

CRANFIELD UNIVERSITY

PASQUALE LAFIOSCA

VISUAL-BASED AUTOMATED AIRCRAFT
INSPECTIONS FOR 3D SKIN DAMAGE

SCHOOL OF AEROSPACE, TRANSPORT AND
MANUFACTURING

PhD

Supervisors:

Dr I. S. Fan

Prof. I. Jennions (2019-2020)

Prof. N. P. Avdelidis (2020-2022)

September 2022

CRANFIELD UNIVERSITY

SCHOOL OF AEROSPACE, TRANSPORT AND
MANUFACTURING

PhD

PASQUALE LAFIOSCA

Visual-based Automated Aircraft Inspections for 3D Skin
Damage

Supervisors:

Dr I. S. Fan

Prof. I. Jennions (2019-2020)

Prof. N. P. Avdelidis (2020-2022)

September 2022

© Cranfield University 2022. All rights reserved. No part of
this publication may be reproduced without the written
permission of the copyright owner.

Abstract

Visual inspection is the common mean to detect damages on the aircraft skin. Standard maintenance programmes require engineers to perform frequent inspections that are costly, time consuming, hazardous and subject to human factors. Engineers are required to inspect all the areas, including crown, wings and vertical stabiliser, and personally evaluate the damage. Dents, in particular, are flaws that are challenging to detect and measure due to undefined boundaries, complex geometry and difficult access. Because of these characteristics, dents cannot be detected by monocular cameras and automation of their inspection has been lacking momentum, generally limited to manually operated 3D scanning tools. Moreover, no solution has been explored to replace the human judgement of the damage. The aim of this work is the design of an automatic system to inspect the aircraft skin, identify dents, measure and report them to the engineer, thus demonstrating the feasibility of such autonomous task via unmanned aerial vehicles. After reviewing the state of the art, data is acquired by means of a single-shot structured-light algorithm for 3D scanning based on Fourier transform profilometry and compatible with the use of unmanned aerial vehicles, yet delivering submillimetre accuracy. Machine learning is then considered for the autonomous identification of dents, implemented through a novel point cloud segmentation algorithm. Finally, a mathematical model is proposed to evaluate dent shapes, replacing the current reporting standards by allowing accurate and comparable dimensional evaluation. The three main contributions operate together to enable autonomous aircraft dent inspections, whose feasibility is demonstrated by experiments with a prototype system. This work paves the way for future automated systems capable to increase safety and advance aircraft inspection reliability, while reducing human

workload, downtime and thus costs.

Keywords

3D scanning; Aircraft inspections; Dents; Maintenance; Point clouds; Segmentation.

Contents

Abstract	III
Contents	V
List of Figures	VIII
List of Tables	XI
List of Abbreviations	XII
Acknowledgements	XIII
Candidate background	XVII
1 Introduction	1
1.1 Aircraft inspections	1
1.2 Skin inspection and its automation	4
1.3 Dents	8
1.4 Considerations for automating MRO inspections	12
1.5 Aim and objectives	13
1.6 Overall thesis structure	15
1.7 Main achievements and contributions to knowledge	17
References	18
2 Review of Non-contact Methods for Automated Aircraft Inspections	22
2.1 Introduction	23
2.2 Pattern recognition	24
2.3 Triangulation	26
2.4 Time of flight	38
2.5 Thermography	41
2.6 Conclusions	43
References	45

3	Rectifying Homographies for Stereo Vision: Analytical Solution for Minimal Distortion	49
3.1	Introduction	50
3.2	Setting the problem	52
3.3	Geometric interpretation and analytical derivation of the minimum	57
3.4	Algorithm summary	64
3.5	Discussion	65
3.6	Conclusions	67
	References	69
	Appendix A	71
4	Automated Aircraft Dent Inspection via a Modified Fourier Transform Profilometry Algorithm	73
4.1	Introduction	74
4.2	Proposed Method	77
4.3	Computer Simulations	87
4.4	Experiments	90
4.5	Discussion	96
4.6	Conclusions	97
	References	98
5	Automatic Segmentation of Aircraft Dents in Point Clouds	102
5.1	Introduction	103
5.2	Previous work	104
5.3	DentNet segmentation strategy	106
5.4	Simulations and experiments	111
5.5	Conclusions	114
	References	116
6	Aircraft Skin Inspections: Towards a New Model for Dent Evaluation	118
6.1	Introduction	119
6.2	Problem statement	120
6.3	Proposed model	122
6.4	Use cases	127
6.5	Conclusions	129
	References	130
7	Discussion and Limitations	132
7.1	Overview of the proposed process	132
7.2	Demonstrative experiment on an aircraft rudder	133

CONTENTS

VII

7.3	Light source and laser speckle effect	138
7.4	Ambient light conditions	143
7.5	Surface reflections	144
7.6	Employment of UAVs inside and outside the hangar	146
	References	147
8	Conclusions and Future Work	150
	Glossary	153

List of Figures

1.1	Aircraft maintenance performed by U.S. Air Force	3
1.2	Using a flashlight helps in detecting surface cracks.	4
1.3	Example of dent and buckle chart of a B737-400	5
1.4	The Aloha accident	6
1.5	B737-400 fuselage skin allowable damage for dents	9
1.6	Light conditions affect dent detection	9
1.7	Corrective actions for dents on the fuselage skin	10
1.8	The use of a flashlight	10
1.9	The boundaries of a dent are not clearly defined	11
1.10	Common depth gauge measuring errors.	11
1.11	Overall structure of the research activity	15
2.1	Basic CNN approach	25
2.2	An example of complete CNN layout	25
2.3	The search for a point is restricted to the epipolar line	27
2.4	Passive stereo system elements	28
2.5	Recovering the depth information from disparity	28
2.6	Local algorithms look for a window w shifted by d_P	30
2.7	Common active triangulation configurations	32
2.8	Point gauges make one-dimension measures	33
2.9	Projecting an identical stripe multiple times causes ambiguity	33
2.10	Common stripe encoding strategies	34
2.11	An example of De Bruijn sequence	35
2.12	Schematics of smart pixels	39
2.13	Time of flight double reflection	40
2.14	Planck's law	41
2.15	Indicative features of the discussed technologies	43
2.16	Comparison of technologies with respect to types of damage	44
3.1	Representation of coordinate systems	53
3.2	Common orientation of the virtual camera pair	57

3.3	Possible trend of the distortion	63
3.4	Images of an synthetic scene with extremely skewed camera positions . .	67
3.5	Images of a synthetic scene	68
3.6	Images of a real scene	69
4.1	System configuration	78
4.2	A fringe pattern with period T_p	79
4.3	Module of the FFT	81
4.4	Virtually built reference image	82
4.5	Illustrative reference and object phases	84
4.6	Wrapped object phase and wrapped phase shift	85
4.7	The unwrapped object phase	86
4.8	Section of the reference dent function	87
4.9	Slower trend of the phase shift	88
4.10	Wrapped phase and phase shift	88
4.11	Unwrapping of the simulated dent signal	89
4.12	Comparison of 3D reconstruction using Itoh's algorithm	90
4.13	Comparison of 3D reconstruction using Estrada's noise-robust algorithm .	91
4.14	Input images for sample A	92
4.15	Calculated and ground truth depth of sample A	93
4.16	Comparison of 3D reconstruction for sample C	94
4.17	Comparison of 3D reconstruction for sample C during movement	95
5.1	Example of noise distribution	108
5.2	Virtual dented surface	108
5.3	Residual image	109
5.4	Network diagram	111
5.5	Prediction results over one of the validation samples	113
5.6	Prediction comparison	115
6.1	Traditional dimensional evaluation of dents	120
6.2	Three illustrative dents with the same length, width and depth	121
6.3	Reference dent function	123
6.4	Expected workflow of an automated dent inspection system.	128
6.5	The complex shape of a dent	129
6.6	Shape which could potentially represent the dent of Fig. 6.5	129
7.1	The dented Eurostar rudder	133
7.2	Detail of the dent on the Eurostar rudder	133
7.3	Demonstrative experiment setup	134

7.4	Scanned point cloud of the dented area	135
7.5	Point cloud detail	135
7.6	Segmentation result with standard threshold	136
7.7	Segmentation result with increased thresholds	137
7.8	Shape manually fitted to the rudder dent	138
7.9	Laser projection causes the speckle effect	139
7.10	Violation of the sampling theorem	140
7.11	Liquid crystal cell	142
7.12	Reduction of speckle contrast	142
7.13	Reflection over Cranfield B737-400 fuselage.	144
7.14	Scan of dark blue fuselage on the Cranfield B737-400.	145
7.15	Netting in place to contain UAVs	147

List of Tables

1.1	Comparing humans and machines performance	13
4.1	RMS reprojection errors of the first configuration	90
4.2	Dimensions of dent samples	91
4.3	Compared MAE with 2D Itoh's unwrapping algorithm	92
4.4	Compared MAE with Estrada's unwrapping algorithm	93
4.5	RMS reprojection error of the second configuration	94
4.6	Compared MAE over both static and dynamic samples	95
4.7	Detection and measures of dents in MRO	96
5.1	Confusion matrix training with Gaussian noise only	112
5.2	Confusion matrix training with both noise signals and Gaussian noise . .	112
5.3	Dimensions of dent samples	113
6.1	Examples of the proposed dent model	127

List of Abbreviations

AC	Alternating current
AI	Artificial intelligence
CCS	Camera coordinate system
CNN	Convolutional neural network
DC	Direct current
DVI	Detailed visual inspection
FCN	Fully convolutional network
FOD	Foreign objects debris
FFT	Fast Fourier transform
FTP	Fourier transform profilometry
FPS	Frames per second
GVI	General visual inspection
ICS	Image coordinate system
IoU	Intersection over union
LC	Liquid crystal
LCD	Liquid-crystal display
MAE	Mean absolute error
MRO	Maintenance, repair and overhaul
NDI	Non-destructive inspection
NDT	Non-destructive testing
RMS	Root mean square
SRM	Structural repair manual
UAV	Unmanned aerial vehicle
WCS	World coordinate system

Acknowledgements

English

Embarking on this adventure has been a great opportunity, made possible by audacity and a good dose of luck. Adventure and research are in fact two similar concepts. Both can never be well planned and generally lead to unexpected results.

I found myself faced with countless challenges while trying to propose solutions to a problem that, until a few months before, was completely unknown to me. And I have learned to consider things from multiple points of view before coming to any conclusion. There were grey days and mistakes, but then always rewarded by some satisfaction. It was only later down the road that I realized that moving between uncertainty and instability is the very nature of a PhD, if not of the whole life.

I must thank my wife Marta for the continuous support and patience (a lot): she has been an advisor and a victim of my abstruse reasoning from the beginning to the end. It is thanks to her that I was able to move forward with great confidence, both professionally and personally. Moreover, during all this intricate process, Tommaso literally swooped between me and the monitor and brought me back down to earth, providing me with a new scale of values. I owe to them what I am today.

I thank my supervisors Dr Fan Ip-Shing, Prof. Ian Jennions and Prof. Nicolas Peter Avdelidis and the many people who have supported my project over the years, in-

cluding: Abdul Abushalla, Mark Droznika, Matteo Migliorini, Dimitri Panagiotakopoulos, Jonathan Pelham, Anna and Richard Plaster, Philipp Reichen, Mauro Righi, Nathan Spiller, Michael Ting, Cengiz Turkoglu.

In conclusion, a special thanks goes to all the anonymous people who have freely shared illuminating explanations and ingenious pieces of code, drops of precious knowledge increasingly hidden in the midst of the overflowing data stream of today's world.

Italiano

Intraprendere questa avventura è stata una grande opportunità, resa possibile per audacia e una buona dose di fortuna. Avventura e ricerca sono infatti due concetti simili. Entrambe non possono mai essere ben pianificate e generalmente conducono a risultati inaspettati.

Mi sono trovato di fronte a innumerevoli sfide mentre cercavo di proporre soluzioni a un problema che fino a pochi mesi prima mi era del tutto sconosciuto. E ho imparato a considerare le cose sempre da più punti di vista prima di arrivare a qualsiasi conclusione. Non sono mancati giorni grigi ed errori, poi però sempre ripagati da qualche soddisfazione. Solo sul tardi ho capito che muoversi tra l'incertezza e l'instabilità è la natura stessa di un dottorato di ricerca, se non dell'intera vita.

Devo ringraziare mia moglie Marta per il continuo supporto e la (tanta) pazienza: lei è stata consigliera e vittima dei miei astrusi ragionamenti dall'inizio alla fine. E' grazie a lei che sono riuscito ad andare avanti con grande fiducia, sia professionalmente sia personalmente. Inoltre, durante tutto questo intricato processo, Tommaso è letteralmente piombato tra me e il monitor e mi ha riportato coi piedi per terra, fornendomi una nuova scala di valori. Devo a loro quello che sono oggi.

Ringrazio i miei supervisor Dr Fan Ip-Shing, Prof. Ian Jennions e Prof. Nicolas

Peter Avdelidis e le tante persone che hanno supportato il mio progetto in questi anni, tra cui: Abdul Abushalla, Mark Droznika, Matteo Migliorini, Dimitri Panagiotakopoulos, Jonathan Pelham, Anna e Richard Plaster, Philipp Reichen, Mauro Righi, Nathan Spiller, Michael Ting, Cengiz Turkoglu.

Per concludere, un ringraziamento speciale va a tutti gli anonimi che hanno liberamente condiviso illuminanti spiegazioni e geniali pezzi di codice, gocce di preziosa conoscenza sempre più nascoste in mezzo allo straripante flusso di dati del mondo odierno.

P. L.

Science is the belief in the ignorance of experts.

(Richard Feynman)

Candidate background

Enthusiastic about all forms of technology since his early years, in 2009 he began Bachelor's studies on Computer Science and Automation Engineer at the Polytechnic University of Bari. In 2011 he decided to put his passion on hold to join the Italian Guardia di Finanza, a peculiar military police force dealing with financial crime.

In 2014, after completing the training integrated by a Bachelor's degree in Business Law, he was assigned to the *Servizio Centrale Investigazione Criminalità Organizzata* (Central Investigation Centre against Organised Crime) in Rome and participated in major criminal investigations covering data analysis, intelligence and technical operations on the field. While committed to the Police, he resumed his Engineering studies part-time, obtaining the BEng degree in 2018.

In 2019 he joined Cranfield University for a funded PhD project to automate aircraft inspections.

Chapter 1

Introduction

1.1 Aircraft inspections

The total number of accidents associated with the operation of an airplane was 393 in the decade 1992-2001, and it slightly decreased to 320 in the decade 2011-2020, as reported by Boeing's statistical summaries [1, 2]. Maintenance, in particular, is reported as the primary cause for 7.4% of hull losses¹ for the period 1959-2020, although decreased to only 3.4% in the decade 2011-2020 for the constant effort towards greater safety. Errors related to maintenance have been associated with up to 15% of major aircraft accidents, exceeded only by pilot errors, with a total of 1474 accidents and 504 fatalities according to reports of the United States National Transportation Safety Board in the period 1988-1997 only [3]. The steady growth of commercial aviation will inevitably affect the absolute number of accidents. In addition, incidents and delays should be added to all the above statistics.

Nevertheless, commercial aviation is the safest mean of transport [4]. This is possible thanks to strict regulations and numerous safety measures in place. Huge costs are involved to implement the latter, with maintenance taking about 10-20% of the aircraft operating costs [5]. As also confirmed by the Department for Business Innovation & Skills of the United Kingdom, maintenance accounts for 13% of the total operating costs of a typical commercial aircraft operator, and it is therefore comparable to fuel that takes 17% [6, 7].

As part of the maintenance process, aircraft inspections are examinations and manual

¹Damage beyond economical repair, resulting in the total loss of the aircraft.

checks that range from a simple walk around to a detailed inspection that involves complete disassembly and complex tools [8]. Inspections are designed to assess and maintain the aircraft *airworthiness*, defined as the aircraft suitability for a safe flight or, more precisely, “the ability of an aircraft or other airborne equipment or system to be operated in flight and on the ground without significant hazard to aircrew, ground crew, passengers or to third parties” [9]. Certification of airworthiness is released in accordance with standards defined by national and international aviation authorities. However, the certification is not to be seen as a one-off objective, instead *continued airworthiness* should be put in place, i.e. keeping the aircraft airworthy throughout its life. This different mindset serves both the purposes of increasing flight safety and reducing costs through an integrated management of the asset, as operating failures and malfunctions are significantly reduced by early detection of wear, minor defects and preventive maintenance.

Continued airworthiness is implemented through a *maintenance programme* that must be produced by the aircraft operator and approved by the aviation authority. It contains a list of tasks to be performed at certain intervals counted as time, flight cycles or flight hours. In this context, inspections are considered the *backbone* of a good maintenance programme [8].

Maintenance programmes are different in function of the aircraft model and its use, however, inspections (or *checks*) generally share a pyramidal order, sometimes referred to as *ABC check* system, although each operator may use its own categorisation [7]. *Light* checks are performed more frequently, often carried out by the operator’s own maintenance team, while *heavy* checks, more labour-intensive, are carried out at longer intervals by means of maintenance, repair and overhaul (MRO) companies in hangars equipped for the scope. An illustrative division of such checks may look like the following:

- *Line maintenance*: Between operations (10 man-hours per week), usually done at night by the operator’s own maintenance team. Inspection of wheels, brakes, fluid levels and any other routine maintenance that may have been reported by aircraft sensors.
- *A checks*: Every 10 weeks (100 man-hours). Replacement of filters and fluids, inspection of hydraulics and emergency equipment. Visual examination of the airframe (structure) and avionics (electronic systems).
- *B checks*: Every 6 months (200 man-hours). It is an extensive version of the previous, involving opening of panels to allow a deeper inspection. Sometimes these tasks are included in the A checks as standard.

- *C and D or “heavy” checks*: Every 2-10 years (3000-10000 man-hours). Depending on the required level, most parts of the aircraft are disassembled to access the internal structure and look for wear, leaks, cracks or corrosion, cabin seats are removed together with any other equipment that limits the access. The same happens for landing gears, engines and electronics. During these occasions, upgrade and repainting may also take place.

In addition to the above, *unscheduled* maintenance is performed following reports of malfunctions or particular events (e.g. lightning, bird strike, hail).

Aircraft passengers often witness to the pilot doing a walk around following a pre-defined path to inspect certain elements of the aircraft, as part of the *pre-flight* checks. Although these are not directly part of the maintenance programme, they represent an important occasion for anomalies to be compared with the maintenance records and take any necessary actions.

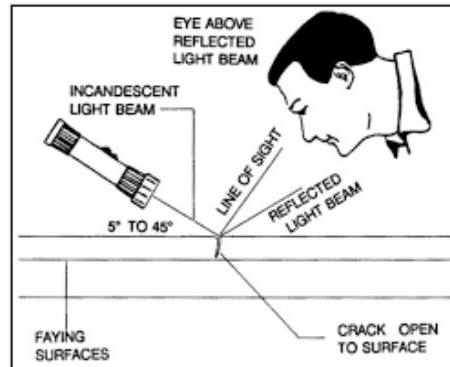
In MRO the *general visual inspection* (GVI) is by far the first and most common inspection method to assess the condition of any aircraft component. By definition GVI is a visual examination made from within touching distance under available light conditions and oriented to detect obvious damage, failure or irregularity [10].



Credits: University of Wisconsin-Milwaukee.

Figure 1.1: Aircraft maintenance performed by U.S. Air Force at Tainan Airport (Taiwan) in 1950s. Today the way of doing GVI is left mostly unchanged.

Any anomalies found with GVI may be investigated further with a *detailed visual inspection* (DET). This is an intensive examination [10] and differs from GVI for the closer distance and the use of artificial light (Fig. 1.2) or inspection aids like mirrors or magnifying lenses.



Credits: EASA.

Figure 1.2: Using a flashlight helps in detecting surface cracks.

Depending on the situation and the inspected part, a *special detailed inspection* (SDI) could be necessary. The difference from the previous is mainly in the use of non destructive testing (NDT) techniques. Common ones include the use of penetrant dye, magnetic particle, eddy current, ultrasonic or X-ray radiography [8].

It must be noted that visual inspections (namely GVI and DET) account for 80%-99% of the total, while the use of NDT is limited to the remaining part [11]. Therefore, visual inspection is the first line of defence against incidents and its improvement can significantly increase reliability and efficiency of early fault detection.

In spite of major differences in how they are scheduled and managed, aircraft inspections represent a significant bottle-neck in terms of safety, costs, downtime and management effort. Therefore, the progressive introduction of automation in aircraft inspection and maintenance sees both operators and manufacturers particularly keen to explore solutions to reduce costs and improve reliability of the operations, while being compliant with regulations.

1.2 Skin inspection and its automation

Visual inspections include the search of a wide variety of anomalies related to structure, engines, electronics, cabin, etc.

In particular, the external surface of the aircraft, or *skin*, is exposed to several causes of damage [12] like hail, birds, rocks on the runway or any other objects found in an inappropriate location, generally known as foreign object debris (FOD). Very often service vehicles hit the aircraft during operations, causing dents and scratches. Common areas of damage include, for example, cargo and passenger doors or the flaps at the back of the wheels.

Due to its extent and its complexity, skin inspection is labour-intensive, hazardous for the human engineer and sometimes unreliable. On the other hand, skin is *directly accessible* without disassembling parts, which is a highly desirable characteristic for the introduction of automation.

During inspections, all the externally visible damages are reported and tracked in the *dent and buckle chart*, as part of the onboard documentation. It generally consists in a diagram of the aircraft with different views and numbered parts, like frames and longerons (Fig. 1.3), followed by a list of damages and their location and status. A damage can be either *within allowable limits*, *temporarily* or *permanently repaired*, but in any case this must be mentioned in the dent and buckle chart.

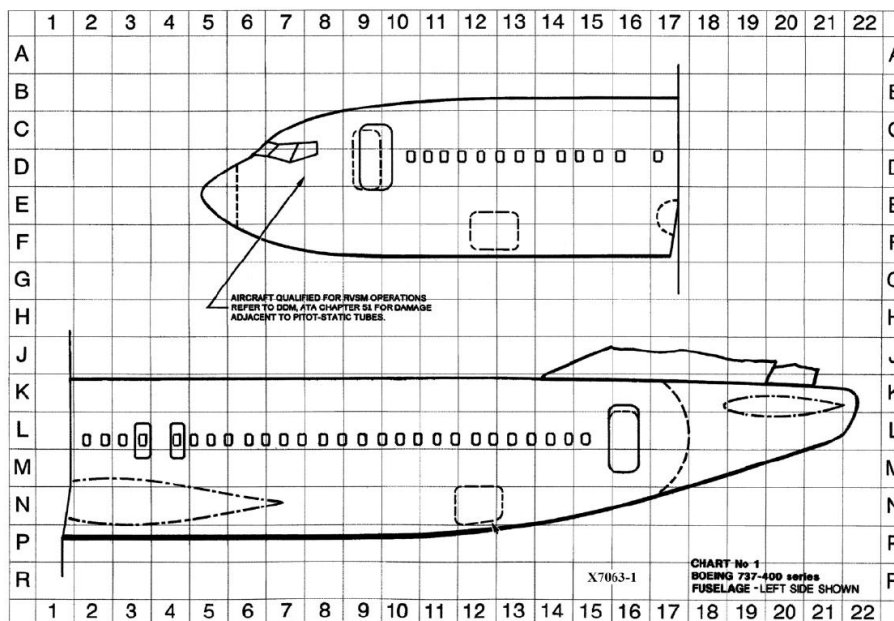


Figure 1.3: Example of dent and buckle chart of a B737-400. Side fuselage view.

Skin inspections are also time-consuming and costly. For example, the wing area of the popular B737-800 is 125 m^2 , while an approximate calculation of the skin area leads to 700 m^2 in total to be inspected for damages. This may be compared to looking for a bunch of penny-size defects over about 3 tennis courts. However in this case the surface to

inspect is not flat nor easily reachable: the maximal height (top of the tail) of a B737-400 is 11.30m, while for an A380-800 it reaches 24.1 m, taller than a five-storey building.

As such, it is no surprise that a GVI of the aircraft skin may take 2-3 full days to be completed. Engineers are required to reach and inspect all the areas, including the upper ones: stairs, platforms, cranes and climbing equipment have to be used to comply with the required *touching distance*, therefore safety becomes a critical aspect to consider during operations.

In this scenario the identification and measurement of damages often result in limited reliability because of the several human factors involved, like stress, fatigue or boredom. Humans, unintentionally, do not provide an objective assessment and it is common to have different outcomes from different inspectors for the same damage or even that the same inspector later disagrees with himself. This further underlines how automation can improve such processes by standardising measurements and ensuring repeatability.

The infamous accident of the Aloha Airlines Flight 243 on 28th April 1988 is exemplary: disbonding and multiple site fatigue cracks went undetected during the previous inspections of the B737-200 and caused an explosive decompression of the fuselage [13]. More fatalities were avoided thanks only to pilot's performance.



Figure 1.4: The Aloha accident: the B737-200 lost its upper part due to undetected damage. A flight attendant lost her life and several people were injured.

Even when not fatal, human errors during inspection and maintenance caused turn-backs, delays and cancellations [14]. The introduction of some automatic processes would reduce both costs and downtimes, increase reliability and safety, and produce a standard inspection outcome, independent from human factors. Consequently, solutions

like robots, sensors and, more recently, flying drones equipped with cameras are increasingly being considered in MRO. Although technology may be available, however, it has not yet translated easily into a direct application on the field and substantial research is still required to meet commercial aviation needs.

Several attempts have been made so far. Back in 1996, a suction crawling robot was developed by the NASA's Jet Propulsion Laboratory for autonomous skin inspection with configurable NDT tools [15]. More recently, a four-wheeled suction-based *Vortex Robot* with a payload of 8 kg was developed for inspecting the surface using ultrasonic and thermal imaging [16]. Automated inspection within a hangar and using fixed commercial cameras has also been proposed [17]. However, the use of unmanned aerial vehicles (UAV) offers greater flexibility of movement and reduced inspection times. For example, a drone was tested by Airbus [18] for indoor visual inspections, equipped with a camera for upper-part inspection and a laser-based obstacle detector. As demonstrated by Mainblades, outdoor airport drone inspection is also possible directly at the active airfield, further reducing downtime [19]. To date, the use of UAV for inspections is a promising field of research.

Nevertheless, the effective introduction of automation in MRO is still very limited. Reasons for this could be traced back not only to the strict regulations in place, but mainly to the lack of comprehensive *end-to-end* solutions capable to actually relieve the human engineer from a certain task with the required level of confidence. Humans, in fact, are really good at considering the current scenario and critically evaluating the object under different aspects. Delamination, for example, is first assessed by engineers with only a *tap test*, which involves the knocking of an object's surface with a small hammer or a coin and judging its integrity by the sound that emanates from it [8].

For machines, the first challenge is to acquire reliable data from outside world. And each technology is capable of detecting only certain types of damages: for example, monocular cameras can detect scratches, lightning and missing parts fairly easily, but generally cannot provide measures and catch smooth shape deformations like dents, nor inner damage like delamination. Also, each method may be differently affected by environmental factors (e.g. thermal imaging by temperature), with some systems being less likely to detect certain types of damage. Non-contact NDT technologies in relation to damage types are discussed in Ch. 2.

Secondly, the data interpretation and analysis that follows should be completely autonomous. The human level of adaptability and judgement in MRO seems still irreplaceable and machines, in general, give lower level of confidence or remit the final evaluation

to a human. Today an exception to this rule could be machine vision, particularly object recognition, which is employed in some of the commercial products oriented to the MRO industry [19]. However, even if its accuracy may resemble the human levels, an immediate human-machine comparison is not straightforward [20]. Therefore, the introduction of automated technologies is expected to be cautious but steady in the foreseeable future.

1.3 Dents

Among all the different types of skin damage, this work addresses the challenge of automatic detection and measurement of dents.

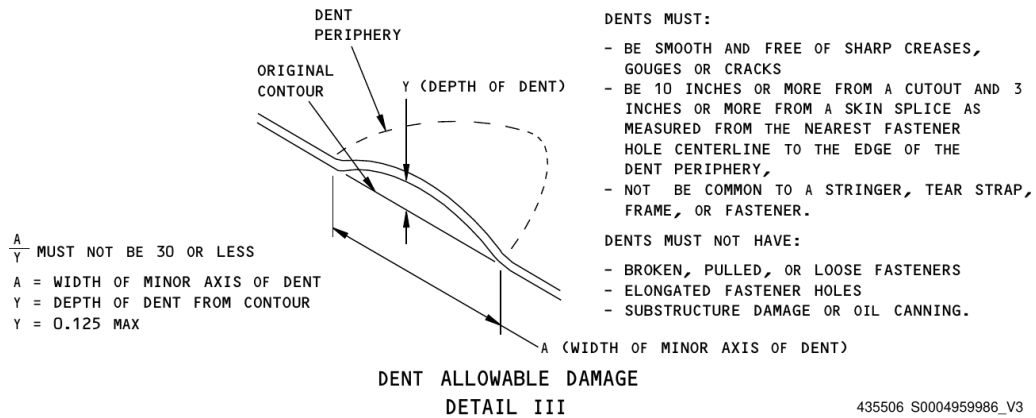
In order to identify the main drivers of this project, visits were arranged to a several MRO facilities in the UK, providing services within international airline operators. This allowed a broad understanding of the environmental factors and the gathering of technical information from experienced engineers, who commonly underlined the importance of dents inspection within the MRO industry, particularly due to high frequency of occurrence and limited employment of advanced tools for their evaluation.

Dents affect structural performance, reducing the amount of time that the material will last before failure or its *fatigue life*. It was found that the presence of dents affects the propagation of cracks and the fatigue life is shorter for deeper dents [21, 22].

Statistics collected on composite parts of B737-800 and B757-200 aircrafts confirm that dents represent a relevant part of the overall damages, their frequency of occurrence being second only to scratches [23]. MRO engineers are well aware of dent inspection workload and associated safety risks. As such, overcoming the challenges that come with the nature of the damage itself could result in a breakthrough in terms of inspection efficiency and safety.

In the B737-400 SRM [24] a dent on metal is defined as “*a damaged area that is pushed in from its normal contour with no change in the cross-sectional area of the material. The edges of the damaged area are smooth. This damage is usually caused by a hit from a smoothly contoured object. The length of the dent is the longest distance from one end to the other end. The width of the dent is the second longest distance across the dent, measured at 90 degrees to the direction of the length*”. For it to be considered a dent, no other kind of damages like gouges or cracks must be present at its location (Fig.1.5).

Similarly, a dent on composite materials is defined as “*a concave depression which*



Credits: The Boeing Company.

Figure 1.5: B737-400 fuselage skin allowable damage for dents.

does not break the fibers. A dent can cause a delamination or disbond”.

Dent detection is a particularly subjective task as a depression may not be spotted or even not reported because not considered a dent, like, for example, a depression across a stringer or a frame. In this regard, often engineers need to use available or artificial sources of light (e.g. a flashlight) to spot dents based on the reflectivity of the surface (Fig. 1.6).

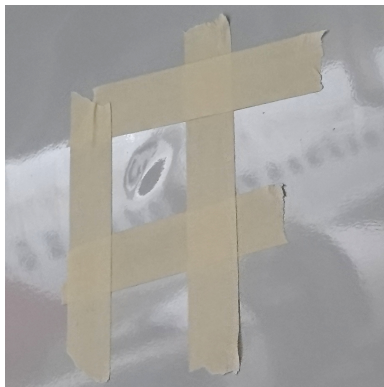
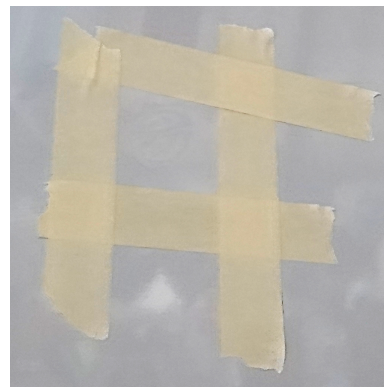


Figure 1.6: Light conditions affect dent detection. The same dent enclosed with tape is seen with and without contrast of the available hangar light.

Moreover, the SRM requires *measures* of a dent, in terms of length, width and depth, to be collected for the evaluation of the damage and the corresponding corrective actions. For example Fig. 1.7 shows the classification reported by the B737-400 SRM. Note that it is the combined evaluation of depth and width/depth ratio (indicated as A/Y by Boeing) that acts as main discriminating factor.

Length and width are mostly collected by a slightly flexible metal ruler. A rather prac-

DEPTH	A/Y RATIO	CORRECTIVE ACTION
LESS THAN 0.125 INCH	GREATER THAN 30	ALLOWABLE DAMAGE. NO REPAIRS NECESSARY.
LESS THAN 0.125 INCH	A/Y RATIO BETWEEN 15 AND 30 ($15 < A/Y < 30$)	DENT MUST BE REPAIRED PERMANENTLY. HOWEVER, YOU CAN DELAY THE REPAIR IF: 1. YOU INITIALLY DO A DETAILED VISUAL INSPECTION OF THE DENT, 2. YOU DO AN HFEC INSPECTION WITHIN 4000 CYCLES, 3. YOU DO A DETAILED VISUAL INSPECTION EVERY 4,000 CYCLES, AND 4. A PERMANENT REPAIR IS MADE BEFORE 24,000 CYCLES.
LESS THAN 0.125 INCH	LESS THAN 15	DENT MUST BE REPAIRED PERMANENTLY. HOWEVER, YOU CAN DELAY THE REPAIR IF: 1. YOU INITIALLY DO A DETAILED VISUAL INSPECTION OF THE DENT, 2. YOU DO AN HFEC INSPECTION WITHIN 4000 CYCLES, 3. YOU DO A DETAILED VISUAL INSPECTION EVERY 4,000 CYCLES, AND 4. A PERMANENT REPAIR IS MADE BEFORE 10,000 CYCLES.
DEPTH BETWEEN 0.125 INCH AND 0.25 INCH ($0.125 < \text{DEPTH} < 0.25$)	A/Y RATIO BETWEEN 15 AND 30 ($15 < A/Y < 30$)	DENT MUST BE REPAIRED PERMANENTLY. HOWEVER, YOU CAN DELAY THE REPAIR IF: 1. YOU INITIALLY DO A DETAILED VISUAL INSPECTION OF THE DENT, 2. YOU DO AN HFEC INSPECTION WITHIN 4000 CYCLES, 3. YOU DO A DETAILED VISUAL INSPECTION EVERY 4,000 CYCLES, AND 4. A PERMANENT REPAIR IS MADE BEFORE 10,000 CYCLES.
DEPTH BETWEEN 0.125 INCH AND 0.25 INCH ($0.125 < \text{DEPTH} < 0.25$)	GREATER THAN 30	DENT MUST BE REPAIRED PERMANENTLY. HOWEVER, YOU CAN DELAY THE REPAIR IF: 1. YOU INITIALLY DO A DETAILED VISUAL INSPECTION OF THE DENT, 2. YOU DO AN HFEC INSPECTION WITHIN 4000 CYCLES, 3. YOU DO A DETAILED VISUAL INSPECTION EVERY 4,000 CYCLES, AND 4. A PERMANENT REPAIR IS MADE BEFORE 16,000 CYCLES.
GREATER THAN 0.125 INCH	LESS THAN 15	DENT MUST BE REPAIRED PERMANENTLY. HOWEVER, YOU CAN DELAY THE REPAIR IF: 1. YOU INITIALLY DO A DETAILED VISUAL INSPECTION OF THE DENT, 2. YOU DO AN HFEC INSPECTION WITHIN 500 CYCLES, 3. YOU DO A DETAILED VISUAL INSPECTION EVERY 500 CYCLES, AND 4. A PERMANENT REPAIR IS MADE BEFORE 6,000 CYCLES.
GREATER THAN 0.25 INCH	ALL	DENT MUST BE REPAIRED PERMANENTLY. HOWEVER, YOU CAN DELAY THE REPAIR IF: 1. YOU INITIALLY DO A DETAILED VISUAL INSPECTION OF THE DENT, 2. YOU DO AN HFEC INSPECTION WITHIN 500 CYCLES, 3. YOU DO A DETAILED VISUAL INSPECTION EVERY 500 CYCLES, AND 4. A PERMANENT REPAIR IS MADE BEFORE 6,000 CYCLES.

Credits: The Boeing Company.

Figure 1.7: *Corrective actions for dents on the fuselage skin with respect to their measures.*

tical method to confirm the presence of a dent and measure its width involves using a torch to highlight the actual gap between the ruler and the surface (Fig. 1.8). Obviously, the human nature greatly affects the final outcome. For example, as represented in Fig. 1.9, the boundaries cannot be safely established and thus the width measure is highly subjective.



Figure 1.8: *The use of a flashlight may help to highlight the gap between surface and ruler.*

The most common way to measure depth is the depth gauge. Although the resolution

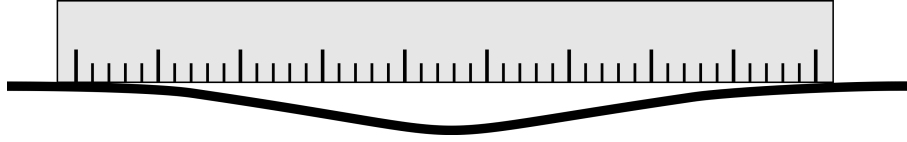
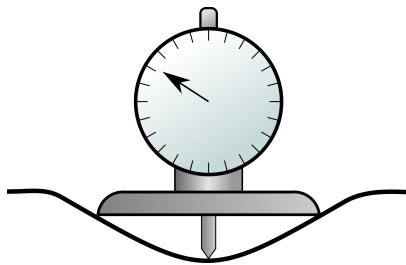
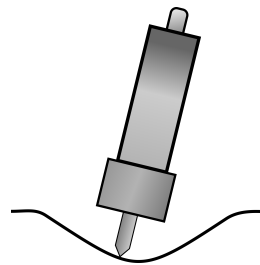


Figure 1.9: *The boundaries of a dent are not clearly defined, thus measures with a ruler are highly subjective.*

is generally at submillimetre level, tilting the gauge or using a too short bar compared to the actual dent width (Fig. 1.10) inevitably affects the measure that, again, is highly dependent on human interpretation.



(a) *The gauge bar is too short with respect to dent width.*



(b) *A wrong depth is measured by inadvertently tilting the gauge.*

Figure 1.10: *Common depth gauge measuring errors.*

In addition, the evaluation of the damage also depends on the dent position with respect to other aircraft elements and other dents. For example, two dents cannot be present in two adjacent bays² at some fuselage zones, even if within the limits of allowable damage [24].

As highlighted by Lawrence C. [25], the widespread use of composite materials increases the difficulties related to dent inspections. Aviation personnel is well experienced in evaluating damages on metal, but the same knowledge does not directly translate to composites. In the same study, the detection reliability was related to damage size, surface colour and finish. After a heavy impact, little or no damage may be visible on composites, while a metal surface subject to the same impact would present more evident signs of damage. However, composites may hide underlying damage that eventually would lead to cracks and structural failure. Therefore, new and more efficient ways to detect such damage must be developed [26].

Regardless of the material, the detection of dents remains challenging, as they lack of clear boundaries and a rigorous shape definition, and are difficult to spot without con-

²A bay is the almost rectangular area falling between stringers and frames.

trolled light or other aids. Even when found, dent measures remains disputable: human factors influence the outcome that in most cases is not repeatable at all.

The state-of-the-art in MRO proposes the introduction of handheld 3D scanning tools. Commercial 3D scanning solutions normally project multiple different patterns to get 3D data from a given scene. Although they offer high accuracy and repeatability³, their use remains limited due to short operating distances and absence of fully automated solutions. In fact, while such devices compensate for movements by using high-speed electronics, they cannot generally be employed with vibrations or in a scenario with fast relative movements, e.g. from a UAV.

Therefore, despite the potential of today's technologies gradually entering MRO, dent detection, measuring and evaluation is still left mainly to the engineer's critical judgement under the more traditional inspection method, which is inevitably subjective even when based on many years of experience. This subjectivity undermines the implementation of process standards that, instead, provide common performance evaluation and interchangeability, increased efficiency, quality and, ultimately, safety of maintenance operations.

The likely introduction of automation in the foreseeable future is thus expected to enhance a large group of aspects, as happened in many other sectors.

1.4 Considerations for automating MRO inspections

In general, the study of autonomous solutions for aircraft inspection and maintenance can be done for two distinct but interconnected reasons: reducing costs by replacing the human inspector and improving the overall performance (including safety). With maintenance accounting for a relevant part of the total operating costs, cost management is the major driver to explore autonomous solutions. Being generally faster, machines can decrease downtime costs, while the resulting standard and repeatable inspections will further reduce indirect costs and lead to a more reliable maintenance scheduling.

Nevertheless, comparing human with machine performance in MRO is not straightforward and many aspects have to be considered (Tab. 1.1): today human engineers are entrusted almost all of the tasks because they are capable of great flexibility and context awareness with respect to machines. Humans are also expensive and valuable resources for a company, not least, when their safety is at stake. On the other hand, tireless and

³See Glossary in Ch. 8.

cheaper machines (often empowered by artificial intelligence) are getting closer and closer to overcoming human performance *for certain tasks* [27]. Their downside is that they may generate catastrophic outcomes in the occurrence of unplanned or unexpected events. As such, careful design is intended to guide the gradual transition to more automated MRO.

	Human	Machine
Pros	Flexible thinking Context-aware	Tireless Faster Cheaper Expendable (in terms of safety) Repeatable standard outcome
Cons	Prone to stress, tiredness and errors Slower Subjective outcome Expensive (training, salary, etc.) Valuable (also in terms of safety)	Limited reasoning Affected by unplanned errors Lack of out-of-the-box thinking

Table 1.1: Comparing humans and machines performance in MRO inspections.

In addition, the increasingly common employment of UAVs was taken into account during the development of the project. The advantages of using drones are already tangible with respect to lightning checks [28]: flexible inspections, reaching difficult areas without climbing equipment or even outdoor [19], in much shorter amount of time and with reduced costs.

Here it is argued that drone versatility can be transferred to dent inspections as well, if the appropriate end-effector is found. In this regard, the recent beginning of a collaboration between *Donecle*, a leading company for drone inspection in aviation, and *8tree*, a company specialised in 3D surface inspection tools, confirms the current interest of the industry in this topic [29].

1.5 Aim and objectives

The overall aim of this research project is the design of a solution to advance aircraft dent inspections, standardise their outcome and reduce human factors in the process. This is achieved by focusing the research on *visual* skin inspections and designing a *fully-automated process for scanning, identifying, measuring and evaluating dents*.

The objectives of this research project were identified in two groups, as follows:

- *Research gap identification*

1. Narrow down the current *industry needs* in the context of aircraft inspections, investigating common damages and critical issues during their inspection and highlighting the potential benefits that automation can introduce.
2. Designate a *target damage* in relation to its relevance and automated inspection feasibility, particularly with respect to available technologies that could be employed for data acquisition in the MRO context.

- *System design and validation*

3. Design a reliable *scanning system*, capable of delivering a satisfactory level of accuracy in the detection of the designated damage, and suitable to fully automate the inspection process (e.g. working on UAVs).
4. Develop and test an algorithm for the *automatic identification* of relevant damage from the acquired raw data.
5. Propose the adoption of a novel method for *inspection reporting* to improve standardisation and effectiveness of MRO inspections.
6. Discuss the *feasibility* and the *limitations* of the proposed solution for its actual application to MRO inspections.

1.6 Overall thesis structure

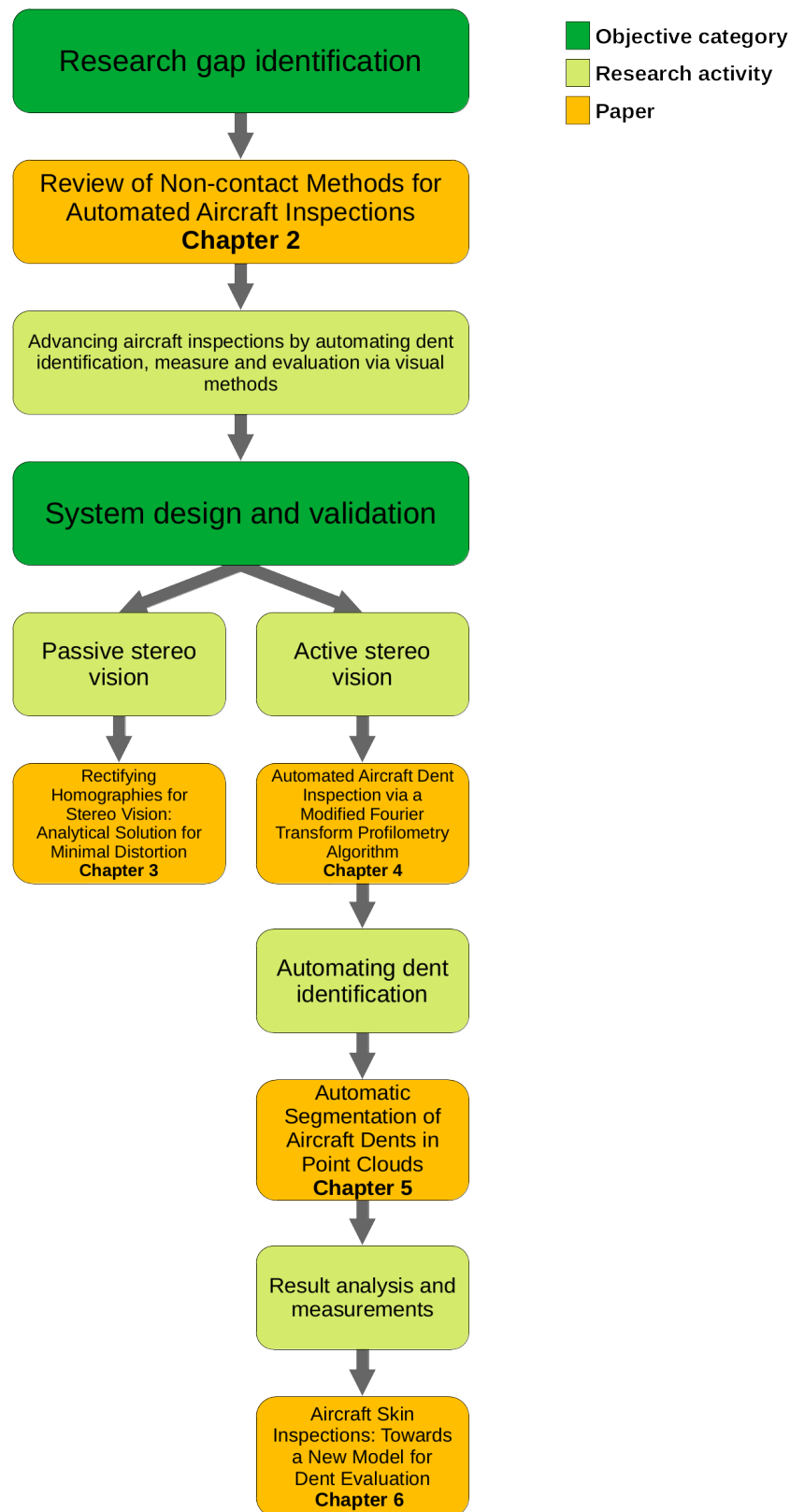


Figure 1.11: Overall structure of the research activity.

