastically decrease the calculation time. However, the requirement of two beams passing the same sample feature to achieve distortion correction implies that distortion correction can never be instantaneous, but instead live updating as feature matching information becomes available.

Fundamentally, the technique of automated frame correlation is similar to that of dynamic time warping (DTW) [28, 29]. DTW is a commonly applied technique used in signal processing to assess the similarity of two time series. The two time series are compared to determine an optimal match with non-linear time, such that the times series can be ‘warped’ in time to determine a measure of similarity. A ‘warped path’ is produced which describes how the two series may be aligned in time. In automated frame correlation, two images are compared to determine an optimal match of individual features, producing a pathway analogous to a ‘warped path’. With one dimension of motion, for instance $v_y = 0$, the pathway produced describes how the two images align in the x-axis, or in time. Both techniques can be used to determine how the two images align in time and can be warped to generate two corrected images. This can be demonstrated by using both methods to estimate a variable velocity profile applied in a simulation of distorted images, shown in Fig. 26. The simulated velocity (blue) is closely estimated by both methods with automated frame correlation (yellow) performing slightly better than DTW (red), especially at the high peaks.

Although both techniques perform similarly for one dimension of actuation, accounting for warping in a second dimension, for example nonzero v_y, is not trivial in DTW. Automated frame correlation as presented in this work could be thought of as two dimensional time warping. Additionally, DTW is typically used to match two time series or see how to distort a sample series to match a reference series whereas in this application two distorted images are compared to create non-distorted images.

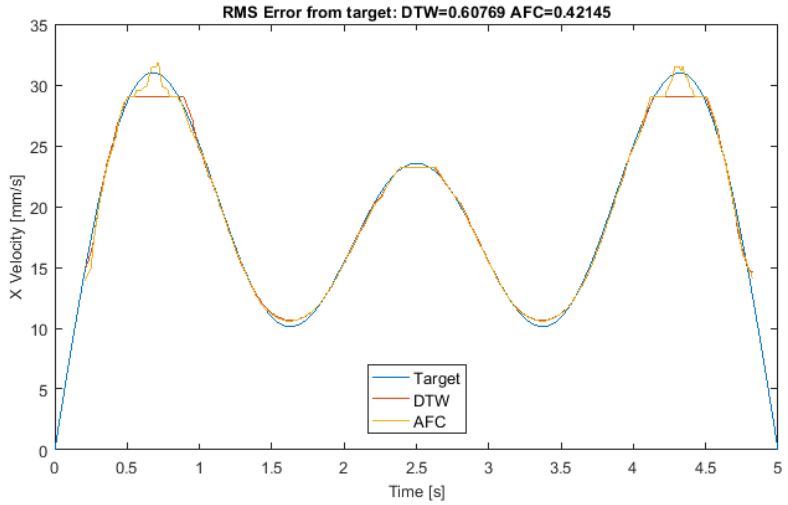

Fig. 26. Comparing DTW and automated frame correlation (AFC) functionality in 1D motion correction. A pair of distorted images were simulated with v_x equal to the variable velocity shown as target [blue] and $v_y = 0$. The velocity was estimated using DTW [red] and AFC [yellow] and compared to the target velocity. AFC performed slightly better than DTW, especially at the higher peaks.

Automated frame correlation utilizes normalized cross correlation as the similarity measure to compare image frames. As both images are captured using the same modality, a linear similarity measure is suitable. Since images A and B can differ in brightness and therefore have different intensity profiles, normalized cross correlation is best suited. However in other applications, different similarity measures may be better suited. For example, if images A and B are of different imaging modalities, then perhaps entropy or mutual information, both non-parametric similarity measures, are more fitting. This method is flexible in the choice of similarity measure, making it accessible for a range of possible applications.

5 Testing

This chapter describes various tests which were performed to assess the performance of the distortion correction algorithm and discusses the corresponding results. First validation tests were made to assess the technique's performance in extracting known constant velocity profiles from biological samples. Second, manual actuation tests were made to assess the technique's performance in extracting unknown and variable velocity profiles. Lastly, *in vivo* application feasibility was evaluated through imaging of a fingerprint.

5.1 Validation

To validate the velocity estimate of the distortion correction algorithm, two different biological samples, beef tongue and kiwi, were imaged while being pushed at known constant velocities, specifically v_x = 10 mm/s and v_y = 0, using a linear stepper motor. By imaging with a known velocity profile, one can infer on the accuracy of the velocity estimation. It is important to note that the linear stepper motor was manually aligned perpendicular to the scan lines such that v_y would be zero. However, it is possible that the alignment was not perfectly orthogonal, resulting in a non-zero v_y component related to the difference in angle from 90 degrees. Error in the v_y estimates due to the algorithm in these validation examples may be confounded with the human error attributed to manual alignment. Furthermore, in these examples, it is assumed that the linear motor in fact pushes at exactly 10 mm/s however there is likely small error in this assumption due to manufacturing and operation tolerances.

5.1.1 Beef tongue

Cow tongues are heavily keratinized for eating and have various forms of papillae covering the surface. The sample used in this work, shown in Fig. 27e), consists of many similarly appearing filiform papillae with the exception of two larger fungiform papillae seen in the middle of the sample. Dark regions in the OCT images, Fig. 27a), correspond to the bright papilla bases in the sample photo and the papilla can be observed protruding from the base, typically downward, in both the OCT images and the sample photo. As there are many similarly appearing features, the cross correlation matrix, projections shown in Fig. 27c), displays relatively consistent correlation values throughout with the feature-matching pairs pathway having only slightly higher correlation values compared to the rest of the volume. The automated frame correlation algorithm is successful in isolating the feature-matching pair pathway from a matrix with similar correlation values. The resulting velocity profiles, shown in Fig. 27d), are close in estimating the actual profiles. Since the velocities of the sample are constant, it is expected that the extracted pathways are straight lines, which is the case in this example with the exception of some noise.

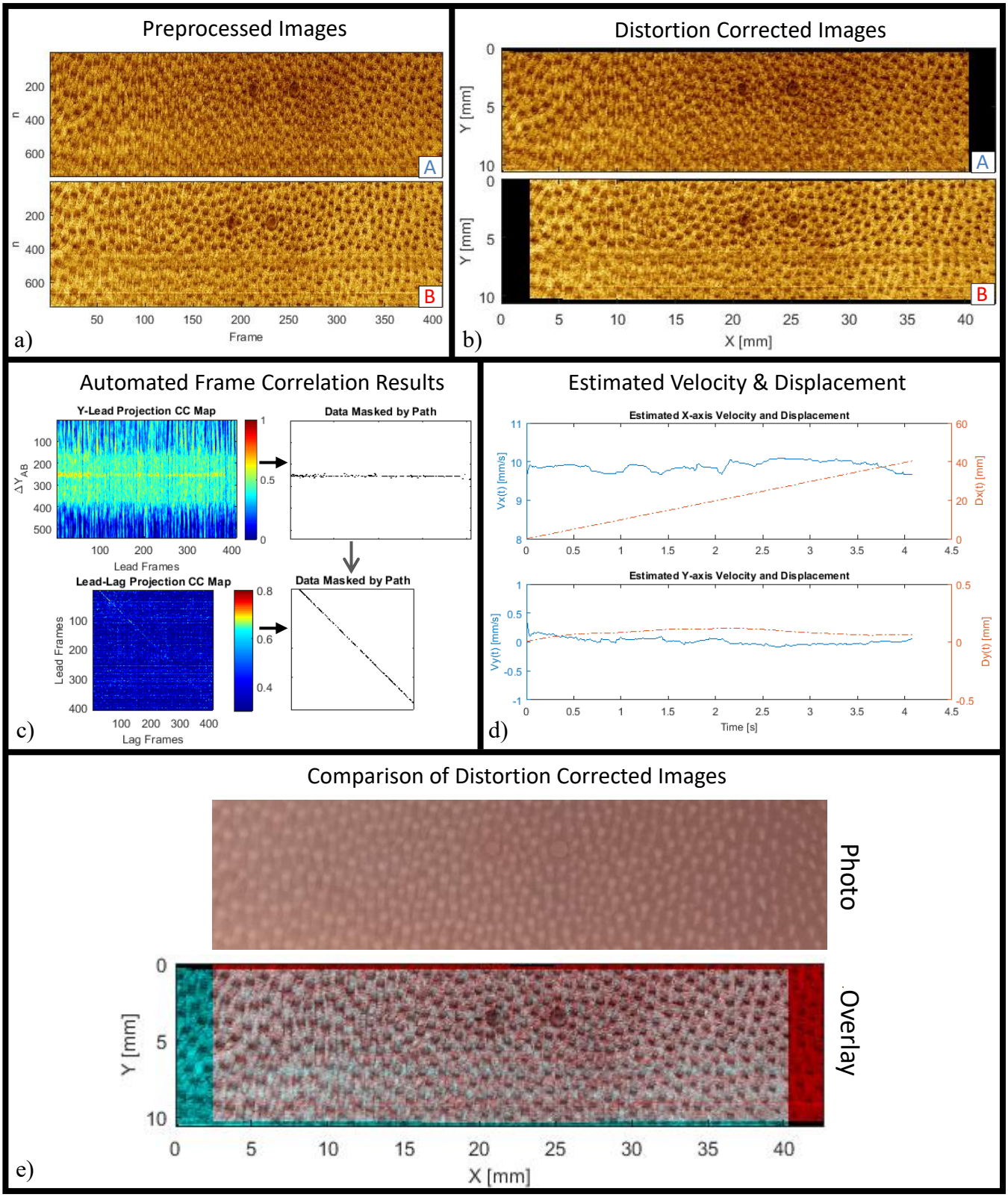

Fig. 27. Distortion correction results of a beef tongue sample imaged at v_x = 10 mm/s perpendicular to the scan lines such that $v_y \approx 0$ mm/s. a) Preprocessed images A and B. b) Resulting distortion corrected images A and B. c) The automated frame correlation results including the two correlation coefficient matrix maximum intensity projections and the resulting isolated pathways. d) The velocity and displacement estimates. e) An overlay of the distortion corrected images to qualitatively infer correction accuracy and compare to the original photo of the sample. Cyan indicates image A, red indicates image B, and in the overlapping region, graysale indicates image similarities

The average estimated v_x is 9.8582 mm/s, corresponding to 1.42% error if it is assumed that the linear stepper motor actuated at exactly 10 mm/s. The average v_y estimated is 0.0029 mm/s is only slightly off compared to the expected 0 mm/s. In this validation example, the algorithm performs well when analyzing similar looking and repetitive features and the result distortion corrected images, Fig. 27b), appear similar, as can be seen in the overlay image shown in Fig. 27e). When comparing the corrected images to the original sample photo, one can see that papilla features appear similar in size and shape.

