Alex Richardson

Effective Temporal Change Detection in Low Altitude Aerial Imagery:

Using 3D Structure and Colour to Detect Scene Change in Models Generated from 2D Imagery

SCHOOL OF AEROSPACE TRANSPORT AND MANUFACTURING

PhD Thesis

CRANFIELD UNIVERSITY

SCHOOL OF AEROSPACE TRANSPORT AND MANUFACTURING

PhD THESIS

Alex Richardson

Effective Temporal Change Detection in Low Altitude Aerial Imagery:

Using 3D Structure and Colour to Detect Scene Change in Models Generated from 2D Imagery

Technical Supervisors:

Dr Toby Breckon (Durham University)

Dr Yifan Zhao (Cranfield University)

December 2019

Abstract

Unmanned Aerial Vehicles (UAVs) are now common place and their sensor solutions are producing ever increasing volumes of data. Typically the data is based around the theme of remote sensing of the Earth, and is gathered by a multitude of sensors for differing applications. The requirement to process the data gathered into useful information grows as does the demand for intelligent systems to assist with this. The most common, cost effective and readily available sensor solution is through standard camera photography, and offers the most usable data format without specialist tools. This also allows for proven methods to process the data gathered by a UAV thorough image processing and computation vision. One consistent theme in computer vision research is the drive for the ability to accurately reconstruct 3D scenes from 2D imagery through the process of Structure from Motion (SfM). This thesis details the research into the use of this 3D imagery, specifically aiding the ability to detect temporal change in dynamic scenes. This work presents a new technique to increase probability of detection and reduce computation required for such a process, the 3D Structure and Colour (3DSAC) differencing technique. The technique also goes to present a visualisation ability that best uses the algorithm for additional end user analysis beyond that of mathematics. Three scenarios where complex non-uniform changes are presented, of which assess and validate this technique to offer a capability to cope with dynamic scenes. The weighted 3DSAC algorithm gives the end user the ability to configure with emphasis being placed more within either structural or colour changes. Finally, through the implementation and evaluation of other current state of the art techniques for describing 3D points, the research shows the 3DSAC technique is more performant with imagery gathered by low altitude UAVs.

Acknowledgements

To my family and friends for their enduring support throughout my research, and their kindness and caring throughout sickness.

Dr Toby Breckon, for his patience and technical guidance both when he was my technical supervisor at Cranfield University and later at Durham University. Also to Dr Yifan Zhao for his guidance as a supervisor at Cranfield University.

I also thank my mentors in my professional career, who have taught me to approach and solve engineering problems and gave me the time to research while also gaining work experience.

Table of Contents

Index of Figures

Chapter 1

Introduction

The presence of Unmanned Aerial Vehicles (UAVs) has grown rapidly in the last decades, both in military and in commercial markets, making them easily accessible. The ability to utilise the data that can be gathered by their various sensors is paramount to their successful use. This could be leisure flying with a simple First-Person View (FPV) camera to advanced military radar and LIDAR systems.

As sensor systems have become more complex, data is gathered with a higher volume, variety, velocity and veracity (the four V's of big data). To maximise the potential effectiveness of data, the systems that process and analyse it must also be developed. This results in a drive for intelligent systems and computational assistance, with many tasks to aiming to become semi or even fully autonomous.

This chapter gives an introduction to some of the topics discussed throughout this thesis, specifically walking through basics of light, colour, and vision, and then into photogrammetry, aerial imagery and computer vision. This gives an understanding for the justification of this research, stating the aims, objectives and constraints faced throughout.

1.1 Light, Colour and Vision

The nature and effects of light have been postulated over the ages from the early Greek philosophers to modern scientists and visual perception is one of our most invaluable senses for understanding our surroundings. Light refers to the visible section (to the human biological eye) of the electromagnetic spectrum (Figure 1-1).

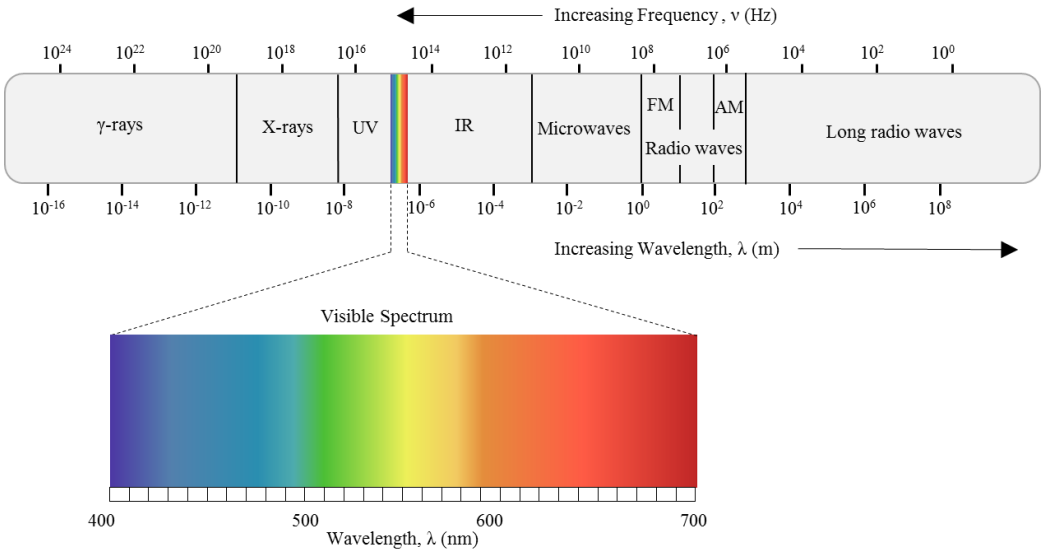


Figure 1-1 - Electromagnetic and visible spectrum

It was in 1021 that Alhazen empirically demonstrated the pinhole camera model through use of the camera obscura (Figure 1-2). He proved that light travels in straight lines by projecting it though a small pinhole. Newton later demonstrated that light is the source of colour, publishing Optiks in 1704. His experiments showed that natural light could be dispersed using a prism into its constituent colours rather than objects themselves being the source of colour. A typical trichromatic human vision system can detect wavelengths of 390 to 700nm. We perceive variation in wavelength (and hence colour) with the cone cells in the eye. There are three main types of cone that are each sensitive to different wavelengths of light; short (blue), medium (green) and long (red). To mathematically quantify and relate this with the visible electromagnetic spectrum, the International Commission on Illumination created the CIE1931 colour spaces; CIE1931 RBG and CIE1931 XYZ (CIE, 1931). The CIE colour standards and colour spaces are discussed in further detail in Section 3.3.1.

1.2 Pinhole Camera Model

As discussed in Section 1.1, it was Alhazen's work in the 11th century that demonstrated the pinhole camera model. The pinhole camera model is the simplest model of the camera, involving no lens and hence no lens distortion. It describes the projecting a 3D scene onto a 2D image plane of which is referred to as the perspective projection.

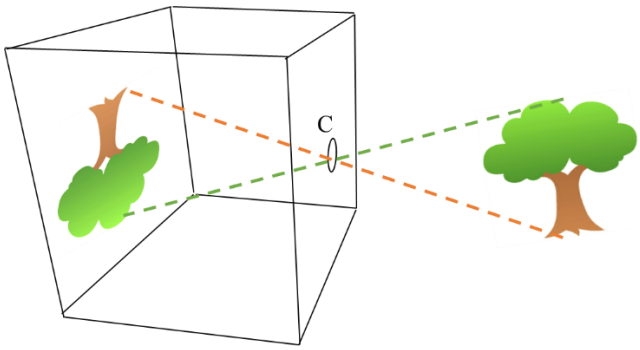


Figure 1-2 – Camera Obscura

Seen above in a diagram of a camera obscura, the size of the object on the image plane is dependent on the distance of this from the pinhole, known as the focal length. The pinhole in this model is referred to as the camera centre or optical centre, C. The line passing through the optical centre that is also perpendicular to the image plane is referred to as the optical axis or principle axis. The intersecting point of this and the image plane is referred to as the principle point. It can be seen that the image shown on the 2D image plane is inverted. To rectify this, the image plane can theoretically be taken forward of the optics/pinhole.

1.3 Photogrammetry, Remote Sensing and Aerial Imagery

The ability to survey scenes from a higher elevation or viewpoint gives a massive advantage to the area that can be observed. This enables sensing the earth (or subject) remotely with applications such as assisting mapping, resource identification, or as a visual early warning.

The first known aerial photograph is dated 1852, where Gaspard-Félix Tournachon, commonly known as "Nadar", used a tethered balloon to give his heavy camera altitude. Unfortunately Nadar's earliest photos did not survive, with the oldest surviving aerial photo attributed to JW Black's view of Boston in 1860 (Figure 1-3).

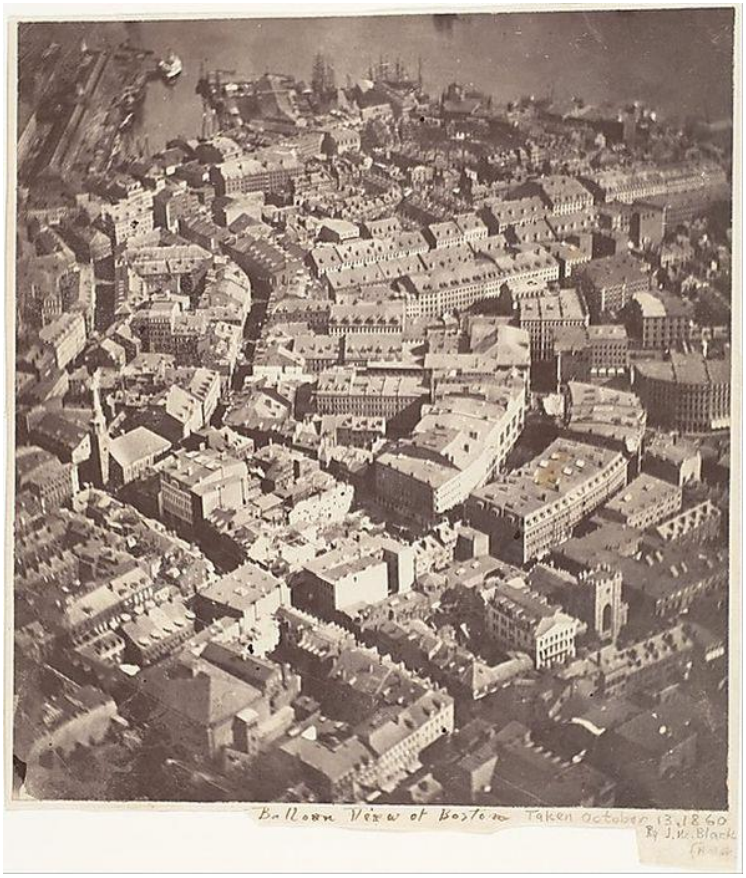


Figure 1-3 - Balloon View of Boston (Black, 1860)

Another very famous early aerial photograph was taken by George Lawrence after the 1906 San Francisco earthquake and fires. He used a series of 17 kites to lift a large format camera to circa 2000 feet altitude. With a curved film plate he captured in exquisite detail

the scene and subsequent devastation. This showed just how useful aerial photography can be in gaining a scene understanding. It wasn't long until the scientific and military potential was realised. Aerial photography and remote sensing became and continues to be a key method for gaining an information advantage.

Figure 1-4 - San Francisco Panorama (Lawrence, 1906)

1.4 Digital Imagery and Computer Vision

The initial concept of the digital camera was borne by Eugene Lally of the Jet Propulsion Laboratory in 1961 when he considered a solution to help navigate in space. It was nearly a decade before sensors were developed to realise the theory. The first of these was the Charged Couple Device (CCD) and was developed by George Smith and Willard Boyle of Bell Laboratories (Boyle and Smith, 1970). A series of light sensitive elements in the CCD detect the light and this is converted into pixel data. The first CCD had a resolution of 100 x 100 pixels and used a simple 8-bit shift register.

By 1975, Kodak engineer Steven Sasson had developed the first black and white digital still camera. This weighed in at nearly 4kg and comprised of a Fairchild 100 x 100 CCD, taking around 23 seconds to capture a digital photograph. This long capture time would constrain the photograph to movements which are relatively static in that time period. The performance and weight of digital cameras now far exceeds that making many more applications viable such as where many images are required or size and weight is a consideration as in this research.

As we know, our brains are incredibly complex and we often take for granted how effortlessly most of us see the world through our vision. It's often said that the eye and visual system are the most complex biological systems. There is a distinct two part process whereby external light enters the eye and is projected onto a plane, and the process of our brains reconstructing this to give depth perception and understanding of the scene.

When this is moved into the digital and computation domain, many overlapping and interlocking disciplines work together to create a computer vision capability. Figure 1-5 gives a non-exhaustive visual overview of some of these. The disciplines cover a wide range of subjects, from classification, feature extraction, signal processing to pattern recognition and object recognition and machine learning to name a few. Research began in the late 60s, and was originally thought of as a relatively simple problem. As more has been researched in to the problem space, the understanding is that vision is actually a highly complex problem with many avenues.

As our understanding has evolved, so have the types of imagery being produced. From simple light sensors, through to sensors of most levels of the electromagnetic spectrum. More recently the advent of 3D imagery has become a hot topic, from the cinema to medical imagery, now people are able to capture this through their smartphones either through multiple cameras of the movement of the camera, known as Structure from Motion.

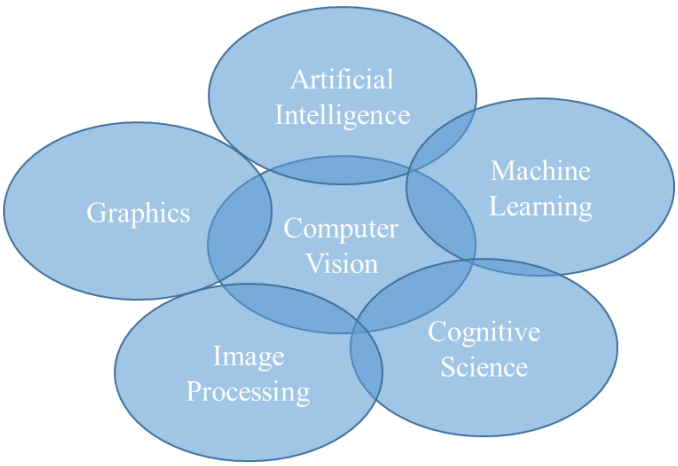


Figure 1-5 – Example of Computer Vision Disciplines

1.5 Structure from Motion (SfM)

SfM allows the reconstruction of a 3D scene that has been captured by a series of 2D images. By using multiple viewpoints of the same point captured in the series of 2D images, a 3D point model is created. The data is transformed from a structured uniform 2D matrix of pixels with colour data, to 3D points in space with colour data.

Figure 1-6 gives a systematic process of a generalised SfM technique. 2D images are captured, the features can be detected and described. From these, the correspondence between those matching in different images can be found and any outliers removed.

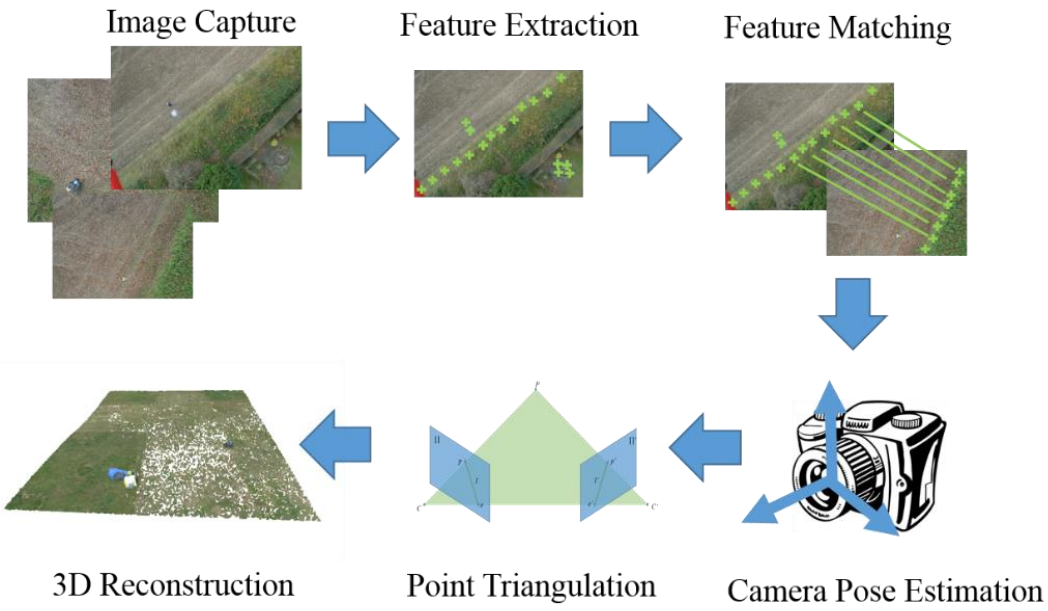


Figure 1-6 – Generalised SfM Process Flow

By taking the projections of these points, the epipolar geometry can be found and the essential and fundamental matrices estimated. Finally the points can be brought together to form a 3D reconstruction of captured scene.

1.5.1 Epipolar Geometry

Epipolar geometry describes the geometric relationship between two perspective viewpoints. This is easiest to understand visually, and Figure 1-7 shows two cameras viewing the same point, *P*, and their relative optical centres, *C* and *C'*. The plane between these three points forms the epipolar plane (green shaded region), and the lines where this intersects the image planes are called the epipolar lines (*l* and *l'* respectively). The line joining the camera centres (points *C* and *C'*) is named the baseline, and the points where it intersects the image plane are called the epipoles (*e* and *e'*).

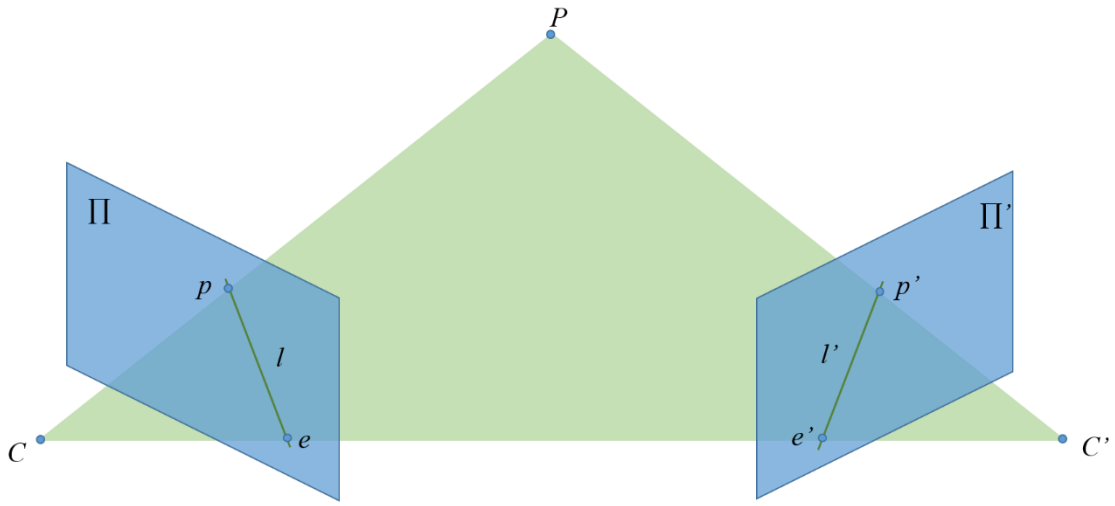


Figure 1-7 - Epipolar Geometry

Epipole *e'* is the projection of camera centre *C* on image plane ∏', and vice versa, *e* is the projection of *C'* on image plane ∏. This means, point *p* must lie on the epipolar line *e* associated with point *p'*, and vice versa. This relationship is called the epipolar constraint and forms the basis for epipolar geometry.

1.5.2 Projection Matrix

In Figure 1-7, the principle point is where the principle axis intersects the image plane. For simplicity in this model, this is located at the centre of the image plane, central projection, and the mapping from 3D to 2D may be expressed by:

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{1}$$

The 3x3 matrix used to map from 3D to 2D is called the projection matrix, *M*, and the homogenous 2D and 3D coordinates can be denoted by *p* and *P* respectively. This enables Equation 1 to be represented simply as:

$$p = MP \tag{2}$$

1.5.3 Essential and Fundamental Matrix

To calculate the epipolar geometry, two matrices need to be derived, the essential matrix (calibrated case) and the fundamental matrix (uncalibrated case). The essential matrix contains the information defining the location of the second camera from the first camera (or vice versa). More specifically, that is, the rotation and translation between the projected points of *P* on the image planes, *p* and *p'* respectively. This gives 5 degrees of freedom. The fundamental matrix relates the rotation and translation of point *p* and *p'*, and the camera intrinsics. This gives a total of 7 degrees of freedom.

Let's first look at the essential matrix. If we define the left camera, *C*, as the origin, camera *C'* is said to be at location *T*. The projection of *P* onto the right hand camera is said to be *p'*, and this can be described as:

$$p' = R(p - T) \tag{3}$$

Using the cross product of *P* and *T*, an equation for all the possible points of *p* through *T* may be defined as:

$$(p - T)^T (Tp) = 0 \tag{4}$$

Rewriting Equation 3 with this in mind gives,

$$(p - T) = R^T p' \tag{5}$$

Substituting this back into Equation 4 yields:

$$(R^T p')^T (Tp) = 0 \tag{6}$$

Rewriting the cross product as a matrix allows us to define matrix *S*:

$$T * p = Sp \Rightarrow S = \begin{bmatrix} 0 & -T_Z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix} \tag{7}$$

Matrix *S* is a singular matrix of rank 2, unless *T = 0*.

Substituting this in to Equation 4 allows us to relate *p* and *p'*:

$$p'^T RSp = 0 \tag{8}$$

This can further be simplified, where *RS* is the essential matrix, *E*:

$$p'^T Ep = 0 \tag{9}$$

Next the fundamental matrix will be defined. This takes the essential matrix and adds the relationships of the camera intrinsics. For this point we look back to Equation 2, and substitute p for q and the intrinsics that relates to it. This shows that point $q = Mp$, and inversely $p = M^{-1}q$. Substituting this into Equation 8 gives:

$$q'^T (M'^{-1})^T E M^{-1} q = 0 \tag{10}$$

Simplifying this to define the fundamental matrix as:

$$F = (M'^{-1})^T E M^{-1} \tag{11}$$

Gives:

$$q'^T F q = 0 \tag{12}$$

The fundamental matrix inherits the cross product *S* matrix (Equation 8) making this also a rank 2 matrix.

1.6 Visual Change Detection

"Detection, a binary decision to determine the presence of a signal (target), based upon sensor measurements." Waltz (1990).

The capability to visually detect change is through a comparison of two or more epochs. In temporally dynamic scenes change detection allows focus of attention. It enables the detecting entity to focus on to only those regions that have changed. Humans naturally detect movement and salient change in a scene, which draws our attention and allows us to cognitively make a decision - if it's pertinent, we analyse that further, or discard it as background and uninteresting noise.

Computationally this is significant, as once change is detected at a macro level, at a micro level it reduces the need to process an entire dataset. An example of this could be analysing a changing coast line and understanding the impacts this has on the environment. Here it is only the change that is relevant, and the rest of the area that is known to be the same is hence irrelevant (or at least within a threshold). Resources can then be focussed into understanding the change, such as what its composition is and how fast it is changing.

Another example could be detecting that an Improvised Explosive Device (IED) has been placed at the side of the road and deciding to deviate a convoy's path. Detecting that something has changed and being able to differentiate this from the normal scene allows us to investigate further. This is paramount to making decisions, it forms the first part to understanding the change and instigate a reaction. This holds true of any advanced artificially intelligent system and is not limited to visual change detection.

Change detection is a converse problem to similarity, whereby something can be detected, recognised and identified from a library data that have already been learnt. That being said, once a change has been detected, using similarity of that change to other objects or change can give it context.

1.7 Research Motivation

As discussed, aerial imagery has become a more prevalent and accessible media, the ability to use this to inform decisions without large amounts of human analysis becomes pertinent. It's evident that the volume of data created through aerial surveying is great, and this in turn requires extensive processing and computation. Detecting change in imagery accurately and consistently in complex environments is a challenging problem and facilitates important decision making processes. From ecological and environmental study through to surveillance and intelligence gathering. Data volumes gathered from UAVs is growing considerably and requires the additional automatic processing to reduce human overhead. Automating this process presents detections that would not easily have been made by human analysis alone.

Traditional 2D imagery is formed by a regular matrix of light sensitive sensors of which quantises data into a matrix of pixels. Converting this data into the 3D point clouds through SfM enriches data with an extra dimension and reduces the quantised effect of a digital camera. This presents a different and enhanced opportunity to utilise that data. Combining aerial photography with knowledge of projective geometry allows for precise measurement in both 2D and 3D. This has become a successful tool in cartography and other surveying. Harwin and Lucieer's research identifies the potential uses for aerial SfM and highlights that this is as a highly desired capability for environmental monitoring (Harwin and Lucieer, 2012).

To date, 3D change detection has focussed solely on structural change, using positional change as the discriminator. It discards useful colour information that could be further used to increase performance in the change detection process. These extra dimensions of colour information could provide vital information when structural change is very small, and hence forms one of the motivations for this research. Importantly, the process of gaining a 3D scene understanding from 2D imagery is a non-perfect one that introduces errors in the form of positional and colour noise and other artefacts such as shadows. Aerial imagery presents further complexities due to the extended distance to the subject and the free movement and vibrations in all axis of the platform.

1.8 Research Aim and Objectives

1.8.1 Research Aim

This research is broken down into a clear and simple aim and subsequent objectives to achieve this. At the very highest level, this research questions the most effective way to detect changes captured through 2D aerial imagery once processed into richer 3D data that has more context.

The ultimate aim is to understand what real change in the presence of noise is and the best way to detect this in a complex environment. This work will contribute to the field of 3D change detection to compare and contrast other point descriptor techniques, and present the most effective way of presenting the end user with true change.

1.8.2 Objectives

To enable and meet this the aim, a series of objectives are presented. These are now defined and discussed.

Objective 1

Develop realistic scenarios of challenging dynamic environments.

To detect change in realistic dynamic real world scenes, this research will go to gather 2D data from highly challenging scenarios whereby change is difficult to detect. Specifically, dynamic real-world environments with small changes. Scenarios should be designed to fulfil each of the combinations of structure and colour to form a detectable change. To enable these specific situations and to constrain other unwanted variables, discrete scenarios should be considered, planned and data gathered as part of this research. That being:

- Scene changes with complex and non-uniform structural change and very little change in colour.
- Scene changes with combinations of structural changes and colour changes.

- Scene changes with very little structural change but changes in colour.

Creating and capturing data allows full control over the magnitude of change, type of change and allows the change to be measured accurately. This is in contrast to publically available datasets where this is not a controlled source and the specific requirements of the scenario are not able to validate the aim.

Objective 2

Process 2D captured data into accurate 3D data that can be used efficiently to allow accurate change detection.

Data gathered from a traditional 2D cameras is required to be processed into 3D imagery in the form of point clouds. This process is a computationally expensive one and introduces noise. An effective method must be researched and developed where necessary. This is then to be applied to create the 3D data required to assess the potential of accurate change detection. Consistency of methodology should be maintained throughout this.

Objective 3

Accurately define ground truth models to enable performance assessments.

As the key facilitator of performance measurement, the ground truth should be of a high fidelity. The creation of this must be carefully considered and performed to enable successful and valid assessment of any of the experiments. For this, a technique should be researched and developed where required. This should demonstrate error and the methods used to reduce this where possible.

Objective 4

Develop or propose a change detection process and framework

The development of a change detection process and framework is required so that each of the algorithms and methods being evaluated against can have a baseline methodology and be unbiased. This will also make it much simpler to use a common framework whereby algorithms can be swapped out for each other. This means data input into the system and the subsequent analysis will be modular and reuse is possible. Where a

consistent process or framework cannot be used, for example due to the inherent nature of the technique of change detection, it should be ensured that each of the evaluations are non-discriminatory and of unbiased.

Objective 5

Develop a novel change detection technique that utilises 3D structure and colour.

The development of a novel change detection technique is to enable the detection of any said change. This should utilise the 3D structure that is derived from the 2D imagery through the structure from motion process and also utilise the colour information. Ideally this should be a tuneable technique that can be adaptable to different situations or scenarios. This could involve the development of a 3D descriptor in its own right or a full end to end change detection technique.

Objective 6

Propose a technique to quantitatively and qualitatively detect change.

A robust quantitative technique should fulfil the capability to detect change in its own right. This on its own will form a solution, but understanding this within the context of the 3D models produced is important, particularly where numbers alone cannot easily describe any outliers or behaviours of an algorithm. Visualising this will offer far more context, particularly when there are such large volumes of multidimensional data.

Objective 7

Compare, contrast and critique the novel technique to other state of the art descriptors.

This entire process will require a robust performance assessment whereby true change and false positive can be measured accurately. Where possible this should use common performance assessment techniques and metrics that can be used in other research pieces. Where these are not suitable or a better technique can be offered, these should be developed and understood. If a new method is proposed to best solve this problem, it should be compared with an evaluation of those most prominent techniques used for 3D description to discriminate change.

1.9 Limitations and Constraints

As with any research or development, limitations and constraints shape what is achievable. The limitations and constraints for this work have been split into three categories as follows:

- Regulatory (safety and law),
- Technical Limitations and Challenges,
- Human Resource.

1.9.1 Regulatory

Regulations exist for flying UAVs in the UK, specifically the Civil Aviation Authority CAP722 (Civil Aviation Authority, 2010). This limits anyone flying UAVs to do so safely and is adhered to through this research. They key point to take from this regulation is that to fly in the UK (where it will be most practicable for this research), there is a maximum altitude above the surface of 120m (400ft).

In simple terms, these regulations state that:

- ✓ you are responsible for flying your UAS in a safe manner
- ✓ you must keep the UAS in your direct sight at all times while it is flying, so that you can ensure that it does not collide with anything, especially other aircraft
- ✓ you must not endanger anyone, or anything with your UAS, including any articles that you drop from it
- ✓ you must not fly more than 400ft above the surface. If flying over hilly/undulating terrain or close to a cliff edge, this may be interpreted as being a requirement to remain within a distance of 400ft from the surface of the earth
- ✓ you must not fly within the Flight Restriction Zone of a protected aerodrome
- ✓ if your UAS weighs more than 7kg, additional rules apply if you fly in certain types of airspace.

Figure 1-8 - CAP 722 Generalisation from CAA

1.9.2 Technical Limitations and Challenges

Technical challenges and limitations also exist, and mostly around the ability to gather and process data.

Gathering data

The gathered data has quality and characteristics that are determined by the quality of apparatus, the accuracy/precision of gathering techniques and the conditions (environmental and geometrical) of the experiment. The quality of the gathered data is determined by the stability of the platform, and this is due to environmental conditions, platform control systems, and platform vibration.

The camera and UAV will be fixed models during this research. This is a high grade of Commercial Off The Shelf (COTS) that can be acquired at the time of research. The camera properties such as its resolution are discussed more fully in Section 3.2.1.

Data gathering techniques will be formed from current best practice and to be consistent. Where possible, the latest known to maximise performance and the efficacy of the UAV and subsequent processing will be used (as discussed in Section 3.2).

Processing data

To be able to process any data, computation is required and as more data and more complex calculations are performed, more computation power is required. This limits the amount of data or datasets that are to be analysed for this research and also ties into the resource section of the limitations. These computational limitations will play a part of this solution, where only what is capable of being processed within a "useful time" for operational use is considered appropriate, i.e. hours not months.

To limit this research to the state of the possible, the full chain of computation is to be on one node (i.e. has to be processed on a single computer). While it could be distributed across a network for collaborative computation, this is out of scope of this project. This research will focus on the computation on only one device rather than employing cloud or high performance computing clusters over multiple nodes.

1.9.3 Human Resource

One of the constraining factors of any research is its duration and the number of hours available of human resource (and skill of that resource). Planning of the research to fulfil the objectives takes this into account and puts direct limitations on the scope of the research. To enable the objectives, some of the key phases of the research are as follows (note these may be on going and in parallel):

- Investigative period (identifying aims, objectives and project scope)
- Literature reviews and research
- Skills identification and learning
- Preliminary Design
- Implementation
- Data production and gathering
- Data processing and analysis
- Evaluation and write-up.

All these subjects take both time and effort. As with any evidence based science, the quantity and quality of data gathered helps prove any theory and detect the envelopes of performance. This research is being performed by a single research student within the required timelines of a PhD research thesis. The number of datasets that can be gathered is limited by this time, but must also be significant to verify and validate the research question and subsequent proposed solution.

1.10 Structure of thesis

This thesis contains six chapters, of which are organised in the following structure:

Chapter 2

This chapter forms a literature review of feature matching, structure from motion, photogrammetry, UAV imagery and change detection. This offers an understanding of the existing research that is relevant to this study.

Chapter 3

A comprehensive discussion into the methodologies used within this research and the decision making process that led to their development and/or use. This chapter gives a full understanding of the how and why of this research.

Chapter 4

This chapter presents the results of the research for each of the datasets. This comprises the entirety of data capture through to analysis and visual representation of data.

Chapter 5

The fifth chapter of this research details the analysis and interpretation of the results and any discussion points around the entire research.

Chapter 6

Finally chapter six concludes the thesis and research. Here, a summary of the findings and the resulting conclusions are presented, and further work is also identified.

Chapter 2

Literature Review

The chapter discusses the relevant works surrounding the topic of this research. It has been broken down into sections to better allow an understanding of the topics of research that have taken place and how they are each extensive in their own right. This thesis goes to discuss those that are state of the art in these fields and those that are most significant in the development of the subject.

2.1 Feature descriptors

A feature descriptor is used to create a portrayal of a point and its surroundings. Usually a set of keypoints should have been detected and are required to be described. Generally this description of the point is through a numerical algorithm and used to recognise that unique point in other images or differentiate from other points. Feature descriptors can usually be split into two categories, local and global. One to describe a point and its local neighbourhood and the other for a global context of any point. This research focusses on point to point analysis and hence local descriptors are of more interest. That being said, global descriptors are also evaluated so that other applications of them can be considered. Descriptors are also split into the type of imagery it is to describe; 2D and 3D. While this research focusses on the analysis of 3D imagery, it is 2D descriptors that are used to create the 3D point clouds. Also these were the first to develop the techniques and their 3D descendants are usually based on their 2D counterparts, and hence are still of significance

to this research. Also the more recent advent of using machine learning to optimise and even learn descriptors is becoming more established. As this is not specifically the fundamental descriptor itself, the topic warrants its own section in this research (Section 2.4).

2.1.1 2D Descriptors

Research into feature descriptors stemmed from 2D imagery analysis, particularly the correspondence of points between images and their flow. This forms part of the SfM pipeline and as such, these will be discussed first. The Horn–Schunck (Horn and Schunck, 1981)and Lucas–Kanade (Lucas and Kanade, 1984) methods were some of the key early techniques to explain the optical flow problem, both being widely used methods that considered pixel and its local neighbourhood. While much has changed in the fields of image processing and computer vision, these presented primary pathways for the need for features to be described. (Shi, 1994) went to look at some of the features that were available and which of those were good to track and is a key work to understand what is required from a feature. At the time, this offered the first look at what is a good descriptor. This was written agnostically to the application and hence still is very relevant today to understanding what a good feature is.

It wasn't until the major breakthrough came in 1999 when David Lowe presented the Scale-Invariant Feature Transform, SIFT (Lowe, 1999) and shifted from relatively simple sum of squared distance techniques. SIFT has turned out to be one of the key research findings in 2D feature descriptors and image processing as a whole, spurring a large amount of other research and feature detector techniques. (Mikolajczyk and Schmid, 2005) gives an extensive study into the performance of feature descriptors, attempting to rationalise the many descriptors that had been presented and was a key baseline paper in feature detection and description performance. The fast moving nature of research into this field quickly rendered this paper out of date, with newer and faster methods being developed with different purposes. One significant method inspired by SIFT, is Speeded Up Robust Features, SURF (Bay et al, 2006). This lowered the computational cost of feature detection, and the SURF's Fast-Hessian detector is said to be about three times faster than the Difference Of Gaussian (DOG) method used in SIFT, and about five times

faster than Hessian-Laplace. SURF offers one of the most usable descriptors, particularly when computational performance is paramount, with little sacrifice in accuracy. (Tuytelaars & Mikolajczyk, 2008) again looked at the performance of feature detection and description, and gives a precise definition of what exactly features and descriptors are. (Juan and Gwun, 2009) also looks at performance, specifically comparing SIFT, SURF and PCA-SIFT (Ke and Sukthankar, 2004). Here, it is concluded that although SURF is by far the fastest, SIFT is more invariant to rotation, scale and affine transformations. Since the arrival of SIFT and SURF, there have been diverse interpretations of implementation resulting in varied performance. This causes obvious problems when considering the performance of descriptors against others. (Abeles, 2012) researched the ambiguities in eight of the most commonly used SURF libraries, and found within CPU implementations an order of magnitude in their computational efficiency. The OpenCV library (Bradski, 2000) provides one of the most complete sets of algorithms for feature detection and description, also giving a very complete documentation into their implementations. (Krig, 2016) provides the most credible and up to date and in-depth survey of 2D feature descriptors and raises discussion on the recent advent of 3D descriptors and the extension of 2D feature descriptors for this. This is the current baseline for 2D descriptor use.

2.1.2 3D Descriptors

LIDAR and SfM has meant that 3D imagery has become a much more common medium in the last decade, and the description of these points is also required. As with 2D descriptors, 3D Descriptors are generally used for the correspondence of points between either a recognition library or for tracking between frames of 3D imagery.

One of the early adopters from 2D description to 3D description was 3DSIFT (Scovanner, Ali and Shah, 2007) for the application of action recognition. 3DSIFT has notably been used successfully for object recognition in CT scanners (Flitton, 2010) and medical imagery (Allaire, 2008). (Tombari, Salti and Di Stefano, 2013) performs an evaluation into the performance of 3D keypoint detectors, with a heavy emphasis on 3D shape retrieval and 3D object recognition. It appears there is a common theme in research led by real world requirements, whereby 3D descriptors are used for retrieval and recognition.

(Zhong, 2009) presents the Intrinsic Shape Signatures (ISS) method, again a technique primed for recognition. ISS is primed to facilitate fast pose estimation though identifying shape patches from different viewpoints. While this proves useful in the SfM process and specific tasks for 3D pose estimation, it appears this is not optimal for understanding differencing. (Filipe, 2014) compares 3D keypoint descriptors in RGB-D datasets, focusing on their invariance to rotation, scale and translation, concluding that 3DSIFT out performs other prevalent 3D descriptors (Harris3D, SUSAN, ISS3D, Lowe, Noble and KLT). An analysis into learning a descriptor specific 3D keypoint detector in (Salti, 2015) found that there is a lack of complimentary detectors and descriptors. The topic of machine learning is discussed in more depth in Section 2.4. (Alexandre, 2012) gives a comparative evaluation of 3D object recognition using 3D descriptors. It gives an excellent overview of the implementation of descriptors and concludes that there are wide variations on descriptor performance depending on the application so they should be carefully matched to the task. The paper identifies colour being of particular use as an extra discriminator and offers a performance advantage for object recognition. It also goes to note that CSHOT or SHOT Colour (Tombari. 2011) provides a good overall balance of recognition performance. From Tombari's CSHOT paper, the addition of colour to the SHOT descriptor (Tombari, 2010) is said to improve the accuracy of SHOT, later confirmed in (Salti, 2014). While developing CSHOT, extensive testing of colour spaces in L_1 norm found the CIELab to be the most effective. This is of particular interest to this research and taking colour into context with the 3D descriptor is key to the investigation. (Blauensteiner, 2006) and (Ganesan, 2010) had earlier confirmed this when reviewing colour spaces for change detection applications.

(Hana *et al.*, 2018) gives a comparative and comprehensive overview of the 3D descriptor field (13 in total - 8 local and 5 global). This gives a very useful indication of the most common descriptors, and discusses their applications and downfalls such as the sensitivity and robustness of those for the application of recognition. That said, it is not as comprehensive as its title suggests, paying attention to only those that are the most common. For most applications though, this gives a really good insight into what is out there. SHOT, CSHOT and PFH (Rusu et al., 2008) are evaluated to offer the best performance in the recognition scenarios. (Guo *et al.*, 2016) looks at a performance evaluation of just local 3D descriptors and is hence more relevant to this research. Also it

extends beyond recognition and retrieval into 3D shape and 3D modelling context. Here, 10 local descriptors are evaluated, finding FPFH (Rusu et al., 2009) to the best for time-critical applications with SHOT again providing a more accurate solution. This paper is the current baseline for 3D descriptor evaluation. While most research and these evaluative research papers focus on recognition tasks, a decision is made to compare and contrast those found to be most performant at that. Hence the following descriptors are chosen in this research for the comparative evaluation in change detection:

- PFH
- FPFH
- SHOT
- CSHOT/SHOT Colour

The Point Cloud Library (PCL) (Rusu et al, 2011) is an open project for 3D image and point cloud processing. It contains many of these most prevalent 3D point descriptors such as PFH, FPFH, SHOT and CSHOT.