Fig. 1.11 shows the flow of the research carried out during the project. Dark green represents objective categories and light green the research activities undertaken. The main outcome is presented in this thesis as a collection of the five papers highlighted in yellow.

A brief summary of the thesis rationale, linking the papers together, is reported hereafter:

- Aircraft inspections and critical MRO aspects were initially investigated. From the considerations already underlined in Sec. 1.1 and 1.2, skin inspection was considered particularly suited for the introduction of fully automated tasks.
- *Non-contact* methods for its inspection were reviewed in Ch. 2. The choice to limit the review to non-contact methods originates from both the need of avoiding further damage during inspections and allowing quick inspections by means of a UAV free to move around the aircraft.

Dents were chosen as primary target of this research activity due their high frequency of occurrence and the limited employment of advanced tools for their evaluation. From the literature review, it appeared clear that dent inspections could benefit of stereo vision technology, already employed in some advanced, yet not fully autonomous, tools in MRO and that the joint use of UAVs with 3D scanning technologies was mostly an unexplored field of research, despite UAVs being already present in MRO.

- Stereo vision fundamentals and the general stereo correspondence problem were then taken into account. Passive stereo setups were firstly considered, as these are easy to realise and suitable for both indoor hangar and outdoor light conditions. Calibration, rectification and stereo matching algorithms were taken into account.

A paper proposing an improved stereo rectification algorithm, is shown in Ch. 3. Experiments highlighted the lack of texture over aircraft skin as a critical issue, showing that passive stereo is incapable to reach the accuracy level required by current regulations.

- *Active* stereo technologies, introducing a controlled light pattern, were then investigated for their capability of reaching higher levels of accuracy in several well-known metrology applications, from quality control to forensic analysis.

An in-depth study of the underlying working principles highlighted the necessity to account for vibrations or fast relative movements between the scanner and the target. The acquisition of a clear projected pattern is paramount to solve the stereo

correspondence problem and lead to a fully automated 3D scanning system enabled by a UAV. These considerations yield to the choice of a *single-shot* approach, that tries to extract as much information as possible to reconstruct the object 3D shape from a single stereo image, acquired by projecting a *fringe pattern*.

Thus, Ch. 4 proposes a novel, single-shot algorithm based on Fourier transform profilometry (FTP). Experiments proved the submillimetre accuracy of the proposed algorithm even at increased distances and during movement.

- Strategies for point cloud *segmentation* were explored to enable the processing of the thousands of point clouds that would normally be produced in a single inspection, and automatically detect damages.

A paper on the design, training and validation of a fully-convolutional neural network, capable of autonomously assign a damaged or undamaged label to elements of the 3D point cloud is presented in Ch. 5.

- The inappropriateness of the current dimensional classification of dents emerged. In fact, while the SRM prescribes a classification limited only to width and depth, a more advanced model was needed to exploit the accuracy of the 3D scanning systems.

A paper formulating a novel mathematical model is proposed in Ch. 6. The model is able to represent the shape of a dent using 7 parameters, enabling a more accurate dimensional evaluation of dents.

- Overall discussion and system limitations are reported in Ch. 7.
- Conclusions and future work follow in Ch. 8.

1.7 Main achievements and contributions to knowledge

This work paves the way for the automation of aircraft dent inspections by the introduction of three main novelties:

- The development and critical evaluation of a novel single-shot 3D scanning algorithm specifically designed for aircraft dent inspections.

A prototype of the scanner was built with off-the-shelf components and experiments proved its capability to acquire a faithful reconstruction of dented surfaces with a

mean absolute error (MAE) of about 0.30 mm at a distance of 120 cm and less than 0.15 mm at a distance of 60 cm.

This also showed the possibility to overcome traditional MRO 3D scanner limitations that, while being capable to reach higher accuracy, have an operating distance of about 30 cm and cannot cope with fast movements and vibrations typically found on UAVs because of their multiple-shot technology.

- The automatic analysis of the 3D point cloud data acquired.

A point cloud segmentation strategy was applied for the first time to the dent inspection scenario. The lack of data for the training of a FCN was overcome by a pseudo-real dataset of damaged surfaces, where scanning noise error was taken into account.

Although the intersection over union⁴ (IoU) decreased from 92.66% to 84.76% by using such noise-added dataset, when the network was applied to real 3D scans of dents it proved high robustness in identifying dent damage and demonstrate the validity of the suggested approach.

- The introduction of a compact but meaningful representation of complex dent shapes by using only 7 dimensional parameters.

The proposed model can potentially establish a novel standards by enhancing the comparability against historical data and structural evaluations, replacing the current way imposed by the SRMs of evaluating dents by only width and depth/width ratio, which is insufficient and ambiguous.

References

- [1] The Boeing Company. *Statistical Summary of Commercial Jet Airplane Accidents, Worldwide Operations, 1959-2000*. 2001.
- [2] The Boeing Company. *Statistical Summary of Commercial Jet Airplane Accidents, Worldwide Operations, 1959-2020*. 2021.
- [3] Scott M. Goldman, Edna R. Fiedler, and Raymond E. King. *General Aviation Maintenance-Related Accidents: A Review of Ten Years of NTSB Data*. Tech. rep. Civil Aerospace Medical Institute - Federal Aviation Administration, 2002.

⁴See Glossary in Ch. 8.

- [4] Ian Savage. “Comparing the fatality risks in United States transportation across modes and over time”. In: *Research in Transportation Economics* (2013). DOI: 10.1016/j.retrec.2012.12.011.
- [5] U PeriyarSelvam, T Tamilselvan, S Thilakan, and M Shanmugaraja. “Analysis on Costs for Aircraft Maintenance”. In: *Advances in Aerospace Science and Applications* (2013).
- [6] Hans-Henrich Altfield. *Commercial Aircraft Projects: Managing the Development of Highly Complex Products*. Ashgate Publishing Limited, 2011.
- [7] Department for Business Innovation & Skills, United Kingdom. *UK Aerospace Maintenance, Repair, Overhaul & Logistics Industry Analysis, BIS RESEARCH PAPER NUMBER 275*. 2016.
- [8] Federal Aviation Administration. *Aviation Maintenance Technician Handbook - General*. US Department of Transportation, 2018.
- [9] Military Aviation Authority. *MAA02: MAA Master Glossary, Issue 9*. 2020.
- [10] European Union Aviation Safety Agency. *Clarification of Definitions for General Visual (GVI), Detailed (DET), and Special Detailed (SDI) Inspections*. 2012. URL: <https://www.easa.europa.eu/sites/default/files/dfu/IP122%20Rev%202.pdf> (visited on 2019-10-22).
- [11] Civil Aviation Authority. *CAP716 - Aviation Maintenance Human Factors*. 2003.
- [12] Yuri DV Yasuda, Fabio AM Cappabianco, Luiz Eduardo G Martins, and Jorge AB Gripp. “Aircraft visual inspection: A systematic literature review”. In: *Computers in Industry* 141 (2022).
- [13] US Government National Transportation Safety Board. *Aircraft Accident Report, NTSB/AAR-89/03*. URL: <https://www.ntsb.gov/investigations/AccidentReports/Reports/AAR8903.pdf>.
- [14] Kara A Latorella and Prasad V Prabhu. “A review of human error in aviation maintenance and inspection”. In: *International Journal of industrial ergonomics* (2000).
- [15] Yoseph Bar-Cohen, Paul G. Backes, and Benjamin Joffe. “Multifunction automated crawling system (MACS)”. In: *Nondestructive Evaluation of Aging Aircraft, Airports, and Aerospace Hardware*. Vol. 2945. International Society for Optics and Photonics. SPIE, 1996. DOI: 10.1117/12.259079.
- [16] Angelica Brusell, George Andrikopoulos, and George Nikolakopoulos. “Vortex Robot Platform for Autonomous Inspection: Modeling and Simulation”. In: *IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society*. 2019. DOI: 10.1109/IECON.2019.8926997.

- [17] Mark Rice, Liyuan Li, Ying Gu, et al. “Automating the Visual Inspection of Aircraft”. In: *Singapore Aerospace Technology and Engineering Conference*. 2018.
- [18] Airbus. *Airbus launches advanced indoor inspection drone to reduce aircraft inspection times and enhance report quality*. 2018. URL: <https://www.airbus.com/newsroom/press-releases/en/2018/04/airbus-launches-advanced-indoor-inspection-drone-to-reduce-aircr.html> (visited on 2021-07-18).
- [19] Airbus. *Mainblades performs autonomous outdoor airport drone inspection*. 2021. URL: <https://mainblades.com/news/a-european-first-mainblades-performs-autonomous-outdoor-airport-drone-inspection/>.
- [20] Chaz Firestone. “Performance vs. competence in human-machine comparisons”. In: *Proceedings of the National Academy of Sciences* 117.43 (2020), pp. 26562–26571.
- [21] Zhigang Li, Rongxin Feng, Yan Wang, and Lishu Wang. “Experimental study on the effect of dents induced by impact on the fatigue life of 2024-T3 aluminum alloy plate”. In: *Engineering Structures* (2017). DOI: 10.1016/j.engstruct.2017.01.048.
- [22] Praveen Shivalli. “The effect of dents on fatigue life and fatigue crack growth of aluminum 2024-T3 bare sheet”. PhD thesis. 2006.
- [23] Xi Chen, He Ren, and Cees Bil. “Inspection Intervals Optimization for Aircraft Composite Structures Considering Dent Damage”. In: *Journal of Aircraft* 51 (2014). DOI: 10.2514/1.C032377.
- [24] The Boeing Company. *Structural repair manual B737-400*. 2015.
- [25] Lawrence Cook. “Visual inspection reliability for composite aircraft structures”. PhD thesis. 2009.
- [26] U. Polimeno and M. Meo. “Detecting barely visible impact damage detection on aircraft composites structures”. In: *Composite Structures* (2009). Sixth International Conference on Composite Science and Technology. DOI: <https://doi.org/10.1016/j.compstruct.2009.04.014>.
- [27] Katja Grace, John Salvatier, Allan Dafoe, Baobao Zhang, and Owain Evans. “When will AI exceed human performance? Evidence from AI experts”. In: *Journal of Artificial Intelligence Research* 62 (2018).
- [28] *Drones Finally Approved For Lightning Checks, Dent Measurement Next*. 2021. URL: <https://aviationweek.com/mro/drones-finally-approved-lightning-checks-dent-measurement-next> (visited on 2021-12-02).

- [29] *Towards dent measurement by drone*. 2020. URL: <https://www.donecle.com/2020/10/towards-dent-measurement-by-drone/> (visited on 2021-12-02).

Chapter 2

Review of Non-contact Methods for Automated Aircraft Inspections

Published as P. Lafiosca, I.S. Fan in: *Insight - Non-Destructive Testing and Condition Monitoring*, Volume 62, Number 12, 2020, doi:10.1784/insi.2020.62.12.692.

Abstract

Damage on the aircraft structure can be caused by lightning strikes, hail, accidental impacts or ageing. Scratches or dents on the aircraft surface are typical indications of impact damage. General visual inspection (GVI) is the primary way in which to detect such forms of damage. The inspection process is time consuming, raises safety concerns for the inspector and is subject to variations due to human factors. Significant inspection automation remains challenging, mainly because GVI requires the critical human ability to assess anomalies. Also, damage specifications in maintenance manuals are influenced by human interpretation. Some automated tools are beginning to be available for aircraft inspection checks. However, none of them is capable of replacing the inspector judgement yet. Humans still need to manually assess the location or the data generated by the tools. Their performance is also affected by different environmental conditions, materials and the overall characteristics of the damage. This review presents the main methods for non-contact visual aircraft inspection, explaining their basic working principles and limitations. Their suitability for automation in aircraft inspection is also discussed.

Keywords: Aircraft inspection; Pattern recognition; Triangulation; Stereo vision; Structured light; Time

of flight; Thermography.

2.1 Introduction

To maintain performance and safety, aircraft are checked both pre-flight and periodically. Pre-flight checks follow a defined checklist comprising an exterior walk around with visual inspection of critical parts such as sensors, probes, engine fans, landing gear and structural components. The pilot looks for leakages, cracks or other abnormal circumstances. Periodic checks, instead, are more extensive. They are carried out by approved maintenance, repair and overhaul (MRO) companies and, depending on type, may take from hours to months to complete.

General visual inspection (GVI) is the primary way to detect anomalies. It is made from within touching distance under available light conditions and oriented to detect obvious damage, failure or irregularity [1]. GVI is frequent, time consuming and affected by environmental circumstances.

For example, in traditional human GVI the probability of detection (POD) of *dents* changes with different surface colour, finish (e.g. gloss surfaces are easier to inspect) and light conditions. Also different depth and width/depth ratio affect the POD [2].

After a potential damage is found, a detailed inspection is performed. Detailed inspections usually require the part to be cleaned up and the use of specific tools to measure the damage against the limits provided by the structural repair manual (SRM), issued by the manufacturer. In both cases the inspector is required to get within touching distance of the aircraft part. This could be at height on the upper-fuselage or wings, raising safety concerns, slowing down the process and increasing costs. Moreover he/she is subjected to a variable psychophysical state (e.g. boredom or fatigue) that may affect results.

Today many tools are readily available for the inspection process, but the final assessment still relies on the inspector's experience. Drones equipped with a camera combined with pattern recognition algorithms are used to evaluate aircraft airworthiness after hail, lightning strike or other accidental damages. Better GVI automation would decrease the downtime of the aircraft while improving safety and reliability, so that the inspector's expertise can be spent on evaluating the actual damage (detailed inspection) without having to search over the whole aircraft structure. The aim of this paper is to evaluate non-destructive inspection (NDI) technologies from the new perspective of their *capability to*

be automated. In this review, only non-contact methods for external NDI of the aircraft surface are included, as they are more suitable for GVI automation in the medium term. Without contact, an end effector composed of multiple sensors could move fast and freely around the aircraft, completing a scan within hours and potentially delivering a report with sufficient level of certainty.

The research methodology is an integrative literature review. Published literature reports a wide range of non-contact inspection technologies. This paper presents the basic working principles of the key methods, and the context in which they are used.

Pattern recognition is briefly illustrated in Sec. 2.2. However this technology cannot be applied to defects such as dents as they do not have sharp boundaries and require to catch the *third dimension*. Automated 3D measurement methods using passive and active triangulation are introduced in Sec. 2.3, while time of flight cameras are presented in Sec. 2.4. Different materials do not react equally to damage. A carbon fibre reinforced composite part, in fact, may show an almost invisible dent compared to an aluminium one exposed to the same impact and yet present a subsurface *delamination* [3], then, again, the judgment of the human inspector is critical. An attempt to detect subsurface damages can be pursued with thermography, as introduced in Sec. 2.5. Despite its pertinence to detect delamination, ultrasound scanning is not discussed here because it generally requires contact. As this review aims to address *in situ* aircraft inspection, methods that require highly controlled environment or complex optical structures [4] are omitted. Sec. 2.6 concludes with the applicability of each technology to different damage types, and the challenges to automate these methods for aircraft inspection.

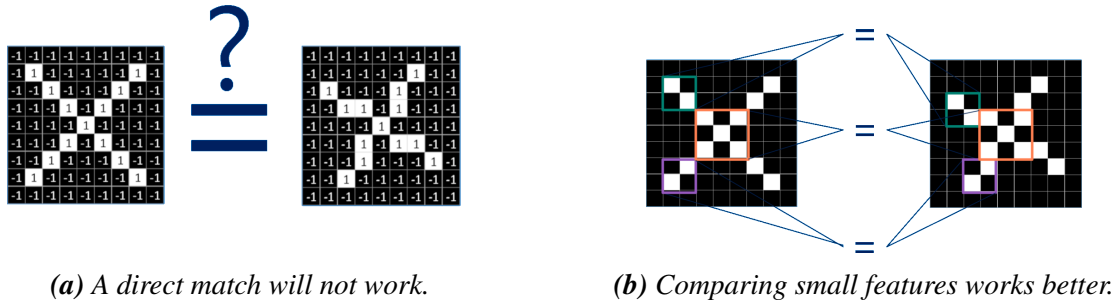
2.2 Pattern recognition

Pattern recognition is the process of classifying image data based on previously gathered knowledge or statistical information extracted from other images, usually using machine learning algorithms.

Pattern recognition has been widely applied in many fields, as well in aircraft inspections. Commercial drones equipped with a camera¹ are already available on the aircraft MRO market, thus allowing to scan the entire aircraft quickly. To some extent, the GVI task can be interpreted as a 2D image classification problem, looking for recurring im-

¹Usually with the addition of a light detection and ranging (LIDAR) sensor for collision avoidance and a positioning system relative to the aircraft itself.

age patterns typically found as positive flaws by the inspector, e.g. lightning strikes or scratches. Although there are several classification algorithms usable for this scope, the state-of-the-art performances are reached using convolutional neural networks (CNN) [5].

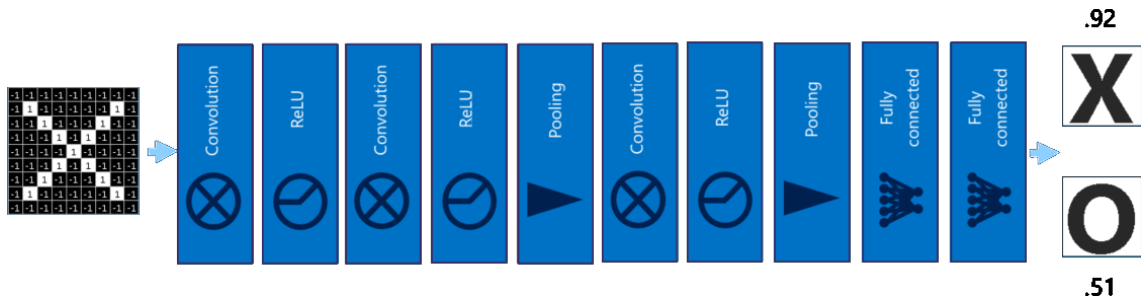


Credits: Brandon Rohrer; CC0 (public domain).

Figure 2.1: Basic CNN approach.

An artificial neural network is a mathematical model made by an ensemble of nodes called neurons and performing a function, interconnected by links called synapses, each one with its own weight that is multiplied with values flowing through the network. Neurons are organised in layers, with *hidden* layers between the first one (input) and the last one (output).

The CNN scheme is particularly adept at visual tasks since is unaffected by position shift and inspired by biological processes [6]. Its approach is to select *relevant* pieces from the original image called *features* (Fig. 2.1), which are smaller two-dimensional matrix that are most able to distinguish e.g. an “X” from a “O” [7]. The layers can be stacked multiple times in every order (Fig. 2.2) and the network may have very different configurations. As a rule of thumb, the more layers (so that the network is *deep*), the more sophisticated features of the image that can be understood. CNNs require a significant



Credits: Brandon Rohrer; CC0 (public domain).

Figure 2.2: An example of complete CNN layout with two outputs.

amount of *training* data available, i.e. “damage” and “non-damage” labeled data so that

the network can iteratively *learn* which features are more important to classify the input. This is done calculating the error at each iteration and doing the *backpropagation* to adapt weights accordingly.

2.2.1 Pros and cons

Using pattern recognition is a relatively cheap technique, with performances above 95% for certain types of flaws [5]. It can drastically reduce the time spent for a GVI and also be integrated with MRO software systems to quickly generate reports. Nevertheless, it cannot be considered a definitive solution because:

- It requires huge amounts of training data. In many situations such data is not easy to gather, for example because that kind of damage does not happen very often.
- It operates well only in diffused light conditions, while could easily fail in presence of non-uniform light.
- Missing the third dimension, it is incapable to detect dents and other shape deformations, while it works good for scratches or lightning strikes.
- The output is only a prediction that a part of an image belongs to classes of flaws within a certain confidence. E.g. an image could have an output of 0.70 in the class “scratch” and, at the same time, 0.40 in the class “lightning strike”. From these values one may infer that the image contains a scratch with a probability of 70% and a lightning strike with a probability of 40%, but in any case this is giving no *quantitative* evaluation of the damage. As a consequence, the inspector would probably need to go through many false positives.

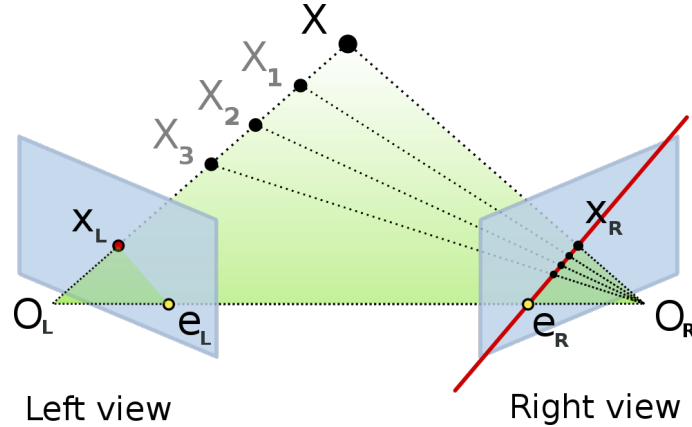
2.3 Triangulation

The basic principle to recover the depth z from images is to measure the triangle among the object located in an unknown world point X and a couple of known pixel coordinates, projections of the X point onto two planes. The two known planes are generally the image planes of two cameras or of a camera and a light source. Using triangulation without manipulating light sources is called *passive* triangulation [8], *active* triangulation otherwise. The nature of the problem between the two is different, because in passive

methods a correspondence needs to be established between features from two different camera images of the same scene, while in active methods the correspondence problem has already been solved using an artificial source of illumination [9].

Each camera is characterised by its *intrinsic* and *extrinsic* parameters. The former are related to the actual camera structure, like the focal length and the principal point, while the latter define the position and orientation of the camera in the 3D space with reference to an origin system. Both intrinsic and extrinsic parameters are given as matrices, output of a calibration procedure of the stereo system [10]. To correct lens distortion both radial and tangential distortion are often taken into account as coefficients and used to correct the image, so that stereo vision can make use of relatively cheap cameras too. While some algorithms can make use of uncalibrated cameras [11], the best accuracy generally requires a well calibrated stereo pair of cameras.

In a given stereo system the so-called *fundamental matrix* establishes a relation between a point X_L in one image and an *epipolar line* onto the other one. Thanks to the epipolar geometry, the *stereo correspondence* is a one-dimension problem that consists to find the point X_R corresponding to X_L on that epipolar line only (Fig. 2.3) without having to span the whole image.



Credits: Arne Nordmann, Wikipedia.org, CC BY-SA 3.0.

Figure 2.3: Given X_L as projection of the world point X on the left image, the search for point X_R , projection of the same point on the right image, is restricted to the epipolar line passing by e_R . If the image is rectified, this line becomes horizontal.

Since epipolar lines are generally oblique, after the acquisition of an image pair it is convenient to *rectify* the two images, i.e. transform the two images as if they were taken with only a horizontal displacement and projected to a common image plane parallel to the baseline. This reduces the matching problem on the x -axis only as the algorithm searches for the corresponding points on a horizontal line [12]. Although rectification is

not a necessary step, it is largely used to speed up the matching process.

The stereo correspondence problem is to match features between two images. Being able to identify the projection of the same point X in both the images, then, means to solve it, as the triangle is completely known. The cameras' focal length, the optical axis orientations and the position of the cameras O_L and O_R are all given by calibration, with $B = \overline{O_L O_R}$ as the baseline.

2.3.1 Passive triangulation

Photogrammetry is a passive triangulation method that involves using many perspective views of the same scene. More than two points of views can be useful in some situations, e.g. where horizontal obstacles are present [13]. *Stereo vision* is the sub-class of photogrammetry that only involves two points of view to extract information (Fig. 2.4).

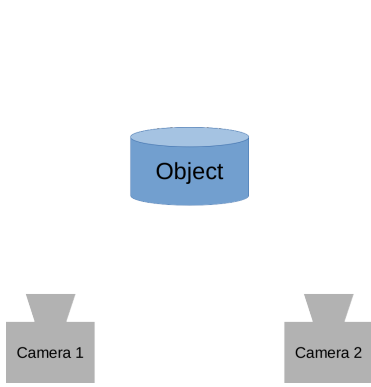


Figure 2.4: Passive stereo system elements.

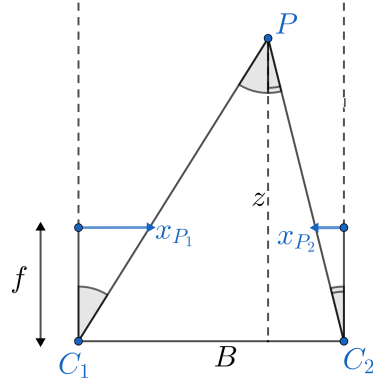


Figure 2.5: Recovering the depth information from disparity is based on similar triangles. The triangle PC_1C_2 is similar to the union of triangles built on the segments X_{P_1} and X_{P_2} .

On a rectified image pair, a point P_1 in the left image and the corresponding point P_2 in the right image are displaced by a number of pixels $d_P = |x_{P_1} - x_{P_2}|$, called *disparity*. The disparity is inversely proportional to the depth, so that it is larger on the foreground and smaller on the background. Disparity information is typically stored in a greyscale image called *disparity map*, where dark pixels encode low disparity (high distances from the camera) and vice versa. In a calibrated stereo system, the disparity map is sufficient to reconstruct a metric 3D model of the recorded scene [14, 12], achieving our final goal.

The depth z of the corresponding pixel point P can be calculated with:

$$z = \frac{f \cdot B}{d_p} \quad (2.1)$$

where f is the camera's focal length² and B is the stereo baseline (Fig. 2.5). Note that in case of cameras with different focal lengths, one can be rescaled to match the other one.

A naive approach to solve the matching problem would be, for every pixel P_1 of the left image, to search along the corresponding horizontal line in the right image, and then select the pixel P_2 whose colour is most similar to that of P_1 as a matching point. This approach leads to very noisy disparity maps, because of the ambiguity of the data (similar colours on different positions). Making the assumption that spatial neighbouring pixels are likely to have similar disparities helps to reduce noise. *Local* and *global* algorithms differentiate in how this assumption is made [15]. In general local methods match small image areas, while global methods try to define a disparity map of the whole image at the same time.

Local algorithms simply assume that all pixels within a certain window have exactly the same disparity. So, given a square window w centered on P_1 , a simplistic algorithm looks for P_2 into a window w' in the right image shifted by d along the x -axis, varying d so that the dissimilarity between w and w' is minimised:

$$d_p = \underset{0 \leq d \leq d_{max}}{\operatorname{argmin}} \sum_{x \in w} \operatorname{diff}(x, x - d) \quad (2.2)$$

where the function $\operatorname{diff}()$ calculates the dissimilarity between a pixel of the left and one of the right image. The difference could be, for example, the sum of absolute differences of RGB colour values. Here d_{max} defines the maximum allowed disparity [14].

Many complex local methods have been developed replacing the window definition and the cost function of Eq. (2.2) to obtain more accuracy and computational efficiency. Most of the local approaches are based on:

- Adaptive windows: the principle is to keep windows large but without overlapping disparity discontinuities, so to get clear edges.
- Weighted pixels: the key idea is to give a different weight to each pixel in influencing the matching process.

²Here the focal length is expressed in pixels, as the calibration process is proportional to the pixel size.

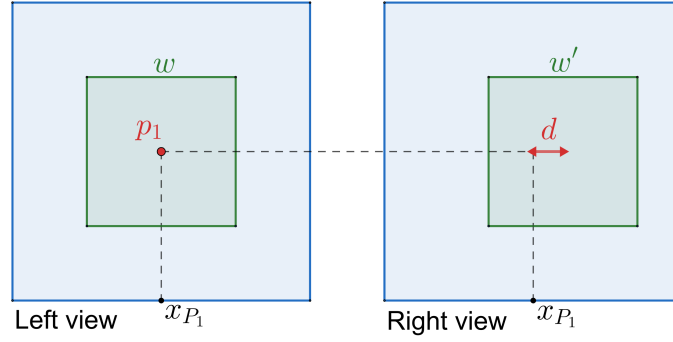


Figure 2.6: Local algorithms look for a window w shifted by d_p so that the overall colour differences between w and w' is minimised.

- Colour: takes into account adjacent colour areas which are likely to have the same disparity.

Nowadays local methods seem to be the most used and best performing, although the same window setting is not suitable for the whole image, e.g. both low textured areas and object borders, so low texture areas and/or edges may remain highly ambiguous.

Global algorithms, instead, introduce an *energy function* that gives a measure of the whole disparity map. The goal is to find the disparity map of lowest energy [14]. The usual form of the energy function is:

$$E(D) = E_{data}(D) + \lambda \cdot E_{smooth}(D) \quad (2.3)$$

where D is a disparity map, λ is a parameter that balances between E_{data} and E_{smooth} . The term E_{data} measures photo consistency, i.e. the absolute difference of intensity values between left image I_l and the corresponding right image I_r pixels:

$$E_{data}(D) = \sum_{p \in I_l} c(p, p - d_p) \quad (2.4)$$

with $c()$ usually being the sum of absolute or squared differences. The E_{smooth} term, instead, makes smooth disparity maps more preferable, assigning lower energy if values are spatially smooth:

$$E_{smooth}(D) = \sum_{p, q \in \eta} s(d_p, d_q) \quad (2.5)$$

where η denotes all pairs of spatially neighbouring pixels in the left image and the $s()$ function assigns a penalty if the disparity is different.

The target of global methods is to find a disparity map that minimises the equa-

tion (2.3), but finding the global minimum is, in general, an np-complete problem [16]. One way to reduce the complexity of the energy equation is to map neighbouring pixels into a tree on the image grid, so that η does not contain cycles and an optimal solution can be computed efficiently via dynamic programming. However, global methods also find their limits because of the energy formulation, which does not efficiently represent the stereo matching problem [14], despite further developments have been made towards better converging algorithms.

Pros and cons

Passive stereo vision (and photogrammetry in general) is an effective way to extract depth information without touching the object and requiring simple hardware. However in NDI it may not lead to an optimal choice for the following issues:

- Texture absence: in case the scene lacks of well-defined textures any matching algorithm will fail³ because no reliable data can be extracted from pixels and the depth remains ambiguous. Unfortunately this is a common case in aircraft inspections.
- Occlusion: not every pixel of the left image may find its correspondence into the right one and vice versa. In fact there may be areas visible from a camera and not visible from the other, so no match can be established.
- Calibration: the conversion from disparity measure to depth assumes a perfect calibrated system. In reality this is not to be taken for granted, especially if submillimetre accuracy is required.
- Distortions, noise, specular surfaces, transparent objects or repetitive patterns could cause matching algorithms to fail.

The correspondence problem has not been completely solved yet, but performances of the newest techniques are encouraging. Moreover, passive sensing keeps its importance where there are overriding circumstantial constraints that prevent the use of artificial illumination (e.g. automated cartography, aircraft navigation) [9] and also because it is much more cheaper compared to other depth estimation methods.

For a benchmark evaluation many of the passive stereo matching algorithms refer to a common evaluation dataset, e.g. KITTI⁴. To date the method with best performance is

³Note that also human eyes are deceived into perceiving the depth in front of a completely white wall.

⁴<http://www.cvlibs.net/datasets/kitti/>

“LEAstereo”, achieving a 1.65% of outliers⁵ over the all ground truth pixels of the test set of the 2015 KITTI dataset.

2.3.2 Active triangulation

Active triangulation introduces the control of light sources. In general the term *active stereo* refers to a configuration with two cameras and one light source (Fig. 2.7a), while *structured light* refers to a configuration with one camera and one light source (Fig. 2.7b). In both cases light is projected with a known pattern onto the object, so that the way how shapes are observed becomes a new source of information.

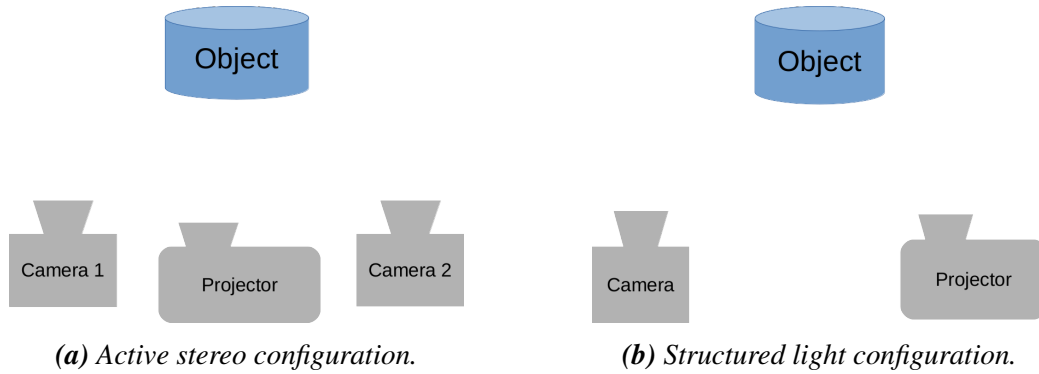


Figure 2.7: Common active triangulation configurations.

Active stereo is more common as it allows to do triangulation between the two cameras as in passive stereo, without the need to calibrate the projector. In fact, even if methods for projector calibration have been designed [17, 18], it remains a laborious operation as the projector is not capable to capture images, thus leading to lower overall accuracy. Anyway the methods discussed below can be found in both configurations.

The basic example of structured light are point gauges to make one dimension measurements [8]. A point gauge system is made by a camera and a light spot. The camera focal axis is directed to the object and the light axis is inclined by θ degrees (Fig. 2.8). So, given the light displacement $d = x_2 - x_1$ between two unknown positions, the depth difference can be easily calculated as:

$$z = \frac{d}{\tan \theta} \quad (2.6)$$

The measure can be converted to metric if the system is appropriately calibrated.

⁵With a certain tolerance. Of course thresholds for inspection applications would be more demanding.

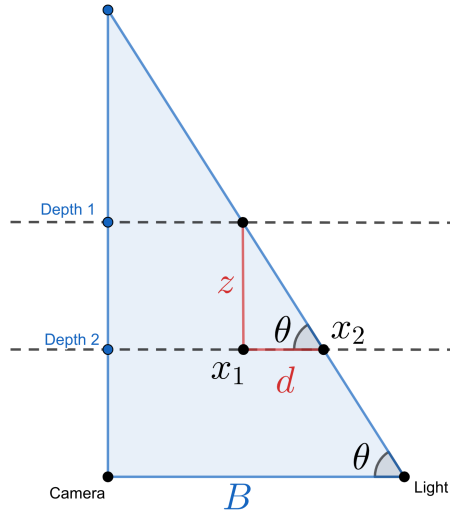


Figure 2.8: Point gauges make one-dimension measures based on similar triangles.

The working principle of structured light is the same of the point gauge but, to match multiple points simultaneously, complex patterns are projected instead of a single spot of light. A naive way to speed up the process would be to project n stripes. That will divide the required time by n but it will cause ambiguity, since the same pattern is repeated multiple times (Fig. 2.9).

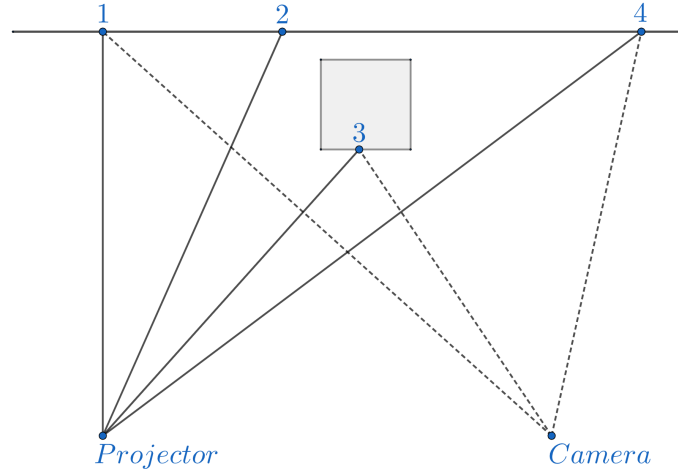


Figure 2.9: Projecting an identical stripe multiple times causes ambiguity from the camera perspective.

Approaches to resolve ambiguity follow *temporally* or *spatially* coding. An effective temporally coded method is *binary coding* [19]: since the correspondence problem is one-dimensional and simplified after rectification, the aim is to represent each vertical stripe on the projector with a unique code. The code is realised by projecting a sequence of black and white patterns, so that the camera knows exactly which vertical stripe the pixel belongs to. In order to do this, binary encoded stripes are projected sequentially

(Fig. 2.10a), such that the stripes have a crescent number from 0 to the maximum value. The resolution is limited to half the width of the smaller stripe.

Higher the number of unique codes, higher the accuracy of the calculated depth, up to physical limits. In fact, one of the problems is that transition pixels create uncertainty areas for the camera: stripes cannot be too thin as the camera would perceive blurred borders. A clever solution is then to minimise transitions in projected patterns, so the *Gray encoding* was developed (Fig. 2.10b). Gray code is less prone to errors due to the lower number of transition lines, still maintaining the same number of projected images and corresponding unique codes generated. Moreover, the width of the stripes is bigger for equal coding capacity, facilitating analysis on oblique surfaces and, since each neighbour code has only 1 bit variation, any wrong decoding introduces a misplacement of one unit only.

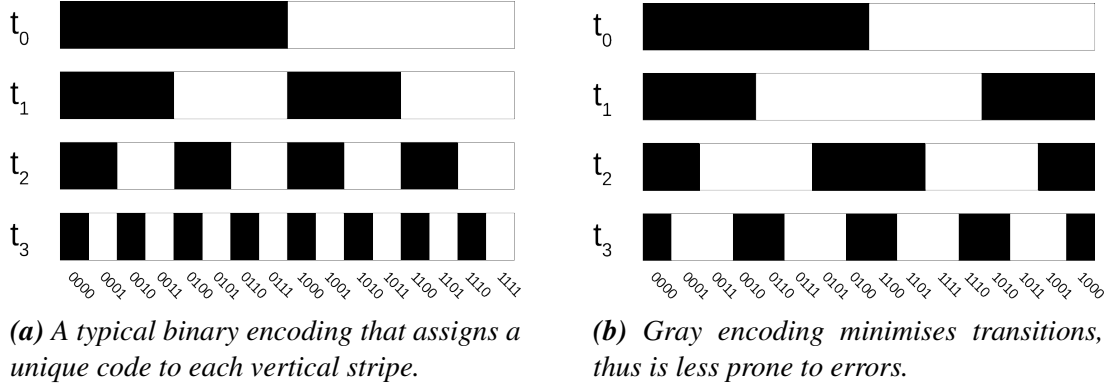


Figure 2.10: Common stripe encoding strategies.

These encodings are widely used in commercial products, despite they are affected by uncertainty and limited in resolution. Moreover, except for expensive hardware solutions, they are slow, e.g. for a 1920 pixel horizontal resolution $\log_2 1920 \cong 11$ pattern projections are required. The acquisition speed is then constrained by projector and camera frame rates, that need to be roughly synchronised. More importantly, the object must stay *still* during the whole process.

In addition, surface texture and reflections could disturb the camera reading. A way to reduce these errors is to project both the pattern and its inverse every time, assigning the unique code with more confidence after comparing on/off thresholds [20]. This approach, evidently, requires twice the time.

More advanced techniques tend to reduce the number of projected patterns by increasing the size of the alphabet: an alphabet of k symbols encodes k^m stripes, with m characters. With binary patterns $k = 2$ only (white and black), but k can be increased

using multiple colours or diversifying the number of intensity levels used to encode the stripes, although more similar the colour or intensity levels, higher the error rate, because diffusion and ambient light affects camera perception.

Pure spatially coding requires larger alphabet, so that a single projected pattern is enough to encode the image. A good example of it is the *De Bruijn sequence* method [21, 22]. Apart from using colours to increase k , it introduces the concept of using neighbour stripes to uniquely determine the associated code, so the algorithm is good at identifying each stripe uniquely, e.g. the couple red-green occurs only once (Fig. 2.11).



Figure 2.11: An example of De Bruijn sequence of order 2 over an alphabet of 4 symbols mapped to 4 colours with no repetitions. A sequence of 2 colours uniquely identifies a position.

The advantage of the De Bruijn sequence method is that it is a “one shot” method (single projected pattern) so it can be applied to moving objects, e.g. to get near real time 3D reconstruction of a moving scene. Apart from transition errors, the major drawback is that it requires many colours to develop a long sequence but a camera does not receive neat colour stripes as they were projected, instead they appear mixed due to the original colour of the object and light scattering.

Instead of remaining into discrete domain, *fringe analysis* introduces a continuous way to map the position [19] projecting sinusoidal structured patterns, called *fringe patterns*, onto the object. Then, recovering the phase from the camera perception image allows to map every point using a continuous function, hence with sub-pixel accuracy. Digital fringe projection (DFP) is the field that deals with the creation of these patterns.

Consider a single sinusoidal fringe pattern represented as:

$$I(x, y) = I'(x, y) + I''(x, y) \cdot \cos[\phi(x, y)] = I'(x, y) + I''(x, y) \cdot \frac{[e^{j\phi(x, y)} + e^{-j\phi(x, y)}]}{2} \quad (2.7)$$

where $I'(x, y)$ is the average or DC component of the signal, $I''(x, y)$ is the intensity modulation and $\phi(x, y)$ the phase captured in the image point (x, y) . Then, to obtain the phase, a high-pass filter can be applied into the frequency domain to filter out the conjugate and the DC component. The remaining signal is:

$$\tilde{I}(x, y) = I''(x, y) \cdot \frac{e^{j\phi(x, y)}}{2} \quad (2.8)$$

From which the phase can be calculated as:

$$\phi(x,y) = \arctan \left\{ \frac{\text{Im}[\tilde{I}(x,y)]}{\text{Re}[\tilde{I}(x,y)]} \right\} \quad (2.9)$$

The calculated phase ϕ ranges in $[-\pi, +\pi)$ as it is the output of the arctangent, so $\phi(x,y)$ it is often called *wrapped* phase map. Since a continuous phase is necessary for 3D model reconstruction, the following process of phase *unwrapping* is needed to remove discontinuities. Substantially, phase unwrapping consists in determining discontinuities location, find the fringe order (i.e. how many 2π to add) and then remove discontinuities [23].

Phase unwrapping algorithms can be *spatial* or *temporal*. Spatial methods analyse the wrapped phase map and determine fringe order for every position. These methods, in practice, always assume a smooth surface: the object geometry must not introduce sudden phase changes between two consecutive points, otherwise a reliable map cannot be built. Temporal methods, instead, require to project two or more sinusoidal patterns with different wavelengths. Two patterns would be sufficient in a noise-free system but, in practice, more the number of patterns used, more the resilience to noise.

All the light patterns can be projected using visible incoherent light, e.g. with a common video projector, infrared, UV light or with a *coherent* source of light (laser) that can be less affected by scattering. Lasers stay focused for longer distances, but typically, there is no way to modulate individual stripes, such as with Gray codes.

Line laser scan technology involves building up an image one line at a time. The line sensor is usually composed by a planar laser and a camera. It slowly moves over the object (or, vice versa, the object moves under the sensor) following known positions on the scanning line. For each position the triangulation is trivial as there will be only one point (along the baseline direction) corresponding to the projected laser line. So, if time is not an issue and the system is well calibrated, *spacetime* analysis [24] can be used. Often, as in many laser optics applications, the laser beam is assumed to have an irradiance profile that follows an ideal Gaussian distribution⁶ [25]. Given the above assumption, for each image the camera will perceive only a few points of a Gaussian curve. With these points the position of the middle point is exactly assigned and used for triangulation. The advantage of this method over the previous ones is that it leads to *sub-pixel* accuracy, however it requires a substantial amount of time during which the object

⁶Nevertheless, all actual laser beams will have some deviation from ideal Gaussian behavior.

must stay completely still [26].

Pros and cons

Theoretically, structured light solves the matching problem because pixels are identified with a unique mark, so that triangulation among object, camera and projector is easily applicable. In fact, if the projected pattern is dense enough, the depth resolution can be very accurate. Moreover, structured vision does not rely on object texture to solve the matching problem, so it is suitable also for smooth single-colour surfaces, as many aircraft surfaces are. The field of structured light has been widely developed and the resulting depth accuracy is generally high ($\sim 50\mu\text{m}$ on current commercial products), especially with the use of laser beams which are less sensitive to ambient light.

Problems that, nevertheless, continue to affect structured light are:

- Short distance: many commercial products require short distances ($\sim 30\text{ cm}$) to guarantee high accuracy.
- Reference targets: high precision devices need reference points, usually in the form of dot reflective stickers, applied to the inspected part.
- Occlusion: if any part of the object can be seen only from the camera and not from the projector, or vice versa, no depth can be calculated;
- Reflections: reflecting surfaces distort the pattern projection, thus preventing the identification of the unique code relative to involved pixels;
- Environment conditions: the use in uncontrolled environment (e.g. outdoor) could drastically reduce performance;
- Time: apart from one-shot methods and fast synchronised devices, it is generally limited to objects that remain static during scanning. For extremely precise systems, like laser line scanning, the time spent may increase significantly up to many hours [26].

Classic scanning systems oriented to manufacturing and reverse engineering usually operate in a controlled environment, with the object staying still and having relatively small dimensions. On the contrary, aircraft inspections are carried on in unconstrained

environments and over large surfaces, so the compromise between accuracy and speed may be difficult to achieve.

So far the problem has concerned the obtainment of depth of the scene and this is only the first step of a comprehensive flaw detection system. In fact the output of triangulation algorithms is a *point cloud*, that can be processed to build a 3D model of the object within a certain degree of confidence. But a subsequent process must be adopted to actually detect flaws and assess the damage. This may be difficult on articulated surfaces and may rely on a comparison with the expected surface model to reduce false positives. Furthermore, it may get computationally expensive as the resolution grows.