5.1.2 Kiwi

The kiwi cross section sample provides high contrast and differently shaped features including the round black seeds, the homogenous (green) flesh, and the bright (white) linear fibrous structures. As can be seen in Fig. 28a), the denser features including the core, seeds, and fibrous structures appear bright while the flesh appears dark in the OCT images. Considering features are inherently high contrast and appear different from one another in shape, it is expected that the correlation matrix would exhibit higher correlation values across the pathway compared to the rest of the matrix as well as that of the beef tongue. This is observed in Fig. 28c) as the correlation pathway appears brighter/hotter comparatively. The estimated velocities, plotted in Fig. 28d), are on par with the expected input velocity profiles, even more so than the beef tongue. The average estimated v_x is 9.9453 mm/s, corresponding to 0.55% error and the average estimated v_y is 0.0012 mm/s compared to an expected 0 mm/s. The resulting distortion corrected images, shown in Fig. 28b), are overlaid in Fig. 28e) to qualitatively show how well the images were corrected. When the corrected images are compared to the original sample photo, one can see the kiwi features align well and are of similar size and shape. Qualitatively and quantitatively, the kiwi results are better compared to the beef tongue.

With this validation experiment, one can conclude that the higher contrast and unique looking features of the kiwi allows for a better velocity estimation compared to the similar and repetitive features of the beef tongue, however both produced reasonably good results. The velocity estimation performs well when extracting constant v_x and zero v_y. For more rigorous testing, one may repeat the experiment multiple times and with different constant velocities to analyze how repeatable the results are and if there are any velocity sensitivities. Instead this was achieved through simulation, later discussed in Chapter 6**.**

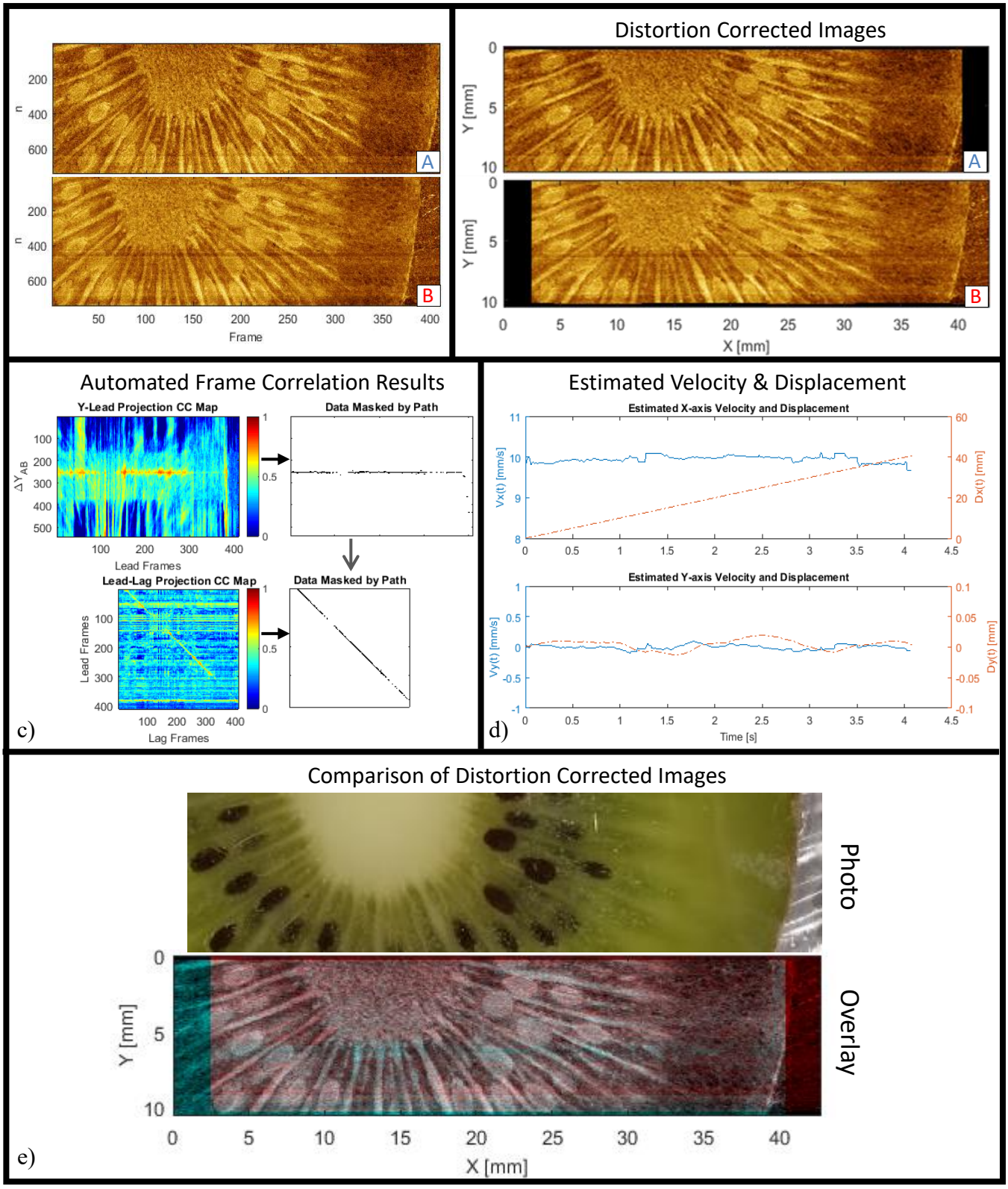

Fig. 28. Distortion correction results of a kiwi cross section sample imaged at v_x = 10 mm/s perpendicular to the scan lines such that $y_v \approx 0$ mm/s. a) Preprocessed images A and B. b) Resulting distortion corrected images A and B. c) The automated frame correlation results including the two correlation coefficient matrix maximum intensity projections and the resulting isolated pathways. d) The velocity and displacement estimates. e) An overlay of the distortion corrected images to qualitatively infer correction accuracy and compare to the original photo of the sample. Cyan indicates image A, red indicates image B, and in the overlapping region, grayscale indicates image similarities

5.2 Manual actuation tests

Manual actuation tests were performed to assess the distortion correction functionality. To achieve this, two experiments were made using a dragon fruit sample: first with 1D actuation and second with 2D actuation. As there is no 'ground truth' of the manual actuated velocity profile to compare to, the metric of correction quality is based on qualitatively assessing the distortion corrected images, their corresponding overlay, and comparing to photographs of the original sample.

5.2.1 1D manual actuation

In the first manual actuation test, the dragon fruit sample was imaged with one dimension of manual actuation. This was achieved by fixing the dragon fruit sample on the stage of a sliding rail which was aligned perpendicular to the scan lines, as the linear stepper motor was in the previous section. With the sample attached to a sliding rail, it can be manually pushed along the x-axis without allowing for motion in the y-axis. Similarly to the linear motor, by manually aligning the rail perpendicular to the scan lines, there may be a nonzero v_y component present which can influence the results.

The images and distortion correction results for the 1D actuation test are summarized in Fig. 29. Similar to the kiwi, dragon fruit provides high contrast and unique looking features, specifically dark round seeds, light flesh, and light fibrous structures. As seen in Fig. 29a), the dragon fruit seeds and fibrous structures appear bright while flesh appears dark, and holes in the flesh, due to displaced seeds, appear very dark in the OCT images. Some seeds deeper into the flesh, which are visible in the photo due to high contrast compared to the lightly colored flesh, are not visible in the *en face* OCT image because they occur at the bottom or below the depth range that is averaged to produce the *en face* image. Furthermore, it is easy to observe in these images that varying velocity profiles result in sample features appearing very different in each image, as can be seen by the stretched seeds between frames 100 and 150.

Based on the conclusions from the validations in the previous section, it is expected that the correlation matrix pathway will contain noticeably higher correlation values compared to the background, and that is observed in Fig. 29c), with the exception of highly correlating regions outside of the pathway, possibly due to regions of roughly homogenous flesh in the images. As expected, the Δy_{AB}-lead pathway is a roughly flat line since v_y should be constant and zero. The lead-lag pathway is a curvy shape related to the variable v_x produced from manually pushing the sample. The corresponding velocity profile estimates are shown in Fig. 29d) and resulting corrected images are shown in Fig. 29b).