2.2 Structure from Motion

As discussed in Chapter 1, SfM is a process of discrete steps that each have their own research space. Although it was generally understood that 3D vision is achieved by binocular disparity in humans, (Wallach and O'Connell, 1953) was the first to fully explore how 3D form could be established using monocular vision. (Ulman, 1979) pioneered motion based reconstruction and the computational application of this. (Oliensis, 2000) reviews modern approaches for SfM and puts forward a framework for designing algorithms within this research space. (Schonberger and Frahm, 2016) again looks to review the entire end to end process, and identifies some of the downfalls with SfM that introduce noise and inconsistency. Specific use cases of SfM for aerial imagery and its subsequent accuracy is discussed in Section 2.3. (Forsyth and Ponce, 2011) gives a comprehensive overview of the generalised process and some of the algorithms involved. (Hartley and Zisserman, 2003) is considered one of the key texts in multiple view geometry and offers in-depth algorithmic level explanation of the process. This is

the most complete text available regarding multiple view geometry and offers comprehensive understanding of a majority of the SfM process. The process follows a structured flow to creating a 3D scene understanding from 2D imagery and detailed in the diagram below (Figure 2-1). Each process forms their own topics of research and are discussed.

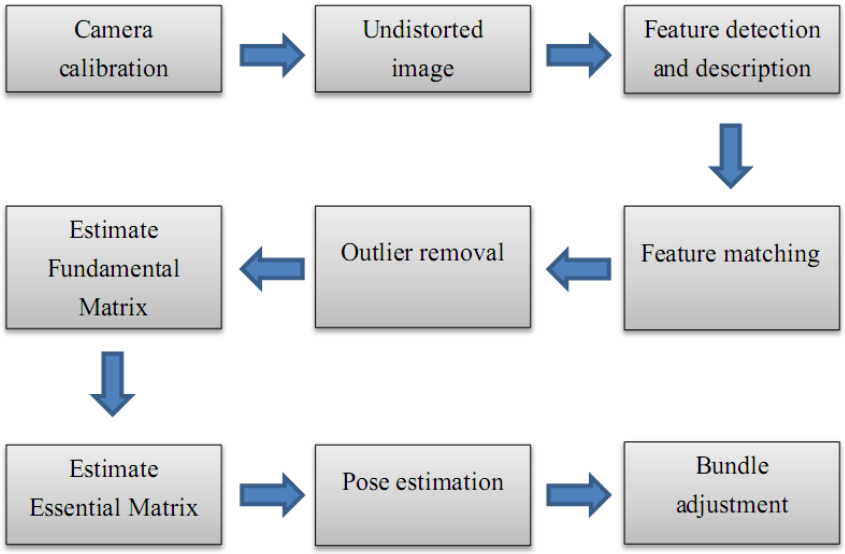


Figure 2-1 - Structure from Motion Process Flow

2.2.1 Feature Matching

Once features have been detected and described in a scene, they can be matched to those from another scene. This is a computationally expensive process and different methods have been presented to try and solve this. (Szeliski, 2010) splits this problem into two components, matching strategy and, data structure and algorithms. The simplest method, linear matching, goes about calculating the distance from one feature point to the other points in the index, keeping an updated history of the nearest point. This is defined as nearest neighbour searching, and a few different techniques have stemmed from this. (Arya et al, 1998) presents a method using the approximate nearest neighbour (ANN), increasing the speed of this practice with only minor loss in the accuracy. Later, (Muja and Lowe, 2009) investigate some of the techniques used for ANN and also propose their own, "*searching hierarchical k-means trees with a priority search order*". In their paper, they conclude that a there is no one best method for ANN, and so propose a system where

the best algorithm can be selected automatically and configured appropriately depending on the dataset. This library was coined, Fast Library for Approximate Nearest Neighbours (FLANN).

2.2.2 Essential and Fundamental matrices

Once features have been detected, described, matched and outliers removed, the arrangement of epipolar geometry can be estimated. This can be described in two ways; by the essential matrix in a calibrated system and by the fundamental matrix in an uncalibrated system. (Longuett-Higgins, 1981) and (Tsai and Huang, 1984) proposed the simplest solution to calculating the essential matrix by means of the 8-point algorithm. Here, 8 or more known point correspondences from two calibrated camera views are used to solve a set of linear equations. The elements in this can then be factorized by Single Value Decomposition (SVD) to give a unique solution (Hartley and Zisserman, 2003). This relatively simple method makes implementation easy and computation fast. (Faugeras, 1992) and (Hartley, 1992) show this method can also be applied to an uncalibrated camera to solve the fundamental matrix. Early analysis in (Bolles et al, 1987) goes to criticise the stability of the 8-point algorithm, stating "his technique has some stability problems and does not seem to have won over many photogrammetrists." This is attributed to the high sensitivity to noise as the algorithms inputs are assumed to be perfect. (Ma et al, 2004) discusses methods for optimal reconstruction by applying least squares criterion the 8-point algorithm. This allows more than 8 points to be considered. (Hartley, 1997) also discusses least squares, and goes to demonstrate that there was a common misconception about the use of the 8-point algorithm. He proposes the noise sensitivity can be solved by a simple preconditioning of the points by transformation, coined the normalized 8-point algorithm, and appropriately titles his paper "In Defence of the 8-point Algorithm." Although the 8-point method is the simplest solution to estimating epipolar geometry, non-linear approaches may be applied to solve the fundamental matrix. As seen in Section 1.5.3, the fundamental matrix has 7 degrees of freedom. This allows it to be solved with only 7 point correspondences enforcing the rank 2 constraint (Luong et al, 1993) and (Hartley, 1997). Although the method applied to this is similar to the 8-point algorithm, the fundamental matrix is always solved to either 1 or

3 solutions when using the 7-point algorithm. One advantage to using fewer points is that there is less chance of selecting noisy correspondences. (Armangué et al, 2002) goes about comparing these methods in calculating the fundamental matrix offering an overall view of the subject.

2.2.3 RANSAC: Inliers and Outliers

As discussed in Section1.5.1, the epipolar geometry of a corresponding pair of features can be used to calculate the essential and fundamental matrices. This assumes that the data being input is correct, while the nature of the Approximate Nearest Neighbour (ANN) matching technique gives light that this is not always true. This brings forth the task of deciding which of the proposed match's inliers are and which of those outliers to this model are. (Fischler and Bolles, 1981) introduced, RANdom SAmple Consensus (RANSAC), ideally suited to this problem within image processing. This is a robust estimation algorithm using an iterative technique for "interpreting/smoothing data containing a significant percentage of gross errors." (Hartley and Zisserman, 2004) gives an overview of this algorithm and also describes other methods such as Least Median of Squares (LMS), along with their advantages and disadvantages. Philip Torr has continued to investigate this field and has proposed MSAC (Torr and Murray, 1997), MLESAC (Torr and Zisserman, 2000) and IMPSAC (Torr and Davidson, 2003), to try and resolve some of RANSACs shortcomings. (Chum and Matas, 2002) looks specifically to reduce the average computation time, proposing Randomized RANSAC (R-RANSAC). (Nistér, 2005) describes R-RANSAC and standard RANSAC as depth-first methods whereby one hypothesis evaluation is completed before the next is initiated. Nistér builds on this and finds another method more appealing whereby only a set number of hypotheses are considered. (Nistér, 2005) describes this as a breadth-first pre-emption method, and coins it, Pre-emptive RANSAC. He proposes this is "powerful enough to enable structure and motion estimation in real-time". Since its arrival in 1981, there have been many proposed extensions and improvements to RANSAC, (Meer, 2006) goes about documenting some of these, and also considers some of RANSACs shortcomings. (Raguram et al, 2008) gives a fuller analysis and comparison of some of the more prominent RANSAC algorithms. Here, pre-emptive RANSAC is noted as being non-adaptive, and being

subject to a good estimation of the inlier ratio. (Raguram et al, 2008) proposes a solution to this, Adaptive Real-Time Random Sample Consensus (ARRSAC), describing the method as partially depth-first. (Frahm et al, 2010) uses this to good effect in large scale scene reconstructions with time constraints in mind. (Choi et al, 2009) also looks at comparing the RANSAC family, and rationally suggests that algorithms should be tailored to their application as there will always be a trade-off between accuracy, speed and robustness (as seen in Figure 2-2).

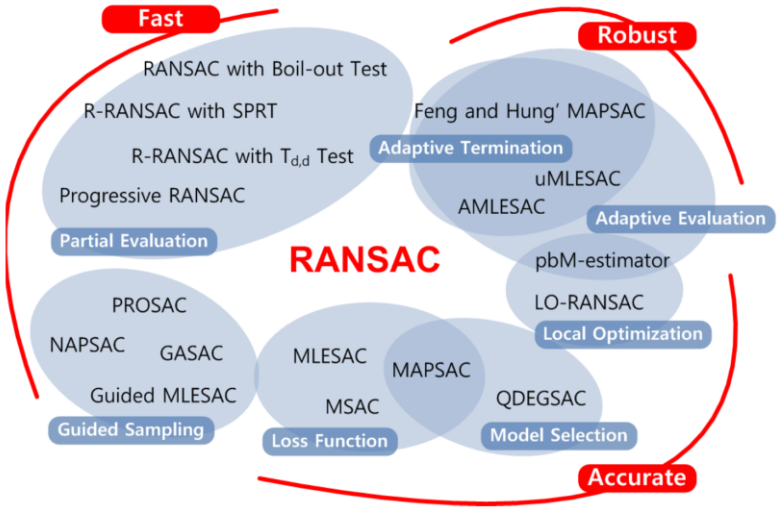


Figure 2-2 - RANSAC Family (taken from Choi, 2009)

2.2.4 Bundle Adjustment

When a large number of images are considered to reconstruct from, the process of bundle adjustment can be applied to reduce pose error across the whole set and is generally said to be the last process in SFM. It refers to the bundles of light rays leaving an object and gathering on the camera centres. To reduce the pose error in the calculation of the camera centres, bundle adjustment uses a process of non-linear least squares. The most popular technique for this is to use the Levenberg–Marquardt (LM) algorithm, (Levenberg, 1944) and (Marquardt, 1963). A key modern paper on bundle adjustment, (Triggs et al, 2000), provides a survey of methods and the theory behind them. One of the aims of the paper was to address a misconception that the bundle adjustment process is computationally expensive. (Lourakis and Argyros, 2009) describes the design and implementation of a generic sparse bundle adjustment software package, SBA, of which has now become very

popular based around the LM algorithm. (Agarwal et al, 2011) provides a look at the use of bundle adjustment in very large scale image collections, and propose a method better suited for this application, although also conclude that SBA is of very high quality.

2.2.5 SfM Software

Since the relatively recent advent of having the computational power to create accurate and large scale models by using SfM, various off the shelf software packages have been presented encapsulating the techniques discussed above and optimising them for use with GPUs. (Gómez-Gutiérrez, 2015) looks at two of the most common of these packages; 123D Catch (Autodesk, 2016) and Agisoft Photoscan (Agisoft, v1.3.3, 2017). Gómez-Gutiérrez found that Photoscan produced more accurate and denser point clouds in 11 out of 12 cases. This paper is discussed further in Section 0 for the techniques used to compare the performance of these point clouds. (Verhoeven, 2011) uses Photoscan for their surveying in aerial archaeological work to good effect but do not go to measure its performance. (Peterson, 2015) presents a whitepaper on SfM for the U.S. Bureau of Reclamation. Here, seven different cases of the use of Photoscan are discussed, from small scale to very large scale. From its aerial survey case, data is gathered using very high quality cameras from helicopters at 800ft altitude. They found accuracy of GPS markers to be as high as 3.4cm rms. (Kersten, 2012) goes to compare SfM software for archaeological work, including open-source Bundler/PMVS2 (Furukawa and Ponce, 2010), Autodesk 123D and Photoscan. There appears to be very little accuracy performance difference between all three but their work seems flawed as only one dataset is computed using Photoscan. (Brutto, 2012) produces a baseline model using LIDAR and compares SfM software. Again Photoscan appears to be consistently most accurate. To maintain consistency in the processing throughout this research, a single solution is required. For this purpose and from the research discussed, Photoscan is down selected to fulfil this. (Bianco, Ciocca and Marelli, 2018) evaluates the performance of SfM pipelines and looks to identify and increase efficiency of its components. Research and Development into SfM is a continuous process currently with regular software updates, including that to Photoscan. To reduce the effect shifting processing accuracy in different

versions, a single baseline version is consistently used throughout this research (Agisoft, v1.3.3, 2017).

2.3 Photogrammetry and Aerial Imagery

The use of SfM for photogrammetry is now a viable approach due to the accessible and affordable software and accuracy of results. The matching of this with small UAVs with lightweight cameras makes for low cost solutions to large area photogrammetry. (Smith, 2015) confirms this, reviewing SfM for photogrammetry in physical geography from aerial imagery. They conclude that SfM has revolutionised photogrammetry and at lower altitudes (under 100m), SfM is comparable to Terrestrial Laser Scanning (TLS). This presents SfM as a viable data gathering technique for 3D imagery in this research. They also note some of the shortcomings of using SfM for photogrammetry, such as areas devoid of features (such as ice and water) and shiny surfaces giving low point densities. Some other particular limitations noted is that they do not favour dynamic environments, such as those with wind effects on vegetation or changes in lighting conditions. This identifies an additional constraint into this research whereby consistent environmental conditions should be used to reduce this effect. That noted, it's unlikely that a small UAV will be flown in conditions where there is any significant wind. (Westoby, 2012) again reviews SfM for aerial photogrammetry as a low cost tool. It notes the significant advantage that aerial imagery has for remote or inaccessible regions. (Ostrowski, 2014) analyses the point cloud generation from three anonymous SfM packages specifically designed for aerial photogrammetry. The 36 Megapixel imagery data is gathered from altitudes of 135m, and while it is frustrating to not know which packages they were, it confirms that point accuracy from this range is around 10cm. (Harwin and Lucieer, 2012) assesses the accuracy and applicability of aerial SfM for photogrammetry in three natural landscape scenarios. The research concludes that using 12 Megapixel images subsampled to 3 Megapixels to reduce computation costs can yield accuracy of 2.5-4cm from a range of 40-50m. It also recognises that SfM based aerial photogrammetry could become very

useful for monitoring changes such as that in coastal erosion. This is highly significant to this research, as it not only identifies SfM being accurate at low altitudes, but also its potential use for monitoring changing scenes. The paper has serious limitations in that the research presents only a qualitative analysis through visualisation. There is no explanation of why this is, but by pure assumption it would either be through lack of time or it could be due the difficulty in creating a credible ground truth model. Nevertheless, the limitation is identified and highlights that this is something that should (and in this research is) quantitatively analysed.

2.4 Change Detection

Change detection allows us to identify differences in an object or phenomenon by observing it at different times (Singh, 1989). The ability to detect change allows us to focus on a specific area rather than looking at a larger picture. This is true for our visual perception and offers the end user of computation systems to identify areas of interest. (Lu, 2004) and later (Hussain, 2013) gives an in-depth review of generalised change detection techniques and breaks them down into 6 levels of increasing complexity as follows; Algebra, Transformation, Classification, Advanced Models, GIS and Visual Analysis. The research then goes to break these levels into further levels, starting from pixel image differencing and ending at human visual interpretation offering a the reader a good understanding of what change detection is and how it can be analysed. Further to this, the selection of thresholds for these methods are discussed, along with their accuracy and finally concludes that all methods of change detection are dependent on the prerequisite geometrical registration, correction and normalisation of their input data.

Due to its high accuracy, there are many recent instances of change detection using LIDAR imagery for geotechnical analysis. (Zavodny, 12) offers a change detection system for aerial LIDAR data in urban environments. In this full study, two methods are considered for detecting the changes; the first uses the octrees to compare segments of the scene and the second looks to compare the points themselves and thresholds an area where change is larger than the rest. (Xiao, 12) also goes to detect changes in aerial

LIDAR imagery, particularly changes in vegetation in urban environments. Here trees are detected and compared by the shape of their convex hulls. (Olsen et al, 12) uses a combination of LIDAR and optical imagery for real-time for earth movements. Here the LIDAR Change Engine (Lichen) framework is used for point-by-point comparison of a scene and a threshold used classify if points are out of bounds. Reviewing this literature, it is found while the technique of change detection in LIDAR imagery will be similar to using optical imagery, many of the constraints of noise are not present. (Barber, 08) again uses aerial LIDAR imagery as a dataset for point cloud comparison. Here octrees are used to detect changes in two models, with "some limited success". The report documents many "insignificant" changes such as that in vegetation growth and change related to errors in data gathering.

(Shi, 2011) identifies how useful low cost aerial imagery can be for the purposes of change detection and proposes the use of 2D SIFT features for object based change detection. This appears effective to detecting the change but no performance metrics are given. Also it is noted that the datasets used in this research appear to have been pre-processed to be well aligned in both scale and rotation. (Champion, 2007) focusses on 2D building change detection and utilises ROC curve analysis for determining the performance of the binary classification problem.

Ground based change detection technology seems to be more refined, both because the platform gathering the data can remain static, and because there are no weight constraints both serving to reduce noise. A key paper in 3D point cloud change detection, (Girardeau-Montaut, 2005) uses ground based LIDAR (Terrestrial Laser Scan) for the change detection in building sites. Three strategies are put forward using octrees as a data structure; average distance, best fitting plane orientation and Hausforff distance. (Lague, 2013) another TLS example, coins this and similar methods Cloud-To-Cloud (C2C) comparison. This research also describes a Cloud-To-Mesh (C2M) used for comparing the resultant point clouds to a reference 3D mesh or theoretical model, stating that his isn't always possible either due to not having a mesh, or the process of creating one too noisy due to surface roughness. The paper aim to solve this by proposing the Multiscale Model to Model Cloud Comparison (M3C2) technique, which takes the local distance between two point clouds along the normal surface direction. This provides a very good

solution to the high fidelity TLS data quoting an order of magnitude reduction in the errors found from C2C methods.

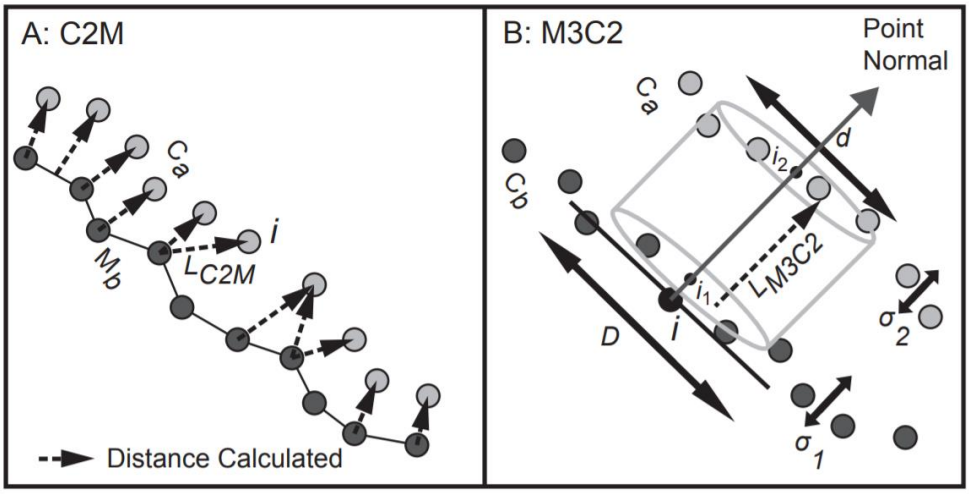


Figure 2-3 –C2M and M3C2 Change Detection Techniques

A) Cloud to Mesh (C2M) B) Multiscale Model to Model Cloud Comparison (M3C2) – (Taken from Barnhart 2013)

(Barnhart, 2013) goes to compare the C2M and M3C2 methods in TLS data and finds that M3C2 offers a better understanding of the uncertainty in roughness and scan registration in TLS imagery. This brings us back to (Gómez-Gutiérrez, 2015) mentioned in Section 2.2.5. This paper no only goes to compare the SfM software packages as discussed, but also compares C2C and M3C2 techniques with both TLS and SfM gathered data. It concludes that both C2C and M3C2 offer similar change detection performance in 12 of the 12 scenarios performed. (Kromer, 2015) later goes to confirm this is with TLS data only.

From the analysis of current literature in the field of change detection, particularly that with reference to SfM and aerial imagery, there appears to be very little research into 3D change detection from UAVs using SfM. The research that has taken place in this field to date has simply looked to use SfM as a tool to gather 3D data from aerial imagery and not further attempt to implement change detection from this.

2.5 Machine Learning

An explosion of developments in Machine Learning (ML) in the last two decades is revolutionising computer vision. In many areas of research and its subsequent applications, there is a shift away from exclusively using statistical and image processing techniques to the additional use of machine learning. This is particularly evident in the field of image and object detection. To better understand the context of machine learning in this research, it is categorised into two types:

- Classification and recognition
- Machine Learning for change detection

2.5.1 Classification and recognition

(Krizhevsky, Sutskever and Hinton, 2012) won the 2012 ILSVRC (ImageNet Large-Scale Visual Recognition Challenge) whereby contestants train on the ImageNet dataset (Deng *et al.*, 2010) for object recognition. This is one of the most significant papers in the field and was the first to use Convolutional Neural Networks (CNN) in the ILSVRC. The performance error of recognition over large datasets was found to be reduced from 26% previously to 15%. Following this, CNNs have become widespread in their use with image processing. The continuing trend of a more active use in recognition tasks is also common in the use of machine learning, and this is emphasised by the inherent nature of machine learning.

(Ioannidou *et al.*, 2017) surveys some of the most recent advances in computer vision with 3D data, again with the specific applications of classification. It identifies that while machine learning had mainly been used for 2D applications in object recognition, the advent of using similar methods in 3D imagery is relevant and viable but at a substantial computation expense. Another key theme identified is that of the use of colour information in 3D description of points and that it presents a significant advantage for recognition of objects. It is noted that this comprehensive analysis of machine learning and 3D data does not discuss any change detection methodologies or application.

2.5.2 Machine learning for change detection

More recently CNNs are appearing to be used to facilitate change detection in imagery. (de Jong and Bosman, 2018) explores Unsupervised Change Detection in Satellite Images Using Convolutional Neural Networks and through image differencing using CNNs achieves a 91% change detection accuracy for very high quality 2D imagery containing little noise. (Varghese *et al.*, 2019) again utilises CNNs for change detection in 2D imagery but for ground based datasets. For a performance assessment, Receiver Operating Characteristic (ROC) curves are used to present a 98% accuracy and 89% area under the ROC curve. These papers present a viable and performant method for change detection in 2D imagery, both from the ground and in aerial imagery.

(Zhang *et al.*, 2018) further expands the use of CNNs in aerial imagery. A unique data creation approach is presented whereby 2D imagery is processed into 3D point clouds through SfM. This is then transformed into a grayscale Digital Surface Model (DSM) and back into 2D image patches. A change detection accuracy of 86% is presented and notes fast computation times in the actual change detection through CNNs, but the arduous transformations are noted and their computation costs described as expensive. The reduction of rich 3D colour models to greyscale 2D image patches allows for the efficient CNN computation, but at a cost of losing informative colour data and introducing additional error in transformation processes.

Chapter 3

Methodology

This chapter describes the analysis, design and decisions made during the research and development process. The techniques used for gathering and producing both synthetic datasets and real world datasets are discussed. Following this, the methodology for describing points in the 3D point clouds are discussed, followed by the 3D Structure and Colour (3DSAC) methodology.

First, an activity chart of the end to end process is presented, breaking down high level activities and their design into four main areas as follows:

- Data Gathering
- 3D Description
- Change Detection
- Performance Assessment

Each of these are comprehensively discussed followed by a visualisation technique for a qualitative assessment.

Finally a full end to end example of the data analysis technique is described. This demonstrates the large quantities of data and analysis that is required for each descriptor.

3.1 Activity Flow Chart

To better understand the process of activities performed from end to end, a simple flow chart is presented below in Figure 3-1. Planning as a flow chart allows for the separation of activities so they can be broken down further into their components for design and implementation. It can be seen that an interchangeable data source can feed into the rest of the change detection process. In this methodology, one being a real world data source (through SfM) and one being a synthetic data source and are discussed in Sections 3.2.1 and 3.2.2 respectively.

A unique combination of both 3D structure and colour (3DSAC) is used with a mean differencing technique to compare the descriptors. A method for measuring this is also presented, whereby the performance can be reduced down to single values and hence descriptors can easily be compared and contrasted. Each of these subjects are discussed in more detail over the following chapter.

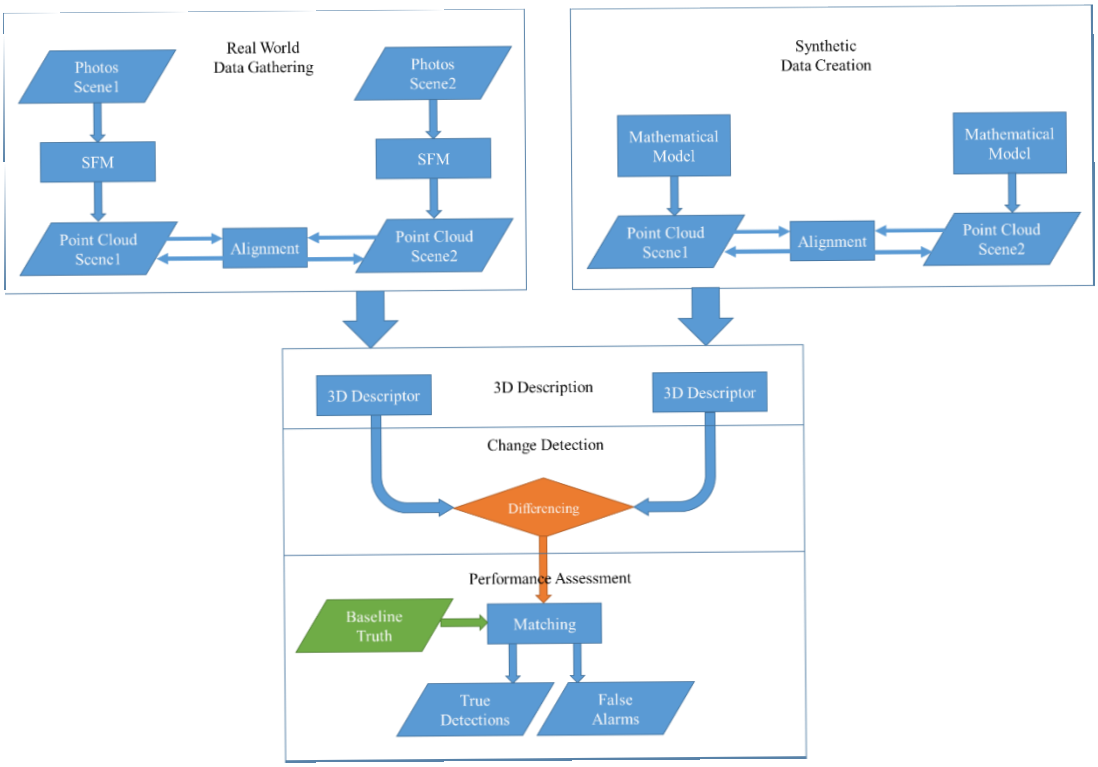


Figure 3-1 - Change Detection Process Flow Chart

3.2 Data Gathering and Production

This section discusses the methods used to gather and produce data. Data that is fed into the rest of the system is either real world data or synthetic modelled data.

It was evident from the review of literature and analysis of datasets that for real world data, most were for object recognition and featured sets of many objects. For the specific use cases defined in this research, data was required to be gathered through carefully planned and bespoke exercises for this research. This enables the exacting understanding of what the changes are and how these can then be converted into truth models.

This section discusses the apparatus and the process used to gather 2D aerial imagery data and then process to 3D point clouds. For the measurement and performance assessment of change detection methodologies, a careful consideration of what change shall be introduced is required. Finally, this section discusses the process used to create a 3D scene reconstruction of each of the datasets.

3.2.1 Synthetic Data Production

To enable fast prototyping and debugging of the software and algorithms, a reduced dataset is produced with the same data structures as expected in the real world datasets (XYZ position and RGB colour). The arrangement of this data set is to show an incremental increase in change so that differing levels of change can be detected, measured, compared and displayed, both in colour and structure. It should be noted that this model is not representative of real world data, and hence purely for development and is not a model for performance evaluation. A synthetic data set allows the system to be fully constrained and with no variables other than those desired. As this is produced computationally, it requires no lengthy data capture process or SFM.

Baseline Model

This model consists of a square 10,000 point matrix (i.e. 100 by 100). All points are coplanar in the z axis and evenly distributed in a Cartesian grid in the x-y axis. All points are defined as only red (i.e. RGB = 255, 0, 0) as seen in the figure below (Figure 3-2).

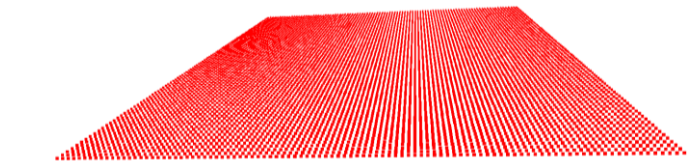


Figure 3-2 - Synthetic Data Baseline Model

Change Model

This synthetic change model again consists of a grid of 10,000 points, although this time a random but increasing weighted value was given to the z values and all points coloured blue (i.e. RGB = 0, 0, 255). This gives a slight wedge shape seen in the diagram below (Figure 3-3). This allows for an understanding of various distances for change detection measurements in a single axis.

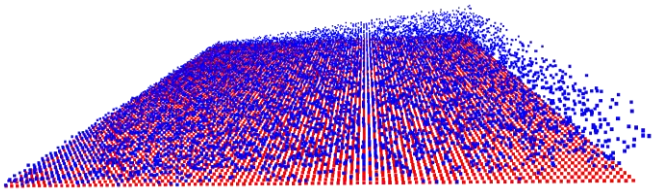


Figure 3-3 - Synthetic Data Change Model

3.2.2 Varying L_p for distance measurement

The synthetic dataset baseline and change models are perfectly aligned and contain exactly the same number of points, making it simple to compare and detect change. Although this is a simple scenario, there are high levels of randomness from point to point and proves valuable in testing the ability to distinguish levels of change. It has been used as a preliminary assessment of the methods of distance measurement and the effect this has on the ability to discriminate distances and change distribution. Here the p-norm is varied between value of one to four and the histograms of these can be seen in Figure 3-4.

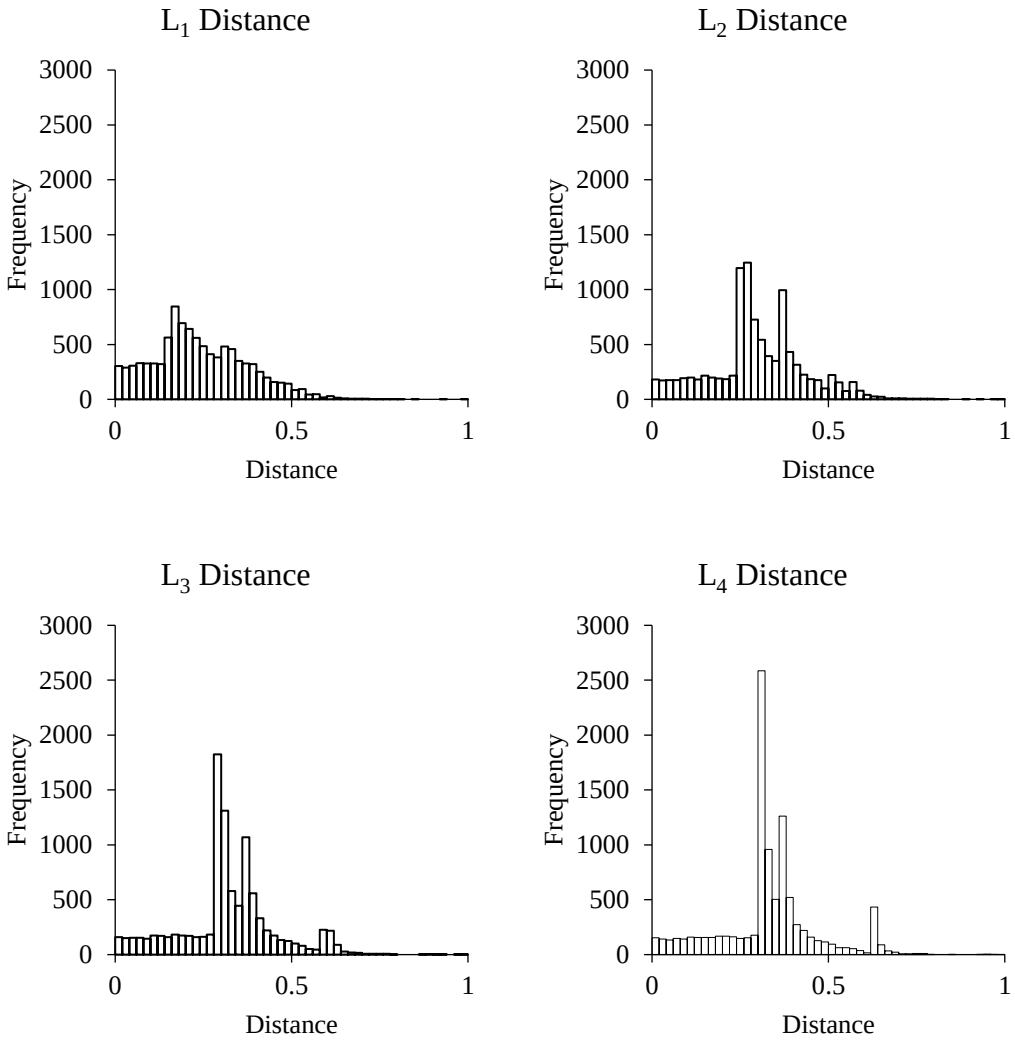


Figure 3-4 - Histograms of variations in L_p distances with Synthetic Data Set

The effect of varying the Lebesque *p* value has the effect of either bunching or spreading the change as seen in the histograms. As the *p* value is increased, the distribution of change is focused, and at a lower p, the values are spread out. As the points measured are in a 3D space, it would be suitable to keep the measurement to the Euclidean norm of L_2. While increasing the *p* value would have the effect of creating a higher change value and a lower noise floor in comparison, it could reduce the ability to detect smaller change and emphasise only larger change. A lower *p* value would have the contradictory effect of increasing the dominance of the smaller change and increase the presence of noise and artefacts. This also offers a logical Euclidean space which is coupled with the space that the 3D points are described in. As identified in Chapter 2 this is the most common method for measuring distance between any two points, and as this has been used elsewhere

presents the best comparative measure to other research. For this reason the L_2 distance will be used throughout for distance measurements.

3.2.3 Real World Data Gathering

As an overview of the process, real world data gathering first takes place where a series of UAV flights are completed gathering the 2D aerial imagery. These datasets of photos are transformed into 3D point clouds through the use of SFM and then aligned. Once 3D point clouds have been obtained, a description of each point and its local neighbourhood may take place and then compared with each other. This presents what is considered change and can then be assessed against what is the true change.

3.2.3.1 Apparatus

The gathering of real world data first appears simple but careful considerations must be taken to reduce as many variables in the data as possible. If too many variables are introduced, (such as different variations on equipment or environments) the analysis becomes complex and solid conclusions harder to make. The apparatus used in this research remains the same throughout all gathering activities. The UAV used to fly the scenarios is an AscTec Falcon 8 as show below.

Figure 3-5 - AscTec Falcon 8 UAV

The relevant AscTec specification for this is as follows:

- Payload of 500 g
- Flight time (max. payload) 16-18 min
- Total flying weight of 1.8 kg

This is a GPS stabilised UAV with accuracy specified as ±3m. This is dependent on the number of satellites detectible to the receiver.

The camera configuration is a Panasonic Lumix LX-5 with a max resolution of 3648 x 2736 (10.1MP) for stills and 1280 x 720 at 25fps for video. These are compact lightweight cameras with built-in image stabilisation that fulfils many of the requirements for UAV use.

Figure 3-6 - Panasonic Lumix LX-5

3.2.3.2 Flight Planning

For the planning of data gathering, the flight time is pertinent as a single flight is preferred to reduce changing factors in the process. To enable the maximum feature correlation between images in the SfM process, overlap is required between images. The more images taken of a particular point, the higher the accuracy of that point can be with the trade-off that the more images there are the higher the computation time. As the sizes of area surveyed are different and of different shape, the numbers of images taken will vary. To enable consistency, where possible a constant velocity of the UAV is maintained with a consistent rate of image gathering.

During the flight planning is where the consideration of what change shall take place. This needs to be carefully planned and controlled so that it can be measured and used for performance determination. For this research, three challenging scenarios have been picked each with different arrangement and magnitude of change.

Linking back to the objectives of this research, they are as follows:

- Scene changes with complex and non-uniform structural change and very little change in colour.
- Scene changes with combinations of structural changes and colour changes.
- Scene changes with very little structural change but changes in colour.

Finally, to gather the data, the UAV must be flown within the regulations of that country. For this research flights are in the UK, so UAV operation is governed by the CAA document CAP722 (Civil Aviation Authority, 2010).

3.2.4 Dataset 1 – Bunker 1

Dataset 1 is planned to implement scene changes with complex and non-uniform structural change and very little change in colour. It is a scenario comprising two epochs of data, 1A and 1B respectively. The scene contains a medium sized golf course bunker on an early March morning in the south of the UK. The surrounding area is rough long grass and the surface sand within the bunker is pitted and hardened from rain. To enable non-uniform changes with little change in colour, the sand is used as fluid material that can be moved without changing colour in comparison to its background.

Imagery data was gathered from an altitude of 50m with a variance of ±5m due to gusts of wind and GPS error.

The first flyover of the area (Figure 3-8A) created an imagery dataset containing 171 photos and the second flyover (Figure 3-8B), 15 minutes later, was gathered containing 187 photos respectively. All photos are of resolution 3648 x 2736 pixels (10MP).

The scene for the second flyover was disturbed by a small hole being dug in the sand. The change in scene consists of the hole itself and the sand displaced by this. The hole while non-uniform in shape was roughly 15cm at its deepest point and 40cm in diameter. The displaced sand took slightly more volume due to the loosening of the sand. Secondly, the spade used to dig the hole was placed away from the hole and partially occluded with sand. The spade is of a low profile, a maximum of 3cm at its deepest point as seen in the figure below and highlighted overleaf in the scene.

Figure 3-7 - Dataset 1 - Spade object

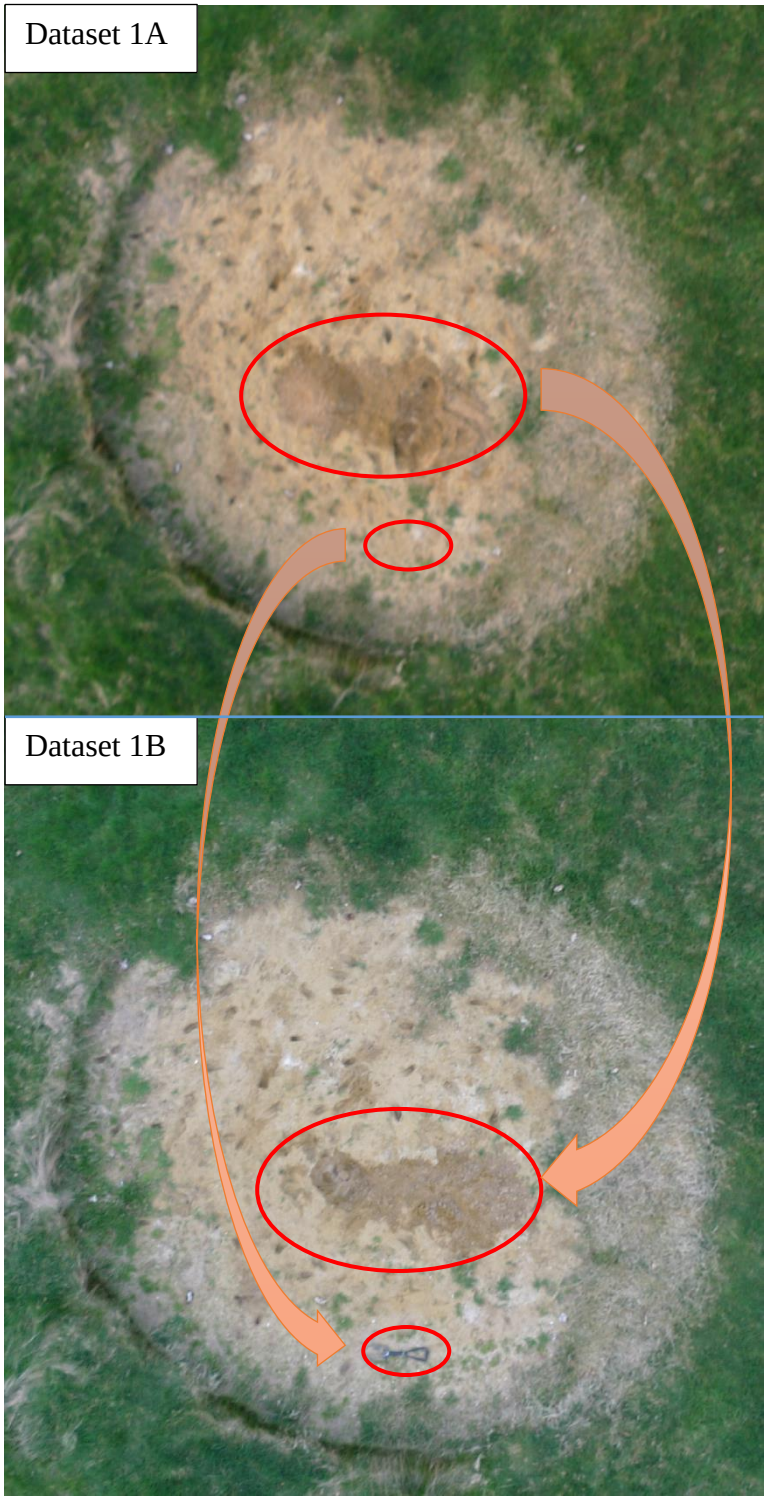


Figure 3-8 - Dataset 1A and 1B - Changes introduced between datasets 1A and 1B

The images above show the first and second flyovers with the changes highlighted. The large ellipse shows the sand change, and the smaller the spade.