2.4 Time of flight

A *time of flight* (TOF) sensor is capable of measuring the light travel distance, through a certain medium, from the source to the object and back. The working principle is similar to a sonar: a wave is emitted and, after having been reflected on an object, it comes back so that given the total journey time Δt and knowing the speed of light in the medium v , the space measure is derived as $v\frac{\Delta t}{2}$. The term TOF camera may include a wide range of systems like LIDAR and motion detection systems, all based on the same principle.

Different methods exist to calculate the time interval Δt , here the followings are summarised:

1. timed pulses (or *direct*).
2. continuous modulation.
3. on/off loop.

The timed pulses method, in theory, would be to send pulses and measure the time interval until reflection. This naive approach requires very fast and expensive electronics, because a clock speed of around 300GHz would be needed to measure a time delay of about 3.33 ps at speed of light, corresponding to a travel distance of 1 mm. Since such frequencies are not easy to achieve in silicon nor cost effective, then the following alternative methods were developed.

Continuous modulation is the basis of most high accuracy TOF cameras currently on the market. It uses a sinusoidal modulated light source and measures the phase shift

of the reflected wave, since the speed in a medium is known very precisely in different conditions and the phase shift directly relates with time shift [27, 28]. The emitted signal frequency f to set depends on the maximum distance to measure λ , also considering the return journey [29]:

$$v = \lambda \cdot f \quad \Rightarrow \quad f = \frac{v}{\lambda} \quad (2.10)$$

where v is the speed on light in the medium. E.g. to measure objects within 7.5 m distance (hence, a $\lambda = 15$ m journey), a wave with $f = 20$ MHz could be used. The reflected light will still be sinusoidal with a different amplitude (due to attenuation and background light) and certain amount of phase shifting. There will be no phase shifting if the full journey is exactly a multiple of λ . A wave coming from farther away will cause ambiguity, that is why TOF cameras usually come with a maximum operating range (which would be 7.5 m with $f = 20$ MHz), unless some phase unwrapping algorithm is used. Note that to actually measure a wave with $f = 20$ MHz, the effective clock frequency of the measuring system speed must be at least $4f$ using the *four-quadrants* method to approximate the arctangent function.

The third method is to provide on/off loops transmitting a 50% duty cycle square wave (stroboscopic light source) and record the amount of returned light in a specific interval using CMOS imaging sensors, called *smart pixels*. These are particular LEDs having two capacitors per pixel [28, 30]. The capacitors are alternately connected to the sensor, using the same clock source of the stroboscopic light. This arrangement allows to directly relate the differential charge in the capacitors to the phase offset, then to the distance as seen above.

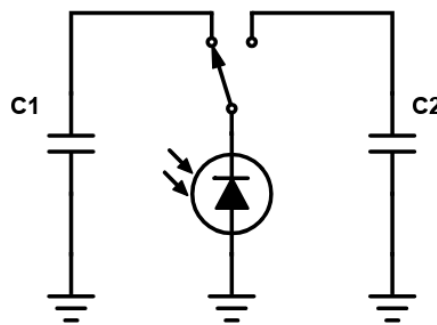


Figure 2.12: Schematics of a smart pixel. This arrangement allows to record the phase shift as electric charge difference.

2.4.1 Pros and cons

TOF cameras are a relatively new type of sensor, but their diffusion is rapidly increasing because of key advantages such as speed, compactness, absence of moving parts, low weight and power consumption. Nevertheless, they suffer of some unsolved issues:

- Limited resolution: to date, commercial TOF cameras have reached 640x480 pixel resolution.
- Internal issues such as imperfections in sensors and lens [31].
- Noise (e.g. interaction with other TOF systems) and background light (e.g. in outdoor environment) have to be filtered out, if possible.
- Ambiguity: the phase estimation is based on the reflected intensity of the signal that may vary. It is known that darker object cause an error that depends on the distance itself, non-linearly related [32].
- Reflections: the complexity of the scene causes double reflection (or *retroreflection*): light enters the sensor after bouncing on two (or more) surfaces instead than only one and mixes with the proper main reflection that becomes indistinguishable [31].

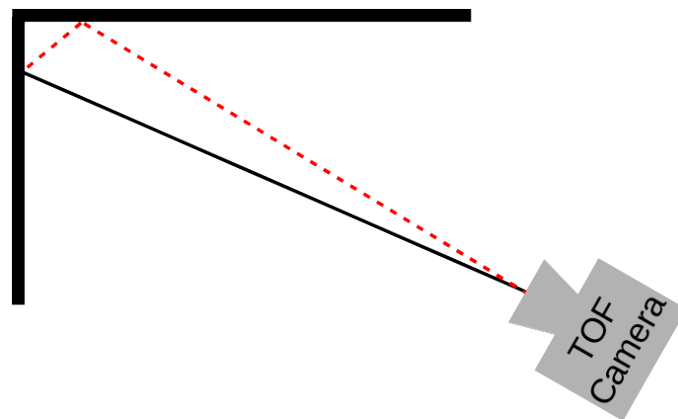


Figure 2.13: Time of flight double reflection happens when the emitted light (black) bounces two or more surfaces before entering the sensor (red, dotted).

2.5 Thermography

All the objects with a temperature above absolute zero emit radiations caused by the thermal motion of particles in the matter. An idealised opaque non-reflective body in thermal equilibrium, called *black body*, emits radiations at different frequencies, with the energy distribution described by Planck's law. At any given temperature T the distribution has its maximum for a specific wavelength λ_{max} , proportional to T . Wien's displacement law assigns the wavelength λ_{max} of the energy peak with:

$$\lambda_{max} = \frac{b}{T} \quad (2.11)$$

where $b = 2.897771955 \dots \times 10^{-3} \text{ m K}$ is the Wien's displacement constant and T is the temperature of the black body. For example, a black body at a temperature of 300 K has its peak of emitted energy at $\sim 9.6 \mu\text{m}$, while a temperature of 400 K corresponds to $\sim 7.2 \mu\text{m}$.

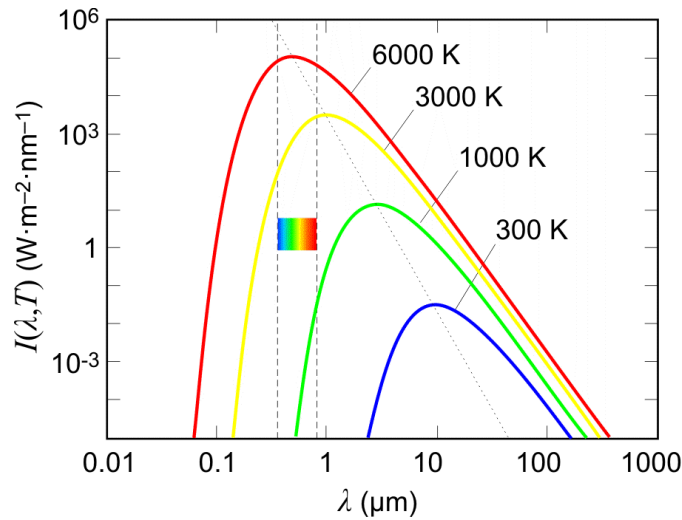


Figure 2.14: Planck's law describes the radiation of a black-body, calculating the curve for each different temperature.

Thermography allows to visualise temperature differences that reveal object structure and is generally a non-contact method, although some applications make use of direct contact with the object [33]. Thermographic cameras are usually made to perceive infrared wavelengths (range 700 nm – 1 mm), in which energy peaks fall more often. Commonly the measures are translated into a pseudo-colour image, called *thermogram*, to better express intensity changes.

The standard sensing structure is made by Focal Plane Arrays (FPA) that measure the

excitation generated by incident photons as change of electrical conductivity or by measuring a generated voltage. Some accurate but expensive technologies require cooling, while uncooled ones are cheaper and compact, hence more popular.

Thermography can be of two types: *active* and *passive* [34]. Passive thermography is limited to observe the object in ambient temperature conditions comparing with surrounding parts. Active thermography, instead, uses an external source to heat or cool the object to be inspected, thus allowing some characteristics to emerge more, also considering transient temperature measurements. Active techniques are the most used in NDI and include *pulse thermography* (PT), *step heating* (SH) and *lock-in thermography* (LT).

PT uses a very short pulse to heat up the inspected object and then observing its temperature decay curve. It only needs from a few microseconds to a few seconds of time, directly proportional to the thermal conductivity of the material inspected (e.g. metal parts are quicker). The presence of a defect reduces the diffusion rate so it appears as an area of higher temperature with respect to the surrounding area. PT is particularly indicated for near-surface defects [34]. SH, also called “long pulse” thermography, inverts the previous approach and, instead of the decay, it monitors the increase of temperature while continuously heating the surface with a low power source. The curve shape is directly related to specimen features. LT stimulates the inspected object with periodic (usually sinusoidal) waves. These cause attenuated, phase-shifted and dispersive *thermal waves* to run inside the material with a certain regime. If the input wave reaches non-homogeneous area (a defect), it is reflected. It is important to note that a defect may appear only if the right pulse ω is chosen. Advanced NDT methods make use of periodic energy stimulation to improve flaws detection [35, 36], even for military aircrafts [37].

2.5.1 Pros and cons

Thermography is very different from the other methods reported so far. In aerospace, it is indeed the most appropriate non-contact way to find structural and subsurface defects, being fast and with results relatively easy to interpret [34]. In particular PT thermography has a fast inspection rate (a few m^2 at a time) and it is safe for the operator as there are no harmful radiations, even if powerful flashes require protections. However thermography comes with some disadvantages such as:

- Difficulties to obtain a uniform thermal stimulation over large surfaces. Defects seen by thermal images usually have high variability [38] due to different environ-

mental temperature (e.g. outdoor), materials, perspective, etc.

- Effect of thermal losses which induces unwanted contrasts (noise).
- Ability to inspect materials of a limited thickness.
- In general, limited to defects that cause a measurable change of thermal properties (e.g. cracks are detected only if they induce thermal resistance) [34].

2.6 Conclusions

The aim of this paper is to consider the *suitability for automation* of aircraft inspections. The previous sections introduced the working principles of the major non-contact NDI methods. Their performance as reported in literature cannot be directly compared because of different equipment, conditions and inspected materials. A *qualitative* comparison is summed up (Fig. 2.15) with respect to speed, cost-effectiveness and ease of use and best use-case accuracy⁷.

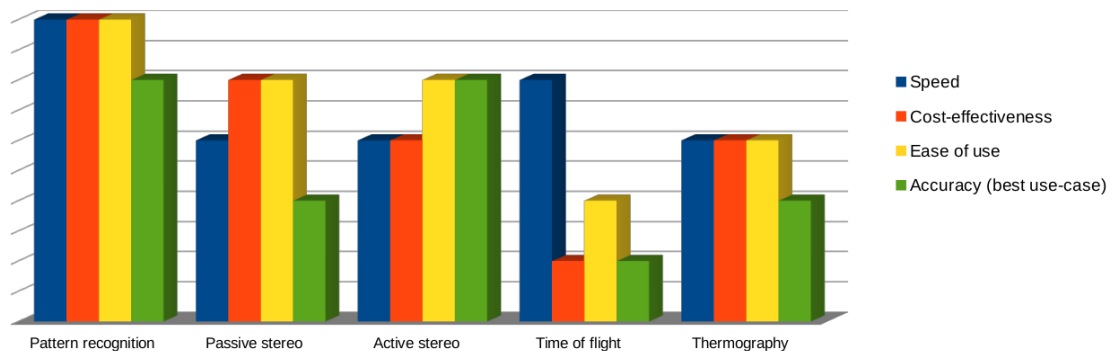


Figure 2.15: Indicative features of the discussed technologies.

Different inspection purposes benefit from different characteristics. In the case of aircraft NDI, accuracy is surely an important factor while higher costs could be overlooked. Apart from the use of high power lasers or some thermography applications, safety is generally not a concern for these NDI techniques. Nevertheless collisions may be an issue if drones are employed in a busy environment (e.g. a MRO hangar), although this is not related to the NDI technologies themselves.

Different NDI technologies work differently with respect to main damage types (Fig. 2.16).

⁷This is an indicative reference as levels may vary significantly among different methods of the same technology.

	Lightning	Scratches	Cracks	Dents	Delamination
Pattern recognition	Best	Good	Poor	N/A	N/A
Passive stereo	N/A	N/A	N/A	Average	N/A
Active stereo	N/A	N/A	Poor	Best	N/A
Time of flight	N/A	N/A	N/A	Average	N/A
Thermography	Average	Poor	Good	N/A	Good

Figure 2.16: Comparison of technologies with respect to types of damage.

Lightning strikes: Pattern recognition is particularly adequate for the identification of lightning strikes, assumed that there is a quality training set to train a CNN. Many commercial solutions make use of this technology, but relying on the final inspector say (also because of regulatory compliance). Pulse thermography is also effective on composites as lightning strikes deteriorate resin structure [38].

Scratches: Similarly to lightning, pattern recognition is a good choice for scratches, if neat contours are visible. Some thermography techniques allow to see differences in the coating of materials, so they are suitable and probably more accurate for this task, but slower and more expensive.

Cracks: A crack may or may not cause a change in thermal conductivity. If it does, thermography can identify it. A crack may also have a small 3D offset that an accurate laser-based active stereo may be able to detect, as successfully experimented for pipeline inspections [39].

Dents: Dents (and other shape deformations) are 3D damages that a single camera cannot see. A millimetre level accuracy or less is required but they come without neat contours and are difficult to be identified even by the human eye. The best option to detect them is to use active stereo [40]. The absence of texture prevents an effective use of passive stereo, while time of flight is still not achieving such levels of detail.

Delamination: Since it often has no noticeable anomaly on the outside, the only practical non-contact method to discover delamination would be thermography. However the results are extremely variable and there are no standard datasets available for training, so automatic detection remains a significant challenge.

In conclusion, aircraft NDI is an important but time-consuming process, a real bottleneck for airlines considering the high downtime costs. Introducing automation in inspections will significantly cut the costs and increase the reliability of inspections that still rely completely on human assessment. Although human inspectors have the best capabilities in understanding the context and interpret the output of different tools, they are also affected by errors and factors such as stress. Automation of current technology has many

challenges. A major hurdle is the lack of quality damages dataset on which algorithms can be trained to detect and assess damages. Substantial development is still needed to get highly automated damage detection techniques that perform better than humans. A combination of the above methods together with high volumes of quality data to train artificial intelligence algorithms is probably the best way forward for future aircraft NDI.

References

- [1] European Union Aviation Safety Agency. *Clarification of Definitions for General Visual (GVI), Detailed (DET), and Special Detailed (SDI) Inspections*. 2012. URL: <https://www.easa.europa.eu/sites/default/files/dfu/IP122%20Rev%202.pdf> (visited on 2019-10-22).
- [2] Lawrence Cook. “Visual inspection reliability for composite aircraft structures”. PhD thesis. 2009.
- [3] L. Cook, A. Boulic, D. Harris, P. Bellamy, and P. E. Irving. *Reliability of Damage Detection in Advanced Composite Aircraft Structures*. Civil Aviation Authority, 2013.
- [4] Feng Gao. “Interferometry for Online/In-Process Surface Inspection”. In: *Optical Interferometry*. 2017. DOI: 10.5772/66530.
- [5] Toubia Malekzadeh, Milad Abdollahzadeh, Hossein Nejati, and Ngai-Man Cheung. *Aircraft Fuselage Defect Detection using Deep Neural Networks*. 2017. URL: <https://arxiv.org/abs/1712.09213> (visited on 2019-10-22).
- [6] Kunihiko Fukushima. “Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position”. In: *Biological Cybernetics* (1980).
- [7] Brandon Rohrer. *How do Convolutional Neural Networks work?* Creative Commons CC0 1.0 Universal Licence. 2016. URL: http://brohrer.github.io/how_convolutional_neural_networks_work.html (visited on 2019-10-23).
- [8] Kevin G. Harding. *Practical Optical Dimensional Metrology*. Bellingham, Washington: SPIE, 2019.
- [9] R. Umesh Dhond and J. K. Aggarwal. “Structure from Stereo - A Review”. In: *Systems, Man, and Cybernetics, IEEE Transactions on* (1989).

- [10] Zhengyou Zhang. “A Flexible New Technique for Camera Calibration”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 22 (2000-12), pp. 1330–1334. DOI: 10.1109/34.888718.
- [11] M. Pilu. “Uncalibrated stereo correspondence by singular value decomposition”. In: *HP Laboratories Technical Report* (1997-01), pp. 3–11.
- [12] L. Nalpantidis, A. Amanatiadis, G. C. Sirakoulis, and A. Gasteratos. “Efficient hierarchical matching algorithm for processing uncalibrated stereo vision images and its hardware architecture”. In: *IET Image Processing* (2011). URL: <https://doi.org/10.1049/iet-ipr.2009.0262>.
- [13] Jan Kallwies and Hans-Joachim Wuensche. “Effective Combination of Vertical and Horizontal Stereo Vision”. In: *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2018-03, pp. 1992–2000. DOI: 10.1109/WACV.2018.00220.
- [14] Michael Bleyer and Christian Breiteneder. “Stereo Matching—State-of-the-Art and Research Challenges”. In: *Advanced Topics in Computer Vision*. Ed. by Giovanni Maria Farinella, Sebastiano Battiato, and Roberto Cipolla. London: Springer London, 2013, pp. 143–179. ISBN: 978-1-4471-5520-1. DOI: 10.1007/978-1-4471-5520-1_6. URL: https://doi.org/10.1007/978-1-4471-5520-1_6.
- [15] Daniel Scharstein and R. Szeliski. “A Taxonomy And Evaluation Of Dense Two-Frame Stereo Correspondence Algorithms”. In: *International Journal of Computer Vision - IJCV* 47 (2000-01), pp. 7–42.
- [16] Yuri Boykov, Olga Veksler, and Ramin Zabih. “Fast Approximate Energy Minimization via Graph Cuts”. In: *IEEE Trans. Pattern Anal. Mach. Intell.* (2001), pp. 1222–1239. ISSN: 0162-8828. URL: <https://doi.org/10.1109/34.969114>.
- [17] Song Zhang and Peisen Huang. “Novel method for structured light system calibration”. In: *Optical Engineering - OPT ENG* 45 (2006-08). DOI: 10.1117/1.2336196.
- [18] Xiaobo Chen, Juntong Xi, Ye Jin, and Jin Sun. “Accurate calibration for a camera–projector measurement system based on structured light projection”. In: *Optics and Lasers in Engineering* 47 (2009-03), pp. 310–319. DOI: 10.1016/j.optlaseng.2007.12.001.
- [19] Song Zang. “High-speed 3D shape measurement with structured light methods: A review”. In: *Optics and Lasers in Engineering* (2018). URL: <https://doi.org/10.1016/j.optlaseng.2018.02.017>.

- [20] Daniel Scharstein and Richard Szeliski. “High-Accuracy Stereo Depth Maps Using Structured Light”. In: *Proceedings of the CVPR 2003* (2003). URL: <https://doi.org/10.1109/CVPR.2003.1211354>.
- [21] Jordi Pages, Joaquim Salvi, and Josep Forest. “A new optimised De Bruijn coding strategy for structured light patterns”. In: (2004-08). DOI: 10.1109/ICPR.2004.70.
- [22] Joaquim Salvi, Sergio Fernandez, Tomislav Pribanic, and Xavier Llado. “A state of the art in structured light patterns for surface profilometry”. In: *Pattern Recognition* 43.8 (2010). ISSN: 0031-3203. URL: <https://doi.org/10.1016/j.patcog.2010.03.004>.
- [23] Song Zhang. “Absolute phase retrieval methods for digital fringe projection profilometry: A review”. In: *Opticas and Lasers in Engineering* 107 (2018).
- [24] Brian Curless and Marc Levoy. “Better Optical Triangulation through Spacetime Analysis”. In: *ICCV*. 1995.
- [25] Rüdiger Paschotta. *Encyclopedia of Laser Physics and Technology*. Wiley, 2008. ISBN: 978-3-527-40828-3.
- [26] Marc Levoy, Kari Pulli, Brian Curless, et al. “The Digital Michelangelo Project: 3D Scanning of Large Statues”. In: *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '00. USA: ACM Press/Addison-Wesley Publishing Co., 2000, pp. 131–144. ISBN: 1581132085. DOI: 10.1145/344779.344849. URL: <https://doi.org/10.1145/344779.344849>.
- [27] Robert Lange and Peter Seitz. “Solid-State Time-of-Flight Range Camera”. In: *IEEE Journal of Quantum Electronics* 37 (2001).
- [28] Bernhard Buettgen, Thierry Oggier, Michael Lehmann, Rolf Kaufmann, and Felix Lustenberger. “CCD/CMOS lock-in pixel for range imaging: Challenges, limitations and state-of-the-art”. In: (2005-01).
- [29] A Dorrington, Andrew Payne, and Michael Cree. “An evaluation of time-of-flight range cameras for close range metrology applications”. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII* (2010).
- [30] Mark Hughes. *How Do Time of Flight Sensors (ToF) Work?* 2019. URL: <https://www.allaboutcircuits.com/technical-articles/how-do-time-of-flight-sensors-work-pmdtechnologies-tof-3D-camera/> (visited on 2019-11-27).

- [31] Sigurjon Ami Gudmundsson, Henrik Aanaes, and Rasmus Larsen. “Environmental Effects on Measurement Uncertainties of Time-of-Flight Cameras”. In: *2007 International Symposium on Signals, Circuits and Systems*. Vol. 1. 2007, pp. 1–4. DOI: 10.1109/ISSCS.2007.4292664.
- [32] Dragos Falie and Vasile Buzuloiu. “Noise Characteristics of 3D Time-of-Flight Cameras”. In: *2007 International Symposium on Signals, Circuits and Systems*. 2007. DOI: 10.1109/ISSCS.2007.4292693.
- [33] Xavier P. Maldague. *Theory and Practice of Infrared Technology for Nondestructive Testing*. John-Wiley & Sons, 2001.
- [34] Xavier P. Maldague. “Introduction to NDT by active infrared thermography”. In: *Materials Evaluation* 60 (2002-09).
- [35] D. Wu, A Salerno, U Malter, et al. “Inspection of aircraft structural components using lockin-thermography”. In: *Quantitative Infrared Thermography* (1997-01).
- [36] D. Wu and G. Busse. “Lock-in thermography for nondestructive evaluation of materials”. In: *International Journal of Thermal Sciences* (1998).
- [37] Pierre Servais. “Development of a New NDT Method Using Thermography for Composite Inspection on Aircraft with Portable Military Thermal Imager”. In: 2006.
- [38] Robert J. Ducar. “Pulsed thermographic inspection and application in commercial aircraft repair”. In: *Thermosense XXI*. Ed. by Dennis H. LeMieux and John R. Snell Jr. Vol. 3700. International Society for Optics and Photonics. SPIE, 1999, pp. 77–83. DOI: 10.1117/12.342277.
- [39] Yang Zhongyuan, Lu Shaohui, Wu Ting, Yuan Gongping, and Tang Yiping. “Detection of morphology defects in pipeline based on 3D active stereo omnidirectional vision sensor”. In: *IET Image Processing* (2018). DOI: 10.1049/iet-ipr.2017.0616.
- [40] Priyan Gunatilake, Mel Siegel, Angel G. Jordan, and Gregg W. Podnar. “Image enhancement and understanding for remote visual inspection of aircraft surface”. In: *Nondestructive Evaluation of Aging Aircraft, Airports, and Aerospace Hardware*. Ed. by Raymond D. Rempt and Alfred L. Broz. Vol. 2945. International Society for Optics and Photonics. SPIE, 1996, pp. 416–427. DOI: 10.1117/12.259117. URL: <https://doi.org/10.1117/12.259117>.

Chapter 3

Rectifying Homographies for Stereo Vision: Analytical Solution for Minimal Distortion

Published as P. Lafiosca, M. Ceccaroni in: *Intelligent Computing*, 2022, doi:10.1007/978-3-031-10464-0_33.

Abstract

Stereo rectification is the determination of two image transformations (or *homographies*) that map *corresponding points* on the two images, projections of the same point in the 3D space, onto the same horizontal line in the transformed images. Rectification is used to simplify the subsequent *stereo correspondence* problem and speeding up the matching process. Rectifying transformations, in general, introduce perspective distortion on the obtained images, which shall be minimised to improve the accuracy of the following algorithm dealing with the stereo correspondence problem. The search for the optimal transformations is usually carried out relying on numerical optimisation. This work proposes a closed-form solution for the rectifying homographies that minimise perspective distortion. The experimental comparison confirms its capability to solve the convergence issues of the previous formulation. Its Python implementation is provided.

Keywords: Epipolar geometry; Image rectification, Stereo image processing; Stereo vision.

3.1 Introduction

Stereo vision gained a prominent role among Computer Vision technologies as it allows machines to perceive depth. Thanks to pinhole camera model, calibration and epipolar geometry, a considerable simplification can be performed before attempting to solve the stereo correspondence. Among these simplifications, rectification is almost always conducted to obtain horizontal and aligned epipolar lines, so that the following stereo matching algorithm can work along the x -axis only, with significant performance increase.

However, rectification transformations generally introduce distortion, which, in turn, impairs the performance of the following stereo matching algorithm. For this reason, minimal-distortion algorithms are proposed in the literature. In this paper the closed-form solution of the rectifying homographies that minimise perspective distortion is derived by means of the metric introduced by Loop and Zhang [1]. This is applied to a calibrated stereo rig, where all the intrinsic and extrinsic parameters are known.

The hereby presented formulation of the minimising solution is enabled by a new geometrical interpretation of the problem, which simplifies the distortion metric expression. The solution found is general and, therefore, valid for every relative position of a given stereo rig, even in extreme configurations and when the previous algorithm [1] fails to provide an initial guess for minimisation (see Sec. 3.5). Furthermore, the proposed method overcomes the need for optimisation libraries and is thus computationally more efficient by avoiding numerical minimisation loops.

3.1.1 Background

Emulating the human vision system, stereo vision derives 3D information by comparison of 2D digital images taken from two distinct camera positions, usually known from calibration [2]. Given a stereo image pair, the third coordinate of each object in the scene is extracted by first solving the *stereo correspondence* problem [3, 4], namely the problem of locating the projections of that object in the two images, and then applying *triangulation* [5] to recover the position of the object in the 3D world (refer to Fig. 2.3). Taking advantage of the geometric relations between the 3D points and their projections onto the 2D images, the so-called *epipolar constraint*, the stereo correspondence problem can be reduced to one dimension. This means that the search of each pair of corresponding points is thus carried on along one, usually oblique, *epipolar* line in each picture. To

further simplify the problem, *image rectifications* can be applied, transforming the stereo correspondence problem into a search along a horizontal line, with a consequent, significant improvement in efficiency.

Image rectifications are a family of 2D transformations that can be applied to a couple of non-coplanar stereo images to re-project them onto a *rectifying plane* such that, on the transformed images, corresponding points will lie on the same horizontal line. The drawback of rectification is that it introduces a certain amount of distortion in the resulting images, that decreases the accuracy of the subsequent stereo matching algorithms. Hence, the importance of identifying, among the rectifying transformations in the family, the ones that minimise distortion. Hereafter, in continuity with the previous algorithm [1], we will refer to *perspective distortion* simply as distortion.

This paper make the rectifications minimising distortion explicit, finding their analytic formulation via a geometrical derivation. The basic notation and metric are given in Sec. 3.2. Our novelties, the geometrical interpretation and derivation of the closed form solution, are reported in Sec. 3.3. The main steps of the algorithm are listed in Sec. 3.4, followed by a discussion and examples in Sec. 3.5. Conclusions follow in Sec. 3.6. The Python 3 implementation of the algorithm is also provided.

3.1.2 Previous work

Several algorithms for stereo image rectification are available in the literature. Among the mostly accepted and applied, the one by Fusiello et al. [6] essentially fixes the rectifying plane using the sole orientation of one of the two cameras. Despite being a simple and compact, this algorithm is based on an arbitrary non-optimal choice of one rectifying plane (see Sec. 3.5.1). This choice leads, for highly skew configurations of the cameras, to the introduction of a massive amount of distortion in the transformed images, with the consequence of worsening the performances of the algorithm dealing with the stereo correspondence.

The importance of minimising the distortion introduced by the rectifying transformations is highlighted by Loop and Zhang [1]. In their work, the authors decompose rectifying homographies as a combination of a *perspective*, a *similarity* and a *shearing* sub-transformations, and minimise the global distortion. A metric for measuring perspective distortion is thus presented. Finally, using an iterative, numerical procedure, the pair of homographies introducing the minimum perspective distortion is found. However, the

initial guess thereby suggested cannot always be found [7].

With the aim to minimise distortions at pixel level, a different distortion metric was proposed by Gluckman and Nayar [8], making use of the Jacobian of the transformation to track changes in local image areas. This method proposes an approximate solution or requires to follow a complex iterative procedure to obtain the optimal one.

A first-attempt of a closed-form solution was proposed by Sun [9]. Here the rectification transformations are readily found estimating the fundamental matrix. However, it leads to a particular solution without the criteria of reducing perspective distortion.

Lately the attention of researchers has been focused on strategies for rectifying *uncalibrated* stereo rigs [10, 11, 12, 13, 14]. These methods do not require a calibration process and try to estimate camera parameters using the scene itself. For example, a handy solution for uncalibrated dual lens cameras [12] relies on key-points that have to be matched in the images. This approach is based on the very limiting assumption of a *small-drift* between the cameras poses. Obviously this does not apply to general camera poses, like verging cameras, and calibration is still important to reach high accuracy and efficiency levels. Different approaches focus on rectifying fisheye images [15, 16, 17]. Uncalibrated and fisheye rectification algorithms are outside the scope of this article. To date, the vast majority of stereo vision systems employs calibration, that allows for a *metric* reconstruction with the subsequent triangulation.

3.2 Setting the problem

This Section summarises the work of Loop and Zhang [1] to provide the basis for the following closed-form solution.

3.2.1 Pinhole cameras and epipolar geometry

Let us consider a calibrated stereo rig composed of two *pinhole* cameras, where distortion caused by lens imperfections has been already corrected. As most definitions hereafter will be analogous for both cameras, we will define them only for one camera (definitions for the second camera will be obtained replacing the subscript with 2). Let $\mathbf{A}_1 \in \mathbb{R}^{3 \times 3}$ be the intrinsic matrix of the left camera, with $\mathbf{R}_1 \in \mathbb{R}^{3 \times 3}$ and $\mathbf{t}_1 \in \mathbb{R}^3$ its rotation matrix and translation vector, respectively, describing the position of the first *Camera Coordinate*

System (CCS) with respect to *World Coordinate System* (WCS), as represented in Fig. 3.1, with a slight abuse of notation for axes names to ease visualisation.

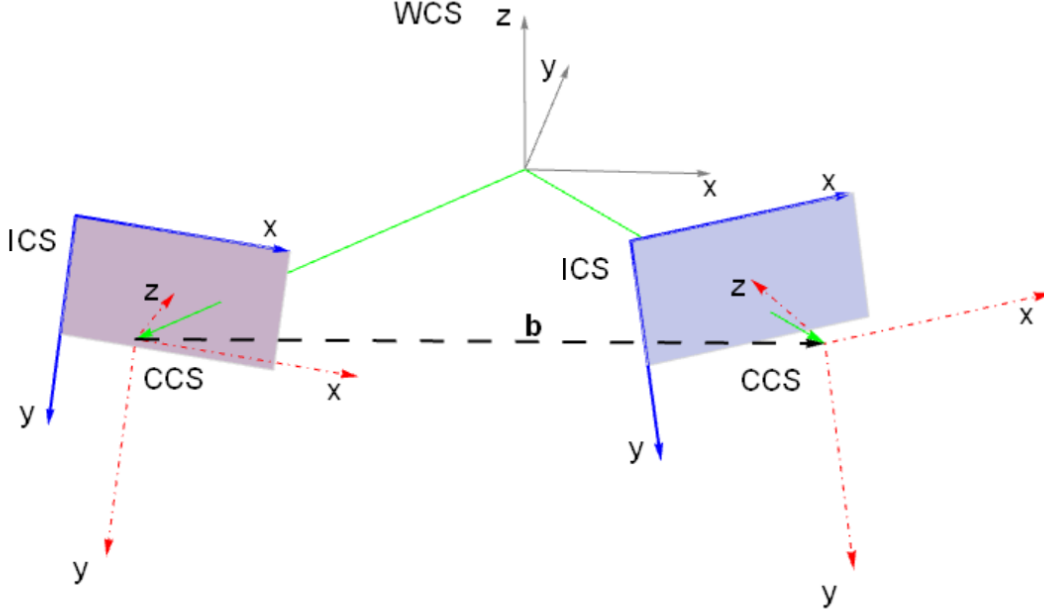


Figure 3.1: Representation of coordinate systems. CCSs are in red, dot-dashed. ICSs are in blue. WCS is in grey and the baseline is dashed black.

Call $\mathbf{o}_1 = -\mathbf{R}_1^{-1} \cdot \mathbf{t}_1$ the position of the optical center in the WCS. Hereafter, unless otherwise stated, elements are expressed in WCS. The *baseline* is the vector $\mathbf{b} = \mathbf{o}_2 - \mathbf{o}_1$ going from the first to the second camera center. Additionally, an *Image Coordinate System* (ICS) is defined on the image plane of each camera, where the left image I_1 forms (located at $z = 1$ in CCS) with the x and y axes parallel to the respective CCS ones and origin in the upper left corner of the image, corresponding to $\begin{bmatrix} -\frac{[\mathbf{A}_1]_{13}}{[\mathbf{A}_1]_{11}} & -\frac{[\mathbf{A}_1]_{23}}{[\mathbf{A}_1]_{22}} & 1 \end{bmatrix}$ in CCS. The notation $[\mathbf{A}_1]_{ij}$ indicates the (i, j) element of $\mathbf{A}_1$. The ICS system is shown in Fig. 3.1 as well.

Given two corresponding image points $\mathbf{p}_1 \in I_1$ and $\mathbf{p}_2 \in I_2$, each one in its ICS and expressed as homogeneous coordinates (i.e. $\mathbf{p}_1 \in \mathbb{R}^3$, with unitary 3rd coordinate), the epipolar geometry constraint is defined as:

$$\mathbf{p}_2^T \cdot \mathbf{F} \cdot \mathbf{p}_1 = 0 \quad (3.1)$$

where $\mathbf{F}$ is the fundamental matrix, a 3×3 matrix of rank 2 that can only be determined up to a scale factor [3, 18], assumed as known. $\mathbf{F}$ maps a point on I_1 to a *line* on I_2 , and

vice versa (using its transpose). Given that we work in a *projective space*, all the points are defined up to an arbitrary scaling factor.

On each image all the epipolar lines will intersect in a single point called *epipole*. Let $\mathbf{e}_1 \in I_1$ and $\mathbf{e}_2 \in I_2$ be the two epipoles in homogeneous ICS, they can be defined as the left and right *kernels* of $\mathbf{F}$, so that:

$$\mathbf{F} \cdot \mathbf{e}_1 = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T = \mathbf{F}^T \cdot \mathbf{e}_2 \quad (3.2)$$

Geometrically, $\mathbf{e}_1$ is the projection of $\mathbf{o}_2$ on the image I_1 . Similarly for $\mathbf{e}_2$. It is worth noticing that each pair of corresponding epipolar lines lies on a same plane together with the baseline.

3.2.2 Rectification

Rectification is the problem of determining two homographies that map corresponding epipolar lines onto parallel horizontal lines sharing the same y-coordinate (i.e. sending the epipoles to ∞). Thus, rectified images must have the new fundamental matrix in the form [3]:

$$\bar{\mathbf{F}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad (3.3)$$

Defining the rectified points $\bar{\mathbf{p}}_1 = \mathbf{H}_1 \cdot \mathbf{p}_1$ and $\bar{\mathbf{p}}_2 = \mathbf{H}_2 \cdot \mathbf{p}_2$, then, from Eq. (3.1), we get:

$$\begin{aligned} \mathbf{p}_2^T \cdot \mathbf{F} \cdot \mathbf{p}_1 &= \bar{\mathbf{p}}_2^T \cdot \mathbf{H}_2^{-T} \cdot \mathbf{F} \cdot \mathbf{H}_1^{-1} \cdot \bar{\mathbf{p}}_1 \\ &= \bar{\mathbf{p}}_2^T \cdot \bar{\mathbf{F}} \cdot \bar{\mathbf{p}}_1 \\ &= 0 \end{aligned} \quad (3.4)$$

where $\mathbf{H}_1$ and $\mathbf{H}_2$ are the left and right sought-after homographies, respectively.

3.2.3 Perspective distortion

Following the original procedure [1], given a generic homography $\mathbf{H}_1$, we scale it by its last element and decompose as:

$$\mathbf{H}_1 = \begin{bmatrix} u_{1a} & u_{1b} & u_{1c} \\ v_{1a} & v_{1b} & v_{1c} \\ w_{1a} & w_{1b} & 1 \end{bmatrix} = \mathbf{H}_{1a} \cdot \mathbf{H}_{1p} \quad (3.5)$$

where $\mathbf{H}_{1a}$ is an affine transformation and $\mathbf{H}_{1p}$ is a purely perspective transformation in the form:

$$\mathbf{H}_{1p} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ w_{1a} & w_{1b} & 1 \end{bmatrix} \quad (3.6)$$

$\mathbf{H}_{1p}$ is the sole responsible for introducing perspective distortion. The affine transformation will then be:

$$\mathbf{H}_{1a} = \begin{bmatrix} u_{1a} - u_{1c}w_{1a} & u_{1b} - u_{1c}w_{1b} & u_{1c} \\ v_{1a} - v_{1c}w_{1a} & v_{1b} - v_{1c}w_{1b} & v_{1c} \\ 0 & 0 & 1 \end{bmatrix} \quad (3.7)$$

$\mathbf{H}_{1a}$ can be decomposed further in a shearing transformation, followed by a similarity transformation [1]. The same decomposition is derived for $\mathbf{H}_2$.

Let $\mathbf{p}_1 = [x_1 \ y_1 \ 1]^T$ be a generic point on I_1 , in homogeneous ICS, then the perspective transformed point is:

$$\begin{aligned} \mathbf{p}'_1 &= \mathbf{H}_{1p} \cdot \mathbf{p}_1 \\ &= \begin{bmatrix} x_1 \\ y_1 \\ w_{1a}x_1 + w_{1b}y_1 + 1 \end{bmatrix} \\ &= \begin{bmatrix} x_1 \\ y_1 \\ \mathbf{w}_1^T \cdot \mathbf{p}_1 \end{bmatrix} \\ &\propto \begin{bmatrix} \frac{x_1}{\mathbf{w}_1^T \cdot \mathbf{p}_1} \\ \frac{y_1}{\mathbf{w}_1^T \cdot \mathbf{p}_1} \\ 1 \end{bmatrix} \end{aligned} \quad (3.8)$$

where $\mathbf{w}_1^T = [w_{1a} \ w_{1b} \ 1]$.

The codomain of I_1 through $\mathbf{H}_{1p}$ must be intended as including the point ∞ , as per the *hyperplane model* defined to handle perspective transformations [19].

Notice that if $w_{1a} = w_{1b} = 0$ there is no perspective component in the rectifying transformations, then $\mathbf{H}_1$ is a purely affine transformation. If this is true also for $\mathbf{H}_2$, then the stereo rig is perfectly frontoparallel and the image pair is already rectified. However this, in general, does not happen.

3.2.4 Distortion metric

We refer to the distortion metric introduced by Loop and Zhang [1]. The aim is to select $\mathbf{H}_1$ and $\mathbf{H}_2$ “as affine as possible”, meaning that the elements w_{1a} , w_{1b} , w_{2a} and w_{2b} should be chosen so to introduce less distortion.

Taking the average point $\mathbf{p}_{1c} = \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{p}_{1i}$ as reference, with n_1 total number of pixels of I_1 , the amount of distortion on I_1 is defined as:

$$\sum_{i=1}^{n_1} \frac{\mathbf{w}_1^T \cdot (\mathbf{p}_{1i} - \mathbf{p}_{1c})}{\mathbf{w}_1^T \cdot \mathbf{p}_{1c}} \quad (3.9)$$

The goal is to find the global minimum of the sum of the distortion of both images:

$$\sum_{i=1}^{n_1} \frac{\mathbf{w}_1^T \cdot (\mathbf{p}_{1i} - \mathbf{p}_{1c})}{\mathbf{w}_1^T \cdot \mathbf{p}_{1c}} + \sum_{i=1}^{n_2} \frac{\mathbf{w}_2^T \cdot (\mathbf{p}_{2i} - \mathbf{p}_{2c})}{\mathbf{w}_2^T \cdot \mathbf{p}_{2c}} \quad (3.10)$$

Then (3.10) can be rewritten in matrix form as:

$$\frac{\mathbf{w}_1^T \cdot \mathbf{P}_1 \cdot \mathbf{P}_1^T \cdot \mathbf{w}_1}{\mathbf{w}_1^T \cdot \mathbf{P}_{c1} \cdot \mathbf{P}_{c1}^T \cdot \mathbf{w}_1} + \frac{\mathbf{w}_2^T \cdot \mathbf{P}_2 \cdot \mathbf{P}_2^T \cdot \mathbf{w}_2}{\mathbf{w}_2^T \cdot \mathbf{P}_{c2} \cdot \mathbf{P}_{c2}^T \cdot \mathbf{w}_2} \quad (3.11)$$

where:

$$\mathbf{P}_1 \cdot \mathbf{P}_1^T = \frac{wh}{12} \begin{bmatrix} w^2 - 1 & 0 & 0 \\ 0 & h^2 - 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (3.12)$$

and:

$$\mathbf{P}_{c1} \cdot \mathbf{P}_{c1}^T = \begin{bmatrix} \frac{(w-1)^2}{4} & \frac{(w-1)(h-1)}{4} & \frac{(w-1)}{2} \\ \frac{(w-1)(h-1)}{4} & \frac{(h-1)^2}{4} & \frac{(h-1)}{2} \\ \frac{(w-1)}{2} & \frac{(h-1)}{2} & 1 \end{bmatrix} \quad (3.13)$$

where w and h are pixel width and height of I_1 . Similarly applies to I_2 .

3.3 Geometric interpretation and analytical derivation of the minimum

A geometric interpretation of the family of rectifying homographies is introduced hereafter, which paves the way to find the analytical formulation for the global minimum of Eq. 3.10 and build the corresponding rectifying homographies that minimise perspective distortion.

3.3.1 Geometric interpretation

It is known that finding rectifying transformations can alternatively be seen as determining a new *common orientation* for a pair of novel *virtual* cameras projecting the images on the same principal plane, such that epipolar lines become horizontal (Fig. 3.2).

In this subsection we will demonstrate that this can only be achieved if the $\hat{x}$ -axis of the new *common orientation* is chosen parallel to the baseline $\mathbf{b}$, while the $\hat{z}$ -axis can be arbitrarily chosen, thus generating the full rectifying family.

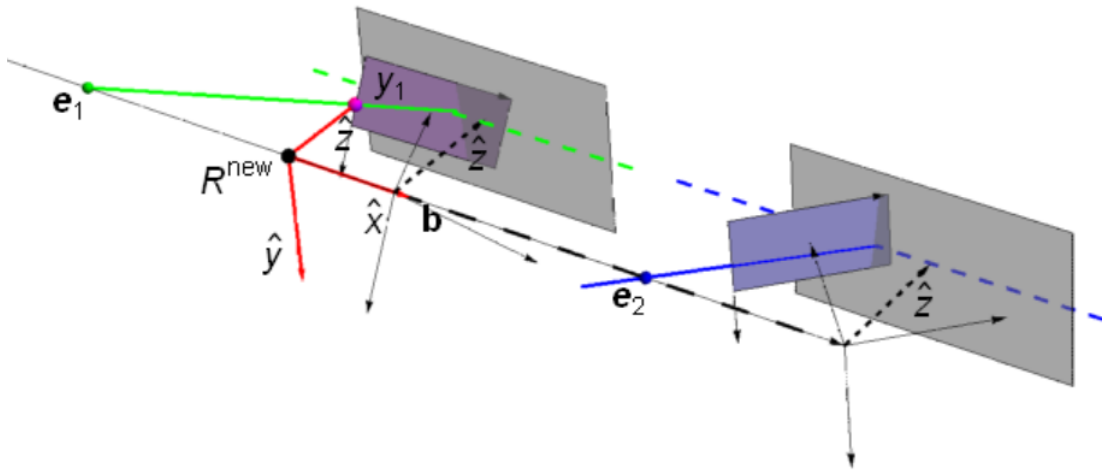


Figure 3.2: Common orientation of the virtual camera pair (red), projecting on a common plane (gray). $\hat{x}$ is parallel to the baseline $\mathbf{b}$ (black, dashed). The corresponding epipolar lines (blue, green) and $\hat{z}$ are identified by y_1 (magenta). Rectified epipolar lines are green and blue, dashed.

The new common orientation will thus be defined as:

$$\mathbf{R}^{new} = \begin{bmatrix} \hat{\mathbf{x}}^T \\ \hat{\mathbf{y}}^T \\ \hat{\mathbf{z}}^T \end{bmatrix} \quad (3.14)$$

where $\hat{\mathbf{x}} = \frac{\mathbf{b}}{\|\mathbf{b}\|}$ (the versor of the baseline) while $\hat{\mathbf{z}}$ and $\hat{\mathbf{y}}$ can be chosen accordingly to form a Cartesian coordinate system.

Each homography will have to cancel the corresponding intrinsic camera matrix and camera rotation, re-orient the camera with the new chosen orientation and apply another *affine* transformation, namely:

$$\begin{aligned} \mathbf{H}_1 &= \mathbf{K}_1 \cdot \mathbf{R}^{new} \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \\ \mathbf{H}_2 &= \mathbf{K}_2 \cdot \mathbf{R}^{new} \cdot (\mathbf{A}_2 \cdot \mathbf{R}_2)^{-1} \end{aligned} \quad (3.15)$$

where $\mathbf{K}_1, \mathbf{K}_2 \in \mathbb{R}^{3 \times 3}$ are arbitrary *affine* transformations.

The geometric interpretation above is based on the following Theorem.

Theorem 1 *The epipolar lines of two cameras are corresponding horizontal lines if and only if the two cameras share the same orientation, with the $\hat{\mathbf{x}}$ -axis parallel to their baseline.*

To prove this theorem we first introduce two Lemmas.