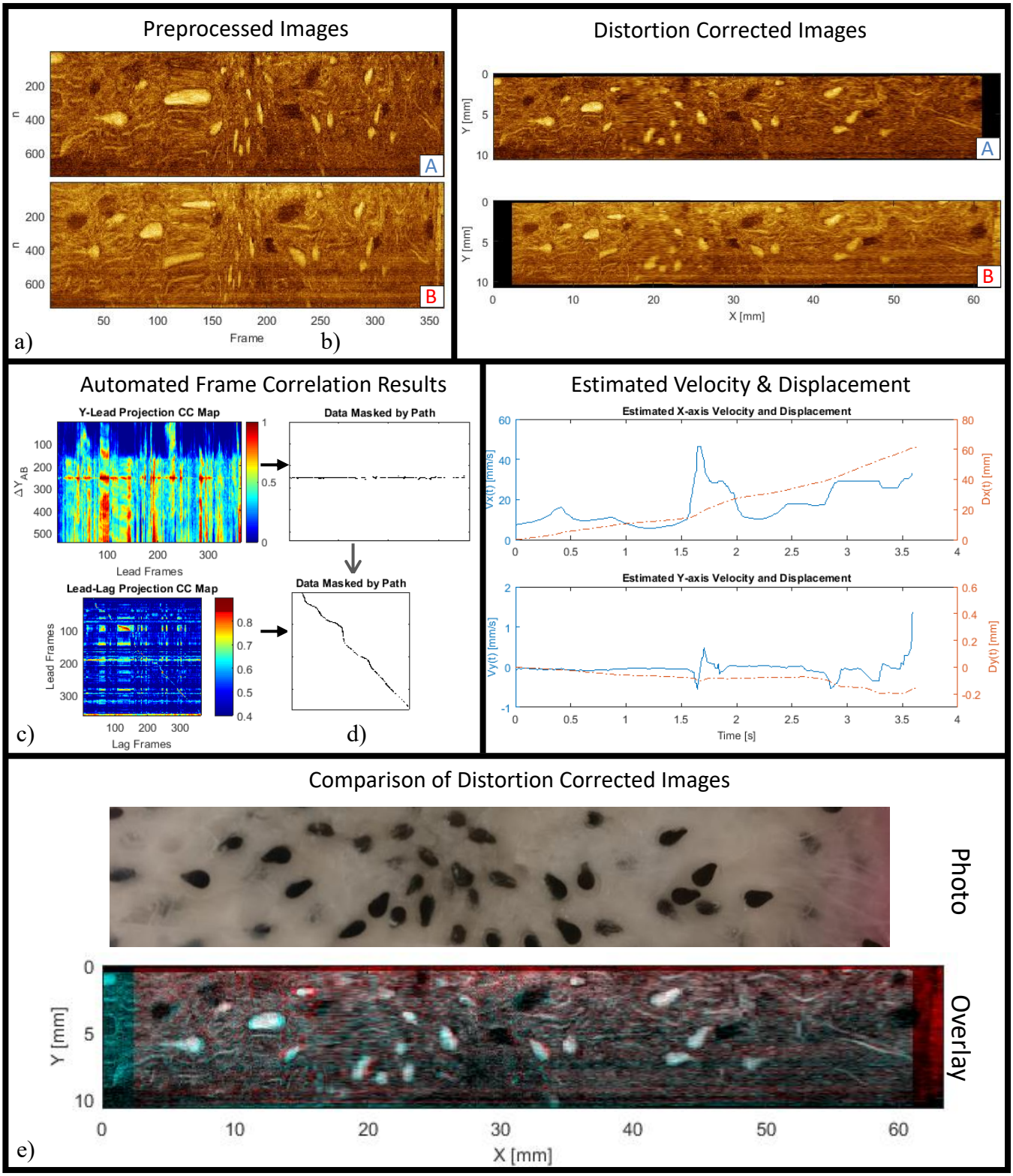

Fig. 29. Dragon fruit sample imaged on a sliding rail with manual actuation in the x-axis and no motion in the y-axis, i.e. $v_y \approx 0$, and the corresponding distortion correction results. a) Preprocessed images A and B. b) Resulting distortion corrected images A and B. c) The automated frame correlation results including the two correlation coefficient matrix maximum intensity projections and the resulting isolated pathways. d) The velocity and displacement estimates. e) An overlay of the distortion corrected images to qualitatively infer correction accuracy and compare to the original photo of the sample. Cyan indicates image A, red indicates image B, and in the overlapping region, grayscale indicates image similarities

When comparing the v_x profile to the extracted lead-lag pathway, one can observe a relationship. At higher v_x, the pathway approaches the diagonal while at lower v_x the pathway travels away from the diagonal. Intuitively this makes sense, as points located near the diagonal imply frames of similar value, i.e. small time interval and therefore higher velocity, whereas further away from the diagonal, the distance between frames increases, i.e. larger time interval and therefore lower velocity.

Qualitatively, the distortion correction procedure performs well in this 1D manual actuation test. The seeds which were visibly stretched between frames 100 and 150 in the distorted images become compressed in the corrected image and appear more natural looking. Similarly, seeds which were visibly compressed between frames 150 and 200 in the raw image have been stretched and have improved in appearance.

Although the appearance of the visibly stretched seeds improved with distortion correction, they still appear distorted in shape when compared to the original sample. This may suggest that this configuration or distortion correction procedure reduces in accuracy when extracting slower speeds. Additionally, features located near the beginning or end of the scan, e.g. the seed located at x≈8mm in the overlay image, appear slightly distorted after correction. This can be attributed to the reduced amount of overlapping velocity information at the edges of the scan and therefore reduced accuracy compared to the middle of the scan.

5.2.2 2D manual actuation

In the second manual actuation test, the dragon fruit was imaged with two dimensions of manual actuation. This was achieved by holding the sample in hand and pushing it underneath the scanner, while maintaining consistent pressure and contact, and allowing for motion in both the x- and y-axis. Care was taken to avoid rotating the sample as rotation adds another degree of freedom for correction which is not addressed with this technique. The distortion correction results of the images acquired through hand held scanning are summarized in Fig. 30.

As there is manual actuation in two dimensions, the resulting pathways, shown in Fig. 30c), and corresponding velocity estimates, shown in Fig. 30d), are variable. Around frame 150 of the distorted images corresponding to around 1.5 s in the velocity plots, spikes in both v_x and v_y can be observed, causing the group of three seeds captured at this point in the scan to appear compressed and sheared respectively. The algorithm performs well in correcting distortion caused by variable v_x and v_y, as can be seen by the improved appearance of the aforementioned group of seeds.

Similar to previous results, although distortion correction performs well, the overlay image shows regions where correction can be improved. Qualitatively, the 1D actuation results are better compared to the 2D actuation results. This could be attributed to a reduction of overlapping sample

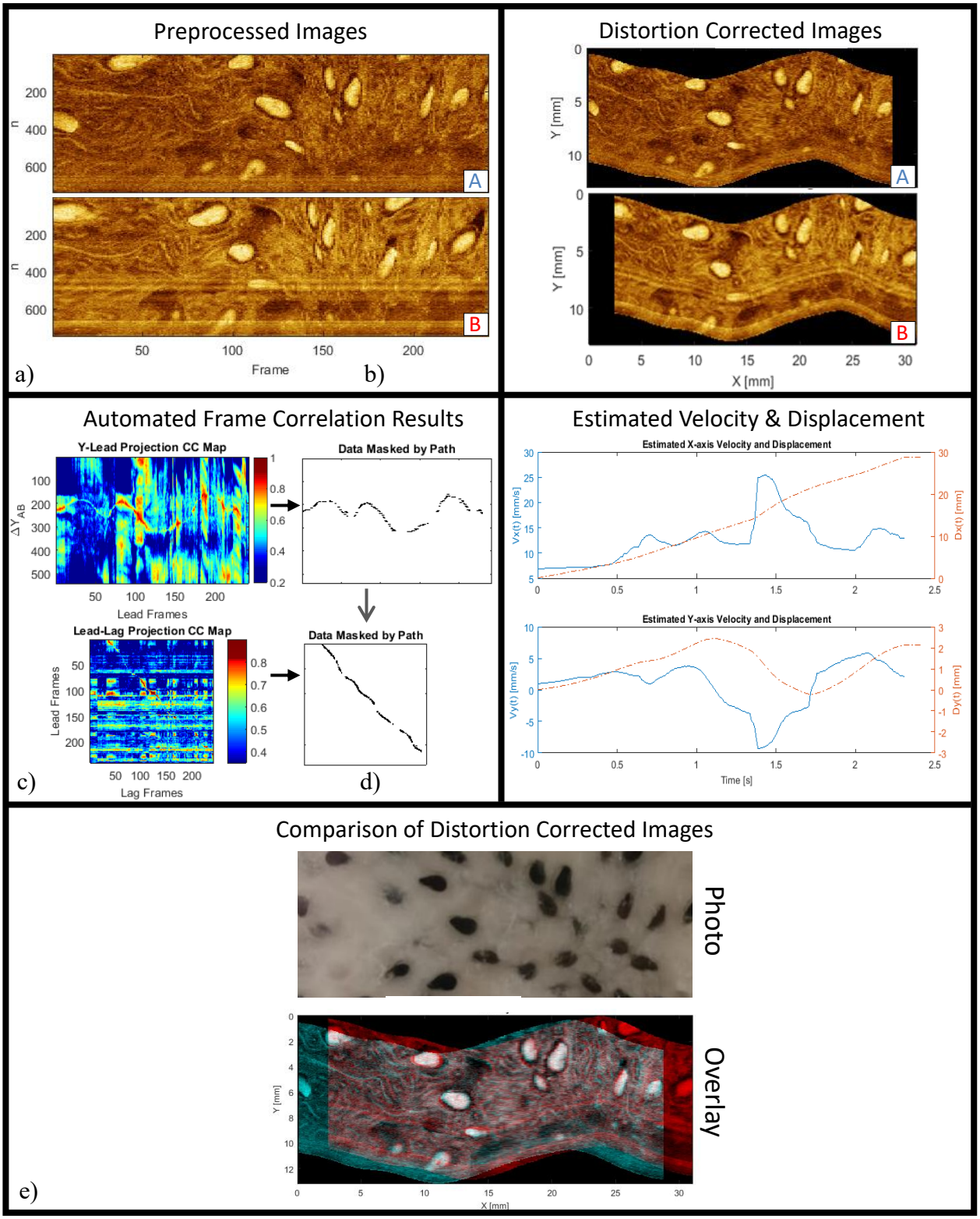

Fig. 30. Dragon fruit sample imaged by hand, with manual actuation in x- and y-axis, and the corresponding distortion correction results. a) Preprocessed images A and B. b) Resulting distortion corrected images A and B. c) The automated frame correlation results including the two correlation coefficient matrix maximum intensity projections and the resulting isolated pathways. d) The velocity and displacement estimates. e) An overlay of the distortion corrected images to qualitatively infer correction accuracy and compare to the original photo of the sample. Cyan indicates image A, red indicates image B, and in the overlapping region, grayscale indicates image similarities.

information in both images due to a non-zero v_y. As can be seen in the 2D actuation overlay image, there are regions of the sample which is only seen by one of the beams, indicating that there is a reduction of overlapping sample information to be used for feature-matching and thus a potential cause for reduced correction accuracy.