3.2.5 Dataset 2 – Bunker 2

Dataset 2 is planned to implement scene changes with combinations of structural changes and colour changes. It features a two epoch dataset comprising 2A and 2B (Figure 3-9). The scene is a large golf bunker that has undergone groundworks for new buildings. This has resulted in undulations and changing textures. The granularity of substrate goes from sub-millimetre sand to chalk rocks of diameter of around 20cm. The parameters of resolution and altitude of flight is the same as Dataset 1 to give continuity of measurement accuracy. Dataset 2A contains a series of 97 photos and 2B a set of 67. Two simultaneous changes were introduced in the second epoch of data, one with a more significant structural change and one with a more significant colour change to the background scene. Firstly a chalk rock was moved. By visual inspection this would be almost impossible to detect due to little change in colour to the background. Secondly a red strap was placed into the scene. While this does not have a large structural footprint and volume, this is to stimulate the colour element of change detection within the scene.

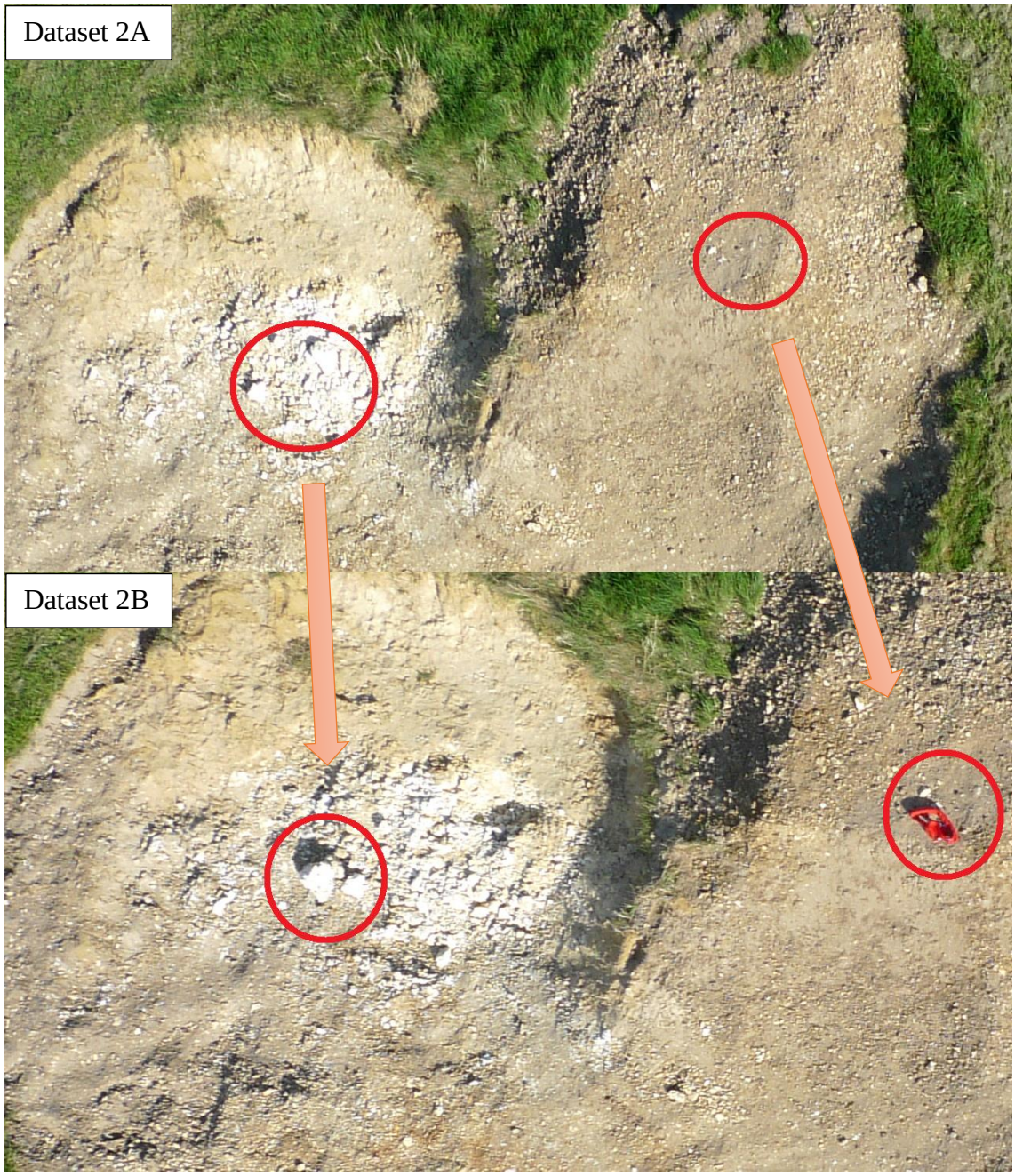


Figure 3-9 - Dataset 2A and 2B - Changes introduced between 2A and 2B

3.2.6 Dataset 3 – Ditch 1

Dataset 3 is planned to implement scene changes with very little structural change but changes in colour. It features a long ditch of about 100m in length, 3m in width and 1m in depth and fairly straight. At either end are small footbridges. As with the other datasets, flight and photographic parameters are kept consistent. Dataset 3A contains a series of 210 images and dataset 3B contains 198 images.

One change was introduced into the second epoch of data whereby a plastic sheet in the ditch was moved further towards the centre of the scene (Figure 3-10). This has a very small structural volume, but a relatively high colour change in the location of the change.

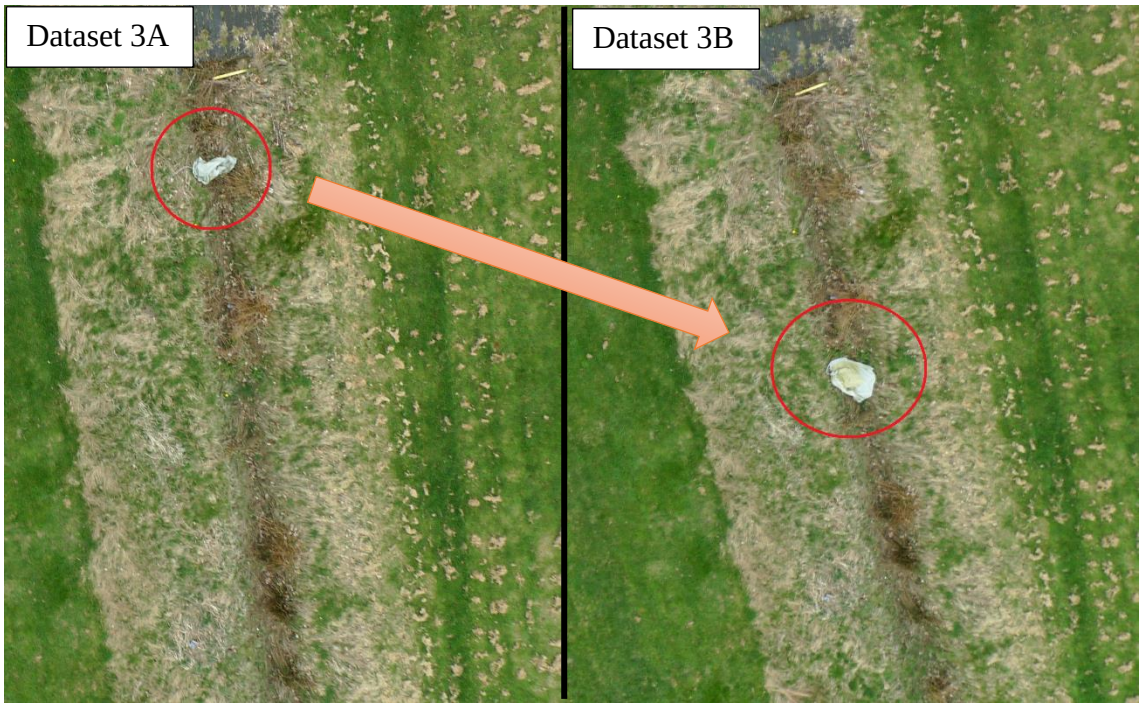


Figure 3-10 - Dataset3A and 3B – Changes introduced between datasets 3A and 3B

3.2.7 3D Reconstruction

Once 2D imagery data has been gathered, it can be go through the SfM process for 3D reconstruction. From the literature review, it became evident that Photoscan was the most widely used and most likely the best software for aerial SfM. As discussed, this software was selected to compute the 3D point clouds due to its accuracy, computation efficiency and other proven uses in aerial imagery. This process involves firstly detecting and matching key points to create a sparse cloud for each epoch of data. The two epochs need to then be aligned and scaled to each other through an Iterative Closest Point (ICP) algorithm.

3.3 3D Description

Once point clouds have been computed, each point and its relevant neighbours can be described. From the process of SfM, XYZ points are computed along with their RGB values.

The Point Cloud Library (PCL) was identified in Section 2.1 to contain the most prevalent 3D point descriptors. Some of these descriptors are designed for the application of global model analysis and some for local model analysis. For this research, we are looking for point to point analysis and hence the local descriptors are those of which have been considered. The following descriptors were specifically identified in the literature review due to their application, precision and computational efficiency:

- PFH
- FPFH
- SHOT
- CSHOT/SHOT Colour

Due to the maturity of library and its community PCL was identified as a credible set of implementations of these algorithms. These implementations will be used to describe the points derived by the data gathering process. As the descriptors are designed for the

application of key points (i.e. subset of points of interest), it was found that this can be computationally expensive and take large amounts of time and as the experiments went on, this was taken into consideration.

This section now discusses the techniques use to derive the supporting data and the specifics around descriptor techniques.

3.3.1 Colour

Each point is attributed a RGB value from the SfM process. While the RBG colour space is convenient for display, the literature review performed in this research noted that this is not an optimal space for understanding the difference between colours. The development of descriptors also utilising colour took this into consideration and employed both the CIE XYZ and CIE Lab colour spaces as a discriminator for changes. CIE XYZ is specifically designed to imitate that of humans colour understanding and so presents a good method for presenting colour. CIE Lab goes to linearly separate colour differences mathematically that we perceive visually. Due to their specified design application for differencing analysis the RBG colours will be converted into CIE XYZ space as the central baseline colour space. The following conversion algorithm achieves this:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [M] \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{13}$$

Where M is the conversion matrix with the following parameters:

$$\begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \tag{14}$$

As stated, this offers a baseline colour space whereby other colour spaces can be calculated. To mathematically measure distance between colours, below is the transform of this into CIE LAB from the XYZ space. To measure the distance between any two CIE Lab points, Euclidean distance is used.

$$\begin{aligned} L &= 116\, f_y - 16 \\ a &= 500(f_x - f_y) \\ b &= 200(f_y - f_z) \end{aligned} \tag{15}$$

3.3.2 Normals

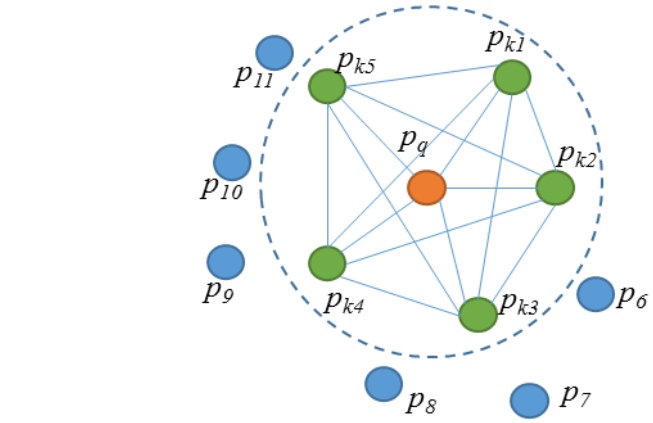
Surface normals form the basis for many of the other point descriptors and are commonly used for computer graphic lighting calculation. From the PCL implementation of normal estimation, the surface normal is approximated by estimating the normal of a plane tangent to the surface. (Rusu, 2010) describes this process in full, but a generalisation is, for each point p_i the covariance matrix C is calculated as follows:

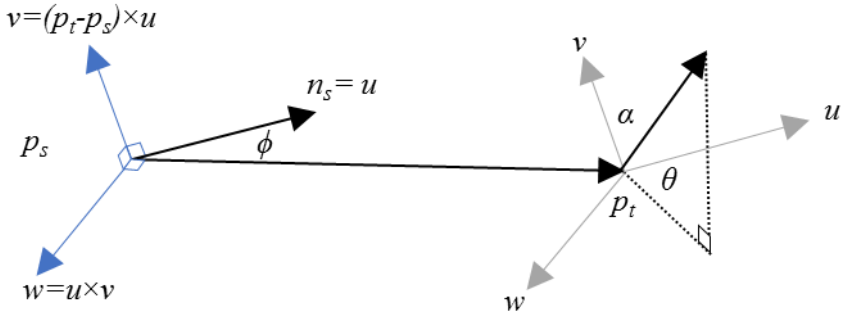
$$C = \frac{1}{k}\sum_{i=1}^{k}.(p_i - \bar{p}).(p_i - \bar{p})^T, C.\vec{v_j} = \lambda_j.\vec{v_j}, j \in \{0,1,2\} \quad (16)$$

Where k is the number of point neighbours to point p_i, and $\bar{p}$ represents the 3D centroid of those neighbours. λ_j is the j-th eigenvalue of the covariance matrix, and $\vec{v_j}$ the j-th eigenvector. The direction of the normal can be resolved to two possible outcomes, and Principal Component Analysis (PCA) is used to solve this.

3.3.3 Point Feature Histogram (PFH)

The PFH descriptor (Rusu, 2008) goes to capture the geometry surrounding a point by analysing the difference between the orientations of their normals (Figure 3-11). First all the points are paired with its neighbours and their neighbours and then for each pair, a fixed Darboux coordinate frame is calculated. From this, the difference between the normal is found and placed with the Euclidean distance between the point pairs. Finally, this is binned into a histogram and then presented by a concatenation of the histogram pairs.





$$u=n_i \qquad v = u \times \frac{(p_j - p_i)}{\|p_j - p_i\|_2} \qquad w=u\times v$$

$$\alpha = \mathrm{v} \cdot n_j \qquad \emptyset = u \cdot \frac{(p_j - p_i)}{d} \qquad \theta = \tan^{-1}(w \cdot n_j, u.n_j)$$

Figure 3-11 - Fixed coordinate frame and angular features (Rusu 2008)

3.3.4 Fast Point Feature Histogram (FPFH)

The FPFH descriptor (Rusu, 2009) builds upon PFH and goes to optimise this calculation due to it being so computationally expensive. It reduces computation by considering only direct connections between key points and their neighbours, removing the second neighbour link. This simplification reduces complexity from O(nk^2) to O(nk) creating a Simplified Point Feature Histogram (SPFH). A second step check the nearest neighbours again and used this to weight the final histogram as follows:

$$FPFH(p_q) = SPFH(p_q) + \frac{1}{k}\sum_{i=1}^{k}\frac{1}{\omega_k}.SPFH(p_k) \qquad (17)$$

Where the weight ω_k represents a distance between the query point p_q and a neighbour point p_k in some given metric space.

3.3.5 Principal Curvatures Estimation (PCE)

PCE estimates the directions (eigenvectors) and magnitudes (eigenvalues) of principal surface curvatures for a given point cloud dataset containing points and its normals. A projection matrix is created from the normal of the query point and nearest neighbours projected onto this plane. The centroid of this plane is computed and then Eigen decomposition is performed to give the resultant directions, k_1 and k_2.

3.3.6 Signature of Histograms Of Orientations (SHOT)

SHOT (Tombari, 2010) builds a spherical ball structure around each query point for local surface description. The search volume is be divided (into 32 to be precise) separate segments and the cosine of the normal of the key point and the points within each volume calculated (Figure 3-12). These angles are then binned into a histogram to create the final descriptor. According to (Tombari, 2010), the result gives a rotation invariant solution due to its dependence on the local reference frame.

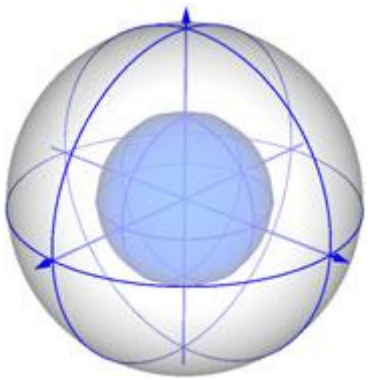


Figure 3-12 - SHOT support structure (Tombari, 2010)

3.3.7 SHOT Colour

SHOT Colour (Tombari, 2010) builds upon the SHOT algorithm, adding an additional context of CIELab colour. This extra dimensionality is integrated into the histogram of orientations.

3.4 Change Detection

Once models have been created using SfM, and all the points described using the different algorithms discussed, change detection may take place. This is generalised in Figure 3-13 below, whereby it can be seen that cloud 1 is represented blue, cloud 2 in orange and the difference between these in green. Where required, this section discusses the design and implementation of each of the processes.

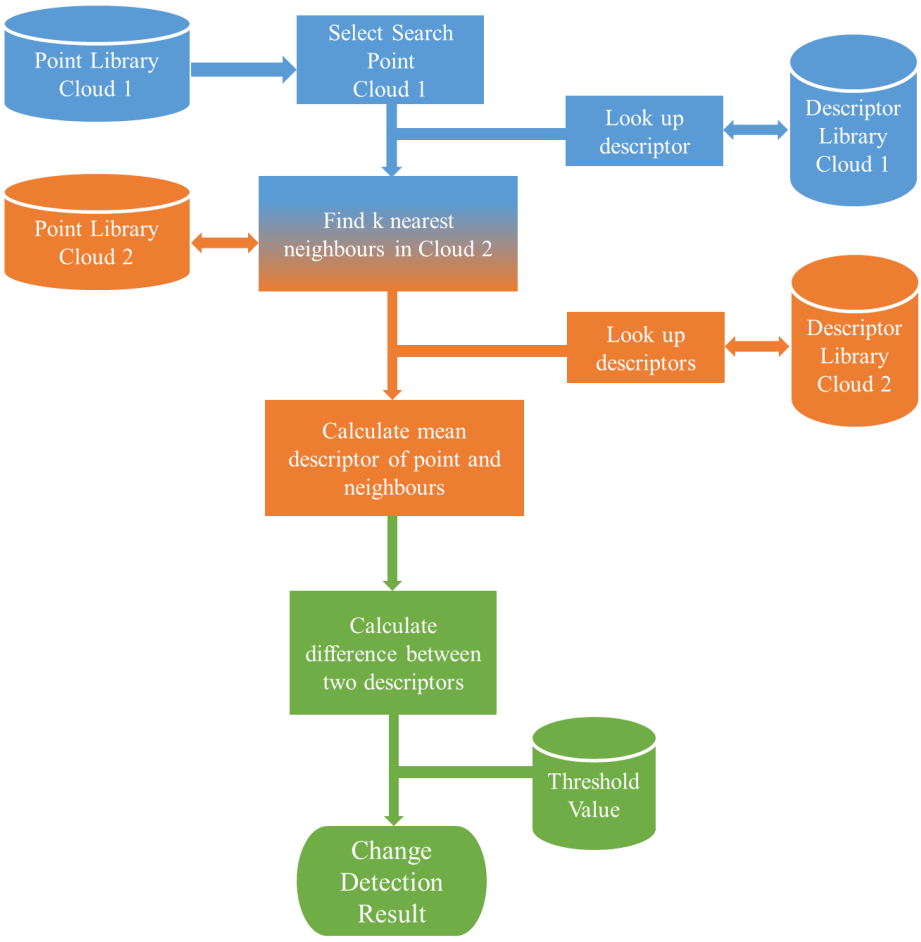


Figure 3-13 - Change Detection Data and Processing Chain

To detect temporal change between two epochs of data, each and every point in the baseline model (cloud 1) is processed to determine its difference from the changed model (cloud 2).

A threshold is then applied to distinguish what is considered change and what is not considered change. This can then be presented visually or statistically.

The process detailed in is broken down as follows:

1) Define Search Point.
2) Find *k* nearest neighbours.
3) Determine difference between descriptors of search point and the nearest neighbours.
4) Create single distance value for difference.
5) Threshold difference to determine if it is change.

This process is iterative, and once a search point is complete, the next is calculated until the entire cloud has been processed.

3.4.1 Finding Nearest Neighbours

Firstly the nearest neighbour points to the search being considered must be identified. As the magnitude of points being compared is in the tens of millions, a computationally optimised approach is required. As expected, this problem has had many solutions proposed, some being more efficient at different dimensionality. In this research, all the points will be in the three-dimensional Euclidean space, which narrows down the applicable solutions. One of particular interest is that of the binary space partitioning solution. This is a method for recursively dividing a space with hyperplanes to create smaller spaces. This allows the organisation of the data points into a tree structure whereby the nearest neighbour can be searched efficiently. One specific binary space partitioning solution, the *k*-d tree, is whereby the hyperplanes are constrained to be perpendicular to the axis of the hypercube and form a tree structure for searching (Figure 3-14). The space is initially split in two by a hyperplane shown a red. The tree defines those points to the left of the plane as the left of the tree, and those to the right as the right of the tree. A second hyperplane is introduced perpendicular to the first, shown as green, that further splits the tree into a series of left and right points, and then a further plane is added, coloured blue. This creates the tree structure whereby the points can be searched efficiently.

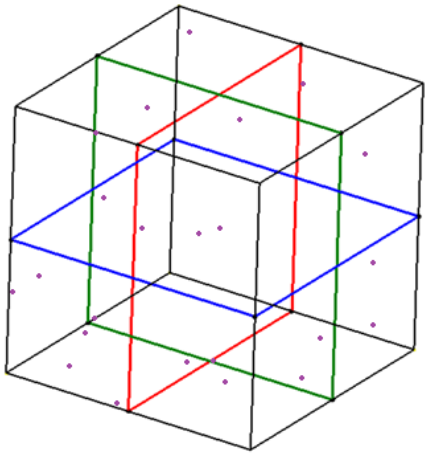


Figure 3-14 - k-d tree hypercube

For each search point, the *k* nearest neighbours (*k*-NN) between each point in the two point clouds are identified (Figure 3-15). The *k* number of nearest neighbours will be varied to enable simple outlier smoothing, and for the purpose of performance assessment. This results in each point having a set of *k* nearest neighbours.

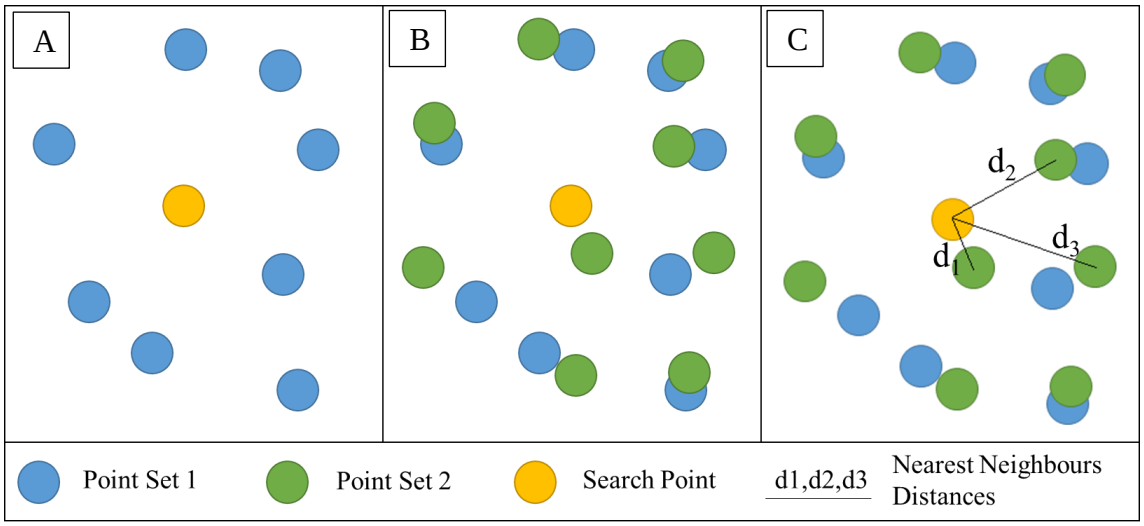


Figure 3-15 - Finding nearest neighbours and distances between two point clouds

A) Point cloud 1 with search point highlighted B) point cloud 1 and 2 overlaid C) nearest neighbours in point cloud 2 from search point in Point cloud 1.

3.4.2 Descriptor Comparison and distance measurement

When the nearest neighbours have been identified for any search point, they are indexed and their descriptors retrieved. The nearest neighbour descriptors are to be compared to the search point descriptor and the relative distances computed. A single distance value is given between each two sets of points, and as discussed earlier, if there is more than one nearest neighbour, the mean of these distances are taken. This is presented as a final distance metric for each of the points. If the first distance is equal to 0, the other distances are negated and an overall distance of 0 is returned due to having a perfect match between the two models. This prevents points that have not moved/changed reporting a changed distance. If there has been a change detected, this is saved as a new point cloud which presents the difference between cloud 1 and cloud 2.

3.4.3 3D Structure And Colour

This research looks at a new method using both the structure and colour data output from the SfM process, coined the 3D Structure and Colour difference (3DSAC).

To achieve this, the Euclidean Squared XYZ distance between two points and the Euclidean Squared distance in CIELab Colour between the two points are calculated. This is a simple sequence of operations. A configurable weighting is applied that allows for a high or lower ratio of XYZ distance to Colour distance to be applied. This permits tuning of the algorithm dependent on the application. This could even take place in real time as does a radar operator adjusting its sensitivity. If a large light variation between the 3D point clouds is expected (e.g. time of data light intensity variation), the colour weighting can be reduced and vice versa. Finally the maximum of these two values is taken as the distance between the two points. This can be described with the equation below:

$$3DSAC = Max(W_1 * \Delta XYZ, W_2 * \Delta CIELab) \quad (18)$$

Where W_1 represents the weighting applied to the XYZ distance and W_2 represents the weighting applied to the CIELab distance.

3.4.4 Detection Thresholding

To decide what can be considered change and what is considered the same, a threshold is set to filter those believed to be change and that which is not.

For the performance assessment, this is governed by the maximum distance measured and forms the maximum detection threshold value. At this point, the 0% threshold, all point are considered to be detections. At the other end of the spectrum is the 100% threshold whereby no points are considered change. This is then split into a thresholding routine of 1000 steps to give granularity over the change detection performance measurement (e.g. 0.0%, 0.1%, 0.2% etc.). For each of the 1000 threshold points, a point cloud is produced so that it can be analysed visually and statistically.

3.5 Performance and Correctness Assessment

To develop, analyse and make conclusions on the effectiveness of any design, an element of performance assessment must take place and may be assessed in two ways; quantitatively and/or qualitatively. This research will take advantage of both methods, with an emphasis on quantitative performance assessment due to the subjective nature of qualitative assessment.

3.5.1 A unique technique for deriving estimated ground truth

To measure the performance of each descriptor, a reliable and effective method must be used to find the true change in the data. This itself is not a simple task, and requires mapping the real-world change to those points in the models that represent that change. To achieve this, the exacting knowledge of the change introduced between the two epochs of data is needed and an accurate method for then deriving those points in the model. Careful attention was paid during the planning and gathering of the data to only introduce measurable change that could be derived.

To estimate the true change, multiple estimates are used in aggregation to reduce the error in this process. The method to derive this takes the 3D point cloud of which has had change introduced, and removes those points that are not change to create a new model. This process is independently repeated to create a series of model estimates of the true change.

During these estimations, it is equally possible to have errors both estimating too many points (larger than the real object) and too few points (smaller than the real object). Figure 3-16 goes so illustrate this whereby each of the following a represented:

- An overestimation i.e. a positive error (outside blue dotted line)
- An underestimate i.e. a negative error (inside blue dotted line)
- The estimated mean truth i.e. zero error (middle green line)

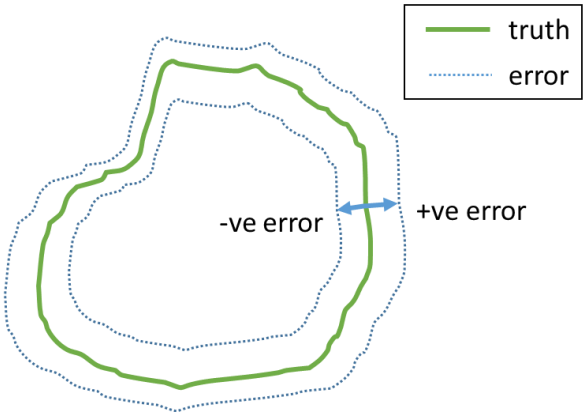


Figure 3-16 - Truth model errors

By multiple estimate instances, there will be a variance of over and under estimations to the true truth. Adding all these up in a single mass of all the points creates a total set of all estimations. This will be vastly over estimated model and represents the all the overestimations compounded into one model. It is also noted that many of the points will be duplicated.

A scoring mechanism is then used to count the number of occurrences of each point in the set. Following this, a threshold is applied to filter only those that have appeared in more than an incrementing number of the estimates.

This will create a series of point clouds. There will be the same number of these point clouds as there were original instances of estimates. The first with a threshold of attaining just 1 instance will contain all the points (i.e. the maximal or +ve error model). As the threshold is increased, less of the points will be present in the models. All the way through the threshold with the maximum number of instances. This will create a set of the point that are present in all the models, and hence common to all. This is the minimal or –ve error model.

By calculating the error of each of the point clouds to both the geographical midpoint and the statistical midpoint of the –ve model and the +ve model, the best estimate of the true truth model can be found. From this exercise it is evident that this quantitative analysis will offer an error margin. As there are more instances, it should be possible to pick the closest model to the midpoint. From this, a final model of true change is made. By independently estimating the true change multiple times, the error in this processes is greatly reduced.

3.5.2 Quantitative Assessment

This section provides some definition to how a point is classified, the rates of these classification and methods for measuring or scoring the accuracy and performance of the algorithms. (Power, 2011) provides an in-depth evaluation on performance measurement techniques covering precision, recall, F-measure to ROC, informedness, markedness and correlation. This section goes to define and discuss those techniques used in this research.

3.5.2.1 True Positive (TP)

The ability to measure when a change has been detected truthfully is paramount to measuring the effectiveness of a binary classifier and forms the basis of measuring the performance of each of the algorithms evaluated.

3.5.2.2 True Negative (TN)

When a point has been correctly identified and not being change, it is deemed a TN.

3.5.2.3 False Positive (FP)

When a point has been attributed as being a change falsely, i.e. not a real change, it is deemed a FP.

3.5.2.4 False Negative (FN)

When a point has been incorrectly identified as not changing, i.e. not detecting change, it is deemed a FN.

3.5.2.5 Confusion Matrix

A confusion matrix simply brings all these four together into one chart as below

TP	FP
FN	TN

Figure 3-17 - Confusion Matrix

3.5.2.6 True Positive Rate (TPR)

TPR, also known as Recall, Sensitivity or Probability of Detection (P_d) in radar theory, gives a measurement of the proportion of positives that have been correctly identified. It can be defined as follows:

$$TPR = \frac{TP}{TP + FN} \tag{19}$$

3.5.2.7 True Negative Rate (TNR)

TNR, also known as specificity, measures the proportion of measuring TNs and can be defined as follows:

$$TNR = \frac{TN}{FN + FP} \tag{20}$$

3.5.2.8 False Positive Rate (FPR)

FPR, also known as Probability of False Alarm (P_{fa}) or a type I error, measures the proportion of incorrect positive measurements. That is, to falsely present a finding when it is truly not there and can be defined as follows:

$$FPR = \frac{FP}{FP + TN} \tag{21}$$

3.5.2.9 False Negative Rate (FNR)

FNR, also known as sensitivity or a type II error, measures the proportion of measuring FNs, or more simply as missing a TP, and can be defined as follows:

$$FNR = \frac{FN}{FN + TP} \tag{22}$$

3.5.2.10 Receiver Operating Characteristic (ROC)

ROC analysis was developed during World War II for classification theory and performance assessment of radar and is now widely used machine learning. A ROC curve gives a visual representation of the relationship between TPR and FPR at various thresholds, offering a simple overview of system performance. As this is a visual performance indicator, is can become subjective and (Fawcett, 2005) goes to analyse and address this by providing a guide to a standardised ROC analysis. As seen in the figure below, as the trend line tends towards the top right i.e. a higher TPR than FPR, the system can be regarded as being better performing and conversely vice versa.

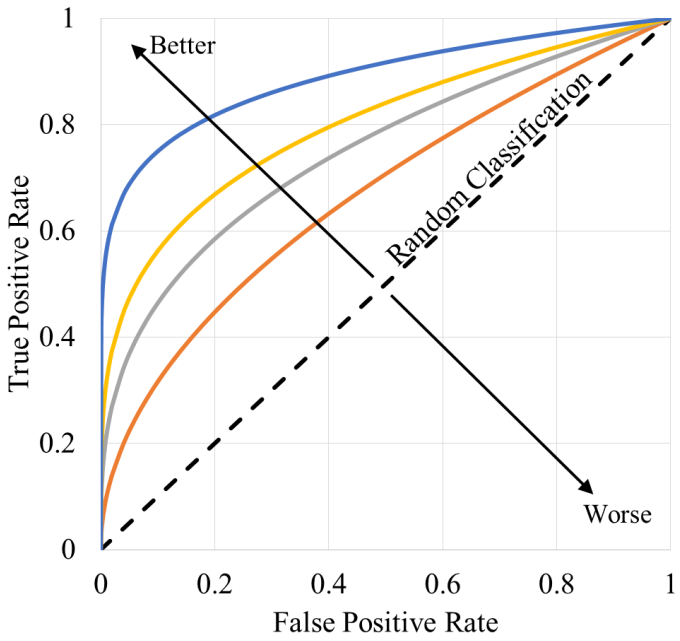


Figure 3-18 - ROC Curve Performance

As discussed, the ROC curve offers a great visual method for assessing performance that is universally used, although can be interpreted subjectively. One way to turn this into a metric is to measure the area under this graph, known as the Area under ROC curve (AuROC). Proposed by (Hanley 1982), this offers a single value that can be used as a discriminator in the performance of systems. A trapezoidal numerical integration is used to calculate the AuROC. This approximates the integration over an interval by the common method of using trapezoids. As we have many discrete values (1000), this proves an accurate method for estimating the AuROC.

3.5.2.11 Matthews Correlation Coefficient (MCC)

The MCC (Matthews, 1975) offers a single measure of classification accuracy taking the entire confusion matrix into account (TP, TN, FP and FN) and hence offers a balanced measure.

There is a substantial calling to use the MCC over other metrics in machine learning due to its ability to handle disproportionate and unbalanced confusion matrices (Boughorbel, Jarray and El-Anbari, 2017).

(Chicco, 2017) goes further and states the following:

"We strongly encourage to evaluate each test performance through the Matthews correlation coefficient (MCC), instead of the accuracy and the F1 score, for any binary classification problem."

It gives a single value between -1 and +1 whereby -1 represents a system with perfect disagreement between truth and prediction, and 1 gives a perfect agreement, with 0 being randomised and calculated as follows:

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}} \qquad (23)$$

3.5.2.12 TPR minus FPR

Finally a very simple metric is proposed, a disparity between the True Positive Rate and the False Positive Rate. This can range from -1 to 1, whereby a -1 represents a system where there is a certain reporting of False Positive, and a 1 whereby a system reports a certain True Positive. This takes all the components of the confusion matrix into consideration and is very easy to interpret. This performance assessment value is given as some variance is expected in the results in the two previous methods. This offers an additional and instantaneous understanding of the difference in the TPR and FPR, also giving assess scores by a quorum. Additionally its computation is very simple and reversed to derive its components if required.

3.5.2.13 Probability of False Alarm at 90% Probability of Detection

This is a single figure statistic that is frequently used to describe the performance of a radar system. It simply offers the probability of a false alarm (False Positive) when there is a 90% probability of an object (or point) being correctly detected. This is commonly used to compare different systems without having to perform ROC analysis.

3.5.3 Qualitative Assessment

The true usefulness of the system is how well the end user can interpret what it outputs. Statistical analysis alone will not be sufficient to understand how well change has been detected and will take an element of visual inspection. The shape and, granularity and distribution of the change across an algorithms surveillance volume gives the end user confidence that the change is true and not a false alarm. With the knowledge of true mathematical change, a visualisation of the data must also form part of the solution. As discussed in the previous section, another element is ROC analysis whereby the shape of the curve will give different characteristics through the use of a detection algorithm.

3.5.3.1 Visualisation of data

Once change between the two models in a dataset has been calculated, displaying it in a useful form gives the processed data context. As the data has been normalised into a scaled distance of 0 to 1, this will make the salient change once visualised. As visual interpretation is a subjective case, it depends on the visual ability and understanding of the viewer.

To give the best understanding, differentiation between high change and low change should be attainable. This means taking the single distance value for each point and assigning it with a colour that enables the viewer to understand the change. Many engineers and scientists use a baseline grayscale for this, as it is and understood baseline so this will form one of the visualisation techniques. Grayscale alone may not offer the best discrimination between changes and suit all end users. For this reason, the visualisation should be customisable and able to be mapped to any colour map. For presentation in this research, two methods are presented, one being a linear grayscale colour map for generic understanding. The other being a 90% of the full HSV colour spectrum to give a maximal spreading of change levels and hence the largest discriminator without risk of colour wheel wrap around.

This can be achieved by converting the distance scale to an angle of 0 to 324 and taking this as the hue variable in the HSV colour space. If the saturation and value variables are set to 1, the 0 to 324 hue values can be seen as the circumference of the top surface in a HSV colour cone. A distance scale of 0 would be represented by red, a distance of 1/3 by

green and 2/3 by blue, and all the colours in between. Once the HSV value is converted to an RGB colour space, these values can be attributed to each point in the baseline model. The resulting models give a colour map of change. The two distance scales and representative colour maps are shown below.

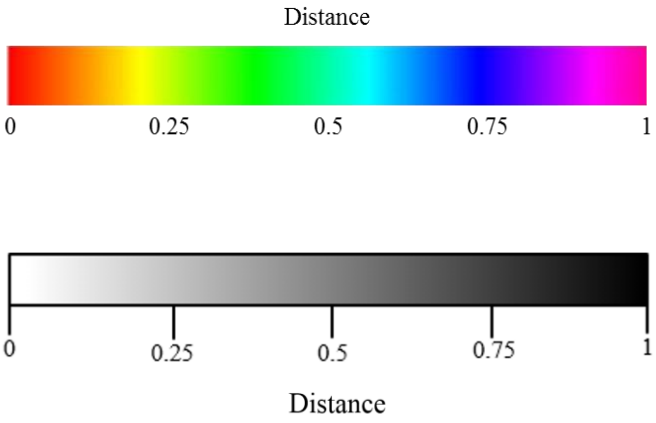


Figure 3-19 – Grayscale and colour bars for distances

3.6 Dataset 1 Example Walkthrough

To give clarity over the full change detection and performance assessment process, an example walkthrough is now given. Due to the vast amount of data and graphing required, it would not be possible to display this for each of the descriptors and dataset in the full results chapter. The output of this consistent and repeatable process are performance metrics defined as earlier in this chapter. The walkthrough details the change detection and then performance assessment using Dataset1 and Euclidean change detection. It starts when two point cloud models have been gathered and a truth model has been determined using the process in Section 3.5.1. Note, this does not consider the analysis of this data, purely the techniques used to produce the results.

Dataset 1 Euclidean Change Detection Walkthrough

Firstly, each search point is iterated over and the nearest neighbour identified. The distance between this and the search point is calculated using the Euclidean distance between the two. The maximum value of distance is then used to normalise all the distances for a value of 0 to 1 for ease of comparison. Figure 3-20 shows the distances relative to their index within the point cloud and also the values sorted by their distance. The position of the change within the point cloud can be identified in the graph as the spikes, and how this compares to the rest of the background. This is then performed for each of the *k* values required for the *k* nearest neighbours. A red example threshold of 0.2 or 20% has been placed on the graph whereby all points above would be considered as detected change.

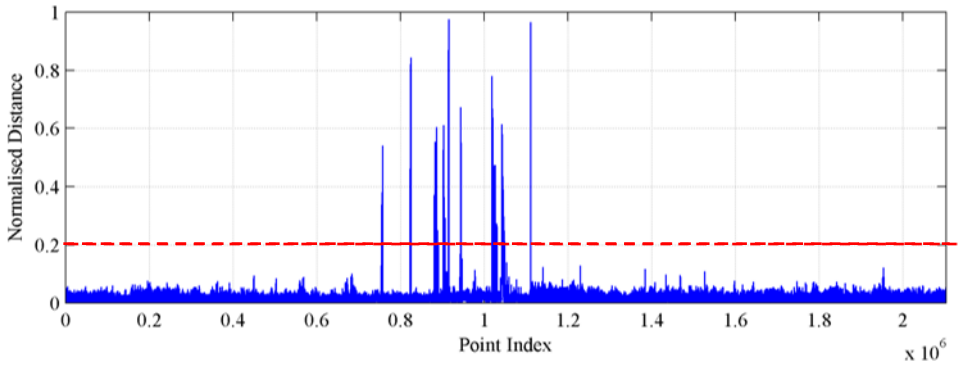


Figure 3-20 – Graph of unsorted distances (threshold = 0.2)

A histogram of distances measured shows that the values that have change are low in number (i.e. a small area of change) and a large in comparison to the background change (Figure 3-21). The large peak here shows these smaller

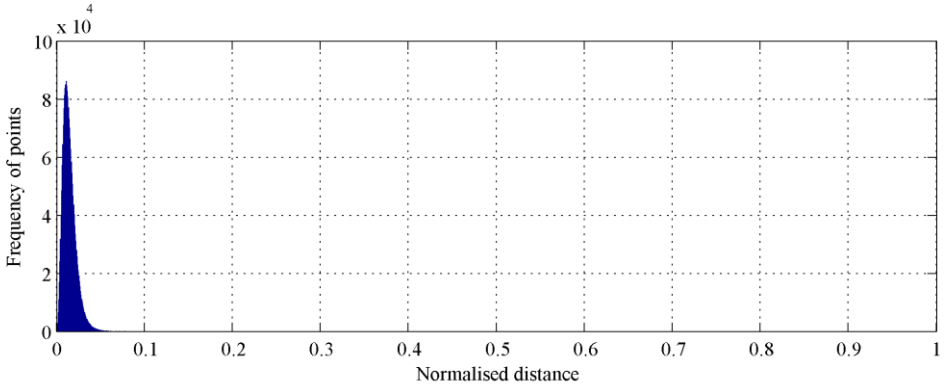


Figure 3-21 - Histogram of Normalised Distances

Next, each of the 1000 threshold values is incremented over, from 0 to 1 and the results output to new subset as a point cloud for analysis. As expected, the number of detections decrease as the threshold is increased as seen below (Figure 3-22).