Theorem 1 *The fundamental matrix $\mathbf{F}$, with an abuse of notation, can be written as:*

$$\mathbf{F} \propto \mathbf{G} \times \mathcal{H} = \begin{bmatrix} [\mathbf{G}]_2 [\mathcal{H}]_{3*} - [\mathbf{G}]_3 [\mathcal{H}]_{2*} \\ [\mathbf{G}]_3 [\mathcal{H}]_{1*} - [\mathbf{G}]_1 [\mathcal{H}]_{3*} \\ [\mathbf{G}]_1 [\mathcal{H}]_{2*} - [\mathbf{G}]_2 [\mathcal{H}]_{1*} \end{bmatrix} \quad (3.16)$$

where $\mathbf{G} = \mathbf{A}_2 \cdot \mathbf{R}_2 \cdot \mathbf{b} \in \mathbb{R}^3$ and $\mathcal{H} = \mathbf{A}_2 \cdot \mathbf{R}_2 \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \in \mathbb{R}^{3 \times 3}$. The i^{th} row of a matrix is denoted as $[\]_{i*}$.

Proof (Lemma 1): let be $\mathbf{X} \in \mathbb{R}^3$ a point in WCS and $\mathbf{x}_1 \in I_1$ and $\mathbf{x}_2 \in I_2$ its projections in ICS, then $\mathbf{X}$ is a non-trivial solution of the linear system [20]:

$$\begin{cases} \mathbf{A}_1 \cdot (\mathbf{R}_1 \cdot \mathbf{X} + \mathbf{t}_1) = \lambda \mathbf{x}_1 \\ \mathbf{A}_2 \cdot (\mathbf{R}_2 \cdot \mathbf{X} + \mathbf{t}_2) = \mu \mathbf{x}_2 \end{cases} \quad (3.17)$$

where $\lambda, \mu \in \mathbb{R}$. Setting $\tilde{\mathbf{X}} = \mathbf{A}_1 \cdot (\mathbf{R}_1 \cdot \mathbf{X} + \mathbf{t}_1)$, yields:

$$\begin{cases} \mathbb{I}_3 \cdot \tilde{\mathbf{X}} &= \lambda \mathbf{x}_1 \\ \mathcal{H} \cdot \tilde{\mathbf{X}} - \mathbf{G} &= \mu \mathbf{x}_2 \end{cases} \quad (3.18)$$

where $\mathbb{I}_3 \in \mathbb{R}^{3 \times 3}$ is the identity matrix. Therefore:

$$\text{Det} \left(\begin{bmatrix} \mathbb{I}_3 & \mathbf{0} & \mathbf{x}_1 & \mathbf{0} \\ \mathcal{H} & \mathbf{G} & \mathbf{0} & \mathbf{x}_2 \end{bmatrix} \right) = 0 \quad (3.19)$$

which is linear in both $\mathbf{x}_1$ and $\mathbf{x}_2$ and can thus be rewritten as:

$$\mathbf{x}_2^T \cdot \mathbf{F} \cdot \mathbf{x}_1 = 0 \quad (3.20)$$

with:

$$[\mathbf{F}]_{ij} = (-1)^{i+j} \text{Det}(\hat{\mathcal{Q}}^{i,j}), \quad \forall i, j = 1, 2, 3. \quad (3.21)$$

where $\hat{\mathcal{Q}}^{i,j} \in \mathbb{R}^{4 \times 4}$ is equal to the matrix $\mathcal{Q} = \begin{bmatrix} \mathbb{I}_3 & \mathbf{0} \\ \mathcal{H} & \mathbf{G} \end{bmatrix} \in \mathbb{R}^{6 \times 6}$ with the i^{th} and $(3+j)^{th}$ rows dropped. Furthermore:

$$[\mathbf{F}]_{ij} = (-1)^{i+1} \text{Det}(\bar{\mathcal{J}}^{i,j}) \quad (3.22)$$

with the first column of $\bar{\mathcal{J}}^{i,j} \in \mathbb{R}^{2 \times 2}$ equal to the j^{th} column of the matrix $\mathcal{H}$ with the i^{th} row dropped, and the second column equal to $\mathbf{G}$ with the i^{th} element dropped. Then Eq. (3.22) is equivalent to Eq. (3.16) $\square$.

Remark: it can be shown that the matrix $\mathcal{H}$ is the matrix representing the axes of the first CCS as seen from the second CCS and that $\mathbf{e}_2 = \frac{\mathbf{G}}{\|\mathbf{G}\|}$. Therefore the fundamental matrix can be expressed as $\mathbf{F} \propto \mathbf{e}_2 \times \mathcal{H}$, similarly to Eq. (3.16).

Theorem 2 *The epipolar lines of the second camera are horizontal if and only if the x-axis of the second camera is parallel to the baseline.*

Proof (Lemma 2): by the definition of the fundamental matrix (see Sec. 3.2.1), the epipolar lines of the second camera are horizontal *if and only if* the first row of the fundamental matrix $\mathbf{F}$ is null. Using *Lemma 1* the first row of $\mathbf{F}$ can be expressed as:

$$[\mathbf{G}]_2 [\mathcal{H}]_{3*} - [\mathbf{G}]_3 [\mathcal{H}]_{2*} \quad (3.23)$$

From the remark above, the rows of $\mathcal{H}$ form a base of linearly independent vectors,

therefore the only linear combination that can set the formula (3.23) to zero is the trivial one, with $[\mathbf{G}]_2 = [\mathbf{G}]_3 = 0$, that is equivalent to state that the x -axis of the second camera is parallel to the baseline $\square$.

We can now demonstrate *Theorem 1*.

Proof (Theorem 1): by *Lemma 1* and *Lemma 2*, if the epipolar lines of the second camera are horizontal, the fundamental matrix can be reduced to the form:

$$\bar{\mathbf{F}} \propto \begin{bmatrix} \mathbf{0} \\ -[\mathcal{H}]_{3*} \\ [\mathcal{H}]_{2*} \end{bmatrix} \quad (3.24)$$

Using Eq. (3.3) we impose:

$$\begin{aligned} [\mathcal{H}]_{2*} &= \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \\ [\mathcal{H}]_{3*} &= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad (3.25)$$

That is equivalent to require that the y and z axes of the first CCS, as seen from the second CCS system of reference (respectively $[\mathcal{H}]_{2*}$ and $[\mathcal{H}]_{3*}$), must be parallel to the corresponding axes of the second camera $\square$.

3.3.2 Analytic derivation of the minimising rectification

By Eqs. (3.14) and (3.15), finding the global minimum of Eq. (3.10) is equivalent to find the optimal common orientation $\mathbf{R}^{new}$ of the virtual cameras. Since $\hat{\mathbf{x}}$ is imposed by the baseline, the problem is reduced to finding the $\hat{\mathbf{z}}$ so that distortion is minimised. Setting $\hat{\mathbf{y}} = \hat{\mathbf{z}} \times \hat{\mathbf{x}}$, will then determine $\mathbf{R}^{new}$.

Choosing the minimising $\hat{\mathbf{z}}$, in turn, is equivalent, by construction, to finding the pair of corresponding epipolar lines to become the new *horizon lines* (see Fig. 3.2), which will lay at the intersection of the plane $\hat{\mathbf{y}} = 0$ and the rectified images. Actually only one *horizon line* must be determined, as the corresponding one is readily found by means of $\mathbf{F}$.

Finally, the minimisation problem is reduced to a single parameter problem, noticing that an epipolar line is uniquely determined by its y -intercept $(0, y_1) \in I_1$ in ICS. Such

y-intercept, in WCS, takes the form:

$$\mathbf{y}_1 = \mathbf{R}_1^{-1} \cdot (\mathbf{A}_1^{-1} \cdot \begin{bmatrix} 0 & y_1 & 1 \end{bmatrix}^T - \mathbf{t}_1) \quad (3.26)$$

The $\hat{\mathbf{z}}$ axes of the virtual cameras will thus be the direction of the line $\mathbf{z}$ perpendicular to $\mathbf{b}$, passing through $\mathbf{y}_1$, calculated as the difference between the vector going from $\mathbf{o}_1$ to $\mathbf{y}_1$ and its projection on the baseline (its derivation is done by means of the outer product $\otimes$ [21]):

$$\begin{aligned} \mathbf{z} &= (\mathbf{y}_1 - \mathbf{o}_1) - [(\mathbf{y}_1 - \mathbf{o}_1)^T \cdot \hat{\mathbf{x}}] \hat{\mathbf{x}} \\ &= (\mathbf{y}_1 - \mathbf{o}_1) - \hat{\mathbf{x}} \otimes \hat{\mathbf{x}} \cdot (\mathbf{y}_1 - \mathbf{o}_1) \end{aligned} \quad (3.27)$$

Then follows:

$$\hat{\mathbf{z}} = \frac{\mathbf{z}}{\|\mathbf{z}\|} \quad (3.28)$$

Using Eqs. (3.14), (3.15) and (3.27), yields:

$$\begin{aligned} \mathbf{w}_1^T &= [\mathbf{H}_1]_{3*} \\ &= [\mathbf{K}_1 \cdot \mathbf{R}^{new} \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1}]_{3*} \\ &= \hat{\mathbf{z}}^T \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \\ &= \frac{\mathbf{z}^T}{\|\mathbf{z}\|} \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \\ &= \left(\mathbf{R}_1^{-1} \cdot \mathbf{A}_1^{-1} \cdot \begin{bmatrix} 0 \\ y_1 \\ 1 \end{bmatrix} - \hat{\mathbf{x}} \otimes \hat{\mathbf{x}} \cdot \mathbf{R}_1^{-1} \cdot \mathbf{A}_1^{-1} \cdot \begin{bmatrix} 0 \\ y_1 \\ 1 \end{bmatrix} \right)^T \\ &\quad \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \\ &= \begin{bmatrix} 0 & y_1 & 1 \end{bmatrix} \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-T} \cdot (\mathbb{I}_3 - \hat{\mathbf{x}} \otimes \hat{\mathbf{x}}) \\ &\quad \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \end{aligned} \quad (3.29)$$

where $\mathbf{K}_1$ has been omitted as it does not affect the 3rd row of what follows (being an affine transformation, its last row is $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$), and $\|\mathbf{z}\|$ has been discarded noticing that

in Eq. (3.11) it appears both at numerator and denominator. Rearranging:

$$\begin{aligned} \mathbf{w}_1 &= (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-T} \cdot (\mathbb{I}_3 - \hat{\mathbf{x}} \otimes \hat{\mathbf{x}}) \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \cdot \begin{bmatrix} 0 \\ y_1 \\ 1 \end{bmatrix} \\ &= \mathbf{L}_1 \cdot \begin{bmatrix} 0 \\ y_1 \\ 1 \end{bmatrix} \end{aligned} \quad (3.30)$$

where the matrix:

$$\mathbf{L}_1 = (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-T} \cdot (\mathbb{I}_3 - \hat{\mathbf{x}} \otimes \hat{\mathbf{x}}) \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \quad (3.31)$$

Similarly for $\mathbf{w}_2$:

$$\mathbf{w}_2 = \mathbf{L}_2 \cdot \begin{bmatrix} 0 \\ y_1 \\ 1 \end{bmatrix} \quad (3.32)$$

where:

$$\mathbf{L}_2 = (\mathbf{A}_2 \cdot \mathbf{R}_2)^{-T} \cdot (\mathbb{I}_3 - \hat{\mathbf{x}} \otimes \hat{\mathbf{x}}) \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \quad (3.33)$$

Our goal is to minimise the distortion from Eq. (3.11), that can now be written as:

$$\begin{aligned} & \frac{\begin{bmatrix} 0 & y_1 & 1 \end{bmatrix} \cdot \mathbf{M}_1 \cdot \begin{bmatrix} 0 & y_1 & 1 \end{bmatrix}^T}{\begin{bmatrix} 0 & y_1 & 1 \end{bmatrix} \cdot \mathbf{C}_1 \cdot \begin{bmatrix} 0 & y_1 & 1 \end{bmatrix}^T} \\ & + \frac{\begin{bmatrix} 0 & y_1 & 1 \end{bmatrix} \cdot \mathbf{M}_2 \cdot \begin{bmatrix} 0 & y_1 & 1 \end{bmatrix}^T}{\begin{bmatrix} 0 & y_1 & 1 \end{bmatrix} \cdot \mathbf{C}_2 \cdot \begin{bmatrix} 0 & y_1 & 1 \end{bmatrix}^T} \end{aligned} \quad (3.34)$$

where:

$$\begin{aligned} \mathbf{M}_1 &= \mathbf{L}_1^T \cdot \mathbf{P}_1 \cdot \mathbf{P}_1^T \cdot \mathbf{L}_1 \\ \mathbf{M}_2 &= \mathbf{L}_2^T \cdot \mathbf{P}_2 \cdot \mathbf{P}_2^T \cdot \mathbf{L}_2 \\ \mathbf{C}_1 &= \mathbf{L}_1^T \cdot \mathbf{P}_{c1} \cdot \mathbf{P}_{c1}^T \cdot \mathbf{L}_1 \\ \mathbf{C}_2 &= \mathbf{L}_2^T \cdot \mathbf{P}_{c2} \cdot \mathbf{P}_{c2}^T \cdot \mathbf{L}_2 \end{aligned} \quad (3.35)$$

The terms of Eq. (3.34) can be now expanded as polynomial expressions in y_1 . It can

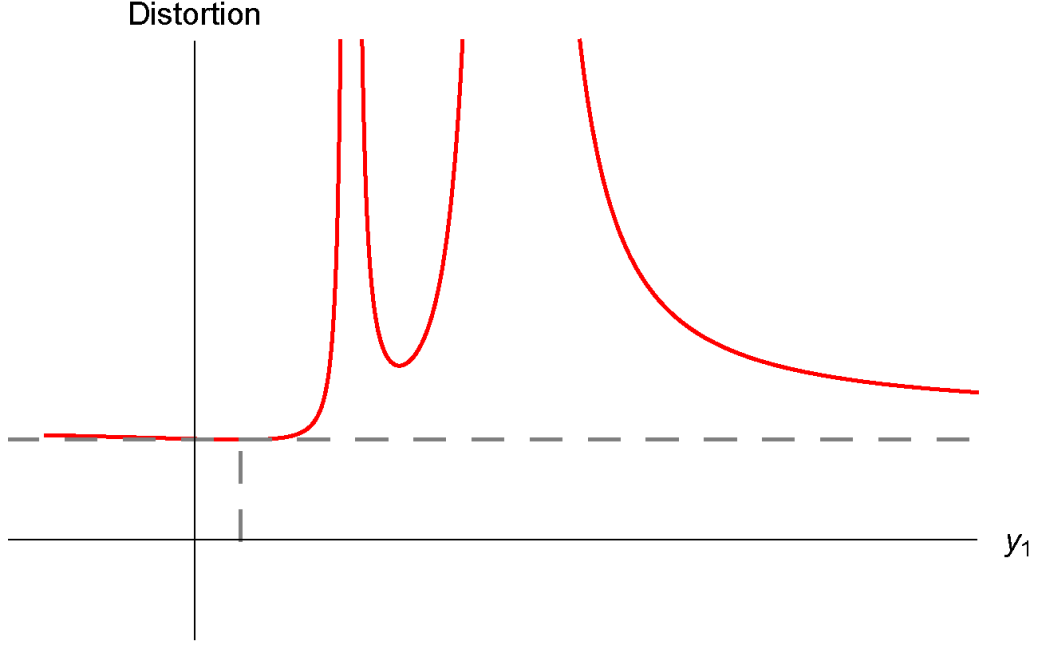


Figure 3.3: Possible trend of the distortion value in Eq. (3.34) as a function of y_1 , value of the y-coordinate on the first camera image (ICS). Global minimum is identified by dashed lines.

be verified that Eq. (3.34) takes the form:

$$\frac{f_1(y_1)}{(g_1(y_1))^2} + \frac{f_2(y_1)}{(g_2(y_1))^2} \quad (3.36)$$

with $f_1 = f_1(y_1)$, $f_2 = f_2(y_1)$ second degree polynomials and $g_1 = g_1(y_1)$, $g_2 = g_2(y_1)$ first degree polynomials. Deriving to find the extreme points yields:

$$\left(\frac{f_1}{g_1^2} + \frac{f_2}{g_2^2} \right)' = \frac{g_1 g_2^3 f_1' + g_1^3 g_2 f_2' - 2f_1 g_2^3 g_1' - 2f_2 g_1^3 g_2'}{g_1^3 g_2^3} \quad (3.37)$$

where the terms of 5th degree cancel out. Discarding the denominator¹, the extreme points are thus found as the solution of a 4th degree polynomial:

$$ay_1^4 + by_1^3 + cy_1^2 + dy_1 + e = 0 \quad (3.38)$$

with $a, b, c, d, e \in \mathbb{R}$. The solutions of Eq. (3.38) can be found using any solving formula for homogeneous polynomials of 4th degree. Full calculations are reported in Appendix A.

¹It can be shown that approaching the roots of the denominator of the distortion function both from the upper and lower limit, the function always goes to $+\infty$, therefore the global minimum never reaches $-\infty$, as in Fig. 3.3.

Fig. 3.3 shows a possible behavior of Eq. (3.34). Among the acceptable real solutions of Eq. (3.38), the one representing the global minimum depends on the initial position of the cameras, and therefore can only be determined by directly comparing the value of the distortion in Eq. (3.34) for each solution.

It must be remarked that, in the very peculiar (unrealistic) case in which the two cameras are identical (i.e. $\mathbf{A}_1 = \mathbf{A}_2$, $\mathbf{P}_1 = \mathbf{P}_2$ and $\mathbf{P}_{c1} = \mathbf{P}_{c2}$) and share the exact same orientation $\mathbf{R}_1 = \mathbf{R}_2$ (not necessarily frontoparallel), then (3.36) takes the form:

$$2 \frac{f(y_1)}{(g(y_1))^2} \quad (3.39)$$

leading, once derived, to a first degree polynomial, thus to a single minimum.

The analytical expression of $\mathbf{w}_1$ and $\mathbf{w}_2$ minimising perspective distortion is therefore found, so that $\mathbf{H}_{1p}$ and $\mathbf{H}_{2p}$ are directly determined. In order to find the complete expression of the rectifying homographies $\mathbf{H}_1$ and $\mathbf{H}_2$, the respective affine components, as shearing and similarity transformations, can be easily calculated [1].

3.4 Algorithm summary

The steps of the direct rectifying algorithm explained above are summarised as follows:

1. The weights $\mathbf{w}_1$ and $\mathbf{w}_2$ of the perspective components of the rectifying homographies are written as polynomial expression in y_1 , the *y*-intercept of the *horizon* epipolar line on I_1 .
2. Distortion is written as function of y_1 and auxiliary matrices are defined to calculate the coefficients of the quartic polynomial.
3. The acceptable solutions (either 2 or 4) are calculated.
4. The global minimum is determined by direct comparison of the distortion values obtained.
5. The minimising $\mathbf{w}_1$ and $\mathbf{w}_2$ are calculated as last row of the matrices in Eq. (3.15).
6. The similarity and shearing transformations are calculated [1].
7. The rectification transformations $\mathbf{H}_1$ and $\mathbf{H}_2$ are fully determined.

The full Python 3 code is made available at <https://github.com/decadenza/DirectStereoRectification>.

3.5 Discussion

Rectifying homographies are transformations that dramatically reduce the search space for correspondences between a couple of stereo images, thus making the stereo matching problem computationally affordable.

The optimal transformations are traditionally found making use of numerical minimisation (see [1]), which prevents the use of rectification when low computational power is available (e.g. space assets). Furthermore, the application of the traditional method is limited by a convergence issue, due to the matrices decomposition used therein and thus not avoidable, that will be discussed in the reminder of this Section. For both these reasons, most present applications [22, 23] prefer to use the arbitrary, non-optimal solution proposed by Fusiello et al. [6], that, for peculiar configurations introduces high distortion levels (Sec. 3.5.1), thus impairing performance of the following stereo matching algorithm.

This paper explicits and demonstrates the formula for the optimal rectifying homographies. The formula is valid for every pair of stereo images, independently from the configuration parameters. Being an exact formulation, it eliminates the need for minimisation, while still providing the optimal transformation, thus enabling the use of rectification in scenarios with very limited computational capabilities, as for example for autonomous navigation of space satellites. Future work will concentrate on such applications.

The reminder of this section discusses a limiting convergence issue of the traditional method proposed by Loop and Zhang [1]. Indeed, the convergence of this algorithm cannot be guaranteed as it depends upon finding a suitable initial guess, derived assuming $\mathbf{A}$ and $\mathbf{A}'$ to be positive-definite, which cannot be guaranteed for all configurations [7].

Consider, for example, the case in which the first CCS is coincident with the WCS and an identical second camera is placed in $\begin{bmatrix} 1 & a & b \end{bmatrix}^T$ (with $a, b \in \mathbb{R}$), oriented as the first one but rotated around the x -axis of an angle θ_x . For all cases in which $b = a \tan \theta_x$, the first element of $\mathbf{A}'$ will be null, thus, as a straightforward consequence of *Sylvester's Theorem*, $\mathbf{A}'$ will not be positive-definite. Then, Loop and Zhang's algorithm fails.

Moreover, numerical simulations confirm that the configurations for which positive-definite assumption is violated are a numerous and therefore relevant in limiting the applicability of the traditional algorithm. Indeed, considering a setup with fixed intrinsic and randomly generating one million of extrinsic parameters (i.e. relative position and rotation between the two cameras), it was found that Loop and Zhang's algorithm failed in over the 50% of the cases, in spite of all the configurations being perfectly legitimate. This happened mostly because of numerical errors in computing Cholesky decomposition of $\mathbf{A}$ and $\mathbf{A}'$. The proposed analytic method, instead, does not need an initial guess and directly provides the optimal solution.

To the best of the author's knowledge, for the cases mentioned above, all algorithms present in the literature fail in providing the optimal rectifying homography (introducing minimum distortion), found instead by the formula derived in this work and not affected by the configuration settings.

Furthermore, Fig. 3.4 shows the optimal solution found by our algorithm in case of extremely skewed camera positions configuration, for which Loop and Zhang's initial guess cannot be retrieved. Finally it must be mentioned that the direct algorithm here presented is capable of providing the optimal solution also in the case of one camera entering the field of view of the other, i.e. when the epipoles lie within image boundaries, which causes convergence issues when employing numerical algorithms.

In summary, the formula here presented guarantees convergence for every configuration, avoids the need of external libraries and saves computational time as it does not make use of minimisation. Indeed, comparing the computational cost, in our implementation on an Intel i7-7500U CPU 2.70GHz, the original algorithm by Loop and Zhang requires 8.899ms on average to calculate the rectifying homographies, while the proposed method takes 7.655ms.

3.5.1 Rectification examples

A synthetic example of rectification is shown in Fig. 3.5. The original left and right images are first shown. The images, as rectified following our direct algorithm, are then listed, where arbitrary horizontal lines have been drawn as a reference. Here the effects of rectification are clearly visible, as corresponding points are aligned. This example is included in the Python code.

The couple of images on the third line of Fig. 3.5, shows the same image pair rectified

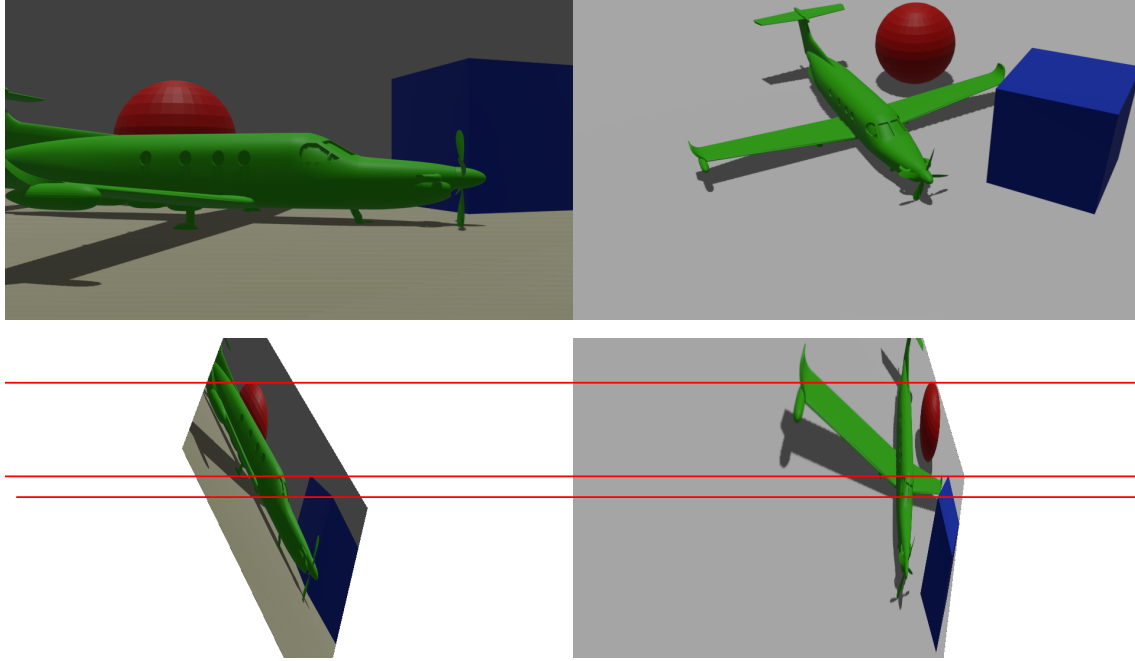


Figure 3.4: Left and right image of a synthetic scene with extremely skewed camera positions (top-left and top-right, respectively). In this case the numerical minimisation of [1] fails. Corresponding rectified image pair using our algorithm is showed (bottom). Horizontal lines are drawn for reference on the rectified image.

following the algorithm in [6]. As expected, while the value of the minimum distortion introduced in the second row is 46252, the homographies found by Fusiello’s algorithm and generating the third pair of images are non optimal and introduce a distortion of 48207, about 4% higher than the minimum.

In Fig. 3.6 a real scene stereo pair is rectified. Despite the cameras are placed almost frontoparallel, corresponding points still lie on different scan-lines, so rectification is needed. In real applications, calibration [2] is required to accurately fit the stereo rig to the pinhole camera model. Apart from lens distortion correction, the algorithm finds no difference between real-case and synthetic images, since it starts from the same premises (i.e. both camera projection matrices).

3.6 Conclusions

A direct rectification algorithm for calibrated stereo rigs has been proposed. Our method improves the well-known approach by Loop and Zhang to calculate the optimal rectifying homographies.

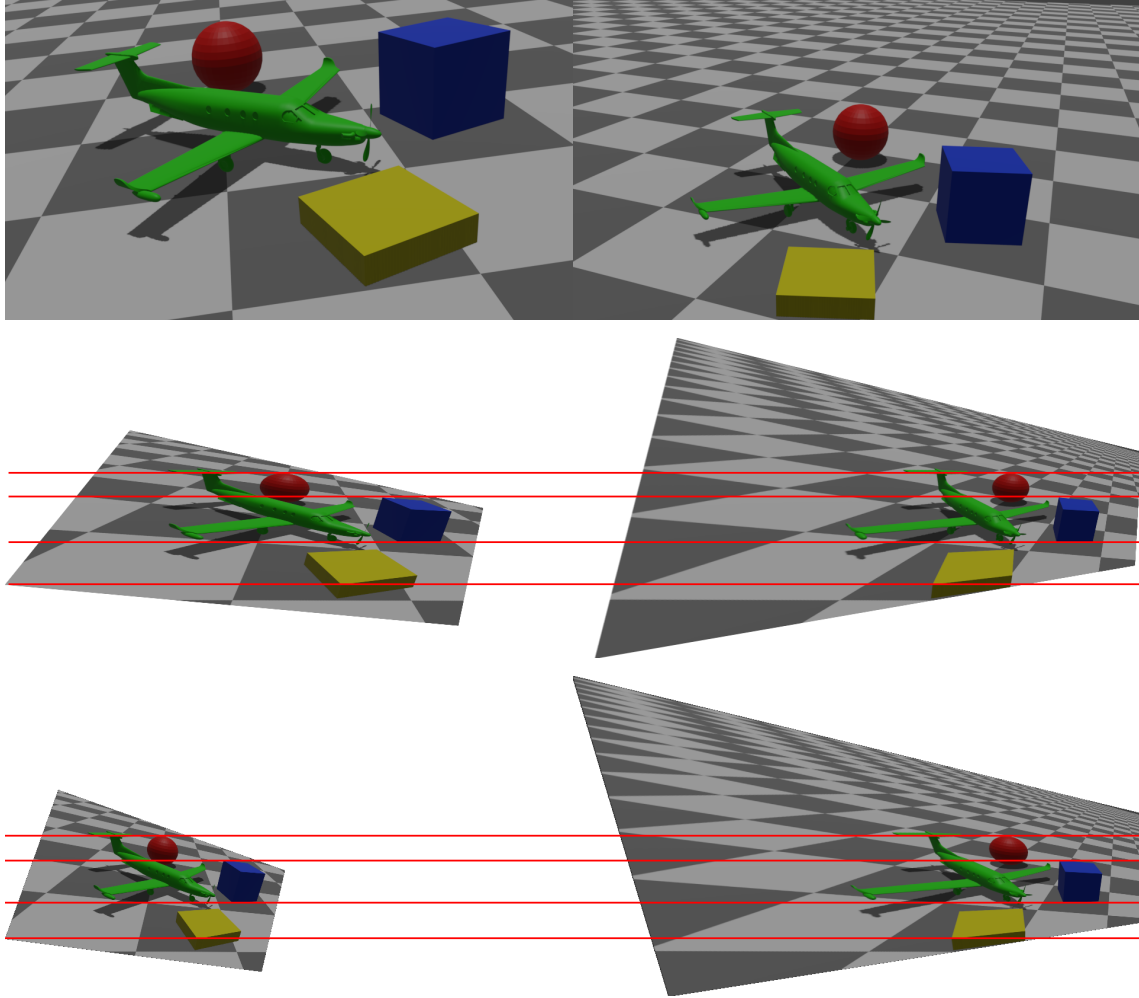


Figure 3.5: Left and right image of a synthetic scene (top-left and top-right, respectively) and corresponding rectified image pair as rectified using the direct algorithm (center) and the algorithm in [6](bottom). Horizontal lines are drawn for reference on the rectified images.

Thanks to an alternative geometrical interpretation of the problem, the proposed algorithm is able to find the formula for the rectifying homographies that introduce the minimal perspective distortion for any camera configurations, even in extremely skewed relative positions, and without depending on minimisation libraries.

The Python 3 code has been made publicly available at <https://github.com/decadenza/DirectStereoRectification>.

Because of the lower computational cost, the value of having an analytic solution may be particularly relevant for hardware-limited applications (e.g. space applications, miniaturised devices), where each change of camera extrinsic or intrinsic parameters would require the computation of new rectifying homographies. Future work might formulate

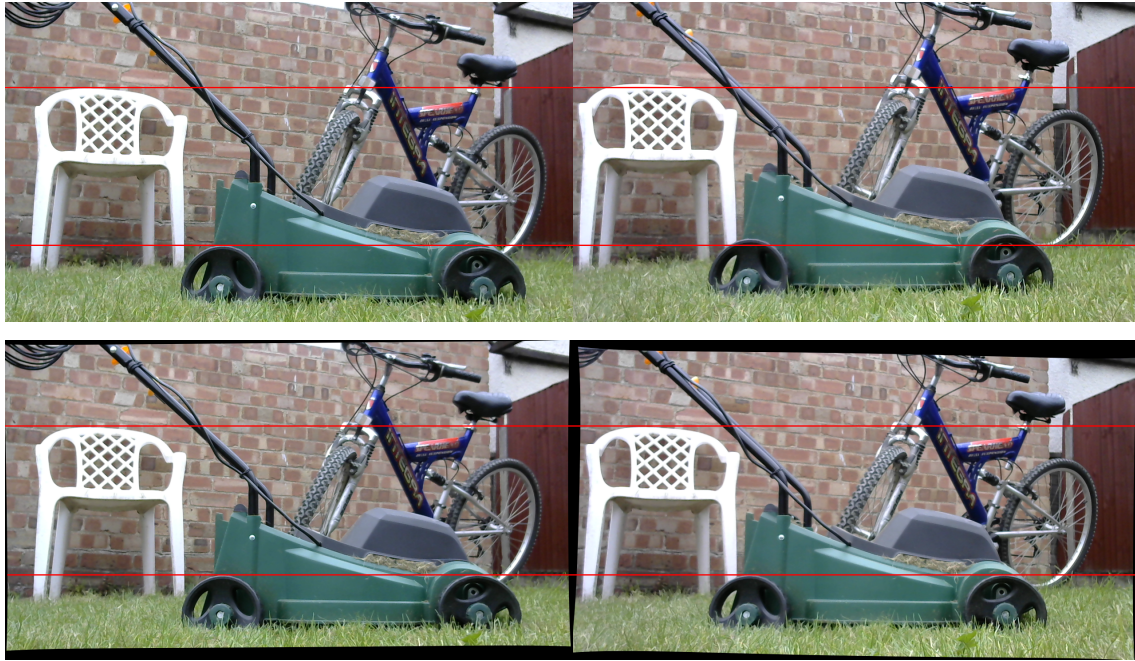


Figure 3.6: Left and right image of a real scene (top-left and top-right, respectively) and corresponding rectified image pair (bottom). Horizontal lines are drawn for reference on both images, showing effects of rectification. Lens distortion correction is also noticeable.

the analytic solution for a distortion metric that includes pixel resampling effects and applications to scenarios with very limited computational capabilities.

References

- [1] Charles Loop and Zhengyou Zhang. “Computing rectifying homographies for stereo vision”. In: *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*. Vol. 1. 1999. ISBN: 0-7695-0149-4. DOI: 10.1109/CVPR.1999.786928.
- [2] Zhengyou Zhang. “A Flexible New Technique for Camera Calibration”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 22 (2000-12), pp. 1330–1334. DOI: 10.1109/34.888718.
- [3] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. 2nd ed. Cambridge, UK: Cambridge University Press, 2003.
- [4] Hornberg. “Handbook of Machine and Computer Vision.” In: Wiley-VCH, 2017.
- [5] P.J. Besl. *Active Optical Range Imaging Sensors*. New York, NY: Springer, 1989.

- [6] Andrea Fusiello, Emanuele Trucco, and Alessandro Verri. “A Compact Algorithm for Rectification of Stereo Pairs”. In: *Machine Vision and Applications* 12 (2000). DOI: 10.1007/s001380050120.
- [7] Liansheng Sui, Zhang Jiulong, and Cui Duwu. “Image Rectification Using Affine Epipolar Geometric Constraint”. In: *International Symposium on Computer Science and Computational Technology*. 2009. DOI: 10.1109/ISCST.2008.76.
- [8] J. Gluckman and S.K. Nayar. “Rectifying transformations that minimize resampling effects”. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. 2001. ISBN: 0-7695-1272-0. DOI: 10.1109/CVPR.2001.990463.
- [9] Changming Sun. “Closed-Form Stereo Image Rectification”. In: Association for Computing Machinery, 2012. DOI: 10.1145/2425836.2425866.
- [10] Zezhi Chen, Chengke Wu, and Hung Tsui. “A New Image Rectification Algorithm”. In: *Pattern Recognition Letters* (2003-01). DOI: 10.1016/S0167-8655(02)00239-8.
- [11] S. Kumar, C. Micheloni, C. Piciarelli, and G.L. Foresti. “Stereo rectification of uncalibrated and heterogeneous images”. In: *Pattern Recognition Letters* (2010). DOI: 10.1016/j.patrec.2010.03.019.
- [12] Xiao Ruichao, Sun Wenxiu, Pang Jiahao, Yan Qiong, and Ren Jimmy. “DSR: Direct Self-Rectification for Uncalibrated Dual-Lens Cameras”. In: *International Conference on 3D Vision* (2018). DOI: 10.1109/3DV.2018.00070.
- [13] R. Deriche, Zhengyou Zhang, Q.-T Luong, and Olivier Faugeras. “Robust Recovery of the Epipolar Geometry for an Uncalibrated Stereo Rig”. In: *Proceedings of the Third European Conference on Computer Vision, vol. 1, Stockholm, Sweden* (2001).
- [14] Pascal Monasse, Jean-Michel Morel, and Zhongwei Tang. “Three-step image rectification”. In: *BMVC 2010-British Machine Vision Conference*. 2010. DOI: 10.5244/C.24.89.
- [15] Steffen Abraham and W. Förstner. “Fish-Eye-Stereo Calibration and Epipolar Rectification”. In: *Isprs Journal of Photogrammetry and Remote Sensing* (2005).
- [16] Xue Zhucun, Nan Xue, Gui-Song Xia, and Weiming Shen. “Learning to Calibrate Straight Lines for Fisheye Image Rectification”. In: (2019-04).
- [17] Xiaoqing Yin, Xinchao Wang, Jun Yu, et al. “Fisheyerecnet: A multi-context collaborative deep network for fisheye image rectification”. In: *Proceedings of the European conference on computer vision (ECCV)*. 2018.

- [18] Quan-Tuan Luong and Olivier Faugeras. “The Fundamental Matrix: Theory, Algorithms, and Stability Analysis”. In: *International Journal of Computer Vision* (1996). DOI: 10.1007/BF00127818.
- [19] Jean Gallier. “Basics of Projective Geometry”. In: *Geometric Methods and Applications: For Computer Science and Engineering*. Springer New York, 2011. ISBN: 978-1-4419-9961-0. DOI: 10.1007/978-1-4419-9961-0_5.
- [20] Cristina Turrini. *Geometria per la ricostruzione tridimensionale*. 2017. URL: http://matematica.unibocconi.it/sites/default/files/turrini_trento.pdf (visited on 2020-05-24).
- [21] Haiping Lu, Konstantinos Plataniotis, and A.N. Venetsanopoulos. *Multilinear Subspace Learning: Dimensionality Reduction of Multidimensional Data*. Chapman & Hall/CRC, 2013-01. DOI: 10.1201/b16252.
- [22] Renbo Xia, Run Su, Jibin Zhao, et al. “An accurate and robust method for the measurement of circular holes based on binocular vision”. In: *Measurement Science and Technology* (2019).
- [23] Timo Hinzmann, Johannes L Schönberger, Marc Pollefeys, and Roland Siegwart. “Mapping on the fly: Real-time 3D dense reconstruction, digital surface map and incremental orthomosaic generation for unmanned aerial vehicles”. In: *Field and Service Robotics*. 2018.

Appendix A

The coefficients of the 4th order polynomial expression in y_1 of Eq. (3.38) are:

$$\begin{aligned}
 a &= m_2 m_4 + m_6 m_8 \\
 b &= m_1 m_4 + 3m_2 m_3 m_4 + m_5 m_8 + 3m_6 m_7 m_8 \\
 c &= 3m_1 m_3 m_4 + 3m_2 m_3^2 m_4 + 3m_5 m_7 m_8 + 3m_6 m_7^2 m_8 \\
 d &= 3m_1 m_3^2 m_4 + m_2 m_3^3 m_4 + 3m_5 m_7^2 m_8 + m_6 m_7^3 m_8 \\
 e &= m_1 m_3^3 m_4 + m_5 m_7^3 m_8
 \end{aligned} \tag{3.40}$$

with:

$$\begin{aligned}
m_1 &= [\mathbf{M}_1]_{23} [\mathbf{C}_1]_{23} - [\mathbf{M}_1]_{33} [\mathbf{C}_1]_{22} \\
m_2 &= [\mathbf{M}_1]_{22} [\mathbf{C}_1]_{23} - [\mathbf{M}_1]_{23} [\mathbf{C}_1]_{22} \\
m_3 &= \frac{[\mathbf{C}_2]_{23}}{[\mathbf{C}_2]_{22}} \\
m_4 &= \frac{[\mathbf{C}_2]_{22}}{[\mathbf{C}_1]_{22}} \\
m_5 &= [\mathbf{M}_2]_{23} [\mathbf{C}_2]_{23} - [\mathbf{M}_2]_{33} [\mathbf{C}_2]_{22} \\
m_6 &= [\mathbf{M}_2]_{22} [\mathbf{C}_2]_{23} - [\mathbf{M}_2]_{23} [\mathbf{C}_2]_{22} \\
m_7 &= \frac{[\mathbf{C}_1]_{23}}{[\mathbf{C}_1]_{22}} \\
m_8 &= \frac{[\mathbf{C}_1]_{22}}{[\mathbf{C}_2]_{22}}
\end{aligned} \tag{3.41}$$

The four roots of the equation are given by:

$$y_1 = \frac{-b}{4a} \pm Q \pm \frac{1}{2} \sqrt{-4Q^2 - 2p + \frac{S}{Q}} \tag{3.42}$$

with:

$$\begin{aligned}
Q &= \frac{1}{2} \sqrt{-\frac{2}{3}p + \frac{1}{3a} \left(\Delta_0 + \frac{q}{\Delta_0} \right)} \\
S &= \frac{8a^2d - 4abc + b^3}{8a^3} \\
\Delta_0 &= \left(\frac{s + \sqrt{s^2 - 4q^3}}{2} \right)^{\frac{1}{3}} \\
p &= \frac{8ac - 3b^2}{8a^2} \\
q &= 12ae - 3bd + c^2 \\
s &= 27ad^2 - 72ace + 27b^2e - 9bcd + 2c^3
\end{aligned} \tag{3.43}$$

Remark: for the case $\mathbf{A}_1 = \mathbf{A}_2$, $\mathbf{P}_1 = \mathbf{P}_2$, $\mathbf{P}_{c1} = \mathbf{P}_{c2}$ and $\mathbf{R}_1 = \mathbf{R}_2$, the solution is given by:

$$y_1 = -\frac{m_1}{m_2} \tag{3.44}$$

Chapter 4

Automated Aircraft Dent Inspection via a Modified Fourier Transform Profilometry Algorithm

Published as P. Lafiosca, I.S. Fan, Avdelidis N.P. in: *Sensors*, Volume 22, Number 2, 2022, doi:10.3390/s22020433.

Abstract

The search for dents is a consistent part of the aircraft inspection workload. The engineer is required to find, measure, and report each dent over the aircraft skin. This process is not only hazardous, but also extremely subject to human factors and environmental conditions. This study discusses the feasibility of automated dent scanning via a single-shot triangular stereo Fourier transform algorithm, designed to be compatible with the use of an unmanned aerial vehicle. The original algorithm is modified introducing two main contributions. First, the automatic estimation of the pass-band filter removes the user interaction in the phase filtering process. Secondly, the employment of a virtual reference plane reduces unwrapping errors, leading to improved accuracy independently of the chosen unwrapping algorithm. Static experiments reached a mean absolute error of ~ 0.1 mm at a distance of 60 cm, while dynamic experiments showed ~ 0.3 mm at a distance of 120 cm. On average, the mean absolute error decreased by $\sim 34\%$, proving the validity of the proposed single-shot 3D reconstruction algorithm and suggesting its applicability for future automated dent inspections.

Keywords: Airworthiness; Fringe patterns; Pinhole camera model; Structured light; Inspections.

4.1 Introduction

Maintenance, repair and overhaul (MRO) companies have an interest in the progressive automation of aircraft inspections. These are not only hazardous, but also costly, time-consuming, labour-intensive, and subject to human error [1], yet they are critical for the *airworthiness* assessment. The potential advantages of automation are evident: improved safety, higher productivity, reduced aircraft downtime, and standard output. Therefore, the general problem of automating inspections has been attracting the attention of many researchers lately, targeting damages on aircraft *skin* for its extent and its desirable characteristic of being directly accessible. Understandably, unmanned aerial vehicles (UAVs) are often considered to reach areas of the skin at a certain height, relieving the engineer from connected safety risks [2]: in 2018, Airbus launched an indoor inspection drone [3] and recently approved UAVs for lightning checks [4]. Similarly, other companies, such as easyJet and Air France Industries-KLM, are pioneering the adoption of UAVs for inspections.

Among inspection technologies that enable the use of UAVs [5], pattern recognition is at the basis of several commercial products recently introduced in MRO. It has the advantage of being easy to implement, not requiring particular hardware except for a high-definition camera, and has been successfully applied for detecting scratches and lightning damage. Thermal imaging has also been proposed for its capability to detect delamination and corrosion [6]. Despite their relevance and frequency of occurrence on the aircraft skin, little has been carried out for dents. Maintenance procedures require periodic and comprehensive inspections during which dents must be found, measured, and reported. However, automation of dent inspections poses new challenges due to the nature of this damage.

A dent is smooth and without well-defined boundaries, with depth easily lower than 1 mm. As the aircraft skin is generally single-coloured and lacking in texture, a dent may not be clearly distinguishable even by trained engineers, who usually rely on reflections or on the light of a torch that seeps under a ruler placed on the surface, to highlight depth variations and spot dent locations during general visual inspections. Furthermore, *measures* of depth and width are required to classify a dent as *allowable damage* or not [7]. In traditional inspections, measures are collected by means of a depth gauge and a ruler.

Although high-accuracy handheld 3D scanning tools are available, quick but less-accurate traditional methods still account for the most part [8, 9] and 3D scanners are reserved for special inspections. The reason is probably to be found in the relatively short operating distances of these devices and in the fact that the aircraft skin includes locations at height normally reached by platforms and climbing equipment. As such, handheld scanning devices not only still require the engineer to reach those locations, they also add practical difficulties for carrying them and increased setup time and scan data analysis effort.

Among the few works addressing dent inspections, Jovančević et al. [10] proposed the use of Air-Cobot equipped with commercial 3D scanners and a region-growing algorithm for the automatic segmentation of dented areas. However, it left open the problem of reaching the upper part of the aircraft. Dođru et al. [11] proposed a convolutional neural network to detect dents, suggesting the use of UAVs to achieve full automation. Still, the use of monocular images not only prevents measuring the damage but also results in low accuracy. The capability to detect a dent with this method is virtually zero for shallow dents over non-textured skin, which are the most common and nevertheless must be found and reported. Hence, the acquisition of 3D data cannot be avoided.

In the above mentioned context, this article discusses the feasibility of automated dent scanning via a single-shot structured-light 3D scanner. With respect to multiple-shot, the choice of a single-shot algorithm makes the system resilient to fast movements and vibration during scanning. With further development, such a system could be integrated on a UAV, enabling quick and effective automated dent inspections.

After a brief review of the available structured-light codification strategies, an algorithm based on the Fourier transform profilometry (FTP) is tested for the scope. The use of a *virtual* reference plane is proposed for:

- The automatic band-pass estimation (Sec. 4.2.2), thus eliminating the user interaction in the filtering process.
- The reduction of errors in the phase unwrapping process via a virtual reference image (Sec. 4.2.4).