5.3 *In vivo* test

In vivo DMDI was evaluated by imaging a finger print, summarized in Fig. 31. Finger ridges appear dark in the OCT images and despite the relatively low resolution of this optical set up, some sweat ducts are visible, in the form of small faint dark dots in between the ridges. The automated frame correlation algorithm performs well at extracting the feature-matching pathway, even with the large patch of high correlation due to the homogeneous region of no sample located at the bottom right of the images. Similar to the beef tongue, the fingerprint correlation matrix exhibits similar and lower correlation values throughout, with the exception of the homogeneous region, with the pathway only slightly higher in value. This is attributed to the repetitive nature and low feature contrast of the fingerprint ridges. Despite this, the distortion correction algorithm performs well, as can be seen by the similarities of the corrected images, shown in Fig. 31b), emphasized in the overlay image in Fig. 31e). This test demonstrates that *in vivo* applications of this implementation of DMDI are feasible, such as imaging the skin or even adapted for the oral cavity.

5.4 Summary

In this chapter, various tests were conducted to assess the functionality of the distortion correction method. The distortion correction method performed well at extracting the feature-matching pathway, thus estimating the velocity profiles, and correcting the distorted images over a range of speeds and for different sample types throughout the tests.

It was observed that the appearance of the correlation matrix is affected by the sample type. For samples with high feature contrast and uniquely shaped features, in the case of kiwi and dragon fruit, the correlation matrix consists of high correlation values across the pathway compared to the rest of the matrix. On the contrary, samples with lower feature contrast and similarly appearing features, in the case of beef tongue and fingerprints, the correlation matrix consists of similarly valued correlation values with the pathway only slightly higher. An exception to these observations is the occurrence of highly correlating regions in the image such as homogeneous, i.e. featureless, regions.

It was observed that correction accuracy is influenced by the sample type and the velocity profile. Samples with higher feature contrast and uniquely shaped features performed better at estimating the velocity profile and therefore correcting the images, compared to samples with lower feature

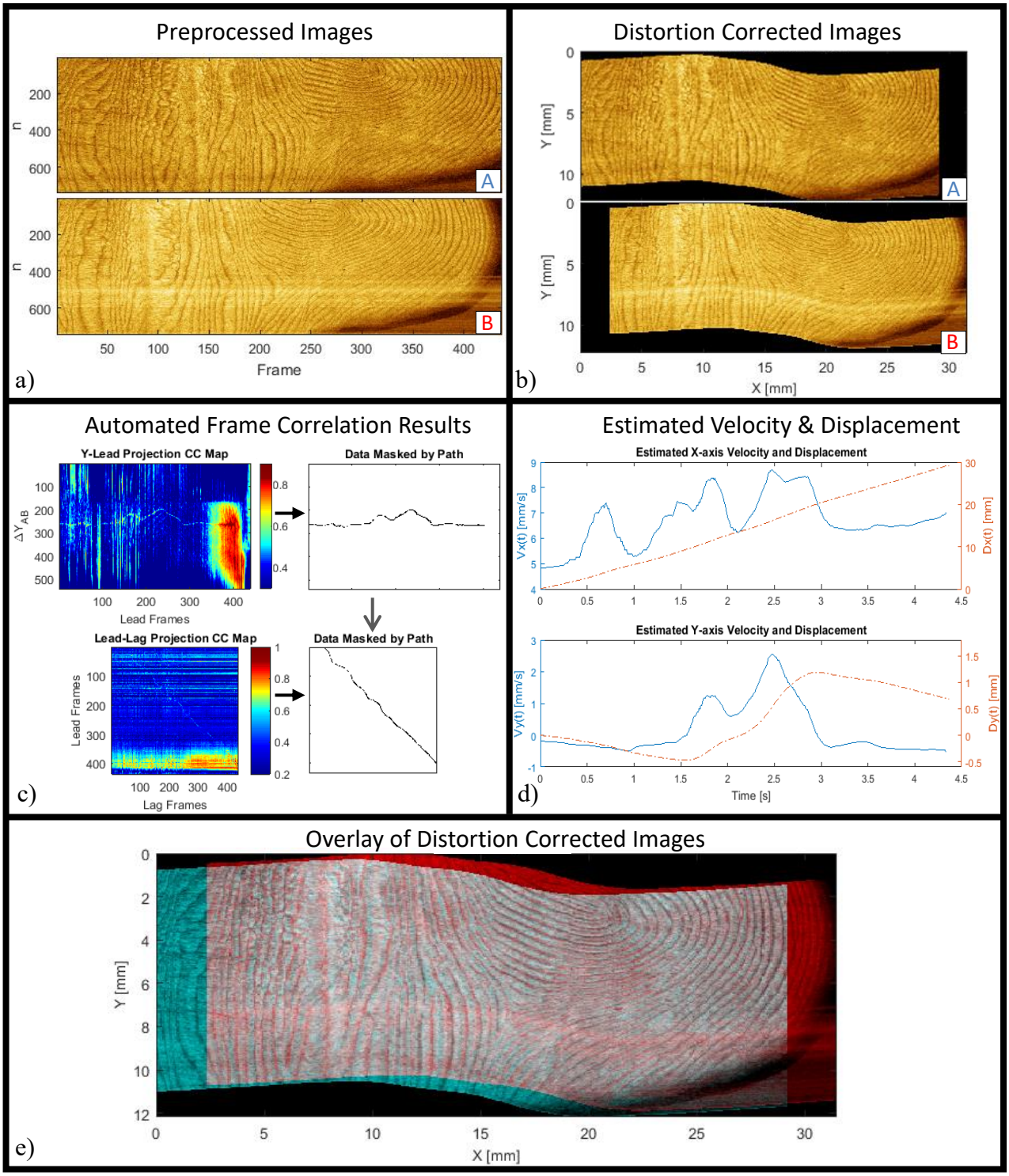

Fig. 31. A scan of a finger print with manual actuation in x- and y-axis, and the corresponding distortion correction results. a) Preprocessed images A and B. b) Resulting distortion corrected images A and B. c) The automated frame correlation results including the two correlation coefficient matrix maximum intensity projections and the resulting isolated pathways. d) The velocity and displacement estimates. e) An overlay of the distortion corrected images to qualitatively infer correction accuracy. Cyan indicates image A, red indicates image B, and in the overlapping region, grayscale indicates image similarities.

contrast and similarly appearing features. The correction accuracy seemed to decrease at lower speeds and when non-zero v_y components are present. Lastly, accuracy appears to be higher in the middle of a scan compared to the beginning and end of a scan, implying that more overlapping feature matching information improves the velocity estimate and therefore improves correction results.

Additionally, as the velocity profiles are estimated by calculating average velocities across the time interval of both beams capturing the same sample feature, the velocity estimation accuracy is inherently related to the beam separation of the optical set up. In all these tests, the beam separation is fixed and therefore it is not possible to evaluate its influence on the correction accuracy. Thus to better understand the influence of different velocity ranges and different beam separations on the accuracy of the correction results, further tests must be done. As changing the beam separation is not possible for this optical setup, this is achieved through simulation in Chapter 6.

6 Simulations

As the beam scanning pattern can be characterized through calibration, the imaging system and its corresponding scanning distortion can be parameterized and simulated. Simulation of the imaging system allows for investigation of changing scanning parameters without the need to adjust or change the physical imaging set up. Furthermore, simulation of the distortion caused by different velocity profiles is possible. Through simulation of the imaging system with the addition of arbitrary velocity profiles, one can investigate the influence of a velocity profile on the correction, as well as quantitatively assess the performance and accuracy of the correction results utilizing an error metric. This chapter describes how simulated distorted images are constructed, the error metric used to assess and compare correction results, and the various results found through simulation tests.

6.1 Simulation of distorted images

Distortion correction of DMDI images is achieved by estimating the effective beam path of two beams onto the sample in order to map and interpolate corrected image pixels from the distorted images. On the contrary, correct images can become distorted by constructing an effective beam path, in other words a scanning pattern with an arbitrary velocity profile, which can be traced onto a correct image to map and interpolate distorted images. These simulated distorted images can in turn be corrected using the automated distortion correction method, and since the exact applied beam path is known, the correction accuracy can be quantitatively assessed.

Simulated distorted images can be constructed from a digital image, for example numerous different and repeating QR codes, or from already corrected images such as the kiwi image from Section 5.1.2. The effective beam paths are determined by applying an arbitrary velocity profile to a simulated scanning pattern. The scanning pattern is constructed such that it matches the scanning pattern determined in Chapter 3 with the exception of parameterizing the beam separation so that it can be changed in order to investigate its effects on the correction accuracy. Also included to construct the scanning pattern, is the laser frequency to dictate the temporal resolution, the total scan time desired, and the starting position of the beams on the sample. Next, an arbitrary velocity profile is applied to the scanning pattern to dictate where the beams scan due to displacement of the sample. The resulting effective beam path is traced onto the correct initial image in order to map or interpolate the distorted image pixels. As an effective beam path with two beams is simulated, two distorted images are constructed. Lastly, to simulate real world imaging conditions, the distorted images are Gaussian blurred with $\sigma = 0.5$ and speckle noise is added to image A with 0.01 variance and image B with 0.02 variance. Speckle noise of different variances are added to the images to prevent the two images from highly correlating with each other due to having the same speckle noise distribution. Further, image B was selected to have higher variance because in the real set up, beam B is noisier than beam A. An example of constructing simulated distorted images is shown in Fig. 32. In this

example effective beam paths A (red) and B (blue) are constructed with a scan time of 5 seconds, v_y = 1 mm/s, and variable v_x with a peak velocity of 10 mm/s.

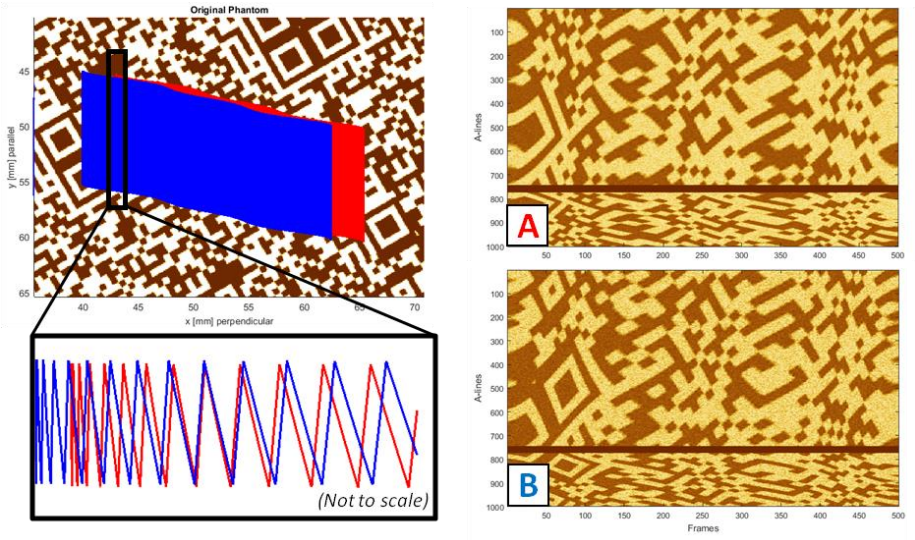

Fig. 32. Example of constructing simulated distortion images. The sample image [left] (in mm) is a tiled QR code digital phantom used as an example. The applied effective beam paths, in red and blue for beams A and B respectively, are constructed with v_y = 1 mm/s and a variable v_x with peak velocities of 10 mm/s. The effective beam paths map the correct sample pixels to the corresponding distorted images [right] (in frames and A-lines). The zoom in box illustrates what the effective beam paths might look like.