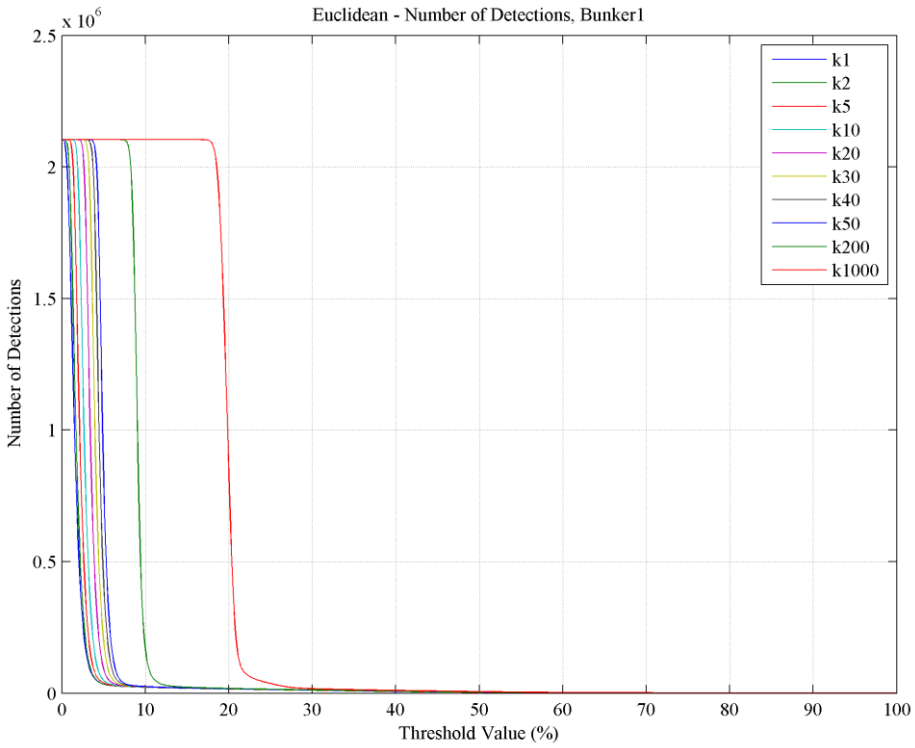


Figure 3-22 – Graph of Total Number of Detections

Dataset 1 Euclidean Performance Assessment Walkthrough

From the previous section, we have all the data required to compare against the Truth Model for performance assessment. Using the Truth Model, the points from each of the 1000 threshold models can be compared. Each of the points in the truth model are iterated over and their nearest neighbours found in the threshold models. If distance = 0, a match is found and can be determined as a TP. If distance > 0, the change detected is a FP. The TNs and FNs can also then be found and a full confusion matrix is available for that threshold. The next threshold is then calculated until all 1000 are complete.

As the number of point of real change is known from the truth model, the probability of detection or TPR is calculated (Figure 3-23).

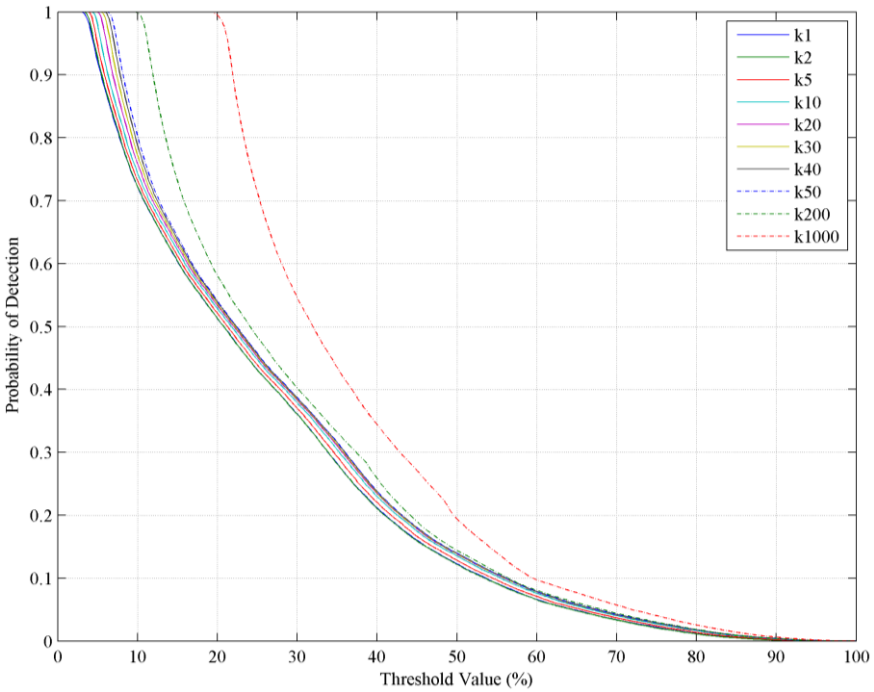


Figure 3-23 - Graph of Probability of Detection / TPR

The converse of this is also required, whereby the FPs are grouped and the Probability of False Alarm / FPR is calculated as seen below (Figure 3-24). It can be seen with a lower threshold there is a much higher probability of false alarm.

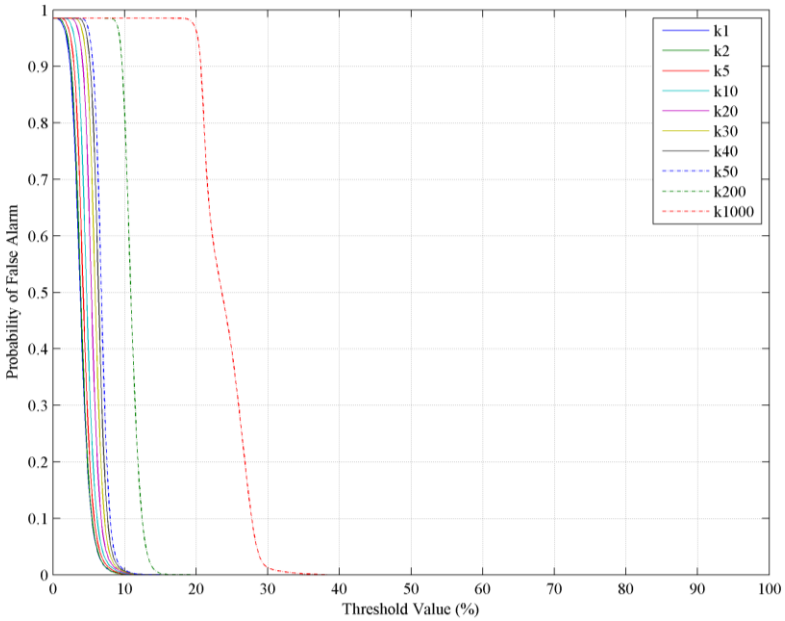


Figure 3-24 - Graph of Probability of False Alarm (PfA) / FPR

Now that both the TPR and FPR are known for all the threshold values, they can be analysed against each other by creating a Receiver Operating Characteristic curve representation of TPR against FPR (Figure 3-25).

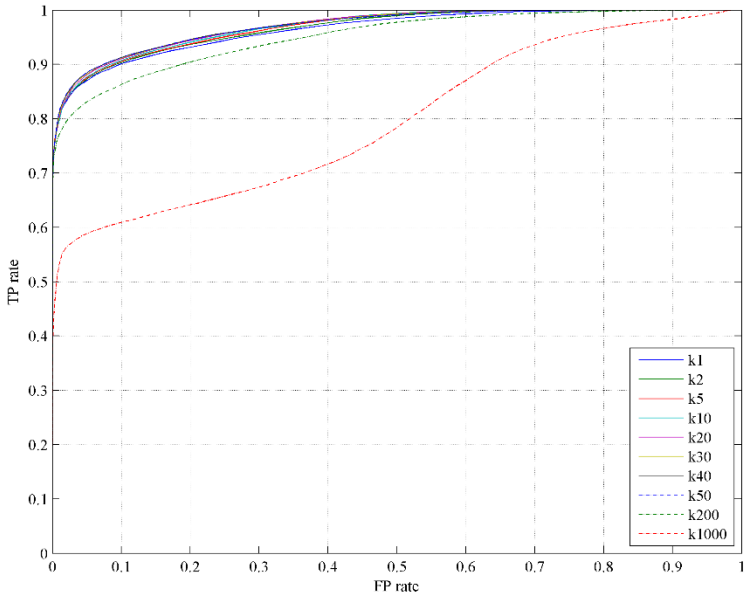


Figure 3-25 - Graph of Receiver Operating Characteristic (ROC) Curve

With the ROC curve graph and an understanding of Figure 3-18, we can immediately contrast the performance of the system for the different *k* values. The area under this curve (Section 3.5.2.10) is computed for each value of *k* and again is graphed for analysis (Figure 3-26). The maximum value of this set of values displayed by the red dotted line is the maximum performance found when using this technique for performance assessment.

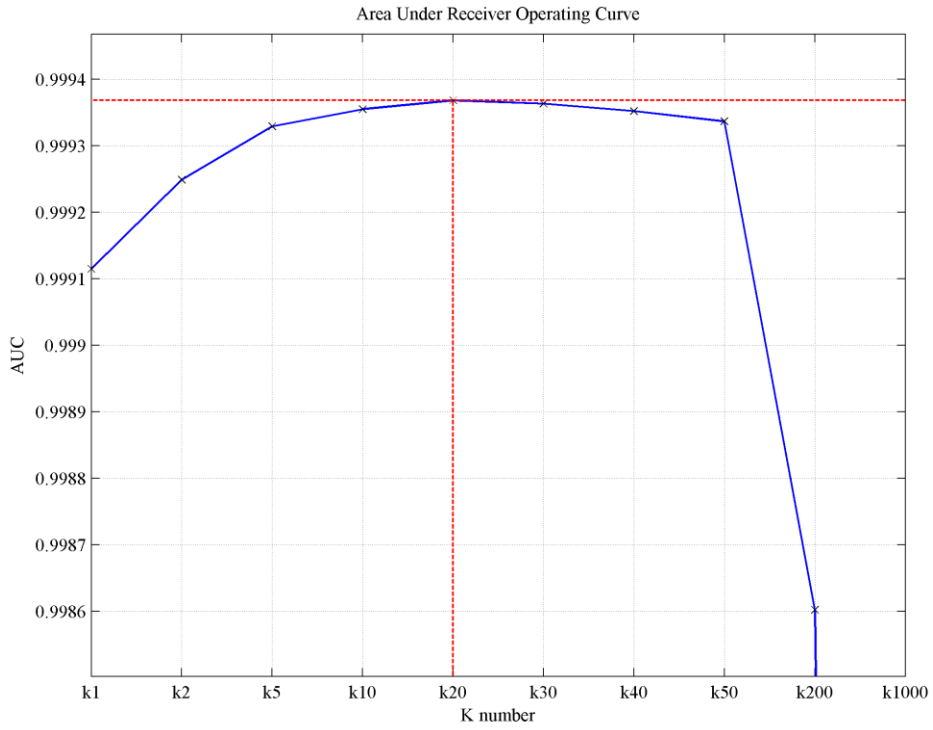


Figure 3-26 - Area under ROC curve, limited axis

As a confusion matrix is available for each threshold value, both the MCC and TPR-FPR can be calculated and their respective maximum values located (Figure 3-27 and Figure 3-28).

To calculate the value of PfA @ 90% Pd (Section 3.5.2.13), the nearest values to the 90% P_{fa} value are located using a nearest neighbour search and then interpolated to estimate the value.

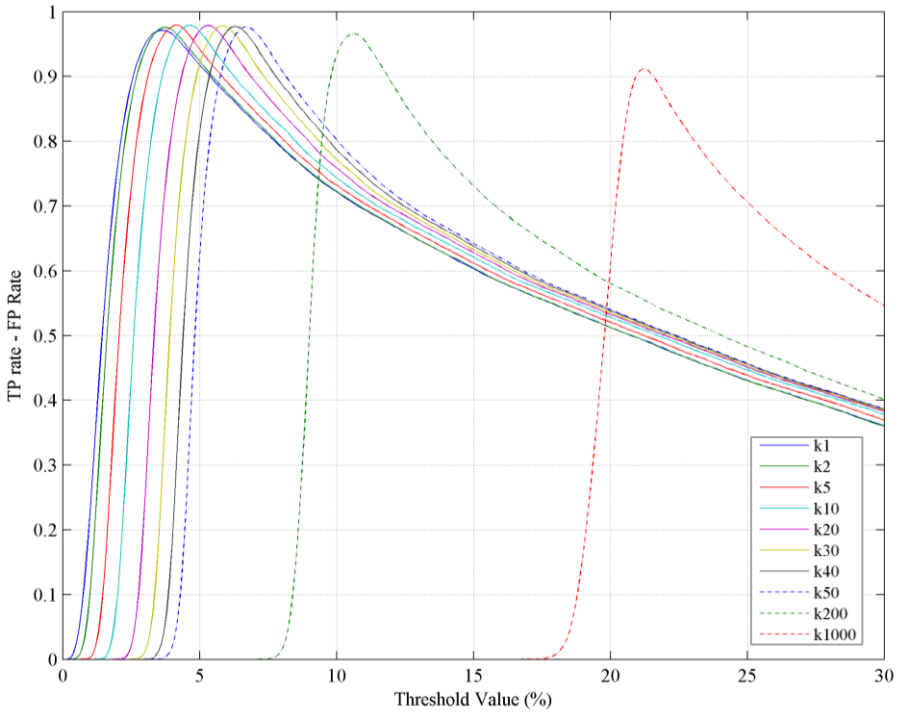


Figure 3-27 - Graph of TPR - FPR

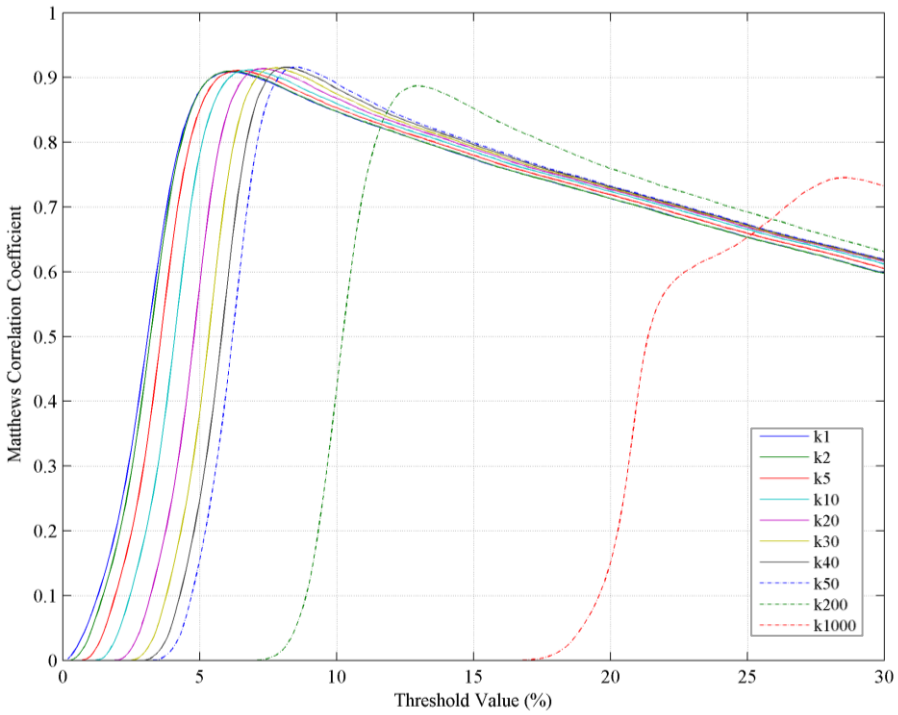


Figure 3-28 - Graph of MCC

Chapter 4

Results

This Chapter displays a summary of results from the methodology described in Chapter 3.

First the truth model results and calculations are presented for each of the three datasets.

Following this, each of the three datasets and corresponding descriptors are sequentially detailed. With the aid of this, each algorithms performance measures are presented statistically with additional examples of visualisations to assist where relevant.

Within each dataset and to simplify understanding, a summary of results limited to only the maximal descriptor performance are displayed in tabular and graphical form (Section 4.6).

The in-depth analysis and discussion of these results follow in Chapter 5.

4.1 Creation of Estimated Ground Truth Models

As discussed in Section 3.5.1, the truth model pays a key part in the credence of any performance assessment. Here the calculations of these truth models are presented with their subsequent errors.

For comparative purposes and to understand error, each truth estimate also has the following statistics displayed:

- Distance to mean.
- % deviation from mean.

This will allow for the highlight of any outliers and to understand the general error when analysed. Here the signage is kept as this shows if this is an over estimation or an underestimation of points as described in Section 3.5.1.

Also the coefficient of variation is presented to give a generalised understanding of relationship between the magnitude of the standard deviation and the mean of the dataset.

4.1.1 Ground Truth - Dataset 1

To calculate the Dataset 1 estimated ground truth model, eight unique estimates of truth were taken in isolation. As detailed in the methodology, each of these consist of a manual analysis of the imagery in combination with analysis of the point clouds generated.

Table 4-1 - Dataset 1 ordered ground truth estimates

numb of points	distance to mean	% deviation from mean
27226	-3647	-11.82
27415	-3458	-11.20
29742	-1131	-3.67
30488	-385	-1.25
31697	823	2.67
32797	1923	6.23
32912	2038	6.60
34713	3839	12.44

Table 4-1 shows the individual iterations of estimated truths. Taking the mean of these presented in Table 4-2, their corresponding distances to the mean are also detailed and the magnitude of this is also presented as a percentage. Visualisation of each of these models is presented in Figure 4-1.

Table 4-2 - Dataset 1 ground truth estimate descriptive statistics

Total sum of all points	246990
Unique points	34713
Mean	30874
standard deviation	2499
Coefficient of variation	0.081

Adding all the points together into a single cloud (with duplicates) gives a total of 246,990 points with 34,713 of these being unique points (Table 4-2).

Using the full sum of all the points with duplicates, the number of instances of each unique point is calculated. A threshold is then applied to filter and create models to those that contain eight instances (is present in every estimate), through to a model that has only one occurrence of each point (the entire set of unique points). The statistics for this are displayed in Table 4-3.

Table 4-3 – Dataset 1 ground truth models with threshold occurrences

threshold occurrences	unique points	probability of unique points	distance to mean	% deviation from mean	distance from half range
1	34713	1.00	3930	12.77	3572.5
2	32912	0.95	2129	6.92	1771.5
3	32798	0.94	2015	6.55	1657.5
4	31238	0.90	455	1.48	97.5
5	30257	0.87	-525	-1.71	-883.5
6	28564	0.82	-2218	-7.21	-2576.5
7	28211	0.81	-2571	-8.35	-2929.5
8	27568	0.79	-3214	-10.44	-3572.5

Table 4-3 shows each of the models created from the entire set when the threshold is applied to individual point occurrences. It can be seen that as a threshold of 1 has the same number of unique points as presented in Table 4-2. As the threshold is increased, the number of points decreases. By taking the mean (Table 4-4), their respective distance to mean and percentage deviation to mean is also presented. Also by calculating the range (max – min) and half of this (Table 4-4), the midpoint geographically between the two extremes can be identified. The respective distance to this for each of the models is also presented. It can be seen that the closest model to the central point (defined as the truth) is when at least 4 occurrences are found of each of the points in the sum model. A visualisation of this can be seen in Figure 4-2. Here an over estimation of points is found at 1.48% to the mean. This give us the error for this truth model to the actual truth.

Table 4-4 - Dataset 1 ground truth models descriptive statistics

mean	30783
standard deviation	6812
coefficient of variance	0.221
range	7145
half range	3573

From Figure 3-16, it can be seen that the model with the least error is closest to the mean. In this instance, a threshold of at 4 occurrences gives the closest to that boundary, and hence presents the estimated ground truth model for Dataset 1. Here there are 31,238 unique points, with a predicted error of 1.48% from the actual truth.

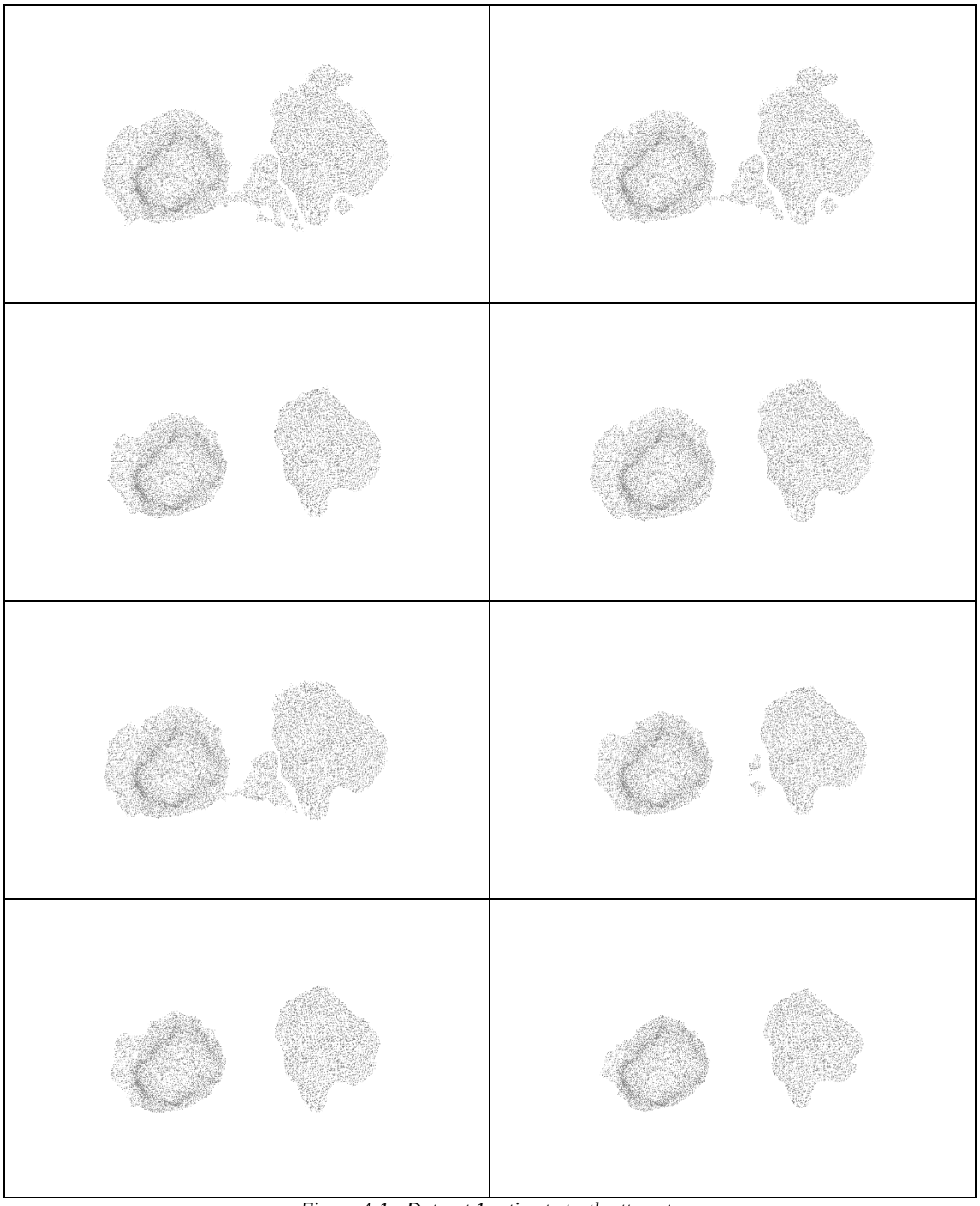


Figure 4-1 - Dataset 1 estimate truth attempts

Figure 4-1 shows each of the eight estimates of truth. The set is ordered by number of points to illustrate the variance in the estimates.

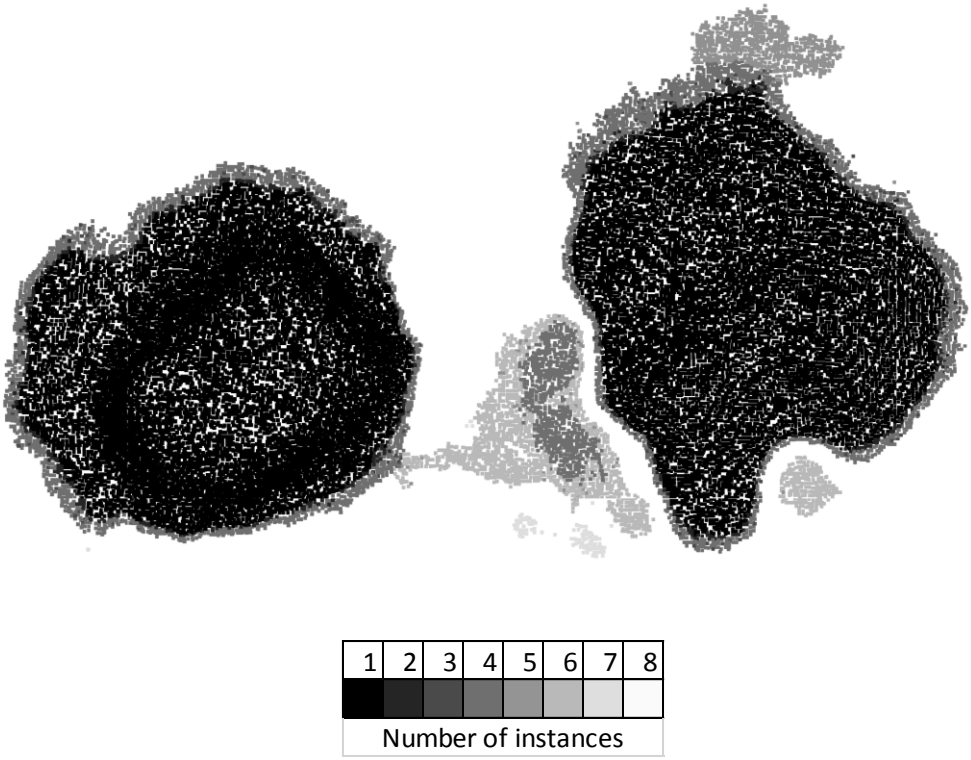


Figure 4-2 - Dataset 1 estimate truth models instances

Figure 4-2 above visualises each of the models created by counting the number of instances of each point and thresholding these into separate models. As the lightness of point colour increases (the white value), the number of instances also increases. This is to show that at the lowest number of instances (minimal model) is smaller than that of all of the points (maximal model).

4.1.2 Ground Truth - Dataset 2

To calculate the Dataset 2 estimated ground truth model, seven unique estimates of truth were taken in isolation. Again as detailed in the methodology, each of these consist of a manual analysis of the imagery in combination with analysis of the point clouds generated.

Table 4-5 - Dataset 2 ordered ground truth estimates

Num of points	distance to mean	% deviation from mean
2293	-695.14	-0.23
2300	-688.14	-0.23
2471	-517.14	-0.17
2906	-82.14	-0.03
2938	-50.14	-0.02
3257	268.86	0.09
4752	1763.86	0.59

Table 4-5 shows the individual iterations of estimated truths. Taking the mean of these presented in Table 4-6, their corresponding distances to the mean are also detailed and the magnitude of this is also presented as a percentage. Visualisation of each of these models is presented in Figure 4-3 and Figure 4-4 (object 1 and 2 respectively).

Table 4-6 - Dataset 2 ground truth estimate descriptive statistics

total sum of all points	20917
unique points	5335
mean	2988
standard deviation	794
Coefficient of variation	0.266

Adding all the points together into a single cloud (with duplicates) gives a total of 20,917 points with 5,335 of these being unique points.

Using the full sum of all the points with duplicates, the number of instances of each unique point is calculated. A threshold is then applied to filter and create models to those that contain at least seven instances (is present in every estimate), through to a model that has at least one occurrence of each point (the entire set of unique points). The statistics for this are displayed in Table 4-7.

Table 4-7 - Dataset 2 ground truth models with threshold occurrences

threshold occurrences	unique points	probability of unique points	distance to mean	% deviation from mean	distance to half range
1	5335	1.00	5168	32.81	3318
2	4285	0.80	4118	26.15	2268
3	3338	0.63	2224	14.12	1321
4	2551	0.48	-137	-0.87	534
5	2292	0.43	-1173	-7.45	275
6	1815	0.34	-3558	-22.59	-202
7	1301	0.24	-6642	-42.17	-716

Table 4-7 shows each of the models created from the entire set when the threshold is applied to individual point occurrences. It can be seen that as a threshold of 1 has the same number of unique points as presented in Table 4-6. As the threshold is increased, the number of points decreases. By taking the mean (Table 4-8), their respective distance to mean and percentage deviation to mean is also presented. Also by calculating the range (max – min) and half of this (Table 4-8), the midpoint geographically between the two extremes can be identified. The respective distance to this for each of the models is also presented. It can be seen that the closest model to the central point (defined as the truth) is when at least 4 occurrences are found of each of the points in the sum model. A visualisation of this can be seen in Figure 4-5 and Figure 4-6 (objects 1 and 2 respectively). Here an under estimation of points is found at 0.87% to the mean. This give us the error for this truth model to the actual truth.

Table 4-8 - Dataset 2 ground truth models descriptive statistics

mean	2988
standard deviation	1321
coefficient of variance	0.442
range	4034
half range	2017

From Figure 3-16, it can be seen that the model with the least error is closest to the mean. In this instance, a threshold of at 4 occurrences gives the closest to that boundary, and hence presents the estimated ground truth model for Dataset 2. Here there are 2,551 unique points, with a predicted error of 0.87% from the actual truth.

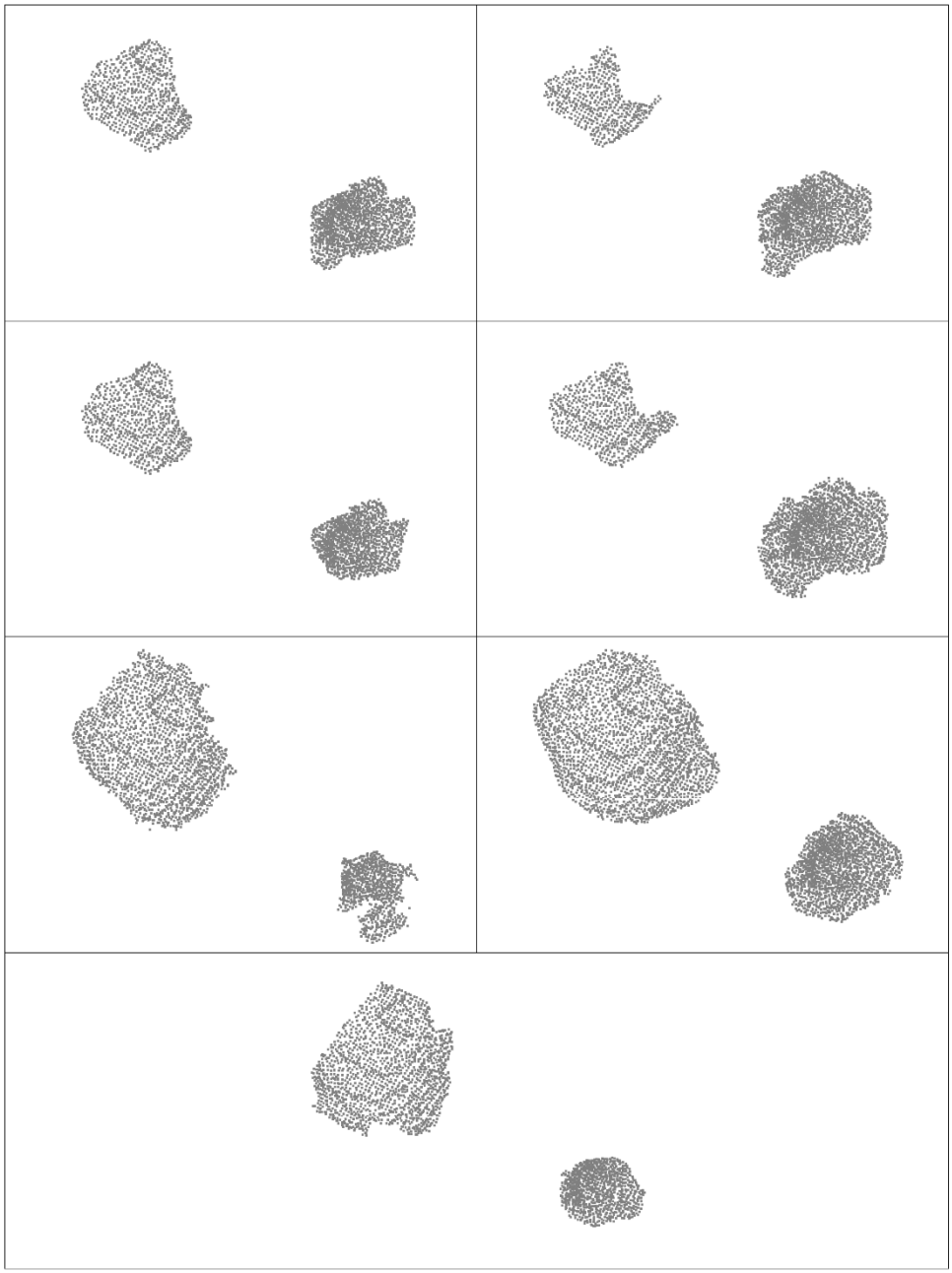


Figure 4-3 - Dataset 2 estimate truth attempts – Object 1

Figure 4-3 shows each of the seven estimates of truth for object 1 (the rock). This has been separated from object 2 due to the geographic separation and hence difficulty to visualise as one image at a suitable scale. The set is ordered by number of points to illustrate the variance in the estimates.

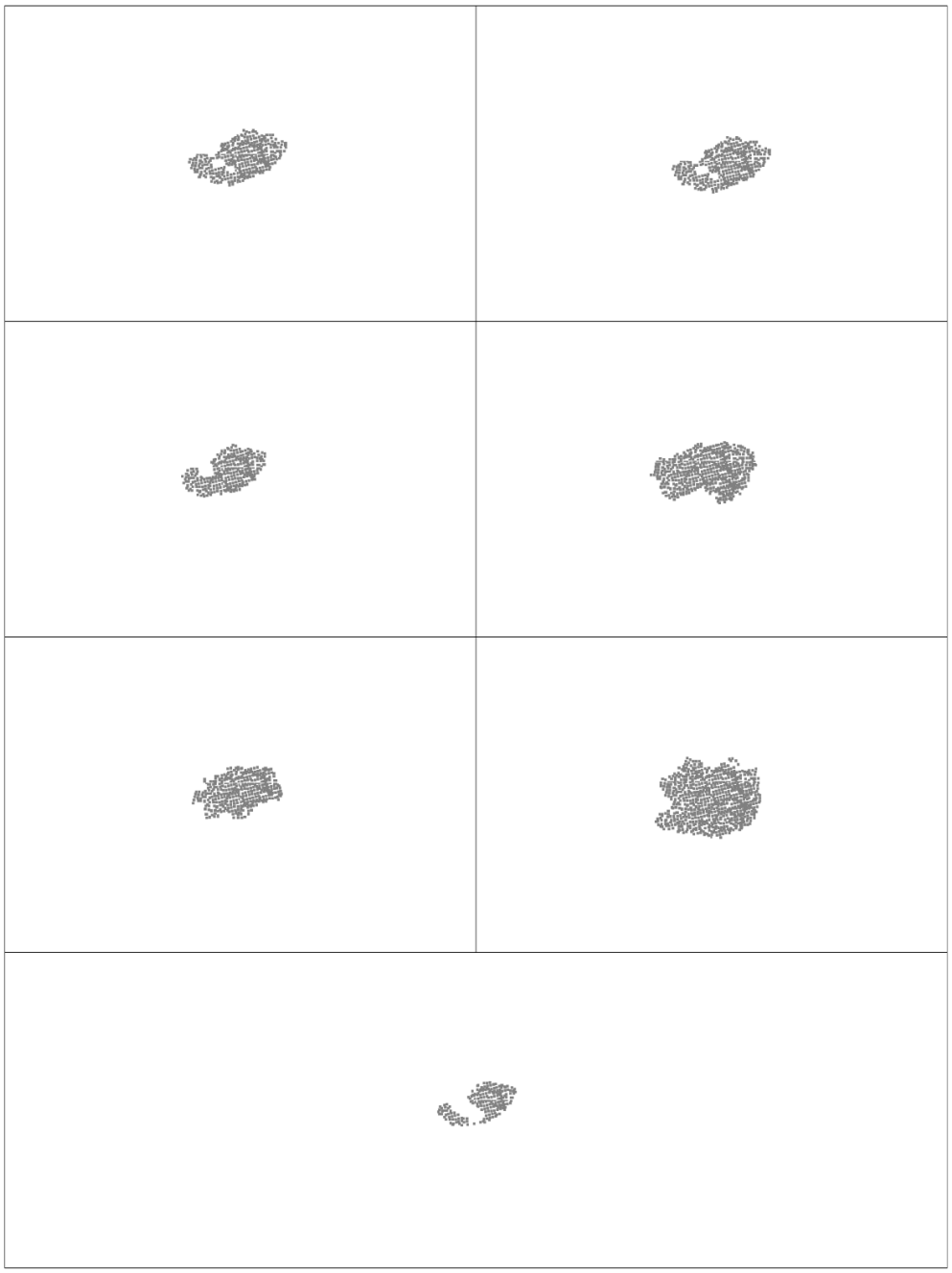


Figure 4-4 - Dataset 2 estimate truth attempts – Object 2

Figure 4-3 shows each of the seven estimates of truth for object 2 (the red strap). As discussed for object 1, this is isolated due to the geographic separation between the objects.

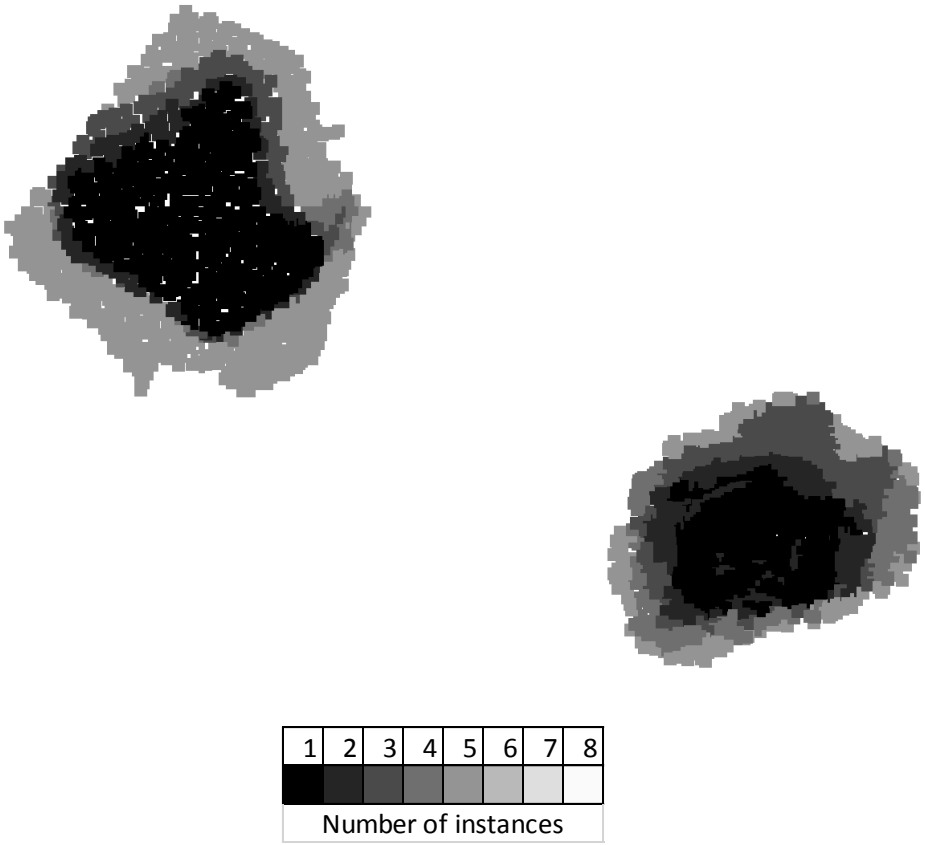


Figure 4-5 - Dataset 2 estimate truth models instances – object 1

Figure 4-5 and Figure 4-6 visualise each of the models created by counting the number of instances of each point and thresholding these into separate models.

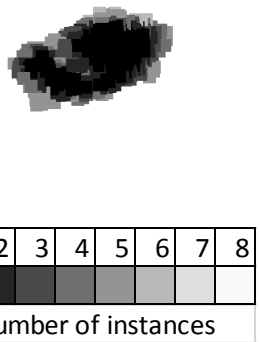


Figure 4-6 - Dataset 2 estimate truth models instances – object 2

4.1.3 Ground Truth - Dataset 3

To calculate the Dataset 3 estimated ground truth model, six unique estimates of truth were taken in isolation. Again as detailed in the methodology, each of these consist of a manual analysis of the imagery in combination with analysis of the point clouds generated.