Simulations (Sec. 4.3) and experiments (Sec. 4.4) on both static and moving surfaces presenting artificial dents prove the validity of the proposed algorithm and show that the use of FTP is promising for the automation of aircraft dent inspections. Discussion and Conclusions follow in Sec. 4.5 and 4.6.

4.1.1 Structured-light codification strategies

Considering the purpose of this work, only the codification strategies applicable to a minimal structured-light system, composed of a camera and a projector, are taken into account. Therefore, strategies involving the use of multiple cameras, additional constraints in the system geometry, or restricted depth range are omitted.

Coding approaches can usually be classified as time-multiplexing or spatial neighbourhood. Time-multiplexing (or *multiple-shot*) is capable of obtaining higher resolution and accuracy. Techniques in this group can be roughly divided into Gray coding and temporal phase shifting, or a combination of the two [12]. The main advantage of phase-based coding over Gray coding is to provide sub-pixel accuracy, thanks to the use of continuous functions. The large number of projected patterns makes time-multiplexing unsuitable for moving objects [13]. Methods using a limited number of patterns (≥ 2) and high-speed devices can be found in the literature [14, 15, 16]; however, their employment on moving objects, outside laboratory conditions, and without high-end hardware is still an active field of research.

Spatial neighbourhood (or *single-shot*) enables the measurement of moving surfaces and is insensitive to vibrational noise [17]. This comes at the price of lower accuracy, as the coding information must be contained in one pattern only, and local smoothness must be assumed for the neighbourhood to be correctly decoded. Spatial methods can use colour or greyscale values. In colour-based approaches, a compromise is to be found between the number of colours and noise sensitivity, while the use of greyscale intensity values results in more robustness [13]. Similarly to time-multiplexing, also spatial coding can be divided into discrete and continuous. The De Bruijn colour sequence [18] is probably the most representative discrete method, from which several techniques have followed [19, 20, 21]. Discrete methods extract *codewords* from local areas, which in most cases require several pixels, limiting resolution. Continuous spatial coding, instead, uses a smooth pattern to obtain a *dense* surface reconstruction.

Among the latter, Fourier transform profilometry (FTP) aims at calculating the continuous phase value from a single projected pattern through a filter in the frequency domain [22, 23]. The filtering is a critical step, and aliasing may prevent the extraction of the fundamental spectrum. It also poses a limit to the phase derivative along the fringe direction, but not directly to the maximum measurable height [22]. With nonstationary signals, the classic FTP may have difficulties in isolating the first harmonic. In order to solve this, windowed FTP and wavelet transform provide a space–frequency represen-

tation demanding high computational cost [17, 24, 25, 26] and the choice of additional system parameters, although approaches for the automatic selection of the window size have been proposed [27, 28].

Finally, in both time-multiplexing approaches and FTP, the phase is used to calculate the object 3D coordinates. However, while time-multiplexing is generally able to obtain the absolute phase directly (via *temporal* unwrapping), FTP can only provide a *wrapped* phase that needs to be processed through a *spatial unwrapping* algorithm and offset to obtain the absolute phase, not without difficulties [29, 30]. The absolute phase can then be used to find a correspondence between camera and projector, and thus proceed with stereo triangulation [31, 32], or the depth can be seen as a function of the phase difference from a physical reference plane [30, 22, 33].

4.2 Proposed Method

In this study, a modified triangular stereo FTP method is implemented to discuss its compatibility and performance towards automated dent inspections. In addition, two contributions are presented.

First, the automatic estimation of the band-pass filter is proposed (Sec. 4.2.2). This feature is essential, as the first-spectrum band varies with the distance from the scanned surface and the system should be able to proceed *without user interaction*.

Secondly, a strategy for the reduction of phase unwrapping errors is shown (Sec. 4.2.4). All phase-based single-shot methods initially output a wrapped (or modulo 2π) phase. After spatial unwrapping, two methods can be found in the literature to relate phase with depth. In systems calibrated using the pinhole model [34], the object phase can be directly used to solve the stereo correspondence problem, thus proceeding with triangulation [31, 32]. Alternatively, the depth can be seen as a function of the *phase shift* of the object, calculated with respect to a *reference plane*, as commonly performed in phase-height mapping methods [22, 33]. While triangular stereo methods are generally more accurate, suitable for extended depth range measurements, and easier to calibrate out of the lab [35, 30], here it is shown that the employment of the phase shift has the benefit of reducing unwrapping errors and that its advantage can be brought to triangular stereo methods by means of a *virtual* reference plane.

4.2.1 Notation and system geometry

Throughout the article, the pinhole camera model is used for both camera and projector [34]. Without loss of generality, the camera optical centre is placed in the world origin $O_c = (0,0,0)$, oriented in the same way so that the world z -axis corresponds to the camera optical axis. $\mathbf{K}_c$ is the camera intrinsic matrix, while its rotation matrix and translation vector, due to its position, are simply $\mathbf{R}_c = \mathbb{I}$ (3×3 identity matrix) and $\mathbf{t}_c = \mathbf{0}$, respectively. The projector optical center is placed in $O_p = (x_p, y_p, z_p)$, freely oriented. Its intrinsic matrix, rotation matrix, and translation vector are identified by $\mathbf{K}_p$, $\mathbf{R}_p$, and $\mathbf{t}_p$, respectively. Lens distortion correction is managed using the radial and tangential model [34], but is not explicitly reported in the following formulae. Camera-projector calibration is assumed as known, following one of the several methods available in the literature [36, 37, 31, 38]. The described system geometry is represented in Figure 4.1.

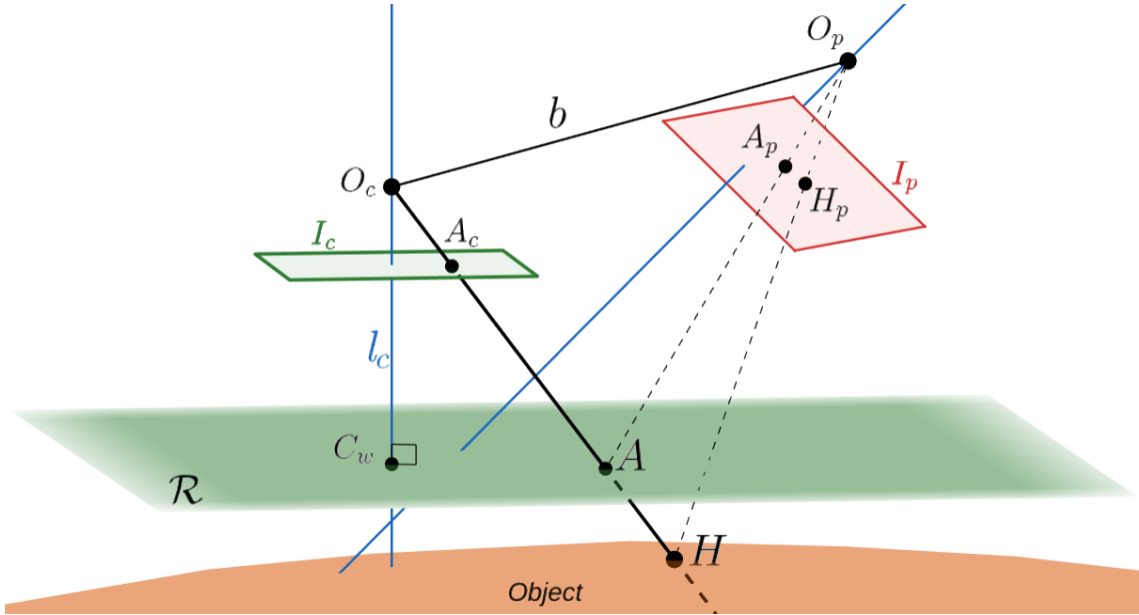


Figure 4.1: System configuration: camera and projector are freely positioned (optical axes in blue solid). Without loss of generality, camera is placed in the world origin and a virtual reference plane $\mathcal{R}$ lays in front of the camera.

I_c and I_p are the camera and projector image planes, respectively. In particular, I_c is the plane where the image acquired by the camera lays *after* distortion correction has been applied, while the distortion-free fringe image chosen as projector input lays on I_p . On each image, a pixel coordinate system is defined, with $(u_c, v_c) \in I_c$ and $(u_p, v_p) \in I_p$ being two generic points laying on camera and projector images, respectively.

A sinusoidal fringe pattern is chosen with fringes perpendicular to the u_p -axis, having period T_p (in pixels) and frequency $f_p = 1/T_p$. A red-coloured stripe is drawn for one cen-

tral period, following the intensity trend (Figure 4.2). After projection, the corresponding points can be identified from the camera using a weighted average of red intensity values over the u -axis, producing sub-pixel accuracy. The use of a similar centerline was also proposed in [31].

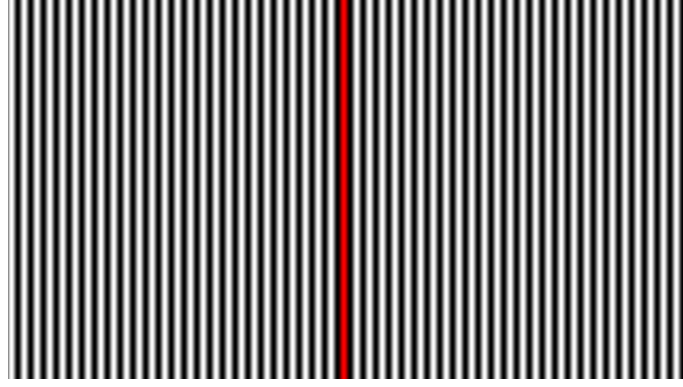


Figure 4.2: A fringe pattern with period T_p , with a central red-coloured stripe.

The equation of the virtual reference plane $\mathcal{R}$ is considered at $z = l_c$, thus laying perfectly in front of the camera, where l_c is chosen as the average depth of the red stripe world points, readily calculated by triangulation.

The aim of FTP is to find the generic 3D point H belonging to the object, observed as $A_c \in I_c$ and $H_p \in I_p$, with A being the intersection of the ray $\overline{O_c H}$ and $\mathcal{R}$.

4.2.2 Automatic band-pass filter

In FTP, a band-pass filter must be applied to isolate the fundamental spectrum and filter out DC component and upper harmonics [22, 17, 29]. The band of interest can be identified by the interval $[f_c - r, f_c + r]$, where f_c is the fundamental frequency, as seen from the camera, and r is an appropriate radius. Experimentally, it can be shown that the fundamental spectrum varies significantly when scanning objects at different distances, requiring the human interaction for the manual selection of the band-pass at each scan. This becomes impractical when dealing with a large number of scans with varying distances. The automatic band-pass filter solves this problem.

The estimation of f_c assumes that the area of major interest is at depth $z = l_c$ (as calculated in Sec. 4.2.1). The intersection of camera optical axis and $\mathcal{R}$ is simply $C_w =$

$(0, 0, l_c)^T$, that is projected on the projector image plane to find $C_p = (u_p, v_p) \in I_p$:

$$\begin{pmatrix} u_p \\ v_p \\ 1 \end{pmatrix} = \text{Normalise} \left(\mathbf{K}_p \cdot [\mathbf{R}_p | \mathbf{t}_p] \begin{pmatrix} 0 \\ 0 \\ l_c \end{pmatrix} \right) \quad (4.1)$$

where $\text{Normalise}()$ is a function that divides a homogeneous coordinate by its 3rd element. On I_p , two points can be identified as

$$C_p^{left} = \begin{pmatrix} u_p - \frac{T_p}{2} \\ v_p \\ 1 \end{pmatrix} \quad C_p^{right} = \begin{pmatrix} u_p + \frac{T_p}{2} \\ v_p \\ 1 \end{pmatrix} \quad (4.2)$$

covering a full period T_p . Then, C_p^{left} is deprojected back to $\mathcal{R}$ and, from here, to the camera:

$$C_c^{left} = \text{Normalise} \left(\mathbf{K}_c \cdot \mathbf{M} \cdot \text{Normalise} \left((\mathbf{K}_p \cdot \mathbf{R}_p)^{-1} \cdot \begin{pmatrix} u_p - \frac{T_p}{2} \\ v_p \\ 1 \end{pmatrix} \right) \right) \quad (4.3)$$

where:

$$\mathbf{M} = \begin{pmatrix} l_c - z_p & 0 & x_p \\ 0 & l_c - z_p & y_p \\ 0 & 0 & l_c \end{pmatrix} \quad (4.4)$$

The same applies to C_p^{right} , giving C_c^{right} . Lens distortion is corrected accordingly in all of the steps. C_c^{left} and C_c^{right} lie on I_c at a distance corresponding to the *projector* period and its orientation. From Pythagorean geometry, T_c is calculated as the projection of $\overline{C_c^{left} C_c^{right}}$ over the u_c -axis. Finally, $f_c = 1/T_c$ and, throughout this paper, $r = f_c/2$.

The band-pass estimation does not require any user input and allows for some automation [35, 29] and extended adaptability of FTP in different system geometries and at varying distances from the object, particularly when scanning surfaces roughly placed in front of the camera, as in the application here discussed. The automatic band-pass can successfully isolate the main spectrum (Fig. 4.3) and it was successfully used in all the following simulations (Sec. 4.3) and experiments (Sec. 4.4).

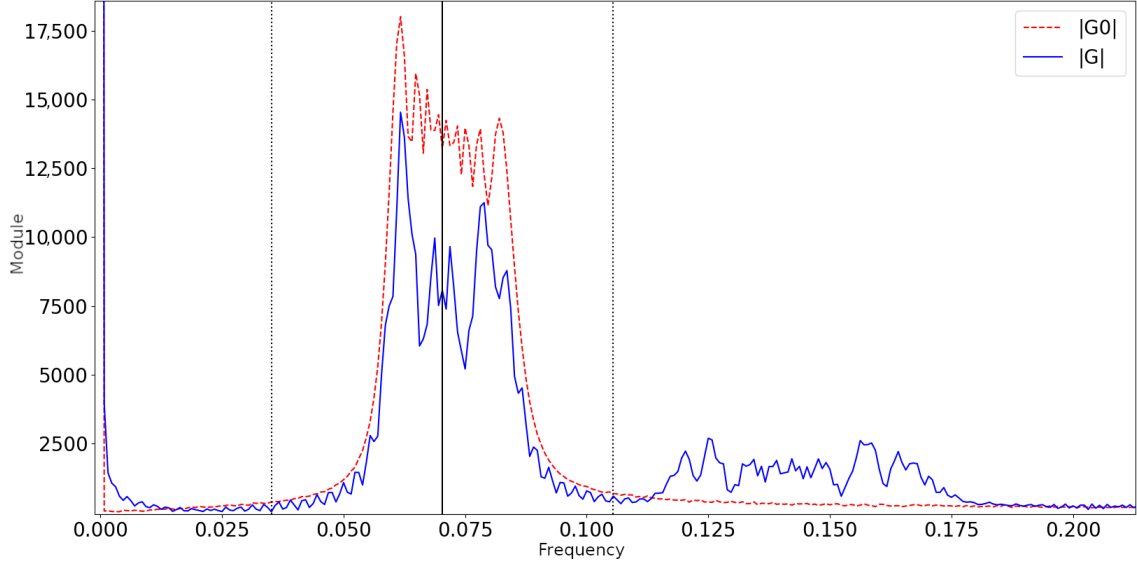


Figure 4.3: Module of the FFT on both reference (dotted red) and object (solid blue) images along middle row for a real dent sample. Spectrum around the fundamental frequency (solid vertical) is isolated with the automatic band-pass filter (dotted vertical).

4.2.3 Virtual reference image

On the domain of the points of $\mathcal{R}$ observed from the camera and illuminated by the projector, the relation between projector and camera image planes is *bijective*. This allows to virtually build a reference image as seen from the camera.

The generic point $A_p = (u_p^A, v_p^A) \in I_p$ is projected to the world point $A \in \mathcal{R}$, which is, in turn, observed from the camera as $A_c = (u_c^A, v_c^A) \in I_c$. A_p is calculated from A_c as

$$\begin{pmatrix} u_p^A \\ v_p^A \\ 1 \end{pmatrix} = \text{Normalise} \left(\mathbf{K}_p \cdot \left[l_c \mathbf{R}_p \mathbf{K}_c^{-1} | \mathbf{t}_p \right] \begin{pmatrix} u_c^A \\ v_c^A \\ 1 \end{pmatrix} \right) \quad (4.5)$$

In other words, the effect of Equation (4.5) is to find the line passing through O_c and A_c that intersects $\mathcal{R}$ at a certain world point A , and then project A onto I_p . As usual, camera and projector coordinates are also corrected for lens distortion.

With this simple strategy, each point of the camera image is mapped to its corresponding point and its intensity value on the projected fringe image on I_p . As the mapping generally produces noninteger numbers, the final camera image is derived with cubic interpolation. The red stripe is not needed in this image.

The phase variation due to perspective is clearly visible in Figure 4.4. As for its frequency content, the virtual reference image is the same as the one that would be acquired placing a real plane at the same position. However, the effects of light diffusion, object albedo, and environment illumination are not introduced.

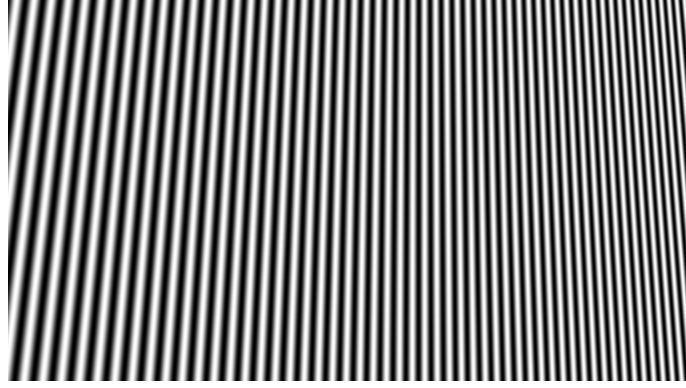


Figure 4.4: Virtually built reference image with sinusoidal fringes. Perspective affects frequency components.

The virtual reference image is then used to calculate the phase shift, similarly to phase-height methods [22, 33] but without a physical plane. The usefulness of this is explained below.

4.2.4 Reduction of noise effect over phase jumps

With single-shot FTP, the choice of unwrapping algorithms is limited to the spatial ones. To calculate the absolute phase Φ , estimated as $\tilde{\Phi}$, spatial phase unwrapping cannot rely on information other than the phase map Φ^w that, due to the presence of noise, may present *genuine* or *fake* 2π -jumps [39]. The two are indistinguishable, making the spatial unwrapping such a challenging task [40].

In both stereo triangulation and phase-height mapping methods, $\tilde{\Phi}$ is related to the world z coordinate. While stereo triangulation directly relates $\tilde{\Phi}$ to a point on I_p [31, 32], phase-height mapping estimates z as a function of the *phase shift* $\Delta\tilde{\Phi}$ from a reference phase Φ_0 calculated over a plane [22, 33].

Although the proposed method uses a stereo triangulation approach, it also exploits the calculation of $\Delta\tilde{\Phi}$ with respect to a *virtual* reference image (Sec. 4.2.3) to reduce the number of 2π -jumps, both genuine and fake ones. Consequently, unwrapping errors are reduced, as there is less probability for the noise to interfere with 2π -jumps.

The virtual reference and the object intensity values observed from the camera can be represented, respectively, as:

$$\begin{aligned} g_0(u_c, v_c) &= A_0(1 + \cos \Phi_0) \\ &= A_0\{1 + \cos[2\pi f_c u_c + \varphi_0(u_c, v_c)]\} \end{aligned} \quad (4.6)$$

and

$$\begin{aligned} g(u_c, v_c) &= A(1 + \cos \Phi) \\ &= A\{1 + \cos[2\pi f_c u_c + \varphi(u_c, v_c)]\} \end{aligned} \quad (4.7)$$

where f_c is the fundamental frequency (Sec. 4.2.2), $A = A(u_c, v_c)$ and $A_0 = A_0(u_c, v_c)$ are intensity modulations, $\varphi_0(u_c, v_c)$ represents the phase modulation over the virtual reference image (caused by perspective only), and $\varphi(u_c, v_c)$ is the phase modulation over the object caused by perspective, object depth, and noise.

After fast Fourier transform (FFT), filtering and inverse FFT (without aliasing) of g , the wrapped object phase can be obtained as:

$$\Phi^w(u_c, v_c) = \arg \left[\frac{A}{2} e^{i\Phi(u_c, v_c)} \right] = W[2\pi f_c u_c + \varphi(u_c, v_c)] \quad (4.8)$$

where $W[\cdot]$ stands for the wrapping operation (or modulo 2π). Then, the classic direct method estimates the unwrapped object phase as:

$$\tilde{\Phi}^r(u_c, v_c) = U[\Phi^w(u_c, v_c)] \quad (4.9)$$

where $U[\cdot]$ represents a generic spatial unwrapping algorithm. Note that $\tilde{\Phi}^r$ is still a *relative* phase, with uncertain fringe order.

Furthermore, as in phase-height mapping methods [22, 33], the wrapped phase shift can be calculated from a new signal obtained as the filtered object signal multiplied by the complex conjugate of the filtered reference [22]:

$$\begin{aligned} \Delta\Phi^w(u_c, v_c) &= \arg \left[\frac{A}{2} e^{i\Phi(u_c, v_c)} \cdot \frac{A_0}{2} e^{-i\Phi_0(u_c, v_c)} \right] \\ &= W[\varphi(u_c, v_c) - \varphi_0(u_c, v_c)] \end{aligned} \quad (4.10)$$

and the unwrapped object phase can be obtained also, as:

$$\begin{aligned} \tilde{\Phi}^r(u_c, v_c) &= \Phi_0(u_c, v_c) + U[\Delta\Phi^w(u_c, v_c)] \\ &= \Phi_0(u_c, v_c) + \Delta\tilde{\Phi}^r(u_c, v_c) \end{aligned} \quad (4.11)$$

where $\Phi_0(u_c, v_c)$ is the absolute reference phase and $\Delta\tilde{\Phi}^r$ is the unwrapped phase shift.

Equations (4.9) and (4.11) should lead to the same result. However, the interaction of noise over 2π -jumps will be greater in Equation (4.9) than in Equation (4.11). In fact, under the general assumption of FTP, φ and φ_0 must vary very slowly compared to the fundamental frequency f_c [22], and as a direct consequence, $\Delta\Phi^w$ presents less 2π -jumps compared to Φ^w . An example is shown in Figure 4.5.

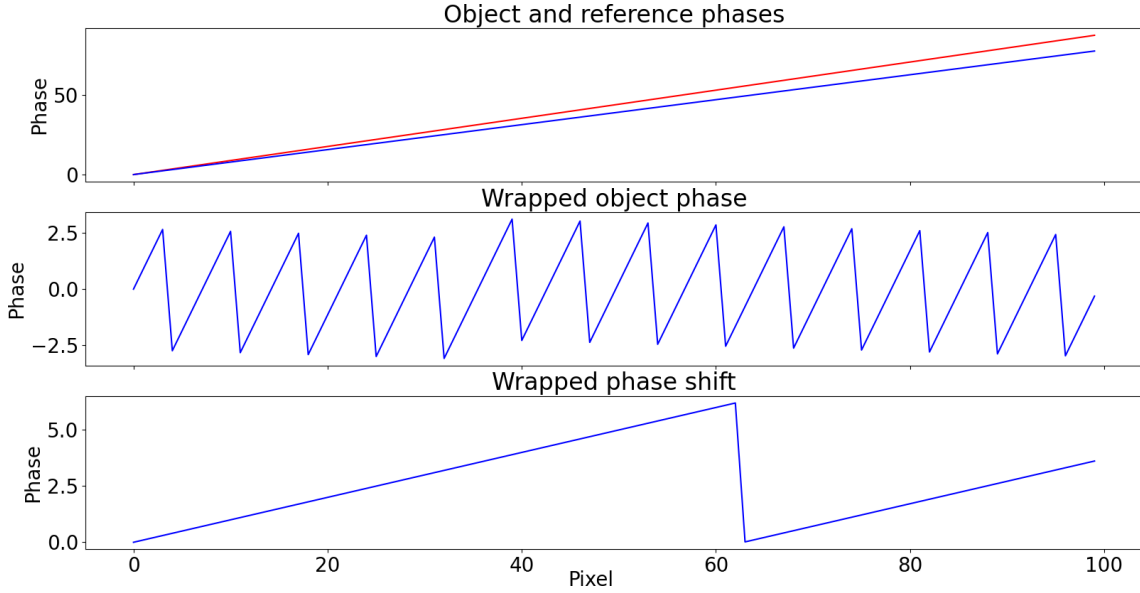


Figure 4.5: (Top) Illustrative reference Φ_0 (blue) and object Φ (red) phases. (Middle) The wrapped object phase Φ^w presents many more 2π -jumps compared to the (bottom) wrapped phase shift $\Delta\Phi^w$.

Therefore, Equation (4.11) allows to deal with the unwrapping of the signal $\Delta\Phi^w$ that contains a slower trend superimposed modulo 2π on the same noise distribution found on Φ^w (which originates from the sole object image, as the virtual reference is noise-free). This leads to a more faithful unwrapped phase.

To show the practical effect of this, a white noise ($\sigma = 0.6$) is added to Φ of the same example of Figure 4.5. The wrapped object phase and phase shift are shown in Figure 4.6: it is clear that the task of unwrapping becomes more challenging due to noise interaction with jumps and that the probability for it to interfere are lower in the second case, due to fewer 2π -jumps.

The unwrapped phase using the algorithm by Itoh [41] is shown in Figure 4.7. Even in this very simple case, noise causes a genuine jump right after $u_c = 20$ to be missed, thus propagating the error to the right. The unwrapping of $\Delta\Phi^w$ in Equation (4.11), instead, is more robust and outputs the correct result.

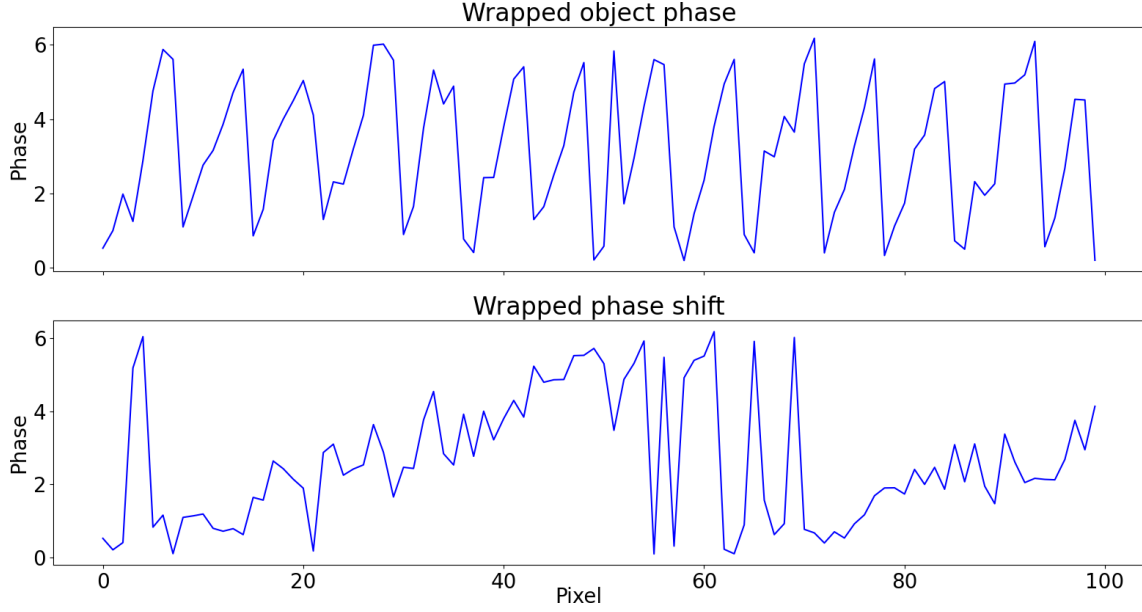


Figure 4.6: (*Top*) Wrapped object phase Φ^w and (*bottom*) wrapped phase shift $\Delta\Phi^w$ with the same white noise ($\sigma = 0.6$) added.

Basic single-shot structured-light applications cannot rely on additional knowledge to eliminate noise, and a spatial unwrapping algorithm cannot count on information other than the wrapped phase itself. Therefore, the reduction of 2π -jumps increases performance, regardless of the unwrapping algorithm used. This was confirmed by the following simulations (Sec. 4.3) and experiments (Sec. 4.4).

Furthermore, simulations in the absence of noise revealed that the use of Equation (4.11) in place of Equation (4.9) increases the numerical precision. This can be explained by the fact that, in computers, complex numbers are generally stored in algebraic form and the argument is extracted by means of the $\arctan2$ function. Numerically, its asymptote regions are more sensitive to machine precision error, and Equation (4.11) leads to reduced error, as the slow-growing signal crosses those regions fewer times. Depending on sampling and period, the error $\|\tilde{\Phi} - \Phi\|$ decreased up to 10^{-6} .

4.2.5 Stereo correspondence

Before proceeding with triangulation, the relative unwrapped phase $\tilde{\Phi}^r$ must be shifted accordingly to obtain the absolute phase $\tilde{\Phi}$, following:

$$\tilde{\Phi} = \tilde{\Phi}^r + 2k\pi \quad (4.12)$$

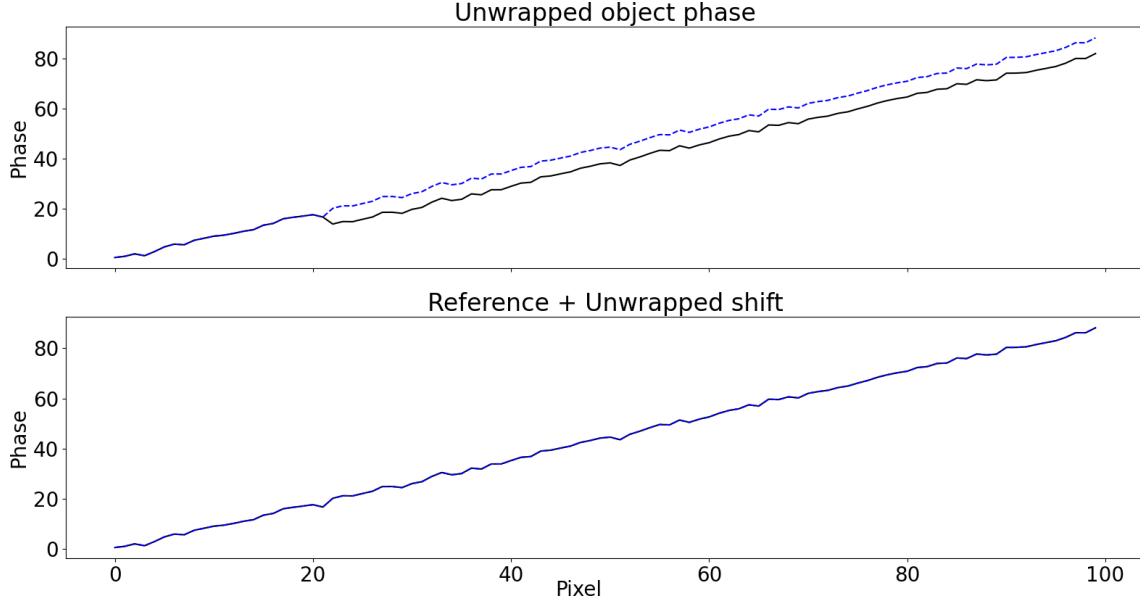


Figure 4.7: The unwrapped object phase $\tilde{\Phi}^r(u_c, v_c)$ using **(top)** Equation (4.9) and **(bottom)** Equation (4.11) with respect to the original phase (dashed blue). In the first case, the noise causes the unwrapping algorithm to discard a genuine jump right after $u_c = 20$, introducing error. This does not happen in the second case due to the reduced presence of 2π -jumps.

where $k \in \mathbb{K}$ is the difference from the correct fringe order [40].

Additional information is needed to find $\tilde{\Phi}$, for example, the absolute phase value in correspondence of the red stripe [31]:

$$\tilde{\Phi}(u_c, v_c) = \tilde{\Phi}^r(u_c, v_c) - \frac{\sum_{n=0}^N \tilde{\Phi}_n(u_c, v_c)}{N} \quad (4.13)$$

where N is the number of pixels of the red stripe and $\tilde{\Phi}_n(u_c, v_c)$ phase values at stripe locations. Then, each $\tilde{\Phi}(u_c, v_c)$ value may be directly associated with its u_p coordinate, as:

$$u_p = \frac{W_p \cdot \tilde{\Phi}(u_c, v_c)}{2\pi} \quad (4.14)$$

where W_p is the projector resolution along the u_p -axis.

Equivalently, Equation (4.12) can be rewritten as

$$2\pi f_p u_p = 2\pi f_p u_p^A + \Delta \tilde{\Phi}^r(u_c^A, v_c^A) + 2k\pi \quad (4.15)$$

where u_p^A is obtained through Equation (4.5). Applying the above relation for each of the red stripe u -coordinate found on I_c allows to resolve for k . The average k is selected

and rounded to the nearest integer. Numerically, the latter approach delivers increased accuracy, as k is correctly forced to be an integer.

Once k is known, Equation (4.15) can be used to retrieve all the u_p points with respect to each camera point. Finally, v_p is calculated via epipolar geometry, and having all the corresponding couples, $(u_c, v_c) \in I_c$ and $(u_p, v_p) \in I_p$, one may proceed with stereo triangulation.

4.3 Computer Simulations

In the aircraft structural repair manual, a dent is defined as a damaged area pushed from its normal contour that presents smooth edges [7]. The length of a dent is the longest distance from one end to the other, while the width is the maximum distance measured at 90° from the direction of the length. The depth is measured at their intersection. A dent is usually classified by its width, depth, and width/depth ratio.

In the absence of a formal shape definition, the following was chosen as representative *dent function*:

$$z = \begin{cases} -e^{-\frac{1}{1-r^2}} & \text{if } |r| < 1 \\ 0 & \text{elsewhere} \end{cases} \quad (4.16)$$

where $r = \sqrt{x^2 + y^2}$. As for real dents, the function presents a smooth trend and vanishing boundaries (Figure 4.8). It can be rescaled and rotated to resemble different dents.

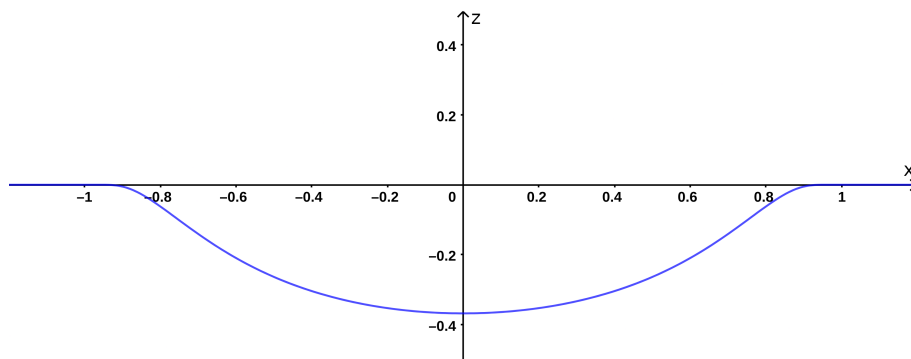
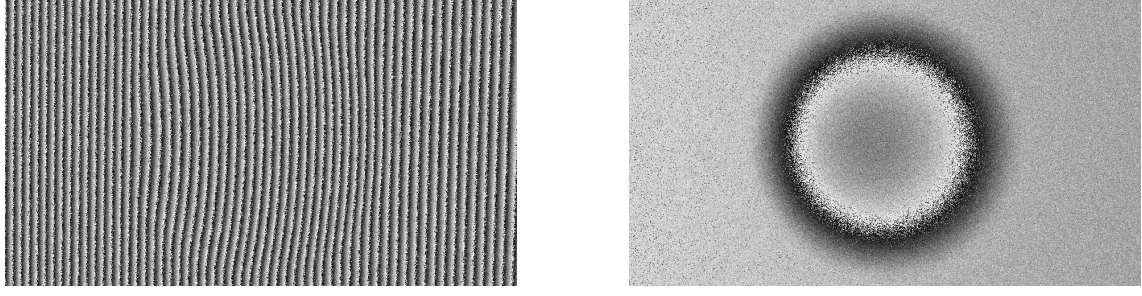


Figure 4.8: Section of the reference dent function. Maximum depth is 0.368 and just 0.005 at $x = 0.9$.

A simulated dent with maximum depth 3 mm and width 30 mm was scanned from 1 m of distance with a resolution of 1280×720 pixels and $T_p = 8$. White noise ($\sigma = 0.6$) was added, similarly to Sec. 4.2.4. The object phase maps and the middle row signals

are shown in Figure 4.9 and 4.10, respectively. Due to noise, Itoh's unwrapping al-



(a) Object phase Φ^w .

(b) Phase shift $\Delta\Phi^w$.

Figure 4.9: Compared to the (a) object phase Φ^w , the slower trend of the (b) phase shift $\Delta\Phi^w$ is clearly visible in a simulated dent with added white noise ($\sigma = 0.6$).

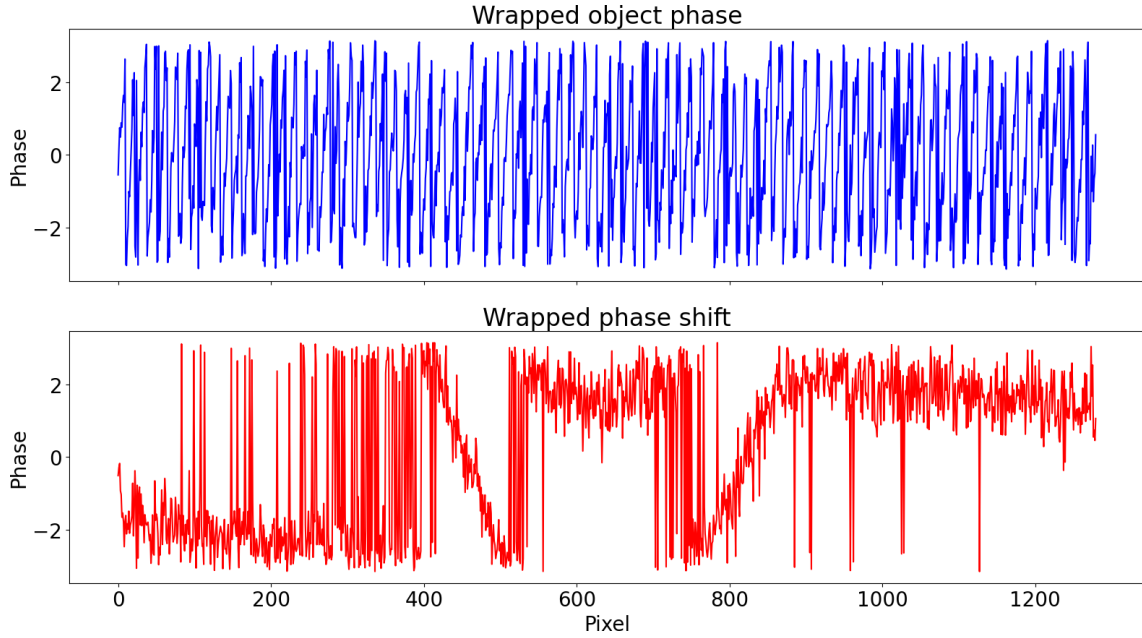


Figure 4.10: (Top) Wrapped phase Φ^w and (bottom) phase shift $\Delta\Phi^w$ at middle row. Unwrapping of Φ^w has to manage many more 2π -jumps.

gorithm missed 2π -jumps on more occasions when using Equation (4.9) compared to Equation (4.11). In Figure 4.11, the detail of a missed genuine 2π -jump is showed.

The mean absolute error (MAE) was chosen as a metric, calculated as the distance between the expected mesh to the point cloud [42]. After 3D reconstruction, the MAE from the original simulated dent was 6.820mm (its standard deviation was $\sigma = 7.174$ mm) with the direct method, and 3.827mm ($\sigma = 7.090$ mm) with the proposed method. Reconstructed 3D points are shown in Figure 4.12.

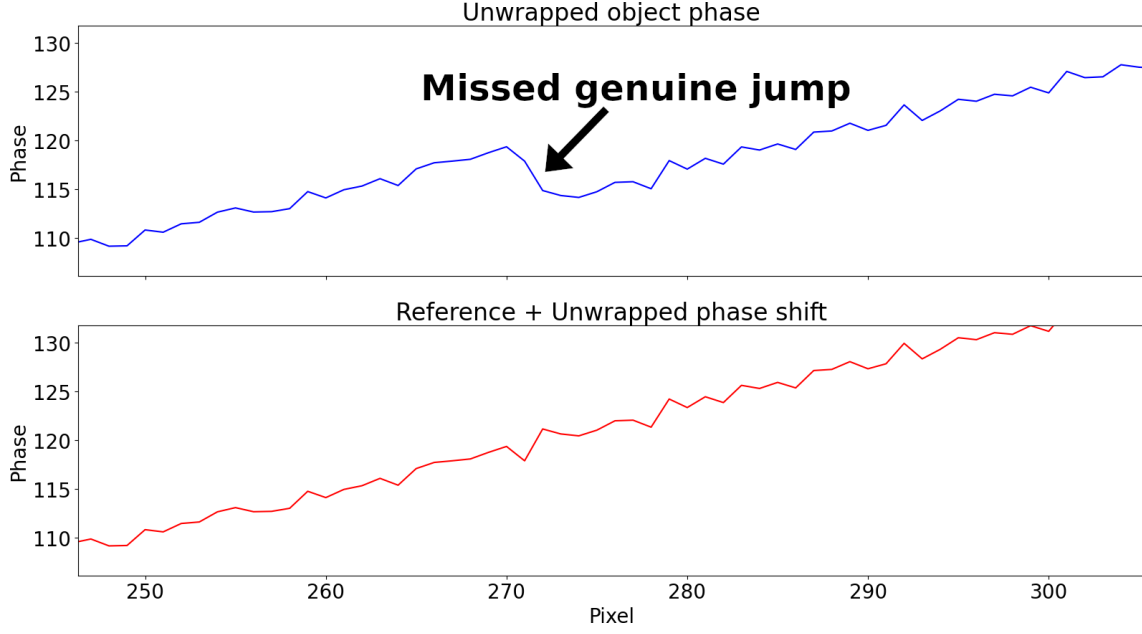


Figure 4.11: Unwrapping of the simulated dent signal with added white noise ($\sigma = 0.6$) using Itoh’s algorithm. Detail of one of the several genuine 2π -jumps missed using Equation (4.9) (top) but not using Equation (4.11) (bottom).

The simulation was repeated using the noise-robust unwrapping algorithm by Estrada et al. [43]. With $\tau = 0.8$, the MAE from the original 3D dent was 7.597 mm ($\sigma = 4.822$ mm) with the direct method and 1.591 mm ($\sigma = 1.339$ mm) with the proposed method.

In this case, the unwrapping was also more effective using the virtual reference. As shown in Figure 4.13, in the first case, Estrada’s algorithm is unable to correctly unwrap the phase, leaving a visible artefact. This does not happen with the proposed method due to a reduced number of 2π -jumps to deal with.

Decreasing to $\tau = 0.7$ for higher noise suppression (or *regularisation*), the reconstruction was free of artefacts in both methods at the price of lower *dynamic range* [43]. The MAEs were 1.241 mm ($\sigma = 1.047$ mm) and 1.221 mm ($\sigma = 1.032$ mm), respectively. Although less evident, the improvement of a virtual reference still applies.

While using Estrada’s noise-robust unwrapping algorithm results in a better reconstruction over Itoh’s algorithm, the proposed virtual reference method visibly improves the final result *independently* from the choice of unwrapping algorithm.

The average running time for the Python implementation was 0.36 s for the direct method and 1.18 s for the proposed method, higher due to the reference image generation and processing.

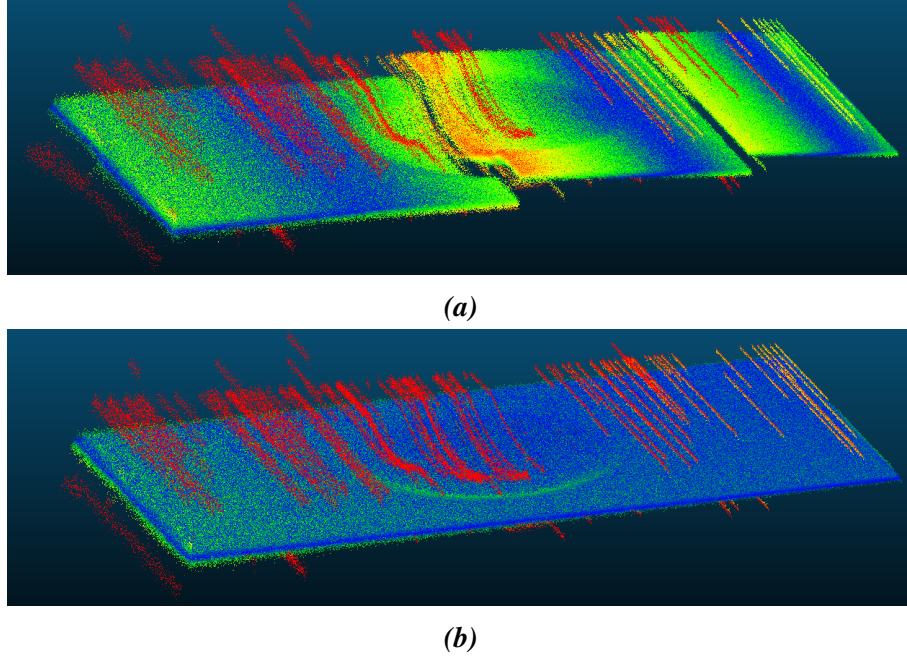


Figure 4.12: Comparison of 3D reconstruction of the simulated dent (as heatmap of MAE) using Itoh's algorithm with the (a) direct mapping and the (b) virtual reference methods.

4.4 Experiments

4.4.1 Static scenario

An inexpensive and lightweight experimental setup was composed of an ELP USB camera (resolution of 1920×1080 pixels at 30 FPS) and an AnyBeam class 1 laser projector (resolution of 1280×720 pixels). To correct lens distortion, three coefficients for the radial distortion and two for the tangential one were used. The system was calibrated using a method similar to [36] and the resulting baseline was $b = 235.040$ mm. Although not indicative of the accuracy over arbitrary objects [38], Table 4.1 shows the RMS reprojection errors.