6.2 Error metric

To assess and compare the accuracy of correction results, root-mean-squared error (RMSE) is calculated. RMSE is an error metric used here to compare how different the estimated velocity profile is from the expected simulated velocity profile. RMSE is given by

$$RMSE = \sqrt{\frac{1}{N_v}\sum_{i=1}^{N_v}\left(v_{est,i} - v_{sim,i}\right)^2}, \qquad N_v, i \in \mathbb{Z} \tag{20}$$

where N_v is the number of velocity points to be compared, v_{est} is the estimated velocity from the distortion correction method, and v_{sim} is the simulated velocity profile. An RMSE value of 0 indicates that the estimated velocity profile exactly matches the simulated velocity profile, which is unlikely to occur in real applications. In general, lower RMSE values indicate better velocity estimation and therefore better distortion correction.

In real scanning scenarios, the ground truth velocity profile is not available to calculate RMSE. Without a ground truth velocity to compare to, it is not straight forward to quantitatively assess the correction accuracy. We hypothesize that the correlation coefficients of the extracted pathway can be used to provide a measure of confidence on the correction accuracy. For example, an extracted pathway consisting of numerous high correlation points would be expected to produce better results compared to a pathway with few, lower correlation points. Thus in these simulations, a confidence metric is

calculated along with the RMSE to investigate how well said confidence can infer correction accuracy. Confidence is given by

$$Confidence = \frac{\sum_{j=1}^{N_{pairs}} CC_j}{N_{frames}}, \qquad j \in \mathbb{Z} \tag{21}$$

where N_{pairs} is the number of feature-matching pairs in the extracted pathway, CC_i is the normalized cross correlation coefficient of feature-matching pair j, and N_{fr} is the total number of frames in the image. The resulting confidence is a value ranging from 0 to 1 where values approaching 1 indicates a pathway with numerous points corresponding to high correlation values.

6.3 Simulation results

Four simulations were run to test various scanning parameters. In the first simulation, the ideal range of v_x is investigated for three scan line separations. In the second simulation, the effects of increasing constant v_x, constant v_y, and beam separation is investigated. In the third simulation the effects of beam separation on estimating variable v_x is investigated. Lastly, in the fourth simulation, two sample types (QR code and kiwi) are compared. This section discusses the simulation results and observations.

6.3.1 Simulation 1: effects of beam separation on $v_{x,max}$

In the first simulation test, a range of constant v_x was applied to investigate the effects of beam separation and v_x on correction accuracy. This was achieved by simulating distorted images from a digitally constructed phantom of numerous and differing QR codes with v_x ranging from 0.1-232 mm/s, v_y fixed to 0 mm/s, and beam separations of 0.7, 2.32, and 3 mm. All scans were originated from the same location on the phantom and the scan time was fixed to 5 seconds, equivalent to 500 frames. With time fixed, at higher v_x the effective beam path scans physically more of the sample than at lower v_x. Thus the digital QR code phantom was constructed long enough to provide information for all effective beam paths applied in this test.

For each combination of beam separation and v_x, two distorted images are constructed and in turn corrected using the automated distortion correction method. RMSE is calculated from the estimated v_x of the correction and the simulated v_x used to construct the distorted images. RMSE is plotted, shown in Fig. 33, along the y-axis at the corresponding simulated v_x along the x-axis and the corresponding beam separation is indicated by curve colour.

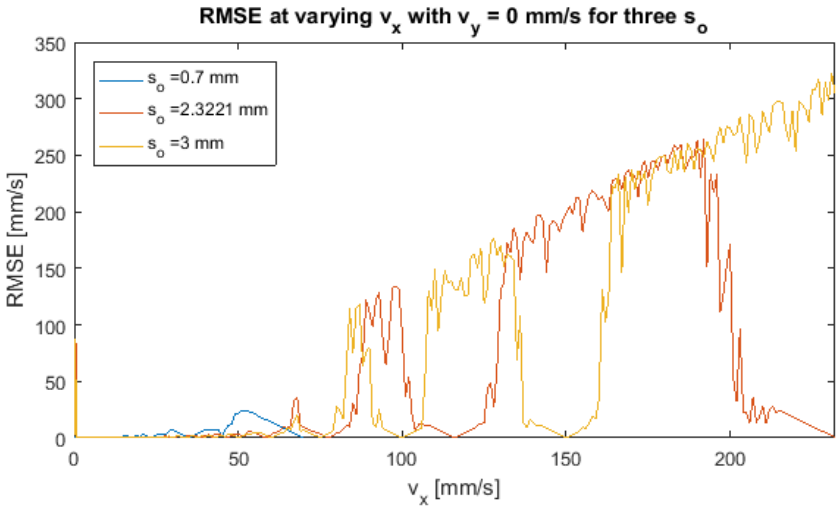

Fig. 33. RMSE is plotted for three beam separations for a range of constant v_x , up until the $v_{x,max}$ corresponding to that beam separation. These results indicate there is an effective ideal range of v_x depending on beam separation to achieve high correction accuracy. Additionally, ideal accuracy is achieved at integer divisions of $v_{x,max}$ corresponding to that beam separation.

Interesting observations can be made from Fig. 33. Firstly, after a certain point, around $v_x \approx 75$ mm/s, RMSE is roughly directly proportional to the input v_x with the exception of some valleys where RMSE becomes very low. It is important to note that $s_o = 0.7$ is only simulated up to its $v_{x,max}$ of 70 mm/s. In this linear region, RMSE is roughly equal to the input v_x indicating that the estimated v_x was essentially zero. In fact the estimated velocities in these regions are essentially zero as there was no discernible pathway in the correlation matrix to estimate a velocity. This indicates that there is an effective $v_{x,max}$ smaller than the theoretical $v_{x,max}$. For example, with a beam separation of 3 mm, $v_{x,max}$ is 300 mm/s from Eq. 17 in Chapter 4. However these results suggest the effective $v_{x,max}$ is approximately 75 mm/s, as after this point there is no feature matching information. The exception, as already mentioned, is the few valleys located at integer division of the theoretical $v_{x,max}$. This indicates that the imaging system and distortion correction technique, is sensitive to velocities which are integer divisions of the theoretical $v_{x,max}$. In other words, feature-matching information is possible when both beams capture the same sample features, and at fast v_x this only occurs when the same sample feature reaches both beams at the same point of a frame cycle. When $v_x = v_{x,max}$, each beam sees the same sample feature with only one frame of separation. The next scenario when both beams capture the exact same sample feature is with two frames of separation, in other words half the $v_{x,max}$, and so on. In between integer divisions of $v_{x,max}$, the two beams do not capture the same sample information and distortion correction is not possible. As the integer division increases, this gap decreases and eventually begins to overlap at lower v_x when both beams are essentially guaranteed to capture the same sample features. This appears to define the effective v_x sensitivity range, which is dependent on the beam separation.

Recall from Chapter 4 that smaller beam separations are expected to perform better at small v_x. To observe this, the plot in Fig. 33 is zoomed in and shown in Fig. 34. From this plot one can observe that

smaller separations register low RMSE, or high accuracy, at lower v_x than larger beam separations. This indicates that smaller beam separations are more sensitive to very small v_x, however at the cost of reduced sensitivity at larger v_x.

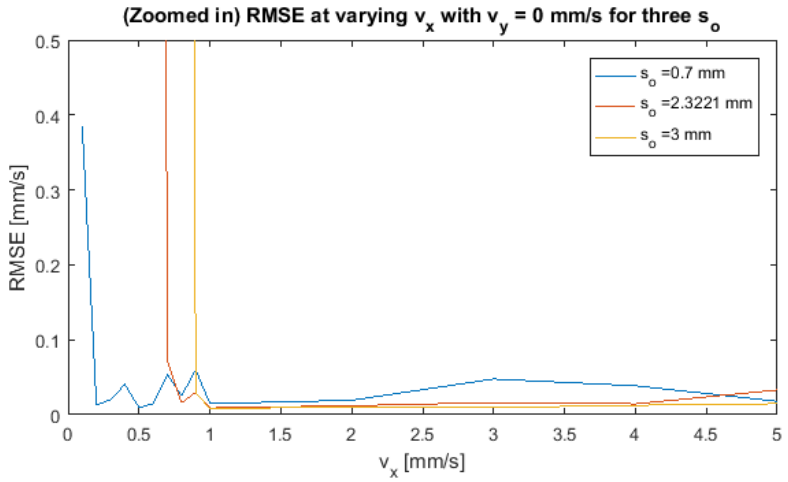

Fig. 34. Zoomed in plot of Fig. 33 to observe the effects of beam separation on correction accuracy. Smaller beam separations are sensitive to smaller v_x but less sensitive to higher v_x compared to larger beam separations.

Confidence of this simulation test is plotted in Fig. 35 allowing for comparison with RMSE. It can be observed that confidence follows similar trends as RMSE. As RMSE increases with increasing v_x, confidence decreases. Similarly, a linear region of poor confidence is reached after 75 mm/s and peaks of high confidence occur at integer divisions of v_x. This indicates that confidence could be used as a measure of correction performance without requiring a ground truth for this QR example. However the confidence measure should be further tested with other samples, specifically biological samples, to validate its functionality as a performance metric.