Table 4-9 - Dataset 3 ordered ground truth estimates

Num of points	distance to mean	% deviation from mean
4777	-4112	-46.26
5426	-3463	-38.96
6765	-2124	-23.90
9389	499	5.62
11597	2707	30.46
15383	6493	73.05

Table 4-9 shows the individual iterations of estimated truths. Taking the mean of these presented in Table 4-10, their corresponding distances to the mean are also detailed and the magnitude of this is also presented as a percentage. Visualisation of each of these models is presented in Figure 4-7. Here the geographic separation can be seen to scale of the object before and after change.

Table 4-10 - Dataset 3 ground truth estimate descriptive statistics

total sum of all points	53337
unique points	18207
mean	8890
standard deviation	3723
coefficient of variance	0.419

Adding all the points together into a single cloud (with duplicates) gives a total of 53,337 points with 18,207 of these being unique points.

Using the full sum of all the points with duplicates, the number of instances of each unique point is calculated. A threshold is then applied to filter and create models to those that contain at six instances (is present in every estimate), through to a model that has only

one occurrence of each point (the entire set of unique points). The statistics for this are displayed in Table 4-11.

Table 4-11 - Dataset 3 ground truth models with threshold occurrences

threshold occurrences	unique points	probability of unique points	distance to mean	% deviation from mean	distance to half range
1	18207	1.00	9318	104.81	10588
2	11484	0.63	2595	29.19	3865
3	8795	0.48	-95	-1.06	1176
4	6717	0.37	-2173	-24.44	-902
5	5165	0.28	-3725	-41.90	-2454
6	2969	0.16	-5921	-66.60	-4650

Table 4-11 shows each of the models created from the entire set when the threshold is applied to individual point occurrences. It can be seen that as a threshold of 1 has the same number of unique points as presented in Table 4-10. As the threshold is increased, the number of points decreases. By taking the mean (Table 4-12), their respective distance to mean and percentage deviation to mean is also presented. Also by calculating the range (max – min) and half of this (Table 4-12), the midpoint geographically between the two extremes can be identified. The respective distance to this for each of the models is also presented. It can be seen that the closest model to the central point (defined as the truth) is when at least 3 occurrences are found of each of the points in the sum model. A visualisation of this can be seen in Figure 4-8 and Figure 4-9. Here an under estimation of points is found at 1.06% to the mean. This give us the error for this truth model to the actual truth.

Table 4-12 - Dataset 3 ground truth models descriptive statistics

mean	8890
standard deviation	4953
coefficient of variance	0.557
range	15238
half range	7619

From Figure 3-16, it can be seen that the model with the least error is closest to the mean. In this instance, a threshold of at 3 occurrences gives the closest to that boundary, and hence presents the estimated ground truth model for Dataset 3. Here there are 8,795 unique points, with a predicted error of 1.06% from the actual truth.

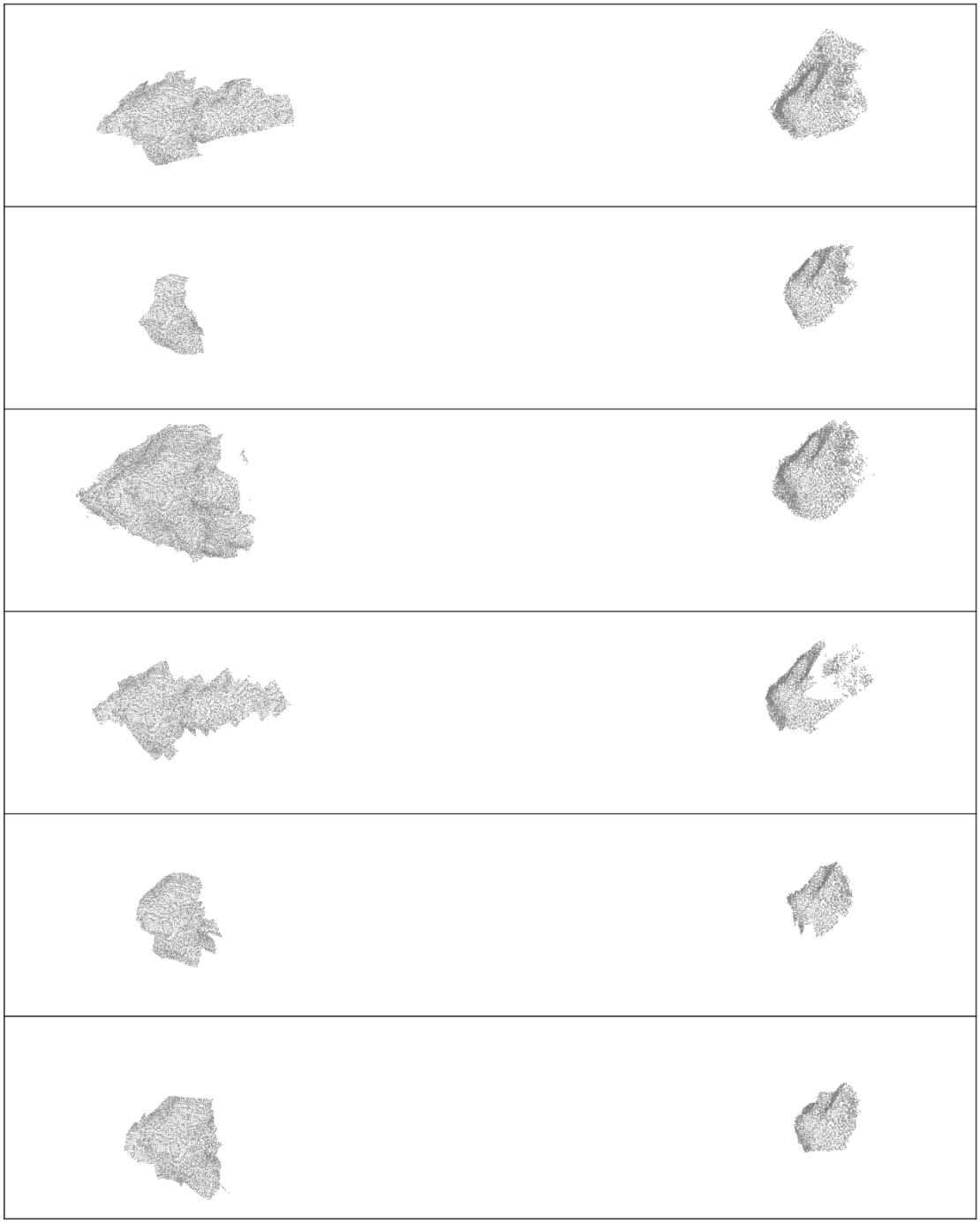


Figure 4-7 - Dataset 3 estimate truth attempts

Figure 4-7 shows each of the six estimates of truth. The separation between the objects has been kept so that the scale between them is understandable.

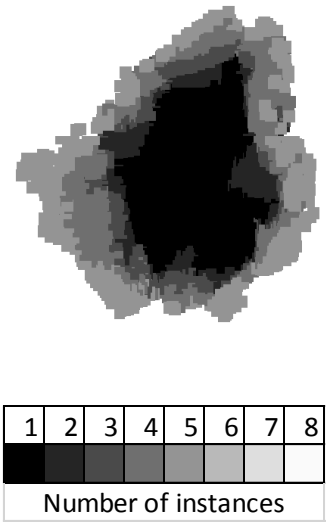


Figure 4-8 - Dataset 3 estimate truth models instances – object before

Figure 4-8 and Figure 4-9 visualise each of the models created by counting the number of instances of each point and thresholding these into separate models. As the lightness of point colour increases (the white value), the number of instances also increases. This is to show that at the lowest number of instances (minimal model) is smaller than that of all of the points (maximal model).

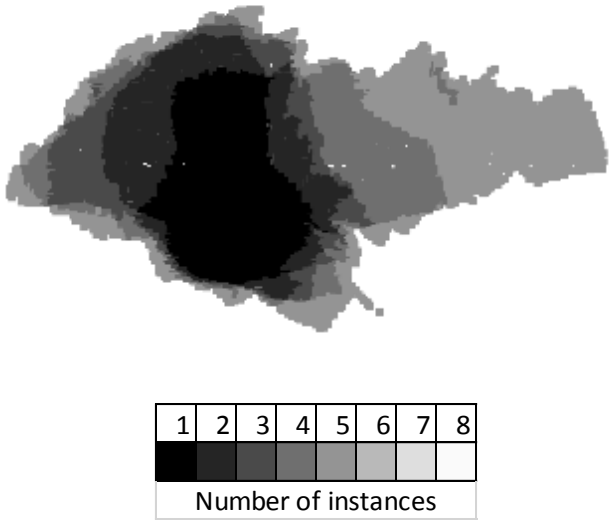


Figure 4-9 - Dataset 3 estimate truth models instances – object after

4.2 3D Model Creation

This section documents the reconstruction results of taking the 2D aerial imagery and applying SfM techniques with Photoscan to transform into 3D point clouds. Each of the datasets models are presented with some statistics of their composition.

4.2.1 Dataset 1

Dataset 1 contained 171 images for the first epoch (1A) and 187 for the second (1B). These were processed through the SfM pipeline to produce an output of 2,103,311 points in 1A and 2,389,536 points in 1B. The coloured point cloud created in Dataset 1A can be seen in the diagram below (Figure 4-10).

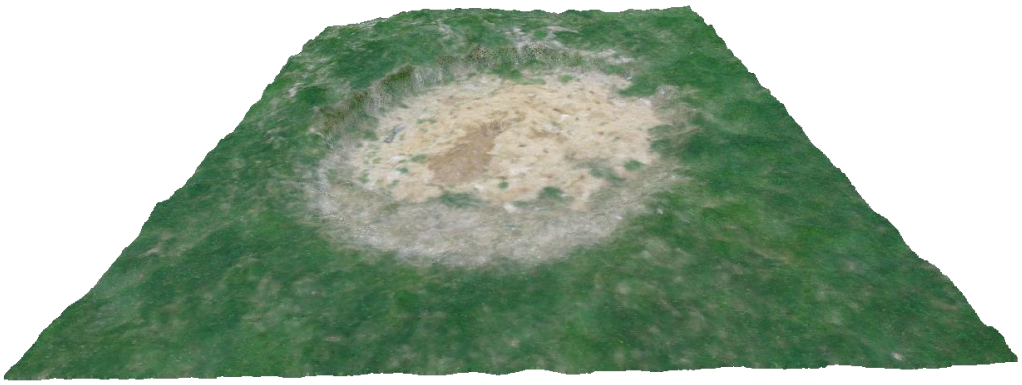


Figure 4-10 - Dataset1A Point cloud

4.2.2 Dataset 2

The 97 photos from Dataset 2A created a point cloud with 2,513,525 points and the 67 photos from 2B created a point cloud with 2,465,878 points respectively. While this is similar to Dataset 1, the area surveyed is larger and hence the resolution of the area represented by each point slightly less. Below (Figure 4-11), shows a bird's eye view of the point cloud from 2A.

Figure 4-11 - Dataset2A Point Cloud Bird's eye view

4.2.3 Dataset 3

The output of the SFM process created 3,973,265 points in 3A and 3,157,423 points in 3B. Below shows the full point clouds of both 3A and 3B and the change identified (Figure 4-12).

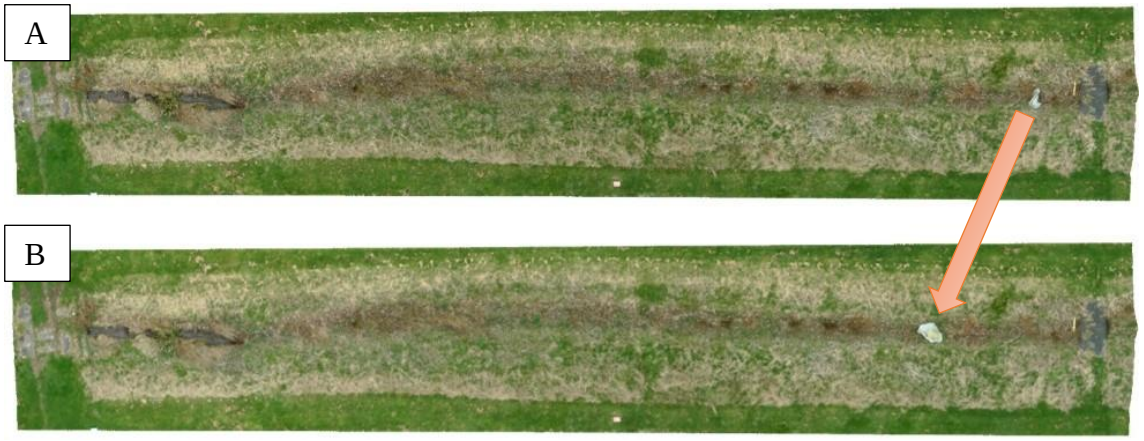


Figure 4-12 - Dataset3 Bird's eye view of both 3A and 3B point clouds

4.3 Change Detection - Dataset 1

This section covers the results for the applications of each of the algorithms identified for change detection in dataset 1. Here both the statistical results and where appropriate visualisations are presented. An initial analysis of these results is provided in this results section, with further analysis and discussion in Chapter 5.

As a visual overview of detected change, Figure 4-13 shows each of the following descriptors from an equivalent view using a colour map to represent that change:

- Colour XYZ
- Euclidean
- FPFH
- 3DSAC
- SHOT
- SHOT Colour

The changes introduced can be clearly identified with varying levels of noise and artefacts surrounding it. As an initial look at this - Colour, Euclidean and 3DSAC present the easiest analysis and identification of the change object. FPFH presents a middle ground where the object can be identified with substantial noise, and SHOT and SHOT Colour much more difficult to interpret. That said, next to the other images, it still offers an understanding of where the change is within the image.

An equivalent view of this is also presented in Figure 4-14 with a greyscale mapping. This is immediately much harder to interpret and the shapes of the changes harder to recognise. With this visualisation it is still possible to identify the change areas of the scene.

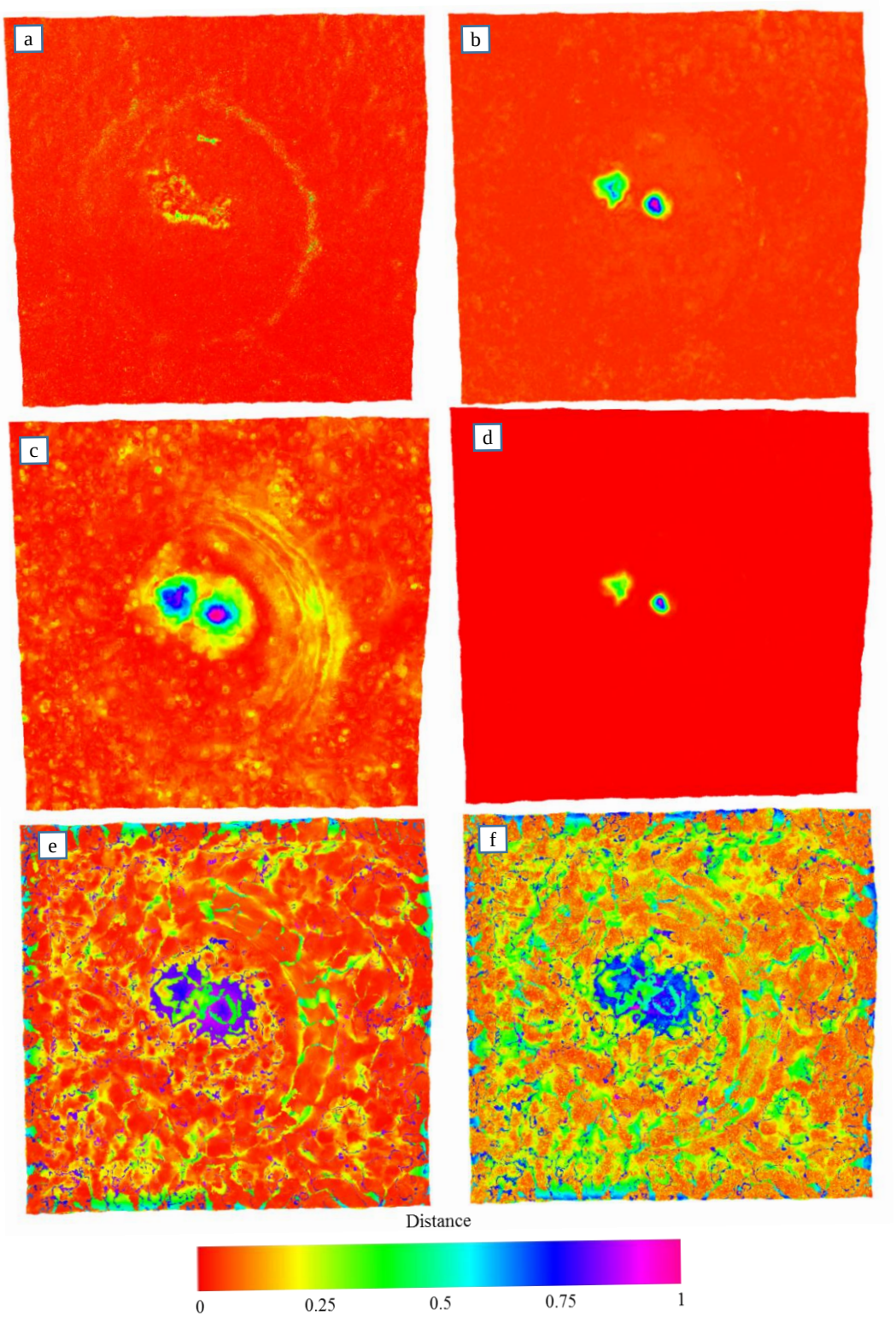


Figure 4-13 - Dataset 1 Descriptors with Colour Map Visualisation.

a) Colour XYZ b) Euclidean c)FPFH d)3DSAC e)SHOT f) SHOT Colour

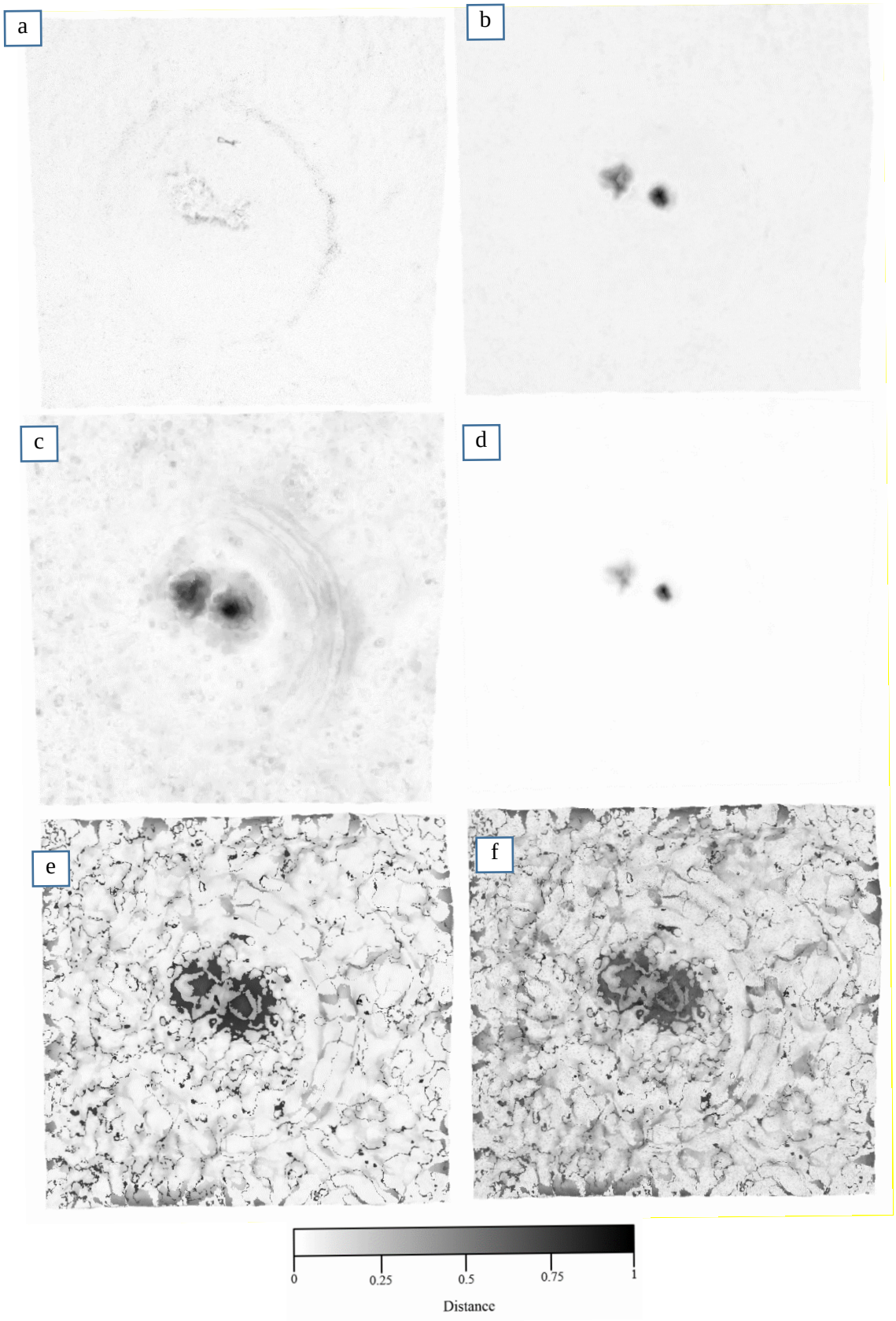


Figure 4-14 - Dataset1 Descriptors with Grayscale Map Visualisation

b) Colour XYZ b) Euclidean c)FPFH d)3DSAC e)SHOT f) SHOT Colour

4.3.1 Euclidean

Following the overview, a more comprehensive result set for the Euclidean descriptor is given. Figure 4-15 presents a graph of the performance metrics against the number of neighbours used to create a mean differencing. The AuROC curve seems to vary very little, followed by the TPR-FPR and then the MCC metrics. By close inspection, a small peak at very low k values can be identified in the TPR-FPR and MCC metrics. This suggests the maximal performance will be here. The FPR@90TRP metric is not comparable to the others, but appears to have little change and a magnitude similar to the AuROC metric.

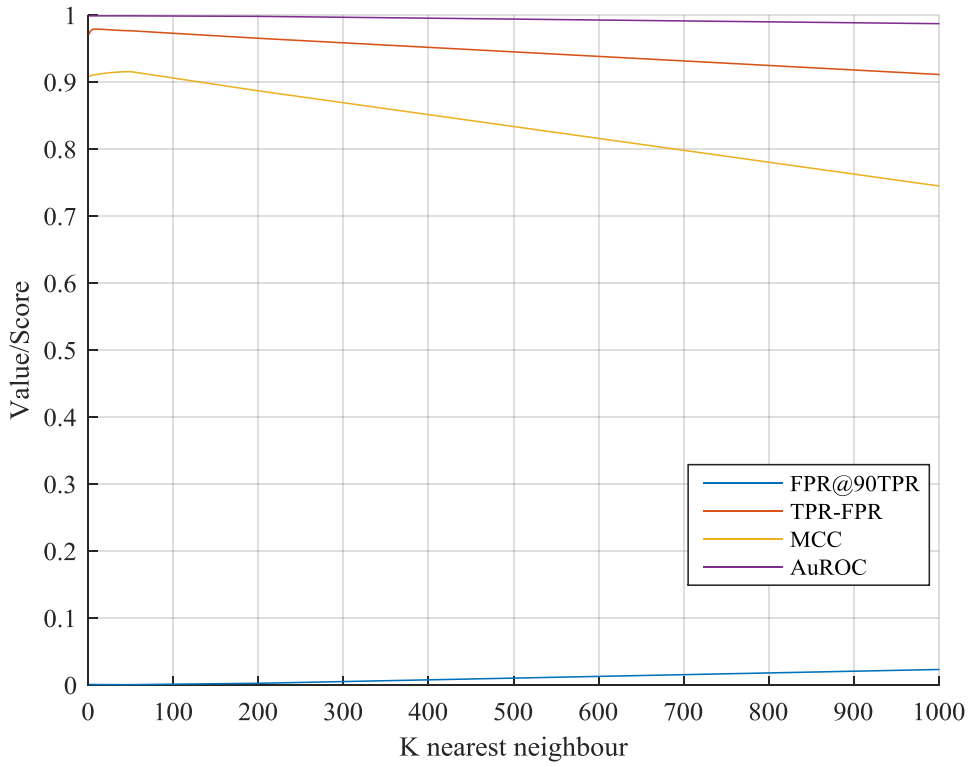


Figure 4-15 - Dataset 1 Euclidean Performance Scores

Table 4-13 shows the corresponding metrics to the graph. Here each of the descriptors numerical values can be seen as the number of nearest neighbours are varied. The variance in these is low but backups the earlier assertion that there is a peak in performance at a lower k value and then performance drops of very slowly.

Table 4-13 – Bunker 1 Euclidean Performance Measures

K nearest neighbours	FPR @ 90% TPR (P_{fa} @ 90%P_d)	Max TP Rate - FP Rate	Max MCC	Max AuROC Curve
1	0.0015	0.9716	0.9088	0.9991
2	0.0014	0.9759	0.9096	0.9992
5	0.0013	0.9793	0.9108	0.9993
10	0.0012	0.9796	0.9119	0.9994
20	0.0012	0.9789	0.9138	0.9994
30	0.0011	0.9781	0.9151	0.9994
40	0.0011	0.9774	0.9158	0.9994
50	0.0011	0.9771	0.916	0.9993
200	0.0031	0.9659	0.8874	0.9986
1000	0.0237	0.9118	0.7454	0.9877

4.3.2 Colour

This section looks at the results for using colour only as a differentiator between the two epochs of data across Dataset 1. Here the three descriptors LAB, RGB and XYZ are compared with the performance metrics for each of these listed. Table 4-14 shows the results for each of these against the chosen metrics. The XYZ colour space appears to offer the best performance numerically across all measurement techniques.

Table 4-14 - Dataset1 Colour Performance Measures

Colour Space	FPR @ 90% TPR (Pfa @ 90%Pd)	Max TPR- FPR	Max MCC	Max AuROC Curve
LAB	0.6958	0.4468	0.1694	0.7756
RGB	0.666	0.4295	0.1517	0.7732
XYZ	0.4331	0.4878	0.1814	0.826

Figure 4-17 shows the point cloud visualisation of these results with the colour maps and corresponding measured differences in the key at the bottom of the figure. Fragments A, C and E show the full data set before thresholding. Fragments B, D and F show a threshold has been applied giving a subset of 15,000 points with the largest distances. It's hard to visually discriminate between the three techniques, but it is noted that the spade object is

much more prevalent than the hole that has been dug. Figure 4-16 shows the specific example of the XYZ differencing with a grayscale mapping. Here it is again evident that the space is more prominent.

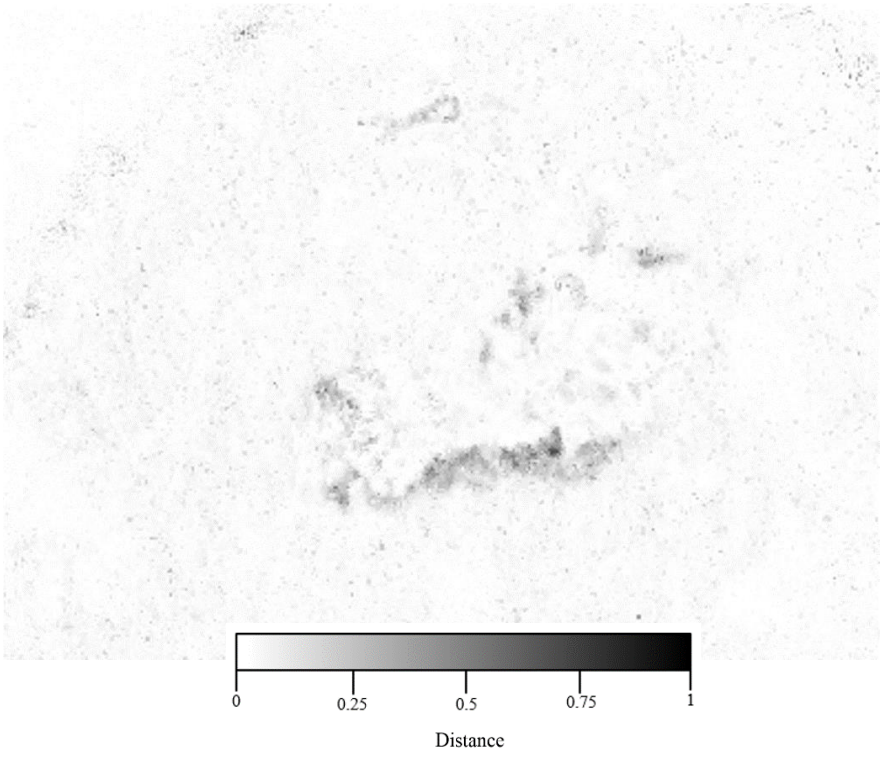


Figure 4-16 - Dataset1 XYZ Difference

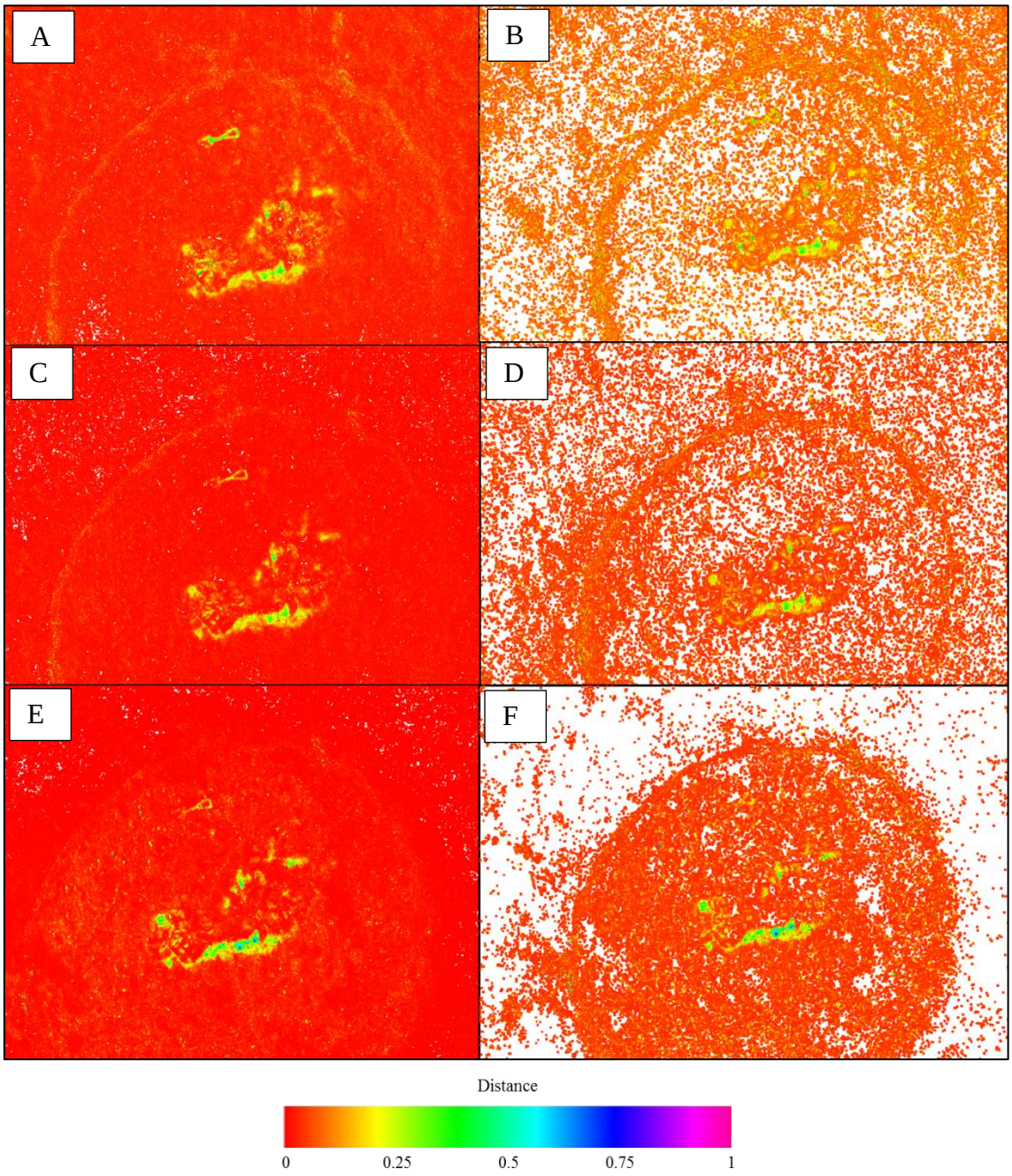


Figure 4-17 - Dataset1 - Colour Differencing

A)LAB– No filter B) LAB – largest 15,000 points C) RGB – No filter D)RGB – largest 15,000 points E) XYZ – No filter F)XYZ – largest 15,000 points

4.3.3 PFH

This section looks at the results for using the PFH descriptor as a differentiator between the two epochs of data across Dataset 1. Figure 4-18 shows a graph of the performance metrics against the number of neighbours used to create a mean differencing. Here, the TPR-FPR, MCC and AuROC metrics start low and dip even further before giving a steady positive gradient. This shows a poor performance at low k values but increasing as more neighbours are used. FPR@90TPR verifies this with a high peak at low k values coincident with the dips of the other metrics.

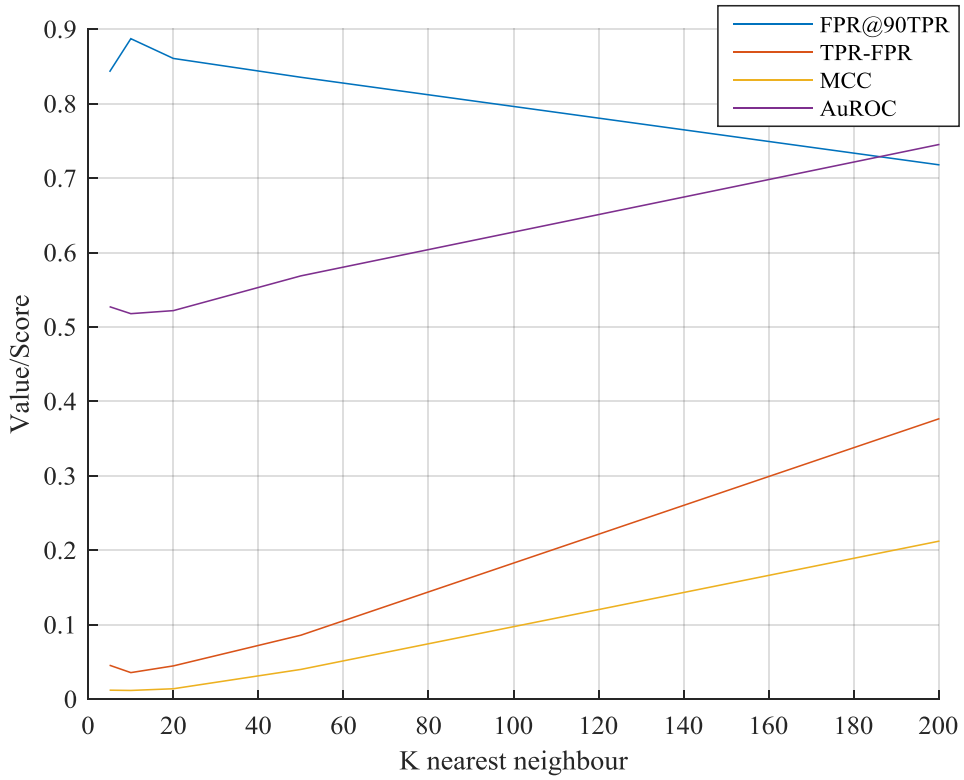


Figure 4-18 - Dataset1 PFH Performance Scores

Table 4-15 shows the corresponding metrics to the graph. Here each of the descriptors numerical values can be seen as the number of nearest neighbours are varied. The TPR-FRP, MCC and AuROC values hit their lowest values at 10 nearest neighbours and then steadily increase with maximums at 200. The extreme computation cost of describing the models with this descriptor limited the number of neighbours that could be considered. It appears that as more neighbours are introduced more performance can be expected. The

limit of where this becomes a maximal performance could not be found due to the limitation of computational costs. The FP@90%TP also mirrors this with a steady decrease as more neighbours are added.

Table 4-15 – Bunker 1 PFH Performance Measures

K nearest neighbours	FP @ 90% TP (P_{fa} @ 90%P_d)	Max TP Rate - FP Rate	Max MCC	Max AuROC Curve
5	0.8428	0.0456	0.0121	0.5273
10	0.8873	0.0357	0.0117	0.5179
20	0.8607	0.0447	0.014	0.522
50	0.8354	0.086	0.04	0.5687
200	0.7179	0.3766	0.2122	0.7451

4.3.4 FPFH

This section looks at the results for using the FPFH descriptor as a differentiator between the two epochs of data across Dataset 1. Figure 4-19 shows a graph of the performance metrics against the number of neighbours used to create a mean differencing. Here, the TPR-FPR, MCC and AuROC show different trends to the performance with all starting off very low representing poor performance at very low numbers of neighbours. The AuROC curve shows an immediate and sharp increase though to about 500 nearest neighbours where it levels off. The TPR-FRP also shows an immediate but slower increase followed by a level off (and a slight dip at 2000 nearest neighbours). The MCC curve has a much more varied shape to its value, with an immediate increase followed by a few peaks and troughs. It appears to steadily decline after a k value of 2000. The FRP@90TRP rapidly declines with a level off mirroring the AuROC curve.

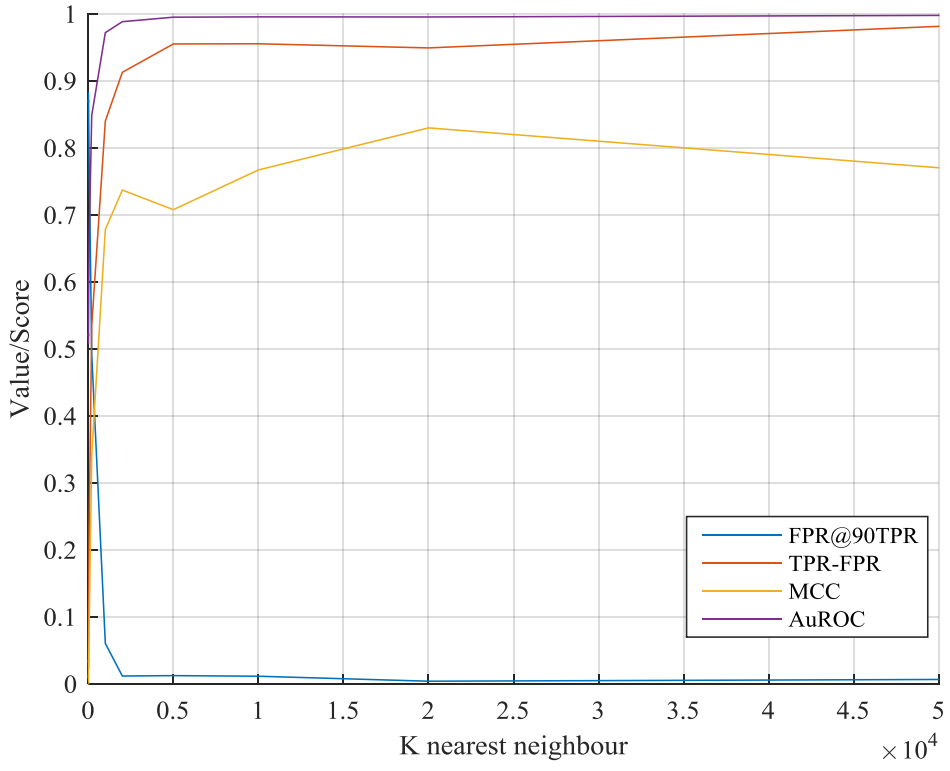


Figure 4-19 - Dataset1 FPFH - Performance Scores

Table 4-16 – Bunker 1 FPFH Performance Measures

K nearest neighbours	FP @ 90% TP (P_{fa} @ 90%P_d)	Max TP Rate - FP Rate	Max MCC	Max AuROC Curve
5	0.8679	0.0354	0.0118	0.5232
10	0.8718	0.0331	0.0104	0.5191
20	0.8492	0.1099	0.0285	0.5766
50	0.779	0.2373	0.0889	0.6652
200	0.4924	0.5348	0.3413	0.8487
1000	0.0606	0.8404	0.6785	0.9722
2000	0.012	0.9131	0.7373	0.9887
5000	0.0126	0.9554	0.708	0.9954
10000	0.0116	0.9557	0.7673	0.9958
20000	0.0042	0.9495	0.8301	0.9957
50000	0.0068	0.9817	0.7706	0.998

Table 4-16 shows the corresponding metrics to the graph. Here each of the descriptors numerical values can be seen as the number of nearest neighbours are varied.

4.3.5 PCE

This section looks at the results for using the PCE descriptor as a differentiator between the two epochs of data across Dataset 1. Figure 4-20 shows a graph of the performance metrics against the number of neighbours used to create a mean differencing. Here, the TPR-FPR and MCC scores maintain a low level throughout, with a very slight incline when k values are more than 200. The AuROC appears to mirror this but at an additional level of around 0.5. The FPR@90TPR starts high and stays high throughout with a slight decline. Taking all these into considerations suggests the performance does not gain much as k values are increased.