Camera	Projector	Stereo
0.477	1.231	2.686

Table 4.1: RMS reprojection errors of the first configuration.

Following the dent function defined above (Egg. 4.16), four different dented surface samples were modeled and 3D printed as $15\text{ cm} \times 15\text{ cm}$ tiles with a white matt finish.

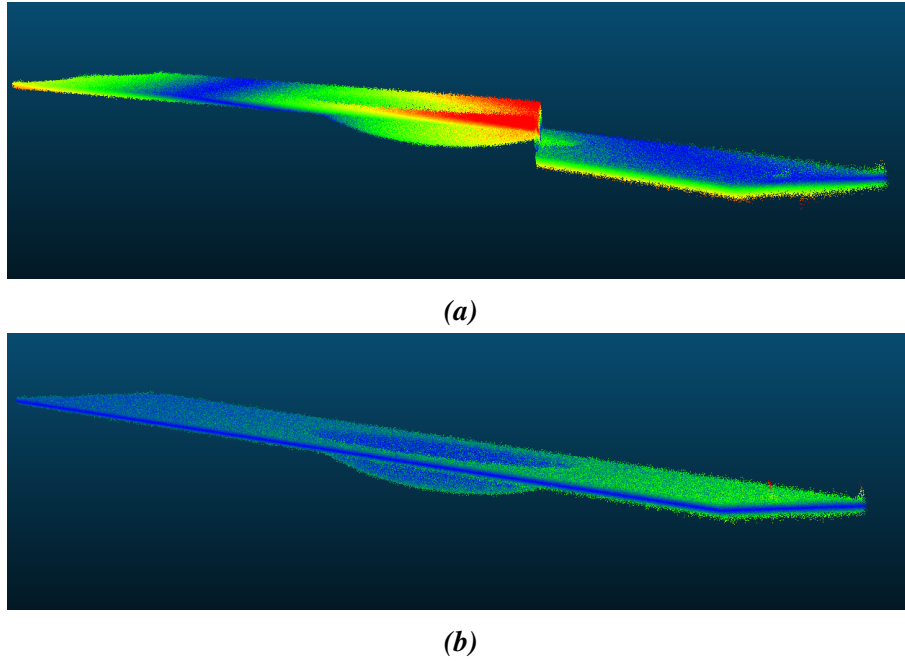


Figure 4.13: Comparison of 3D reconstruction of the simulated dent (as heatmap of MAE) using Estrada's noise-robust algorithm ($\tau = 0.8$) with the (a) direct mapping method and the (b) proposed virtual reference method.

The samples were scanned from a distance of circa 60cm using $T_p = 8$. Images were cropped to 760×640 pixels to select only the tile and remove the background. The characteristics of the samples are reported in Table 4.2. Captured image and calculated virtual reference plane for sample A are showed in Figure 4.14.

Sample	Length (mm)	Width (mm)	Depth (mm)
A	60	40	2
B	120	100	2
C	100	80	3
D	120	80	1

Table 4.2: Dimensions of dent samples.

Figure 4.15 reports the depth and its error along the middle row of sample A. The causes of the shown noise distribution include reflections, laser speckle, sampling, spectral leakage, border artefacts in the Fourier transform, and systematic error of the 3D printing process.

The MAE and its standard deviation are reported for each sample, comparing the direct mapping and the virtual reference methods using 2D Itoh's (Table 4.3) and Estrada's (Table 4.4) unwrapping algorithms.

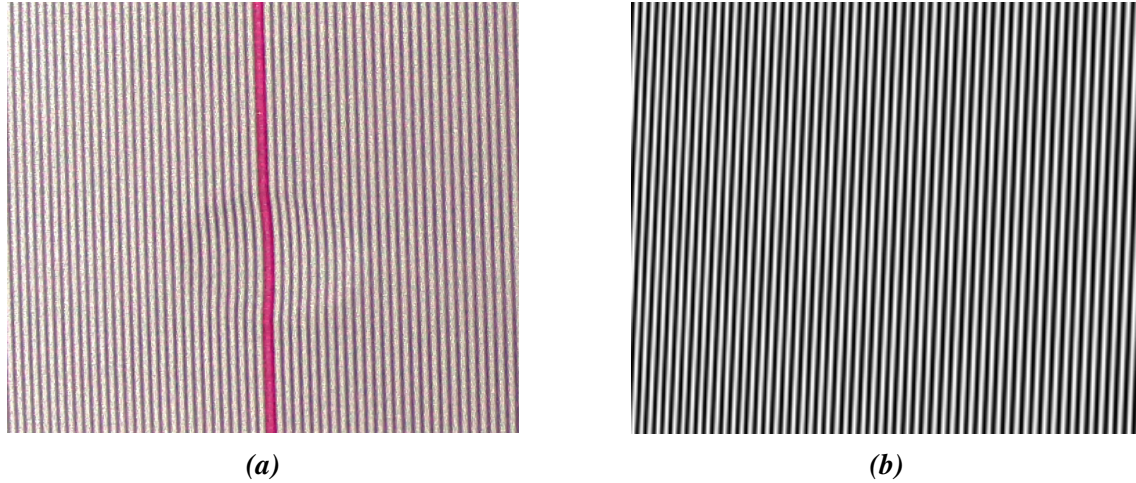


Figure 4.14: Input images for sample A. (a) Captured object image. Red values are converted to greyscale before FFT. (b) Virtual reference image.

Sample	Direct mapping MAE (std) in mm	Proposed MAE (std) in mm	Difference mm (%)
A	0.158 (0.176)	0.126 (0.115)	0.032 (22.5%)
B	0.167 (0.202)	0.144 (0.151)	0.023 (10.1%)
C	0.259 (0.616)	0.154 (0.494)	0.105 (50.8%)
D	0.158 (0.566)	0.144 (0.537)	0.014 (9.3%)

Table 4.3: Compared MAE (and standard deviation) of the direct mapping versus the proposed virtual reference method with 2D Itoh's unwrapping algorithm. Last column reports the absolute and percentage difference.

Independently from the unwrapping algorithm, the proposed method can generally deliver more accurate 3D reconstructions. The difference is visible, for example, in Figure 4.16, showing the sample C reconstruction using Estrada's unwrapping. Border artefacts were responsible for the most of the outliers, accounting for approximately 10% of MAE. In future, the windowed FTP or the wavelet transform could be considered to reduce these errors.

4.4.2 Dynamic scenario

The class 1 laser projector employed above has the advantage of being focus-free, that is, the projected image is always *in focus* at any distance from the object. This technology is made possible by the use of microelectromechanical systems (MEMS) that drive laser beams with a certain refresh rate, independent from camera exposure time.

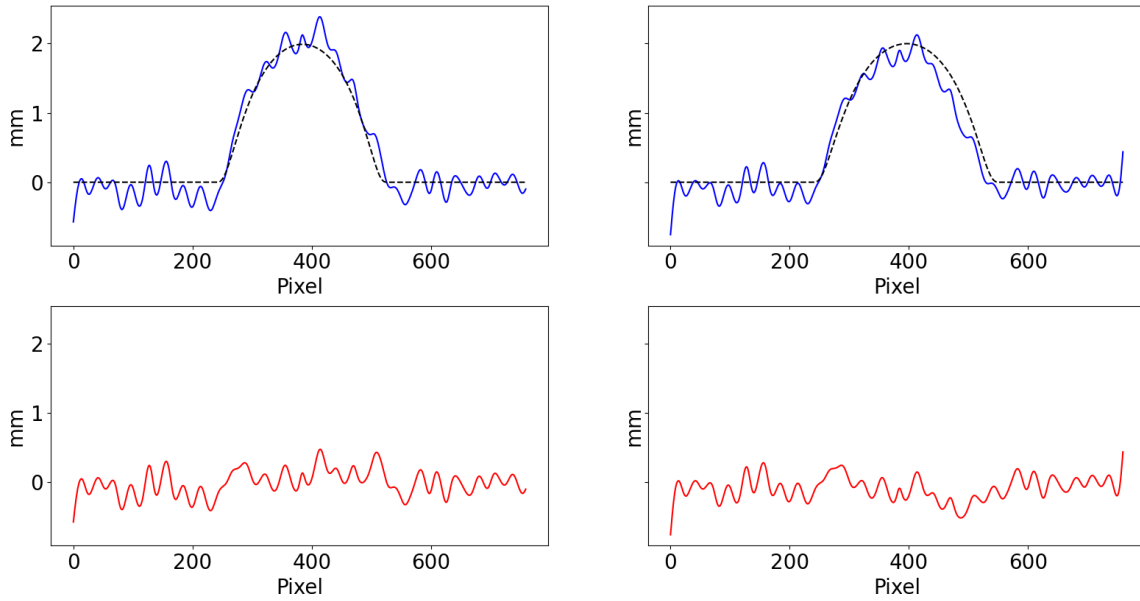


Figure 4.15: Calculated (solid) and ground truth (dashed) depth of sample A (*top*) and error (*bottom*) along the middle section using direct mapping (*left*) and proposed virtual reference method (*right*). Overall MAE was 0.158 mm and 0.126 mm, respectively.

Sample	Direct mapping MAE (std) in mm	Proposed MAE (std) in mm	Difference mm (%)
A	0.153 (0.174)	0.120 (0.110)	0.033 (24.2%)
B	0.165 (0.190)	0.135 (0.130)	0.030 (20.0%)
C	0.202 (0.162)	0.115 (0.114)	0.087 (54.9%)
D	0.213 (0.368)	0.124 (0.131)	0.089 (52.8%)

Table 4.4: Compared MAE (and standard deviation) of the direct mapping versus the proposed virtual reference method with Estrada’s unwrapping algorithm ($\tau = 0.8$). Last column reports the absolute and percentage difference.

This also causes a flickering effect in the camera image, that in the experiments above was compensated by setting high exposure times. Consequently, such a system cannot be used to scan moving objects as is. To maintain the focus-free benefit, the best solution would be to employ a synchronised hardware, so that the camera exposure time corresponds to one projector refresh cycle.

Unfortunately, such technology was not available during this work, thus the projector was replaced with a common LCD one of the same resolution, and the system was recalibrated to test the algorithm over moving objects. Table 4.5 shows the RMS reprojection errors in this second configuration ($b = 266.784$ mm).

Samples were scanned while moving at a speed of about 0.1 m/s along the x -axis at

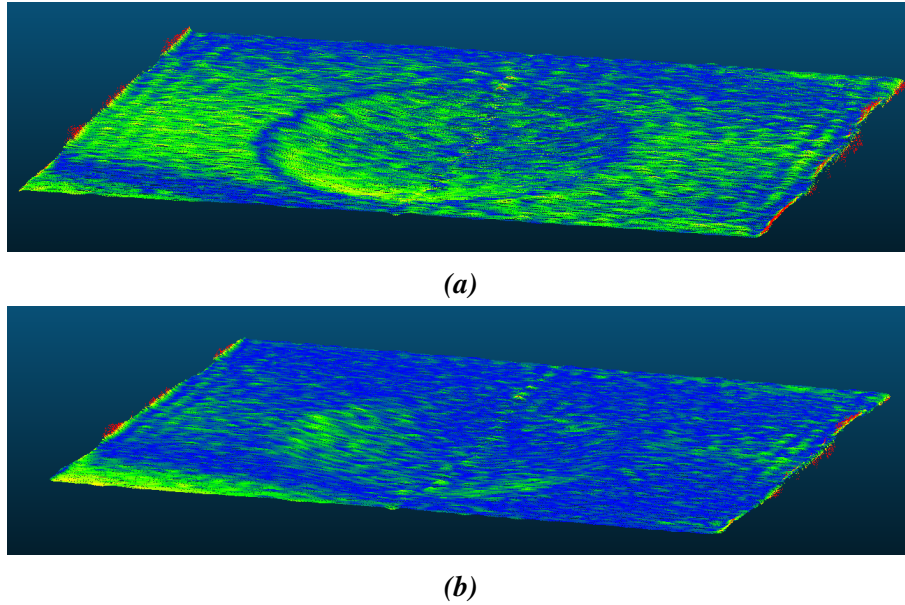


Figure 4.16: Comparison of 3D reconstruction for sample C (as heatmap of MAE) using Estrada's unwrapping algorithm with the (a) direct mapping and the (b) proposed virtual reference methods. Blue areas correspond to lower error than green ones. The border artefacts of the Fourier transform are visible as high-error areas (yellow and red).

Camera	Projector	Stereo
0.282	0.772	2.625

Table 4.5: RMS reprojection error of the second configuration.

a distance of circa 120cm. As before, compared results between the two methods are reported in Table 4.6. This time, only Estrada's unwrapping algorithm was used, and, for reference, all the values from dynamic scenes were compared with the corresponding static ones of the current setup, to highlight the effect of movement. Fig. 4.17 shows the results for sample C.

As expected, dynamic MAE does not consistently increase when compared with the static one, as the single-shot acquisition is not significantly affected by movement and vibration, provided that motion blur is suppressed by a short exposure time with respect to the speed of the object. MAE is much more affected by other factors, such as object albedo and distance, which are independent of movement.

Overall, if compared with the first configuration, the reduced performance is due to the change of projector and the increased distance. In addition, for this setup, the use of the phase shift, calculated with respect to the virtual reference, performed better than the direct mapping in all the cases. On average, the MAE difference was 0.138mm, even

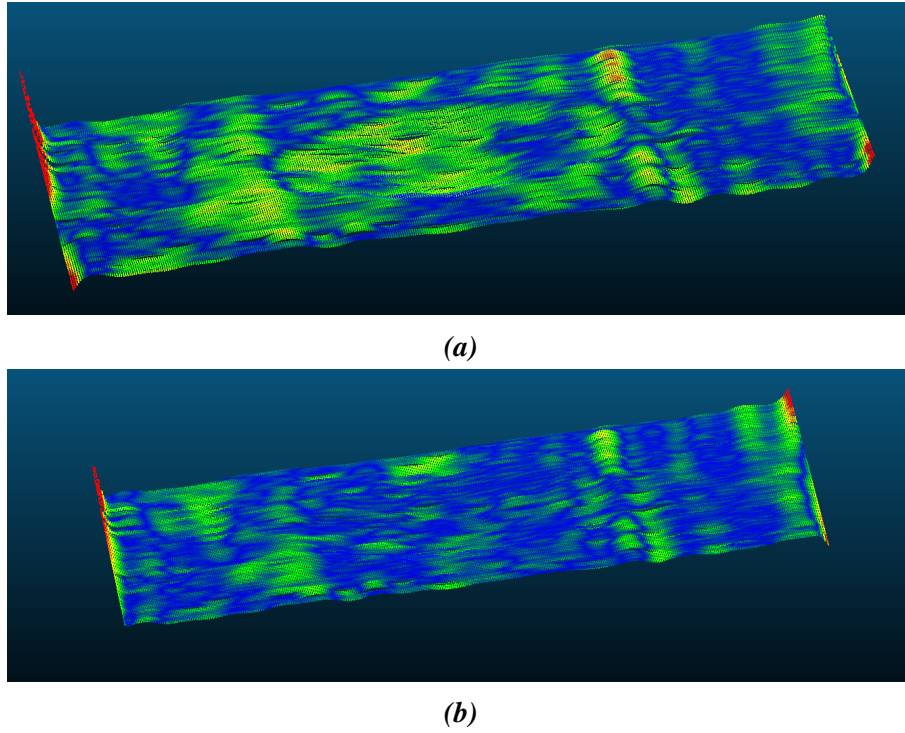


Figure 4.17: Comparison of 3D reconstruction for sample C acquired during movement using Estrada's unwrapping algorithm with the (a) direct mapping and the (b) proposed virtual reference methods. Blue areas correspond to lower error than green ones. The border artefacts of the Fourier transform are visible as high-error areas (yellow and red).

Sample	Type	Direct mapping MAE (std) in mm	Proposed MAE (std) in mm
A	Dynamic	0.381 (0.456)	0.280 (0.307)
	Static	0.504 (0.486)	0.285 (0.217)
B	Dynamic	0.419 (0.438)	0.309 (0.396)
	Static	0.432 (0.329)	0.287 (0.255)
C	Dynamic	0.436 (0.443)	0.271 (0.265)
	Static	0.493 (0.546)	0.320 (0.468)
D	Dynamic	0.411 (0.394)	0.295 (0.294)
	Static	0.370 (0.357)	0.293 (0.401)

Table 4.6: Compared MAE (and standard deviation) of the direct mapping versus the proposed virtual reference method with Estrada's unwrapping algorithm ($\tau = 0.8$) over both static and dynamic samples.

more relevant than in the first configuration.

4.5 Discussion

To date, most of the UAV-based systems developed for aircraft inspections make use of a monocular camera as a primary acquisition device and are thus incapable of detecting shallow dents or collecting measures [11]. As the acquisition of measures is essential for damage evaluation [7], 3D scanning must be considered for proper automation. However, the employment of 3D scanning technology on UAVs is not straightforward: high-accuracy 3D scanners are generally based on multiple-shot algorithms, as the projection of many different patterns allows for detailed codification, and are incompatible with fast movements and vibration that would be experienced while scanning from a flying drone.

This work showed that the FTP should be considered to automate aircraft dent inspections. The single-shot nature of the method, together with its high phase sensitivity to depth variations, constitutes an appropriate solution for the integration with UAVs, ultimately providing fast and effective inspections. Table 4.7 summarises existing dent inspection approaches, comparing their pros and cons against the proposed one.

Approach	Pros	Cons
Traditional	No special hardware requirements Straightforward	Hazardous Subjective output Time-consuming
Handheld 3D scanner	High accuracy Repeatable	Hazardous Time-consuming
Monocular camera UAV	Fast Repeatable Safe	No measures Shallow dents not detectable
UAV equipped with FTP	Fast Repeatable Safe Measures collected	Complex light control Smoothness assumption

Table 4.7: *Detection and measures of dents in MRO. Summary of pros and cons for different inspection approaches.*

FTP is naturally resilient to movement and vibration, although codification information must be condensed into a single pattern, implying local smoothness. While the local smoothness assumption is true for most of the aircraft skin, the major challenge for effective application in the field is the increase of signal-to-noise ratio, especially from

longer distances. This translates into overcoming surface albedo and employing a projection technology capable of operating in common indoor or even outdoor light conditions, generally not dark enough.

The use of light sources outside the visible spectrum or higher class lasers might be considered. However, while the use of a laser projector highly improves the contrast of the projected pattern, it normally introduces the speckle phenomenon, presented as a granular pattern over the camera image that affects the filtered phase and, consequently, produces noise in the final measures. Speckle can be corrected with different types of *despecklers* [44].

4.6 Conclusions

The aviation industry is increasingly making use of UAVs for aircraft inspections. To date, a UAV-based solution for the detection and measurement of dents has yet to be found, as high-accuracy 3D scanners usually work with multiple-shot algorithms and are thus incompatible with fast movements and vibration.

Triangular stereo Fourier transform profilometry (FTP) was proposed here to evaluate its potential towards automated aircraft dent inspections. Additionally, two modifications to the FTP algorithm were introduced. First, the automatic band-pass estimation was presented via a *virtual* reference plane, thus eliminating the user interaction in the filtering process. Secondly, after showing that the interaction of noise over 2π -jumps is lower using the phase shift, a *virtual* reference image was employed to increase accuracy.

Overall, the proposed method was able to faithfully reconstruct the 3D point clouds of demonstrative dents at distances of 60cm and 120cm, delivering submillimetre accuracy despite the use of unsophisticated hardware. Simulations and experiments with 3D-printed dent samples showed that the employment of the virtual reference delivers better results, independent of the choice of unwrapping algorithm.

With additional research, an FTP-based system could be integrated with a UAV, enabling quick and effective dent inspection. Solutions to deal with light conditions in common operational environments are to be found. It must also be noted that a successful 3D data acquisition is just the first step for comprehensive automation, as the difficulties in detecting dents in the physical world are somehow translated into the digital world. As such, an end-to-end automatic system for dent detection and measurement remains an

open challenge in MRO.

References

- [1] Kara Latorella and Prasad Prabhu. “A review of human error in aviation maintenance and inspection”. In: *International Journal of Industrial Ergonomics* 14653 (2000). DOI: 10.1016/S0169-8141(99)00063-3.
- [2] Umberto Papa and Salvatore Ponte. “Preliminary design of an unmanned aircraft system for aircraft general visual inspection”. In: *Electronics* (2018).
- [3] *Airbus launches advanced indoor inspection drone*. 2018. URL: <https://www.airbus.com/en/newsroom/press-releases/2018-04-airbus-launches-advanced-indoor-inspection-drone-to-reduce-aircraft> (visited on 2021-12-19).
- [4] *Drones Finally Approved For Lightning Checks*. 2021. URL: <https://aviationweek.com/mro/drones-finally-approved-lightning-checks-dent-measurement-next> (visited on 2021-12-19).
- [5] Pasquale Lafiosca and Ip-Shing Fan. “Review of non-contact methods for automated aircraft inspections”. In: *Insight-Non-Destructive Testing and Condition Monitoring* (2020).
- [6] Shakeb Deane, Nicolas P Avdelidis, Clemente Ibarra-Castanedo, et al. “Application of NDT thermographic imaging of aerospace structures”. In: *Infrared Physics & Technology* (2019). DOI: 10.1016/j.infrared.2019.02.002.
- [7] The Boeing Company. *Structural repair manual B737-400*. 2015.
- [8] Civil Aviation Authority. *CAP716 - Aviation Maintenance Human Factors*. 2003.
- [9] Civil Aviation Authority. *Paper 2013/03 - Reliability of Damage Detection in Advanced Composite Aircraft Structures*. United Kingdom: Civil Aviation Authority, 2013.
- [10] Igor Jovančević, Huy-Hieu Pham, Jean-José Orteu, et al. “3D point cloud analysis for detection and characterization of defects on airplane exterior surface”. In: *Journal of Nondestructive Evaluation* (2017).
- [11] Anil Doğru, Soufiane Bouarfa, Ridwan Arizar, and Reyhan Aydoğan. “Using convolutional neural networks to automate aircraft maintenance visual inspection”. In: *Aerospace* (2020). DOI: 10.3390/aerospace7120171.

- [12] Jens Guehring. “Dense 3-D surface acquisition by structured light using off-the-shelf components”. In: *Proceedings of SPIE - The International Society for Optical Engineering* (2001). DOI: 10.1117/12.410877.
- [13] Joaquim Salvi, Jordi Pages, and Joan Batlle. “Pattern codification strategies in structure light systems”. In: *Pattern Recognition* (2004). DOI: 10.1016/j.patcog.2003.10.002.
- [14] Chao Zuo, Tianyang Tao, Shijie Feng, et al. “Micro Fourier Transform Profilometry (FTP): 3D shape measurement at 10,000 frames per second”. In: *Optics and Lasers in Engineering* (2017). DOI: 10.1016/j.optlaseng.2017.10.013.
- [15] Yatong An, Jae-Sang Hyun, and Song Zhang. “Pixel-wise absolute phase unwrapping using geometric constraints of structured light system”. In: *Optics Express* (2016). DOI: 10.1364/OE.24.018445.
- [16] Yang Chen, Baoan Song, Renji He, Hailing Hu, and Simao Chen. “Fast two-step phase-shifting method for measuring the three-dimensional contour of objects”. In: *Optical Engineering* (2021). DOI: 10.1117/1.OE.60.9.094104.
- [17] Zhonghua Zhang. “Review of single-shot 3D shape measurement by phase calculation-based fringe projection techniques”. In: *Optics and Lasers in Engineering* (2012). URL: <https://doi.org/10.1016/j.optlaseng.2012.01.007>.
- [18] Heinz Hügli and Gilbert Maître. “Generation And Use Of Color Pseudo Random Sequences For Coding Structured Light In Active Ranging”. In: *Proceedings of SPIE - The International Society for Optical Engineering* (1989-02). DOI: 10.1117/12.949215.
- [19] P. Vuylsteke and A. Oosterlinck. “Range image acquisition with a single binary-encoded light pattern”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1990). DOI: 10.1109/34.44402.
- [20] Ali Ulusoy, Fatih Calakli, and Gabriel Taubin. “One-shot scanning using De Bruijn spaced grids”. In: *IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*. 2009. DOI: 10.1109/ICCVW.2009.5457499.
- [21] Li Zhang, Brian Curless, and Steven Seitz. “Rapid Shape Acquisition Using Color Structured Light and Multi-pass Dynamic Programming”. In: *Proceedings. First International Symposium on 3D Data Processing Visualization and Transmission*. 2002. DOI: 10.1109/TDPVT.2002.1024035.
- [22] Mitsuo Takeda and Kazuhiro Mutoh. “Fourier transform profilometry for the automatic measurement of 3-D object shapes”. In: *Applied Optics* (1983). DOI: 10.1364/AO.22.003977.

- [23] Joaquim Salvi, Sergio Fernandez, Tomislav Pribanic, and Xavier Llado. “A state of the art in structured light patterns for surface profilometry”. In: *Pattern Recognition* 43.8 (2010). ISSN: 0031-3203. URL: <https://doi.org/10.1016/j.patcog.2010.03.004>.
- [24] Qian Kemao, Haixia Wang, and Gao Wenjing. “Windowed Fourier transform for fringe pattern analysis: Theoretical analyses”. In: *Applied optics* 47 (2008). DOI: 10.1364/AO.47.005408.
- [25] Lei HUANG, Qian Kemao, Bing Pan, and Anand Asundi. “Comparison of Fourier transform, windowed Fourier transform, and wavelet transform methods for phase extraction from a single fringe pattern in fringe projection profilometry”. In: *Optics and Lasers in Engineering* (2010). DOI: 10.1016/j.optlaseng.2009.04.003.
- [26] Lionel Watkins. “Phase Recovery Using The Wavelet Transform”. In: *AIP Conference Proceedings* (2010). DOI: 10.1063/1.3426093.
- [27] Jingang Zhong and Huiping Zeng. “Multiscale windowed Fourier transform for phase extraction of fringe patterns”. In: *Appl. Opt.* (2007). DOI: 10.1364/AO.46.002670.
- [28] Sergio Fernandez, M. Gdeisat, J. Salvi, and D. Burton. “Automatic window size selection in Windowed Fourier Transform for 3D reconstruction using adapted mother wavelets”. In: *Optics Communications* (2011). DOI: 10.1016/j.optcom.2011.01.068.
- [29] Jing Xu and Song Zhang. “Status, challenges, and future perspectives of fringe projection profilometry”. In: *Optics and Lasers in Engineering* (2020). DOI: 10.1016/j.optlaseng.2020.106193.
- [30] Shijie Feng, Chao Zuo, Liang Zhang, et al. “Calibration of fringe projection profilometry: A comparative review”. In: *Optics and Lasers in Engineering* (2021). DOI: 10.1016/j.optlaseng.2021.106622.
- [31] Song Zhang and Peisen Huang. “Novel method for structured light system calibration”. In: *Optical Engineering* (2006). DOI: 10.1117/1.2336196.
- [32] Beiwen Li and Song Zhang. “Structured light system calibration method with optimal fringe angle”. In: *Applied optics* (2014). DOI: 10.1364/AO.53.007942.
- [33] Yongfu Wen, Sikun Li, Haobo Cheng, Xianyu Su, and Qican Zhang. “Universal calculation formula and calibration method in Fourier transform profilometry”. In: *Applied optics* (2010). DOI: 10.1364/AO.49.006563.

- [34] Zhengyou Zhang. “A Flexible New Technique for Camera Calibration”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* (2000). DOI: 10.1109/34.888718.
- [35] Song Zhang. “High-speed 3D shape measurement with structured light methods: A review”. In: *Optics and Lasers in Engineering* (2018). DOI: 10.1016/j.optlaseng.2018.02.017.
- [36] Zhongwei Li, Yusheng Shi, Congjun Wang, and Yuanyuan Wang. “Accurate calibration method for a structured light system”. In: *Optical Engineering* (2008). DOI: 10.1117/1.2931517.
- [37] Ivan Martynov, Joni-Kristian Kamarainen, and Lasse Lensu. “Projector Calibration by “Inverse Camera Calibration””. In: *Lecture Notes in Computer Science*. 2011. DOI: 10.1007/978-3-642-21227-7_50.
- [38] Daniel Moreno and Gabriel Taubin. “Simple, Accurate, and Robust Projector-Camera Calibration”. In: *Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*. 2012. DOI: 10.1109/3DIMPVT.2012.77.
- [39] Munther Gdeisat and Francis Lilley. *One-Dimensional Phase Unwrapping Problem*. 2011. URL: https://www.ljmu.ac.uk/~media/files/ljmu/about-us/faculties-and-schools/fet/geri/onedimensionalphaseunwrapping_finalpdf.pdf (visited on 2021-11-17).
- [40] Song Zhang. “Absolute phase retrieval methods for digital fringe projection profilometry: A review”. In: *Optics and Lasers in Engineering* 107 (2018). DOI: 10.1016/j.optlaseng.2018.03.003.
- [41] Kazuyoshi Itoh. “Analysis of the phase unwrapping algorithm”. In: *Appl. Opt.* (1982). DOI: 10.1364/AO.21.002470.
- [42] Fraser Dickin, Stephen Pollard, and Guy Adams. “Mapping and correcting the distortion of 3D structured light scanners”. In: *Precision Engineering* (2021). DOI: 10.1016/j.precisioneng.2021.06.001.
- [43] Julio C. Estrada, Manuel Servin, and Juan A. Quiroga. “Noise robust linear dynamic system for phase unwrapping and smoothing”. In: *Opt. Express* (2011). DOI: 10.1364/OE.19.005126.
- [44] Igor Kompanets and Nikolay Zalyapin. “Methods and Devices of Speckle-Noise Suppression”. In: *Optics and Photonics Journal* (2020).

Chapter 5

Automatic Segmentation of Aircraft Dents in Point Clouds

Published as P. Lafiosca, I.S. Fan, N.P. Avdelidis in: *SAE Technical Paper - AeroTech*, doi:10.4271/2022-01-0022.

Abstract

Dents on the aircraft skin are frequent and may easily go undetected during airworthiness checks, as their inspection process is tedious and extremely subject to human factors and environmental conditions. Nowadays, 3D scanning technologies are being proposed for more reliable, human-independent measurements, yet the process of inspection and reporting remains laborious and time consuming because data acquisition and validation are still carried out by the engineer. For full automation of dent inspection, the acquired point cloud data must be analysed via a reliable segmentation algorithm, releasing humans from the search and evaluation of damage. This paper reports on two developments towards automated dent inspection. The first is a method to generate a synthetic dataset of dented surfaces to train a fully convolutional neural network. The training of machine learning algorithms needs a substantial volume of dent data, which is not readily available. Dents are thus simulated in random positions and shapes, within criteria and definitions of a Boeing 737 structural repair manual. The noise distribution from the scanning apparatus is then added to reflect the complete process of 3D point acquisition on the training. The second proposition is a surface fitting strategy to convert 3D point clouds to 2.5D. This allows higher resolution point clouds to be processed with a small

amount of memory compared with state-of-the-art methods involving 3D sampling approaches. Simulations with available ground truth data show that the proposed technique reaches an intersection-over-union of over 80%. Experiments over dent samples prove an effective detection of dents with a speed of over 500 000 points per second.

Keywords: Aircraft inspection; Fully convolutional network; Point cloud segmentation.

5.1 Introduction

Aircraft visual inspections are costly, time-consuming and sometimes hazardous procedures, yet essential for the airworthiness assessment. Research also reveals the subjective nature of engineers' evaluations [1, 2]. Machine vision, on the other hand, is progressing at an incredible rate, surpassing humans in many occasions [3], with great potential to increase both efficiency and efficacy of inspections.

Nevertheless, the automation of aircraft visual inspections is still very limited. Automation of dent inspections, in particular, is challenging due to the nature of the damage itself. A dent is smooth, without clear boundaries and difficult to identify even by trained engineers. In addition, its evaluation demands to collect measures as prescribed by the structural repair manual (SRM). Hence, 2D pattern recognition of monocular images is useless for the scope, because metric information is absent.

Even if it solves a part of the problem producing metric data, the use of 3D scanning technologies leaves the operator with a substantial amount of raw data, or *point clouds*, from which information has to be extracted manually. As such, the correct labeling of each point as belonging to a dent or not, also known as *semantic segmentation*, becomes essential for comprehensive automation. With the segmentation strategy here proposed, dent damage can be automatically found in 3D data points, so that additional algorithms may be able to measure and evaluate its dimensions and other characteristics with respect to the rules dictated by each aircraft SRM.

While the use of *ad hoc* rule-based algorithms is anachronistic, the actual application of machine learning technologies is not straightforward, mainly for two reasons. First, the lack of a publicly accessible, *labelled* and substantial dent dataset required for the training and validation. Secondly, because dent segmentation is very different from the segmentation of general objects, as local features are not particularly distinctive over the aircraft skin.

This study addresses the problem of semantic segmentation in aircraft dent inspections, proposing two concepts. First, an innovative process to build a *virtual* dataset: dented surfaces are randomly generated according to the model explained below and mixed with *noise signals* extracted from the chosen 3D scanning apparatus, adding robustness during training. To the best of our knowledge, no similar approach was found in the literature to address the lack of real data.

The dataset generated in such way is accompanied by its *ground truth data*, i.e. the true classification for each one of the points, labelled as dented or not dented. Being virtually generated, this classification is exact and not subject to measuring or evaluation errors, and provides the basis for error backpropagation during the training of a machine learning algorithm. Secondly, the problem dimensionality is reduced by the employment of an effective surface fitting strategy, where 3D points are reduced to 2D by projection onto the fitted bivariate quadric surface. Then, for each point is possible to calculate a residual or distance from the fitted surface. The 2D residuals are then fed to a fully convolutional network (FCN) inspired from U-Net [4] for quick, lightweight and accurate surface segmentation.

Simulations with available ground truth data show that the proposed segmentation technique reaches an intersection-over-union (IoU) (or *Jaccard index*) of over 80%. Experiments over dent samples prove the effective segmentation of dents with a speed of over 500000 points per second.

5.2 Previous work

The general problem of segmentation has been extensively discussed in the literature, while only a few articles dealing with dent segmentation. Methods can be classified as 2D or 3D, according to their input dimensionality, and considering whether they employ machine learning or *ad hoc* algorithms based on handcrafted features and rules [5].

Among the latter, region growing has been applied to both 2D and 3D data and uses criteria to combine features of local points, merging them together if they have similar properties. Jovančević et al. [6] proposed a method to detect and characterise aircraft dents based on the estimation of normal and curvature values. The defect is characterised after 2D projection and further processing. As point-wise calculation is generally slow, the processing time reported was 20s for 30000 points, and up to 120s for denser point clouds (on a PC with 2.4 GHz Core i7 CPU, 8GB RAM). Other non-AI approaches are,

for example, clustering and model fitting [5]. For all non-AI methods, preprocessing parameters and thresholds affect the algorithm sensitivity and must be tuned accordingly for each scenario. Since the user interaction is continuously required, no proper automation can be achieved. As such, recent research is mostly oriented towards machine learning.

In [7] a Mask R-CNN [8] was used to detect and segment pictures of aircraft dents with average IoU of 36.54%. The use of R-CNN was also tested for the sole detection of car dents [9]. In this case light bands were projected on the surface to highlight its deformations. Although methods processing 2D images via CNNs are the *de facto* standard for image classification, object detection and segmentation [4], they come with two main problems when applied to aircraft inspections. First, *measures* cannot generally be provided as absent in the original input, usually consisting in monocular images. Since measures are the discriminating factor for evaluating defects as allowable damage or not, as prescribed by the SRM, the usefulness of doing an inspection without acquiring them is questionable. Secondly, while scratches and lightning may be detected reasonably well, the lack of texture and distinctive features over the single-coloured aircraft skin makes the detection of a dent particularly challenging. Therefore, methods acquiring and processing 3D metric data are more interesting for the scope.

Convolution cannot be applied directly to unordered point clouds [5]. The most straightforward way to apply 2D CNN to 3D data are projection-based approaches. While these allow seamlessly to transfer the same technology to 3D data, projecting in 2D generally causes loss of information and the results are highly dependent on the chosen points of view. Alternatively, *voxels*¹ address the need of having an ordered structure to apply convolution, creating low resolution, equally-spaced data points. These strategies organise all the volume, including empty space, into some structure [10, 11], resulting in high computational cost and memory requirements, often beyond the availability of the average user. A third and innovative approach is to directly process the point coordinates. PointNet [12] pioneered the use of a symmetric function achieving permutation invariance. Also, the input size is often predefined because the data is processed via multilayer perceptron (MLP) layers. As PointNet is not effective at capturing local features, its derivations introduce expensive preprocessing and neighbouring operations [13, 14]. A substantial number of architectures based on a combination of the above principles is present in the literature, summarised in [5, 15], however, a general approach that outperforms the others is yet to be found.

¹A voxels represents an element of an equally-spaced three dimensional grid in the space. It may also be considered as the 3D version of a pixel.

5.3 DentNet segmentation strategy

The B737-400 SRM defines a dent as “a damaged area that is pushed in from its normal contour” and has “smooth” edges [16]. However, no formal shape definition is given. For the scope of this work, a pseudo-realistic synthetic dataset of surfaces is created: dents are simulated in random positions and shapes, following an analytical model within the above SRM criteria and definitions. To increase robustness, the noise distribution from the scanning apparatus is then acquired and added over the synthetic data during the training of the FCN.

The network can thus be used for segmentation of real 3D scan data. The proposed method exploits the nature of surface inspection to deal with the sparsity of point cloud data, making use of a 2.5D strategy. This allows higher resolution point clouds to be processed with a small amount of memory compared with state-of-the-art methods.

5.3.1 Virtual dataset

As anticipated above, the virtual dataset generation is necessary to overcome the lack of real data. Curved surfaces are considered with shapes compatible with most of the aircraft areas. These are then virtually dented.

The world xy -dimensions of the surface are fixed as $X \times Y$ (metric units), sampled over a grid of $w \times h$ pixels. For each sample, xy -values are fed to a parabola function:

$$z_s = \alpha x^2 + \beta y^2 \quad (5.1)$$

with α and β randomly chosen within a certain range, compatible with the curvature of the aircraft surfaces to be inspected. Later, noise from a uniform distribution is added to xy -values.

A dent is modeled by:

$$z_d = \begin{cases} -e^{-\frac{1}{1-r^2}} & \text{if } |r| < 1 \\ 0 & \text{elsewhere} \end{cases} \quad (5.2)$$

where $r = \sqrt{x^2 + y^2}$. The function is randomly rescaled in x and y to cover different dent dimensions and proportions, rotated and shifted with respect to the surface. A dent is

added to the surface with a certain probability p . Further dents are added to the same surface with exponential decreasing probability.

The depth of the sample is finally obtained as:

$$z = z_s + \max \left(z_d^{(1)}, \dots, z_d^{(n)} \right) + z_w \quad (5.3)$$

where n is the total number of dents and z_w represents a Gaussian white noise distribution, added for extra robustness during training. xy -values and z -values are joined together to form the final dented surface, which is then randomly rotated along all the three world axes, covering the possible angles of a real scanning scenario. The surface here generated is thus represented by 3D coordinates, still ordered into a $w \times h$ matrix that preserves relative proximity of points.

5.3.2 Noise signals

The noise distribution when scanning real samples is not necessarily Gaussian and depends on the selected scanning apparatus and its 3D reconstruction algorithm. The noise may include frequency filtering and quantisation effects, imperfect calibration and distortion correction of both camera and projector, light scattering, laser speckle and other unwanted effects. Although mixing virtual and real data for training is not a new concept in literature [17, 18], the innovative procedure here proposed consists in mixing virtual data with only *noise signals* from real world.

The signals are obtained scanning flat a white board several times and in different positions with the 3D scanner. In this case a system based on Fourier transform profilometry [19, 20] was chosen. An example of its noise distribution over a flat board about 40cm wide is shown in Fig. 5.1.

The acquired signals are augmented via random cropping and flipping, and then added to the z value of Eq. (5.3) before proceeding with preprocessing and training. An example of resulting dented surface is shown in Fig. 5.2.

The above strategy allows to produce more plausible data (accompanied by its ground truth), which is used to train a neural network that can eventually translate its knowledge to real dent samples. Compared to the addition of only Gaussian noise, this strategy leads to a considerable improvement of the predictions when scanning real samples with the same apparatus, as confirmed by the following experiments.

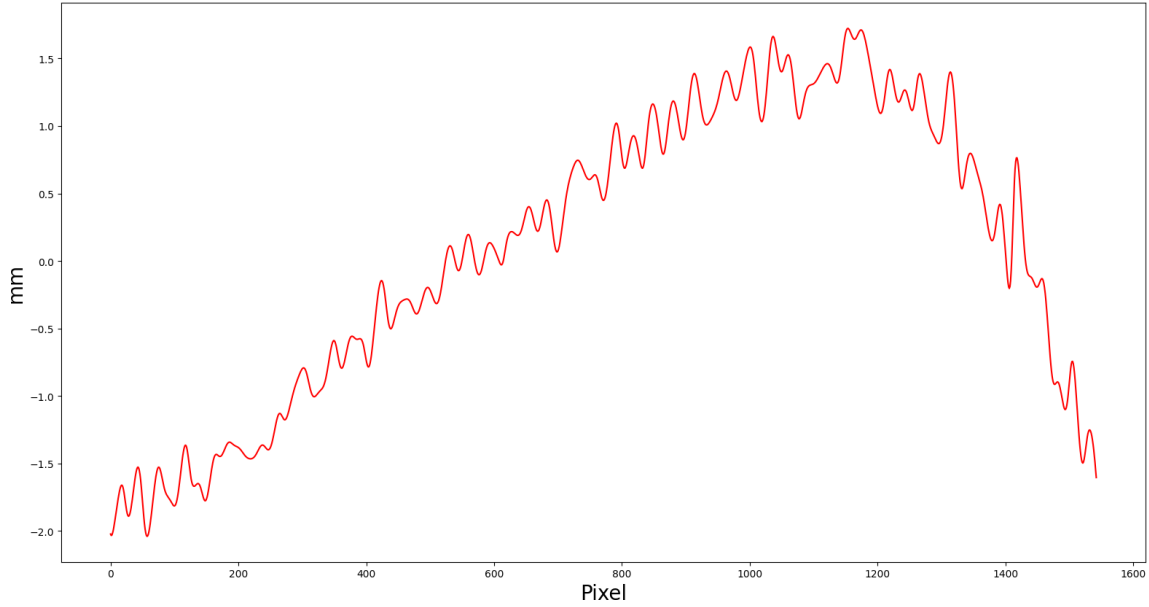


Figure 5.1: Example of noise distribution of the scanning apparatus.

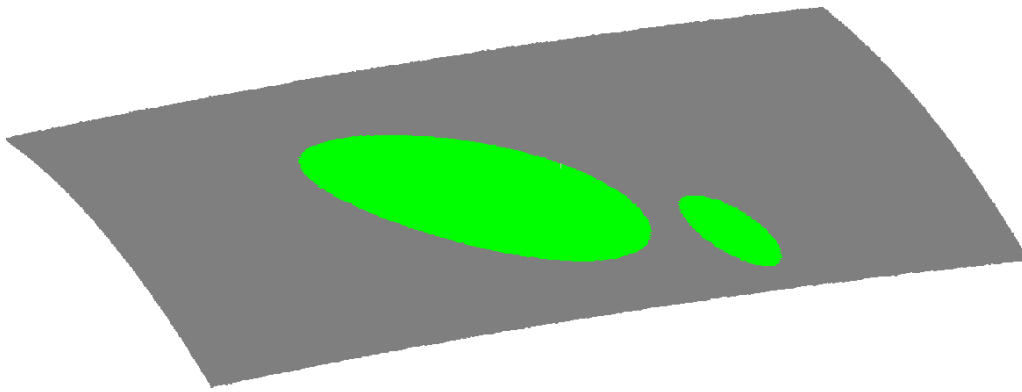


Figure 5.2: Virtual dented surface with ground truth (dents in green).

5.3.3 Input preprocessing

Input 3D data consist in a point cloud, i.e. a list of xyz coordinates derived from any 3D scanning device. For quick and effective segmentation, a bivariate quadratic function is first fitted to the 3D points, then the residuals are calculated as distances along z -axis, *de facto* providing 2D data that can be used as input to a FCN. The data preprocessing pipeline consists in:

1. Canonical rotation of the surface;
2. Bivariate quadratic function fitting;
3. Extraction of the 2D matrix of residuals.

First, the surface is oriented so that its average normal direction $\hat{n}$ is parallel the world

direction $\hat{z}$ (coincident with the scanner optical axis). Given that the scanned objects are rather wide surfaces, $\hat{n}$ is calculated by simply fitting a plane. This expedient ensures that the depth information lays on the new z -axis, that later will be mapped to an intensity value over an conventional greyscale image. Thus, the rotation to be applied to the original point cloud is found as:

$$\mathbf{R} = \mathbb{I} + [\hat{z} \times \hat{n}]_{\times} + [\hat{z} \times \hat{n}]_{\times}^2 \frac{1}{1 + \hat{z} \cdot \hat{n}} \quad (5.4)$$

where $\mathbb{I}$ is the 3×3 identity matrix and $[\]_{\times}$ indicates the cross-product skew-symmetric matrix. In our setup is always $\hat{z} \cdot \hat{n} \neq -1$.

Then, a bivariate quadratic function expressed as:

$$z(x, y) = a + bx + cy + dx^2 + exy + dy^2 \quad (5.5)$$

is fitted to the surface. $z(x, y)$ is general enough to represent the shape of most of the aircraft skin, *neglecting* the potential dents there present and following its baseline. The residuals (or distances from the fitted surface) of this fitting are thus used as input for the FCN in the form of a $w \times h$ matrix. In such way dents will appear as consistent high-residual areas, which the network should learn to segment comparing its predictions with the ground truth and ignoring noise (Fig. 5.3).

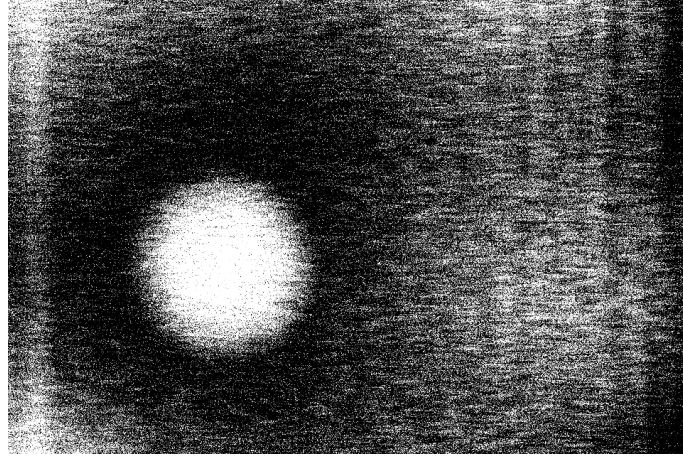


Figure 5.3: Residual image as a result of the pre-processing step, allowing depth information to be quantised as an intensity value and then to be fed to the machine learning algorithm.