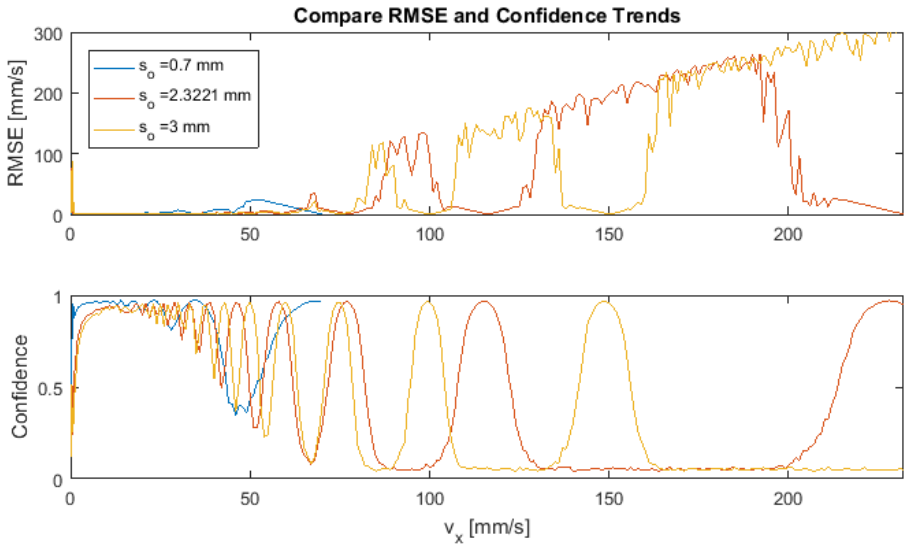

Fig. 35. Confidence [bottom] inversely follows the same trends as RMSE [top], suggesting confidence may be used as a measure of performance when ground truth is not available.

6.3.2 Simulation 2: effects of v_y on $v_{x,min}$

In the second simulation test, multiple combinations of beam separation, constant v_x, and constant v_y were simulated to investigate the effects of introducing nonzero v_y as well as further investigate the effects of beam separation and v_x on correction accuracy. As in the first simulation test, two distorted images were constructed from the same QR code phantom with the same fixed scan time, laser frequency, and physical starting position. However in this test, four beam separations (s_o = 0.5, 1.5, 3, and 4 mm), two y velocities (v_y = 0 and -5 mm/s), and v_x ranging from 0.01-100 mm/s were simulated. For each combination, three metrics were calculated and plotted in Fig. 36, specifically RMSE of the estimated v_x (top row), RMSE of the estimated v_y (middle row), and confidence (bottom row). The plots in Fig. 36 illustrate the effects of v_x, v_y, and beam separation on the accuracy of estimating the v_x and v_y profiles.

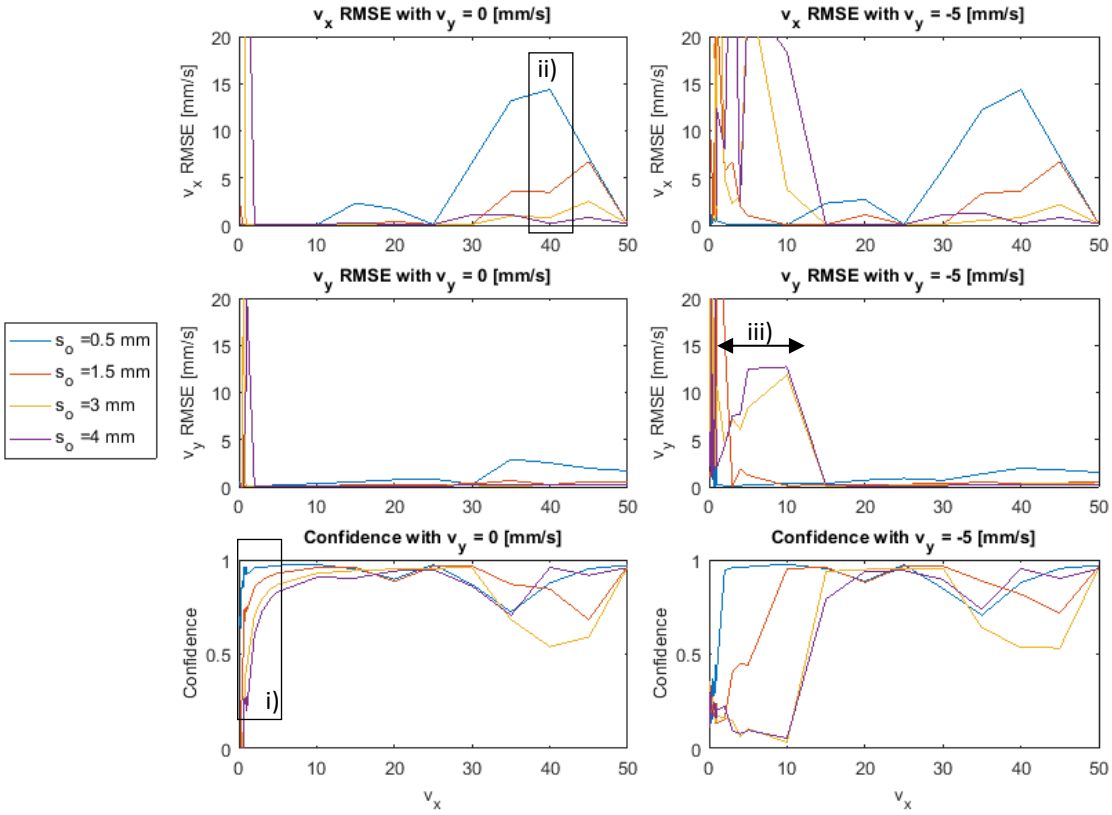

Fig. 36. Compare correction results for two cases: v_y = 0mm/s [left] and v_y = -5 mm/s [right]. RMSE is calculated and plotted for both v_x estimates [top] and v_y estimates [middle] for each simulation as well as the confidence [bottom]. i) Small beam separations are more accurate at very small v_x while ii) large beam separations are more accurate at high v_x. iii) Increasing v_y appears to increase the minimum detectable v_x.

Introducing a non-zero v_y appears to reduce the accuracy at smaller v_x, especially for larger beam separations, indicated by arrows in Fig. 36iii). This suggests that non-zero v_y increases the minimum detectable v_x dependent on the beam separation. An influence of v_y on correction accuracy can be attributed to a reduction of frame overlap for in feature matching. With increasing v_y, the sample may displace sufficiently enough between the times of each beam passing the same part of the sample

such that the frames from each beam contain less overlapping information. With less overlapping information, the feature-matching and velocity estimate accuracy may degrade. These results also nicely illustrate the effect of beam separation on accuracy at ranges of v_x. Smaller beam separations perform better at low v_x, highlighted in box i), while large beam separation perform better at high v_x, highlighted in box ii).

6.3.3 Simulation 3: effects of beam separation estimating a varying v_x

In the third simulation test, one variable v_x profile was applied with varying beam separation to investigate how beam separation influences the accuracy in a variable velocity profile. As in the first two simulation tests, the same QR code and scanning parameters with the exception of using five beam separations (s_o = 1, 1.5, 2.3, 3, 5 mm), v_y = 0, and v_x equal to the variable velocity shown in black in Fig. 37. The estimated velocity profiles corresponding to each beam separation are plotted to show how well each performs at estimating the variable v_x profile.

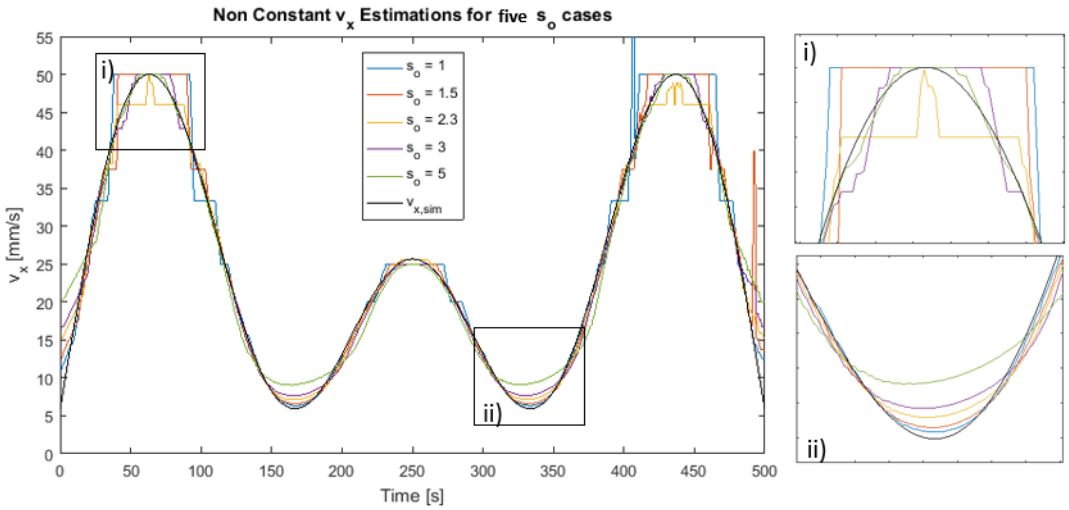

Fig. 37. Estimated variable v_x profile with different beam separations. The simulated v_x profile is a modified sinusoidal function plotted in black and the estimated velocity profile corresponding to different beam separations are plotted in colour. i) Large beam separations appear to perform better at estimating high v_x peaks. ii) Smaller beam separations appear to perform better at estimating low v_x peaks.

In general, due to the averaging approach used to estimate the velocity profiles, all cases show error at the peaks of the varying v_x profile. As can be seen in Fig. 37i), larger beam separations appear to extract the high v_x peak better than smaller beam separations. On the contrary, as can be seen in Fig. 37ii), smaller beam separations appear to extract the low v_x peak better than large beam separations.