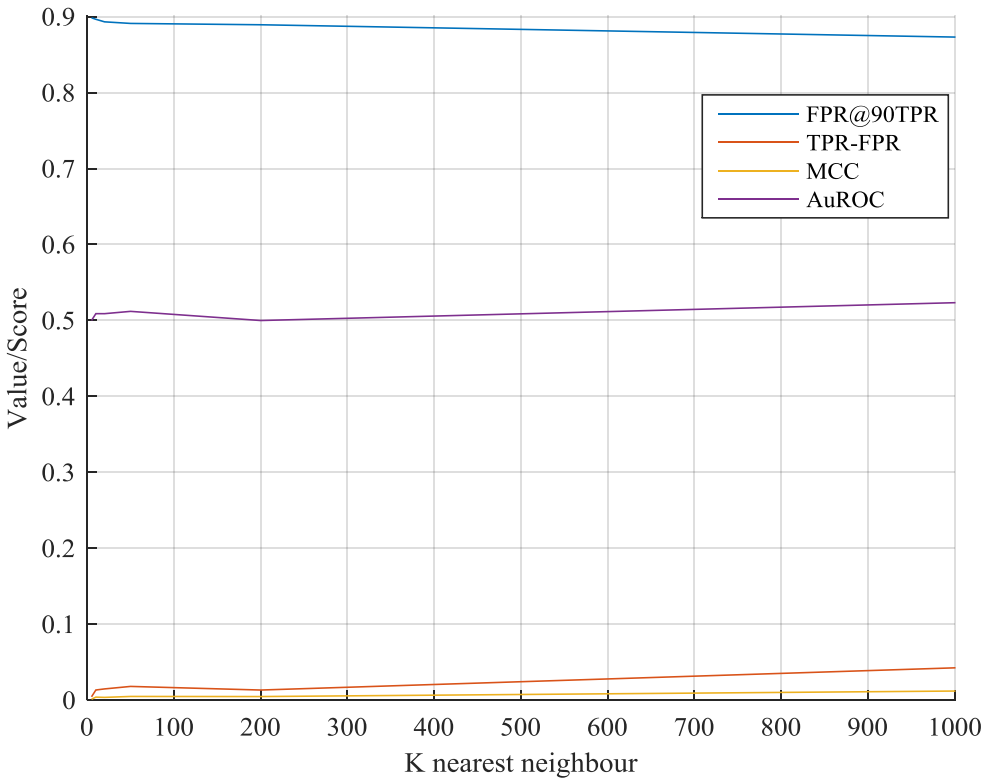


Figure 4-20 - Dataset1 PCE – Performance Scores

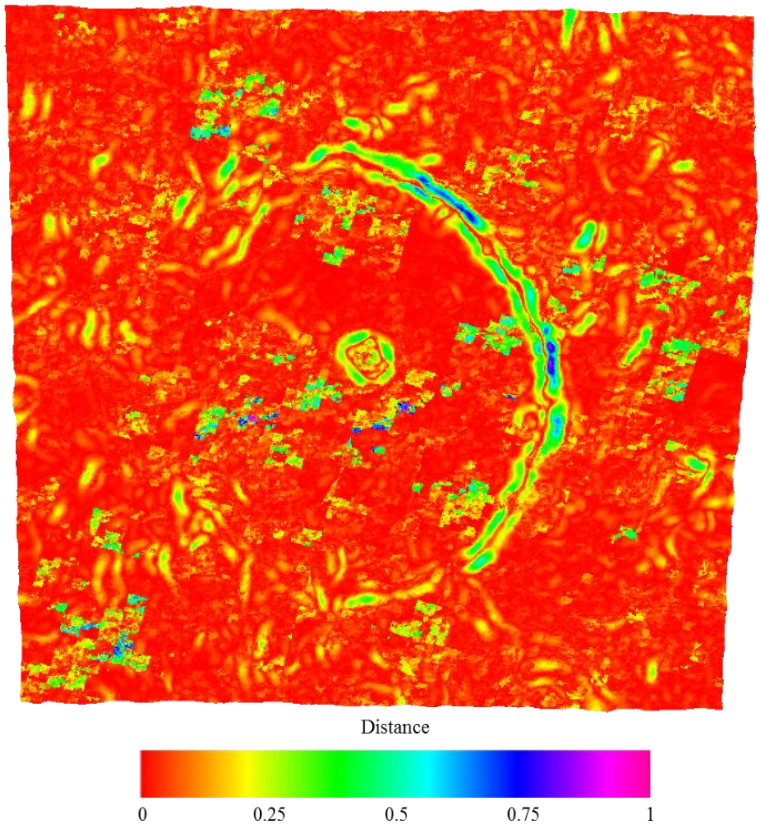


Figure 4-21 - Dataset 1 PCE Visualisation

A visual representation in Figure 4-21 shows a pattern of higher change detected around the banks of the golf bunker. It is still possible to recognise the change introduced but this is less prevalent than that of the banks.

Table 4-17 – Dataset 1 PCE Performance Measures

K nearest neighbours	FP @ 90% TP (P_{fa} @ 90%P_d)	Max TP Rate - FP Rate	Max MCC	Max AuROC Curve
5	0.8985	0.0043	0.0011	0.4999
10	0.8965	0.0133	0.0039	0.5089
20	0.8932	0.0147	0.0036	0.5088
50	0.8911	0.018	0.0048	0.5119
200	0.8893	0.0132	0.0047	0.4998
1000	0.8731	0.0425	0.0119	0.5233

Table 4-17 shows the corresponding metrics to the graph and visualisation. Here each of the descriptors numerical values can be seen as the number of nearest neighbours are varied.

4.3.6 SHOT

This section looks at the results for using the SHOT descriptor as a differentiator between the two epochs of data across Dataset 1. Figure 4-22 shows a graph of the AuROC performance metric against the number of neighbours used to create a mean differencing across various radii. Here the initial peak is seen when the k value is 2 and then decreases through to k values of 10 where it stabilises and levels off. .The highest performance is seen at a radius of 0.5 and this decreases as the radius is decreased. There is a crossover between the performance of descriptor using radius of 0.01 and 002. Table 4-18 shows the corresponding data to this graph.

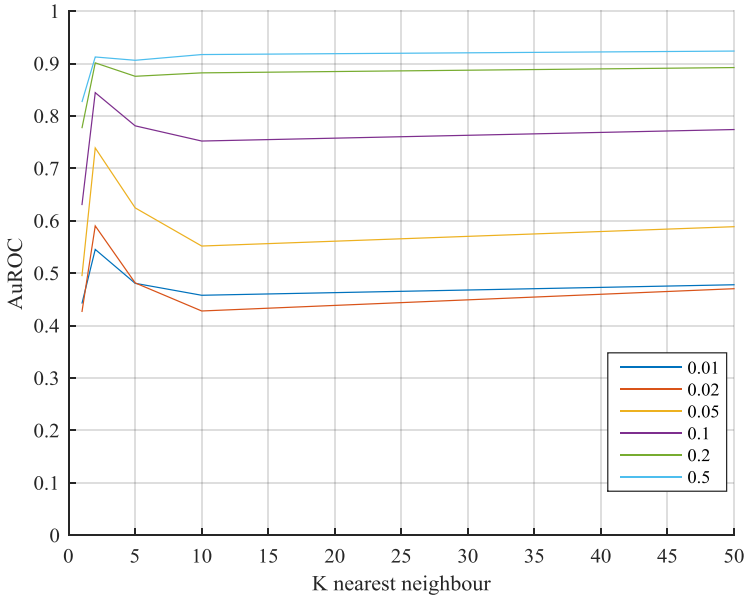


Figure 4-22 Dataset1 SHOT - Graph of AuROC Curves

Table 4-18 – Dataset1 SHOT AuROC Values

AuROC	Radius					
K Nearest Neighbour	0.01	0.02	0.05	0.1	0.2	0.5
1	0.4421	0.4261	0.4944	0.6299	0.7767	0.8266
2	0.5454	0.5903	0.739	0.8445	0.9013	0.9126
5	0.4808	0.4816	0.6247	0.7812	0.8757	0.9062
10	0.4579	0.428	0.5519	0.7522	0.8823	0.9172
50	0.478	0.4705	0.5888	0.7741	0.8923	0.9238

Figure 4-23 shows a graph of the TPR-FPR performance metric against the number of neighbours used to create a mean differencing across various radii. Here again the initial peak is seen when the k value is 2 and then decreases through to k values of 10 where it stabilises and levels off. .Table 4-18 shows the corresponding data to this graph. Here it is evident again that maximal performance is found at a higher radius value and this is decreases as the radius is decreased. Table 4-19 shows the corresponding data to this graph.

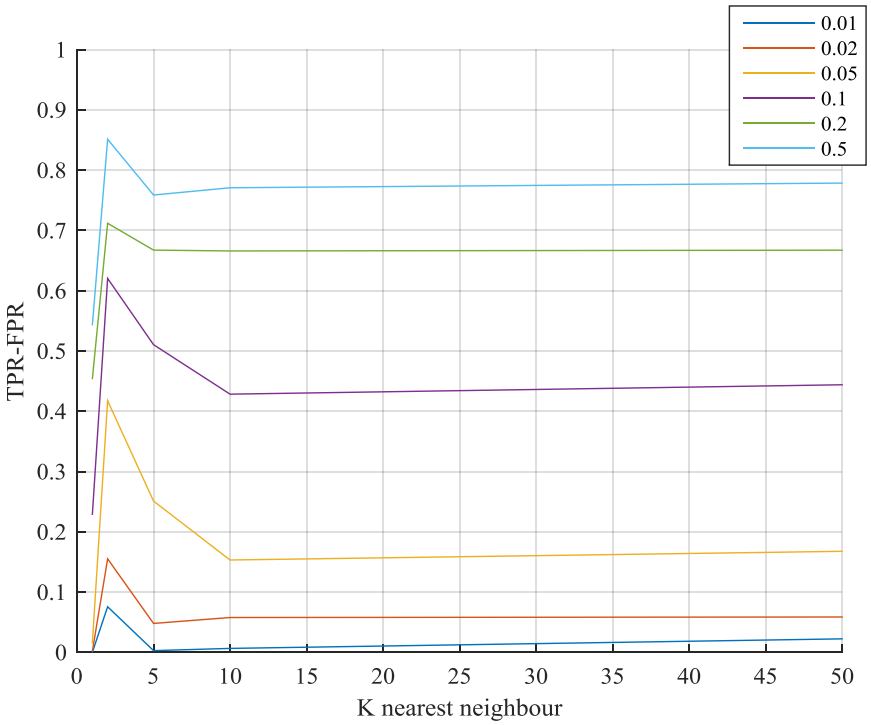


Figure 4-23 Dataset1 SHOT - Graph of TPR-TPR metrics

Table 4-19 – Dataset1 SHOT Max TP Rate – FP Rate

TP Rate - FP Rate	Radius					
K Nearest Neighbour	0.01	0.02	0.05	0.1	0.2	0.5
1	0	0.0016	0.0148	0.228	0.4532	0.5424
2	0.0757	0.1551	0.4174	0.6204	0.7119	0.8513
5	0.0029	0.048	0.251	0.5105	0.6674	0.7588
10	0.0066	0.0579	0.1533	0.4284	0.666	0.7708
50	0.0224	0.0587	0.1677	0.444	0.6673	0.7786

Figure 4-24 shows a graph of the MCC performance metric against the number of neighbours used to create a mean differencing across various radii. Here the initial peak is seen when the k value is 2 and then mostly decreases through to k values of 10 where it stabilises and levels off. .Its noted this are consistently low. Table 4-20 shows the corresponding data to this graph.

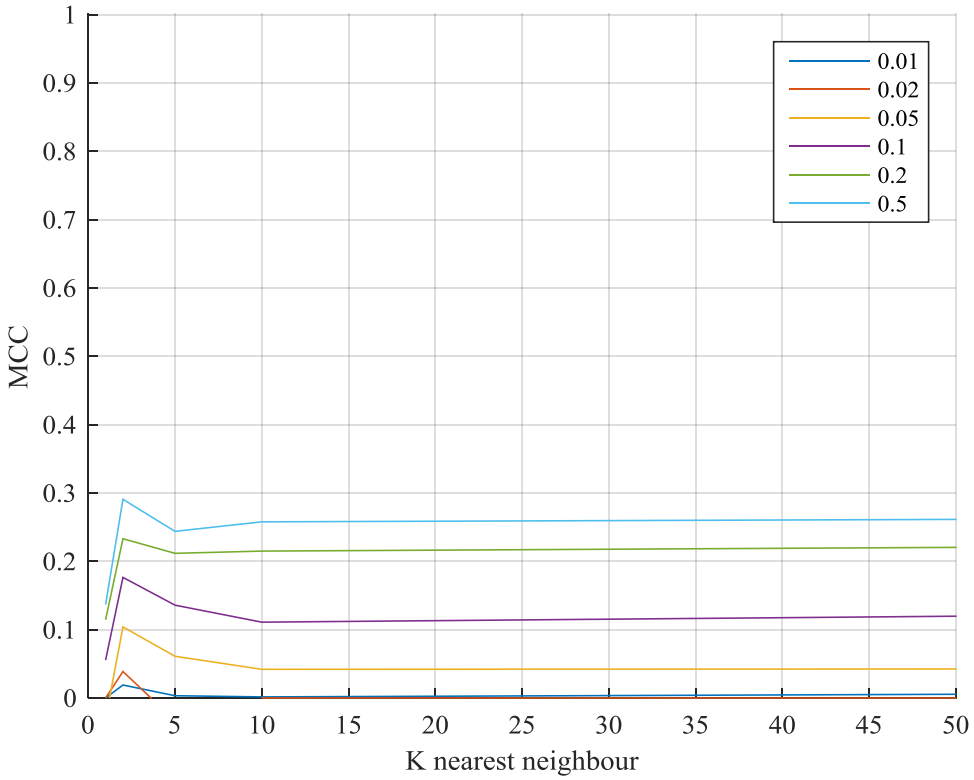


Figure 4-24 Dataset1 SHOT - Graph of MCC metrics

Table 4-20 – Dataset1 SHOT Max MCC

MCC	Radius					
K Nearest Neighbour	0.01	0.02	0.05	0.1	0.2	0.5
1	0.0001	0	-0.0304	0.0555	0.1146	0.1369
2	0.019	0.0388	0.104	0.1765	0.233	0.2908
5	0.0033	-0.0332	0.061	0.1358	0.2117	0.2438
10	0.0016	0	0.042	0.111	0.2149	0.2578
50	0.0055	0	0.0425	0.1197	0.2204	0.2613

Figure 4-24 shows a graph of the FPR@90%TPR performance metric against the number of neighbours used to create a mean differencing across various radii. Here the initial trough is seen when the k value is 2 and then mostly decreases through to k values of 10 where it stabilises and levels off. Table 4-21 shows the corresponding data to this graph.

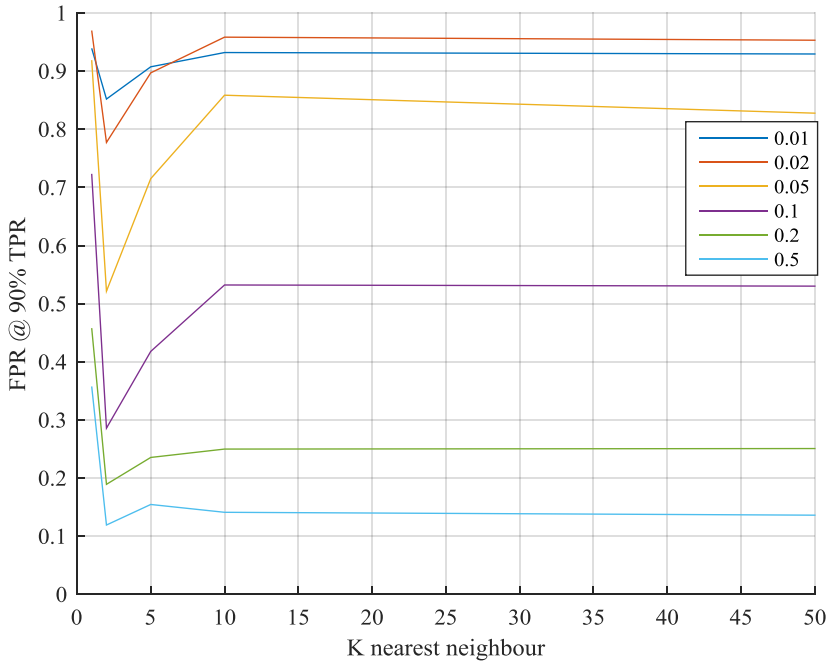


Figure 4-25 Dataset1 SHOT - Graph of Area FPR @ 90% TPR metrics

Table 4-21 – Dataset1 SHOT FPR @ 90% TPR

FPR @ 90% TPR (P_{fa} @ 90% P_d)	Radius					
K Nearest Neighbour	0.01	0.02	0.05	0.1	0.2	0.5
1	0.9393	0.9699	0.919	0.7234	0.458	0.3577
2	0.852	0.7774	0.5216	0.286	0.1893	0.1193
5	0.9075	0.8971	0.7151	0.4179	0.2355	0.1546
10	0.932	0.9585	0.8587	0.5322	0.2499	0.1412
50	0.9296	0.9533	0.8278	0.5302	0.2509	0.1362

Figure 4-26 shows each of the performance metrics on one graph for comparative analysis for a radius of 0.5. As seen in the previous graph, the performance has a slight peak at k values of 2 and then levels off. This suggests there is little extra performance to be gained by additional nearest neighbours.

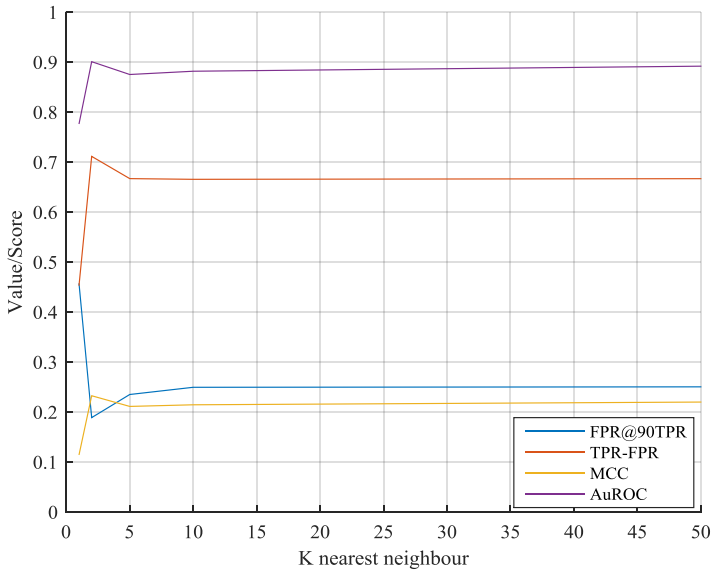


Figure 4-26 Dataset 1 SHOT - Performance Scores @ Radius 0.5

Figure 4-27 shows the visual representation of the change detection data. Here the pre-threshold data is displayed in image 27a and a threshold applied in 27b. It can be seen that most of the change is present around contours in the image but no additional present around that of the changes implemented.

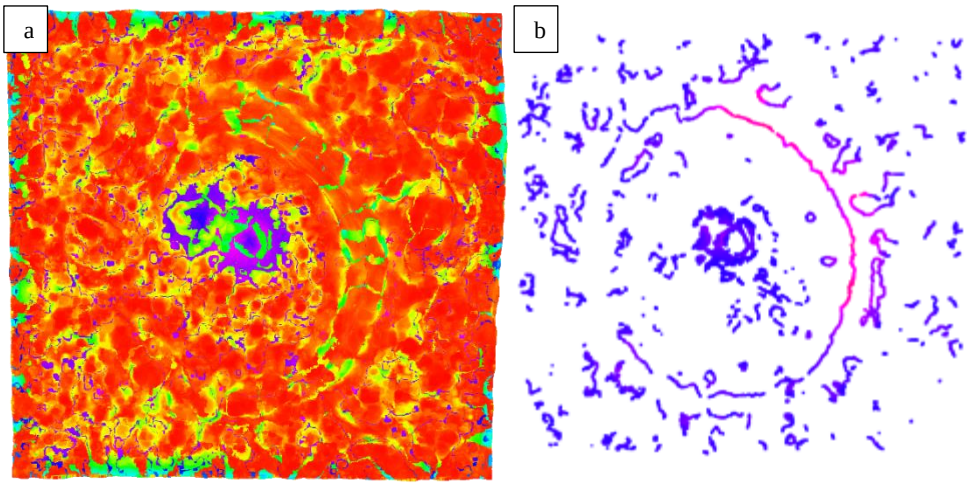


Figure 4-27 Dataset1 SHOT - Change Visualisation at radius 0.5

a) no threshold applied b) threshold applied

4.3.7 SHOT Colour

This section looks at the results for using the SHOT Colour descriptor as a differentiator between the two epochs of data across Dataset 1. Figure 4-28 shows a graph of the FPR@90TPR performance metric against the number of neighbours used to create a mean differencing with varied radii. Here, instability a very low k numbers can be seen and then a levelling effect through the rest of the data. The best performance appears to be from a radius of 0.5 with decreasing performance as the radius is decreased with exception to radius 0.01 and 0.02 being reversed. Table 4-22 shows the corresponding data to this graph.

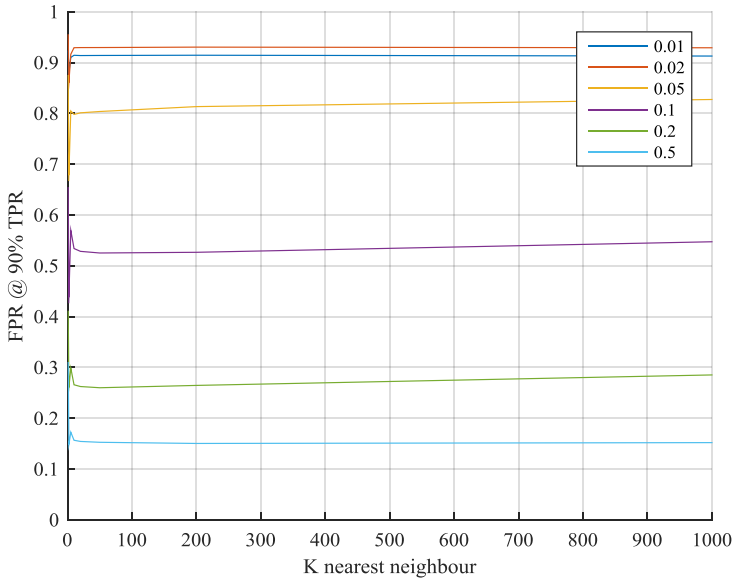


Figure 4-28 Dataset1 SHOT Colour - Graph of FPR @ 90% TPR

Table 4-22 Dataset1 SHOT Colour - Table of FPR @ 90% TPR

FPR @ 90% TPR (P_{fa} @ 90% P_d)	Radius					
k	0.01	0.02	0.05	0.1	0.2	0.5
1	0.9236	0.9556	0.8755	0.6548	0.4115	0.3106
2	0.8788	0.8598	0.6783	0.4378	0.2601	0.1483
5	0.9101	0.9163	0.8047	0.5697	0.2987	0.1725
10	0.9142	0.9291	0.7978	0.5339	0.2658	0.1566
20	0.9137	0.9294	0.8009	0.5284	0.2623	0.1543
50	0.9139	0.9295	0.8036	0.5251	0.2599	0.1526

Figure 4-29 shows a graph of the TPR-FPR performance metric against the number of neighbours used to create a mean differencing with varied radii. Again instability a very low k numbers can be seen and then a levelling effect through the rest of the data. The best performance appears to be from a radius of 0.5 with decreasing performance as the radius is decreased. Table 4-23 shows the corresponding data to this graph.

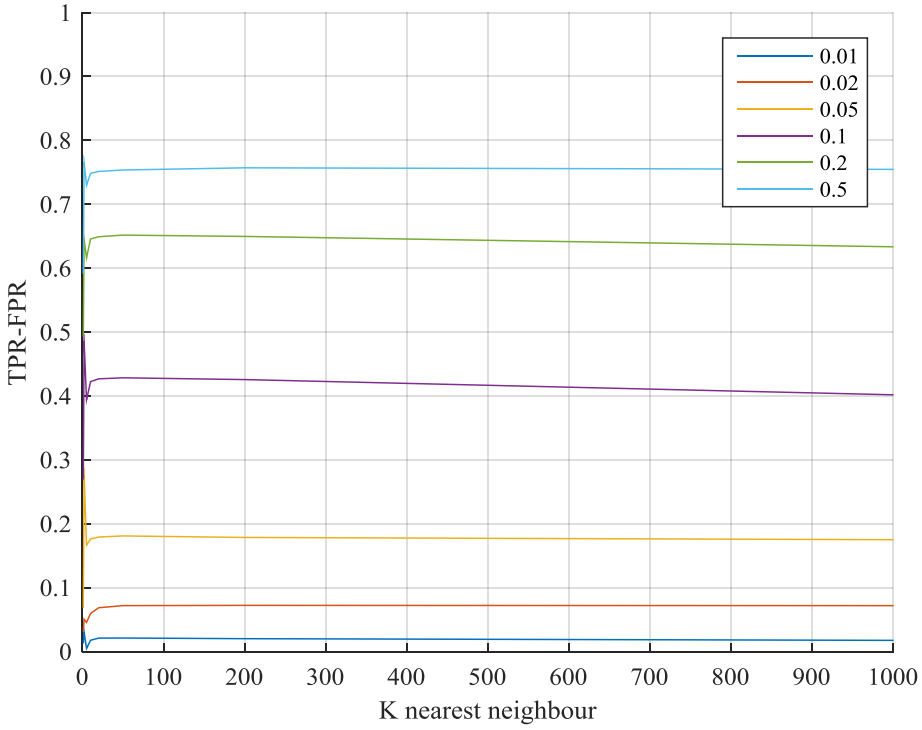


Figure 4-29 Dataset1 SHOT Colour- Graph of TPR-FPR metrics

Table 4-23 Dataset1 SHOT Colour - Table of TPR-FPR metrics

Max TPR - FPR	Radius					
k	0.01	0.02	0.05	0.1	0.2	0.5
1	0.0127	0.0316	0.0679	0.2693	0.4932	0.5914
2	0.0304	0.0505	0.2762	0.4855	0.6447	0.7652
5	0.0050	0.0457	0.1667	0.3941	0.6163	0.7299
10	0.0178	0.0597	0.1762	0.4225	0.6458	0.7485
20	0.0211	0.0687	0.1793	0.4269	0.6492	0.7515
50	0.0213	0.0720	0.1810	0.4285	0.6518	0.7537

Figure 4-30 shows a graph of the MCC performance metric against the number of neighbours used to create a mean differencing with varied radii. The instability a very low k numbers can be seen and then a levelling effect through the rest of the data. Again, the best performance appears to be from a radius of 0.5 with decreasing performance as the radius is decreased. Table 4-24 shows the corresponding data to this graph.

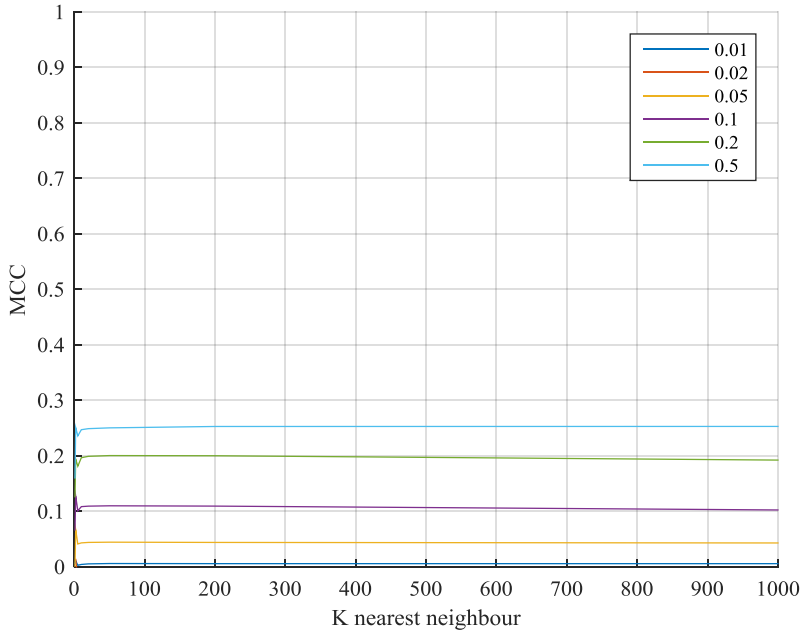


Figure 4-30 Dataset1 SHOT Colour - Graph of MCC metrics

Table 4-24 Dataset1 SHOT Colour - Table of MCC metrics

MCC	Radius					
k	0.01	0.02	0.05	0.1	0.2	0.5
1	0.0051	NaN	0.0000	0.0653	0.1252	0.1591
2	0.0111	0.0144	0.0669	0.1275	0.1940	0.2509
5	0.0028	-0.0225	0.0412	0.1010	0.1813	0.2359
10	0.0043	-0.0509	0.0431	0.1085	0.1966	0.2471
20	0.0053	NaN	0.0440	0.1094	0.1992	0.2490
50	0.0060	-0.0991	0.0444	0.1100	0.2005	0.2504

Figure 4-31 shows a graph of the AuROC performance metric against the number of neighbours used to create a mean differencing with varied radii. The instability a very low k numbers can be seen and then a levelling effect through the rest of the data. Again, the best performance appears to be from a radius of 0.5 with decreasing performance as the radius is decreased. Table 4-25 shows the corresponding data to this graph.

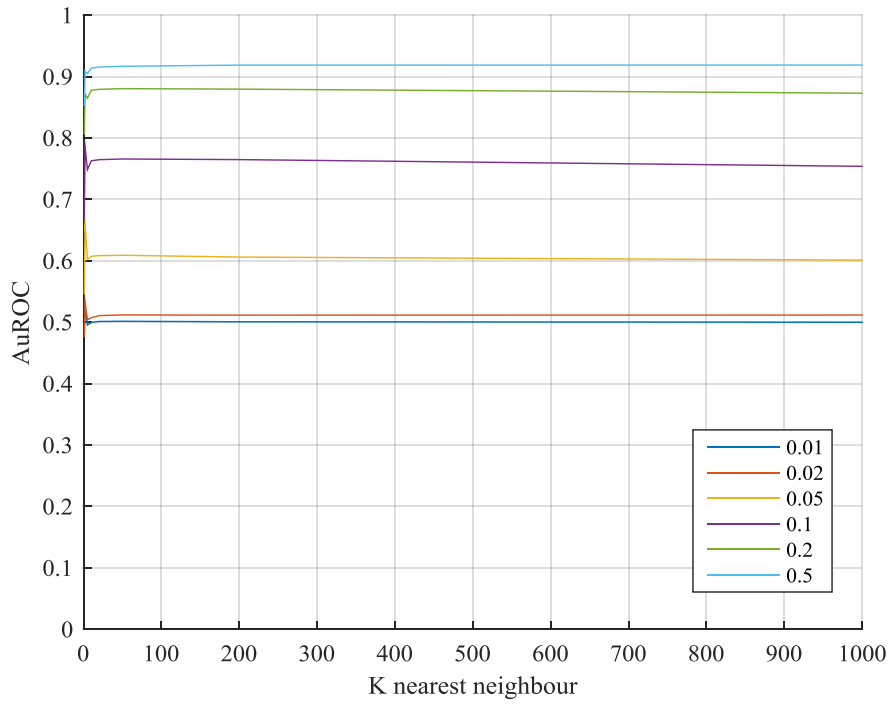


Figure 4-31 Dataset1 SHOT Colour - Graph of AuROC metrics

Table 4-25 Dataset1 SHOT Colour - Table of AuROC metrics

AUC	Radius					
k	0.01	0.02	0.05	0.1	0.2	0.5
1	0.4839	0.4758	0.5451	0.6697	0.8057	0.8522
2	0.5225	0.5331	0.6600	0.7859	0.8717	0.9091
5	0.4955	0.5042	0.6023	0.7486	0.8649	0.9047
10	0.4997	0.5074	0.6075	0.7630	0.8777	0.9140
20	0.5013	0.5109	0.6085	0.7648	0.8794	0.9156
50	0.5017	0.5120	0.6091	0.7659	0.8805	0.9169

Figure 4-32 shows each of the performance metrics on one graph for comparative analysis for a radius of 0.5. As seen in the previous graphs, instability to low k number is seen with a levelling of this through the rest of the series. This suggests there is little extra performance to be gained by additional nearest neighbours.

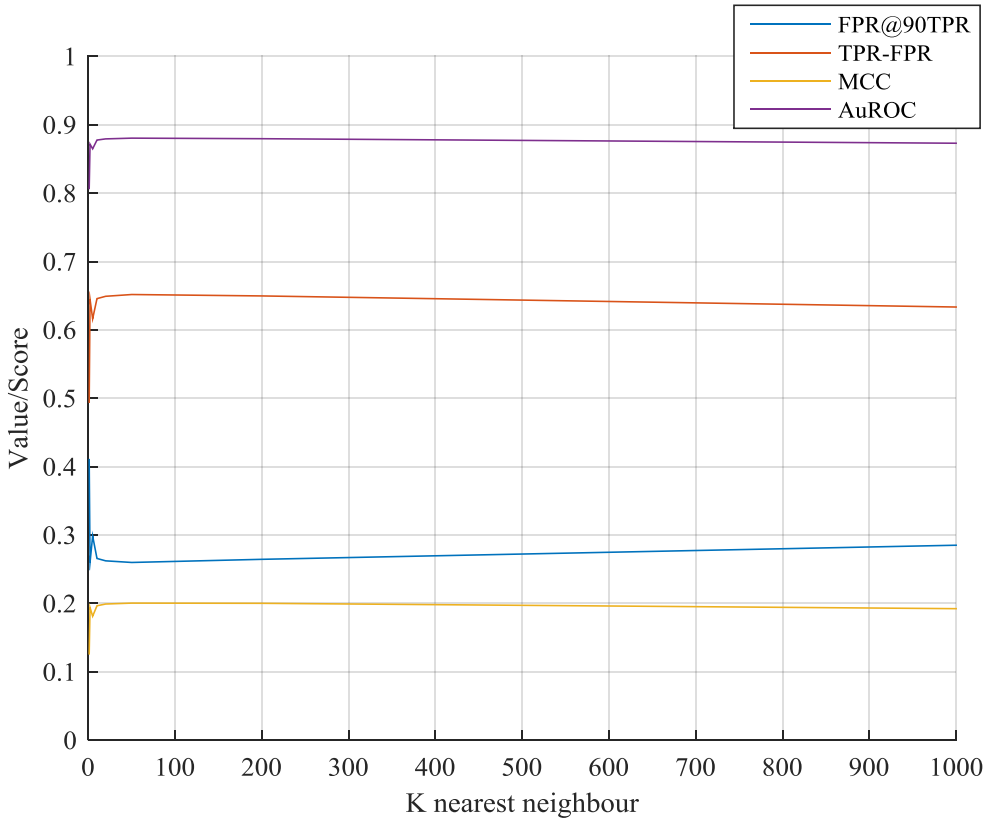


Figure 4-32 - Dataset 1 SHOT Colour – Performance Scores @ Radius 0.5

4.3.8 3DSAC

This section looks at the results for using the 3DSAC descriptor as a differentiator between the two epochs of data across Dataset 1. Figure 4-33 visualises the change detection results across varied weighting of colour to structure. Figure 4-33A shows the changes of the hole very clearly with little background noise surrounding it. With this, it is identified that the presence of the space cannot be achieved. As colour is introduced in Figure 4-33B the spade can be identified and some presence of noise and artefacts around the hole can be appreciated. Through Figure 4-33C and Figure 4-33D increases noise can be seen with the spade becoming more prevalent. In Figure 4-33E the hole appears to be less identifiable with less change detected around this area and increased noise around the scene. Finally in Figure 4-33F where no structure weighting is given in place of maximum colour weighting, the hole is almost unrecognisable but the spade is easily identified.

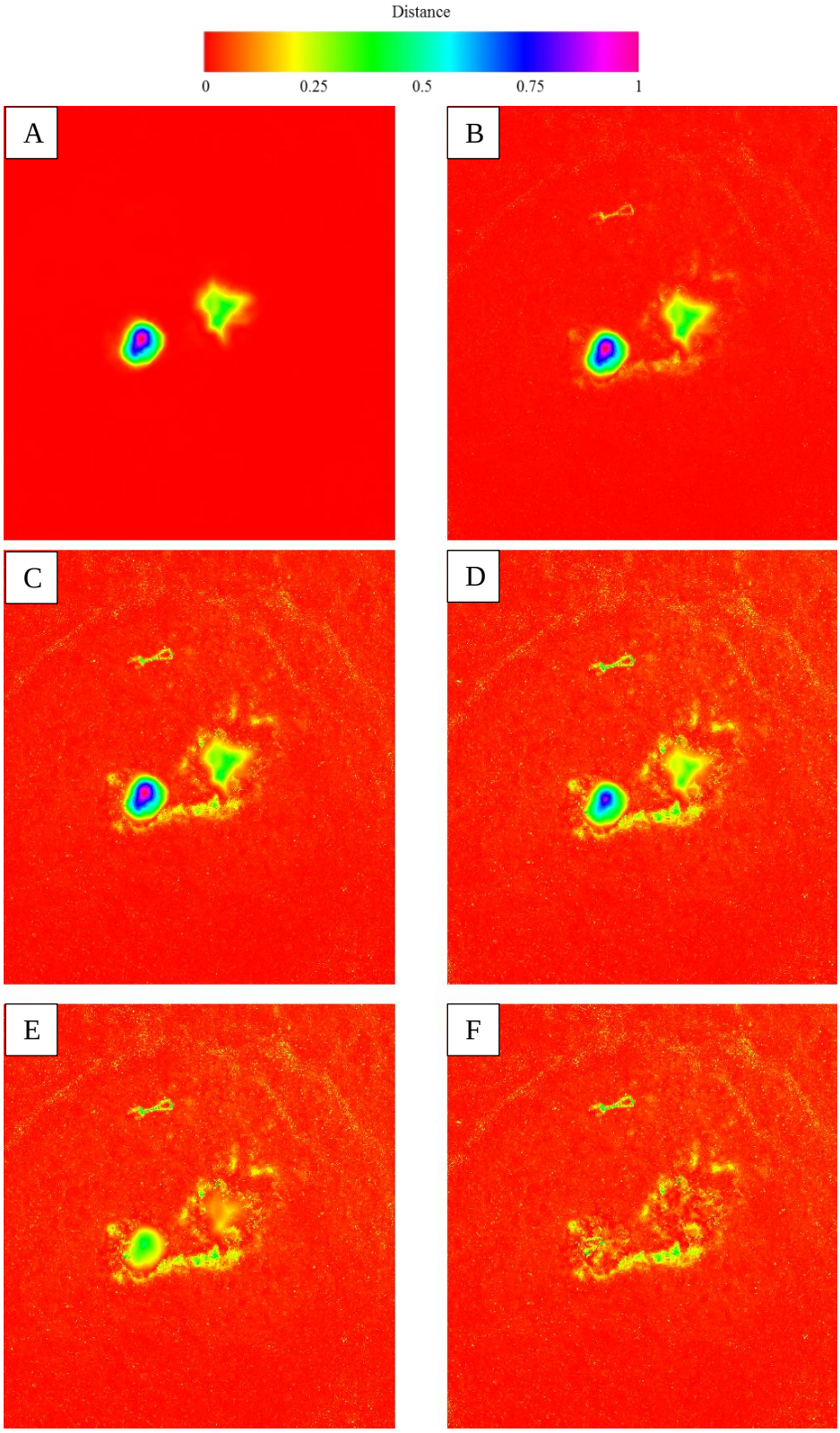


Figure 4-33 – 3DSAC weighting changes

A) 100% Structure 0% Colour B) 100% Structure 40% Colour C) 100% Structure 80% Colour D) 80% Structure 100% Colour E) 40% Structure 100% Colour F) 0% Structure 100% Colour

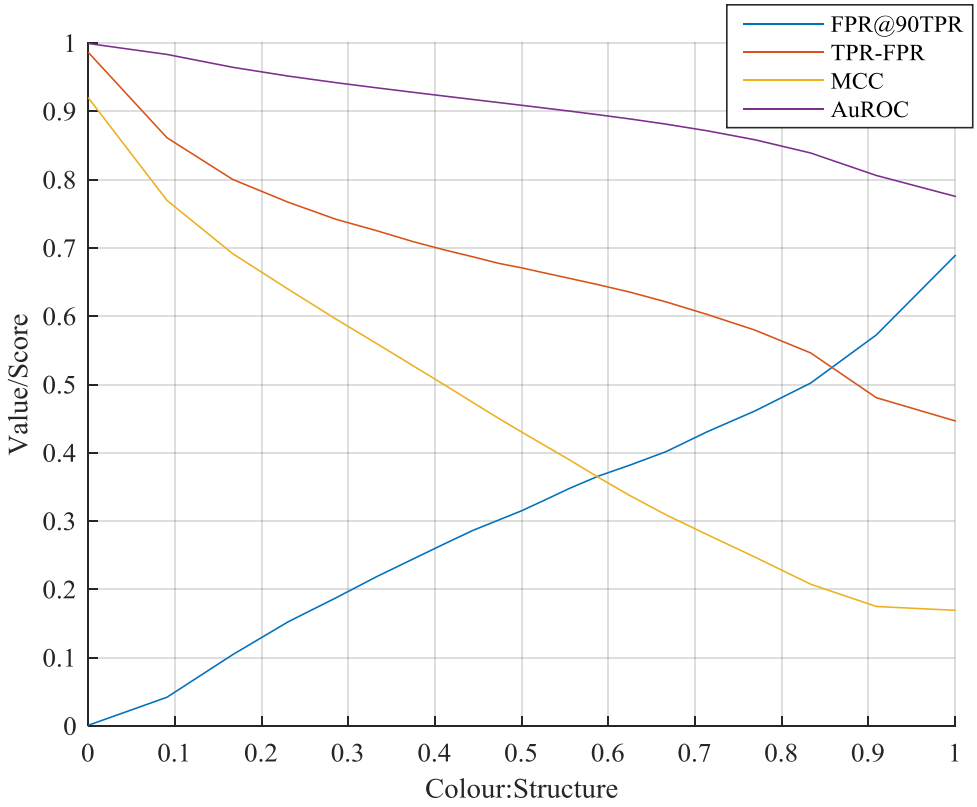


Figure 4-34 Dataset1 3DSAC - Graph of performance metrics

Figure 4-34 shows a graph of the performance metrics against the weighting used to create a differencing for 3DSAC. Here, the TPR-FPR, MCC and AuROC metrics start very high when maximum structural change weighting is defined, with a steady but significant decrease as colour is introduced in the place of structure. This is mirrored in the FPR@90%TPR where little false positive is seen at high structural weighting. This decreases as structural weighting is replaced with colour. Table 4-26 shows the corresponding data to this graph.