Dents do not have distinctive characteristics and the conversion of 3D points to distances from the quadratic function is essential to highlight deformations that the FCN can discriminate. Furthermore, point data is reduced to 2D for quick and lightweight processing. No expensive neighbouring searching mechanism is used and the preprocessing is

reasonably fast, taking on average 0.4 s for 614 400 points (Intel i7-7500U CPU, 12 GB of RAM).

5.3.4 Network architecture

U-Net is a FCN that was developed for medical image segmentation [4]. It is composed by a downsampling part that converts the input image to a compressed set of features, similarly to what is done for image classification where these are then linked to labels. However, with U-Net these features are not linked to labels but fed to an upsampling part, symmetric to the previous one, that expands the dimensions back to the original ones, delivering a label for each input value (segmentation).

The proposed network is fully convolutional, inspired from U-Net and with a similar number of trainable parameters, although with some key differences. The encoder part consists in the repeated application of two convolution operations with kernel size 3×3 and padding size 1. The first, identified with *Conv*, has stride 1. The second one, identified with *DownConv*, has stride 2: it halves the dimensions while doubling the number of channels. The decoder part has repeated application of *Conv* followed by a transposed convolution layer, or *UpConv*, having kernel size 3×3 and both padding and output padding set to 1, doubling the dimensions and halving the number of channels.

Each convolution is followed by a batch normalisation (BN), that enables faster and stable training [21], and a rectifier linear unit (ReLU) as activation function. In the output layer, the latter is replaced by a sigmoid. The addition of a dropout layer after each *Conv* was found to impair performance, thus it was omitted. For better localisation, a skip connection is introduced after each *Conv* in the encoder to increase localisation capabilities. The number of channels is reduced via a 1×1 convolution and the features are concatenated at the same level in the decoder part. The network diagram is shown in Fig. 5.4.

The training is performed by using binary cross entropy (BCE) as loss function. However, the input data may be *unbalanced*: dented areas (positive) are generally much less than non-dented areas (negative). Thus, a balancing weight is estimated from the input data and applied during loss calculation, increasing recall. The network can accept any input size with dimensions multiple of 16, nevertheless, the point *density* (points per mm) of the test set is expected to be roughly the same as the one in the training set.

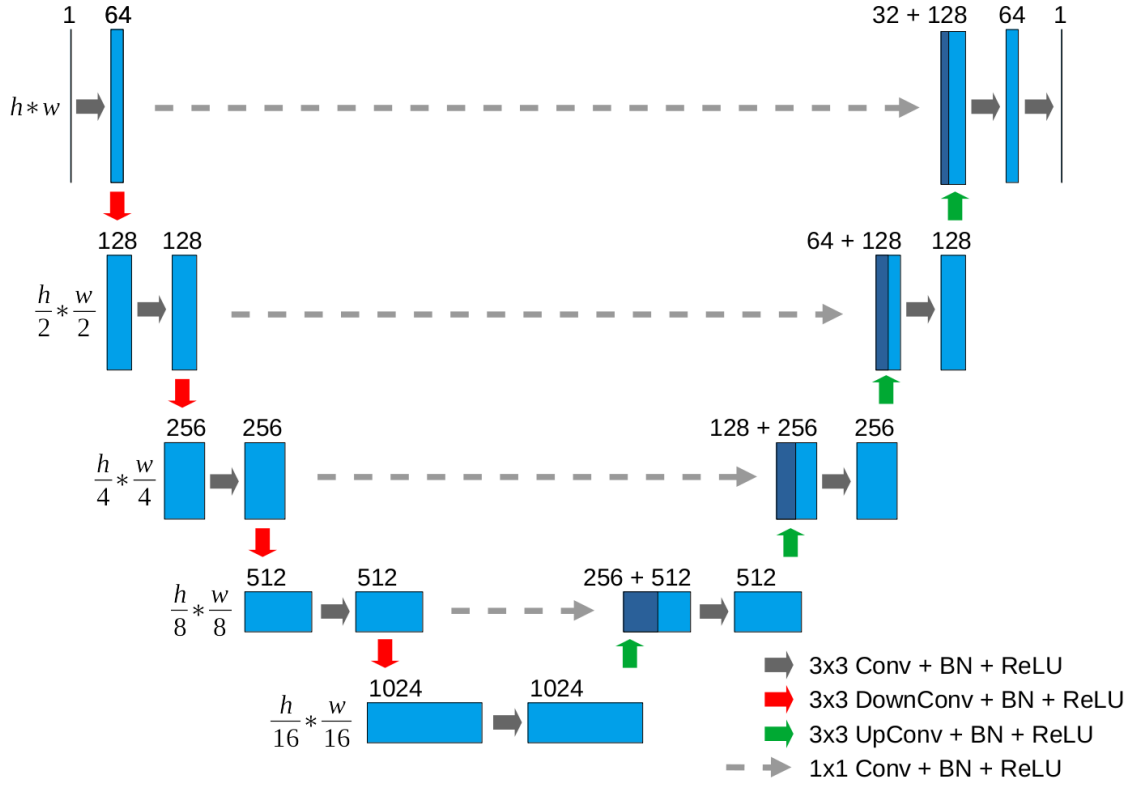


Figure 5.4: Network diagram.

5.4 Simulations and experiments

A dataset of 30000 dented surfaces was generated as shown above, each one containing 960×640 points over a spatial dimension of 50×33 cm. The dataset was split assigning 80% for training and 20% for validation. The network was trained using two Nvidia V100 64GB GPUs and each epoch trained in about 2h. The goal was to achieve the maximum IoU over the validation set.

For comparison, the network was trained first with Gaussian noise only ($\sigma = 2$ mm), without the use of noise signals. The confusion matrix is shown in Table 5.1. In this case the network reached a IoU of 92.66%. Precision and recall were 94.88% and 97.54%, respectively. This result, however, is affected by the more predictable noise distribution and does not necessarily translate to better performance when the network is used on *real* dent samples, as shown in the experiments below.

Afterwards, 20 flat surfaces with a resolution of circa 1800×900 pixels were acquired with the scanning apparatus and their error values calculated and augmented as for classic 2D images, then the network was trained again with both these noise signals and Gaussian

		Prediction	
		Positive	Negative
Actual	Positive (dented)	11.49%	0.29%
	Negative (non dented)	0.62%	87.60%

Table 5.1: Confusion matrix resulting from training with Gaussian noise only. IoU was 92.66%.

noise ($\sigma = 1$ mm) added to the virtual dents. The xy noise and all the other parameters were left unchanged.

In this second case, the IoU reached was 84.76%, lower, as expected, due to the presence of the irregular noise signals that the network has to model. Precision and recall were 89.35% and 94.25%, respectively, while the accuracy 98.05%. The confusion matrix of this second case is shown in Table 5.2.

		Prediction	
		Positive	Negative
Actual	Positive (dented)	10.82%	0.66%
	Negative (non dented)	1.29%	87.24%

Table 5.2: Confusion matrix resulting from training with both noise signals and Gaussian noise. IoU was 84.76%.

The prediction over one of the virtual surfaces in the validation set is shown in Fig. 5.5. It presents a rather shallow and wide dent, generally the kind more difficult to detect by humans [2]. In this case, results can be compared with ground-truth data and, like for human engineers, the boundaries result as the most challenging to identify.

However, the most interesting fact is to see if the knowledge learnt by the network is applicable to real scan data. For this experiment, four different surface samples were modeled, each one with a dent of different size (Table 5.3) following the shape of Eq. (5.2), then 3D printed as 15×15 cm tiles, sanded and painted. The samples were scanned using the same apparatus used to acquire the noise signals above. The comparison of the output predictions by means of the network trained without and with the noise signals is showed in Fig. 5.6. Regions predicted as dented are in green. The average processing time was only 0.74 s for 481 280 points (Intel i7-7500U CPU, 12 GB of RAM).

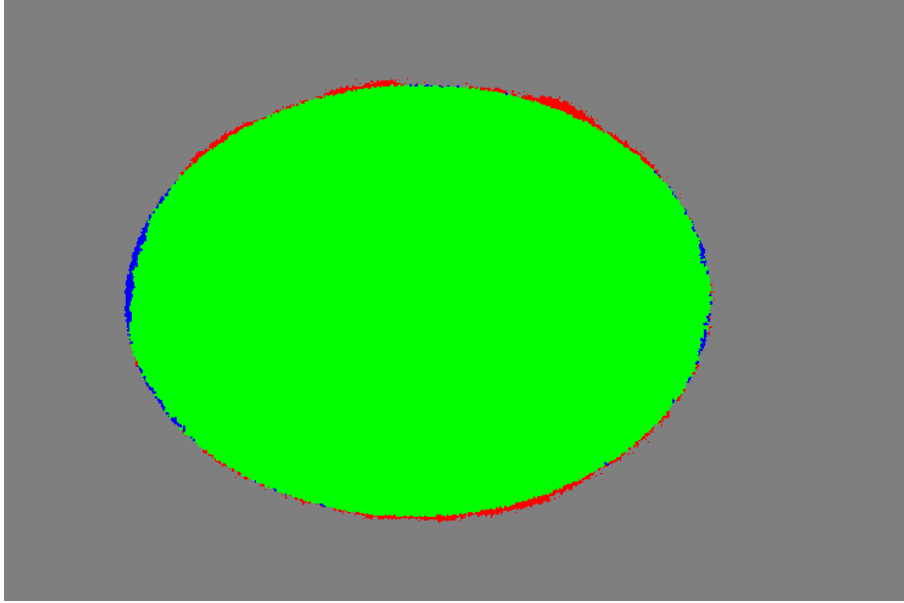


Figure 5.5: Prediction results over one of the validation samples. True positive are in green, false positive in red and false negative in blue.

Sample	Length (mm)	Width (mm)	Depth (mm)
A	60	40	2
B	120	100	2
C	100	80	3
D	120	80	1

Table 5.3: Dimensions of dent samples.

Although no rigorous ground truth is available for these samples, the network trained using the noise signals was able to correctly segment the dented areas almost completely. The network trained using only Gaussian noise, instead, performed poorly in all the cases.

As it happens for human engineers, the identification of the dent boundaries remains a challenging task. In fact, the samples B and D produced less than clear boundaries. This seems to confirm the thesis that, even for artificial networks, the wider and longer the dent, the more difficult its detection [2]. The sample D, in particular, presented a maximum depth of only 1 mm and the higher width/depth ratio, hence its segmentation is harder and some false negative spots are clearly visible. Some minor false positive spots are also noticeable in sample A and D. Nevertheless, for the most part, the proposed method correctly distinguished dents from the undamaged surface with evident difference of performance compared to the first approach not involving the acquired noise signals.

5.5 Conclusions

The automation of aircraft dent inspection can count on modern and accurate 3D scanning technologies to acquire raw data in the form of point clouds. Without comprehensive automation, however, the task of analysing them is left to the operator. As such, point cloud segmentation plays a fundamental role for the future of automated inspections. Dent segmentation, in particular, is difficult because, similarly to real world inspections, the dent does not present distinctive features. Another issue is the lack of a substantial dataset containing annotated point clouds of dented surfaces.

The proposed method reduces the problem to 2D segmentation by the fitting of a quadratic function. The matrix of residuals is then fed to a fully convolutional network, with an overall processing speed of over 500000 points per second on an Intel i7-7500U CPU. The lack of a suitable dataset was addressed by using virtual data. To prepare the network for real case scenarios, noise signals were acquired by the chosen scanning apparatus and added to the dataset during training. Thus, the network effectively learns to filter out the scanner noise and is able to correctly predict dents in real scan data, even if trained with virtual samples only.

Once the dent is correctly segmented, further post-processing algorithms may be applied to extract measures and show a full report to the engineer for the final damage evaluation. When available, the use of a labelled dataset of real dent scans could be used for fine tuning of the model, improving performance and robustness for the effective field application.

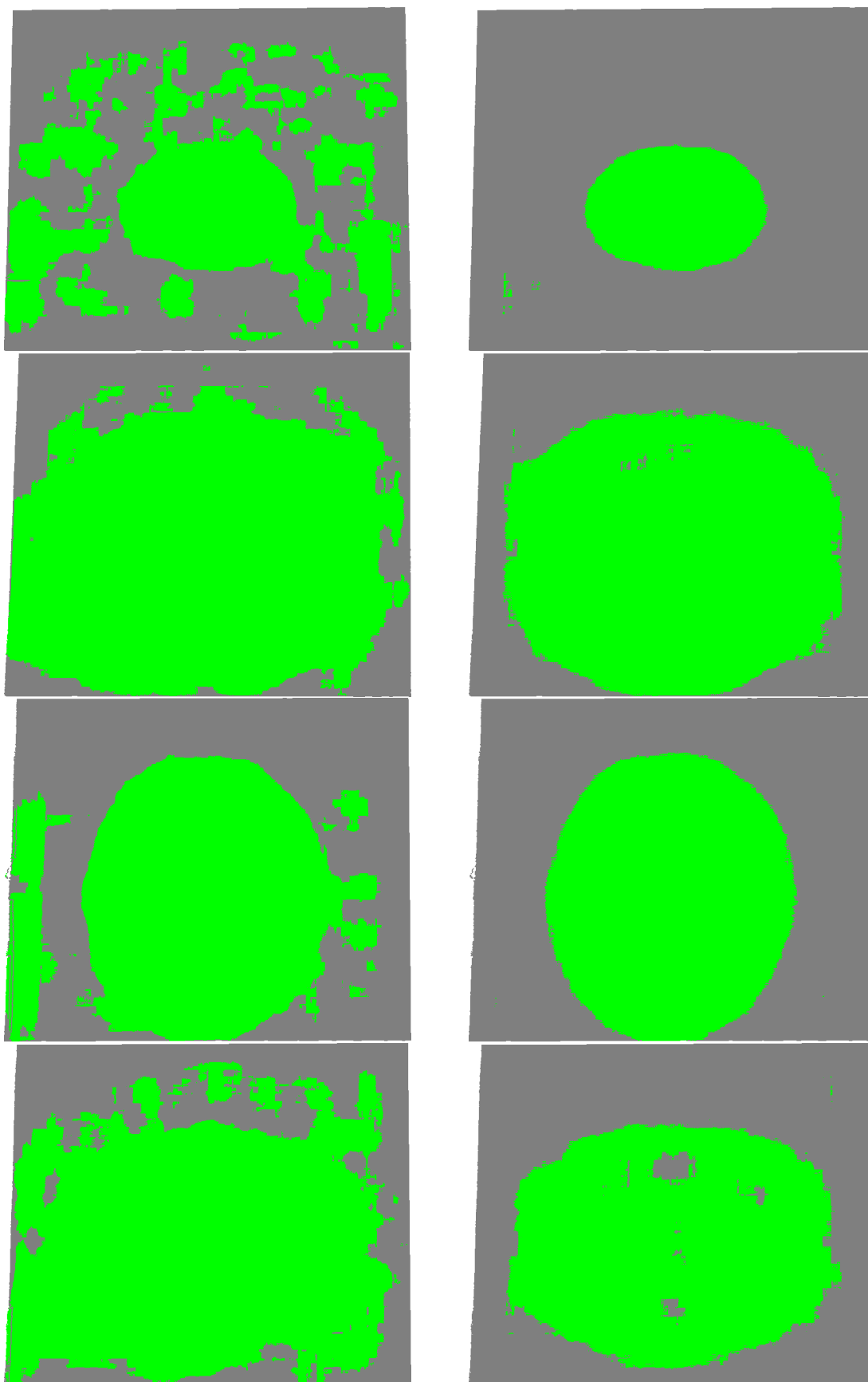


Figure 5.6: Prediction comparison between the network trained without (left column) and with (right column) noise signals of samples A,B,C and D (from top to bottom).

References

- [1] Civil Aviation Authority. *CAP716 - Aviation Maintenance Human Factors*. 2003.
- [2] Civil Aviation Authority. *Paper 2013/03 - Reliability of Damage Detection in Advanced Composite Aircraft Structures*. United Kingdom: Civil Aviation Authority, 2013.
- [3] Valerie C Coffey. “Machine vision: The eyes of industry 4.0”. In: *Optics and photonics news* 29.7 (2018), pp. 42–49.
- [4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-net: Convolutional networks for biomedical image segmentation”. In: *International Conference on Medical image computing and computer-assisted intervention*. 2015.
- [5] Yuxing Xie, Jiaojiao Tian, and Xiao Xiang Zhu. “Linking Points With Labels in 3D: A Review of Point Cloud Semantic Segmentation”. In: *IEEE Geoscience and Remote Sensing Magazine* (2020). DOI: 10.1109/MGRS.2019.2937630.
- [6] Igor Jovančević, Huy-Hieu Pham, Jean-José Orteu, et al. “3D point cloud analysis for detection and characterization of defects on airplane exterior surface”. In: *Journal of Nondestructive Evaluation* (2017).
- [7] Soufiane Bouarfa, A Dogru, Ridwan Arizar, R Aydogan, and J Serafico. “Towards Automated Aircraft Maintenance Inspection. A use case of detecting aircraft dents using Mask R-CNN”. In: *AIAA Scitech 2020 Forum*. 2020.
- [8] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. “Mask r-cnn”. In: *Proceedings of the IEEE international conference on computer vision*. 2017.
- [9] Sung Hyun Park, Amir Tjolleng, Joonho Chang, et al. “Detecting and Localizing Dents on Vehicle Bodies Using Region-Based Convolutional Neural Network”. In: *Applied Sciences* ().
- [10] Gernot Riegler, Ali Osman Ulusoy, and Andreas Geiger. “Octnet: Learning deep 3d representations at high resolutions”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- [11] Dario Rethage, Johanna Wald, Jurgen Sturm, Nassir Navab, and Federico Tombari. “Fully-convolutional point networks for large-scale point clouds”. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.
- [12] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. “Pointnet: Deep learning on point sets for 3d classification and segmentation”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

- [13] Charles R Qi, Li Yi, Hao Su, and Leonidas J Guibas. “Pointnet++: Deep hierarchical feature learning on point sets in a metric space”. In: *arXiv preprint arXiv:1706.02413* (2017).
- [14] Zhijian Liu, Haotian Tang, Yujun Lin, and Song Han. “Point-voxel cnn for efficient 3d deep learning”. In: *arXiv preprint arXiv:1907.03739* (2019).
- [15] Yulan Guo, Hanyun Wang, Qingyong Hu, et al. “Deep learning for 3d point clouds: A survey”. In: *IEEE transactions on pattern analysis and machine intelligence* (2020).
- [16] The Boeing Company. *Structural repair manual B737-400*. 2015.
- [17] Seokho Kang. “On Effectiveness of Transfer Learning Approach for Neural Network-Based Virtual Metrology Modeling”. In: *IEEE Transactions on Semiconductor Manufacturing* (2018). DOI: 10.1109/TSM.2017.2787550.
- [18] Yonglin Tian, Xuan Li, Kunfeng Wang, and Fei-Yue Wang. “Training and testing object detectors with virtual images”. In: *IEEE/CAA Journal of Automatica Sinica* (2018).
- [19] Shijie Feng, Chao Zuo, Liang Zhang, et al. “Calibration of fringe projection profilometry: A comparative review”. In: *Optics and Lasers in Engineering* (2021).
- [20] Pasquale Lafiosca, Ip-Shing Fan, and Nicolas P. Avdelidis. “Automated Aircraft Dent Inspection via a Modified Fourier Transform Profilometry Algorithm”. In: *Sensors* (2022). ISSN: 1424-8220. DOI: 10.3390/s22020433. URL: <https://www.mdpi.com/1424-8220/22/2/433>.
- [21] Shibani Santurkar, Dimitris Tsipras, Andrew Ilyas, and Aleksander Madry. “How does batch normalization help optimization?” In: *Proceedings of the 32nd international conference on neural information processing systems*. 2018.

Chapter 6

Aircraft Skin Inspections: Towards a New Model for Dent Evaluation

Submitted for publication as P. Lafiosca, I.S. Fan, N.P. Avdelidis.

Abstract

Aircraft maintenance, repair and overhaul (MRO) industry is gradually switching to 3D scanning for dent inspection. High-accuracy devices allow quick and repeatable measurements, which translate into efficient reporting and a more objective damage evaluation. However, the potential of 3D scanners is not fully exploited yet. This is due to the traditional way in which the structural repair manual (SRM) deals with dents: it defines length, width and depth as the only relevant measures and evaluates a dent by its width and width-over-depth ratio. This causes high degrees of ambiguity, with very different dent shapes (and corresponding fatigue life) being classified as the same, and nullifies the effort of acquiring such great amount of information from high-accuracy 3D scanners. Benefiting from the combination with 3D scanning technologies, the model here introduced proposes using 7 dimensional parameters to better describe the actual shape of the damage. The compact set of values can then be compared against historical data and structural evaluations based on the same model. Thus, the proposed approach can bring increased capability to evaluate damage, enabling more targeted interventions and, ultimately, saving costs.

6.1 Introduction

Aircraft skin is generally built with aluminium alloys or fibre-reinforced composites. Regardless of the differences in weight, fatigue performance, corrosion resistance and other material properties, the skin is exposed to various types of damage [1, 2, 3].

Dents, in particular, are extremely common during operations due to accidental impact with service vehicles, hail, small rocks on the runway or any other object that may cause damage, also known as *foreign object debris* (FOD). They affect structural performance of the thin aircraft skin, shortening fatigue life and contributing in the propagation of cracks [4, 5, 6]. For this reason, during maintenance tasks, the skin is thoroughly inspected, seeking for dents among many other anomalies that must be carefully evaluated.

The *structural repair manual* (SRM) is an important document containing information to identify and repair damages on primary and secondary structures. The former includes all those structures that contribute significantly to carrying flight, ground and pressurisation loads, with most of the skin usually falling into this category, as for example stated in the Boeing 737-400 SRM [7], here used as reference.

The Boeing 737 SRM defines a dent on metal as “*a damaged area that is pushed in from its normal contour with no change in the cross-sectional area of the material. The edges of the damaged area are smooth. This damage is usually caused by a hit from a smoothly contoured object. The length of the dent is the longest distance from one end to the other end. The width of the dent is the second longest distance across the dent, measured at 90 degrees to the direction of the length*” (see Fig. 6.1). The depth is measured where the width is taken. There must be no other kind of damages like gouges or cracks at the same location. Similarly, a dent on composites is *a concave depression which does not break the fibers*.

During MRO operations dents must be *found*, *measured* and *evaluated* in order to identify the necessary repair actions, if any [8]. In the event that the appropriate repair procedure is not contained in the SRM, the repair instructions must be obtained from the aircraft manufacturer. In this case, the need of a reliable damage representation is even more evident.

However, the traditional dimensional parameters are not sufficient to represent the complexity of dent damage: in particular, they cannot properly represent dents where the boundary is irregular or when the deepest point is not in the centre or, also, indicate how

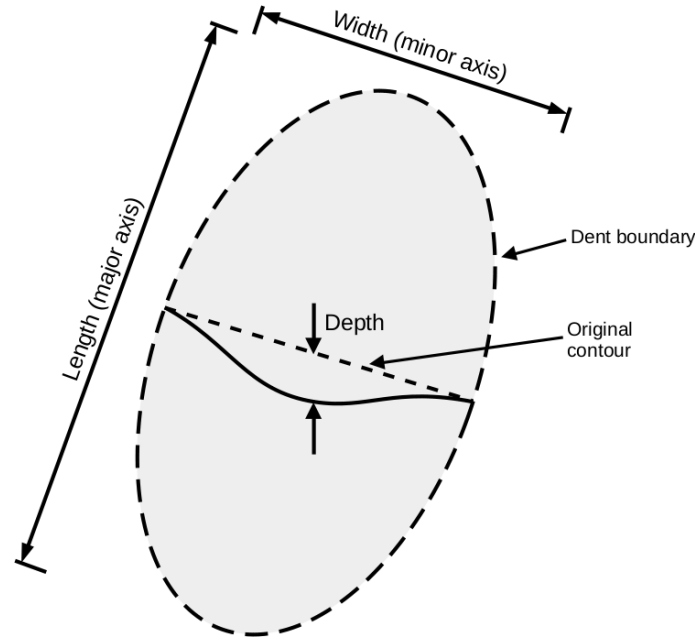


Figure 6.1: Traditional dimensional evaluation of dents.

quickly the maximum depth is reached. This work addresses the clear need for a more accurate model for dent evaluation in the current context of MRO, where new 3D scanning devices are being introduced to improve inspection and maintenance processes, while the way in which dents are evaluated has not changed yet.

6.2 Problem statement

As anticipated above, the SRM first classifies a dent by its *dimensional* characteristics, namely length, width and depth. Apart from other properties, like the type of material and the position with respect to certain aircraft parts or other dents, the measures of width and width/depth ratio are the only dimensional ones used to classify the damage as *allowable* or *not allowable*, the latter meaning that a repair is needed, immediately or within a certain number of cycles (if some conditions are met). This information is then reported in the *dent and buckle chart*. Often a representation of the damage is sent to the aircraft manufacturer, who provides detailed instructions about how to conduct the repair. Despite the dimensional characteristics being paramount for a proper evaluation, their reliability is undermined by many factors [2].

In fact, in the vast majority of cases, measures are collected manually by engineers by using a depth gauge and a ruler, thus a certain degree of subjectivity and error is always present. In attempt to introduce measurement repeatability, 3D scanning devices are being

gradually introduced in MRO. These devices produce a point cloud representation of the scanned surface and help the engineer producing a report. For example, in a study carried out by the company *8tree* the average standard deviation of traditional measuring methods was 23.2, and only 1.3 when using their 3D scanning device [9].

Nevertheless, 3D scanning does not entirely solve the problem: while the subjectivity of the measure is eliminated and the damage representation is moved from the physical to the digital world as 3D point cloud, most of the shape information is lost when comparing against the simplistic dimensional description and evaluation procedure prescribed by the SRM. In other words, despite the remarkable accuracy reached by 3D scanners for MRO (currently in the range of $25\mu\text{m}$), the assessment of the damage follows a *model* that does not take into account the actual shape and, instead, is typically made according to depth and width, nullifying the effort to acquire such great level of detail.

Indeed, a *model* provides a simplified representation of reality, that does not include all the real attributes, but only those that are relevant to the issue under consideration, within a given tolerance [10]. The current SRM dent model, however, is ambiguous and insufficient for an objective evaluation of the damage, especially with respect to the amount of information available by 3D scanning.

Consider, for example, the three illustrative dents represented in Fig. 6.2. By following

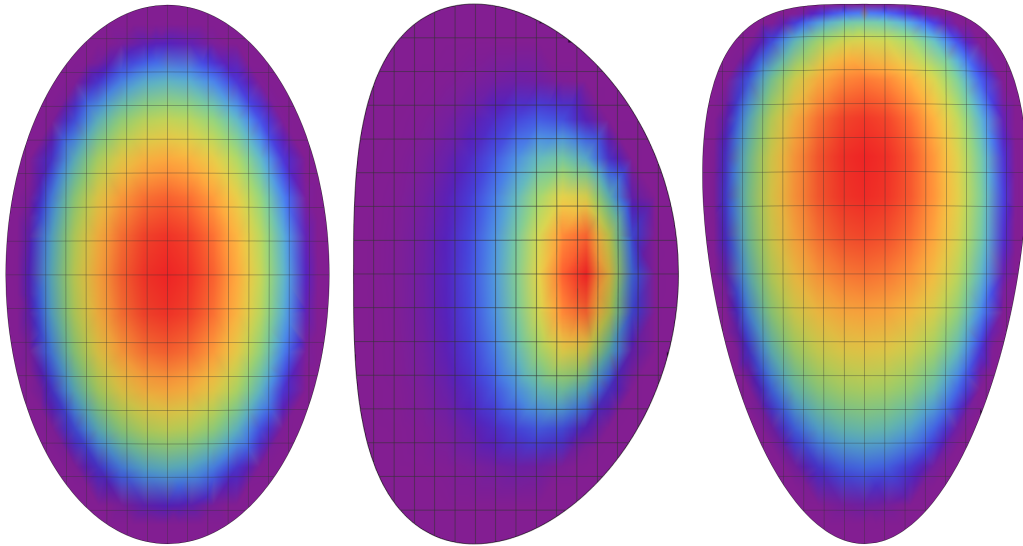


Figure 6.2: These three illustrative dents have all the same length, width and depth (as heatmap) and are thus classified in the same way by the SRM. However, they have evidently different shapes.

the SRM to the letter, these dents should be classified in the exact same way, yet they have

completely different shapes. In all likelihood, also their structural characteristics will be different.

Hence it follows the need for a more accurate model, against which tailored structural evaluations may be carried out, redefining the allowable damage thresholds accordingly. In practice, this translates into avoiding unnecessary repairs and, ultimately, reduced costs.

The next section proposes a mathematical model for the shape of a generic dent. When coupled with 3D scanning, it paves the way towards a new benchmark for dent classification and evaluation, and it may also facilitate the introduction of noise-robust automated measuring, as covered in later section.

6.3 Proposed model

6.3.1 Definition

The model here reported is an elaboration of the so-called *bump* function [11], having characteristics ascribable to a dent: smooth shape and vanishing boundaries. While a simpler derivation has been used to conduct simulations and build specimens [12], it does not have the capability to resemble asymmetric and irregular shapes. The general model here presented, instead, introduces *egg-like* boundaries, *xy*-shifted maximal depth and variable exponential base for a better representation of the dimensional characteristics, providing 4 extra parameters in addition to the traditional 3 (length, width and depth) already employed in the SRM.

The proposed dent model is first introduced by defining a *reference* dent:

$$\text{refDent}(x,y) \triangleq \begin{cases} b \left[1 - \frac{1}{1-r^2(x,y)} \right], & \begin{aligned} & -f(x) < y < f(x) \\ & \wedge s_x = 0 \wedge s_y = 0 \end{aligned} \\ b \left[-\frac{1}{1-r^2(x,y)} + (x-s_x) \frac{2r(s_x,s_y) \frac{\partial r}{\partial x}(s_x,s_y)}{(1-r^2(x,y))^2} \right. \\ \left. (y-s_y) \frac{2r(s_x,s_y) \frac{\partial r}{\partial y}(s_x,s_y)}{(1-r^2(x,y))^2} + \frac{1}{1-r^2(s_x,s_y)} \right], & \begin{aligned} & -f(x) < y < f(x) \\ & \wedge (s_x \neq 0 \vee s_y \neq 0) \end{aligned} \\ \text{undefined}, & \text{elsewhere} \end{cases} \quad (6.1)$$

The depth is always zero along the boundary, beyond which the dent ceases to be such and the function is considered undefined. The helper function $r(x,y)$ is defined as:

$$r(x,y) \triangleq \frac{\sqrt{x^2 + y^2}}{\sqrt{x^2 + f^2(x)}} \quad (6.2)$$

with $\frac{\partial r}{\partial x}(x,y)$ and $\frac{\partial r}{\partial y}(x,y)$ its partial derivatives and:

$$f(x) \triangleq \sqrt{0.25 - ((x + 0.5)^p - 0.5)^2} \quad (6.3)$$

The following 4 extra parameters are thus employed to characterise a dent:

- The base $b \in (1, \infty)$ is responsible for how fast the depth increases, going inward. Faster for values close to 1;
- The *egg-factor* $p \in (0, 2)$ produces the egg-shaped boundary and it is adapted from the work of T.C. Carter [13]. Note that for $p = 1$ the boundary is a circumference and its behaviour is not symmetric between lower and greater values;
- $s_x \in (-0.5, 0.5)$ is the shift of the deepest point along the x -axis;
- $s_y \in (-0.5, 0.5)$ is the shift of the deepest point along the y -axis.

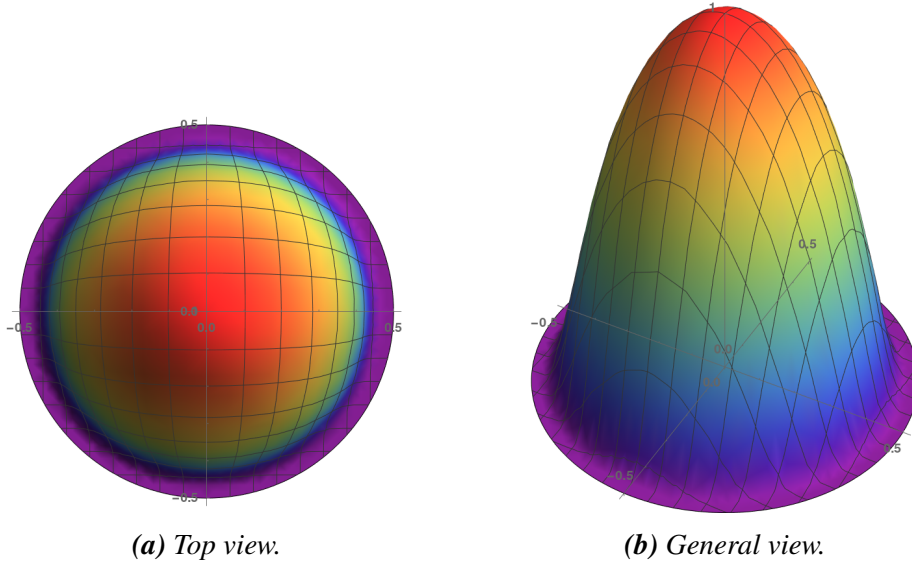


Figure 6.3: Reference dent function $\text{refDent}(x,y)$ with unitary xyz dimensions, $p = 1$, $s_x = s_y = 0$ and $b = e$.

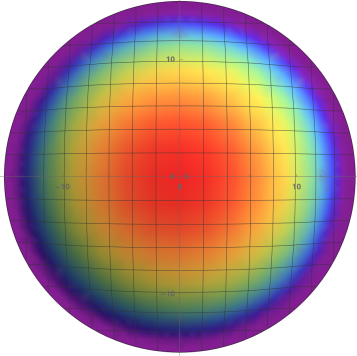
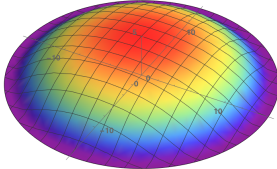
The reference function $\text{refDent}(x,y)$ showed in Fig. 6.3 has unitary dimensions along the three axis and can be easily rescaled to produce a general dent:

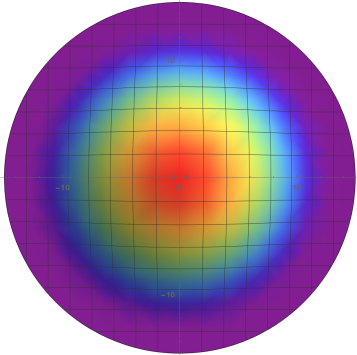
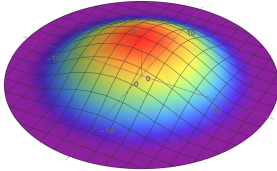
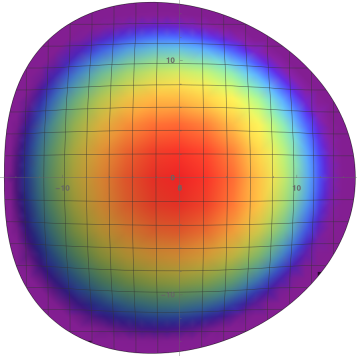
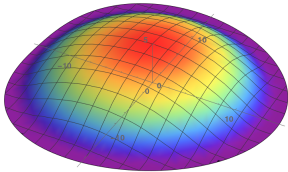
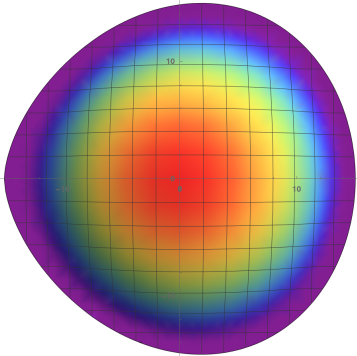
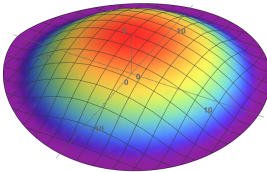
$$\text{dent}(x,y) \triangleq d \cdot \text{refDent}\left(\frac{x}{l}, \frac{y}{w}\right) \tag{6.4}$$

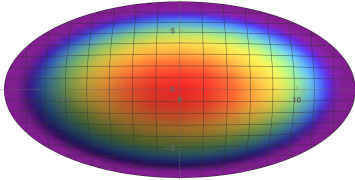
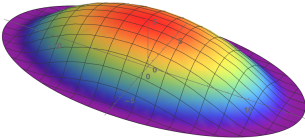
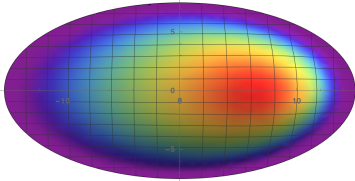
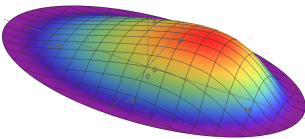
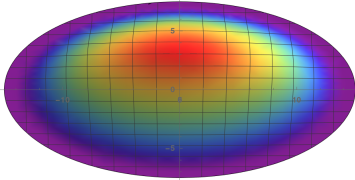
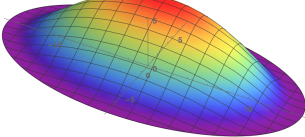
where l , w and d are the length, width and depth of the dent, respectively. Here the terms *length* and *width* are used for the dimensions along the x -axis and y -axis, respectively, and may be swapped with respect to the SRM definition (where the length is always the *longest* dimension).

6.3.2 Examples

Before discussing the possible applications in the following section, some shapes obtained with the proposed model are shown in Table 6.1, accompanied by their corresponding parameters. Depth values have been intentionally exaggerated for visualisation purposes.

Top view	Overall view	Parameters
		$l = 30$ $w = 30$ $d = 5$ $b = 2$ $p = 1$ $s_x = 0$ $s_y = 0$

Top view	Overall view	Parameters
		$l = 30$ $w = 30$ $d = 5$ $b = 10$ $p = 1$ $s_x = 0$ $s_y = 0$
		$l = 30$ $w = 30$ $d = 5$ $b = e$ $p = 0.8$ $s_x = 0$ $s_y = 0$
		$l = 30$ $w = 30$ $d = 5$ $b = e$ $p = 1.2$ $s_x = 0$ $s_y = 0$

Top view	Overall view	Parameters
		$l = 30$ $w = 15$ $d = 5$ $b = e$ $p = 1$ $s_x = 0$ $s_y = 0$
		$l = 30$ $w = 15$ $d = 5$ $b = e$ $p = 1$ $s_x = 0.2$ $s_y = 0$
		$l = 30$ $w = 15$ $d = 5$ $b = e$ $p = 1$ $s_x = 0$ $s_y = 0.2$

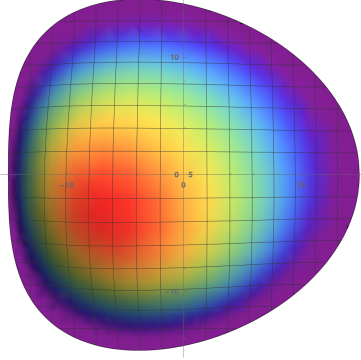
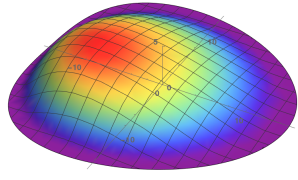
Top view	Overall view	Parameters
		$l = 30$ $w = 30$ $d = 5$ $b = e$ $p = 0.7$ $s_x = -0.2$ $s_y = -0.1$

Table 6.1: Examples of the proposed dent model for different parameter values, where e stands for the Euler's number.

6.4 Use cases

The model introduced in the previous section proposes using 7 parameters that all have a physical meaning. However, they cannot be easily measured manually (e.g. with rules, gauges, etc.) and, thus, the model benefits from the combination with 3D scanning technologies. The intended workflow is displayed in Fig. 6.4.

Starting from the physical surface to be assessed, its point cloud representation is acquired. Then a *segmentation* of some sort must be applied in order to actually isolate each dent by separating dented and non-dented areas [12]. Up to this point, this procedure is exactly the same as for most MRO products employing 3D scanning.

The next step aims at reducing the amount of information and making the measure intelligible and comparable. To date this is done by just extracting length, width and depth as prescribed by the SRM [14]. However, it was shown above that this does not adequately represent the actual shape and creates high degree of ambiguity. A more accurate alternative, instead, is available by means of the model of Eq. (6.4), which can be fitted by numerical optimisation to each group of points representing a dent, providing the

optimal set of 7 parameters representing that particular dent damage. This compact set of values can then be *compared* against historical data and structural evaluations that had been based on the same model, while such comparison would not be possible by directly using raw point cloud data.

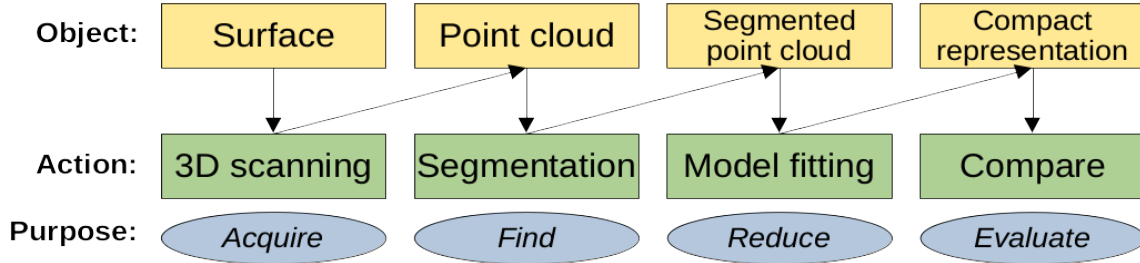


Figure 6.4: Expected workflow of an automated dent inspection system.

The added benefit is that storing or sharing the raw point cloud can be avoided, and only 7 numbers are enough for a fairly accurate representation of the real shape of the damage.

For example, Fig. 6.5 shows a dent manually marked by the engineers, photographed and then added into a graphical report together with its measures. While this method may appear approximated and cumbersome, it responds to the need for communicating more about the actual dent shape, and thus obtain specific instructions for repair, than simply reporting length, width and depth, which are insufficient for the ultimate goal. On the other hand, one may represent the damage by its full 3D point cloud. While this is rich of data, it is not easily comparable and manageable.

Instead, the shape calculated by means of Eq. (6.4) and showed in Fig. 6.6 easily approximates the actual one and with an intelligible set of only 7 parameters. In fact, while length, width and depth are the same as in Fig. 6.5, the 4 extra degrees of freedom address the necessity of an accurate and compact description that was previously attempted by manual drawings only.

The approach proposed in Fig. 6.4 can lead to more targeted interventions, ultimately saving costs. In order to accomplish this, the current practices are supposed to be revised and updated to embrace the increased capabilities brought by 3D scanning. This process of change has already started with the introduction of new devices, however, as part of a highly regulated and safety-critical industry, it is likely to take several years.

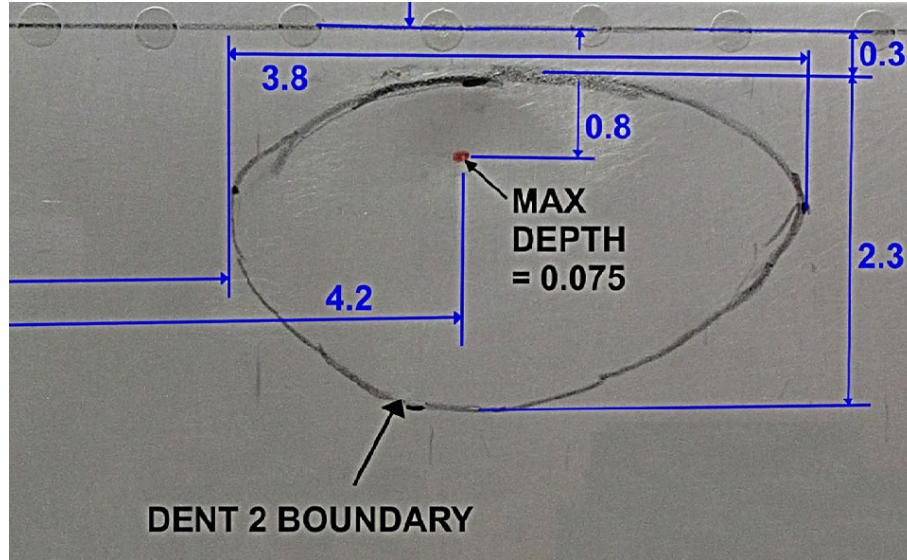


Figure 6.5: The complex shape of a dent cannot be expressed by only a combination of length, width and depth. In this picture the engineer relied on drawings to express the actual shape and the position of maximum depth, so that a better structural evaluation could be carried on. Even using 3D scanning, such great amount of data is ineffective if comparison is then done by means of depth and width/depth ratio only.

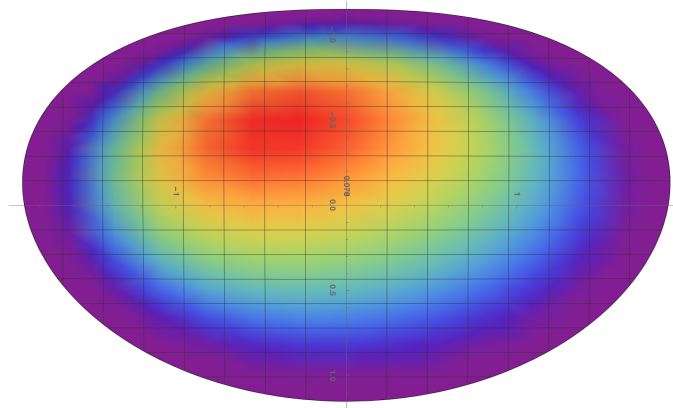


Figure 6.6: Shape obtained with $b = 2.7$, $l = 2.3$, $w = 3.8$, $d = 0.075$, $p = 0.85$, $s_x = -0.2$ and $s_y = -0.12$ which could potentially represent the dent of Fig. 6.5.

6.5 Conclusions

The MRO industry is gradually embracing 3D scanning devices for dent inspections. However, the accuracy and the amount of data that such devices offer has not been fully exploited yet. This is due to the way in which the dimensional properties of dents are evaluated by the SRM, namely by width and width/depth ratio only.