6.3.4 Simulation 4: confidence functionality for a biological sample

In the fourth simulation test, two different samples of the same physical size are simulated with varying constant v_x and v_y = 0 to investigate the functionality of confidence for a biological sample. In this simulation the same digitally constructed QR code is used as well as the corrected images A and B of the kiwi from section 5.1.2. Two sets of distorted images are constructed with only one beam separation of 1.5 mm and one v_y of 0 mm/s, with v_x varying from 0.1 - 74 mm/s, the maximum

allowable velocity. The maximum allowable velocity is limited by the length of the kiwi data available and the decision to have at least 50 frames in the distorted images. With the sample size fixed to the length of the kiwi data available, the total scan time and therefore number of frames is chosen such that the whole sample is imaged at a given v_x. Thus at higher v_x, the distorted images have less frames. Again, the RMSE and confidence of the estimated v_x were calculated for both sample types. The results are plotted in Fig. 38 with RMSE on the left y-axis shown in solid lines, confidence on the right y-axis shown in dashed lines, QR sample shown in blue, and kiwi shown in red.

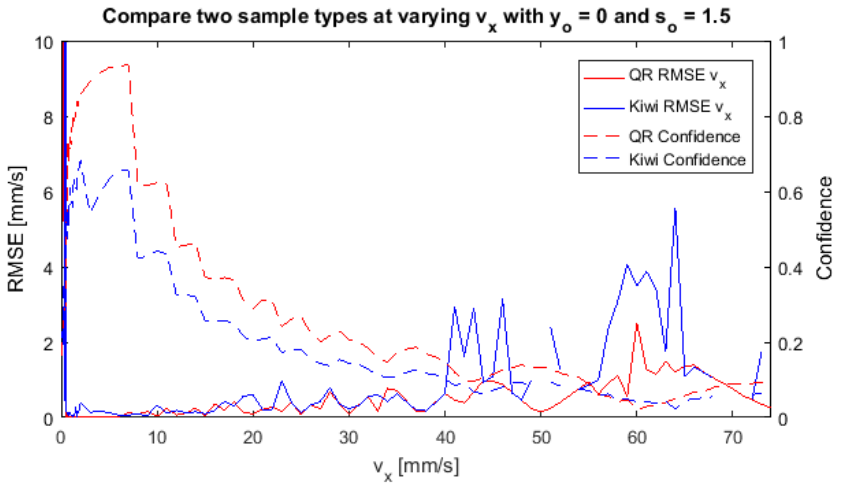

Fig. 38. Comparing correction metrics for QR and biological sample. Confidence inversely follows the same trends as RMSE. Further, confidence appears to follow the same trends for both the digitally constructed QR sample and the real images of the biological sample, i.e. kiwi, however at a lower offset.

From the plot in Fig. 38, it can be seen that confidence from the kiwi sample follows closely to that of the QR code but with an offset. This is expected as a real image of a biological sample will have lower correlation values compared to a high contrast black and white phantom. This implies that confidence could be used as a performance metric however it may be difficult, and also application or sample dependent, to determine where to place the threshold of what is good or bad confidence.

The downward trend of confidence in this test appears to follow the increasing RMSE trend but more drastically. This can be attributed to the reduction of frames therefore reduction of information in the correlation matrix.

6.4 Summary

Simulation allowed for fast investigations of scanning parameters without the need to acquire and process numerous images or change the optical set up in any way. In general it was found that at very small v_x, smaller beam separations perform better than larger beam separations at accurately estimating v_x. However, larger beam separations perform better throughout a larger range of v_x, especially at higher v_x where the accuracy of small beam separations quickly degrades. Additionally it was observed that velocity estimate accuracy is most ideal at integer divisions of $v_{x,max}$ which is dependent on the beam separation.

Introducing non-zero v_y appears to decrease the range of detectable v_x, specifically it appears to increase the minimum detectable v_x, especially for larger beam separations. In addition, smaller beam separations perform well at estimating a varying v_x profile at lower v_x magnitudes whereas large beam separations estimate a varying v_x profile better at higher v_x magnitudes.

In these simulations, confidence was observed to inversely follow the same trends as RMSE. This suggests that confidence may be a potential performance metric used to quantitatively assess correction results when a ground truth is not available. Further work is needed to determine at what confidence threshold is a correction considered good or bad as it is possible this threshold may be dependent on the sample type.

Further work can be done to improve the simulations. For example, the simulations can become more realistic and comparable to the real imaging system by improving how noise is added to the images. The noise construction can be improved to better simulate the actual noise and reflection artefacts caused by the real imaging system. This will allow for more representative data and more accurate testing conditions.

Additionally, with the imaging system parameterized, many more scanning parameters can be investigated. For example, alternative galvo scanner waveforms could be simulated to investigate what waveform is most ideal for a given application. Further, instead of two parallel lines, other scanning geometries could be simulated to investigate their corresponding scanning pattern distortion and evaluate their usability as a DMDI implementation. Another example could be to investigate if there is an ideal sample feature type or density to help dictate future applications.

These and future simulations can be used to evaluate the performance of an imaging system and the distortion correction method. The results can be used to determine the ideal scanning parameters and conditions of an imaging set up for a particular application, thus allowing for application specific system optimization. System optimization can be advantageous for allowing numerous and various future applications of DMDI.

7 Summary and Future Outlook

In this thesis, a new implementation of DMDI was investigated with automated distortion correction. DMDI uses only the information from two distorted images and calibrated scanning pattern of two beams to produce two corrected images of the sample. By utilizing a single-axis galvanometer, the resulting scanning pattern of two roughly parallel lines simplifies the correction algorithm and provides an opportunity for automated feature matching, a key pitfall to the first implementation of DMDI using a DBMC. By implementing automated distortion correction, the image analysis time reduces from on the order of hours to <2 mins. In theory, the technique can be further optimized and developed to operate at real-time with an inherent delay attributed to the scan line separation and the time required to have both beams capture the same sample feature. Practically, a distortion correction algorithm for DMDI in real-time would take the form of live updating a distortion corrected image as feature-matching data becomes available during the scan. Currently a manuscript for this work is in preparation to be submitted to Biomedical Optics Express Journal.

The optical set up used in this work allowed for planar, *en face* OCT imaging of various samples with a resolution of 80x80 µm FWHM. Compared to typical OCT resolutions on the order of 10 µm, there is opportunity to change the configuration to improve the resolution. For example, increasing NA at the scan lens will result in improved focus and resolution at the sample. A configuration with two separate fibers with collimators can achieve this but with the added challenge of ensuring the two beams are properly aligned to the mirror.

In this work, distortion correction using DMDI was demonstrated with 2D *en face* images constructed from 3D-OCT data sets. However it should be possible to generate distortion corrected 3D volumes by applying the same correction on the corresponding 3D-OCT data set. This is only possible when both beams penetrate the sample at the exact same angle of incidence and therefore capture the same sample information when scanning at the same location. In theory of this implementation of DMDI, both beams should scan normal to the sample plane due to the scanning lens and therefore penetrate the sample at the same angle of incidence. This can be confirmed by profiling the two beams at various depths from the focal point to see if the beams are in fact normal to the sample plane. If the beams are not normal to the sample plane, then the depth dimension must be included in the calibration to account for the depth scanning pattern and allow for 3D volumetric distortion correction.

The automated distortion correction method developed in this work performs well at matching sample features, estimating the velocity, and generating distortion corrected images. As demonstrated in the validation tests, the method extracts known constant velocities with a high level of accuracy. Through manual actuation tests, the method was qualitatively assessed and shown to estimate and correct for variable motion in two dimensions of actuation reasonably well. Future work can be made to further refine and improve the correction results. One example is to utilize the retrace

period of the image which will reduce the loss of information and therefore increase the amount of information to be used for an improved velocity estimate. To aid in utilizing the retrace period, the galvo scanner waveform can be adapted, for example from a modified saw tooth to a triangular waveform, to utilize the retrace period more efficiently.

A potential method to further improve correction results would be to use machine learning and iteratively converge to an improved velocity estimate using the results of the method developed in this work as an initial seed estimate. As DMDI utilizes only the information from two distorted images to produce corrected images, the ending condition of this iterative approach would need to be decided from the corrected images. For example, a perfect correction could be identified when correct images A and B are structurally identical. However comparing the two images and setting an acceptable threshold is nontrivial and requires considerable work to implement robustly.

In this work, sample motion was restricted to monotonic v_x and without rotation of the sample in the x-y plane. A non-monotonic v_x would appear in the correlation matrix as a pathway which occurs on either side of the A-B diagonal. It should be feasible to refine and adapt the pathway extraction of the distortion correction algorithm to detect changes of direction in the x-axis. Additionally, rotation of the sample would appear as sample features in the both images occurring at different portions of multiple frames. Rotation of the sample could possibly be detected by comparing various segments or groups of frames between both images as opposed to comparing single frames directly. Further work can be done to detect and correct changing directions in v_x and rotation of the sample.

DMDI requires two spatially separated beams to capture sample features and allow for manual actuation distortion correction. However the two beams do not necessarily need to be the same imaging modality, they just need common identifiable sample features. For example, one beam could allow for regular OCT and the second could allow for autofluoresence-OCT (AFI-OCT) [2]. The imaging quality of AFI-OCT is inferior to OCT, however it should be sufficient enough to provide feature matching information to allow for distortion correction in addition to providing additional sample information not possible with OCT imaging alone. The automated distortion correction method can potentially be applied to multi-modal DMDI by choosing a similarity measure that is best suited to compare the modalities. Thus automated multi-modal DMDI is a plausible future imaging tool.

Although a bench top galvanometer scanner setup was used in this study, this implementation could be adapted for imaging body sites such as the oral cavity or skin. Furthermore, miniaturization of the imaging head or implementing an alternative optical setup to achieve linear scanning geometry are potential solutions to reach different imaging sites with automated distortion correction capabilities.

Simulation of the imaging system and actuation velocities allowed for fast investigation on the effects of scanning parameters and velocity profiles on correction accuracy without the need to image and

process numerous tests or make changes to, or construct a new, optical set up. Simulation results showed that correction accuracy is optimal at integer divisions of the maximum detectable x velocity, which is dependent on the beam separation. Additionally, smaller beam separations appear to perform better at extracting very small v_x compared to larger beam separations whereas larger beam separations perform better across a larger range of v_x compared to smaller beam separations. Lastly, the confidence, a potential performance metric which does not require a ground truth, appeared to inversely follow the same trends as RMSE, a direct comparison to ground truth. Further work will need to be done to set a threshold in confidence that can describe what is considered a good or bad correction, however the metric shows promise for providing quantitative assessment of manually actuated images.