Table 4-26 - Dataset1 3DSAC - Table of performance metrics

Colour to Structure Ratio	FPR @ 90% TPR	TPR-FPR	MCC	AuROC
0.0000	0.0010	0.9866	0.9202	0.9995
0.0909	0.0421	0.8614	0.7696	0.9834
0.1667	0.1044	0.8004	0.6917	0.9646
0.2308	0.1526	0.7670	0.6396	0.9517
0.2857	0.1876	0.7420	0.5959	0.9421
0.3333	0.2191	0.7251	0.5596	0.9344
0.3750	0.2447	0.7092	0.5273	0.9279
0.4118	0.2672	0.6969	0.4989	0.9223
0.4444	0.2868	0.6868	0.4732	0.9173
0.4737	0.3017	0.6774	0.4503	0.9129
0.5000	0.3152	0.6709	0.4305	0.9089
0.5263	0.3307	0.6635	0.4110	0.9049
0.5556	0.3482	0.6554	0.3894	0.9004
0.5882	0.3658	0.6463	0.3643	0.8951
0.6250	0.3821	0.6353	0.3372	0.8889
0.6667	0.4019	0.6210	0.3091	0.8813
0.7143	0.4310	0.6025	0.2801	0.8715
0.7692	0.4614	0.5796	0.2472	0.8584
0.8333	0.5022	0.5463	0.2076	0.8391
0.9091	0.5727	0.4807	0.1750	0.8062
1.0000	0.6888	0.4468	0.1694	0.7756

4.4 Change Detection - Dataset 2

This section covers the results for the applications of each of the algorithms identified for change detection in dataset 2. Here both the statistical results and where appropriate visualisations are presented. An initial analysis of these results is provided in this results section, with further analysis and discussion in Chapter 5.

As a visual overview of detected change, Figure 4-35 shows each of the following descriptors from an equivalent view using a colour map to represent that change:

- Colour XYZ
- Euclidean
- FPFH
- 3DSAC
- SHOT
- SHOT Colour

The change introduced of the rock is most clearly visible in Figure 4-35d representing the 3DSAC descriptor. The changes are also viable in Figure 4-35a also but much less prevalent, and it is noted that the second object (the red strap) can also be detected. Figure 4-35b enables the visual detection of the rock object with more prevalence than with Colour XYZ but less so than 3DSAC. In Figure 4-35c the structural change of the rock can be detected amongst large amounts of other noise and artefacts. An equivalent view of this is also presented in Figure 4-36 with a greyscale mapping. Again, this is harder to interpret and the shapes of the changes harder to recognise. With this visualisation it is still possible to identify the change areas of the scene and it is noticed the structural change of the rock is most easily visually detected through FPFH and 3DSAC.

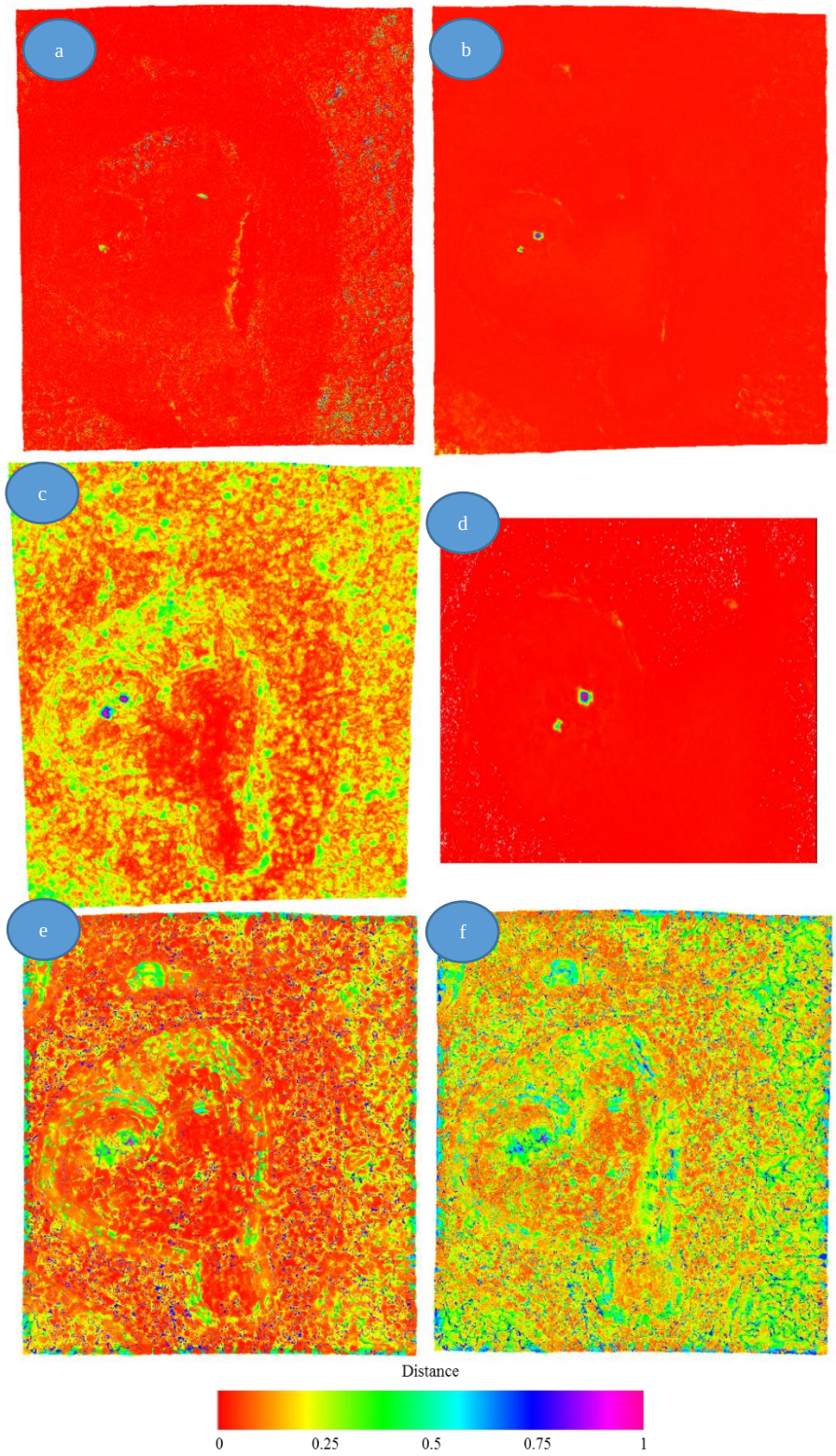


Figure 4-35 - Dataset2 Descriptors with Colour Map Visualisation.

a) Colour XYZ b) Euclidean c)FPFH d)3DSAC e)SHOT f) SHOT Colour

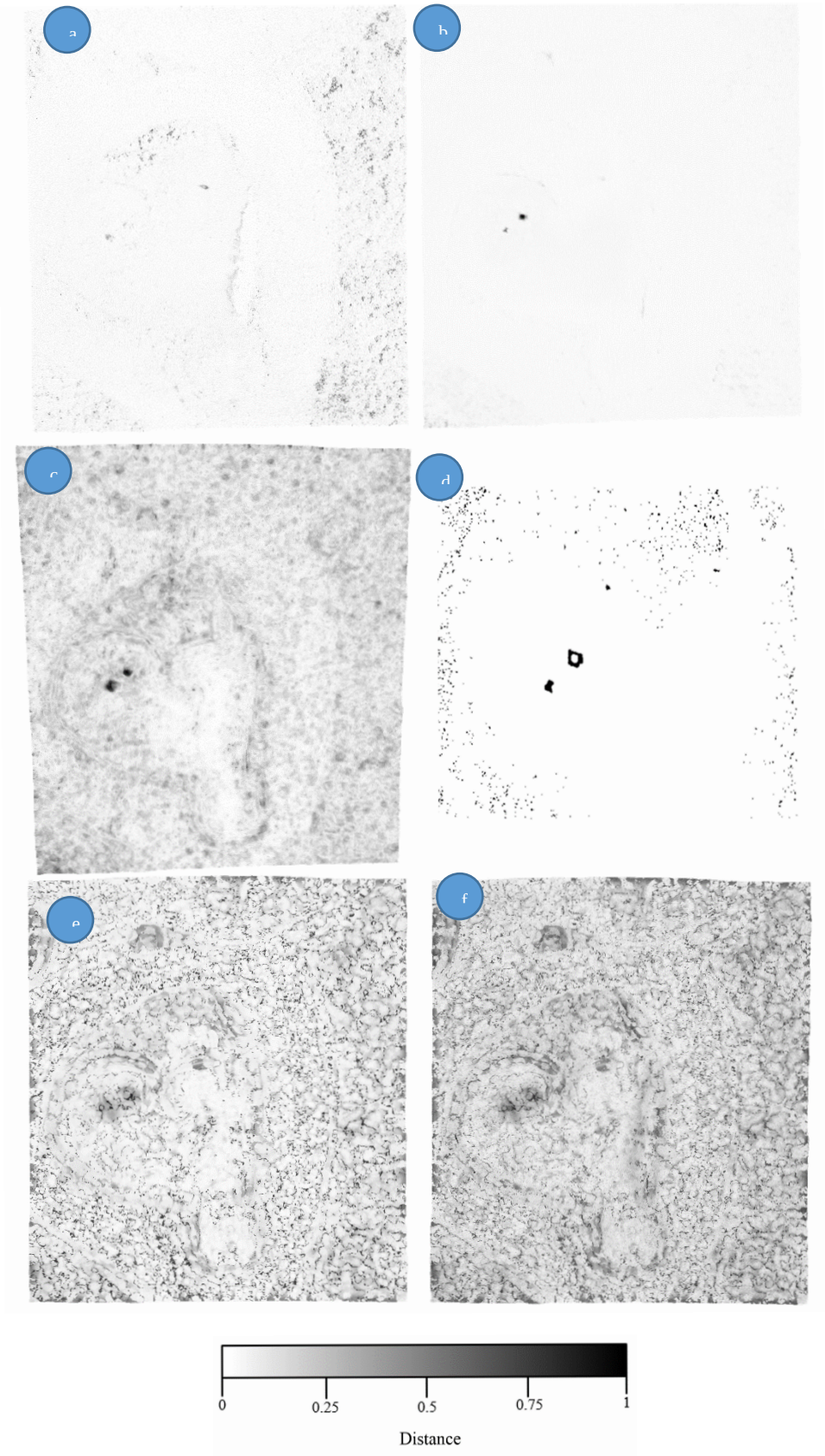


Figure 4-36 - Dataset2 Dataset1 Descriptors with Grayscale Map Visualisation.

a) Colour XYZ b) Euclidean c)FPFH d)3DSAC e)SHOT f) SHOT Colour

4.4.1 Euclidean

Following the overview, a more comprehensive result set for the Euclidean descriptor is given. Figure 4-37 presents a graph of the performance metrics against the number of neighbours used to create a mean differencing. The AuROC and TPR-FPR metrics start high with a low k number and gradually decrease in performance as more neighbours are introduced. The MCC metric follows a similar trend with a slightly steeper gradient and a lower starting position at low k numbers. The FPR@90%TPR start off very low and level until a k value of around 50 whereby it increases gradually to almost 1 by the time k is 1000. Table 4-27 shows the corresponding data to this graph.

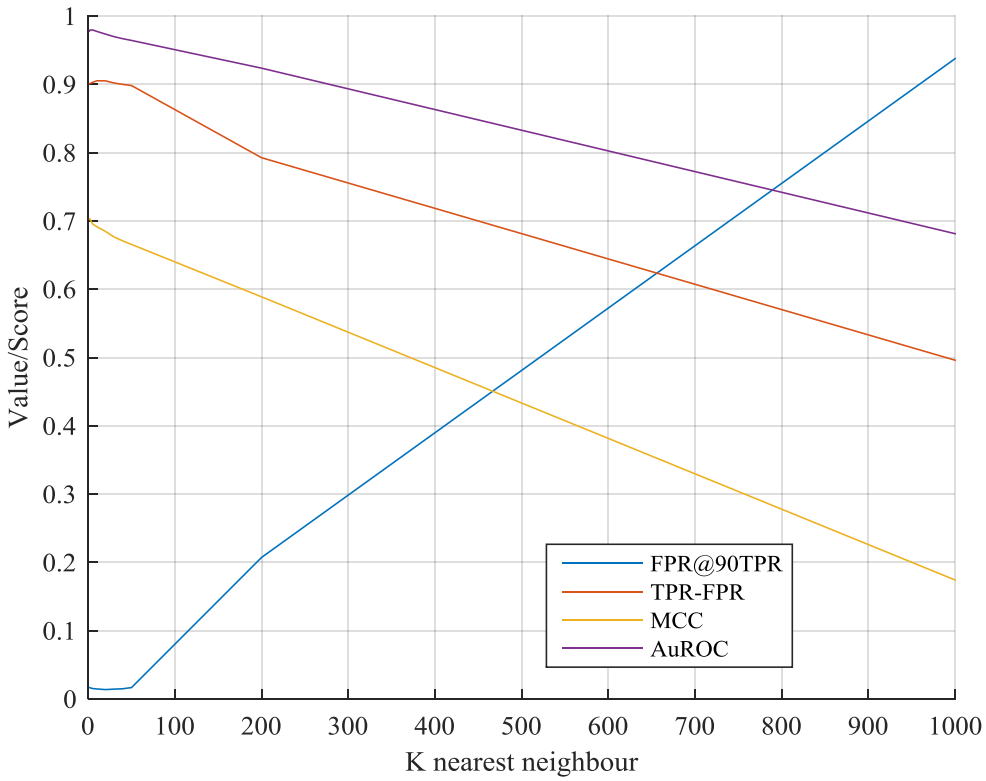


Figure 4-37 - Dataset2 Euclidean – Graph of performance metrics

Table 4-27- Dataset2 Euclidean - Table of performance metrics

K nearest neighbours	FP @ 90% TP (P_{fa} @ 90%P_d)	Max TP Rate - FP Rate	Max MCC	Max AuROC Curve
1	0.0173	0.9007	0.7022	0.9770
2	0.0166	0.9013	0.7032	0.9795
5	0.0154	0.9030	0.6957	0.9799
10	0.0148	0.9054	0.6918	0.9777
20	0.0139	0.9052	0.6848	0.9737
30	0.0145	0.9021	0.6766	0.9699
40	0.0151	0.9001	0.6708	0.9669
50	0.0168	0.8984	0.6658	0.9645
200	0.2074	0.7927	0.5890	0.9239
1000	0.9379	0.4962	0.1743	0.6815

4.4.2 Colour

This section looks at the results for using colour only as a differentiator between the two epochs of data across Dataset 2. Here the three descriptors LAB, RGB and XYZ are compared with the performance metrics for each of these listed. Table 4-28 shows the results for each of these against the chosen metrics. The XYZ colour space appears to offer the best performance numerically across all measurement techniques.

Table 4-28- Dataset2 Colour - Table of performance metrics

	FPR @ 90% TPR	TPR-FPR	MCC	AUC
LAB	0.6789	0.3977	0.0581	0.7559
RGB	0.6801	0.4714	0.1874	0.7908
XYZ	0.3318	0.6288	0.1940	0.9002

Figure 4-39 shows the point cloud visualisation of these results with the colour maps and corresponding measured differences in the key at the bottom of the figure. Fragments A, C and E show the full data set before thresholding. Fragments B, D and F show a threshold has been applied giving a subset of 130,000 points with the largest distances. Figure 4-39 show a grayscale representation of the XYZ change space. Here the structural change of the rock can be identified and also that of the colour change in the red strap.

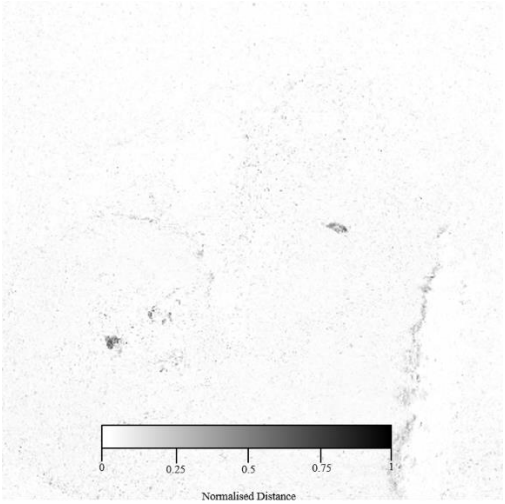


Figure 4-38 - Dataset2 XYZ Colour Difference Visualisation

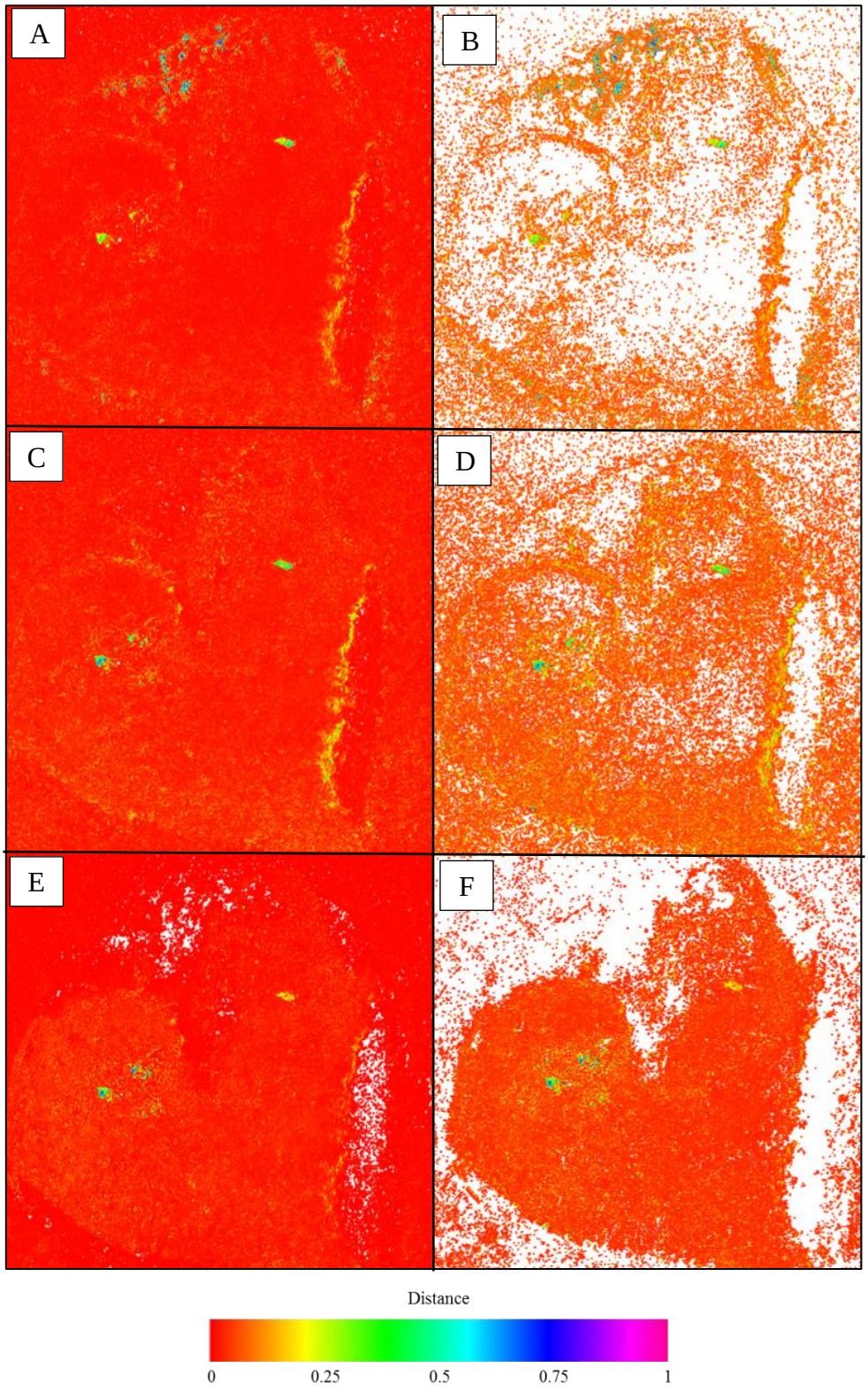


Figure 4-39 - Dataset2 Colour Point Clouds

A)CIE– No filter B) CIE – largest 130000 points C) RGB – No filter D)RGB – largest 130000 points E) XYZ – No filter F)XYZ – largest 130000 points

4.4.3 FPFH

This section looks at the results for using the FPFH descriptor as a differentiator between the two epochs of data across Dataset 2. Figure 4-40 shows a graph of the performance metrics against the number of neighbours used to create a mean differencing. Here, the TPR-FPR, MCC and AuROC all peak at around k values of 1000 with gradual declines thereafter. FPR@90TPR seems to mirror this until a k value of 5000 where it begins to decline suggesting false positive rate is falling. Table 4-29 shows the corresponding data to this graph.

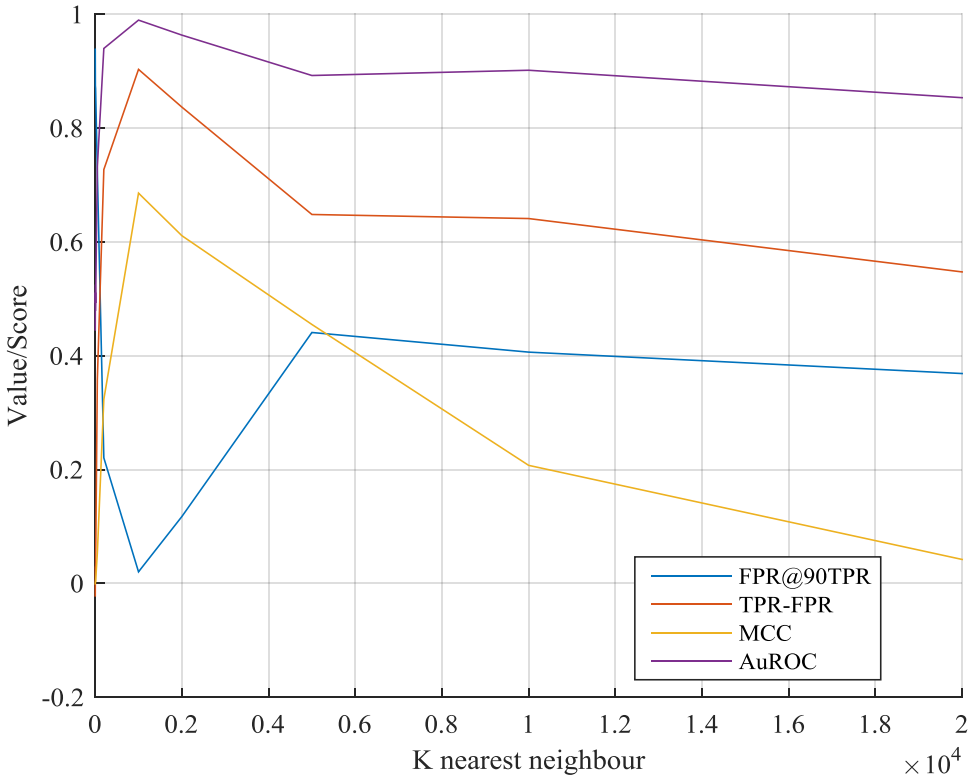


Figure 4-40 - Dataset2 FPFH – Graph of performance metrics

Table 4-29- Dataset2 FPFH - Table of performance metrics

K nearest neighbours	FP @ 90% TP (P_{fa} @ 90%P_d)	Max TP Rate - FP Rate	Max MCC	Max AuROC Curve
2	0.9401	-0.0232	-0.0020	0.4442
5	0.8748	0.0289	0.0024	0.5114
10	0.8697	0.0314	0.0033	0.4929
50	0.7244	0.3579	0.0532	0.7321
200	0.2207	0.7274	0.3237	0.9402
1000	0.0205	0.9034	0.6860	0.9900
2000	0.1176	0.8371	0.6112	0.9636
5000	0.4411	0.6484	0.4551	0.8927
10000	0.4066	0.6413	0.2076	0.9019
20000	0.3689	0.5475	0.0422	0.8536

Figure 4-41 below shows the FPFH descriptor point cloud visualisation with the corresponding measured differences in the key at the bottom of the figure. Here two values of *k* for the *k* nearest neighbours are displayed (200 and 1000 respectively). The structural change of the rock can clearly be distinguished in both images, with addition of the colour change object in the red strap being viable where k is 200.

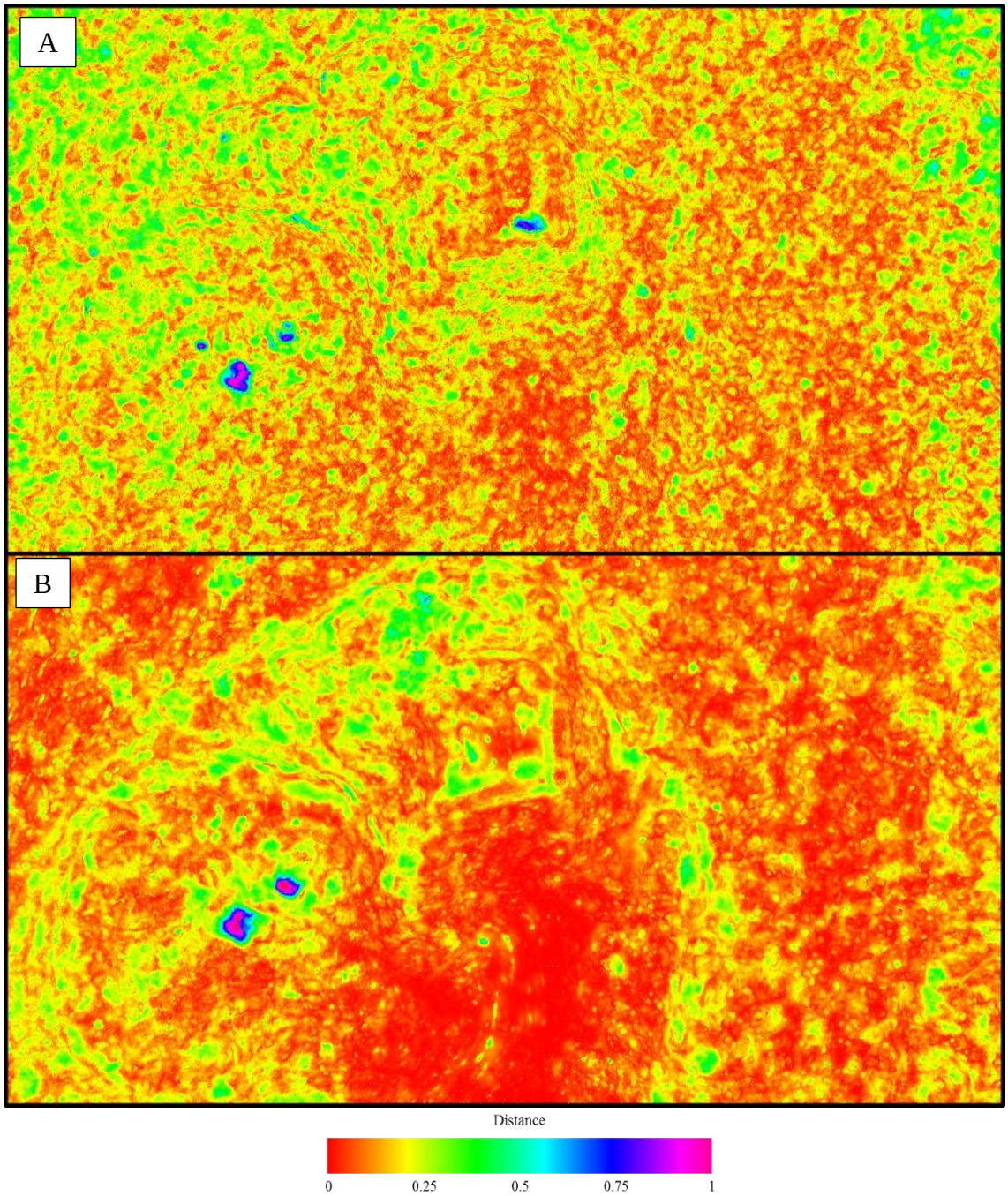


Figure 4-41 - Dataset2 FPFH A) k =200 B) k = 1000

4.4.4 SHOT

This section looks at the results for using the SHOT descriptor as a differentiator between the two epochs of data across Dataset 2. Figure 4-42 shows a graph of the TPR-FPR performance metric against the number of neighbours used to create a mean differencing across various radii. For lower radii the initial peak is seen when the k value is 20 and then decreases. The highest performance is seen at a radius of 0.2 and this decreases as the radius is decreased. Table 4-30 shows the corresponding data to this graph.

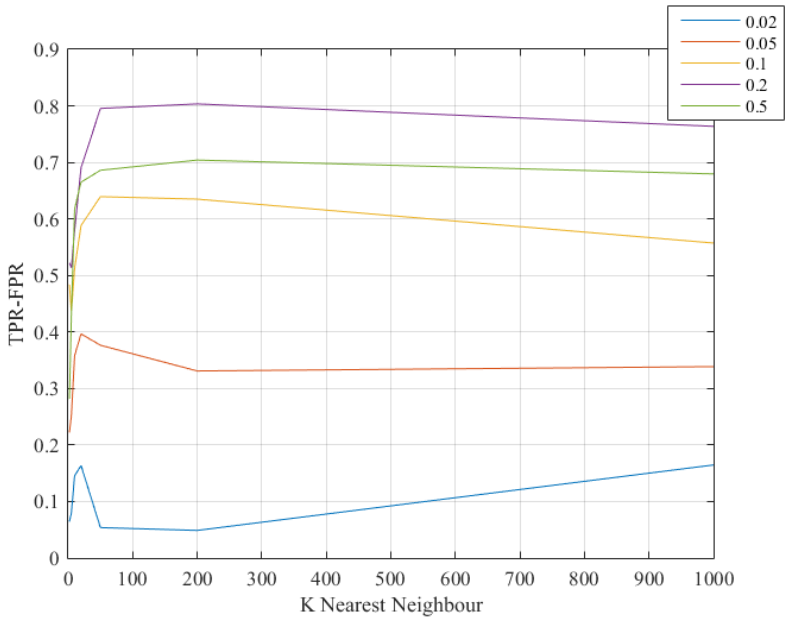


Figure 4-42 Dataset2 SHOT - Graph of TPR – FPR

Table 4-30- Dataset2 SHOT - Table of TPR - FPR metrics

TPR-FPR	Radius				
K nearest neighbour	0.02	0.05	0.1	0.2	0.5
2	0.0645	0.2222	0.4840	0.5224	0.2815
5	0.0794	0.2516	0.4354	0.5136	0.4308
10	0.1455	0.3580	0.5122	0.5792	0.6178
20	0.1637	0.3966	0.5883	0.6903	0.6650
50	0.0539	0.3765	0.6394	0.7955	0.6861
200	0.0489	0.3310	0.6350	0.8035	0.7041
1000	0.1647	0.3387	0.5573	0.7638	0.6796

Figure 4-43 shows a graph of the MCC performance metric against the number of neighbours used to create a mean differencing across various radii. Here again the initial peak is seen when the k value is 20 and then decreases fairly rapidly. Table 4-31 shows the corresponding data to this graph. Here it is evident again that maximal performance is found at a higher radius value and this is decreases as the radius is decreased.

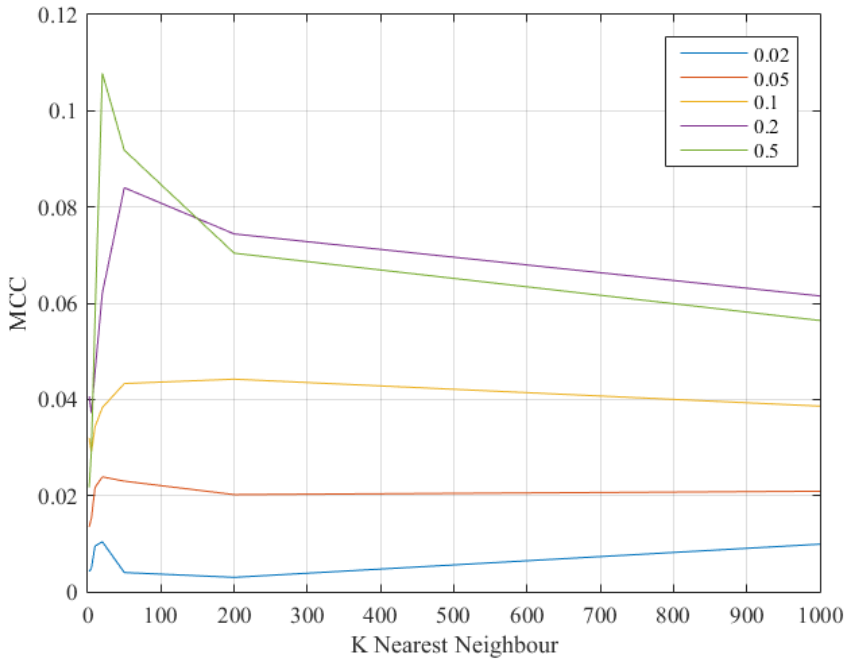


Figure 4-43 Dataset2 SHOT - Graph of MCC metrics

Table 4-31- Dataset2 SHOT - Table of MCC metrics

MCC Score	Radius				
K nearest neighbour	0.02	0.05	0.1	0.2	0.5
2	0.0042	0.0135	0.0320	0.0406	0.0217
5	0.0049	0.0152	0.0290	0.0372	0.0293
10	0.0095	0.0217	0.0343	0.0458	0.0571
20	0.0104	0.0239	0.0383	0.0621	0.1078
50	0.0040	0.0230	0.0433	0.0840	0.0918
200	0.0030	0.0202	0.0442	0.0744	0.0704
1000	0.0099	0.0209	0.0386	0.0615	0.0564

Figure 4-44 shows a graph of the AuROC performance metric against the number of neighbours used to create a mean differencing across various radii. Here again the initial peak is seen when the k value is 20 and then levels off in the case of the highest performer (radius 0.2). Other radius value seem to decrease in performance more rapidly as k values are increased. Table 4-32 shows the corresponding data to this graph. Here it is evident again that maximal performance is found at a higher radius value and this is decreases as the radius is decreased.

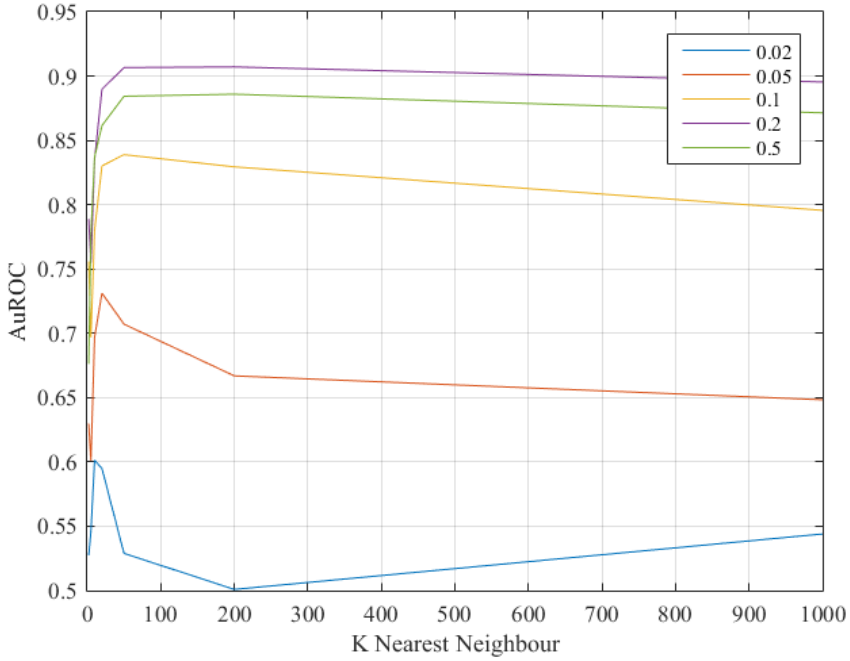


Figure 4-44 Dataset2 SHOT - Graph of AuROC metrics

Table 4-32- Dataset2 SHOT - Table of AuROC metrics

AuROC	Radius				
K nearest neighbour	0.02	0.05	0.1	0.2	0.5
2	0.5273	0.6299	0.7559	0.7889	0.6762
5	0.5457	0.6010	0.6967	0.7549	0.7769
10	0.6013	0.6969	0.7800	0.8350	0.8371
20	0.5949	0.7312	0.8301	0.8897	0.8614
50	0.5289	0.7071	0.8389	0.9066	0.8843
200	0.5009	0.6669	0.8294	0.9071	0.8859
1000	0.5440	0.6483	0.7956	0.8954	0.8714

Figure 4-45 shows a graph of the FPR@90%TRP performance metric against the number of neighbours used to create a mean differencing across various radii. Here again the initial peak of performance is seen when the k value is 20 and then levels off in the case of the highest performer (radius 0.2). Other radius value seem to decrease in performance more rapidly as k values are increased. Table 4-33 shows the corresponding data to this graph. Here it is evident again that maximal performance is found at a higher radius value and this is decreases as the radius is decreased.

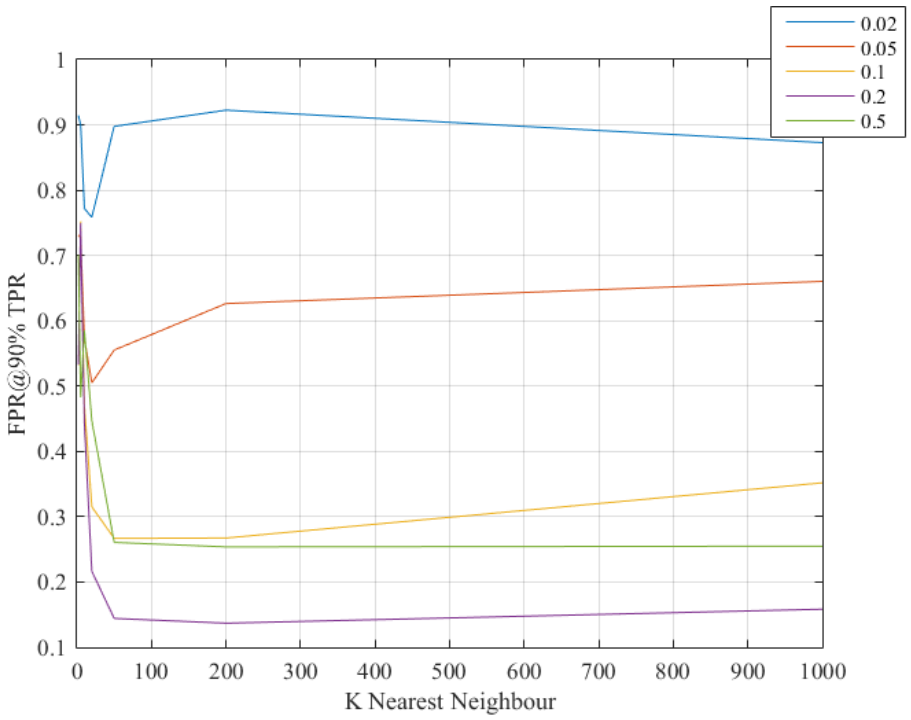


Figure 4-45 Dataset2 SHOT - Graph of FPR @ 90% TPR

Table 4-33- Dataset2 SHOT - - Table of FPR @ 90% TPR

FPR @ 90% TPR	Radius				
K nearest neighbour	0.02	0.05	0.1	0.2	0.5
2	0.9145	0.7313	0.5322	0.5328	0.7011
5	0.9006	0.7266	0.7524	0.7494	0.4828
10	0.7721	0.5736	0.4699	0.4434	0.5853
20	0.7584	0.5050	0.3151	0.2164	0.4468
50	0.8977	0.5552	0.2668	0.1443	0.2607
200	0.9225	0.6264	0.2674	0.1370	0.2538
1000	0.8728	0.6604	0.3519	0.1584	0.2549

4.4.5 SHOT Colour

This section looks at the results for using the SHOT Colour descriptor as a differentiator between the two epochs of data across Dataset 2. Figure 4-46 shows a graph of the FPR@90%TPR performance metric against the number of neighbours used to create a mean differencing with varied radii. Here, instability a low k numbers can be seen and then a levelling effect through the rest of the data. The best performance appears to be from a radius of 0.5 and 0.2 with decreasing performance as the radius is decreased. Table 4-34 shows the corresponding data to this graph.