While complex shapes are easily acquired by 3D scanners, this valuable information

is lost when dents are evaluated only by their basic measures. Furthermore, a full point cloud or a drawing representing the dent is not readily intelligible nor easily comparable, because it cannot be represented as a compact set of parameters.

In this work a new model enables the representation of complex dent shapes by adding 4 dimensional parameters (b , p , s_x and s_y) to the traditional 3 defined in the SRM, namely length, width and depth. The optimal parameters are also suitable to be automatically calculated by numerical optimisation from a point cloud, derived, in turn, by any 3D scanning system.

The model here introduced proposes a compact and convenient way to describe dent damage that can potentially pave the way for a novel dent definition in the SRM and an efficient reformulation of allowable damage thresholds. This, in turn, would allow more accurate reporting, more targeted repairs and, ultimately, reduced costs. In future, the model may be revised and improved to represent other properties of dent damage.

References

- [1] Tolga Dursun and Costas Soutis. “Recent developments in advanced aircraft aluminium alloys”. In: *Materials & Design (1980-2015)* 56 (2014). DOI: 10.1016/j.matdes.2013.12.002.
- [2] Lawrence Cook. “Visual inspection reliability for composite aircraft structures”. PhD thesis. 2009.
- [3] Xi Chen, He Ren, and Cees Bil. “Inspection intervals optimization for aircraft composite structures considering dent damage”. In: *Journal of Aircraft* (2014).
- [4] Zhigang Li, Rongxin Feng, Yan Wang, and Lishu Wang. “Experimental study on the effect of dents induced by impact on the fatigue life of 2024-T3 aluminum alloy plate”. In: *Engineering Structures* (2017). DOI: 10.1016/j.engstruct.2017.01.048.
- [5] X Zhang, B Chan, S Lama, et al. “Influence of impact dents on the fatigue strength of aluminium alloy friction stir welds”. In: *Procedia Engineering* 2.1 (2010), pp. 1691–1700.
- [6] Praveen Shivalli. “The effect of dents on fatigue life and fatigue crack growth of aluminum 2024-T3 bare sheet”. PhD thesis. 2006.
- [7] The Boeing Company. *Structural repair manual B737-400*. 2015.

- [8] Ulf Paul Breuer. *Commercial aircraft composite technology*. Springer, 2016. DOI: 10.1007/978-3-319-31918-6_6.
- [9] Arun Chhabra. *How Did 100 Different Manual Measurements Stack Up?* 2019. URL: <https://www.8-tree.com/insights/how-did-100-different-manual-measurements-stack-up/> (visited on 2022-04-02).
- [10] H. Stachowiak. *Allgemeine Modelltheorie*. Springer, 1973. ISBN: 9783211811061.
- [11] Steven G Johnson. “Saddle-point integration of C_∞ ‘bump’ functions”. In: *arXiv preprint arXiv:1508.04376* (2015).
- [12] Pasquale Lafiosca, Ip-Shing Fan, and Nicolas P. Avdelidis. “Automatic Segmentation of Aircraft Dents in Point Clouds”. In: SAE International, 2022. DOI: 10.4271/2022-01-0022.
- [13] TC Carter. “The hen’s egg: a mathematical model with three parameters”. In: *British Poultry Science* 9.2 (1968), pp. 165–171.
- [14] T Reyno, C Marsden, and D Wowk. “Surface damage evaluation of honeycomb sandwich aircraft panels using 3D scanning technology”. In: *NDT & E International* (2018). DOI: 10.1016/j.ndteint.2018.03.007.

Chapter 7

Discussion and Limitations

7.1 Overview of the proposed process

Each chapter of this work addresses a specific aspect related to the automation of aircraft dent inspection. In Ch. 2 the attention was focussed on data acquisition technologies compatible with the use of UAVs. Although this choice imposed the design of a single-shot 3D scanning algorithm, it also allows portability and flexibility for the MRO engineer, who would be the final user.

In this regard, Ch. 4 proposed a new single-shot algorithm, i.e. capable to collect 3D point clouds from a single image. The immediate advantage is that data can be acquired even when fast movement and vibrations are preset, conditions that are common from the point of view of a UAV.

Ch. 5 showed how 3D point cloud data can be automatically analysed, or segmented, to extract damaged areas by means of a machine learning algorithm. After this stage the damage can be evaluated and measured, as point cloud data is in metric units.

Finally, Ch. 6 challenges the current way in which dents are evaluated in the SRM, taking the B737-400 SRM as a reference. Since the traditional evaluation is subject to ambiguity and cannot properly communicate the damage shape, also the following evaluation and repair instructions are affected by this gross approximation. Thus a new mathematical model to describe the shape of a general dent was proposed.

All these steps can be then integrated together into a new system for automatic, safe and non-invasive inspection of aircraft dents.

7.2 Demonstrative experiment on an aircraft rudder

In order to show the full working process proposed in this thesis, a demonstrative experiment was carried out on a rudder belonging to an Evektor-Aerotechnik Eurostar aircraft (Fig. 7.1) and available in the Accident Investigation Laboratory of Cranfield University. Other than several minor damages, the rudder presented a relatively large dent (Fig. 7.2) with a length of 65 mm, a width of 37 mm and a maximum depth of 2.5 mm approximately.



Figure 7.1: The dented Eurostar rudder used in the experiment.



Figure 7.2: Detail of the dent on the Eurostar rudder.

7.2.1 Scanning

Due to the partially reflective surface of the unpainted rudder, an anti-reflective spray¹ was applied before scanning. Camera and projector were placed at a distance of about 70 cm and the surface was scanned using the algorithm presented in Ch. 4 under the available soft light of the hangar (Fig. 7.3), as it was not possible to operate in full dark condition.



Figure 7.3: *Demonstrative experiment setup in the Accident Investigation Laboratory.*

Despite the fact that the conditions were far from a controlled scenario, the single-shot scanning algorithm was able to capture the resulting point cloud shown in Fig. 7.4, where the points are presented as heatmap in function of the deviation from a quadratic function and the main dented area is clearly noticeable. This expedient is necessary for visualisation purposes, as the original undamaged (undented) surface itself is not perfectly flat.

As the shape (perfect 3D model) of the scanned surface is unknown, no mean absolute error (MAE) can be rigorously calculated to evaluate the performance of the scanner; contrary to what has been done during the evaluation of the algorithm in Ch. 4, where

¹AESUB white anti-reflective spray for 3D laser scanning.

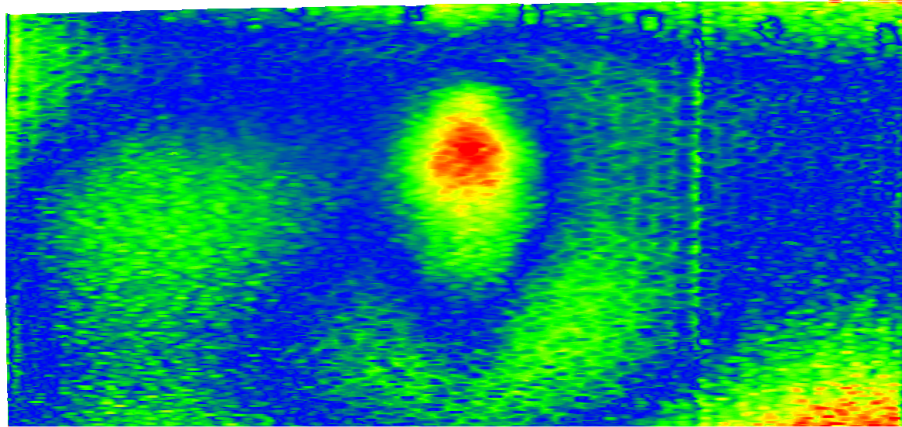


Figure 7.4: Scanned point cloud of the dented area on the Eurostar rudder coloured as distance error from a fitted quadric (blue = 0mm and red = 3 mm).

known 3D printed target dents were scanned.

A detail of the same point cloud is also shown in Fig. 7.5. The evident wave-shaped irregularities correspond to imperfect phase reconstruction (see Ch. 4), mostly due to *laser speckle*, responsible for introducing high-frequency noise as discussed in the following Sec. 7.3.

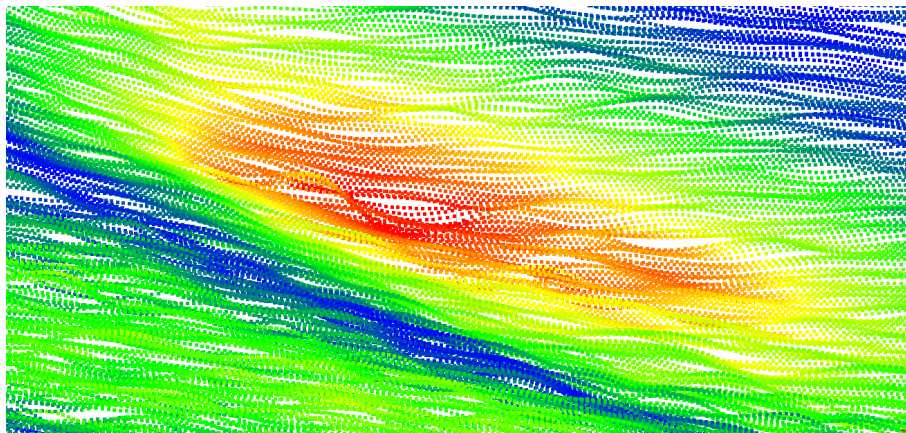


Figure 7.5: Point cloud detail of the same dent. Error trend is visible.

7.2.2 Segmenting

Since it would be unfeasible to manually evaluate thousands of point clouds during the inspection of an entire aircraft, the following step of the process proposed in this thesis is the automatic identification of the damage. Thus, the aim of point cloud segmentation is to bring only the relevant damage to the attention of the engineer.

As shown in Ch. 5, the fully convolutional network (FCN) was trained on pseudo-real data and it was fed with data coming from a real scan. Compared to the 3D printed samples scanned earlier, this time the task was harder as the rudder is much more different from the ideal surface used in training. In fact, by directly applying the standard 0.5 threshold to the prediction of the FCN, the unsatisfactory result of Fig. 7.6 is obtained.

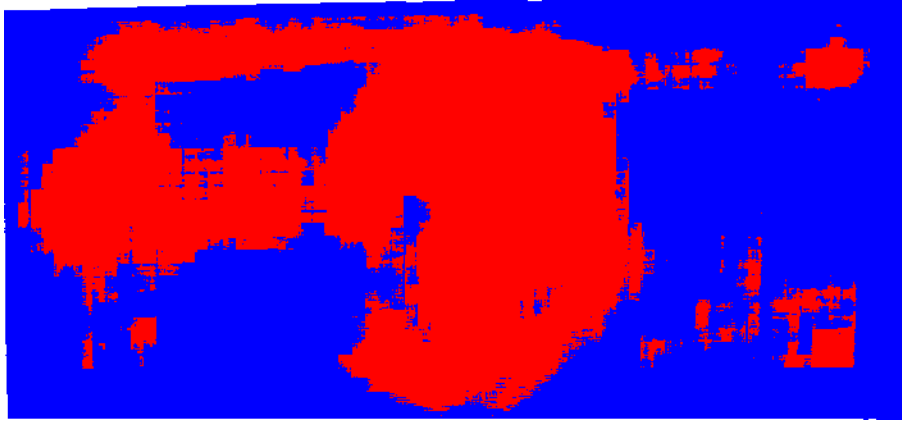


Figure 7.6: Segmentation result with standard threshold of 0.5. Blue corresponds to areas predicted as undamaged, red corresponds to damage prediction.

Input data was imbalanced during the training of the FCN (because dented areas are normally fewer than regular surface) and a manual rescaling weight was applied to compensate for this in the total loss calculation². While this seemed to work well for the experiments conducted in Ch. 5, the high rate of false positive found on the rudder point cloud (Fig. 7.6) shows that the network cannot identify the irregular dent shape there present and is deceived by the irregularities of the surface, as visible in Fig. 7.4.

By analysing the predictions for each 3D point, it was evident that the distribution of values was uneven, with non-damaged areas very close to 0 and other areas rapidly reaching values around 0.99 (where 1 means 100% probability of damage). Therefore, a simple threshold-moving strategy [1, 2] was considered. By moving the threshold to 0.99 or even 0.9999, the rate of false positive is decreased and the damage prediction is generally improved, as shown in Fig. 7.7.

As previously suggested, a fine tuning of the FCN by adding a reasonable amount of manually-labelled dents in both the training and validation sets will improve the final segmentation performance. Despite the insufficient data available during this research, the machine learning approach is nevertheless considered worth pursuing in the future development to enable proper automated damage identification.

²Refer to `pytorch.nn.BCEWithLogitsLoss()` documentation.

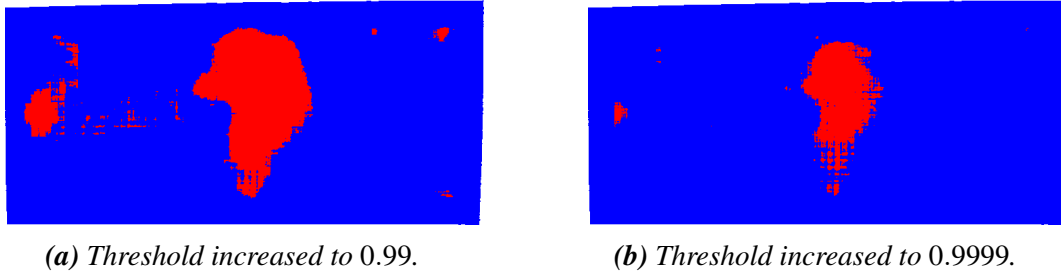


Figure 7.7: Segmentation result with increased thresholds. Blue corresponds to areas predicted as undamaged, red corresponds to damage prediction.

7.2.3 Measuring

Finally, supposing that all the points belonging to the dent are correctly classified as damage and isolated from the rest of the regular undamaged surface, measures must be extracted to comply with the SRM instructions.

The disadvantages and the subjectivity of manual measurements have been already discussed in Sec. 1.3. In addition, Ch. 6 highlighted how traditional measures do not allow a proper evaluation of the damage, and are instead neglecting important information about its shape. In fact, even when using high-accuracy 3D scanners, most of the information acquired in the point cloud is discarded because dents are evaluated only by their depth and width (and the ratio between the two), thus very different shapes may actually produce the same evaluation.

Consequently, the alternative model proposed in Ch. 6 represents dent shapes with a limited number of parameters, which is very well applicable in the event of 3D scanning inspections, not only in the process suggested by this thesis but also with today's handheld devices. The model is based on a mathematical function that can represent the shape of a dent more closely by taking in input 4 more parameters, in addition to length, width and depth as defined in the SRM.

By manually fitting the function with values of length, width and depth of 62.20mm, 42.60mm and 2.23mm, respectively, and extra parameters $p = 0.90$, $s_x = -0.12$, $s_y = 0.00$ and $b = 2.77$ the shape of Fig. 7.8 is more faithfully resembled. In fact, by taking into account common irregularities as the non-elliptical boundary, a decentralised deepest point and different curvature, the model provides a better approximation of the actual dent shape found on the rudder than the traditional approach.

Although the parameters used in the proposed model have all a physical meaning,

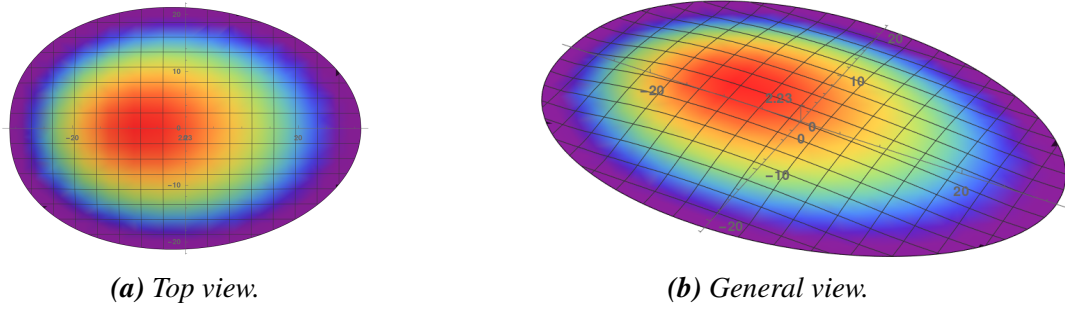


Figure 7.8: Shape manually fitted to the rudder dent showed in Fig 7.2. Red corresponds to the deepest point (2.23 mm).

it is unrealistic to manually measure them for each dent: the function can be instead numerically optimised to fit the damage point cloud. Consequently, further work beyond the scope of this research is required to design a proper cost function and employ an optimisation algorithm [3, 4, 5], so that measures are obtained automatically with such fitting strategy and eventually reported to the engineer in the desired format.

7.3 Light source and laser speckle effect

Part of the content of this section was contributed by Nathan Spiller, Department of Engineering Science, Oxford University.

The use of an incoherent light source while scanning, like common LCD projectors, generally results in short depth of field, as the projected pattern is sharp within a limited interval of distance. Consequently the operating range of the whole scanner is limited, adding practical difficulties that jeopardise system flexibility. Strategies that employ an out-of-focus projector have been proposed, although they may decrease the accuracy and introduce some other constraints [6].

On the other hand, the use of a *coherent* light source (laser) greatly extends the working range of the structured light system: the projector has *no focal length setting* and the pattern stays in focus at much greater distances. Overall, this leads to increased depth of field for the whole system, limited only by the camera settings.

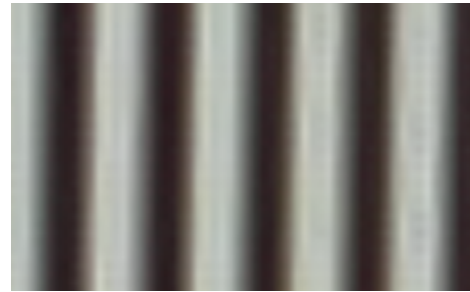
However, the use of laser introduces two technical challenges. First, the refreshing period of the laser pointer must be synchronised with camera exposure time: if this is not done, the camera image would present artifacts in the form of horizontal bandings (*flickering*). This technical issue can be addressed by building integrated hardware that shares the clock signal between camera and projector module, so that the camera exposure

time is an integer multiple of the refreshing period.

Secondly and more importantly, the projected pattern is affected by *laser speckle*, a granular distortion present when a coherent light source is scattered from an optically rough surface [7]. Indeed most surfaces are rough when compared with the optical wavelength of the light source, and the aircraft skin is no exception. Any observer point may be subject to incident light from multiple spatially separated points on the scattering surface. As a result of the coherence, light contributions will interfere constructively and destructively and the camera will thus observe a high contrast, time invariant, high spatial frequency speckle pattern superimposed upon the desired image (Fig. 7.9a). This phenomenon is unwanted because it interferes with the original frequency components of the projected pattern, lowering scanner accuracy.



(a) Detail of a fringe pattern projected by a class 1 laser projector.



(b) Detail of the same pattern projected by a common LCD projector.

Figure 7.9: Although focus-free, laser projection causes the speckle effect, which degrades pattern definition.

More specifically, each pixel, the lowest sampling unit for the camera, has an intensity value obtained as the average of the light rays focusing on that particular sensor. Such unwanted noise overlaps onto the projected fringe. Due to its random nature, the speckle signal has high frequency components that can easily be equal or higher than $0.5/\text{pixels}$, whereas the sampling frequency of the image is fixed at $1/\text{pixels}$.

The Nyquist–Shannon sampling theorem states that a sampling frequency f_s at least twice the maximum frequency f_{max} contained in a band-limited signal is required to reconstruct the original signal. Besides the fact that the band-limited assumption cannot be guaranteed, the theorem is thus violated even before the application of the algorithm presented in Ch. 4, and high frequencies appear as if they were lower ones. An illustrative example of this effect is presented in Fig. 7.10.

Hence, if laser source is used for the discussed scanning application, low-frequency artifacts enter the band filtered during the application of the FTP algorithm presented in

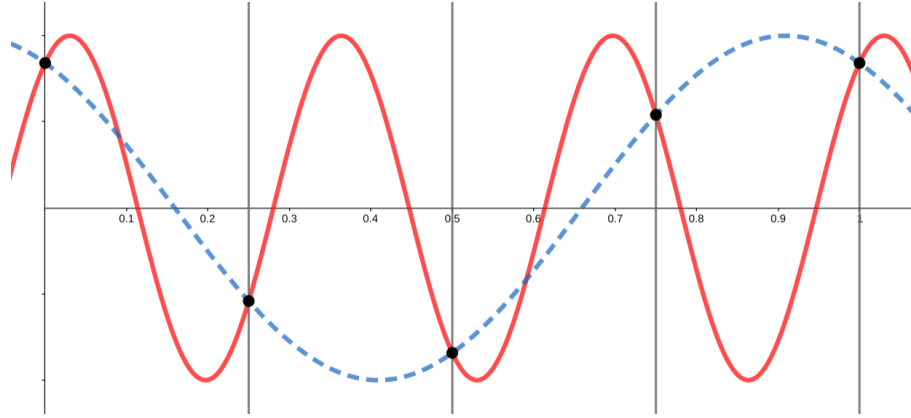


Figure 7.10: Violation of the sampling theorem: original sinusoidal signal (solid red) having $f_o = 3$ is sampled at the insufficient frequency $f_s = 4$. The reconstructed signal having $f_r = 1$ (dashed blue) is incorrect.

Ch. 4 (see also Fig. 4.3), resulting in smooth noise patterns noticeable, for example, over the point cloud of Fig. 7.5.

For an image of uniform intensity, the level of speckle present in an image is quantified by the speckle contrast parameter:

$$C = \frac{\sigma_I}{\bar{I}} \quad (7.1)$$

where σ_I represents the standard deviation of the intensities, I represents the individual pixel intensities and $\bar{I}$ represents the mean intensity of the image. The parameter can take any value between 0 and 1, where 1 represents fully developed speckle where the intensity fluctuations are on the same order as the average intensity.

Speckle is a distortion linked to the fundamental coherence properties of the light source and it cannot be removed without destroying the coherence of the source (such as using an LED instead of a laser). General methods to reduce the perception of speckle exploit the finite temporal response and spatial resolution of the detector, while spatial filtering of the image would be ineffective, as it also causes loss of information from the desired signal. The appropriate solution is rather to equip the scanning system with a *despeckler* [8, 9], a device that can reduce the speckle contrast by partially breaking the coherence of the laser beam, so that its effect becomes negligible.

The common approach to reducing speckle comes from averaging N statistically independent (non-interfering and decorrelated) speckle patterns. For the case where the speckle patterns have equal mean intensities, the resulting speckle contrast is [10]:

$$C = \frac{1}{\sqrt{N}} \quad (7.2)$$

Therefore, the speckle contrast decreases as the number of independent speckle patterns averaged in the integration time of the detector increases. Typically, there are 3 methods to generating statistically independent speckle patterns by diversification of three parameters of the light [11]:

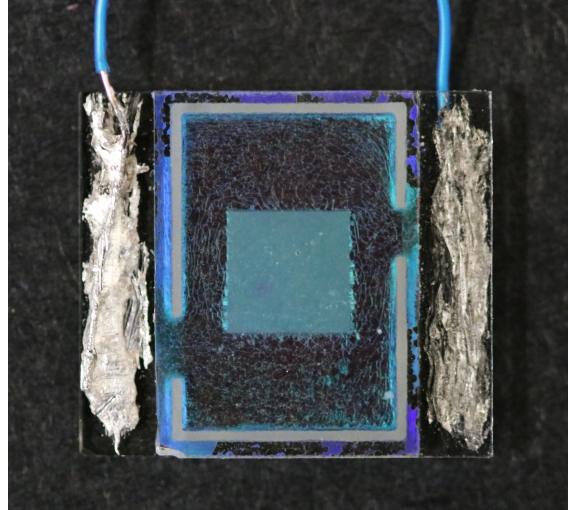
- *Angle diversity*: speckle can be reduced by altering the angle of the illumination of the light source.
- *Polarisation diversity*: when a polarised light source (such as a laser) is scattered it will experience depolarising. The resulting speckle pattern can be decomposed into two orthogonal and independent polarisation states, thus a reduction of $1/\sqrt{2}$ is gained.
- *Wavelength diversity*: speckle patterns from two light sources with different wavelengths will become uncorrelated if the wavelengths are separated by more than a threshold distance.

Practical methods of reducing speckle tend to focus on the first approach. Common examples include rotating (or vibrating) a ground glass diffuser, rotating a rectangular light pipe and vibrating a diffractive beam shaper with the aim to create lots of random patterns in the integration time of the detector (camera). If their number is large enough, speckle will be thus reduced through averaging.

These methods all suffer from some drawbacks: they introduce vibrations into the system, use bulky power consuming parts and may be sophisticated or expensive.

A novel approach that seemed noteworthy and particularly suited for the discussed application was instead proposed by the Soft Matter Photonics Group at the University of Oxford [12, 13]. Due to the anisotropy of liquid crystal (LC) molecules, and their ability to flow, LC devices can act as a dynamic scatterer (similarly to some of the above methods such as a rotating ground glass diffuser) without the need for moving parts. This is achieved by applying an AC electric field across the device. Additionally, these devices require little power and can be very compact (LC layer is on the order of a few μm) (Fig. 7.11). Experiments showed that in the appropriate conditions the LC despeckler can reduce speckle contrast by 80% or more [13]. The practical effect is visible in Fig. 7.12.

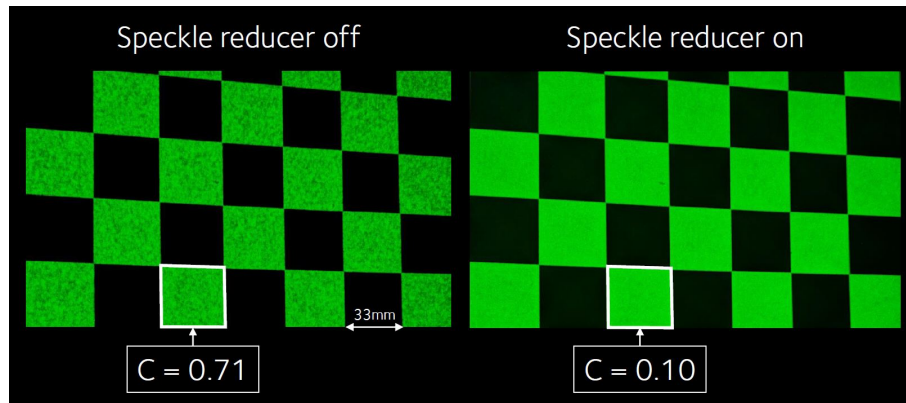
All these advantages mean that speckle-reducing LC devices would be compatible with precision measurements or drone based optical systems. However, for the case of a scanning laser projector new ideas need to be developed as both the traditional methods



Credits: Department of Engineering Science, Oxford University.

Figure 7.11: Liquid crystal cell with an active region of 5 mm^2 and a few μm thickness. The field applied between the two electrodes causes the light to produce different speckle patterns.

involving rotating a glass diffuser and the previous volume scattering LC cells aim at scattering the light, and this inevitably results in the final image becoming blurred. As such, other methods involving a mixture of angle and polarisation diversity need to be explored for maintaining the spatial shape of the beam.



Credits: Department of Engineering Science, Oxford University.

Figure 7.12: Reduction of speckle contrast by means of the LC despeckler can reach 80% in the appropriate conditions.

7.4 Ambient light conditions

Ambient light is one of the factors that most influence the final output accuracy of any structured light scanner [14]. Even if the use of laser source may improve the contrast of the projected pattern (see Sec. 7.3), camera input image will always contain a certain amount of ambient light noise.

The quantification of such factor is not straightforward as the effect of ambient light mixes with random scanning errors. In this regard, Li et al. [15] showed that the structured light system GOM ATOS III Triple Scan used in the experiment is capable to achieve similar performance under different ambient illuminance values (0, 280 and 660lx), with ambient light error being negligible with respect to random one. However, many are the factors to consider for a reliable evaluation, like algorithm type, distance from the object or light source.

In relation to the specific algorithm proposed during this work, its calculation are based on the phase shift rather than absolute intensity values. As such, at least in theory, uniform diffused light should not affect phase extraction via Fourier transform. Nevertheless, digital quantisation and camera exposure play a big role for the final quality of the image.

Furthermore, latest commercial scanners use a *blue light* source, believed to reduce errors as its shorter wavelength can be more easily filtered from the rest of the ambient light [16]. While this has been investigated for dental applications [17], an experimental confirmation for general applications could not be found in the literature.

Today's use of such scanning systems in MRO hangars, while at relatively short distances, seems to suggest their effectiveness for the *indoor* use cases. Further experiments and solutions are to be found, instead, for an *outdoor* use, where sunlight comes into play [18]. In this last regard, possible areas of research may involve *polarised light*, sometimes considered for NDT [19], or even particular wavelengths, like for example *ultraviolet* (UV): as the sunlight passes through the atmosphere most of its UV light is filtered out, meaning that the outdoor scenario is almost black at those wavelengths. As such, a UV projector may be able to provide a clear pattern even outdoor. To date, research appears to have covered these important aspects only marginally [20].

7.5 Surface reflections

While human inspectors have a distinct advantage in spotting dents or other anomalies over highly reflective surfaces [21], these have always been a challenge for structured light 3D scanners.

Structured light scanners work well when the scanned object exhibits roughly *Lambertian reflectance*, an ideal matte finish that reflects the incident light uniformly in all directions. Most of the real world objects, however, are not Lambertian reflectors, and indeed many applications require to scan highly reflective materials [22]. Also the aircraft skin, depending on the particular finish, may have high reflectivity levels that make the scanned point cloud partially or completely unreliable, as confirmed by experiments (Fig. 7.13).



Figure 7.13: Reflection over Cranfield B737-400 fuselage. Detail of the projected pattern that becomes unreadable in the reflection area.

Several approaches have been tried to deal with this issue. Generally the use of fewer grey levels when choosing the projected pattern showed better results for reflective objects [23], but it is far from being the definitive solution. Researchers also proposed a method to automatically adapt camera exposure based on surface reflectivity [24]. A more geometric approach is to use multiple cameras and projectors: since light reflections depend on the angles, this simple technique increases the likelihood to capture reflection-free areas [22].

Some solutions, possibly based on the above approaches, have been implemented in commercial products. GOM ATOS scanners, for example, makes use of two cameras and

a projector and are successfully employed to scan reflective car bodies. Other suggestions from commercial providers involve setting scanner sensitivity or preparing the surface with scanning spray.

The use case scenario proposed in this work does not allow any surface preparation and the one-camera, one-projector setup is motivated by obvious compactness constraints, due to the limited payload of UAVs. Therefore, the automatic camera exposure adaptation is likely to be the most viable and straightforward method to reduce the errors caused by reflections. Where this may not be sufficient, the proposed use case scenario may itself help in mitigating the problem.

In fact, during the continuous scanning, the same area of the skin is actually scanned multiple times from different point of views. Thus, a practical strategy for future development could be to programmatically identify and discard those areas presenting high reflections, so that the final point cloud can be build by combining acceptable areas only.

Dark non-reflecting surfaces introduce issues as well as high reflective ones. If the surface absorbs most of the projected light (or the projecting light is not powerful enough) the pattern is not clearly visible and accurate 3D scanning is not possible.

In the conditions of one of the experiments the two issues appeared even at the same time, like shown in Fig. 7.14.

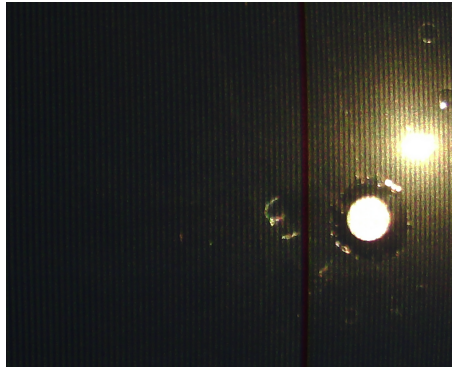


Figure 7.14: Scan of dark blue fuselage on the Cranfield B737-400 resulted in poor image quality, from which 3D information cannot be extracted. The surface is dark and highly reflective at the same time. Reflections disturb the right part of the image in several points. The white circle is a reflective numbered sticker applied to identify damages during previous inspections.

7.6 Employment of UAVs inside and outside the hangar

A UAV can inspect the entire aircraft body efficiently and quickly, saving a considerable amount of time. Companies like *Mainblades*³, *Donecle*⁴ or *Rapid*⁵ are innovating visual MRO inspections by reducing inspection time to as little as one hour and simplifying the reporting, although to date this is limited to certain types of damage like lightning. At the same time, the combination of UAVs with advanced solutions is being studied to cover more elusive damages like dents. Due to the commercial relevance, it is reasonable to expect such solutions to reach the market over the next few years, if not months.

Nevertheless, the employment of UAV side by side with humans during operations has to be considered in the context of safety and compliance [25, 26], even more so when dealing with autonomous systems and their unintended consequences.

Current regulations impose limits to UAV weight and flight radius [27], together with licences and insurance policies. UAVs are generally forbidden from certain areas. Airports, airfields and spaceports are restricted zones where permission to fly a drone must be granted by the airport itself. Flying a UAV indoor, or in any case without the possibility for the drone to escape, relaxes most of the rules, as it is not subject to air navigation legislation and cannot affect air traffic. However, general health and safety provisions still apply.

As such, issues that may arise in operating UAVs during MRO operations can be related to the following factors:

- Compliance: being able to comply with the strict regulations and standards of the aviation industry.
- Safety and collision avoidance: control the UAVs in the busy operating environment while avoiding accidents that may result in relevant threat and serious damage.
- Completeness: ensure that the inspection carried out by the UAV covers the intended areas, without producing false negative reports and being aware of occluded areas.
- Human resources: qualified and trained personnel is required to supervise the process and ensure system operability.

³<https://mainblades.com/>

⁴<https://www.donecle.com/>

⁵<http://www.mrodrone.net/>



Figure 7.15: Netting in place to contain UAVs during experiments at the Digital Aviation and Technology Centre, Cranfield University.

Today's increasing presence of UAVs in MRO shows that these issues are being overcome and that UAVs are a winning bet for aircraft inspection optimisation.

References

- [1] Guillem Collell, Drazen Prelec, and Kaustubh R Patil. "A simple plug-in bagging ensemble based on threshold-moving for classifying binary and multiclass imbalanced data". In: *Neurocomputing* (2018).
- [2] Alberto Fernández, Salvador García, Mikel Galar, et al. *Learning from imbalanced data sets*. Vol. 10. Springer, 2018.
- [3] Carsten Stoll, Zachi Karni, Christian Rössl, Hitoshi Yamauchi, and Hans-Peter Seidel. "Template Deformation for Point Cloud Fitting." In: *PBG@ SIGGRAPH*. 2006.
- [4] Qian Xie, Dening Lu, Kunpeng Du, et al. "Aircraft Skin Rivet Detection Based on 3D Point Cloud via Multiple Structures Fitting". In: *Computer-Aided Design* (2020). DOI: 10.1016/j.cad.2019.102805.
- [5] Jiajia Dai, Mingqiang Wei, Qian Xie, and Jun Wang. "Aircraft seam feature extraction from 3D raw point cloud via hierarchical multi-structure fitting". In: *Computer-Aided Design* (2021).

- [6] Beiwen Li, Yajun Wang, Junfei Dai, William Lohry, and Song Zhang. “Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques”. In: *Optics and Lasers in Engineering* (2014).
- [7] J Christopher Dainty. *Laser speckle and related phenomena*. Vol. 9. Springer science & business Media, 2013.
- [8] Igor Kompanets and Nikolay Zalyapin. “Methods and Devices of Speckle-Noise Suppression”. In: *Optics and Photonics Journal* (2020).
- [9] Weston H Thomas. “Multi-wavelength speckle reduction for laser pico-projectors using diffractive optics”. PhD thesis. 2014.
- [10] Joseph W Goodman. *Speckle phenomena in optics: theory and applications*. Roberts and Company Publishers, 2007.
- [11] Jahja I Trisnadi. “Speckle contrast reduction in laser projection displays”. In: *Projection displays VIII*. 2002. DOI: 10.1117/12.463781.
- [12] David J Hansford, Julian AJ Fells, Steve J Elston, and Stephen M Morris. “Speckle contrast reduction of laser light using a chiral nematic liquid crystal diffuser”. In: *Applied Physics Letters* (2016). DOI: 10.1063/1.4971997.
- [13] David J Hansford, Yihan Jin, Steve J Elston, and Stephen M Morris. “Enhancing laser speckle reduction by decreasing the pitch of a chiral nematic liquid crystal diffuser”. In: *Scientific reports* (2021). DOI: 10.1038/s41598-021-83860-3.
- [14] Samir Lemeš and Nermina Zaimović-Uzunović. “Study of ambient light influence on laser 3D scanning”. In: *Proceedings of the 7th International Conference on Industrial Tools and Material Processing Technologies*. 2009.
- [15] Feng Li, David Stoddart, and Iwona Zwierzak. “A performance test for a fringe projection scanner in various ambient light conditions”. In: *Procedia CIRP* (2017). DOI: 10.1016/j.procir.2016.06.080.
- [16] Prem Rachakonda, Bala Muralikrishnan, and Daniel Sawyer. “Sources of errors in structured light 3D scanners”. In: *Dimensional Optical Metrology and Inspection for Practical Applications VIII*. SPIE. 2019.
- [17] Jin-Hun Jeon, Byeong-Yeol Choi, Chong-Myeong Kim, et al. “Three-dimensional evaluation of the repeatability of scanned conventional impressions of prepared teeth generated with white-and blue-light scanners”. In: *The Journal of prosthetic dentistry* (2015). DOI: 10.1016/j.prosdent.2015.04.019.
- [18] Mohit Gupta, Qi Yin, and Shree K Nayar. “Structured light in sunlight”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2013.

- [19] Richard J Fryer and J Miller. “Imaging polarimetry for industrial inspection”. In: *BMVC91*. 1991. DOI: 10.5244/C.5.47.
- [20] Jostein Thorstensen, Jon Tschudi, Karl Henrik Haugholt, Gregory Bouquet, and Trine Kirkhus. “Structured light imaging under sunlight conditions”. In: *3D Image Acquisition and Display: Technology, Perception and Applications*. Optical Society of America. 2018, 3M3G–3.
- [21] Lawrence Cook. “Visual inspection reliability for composite aircraft structures”. PhD thesis. 2009.
- [22] Jeremy Gribben and Alan R. Boate. “Structured light 3D measurement of reflective objects using multiple DMD projectors”. In: *Emerging Digital Micromirror Device Based Systems and Applications XI*. Vol. 10932. International Society for Optics and Photonics. 2019. DOI: 10.1117/12.2510455.
- [23] Jing Xu, Ning Xi, Chi Zhang, et al. “Rapid 3D surface profile measurement of industrial parts using two-level structured light patterns”. In: *Optics and Lasers in Engineering* (2011). DOI: 10.1016/j.optlaseng.2011.02.010.
- [24] Laura Ekstrand and Song Zhang. “Autoexposure for three-dimensional shape measurement using a digital-light-processing projector”. In: *Optical Engineering* (2011). DOI: 10.1117/1.3662387.
- [25] Andrej Novák, Alena Novák Sedláčková, Martin Bugaj, Branislav Kandra, and Tomasz Lusiak. “Use of unmanned aerial vehicles in aircraft maintenance”. In: *Transportation Research Procedia* 51 (2020).
- [26] Efstratios Kakaletsis, Charalampos Symeonidis, Maria Tzelepi, et al. “Computer Vision for Autonomous UAV Flight Safety: An Overview and a Vision-based Safe Landing Pipeline Example”. In: *ACM Computing Surveys (CSUR)* (2021).
- [27] Civil Aviation Authority. *CAP1789A - Unmanned Aircraft Systems*. 2022.

Chapter 8

Conclusions and Future Work

Aircraft inspections to date lack a fully automated end-to-end system, capable of autonomously scan, identify, measure and report any damages found on the skin. The advantages of automating such tasks range from reduced costs to increased safety, from higher accuracy to repeatability of measurements. These far outweigh the challenges to research, develop and introduce such new technologies while being compliant with stringent regulations. As such, aircraft inspections are set to see increased automation, as occurred to many other industrial processes.

Among the numerous aspects covered by aircraft inspections, this work treated the specific and rather frequent dent damage on the aircraft skin. It showed that the tedious and highly subjective task of manually inspecting, evaluating and reporting dents on the skin can in principle be fully automated by developing an end-to-end system.

In order to do so, a quick and convenient way of moving around the entire surface must be coupled with the appropriate set of algorithms responsible for extracting and processing data. The practical advantage brought to MRO industry by UAVs is undeniable but, to date, limited to monocular vision applications (e.g. lightning damage detection) and thus leaving important questions open, like the detection of dents or other 3D damages and, in general, the reliable acquisition of measures discussed in this work.

Therefore, an improved Fourier-based 3D scanning algorithm was developed for the scope and tested in different conditions. The proposed single-shot approach inverts the paradigm: it tries to extract as much information as possible from a single picture, instead of relying on the multiple pattern acquisitions as commonly done by commercial 3D scanners. This immediately provides a great advantage for UAV applications, as the

approach is directly compatible with fast movements and vibrations.

Experiments conducted with a prototype of the proposed scanner showed that the system is already able to scan representative dents with submillimetre accuracy even during movement, with inexpensive equipment and at increased operating distance when compared to commercial 3D scanners, proving the compatibility of the algorithm for the suggested use case. Future results can also be enhanced by improving quality of the hardware and, with further research, by combining single-shot acquisitions from different poses or considering more advanced acquisition technologies. The Python code was made publicly available¹.

Acquisition of accurate data does not directly imply having quality information. Since the amount of data extracted from the skin of an average aircraft is potentially huge, a way to automatically segment 3D point clouds was proposed by means of a FCN. To overcome the lack of real data, a novel strategy was found to build a pseudo-real dataset, in which synthetic point clouds were mixed with noise signals extracted from the actual scanning system.

The new dent segmentation algorithm showed IoU higher than 80% and fast execution times, despite the non-distinctive nature of dent damage. The trained network was able to transfer such segmentation ability when real 3D points were given in input and, by fine tuning the network with a small labelled dataset, performance is set to increase to an even higher level of confidence, sufficient to programmatically identify damages from a large amount of point clouds and highlight only those relevant areas to the attention of the engineer.

During this work, the way that aircraft dents are traditionally evaluated was also critically revised, especially because the MRO industry has not yet taken full advantage of the high levels of accuracy brought by 3D scanning technologies. Private discussions with MRO engineers confirmed how in many occasions the use of high-accuracy handheld 3D scanning was eventually limited to special situations or dropped altogether, as the amount of data and the increased complexity was not justifiable in terms of efficiency.

It was argued that a high detailed 3D point cloud is difficult to evaluate by comparison with some standards or even previous damages while, on the other hand, the dimensional properties prescribed by the SRM are not representative of the actual damage. Therefore, a novel model for dent evaluation was introduced by providing a purely mathematical function that can represent complex dent-like shapes by means of 7 parameters, accom-

¹GitHub repository: <https://github.com/decadenza/SimpleStereo>

panied by some practical examples.

Overall, this study concluded the feasibility of automated dent inspections and pointed towards possible ways to overcome the shortcomings of the actual application in MRO.

Several future work directions may arise from this study. First and foremost, the data acquisition technology should deal with sub-optimal environmental conditions, affected by several sources of noise like, for example, natural light and laser speckle phenomenon. Such improvements may be achieved by considering custom hardware solutions, like high power projectors or despecklers discussed above, and also by exploring alternative algorithms for data retrieval that present reduced spectral leakage, unavoidable effect in the discrete Fourier transform.

The use of machine learning for the analysis of 3D data is yet another area where a multitude of algorithms may be introduced for the scope of this work. In fact, while the 2.5D strategy proposed with the U-Net segmentation showed the capability to correctly predict dents in point clouds, it also introduced an approximation of the input data when projecting over a plane. In the future this may be avoided by more advanced machine learning algorithms capable to process 3D data directly, however, these solutions introduce additional challenges in terms of convergence.

The literature on how damage measurements are affected by human factors is still limited and lacking a quantification of the tolerances involved, so that corrective actions may be considered by the operators, like the introduction of more autonomous solutions. Furthermore, the model proposed to describe dent shapes shall be integrated by a rigorous finite element analysis, aimed at the correlation between the shape and the real structural performance and fatigue life, also by means of structural laboratory trials on different materials.

Finally, all the above processes are to be collected into a data pipeline and integrated in routine MRO processes, resulting in a higher technology readiness level and thus enabling safer, more reliable and cost-effective autonomous aircraft inspections.

Glossary

Accuracy, closeness of agreement between a measured value and a reference value of a measurand. The measured value is generally considered to be a quantity that can be evaluated [1]. The actual method with which the measure is done may vary considering the properties of such quantity.

Epipolar constraint, geometrical property of a pair of stereo images, where two points taken on each image and corresponding to the same world object, together with the camera centers define a plane, that can be represented as an epipolar line on each image and can be used to simplify the problem of stereo correspondence [2]. See also Fig. 2.3.

Fatigue life, number of loading (stress) cycles or time span that a specimen can sustain without presenting failure or damage. It includes mechanical, thermal loads and more [3].

Image plane, in optics and photography, that plane in the world space that contains the projection of the object, generally coincident with the acquisition sensor. Similarly, in computer graphics, it is that plane where the image is being rendered.

Intersection over union (or Jaccard index), a statistic used for gauging the similarity and diversity of sample sets [4], often used to evaluate object detection and segmentation algorithms. It can be defined as the ratio between the size of the intersection divided by the size of the union, or:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (8.1)$$

Mean absolute error, in statistics, the measure of error between paired observations of the same phenomenon calculated as the sum of the absolute errors divided by the sample size.

Repeatability, condition of measurement that is held fixed while performing two or more measurements over a short period of time [1].

Bibliography

- [1] Joint Committee for Guides in Metrology. *International Vocabulary of Metrology*. 2021.
- [2] Song Zhang. *Handbook of 3D Machine Vision - Optical Metrology and Imaging*. CRC Press, 2013. DOI: 10.1201/b13856.
- [3] M. Kamal and M.M. Rahman. “Advances in fatigue life modeling: A review”. In: *Renewable and Sustainable Energy Reviews* (2018). DOI: 10.1016/j.rser.2017.09.047.
- [4] Wikipedia.org. *Jaccard index*. URL: https://en.wikipedia.org/wiki/Jaccard_index (visited on 2023-02-25).