With the imaging system parameterized, simulation allows for investigations of many additional scanning parameters that were not addressed in this thesis. Plenty of future work can be done in investigating these parameters and determine the ideal combination of scanning parameters and actuation speeds for a given application. Additionally, new imaging system configurations can be explored to assess their applicability for DMDI imaging. For example, as beam separation influences the range of velocities the imaging system is sensitive to, perhaps a 'V' shaped scanning pattern, allowing for a range of beam separations, would have a higher range of sensitivity. Simulation results can help guide development to optimize and refine the imaging set up and correction methods to further improve results and robustness.

Despite the limitations mentioned, the technique developed in this thesis produces favorable results and puts DMDI in a more attractive position as a manual scanning solution for clinical applications. With this work as a foundation, the future potential for this imaging modality is very promising.

8 References

[1] A. M. D. Lee, L. Cahill, K. Liu, C. MacAulay, C. Poh, and P. Lane, "Wide-field in vivo oral OCT imaging," *Biomedical Optics Express,* Article vol. 6, no. 7, pp. 2664-2674, Jul 2015.

[2] H. Pahlevaninezhad *et al.*, "A high-efficiency fiber-based imaging system for co-registered autofluorescence and optical coherence tomography," *Biomedical Optics Express,* Article vol. 5, no. 9, pp. 2978-2987, Sep 2014.

[3] W. G. Jung *et al.*, "Three-dimensional optical coherence tomography employing a 2-axis microelectromechanical scanning mirror," *IEEE Journal of Selected Topics in Quantum Electronics,* Article vol. 11, no. 4, pp. 806-810, Jul-Aug 2005.

[4] J. J. Sun *et al.*, "3D In Vivo optical coherence tomography based on a low-voltage, large-scan-range 2D MEMS mirror," *Optics Express,* Article vol. 18, no. 12, pp. 12065-12075, Jun 2010.

[5] G. Muller, S. Meissner, J. Walther, E. Koch, and H. Morawietz, "In vivo imaging of murine vasodynamics analyzing different mouse strains by optical coherence tomography," *Atherosclerosis Supplements,* Article vol. 30, pp. 311-318, Nov 2017.

[6] D. Schaefer, J. Gottmann, M. Hermans, J. Ortmann, and I. Kelbassa, "High speed micro scanner for 3D in-volume laser micro processing," in *Conference on Laser-Based Micro- and Nanopackaging and Assembly VII (LBMP) as a part of LASE Photonics West*, San Francisco, CA, 2013, vol. 8608, BELLINGHAM: Spie-Int Soc Optical Engineering, 2013.

[7] J. W. Li *et al.*, "A portable confocal hyperspectral microscope without any scan or tube lens and its application in fluorescence and Raman spectral imaging," *Optics Communications,* Article vol. 392, pp. 1-6, Jun 2017.

[8] S. Optronics. (19.02). *Scan Heads (Marking Heads)*. Available: http://www.sintecoptronics.com.sg/index.php/product/index/id/171.html

[9] G. Huszka and M. A. M. Gijs, "Turning a normal microscope into a super-resolution instrument using a scanning microlens array," *Scientific Reports,* Article vol. 8, p. 8, Jan 2018, Art. no. 601.

[10] F. B. Wang, P. Tu, C. Wu, L. Chen, and D. Feng, "Multi-image mosaic with SIFT and vision measurement for microscale structures processed by femtosecond laser," *Optics and Lasers in Engineering,* Article vol. 100, pp. 124-130, Jan 2018.

[11] K. C. Liang *et al.*, "Cycloid scanning for wide field optical coherence tomography endomicroscopy and angiography in vivo," *Optica,* Article vol. 5, no. 1, pp. 36-43, Jan 2018.

[12] M. J. Gora *et al.*, "Tethered capsule endomicroscopy enables less invasive imaging of gastrointestinal tract microstructure," *Nature Medicine,* Article vol. 19, no. 2, pp. 238-240, Feb 2013.

[13] K. C. Liang *et al.*, "Ultrahigh speed en face OCT capsule for endoscopic imaging," *Biomedical Optics Express,* Article vol. 6, no. 4, pp. 1146-1163, Apr 2015.

[14] B. Lau, R. A. McLaughlin, A. Curatolo, R. W. Kirk, D. K. Gerstmann, and D. D. Sampson, "Imaging true 3D endoscopic anatomy by incorporating magnetic tracking with optical coherence tomography: proof-of-principle for airways," *Optics Express,* Article vol. 18, no. 26, pp. 27173-27180, Dec 2010.

[15] B. Y. Yeo, R. A. McLaughlin, R. W. Kirk, and D. D. Sampson, "Enabling freehand lateral scanning of optical coherence tomography needle probes with a magnetic tracking system," *Biomedical Optics Express,* Article vol. 3, no. 7, pp. 1565-1578, Jul 2012.

[16] J. Ren, J. G. Wu, E. J. McDowell, and C. H. Yang, "Manual-scanning optical coherence tomography probe based on position tracking," *Optics Letters,* Article vol. 34, no. 21, pp. 3400-3402, Nov 2009.

[17] N. Iftimia, G. Maguluri, E. W. Chang, S. Chang, J. Magill, and W. Brugge, "Hand scanning optical coherence tomography imaging using encoder feedback," *Optics Letters,* Article vol. 39, no. 24, pp. 6807-6810, Dec 2014.

[18] P. Pande, G. L. Monroy, R. M. Nolan, R. L. Shelton, and S. A. Boppart, "Sensor-Based Technique for Manually Scanned Hand-Held Optical Coherence Tomography Imaging," *Journal of Sensors,* Article p. 7, 2016, Art. no. 8154809.

[19] P. Hassenpflug, R. W. Prager, G. M. Treece, and A. H. Gee, "Speckle classification for sensorless freehand 3-D ultrasound," *Ultrasound in Medicine and Biology,* Article vol. 31, no. 11, pp. 1499-1508, Nov 2005.

[20] A. Ahmad, S. G. Adie, E. J. Chaney, U. Sharma, and S. A. Boppart, "Cross-correlation-based image acquisition technique for manually-scanned optical coherence tomography," *Optics Express,* Article vol. 17, no. 10, pp. 8125-8136, May 2009.

[21] X. Liu, Y. Huang, and J. U. Kang, "Distortion-free freehand-scanning OCT implemented with real-time scanning speed variance correction," *Optics Express,* Article vol. 20, no. 15, pp. 16567-16583, Jul 2012.

[22] N. Uribe-Patarroyo and B. E. Bouma, "Rotational distortion correction in endoscopic optical coherence tomography based on speckle decorrelation," *Optics Letters,* Article vol. 40, no. 23, pp. 5518-5521, Dec 2015.

[23] Y. W. Wang, Y. H. Wang, A. Akansu, K. D. Belfield, B. Hubbi, and X. Liu, "Robust motion tracking based on adaptive speckle decorrelation analysis of OCT signal," *Biomedical Optics Express,* Article vol. 6, no. 11, pp. 4302-4316, Nov 2015.

[24] A. M. D. Lee, G. Hohert, P. T. Angkiriwang, C. MacAulay, and P. Lane, "Dual-beam manually-actuated distortioncorrected imaging (DMDI) with micromotor catheters," *Optics Express,* Article vol. 25, no. 18, pp. 22164-22177, Sep 2017.

[25] J. G. Fujimoto and W. Drexler, "Introduction to OCT," in *Optical Coherence Tomography: Technology and Applications*, W. Drexler and J. G. Fujimoto, Eds. Cham: Springer International Publishing, 2015, pp. 3-64.

[26] Arges. (2018, 24.02.18). *Objectives, telescope lenses and focus translators*. Available: https://www.arges.de/industrial-products/optical-components/lenses-and-optics/

[27] MATLAB, *version 9.1.0 (R2016b) Image Processing Toolbox - normxcorr2*. Natick, Massachusetts: The MathWorks Inc., 2016.

[28] MATLAB, *version 9.1.0 (R2016b) Signal Processing Tool Box - DTW*. Natick, Massachusetts: The MathWorks Inc., 2016.

[29] T. Mondal, N. Ragot, J. Y. Ramel, and U. Pal, "Comparative study of conventional time series matching techniques for word spotting," *Pattern Recognition,* Article vol. 73, pp. 47-64, Jan 2018.

9 Appendix A

Imaging Head Preparation Steps

1. Glue GRIN into 10 mm ferrule with the 8 degree face of the GRIN inside the ferrule and the flat face protruding from the ferrule a couple mm
2. Place ferrule with GRIN into the clamp which is situated in front of beam profiler
3. Slide pigtail into the ferrule and tape down the fiber with some tension to hold in place
4. Slide the pigtail back and forth inside the ferrule while twisting to find the orientation that aligns the 8 degree faces.
5. Adjust z of the profiler to give show two beams (ensuring you do not hit the GRIN with the relay lens). Likely it is required to rotate the clamp (and therefore the ferrule) until the two beams are horizontally aligned.
6. Adjust the center of the beam in the profiler to the relay center written on the lens (x=118 y=159). Take note of the beam width.
7. Once horizontally aligned and 8 degree faces flush, tape the fiber to the stage situated distal to the clamp and beam profiler. Keep some slack to ensure faces are still flush
8. Slowly move the stage, and therefore the pigtail, distal to the clamp and ferrule. Observe the beam width. The beam width should remain constant until the pigtail and GRIN start to separate. Make note of this location as the zero point.
9. Rotate 9.5 notches (0.0095 in ~ 0.242 mm) from the zero point. This should be the separation between the GRIN and pigtail to provide collimation. But test around this area to find the point of collimation. Test this by changing the z of the beam profiler and observing the beam width.
10. Once pigtail is in the correct location, use syringe to add glue then cure it while still held down by the clamp and tape. Cover the beam profiler to prevent damage from the UV light.
11. If secured, remove the probe and add more glue where needed

Notes:

- Locating the relay lens ~ 3.6mm from the front of the GRIN should display the intersection of the two beams. At this point (precisely located by minimizing the beam waist), I reset the beam profiler position to the relay center (118,159).
- I moved the beam profiler backwards away from the GRIN to observe the beam separation. After 4 tenths of an inch I started getting beam separation and recorded the beam waist and then observed and recorded the center and beam waist (both u and v) of each beam individually (only A then only B)
- I moved back a tenth of an inch two more times and recorded again.