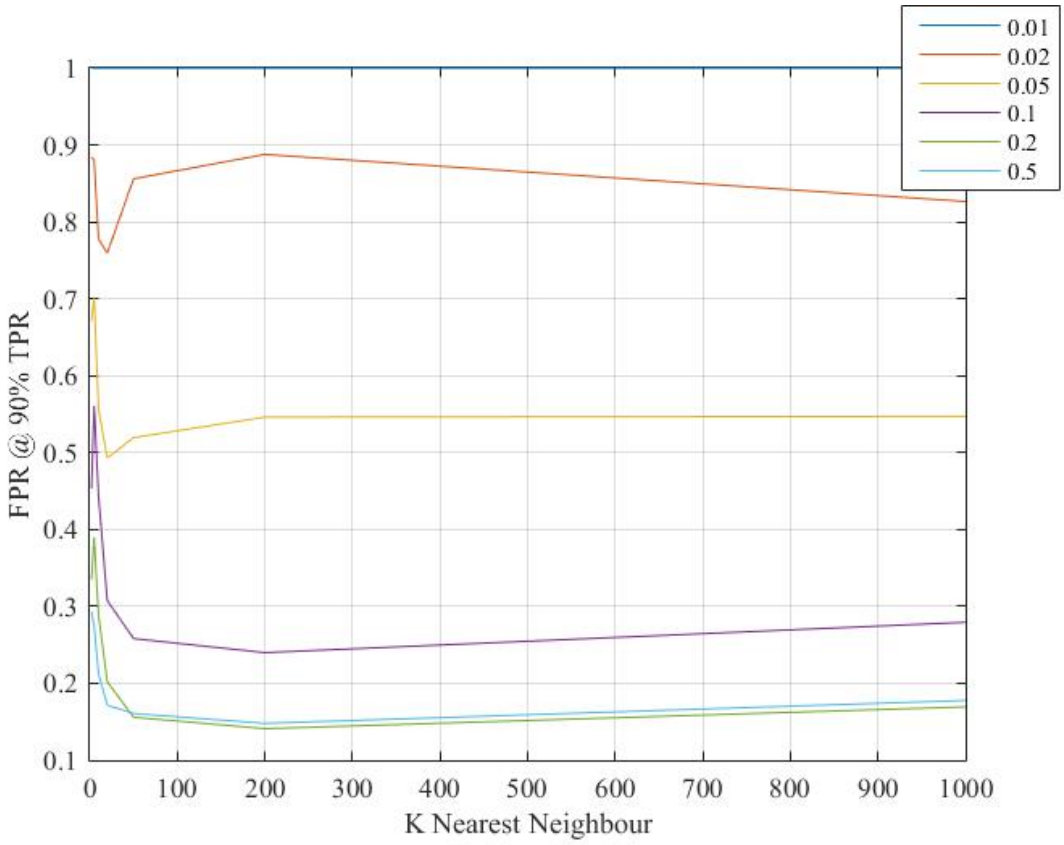


Figure 4-46 Dataset2 SHOT Colour - Graph of FPR @ 90% TPR

Table 4-34 – Dataset2 SHOT Colour - Table of FPR @ 90% TPR

FPR @ 90% TPR	Radius					
K Nearest Neighbour	0.01	0.02	0.05	0.1	0.2	0.5
2	1	0.8838	0.6711	0.4532	0.3347	0.2927
5	1	0.8807	0.7021	0.5611	0.3903	0.2746
10	1	0.7776	0.5567	0.4433	0.2886	0.2127
20	1	0.7595	0.4928	0.3075	0.2019	0.1716
50	1	0.8560	0.5194	0.2581	0.1558	0.1607
200	1	0.8877	0.5461	0.2399	0.1410	0.1479
1000	1	0.8265	0.5471	0.2793	0.1693	0.1777

Figure 4-47 shows a graph of the TPR-FPR performance metric against the number of neighbours used to create a mean differencing with varied radii. Here, instability a low k numbers can be seen and then a levelling effect through the rest of the data. The best performance appears to be from a radius of 0.5 and 0.2 with decreasing performance as the radius is decreased. Table 4-34 shows the corresponding data to this graph.

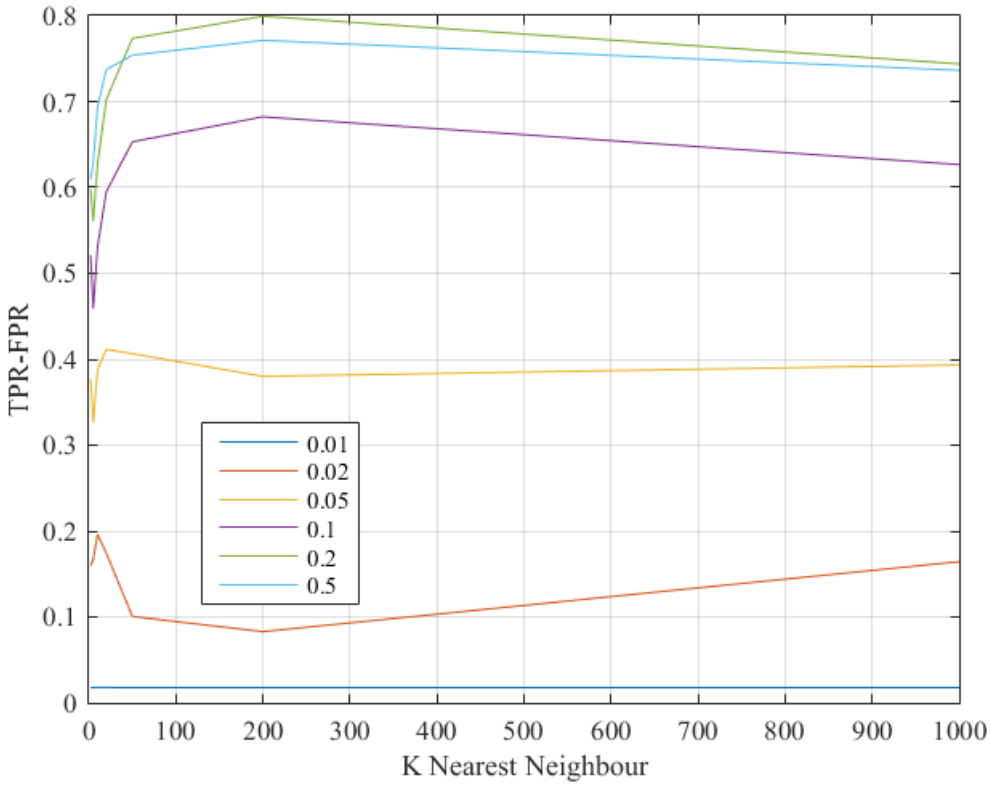


Figure 4-47 Dataset2 SHOT Colour- Graph of TPR-FPR metrics

Table 4-35- Dataset2 SHOT Colour- Table of TPR-FPR metrics

	Radius					
TPR-FPR	0.01	0.02	0.05	0.1	0.2	0.5
2	0.0179	0.1596	0.3766	0.5210	0.5983	0.6096
5	0.0178	0.1668	0.3263	0.4589	0.5608	0.6286
10	0.0179	0.1960	0.3878	0.5307	0.6277	0.6934
20	0.0179	0.1745	0.4116	0.5948	0.7022	0.7372
50	0.0178	0.1005	0.4065	0.6531	0.7735	0.7539
200	0.0178	0.0829	0.3801	0.6822	0.7993	0.7712
1000	0.0178	0.1644	0.3933	0.6266	0.7438	0.7363

Figure 4-48 shows a graph of the MCC performance metric against the number of neighbours used to create a mean differencing with varied radii. Here, instability a low k numbers can be seen and then a levelling effect and steady decline in performance. The best performance appears to be from a radius of 0.2 with decreasing performance as the radius is decreased. Table 4-46 shows the corresponding data to this graph.

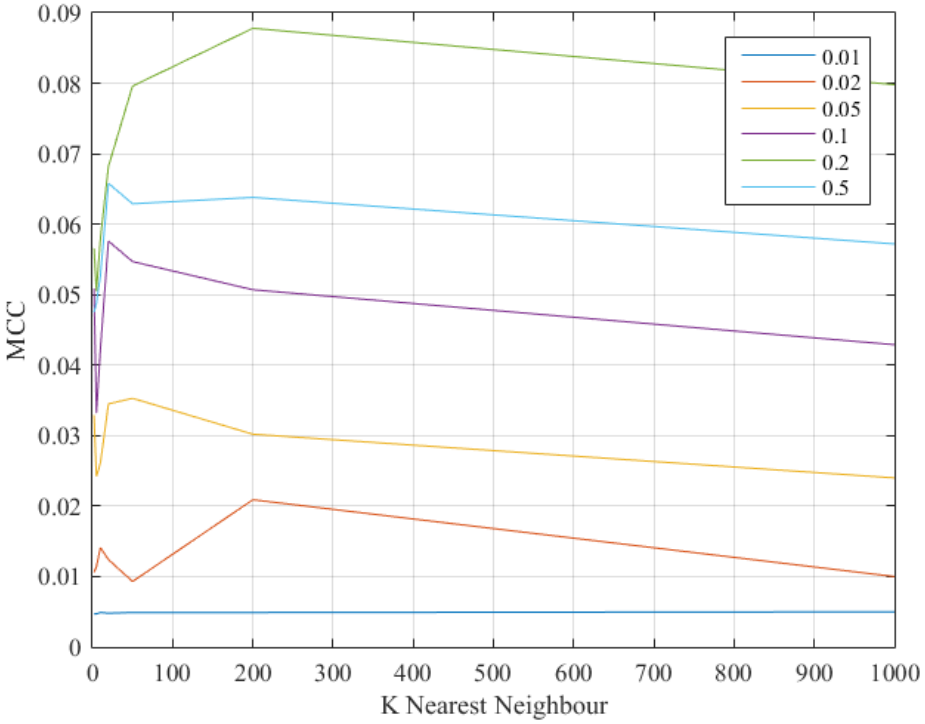


Figure 4-48 Dataset2 SHOT Colour- Graph of MCC metrics

Table 4-36- Dataset2 SHOT Colour- Table of MCC metrics

	Radius					
MCC	0.01	0.02	0.05	0.1	0.2	0.5
2	0.0047	0.0106	0.0329	0.0509	0.0566	0.0475
5	0.0047	0.0114	0.0242	0.0332	0.0505	0.0488
10	0.0049	0.0141	0.0261	0.0423	0.0580	0.0527
20	0.0048	0.0124	0.0345	0.0576	0.0682	0.0658
50	0.0049	0.0093	0.0353	0.0547	0.0796	0.0629
200	0.0049	0.0209	0.0302	0.0507	0.0878	0.0638
1000	0.0050	0.0100	0.0240	0.0429	0.0798	0.0572

Figure 4-49 shows a graph of the AuROC performance metric against the number of neighbours used to create a mean differencing with varied radii. Here, instability a low k numbers can be seen and then a levelling effect and steady decline in performance. The best performance appears to be from a radius of 0.5 and 0.2 with decreasing performance as the radius is decreased. Table 4-37 shows the corresponding data to this graph.

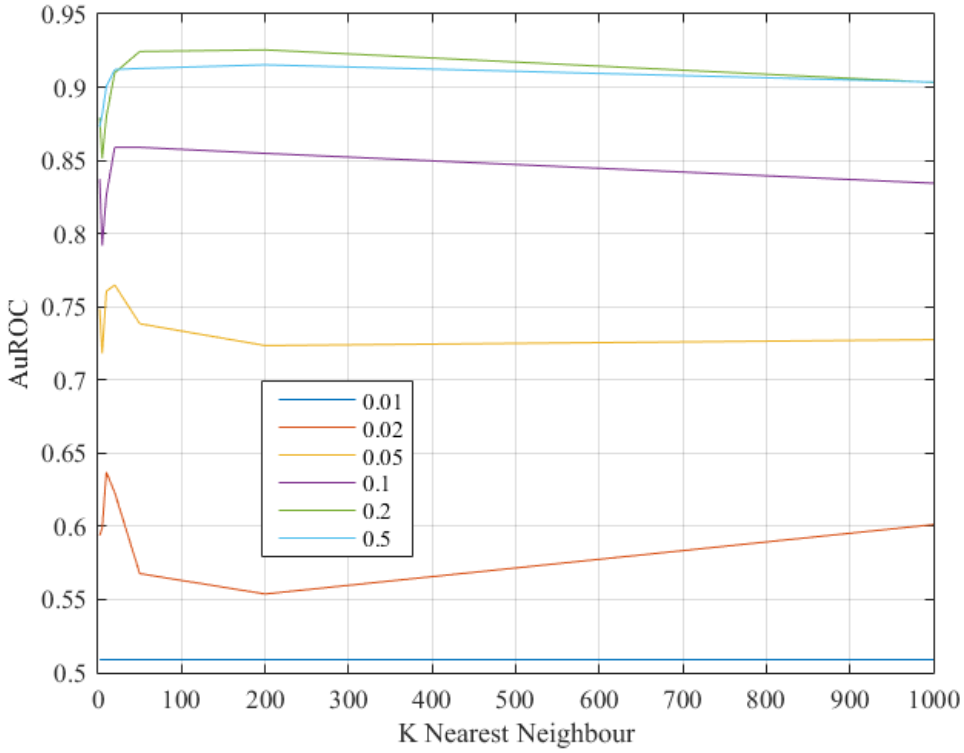


Figure 4-49 Dataset2 SHOT Colour- Graph of AuROC metrics

Table 4-37- Dataset2 SHOT Colour- Table of AuROC metrics

AuROC	0.01	0.02	0.05	0.1	0.2	0.5
2	0.5089	0.5937	0.7485	0.8372	0.8791	0.8730
5	0.5089	0.5986	0.7182	0.7918	0.8513	0.8812
10	0.5089	0.6370	0.7606	0.8262	0.8799	0.9005
20	0.5089	0.6231	0.7649	0.8589	0.9098	0.9121
50	0.5089	0.5677	0.7384	0.8589	0.9244	0.9128
200	0.5089	0.5538	0.7235	0.8548	0.9255	0.9153
1000	0.5089	0.6012	0.7275	0.8344	0.9033	0.9035

4.4.6 3DSAC

This section looks at the results for using the 3DSAC descriptor as a differentiator between the two epochs of data across Dataset 2. Figure 4-51 visualises the change detection results across varied weighting of colour to structure. Figure 4-51A shows 100% structural weighting and the structural changes of the rock can be seen very clearly with little background noise surrounding it. The colour change in the red strap is much harder to recognise here. As colour weighting is increased through the series of images, it become much easier to distinguish. This has the adverse effect that the rock becomes harder to detect. By analysis of Figure 4-51F where no structural changes are weighted in to the calculation, the rock is almost undetectable.

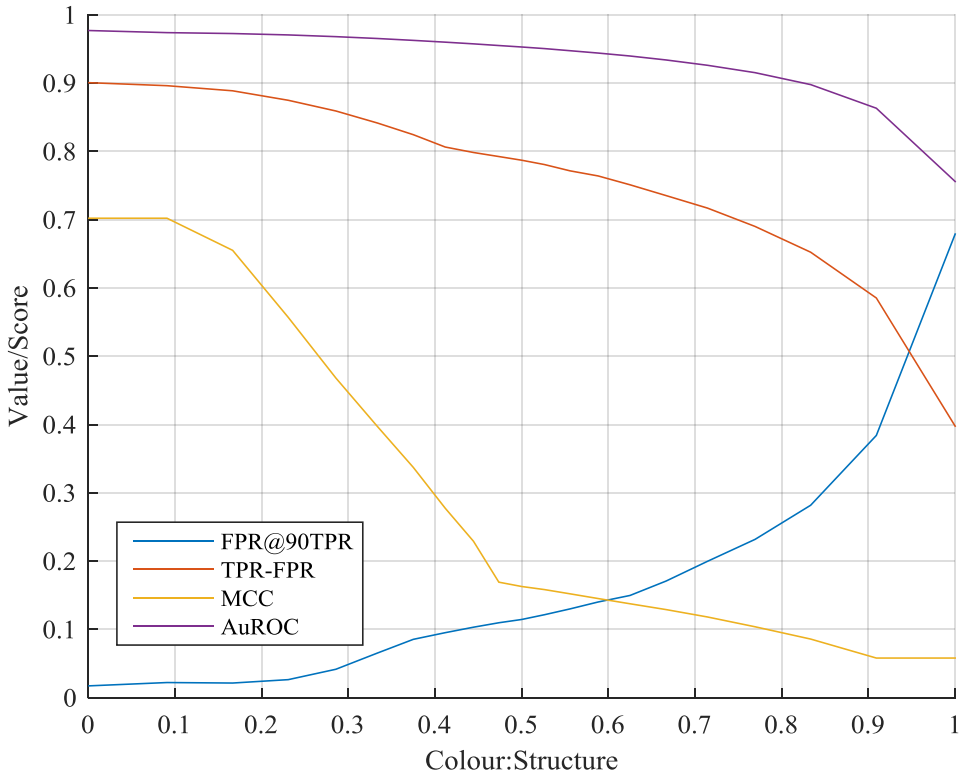


Figure 4-50 - Dataset2 3DSAC Performance Metrics

Figure 4-50 shows a graph of the performance metrics against the weighting used to create a differencing for 3DSAC. Here, the TPR-FPR, MCC and AuROC metrics start very high when maximum structural change weighting is defined, with a steady but significant decrease as colour is introduced in the place of structure. The MCC stands out a little with

a slightly varied performance of which decreases rapidly as more colour in introduced. This levels out around the midpoint between structure and colour. FPR@90TPR appears to mirror the TPR-FPR metric. Table 4-38 shows the corresponding data to this graph.

Table 4-38 - Dataset2 3DSAC Performance Statistics

Colour : Structure Ratio	P_{fa}	TPR-FPR	MCC	AuROC
0.0000	0.0173	0.9007	0.7022	0.9770
0.0909	0.0222	0.8962	0.7022	0.9739
0.1667	0.0215	0.8887	0.6551	0.9726
0.2308	0.0265	0.8749	0.5570	0.9706
0.2857	0.0417	0.8592	0.4681	0.9680
0.3333	0.0653	0.8417	0.3976	0.9654
0.3750	0.0855	0.8245	0.3373	0.9626
0.4118	0.0952	0.8064	0.2776	0.9600
0.4444	0.1033	0.7984	0.2291	0.9575
0.4737	0.1098	0.7925	0.1692	0.9550
0.5000	0.1146	0.7871	0.1629	0.9529
0.5263	0.1215	0.7807	0.1583	0.9506
0.5556	0.1300	0.7716	0.1522	0.9475
0.5882	0.1402	0.7641	0.1452	0.9440
0.6250	0.1497	0.7511	0.1376	0.9397
0.6667	0.1709	0.7353	0.1290	0.9339
0.7143	0.1997	0.7170	0.1182	0.9261
0.7692	0.2319	0.6901	0.1039	0.9153
0.8333	0.2818	0.6524	0.0859	0.8978
0.9091	0.3840	0.5853	0.0581	0.8632
1.0000	0.6789	0.3977	0.0581	0.7559

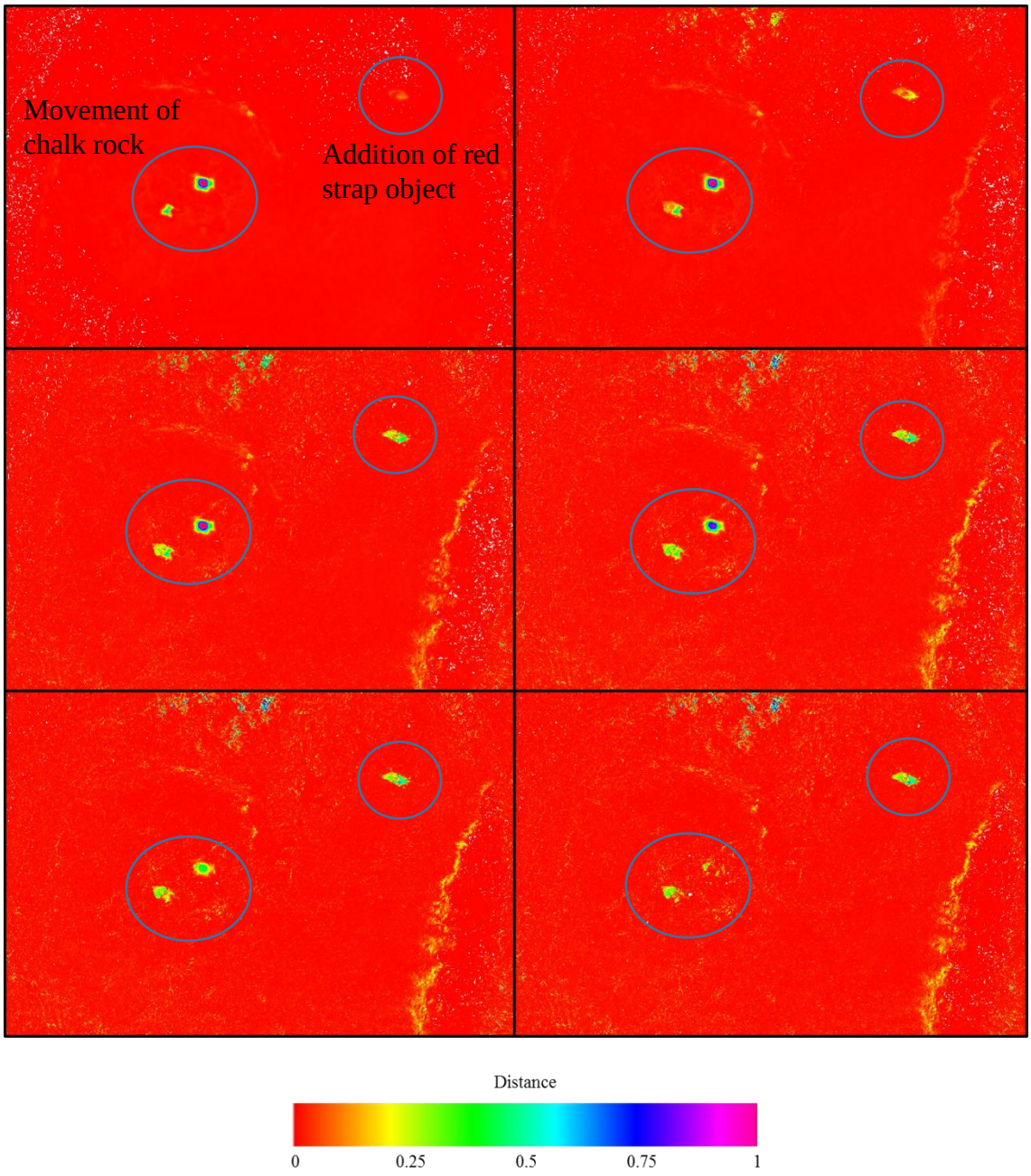


Figure 4-51 - Dataset2 3DSAC Change Visualisation

A) 100% Structure 0% Colour B) 100% Structure 40% Colour C) 100% Structure 80% Colour D) 80% Structure 100% Colour E) 40% Structure 100% Colour F) 0% Structure 100% Colour

4.5 Dataset 3

This section covers the results for the applications of each of the algorithms identified for change detection in dataset 3. Here both the statistical results and where appropriate visualisations are presented. An initial analysis of these results is provided in this results section, with further analysis and discussion in Chapter 5.

As a visual overview of detected change, Figure 4-52 shows each of the following descriptors from an equivalent view using a colour map to represent that change:

- Colour XYZ
- Euclidean
- FPFH
- 3DSAC
- SHOT
- SHOT Colour

The change introduced is most clearly visible in Figure 4-52b and Figure 4-52d representing the Euclidean and 3DSAC descriptors respectively. The changes are also viable in Figure 4-52a also but less prevalent. Figure 4-52c enables the visual detection of the object with more prevalence than with but shows a much larger area of change. An equivalent view of this is also presented in

Figure 4-53 with a greyscale mapping. Again, this is harder to interpret and the shapes of the changes harder to recognise. With this visualisation it is still possible to identify the change areas of the scene and it is noticed the change is most easily visually detected through Euclidean and 3DSAC techniques.

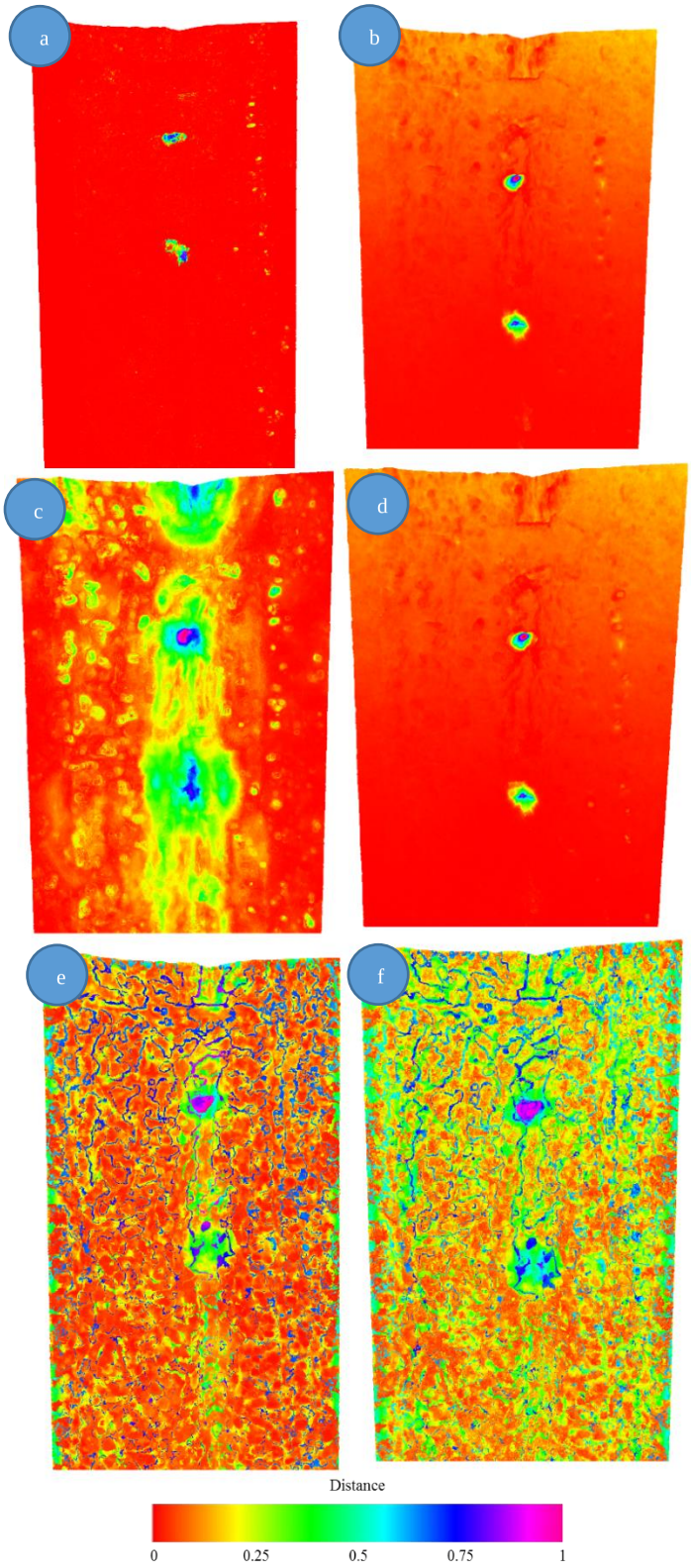


Figure 4-52 - Dataset 3 Descriptors with Colour Map Visualisation.

a) Colour XYZ b) Euclidean c)FPFH d)3DSAC e)SHOT f) SHOT Colour

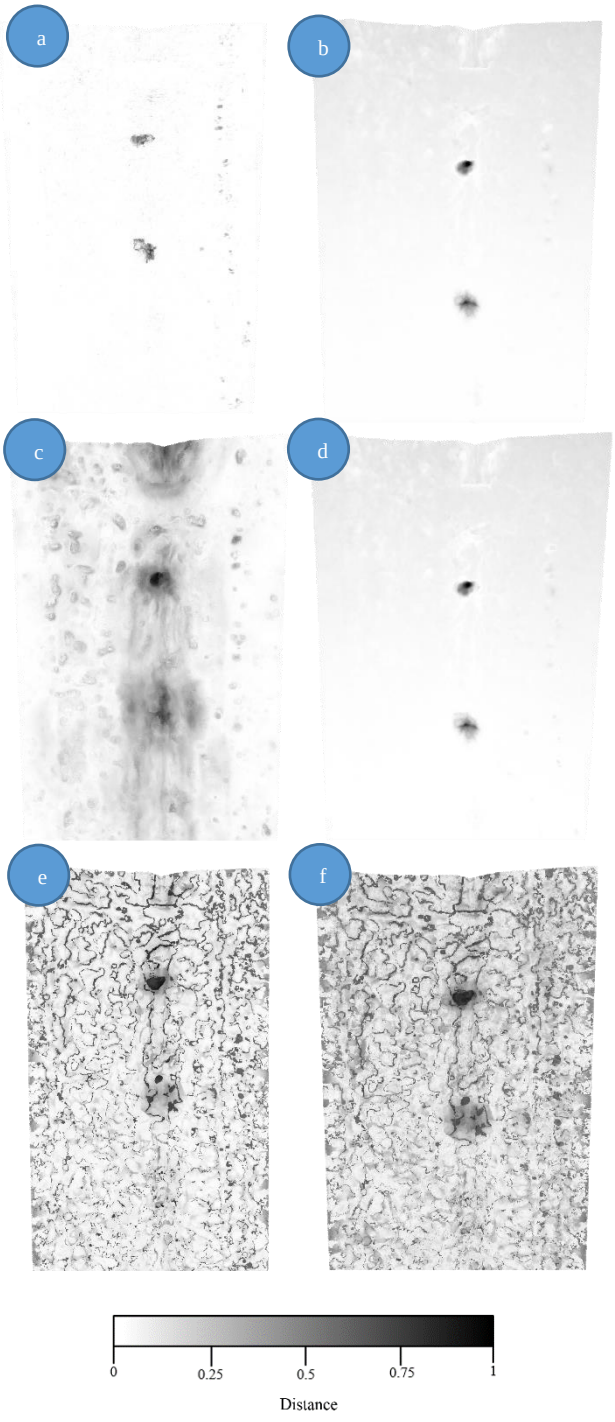


Figure 4-53 - Dataset3 Descriptors with Grayscale Map Visualisation.

a) Colour XYZ b) Euclidean c)FPFH d)3DSAC e)SHOT f) SHOT Colour

4.5.1 Euclidean

Following the overview, a more comprehensive result set for the Euclidean descriptor is given. Figure 4-54 presents a graph of the performance metrics against the number of neighbours used to create a mean differencing. The AuROC and TPR-FPR metrics start high with a low k number and very slowly gradually decrease in performance as more neighbours are introduced. The MCC metric follows a similar trend with a slightly steeper gradient and a lower starting position at low k numbers and more rapidly declines after a k value of 200. The FPR@90%TPR start off very low and mains this throughout. Table 4-39 shows the corresponding data to this graph.

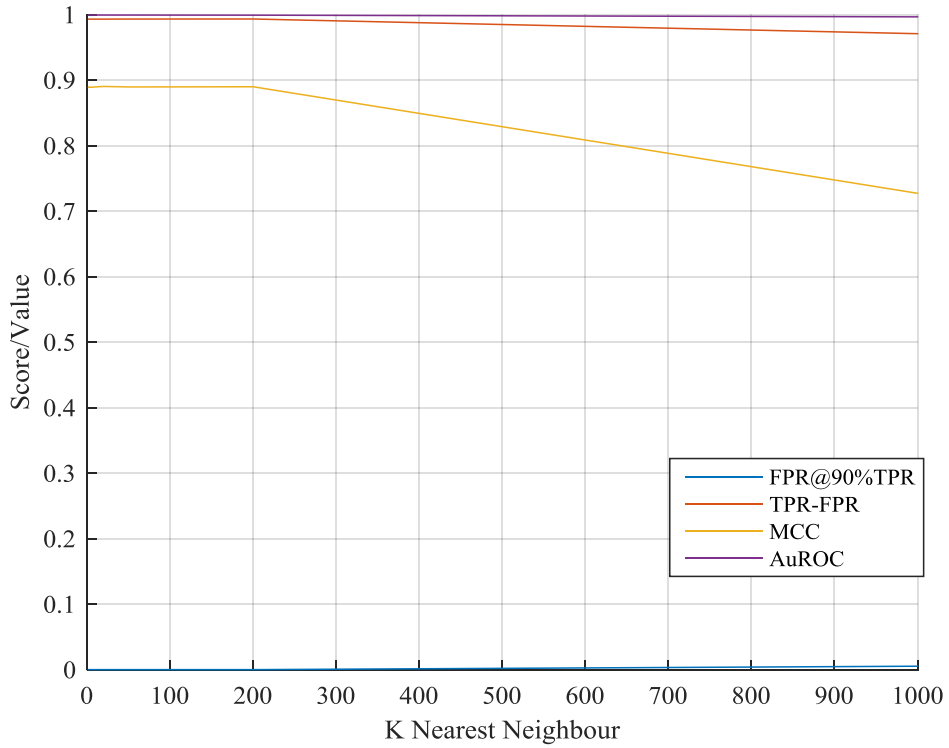


Figure 4-54 - Dataset 3 Euclidean - Graph of performance metrics

Table 4-39 - Dataset3 Euclidean - Table of performance metrics

K nearest neighbours	FPR @ 90% TPR (P_{fa} @ 90%P_d)	Max TP Rate - FP Rate	Max MCC	Max AuROC Curve
1	0.0002	0.9932	0.8899	0.9995
2	0.0002	0.9932	0.8893	0.9995
5	0.0002	0.9933	0.8893	0.9995
10	0.0002	0.9932	0.8898	0.9995
20	0.0002	0.9933	0.8905	0.9995
30	0.0002	0.9933	0.8902	0.9995
40	0.0002	0.9934	0.8901	0.9995
50	0.0002	0.9934	0.8897	0.9995
200	0.0002	0.9935	0.8900	0.9993
1000	0.0054	0.9710	0.7274	0.9967

4.5.2 Colour

This section looks at the results for using colour only as a differentiator between the two epochs of data across Dataset 3. Here the three descriptors LAB, RGB and XYZ are compared with the performance metrics for each of these listed. Table 4-40 shows the results for each of these against the chosen metrics. The LAB colour space appears to offer the best performance numerically for TPR-FPR, AuROC and FPR@90%TPR. With XYZ offering the best performance using the MCC metric.

Table 4-40 Dataset3 - Colour performance metrics

K nearest neighbours	FPR @ 90% TPR (P_{fa} @ 90%P_d)	Max TP Rate - FP Rate	Max MCC	Max AuROC Curve
LAB	0.0050	0.9341	0.5256	0.9931
XYZ	0.0066	0.9142	0.5544	0.9872
RGB	0.0061	0.9008	0.5523	0.9804

4.5.3 FPFH

This section looks at the results for using the FPFH descriptor as a differentiator between the two epochs of data across Dataset 3. Figure 4-55 shows a graph of the performance metrics against the number of neighbours used to create a mean differencing. Here, the TPR-FPR and AuROC rapidly increase but start to level off around at k values of 5000. FPR@90TPR seems to remain flat throughout. The MCC metric offers the largest discriminator with a rapid increase and levelling off at 5000 but then a steady increase again following k values of 10000. Table 4-41 shows the corresponding data to this graph.

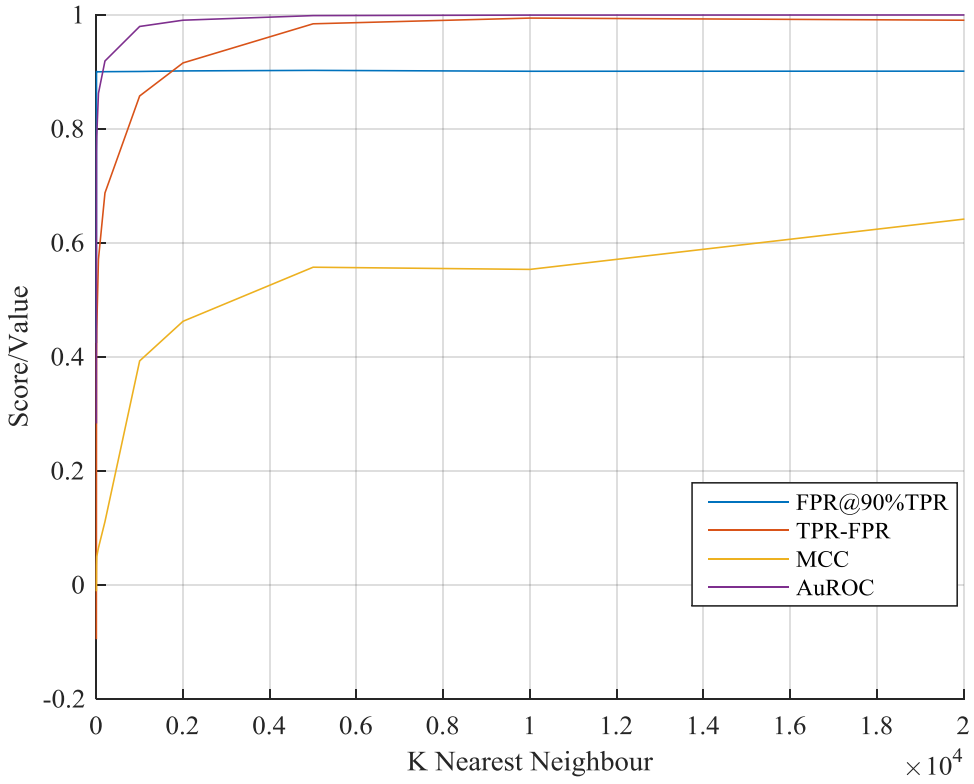


Figure 4-55- Dataset3 FPFH - Graph of performance metrics

Table 4-41 - Dataset3 FPFH - Table of performance metrics

K nearest neighbours	FPR @ 90% TPR (P_{fa} @ 90%P_d)	Max TP Rate - FP Rate	Max MCC	Max AuROC Curve
2	0.9020	-0.0958	-0.0117	0.2827
5	0.9011	0.3984	0.0329	0.7595
10	0.9007	0.4480	0.0491	0.7860
50	0.9005	0.5709	0.0645	0.8619
200	0.9001	0.6870	0.1104	0.9185
1000	0.8999	0.8571	0.3924	0.9791
2000	0.8999	0.9151	0.4619	0.9900
5000	0.8997	0.9837	0.5567	0.9981
10000	0.8995	0.9937	0.5529	0.9990
20000	0.4238	0.9900	0.6410	0.9992

4.5.4 SHOT

This section looks at the results for using the SHOT descriptor as a differentiator between the two epochs of data across Dataset 3. Figure 4-56 shows a graph of the TPR-FPR performance metric against the number of neighbours used to create a mean differencing across various radii. For lower radii the initial peak is seen when the k value is 5 and then decreases. The highest performance is seen at a radius of 0.5 and this decreases as the radius is decreased. Table 4-42 shows the corresponding data to this graph.

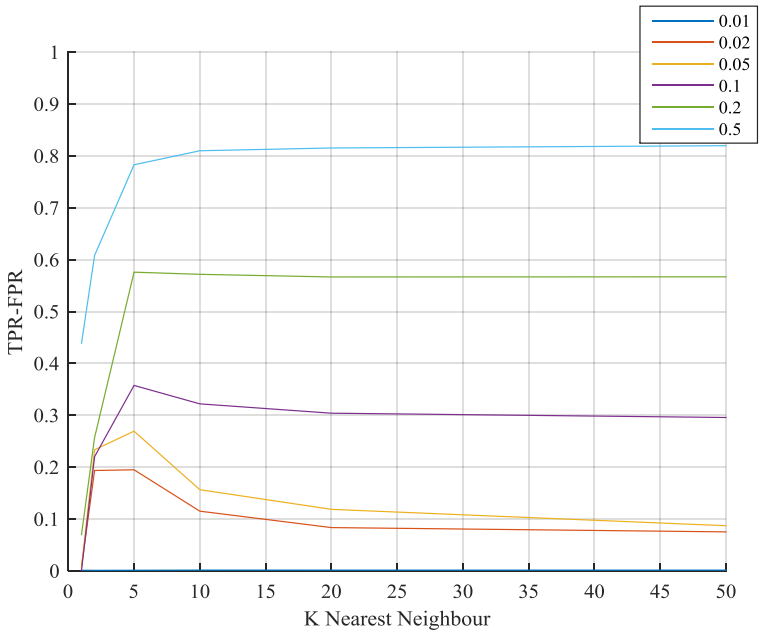


Figure 4-56- Dataset3 SHOT - Graph of TPR – FPR performance metrics

Table 4-42- Dataset3 SHOT - Table of TPR - FPR performance metric

TPR - FPR	Radius					
K nearest neighbour	0.01	0.02	0.05	0.1	0.2	0.5
1	0.0014	0.0070	0.0071	0.0001	0.0690	0.4380
2	0.0012	0.1937	0.2337	0.2204	0.2572	0.6084
5	0.0014	0.1951	0.2695	0.3577	0.5760	0.7830
10	0.0017	0.1153	0.1567	0.3222	0.5719	0.8101
20	0.0016	0.0837	0.1188	0.3042	0.5668	0.8154
50	0.0016	0.0754	0.0873	0.2959	0.5671	0.8198

Figure 4-57 shows a graph of the MCC performance metric against the number of neighbours used to create a mean differencing across various radii. For lower radii the initial peak is seen when the k value is 5 and then decreases. The highest performance is seen at a radius of 0.5 of which is substantially higher performing and this decreases as the radius is decreased. Table 4-43 shows the corresponding data to this graph.

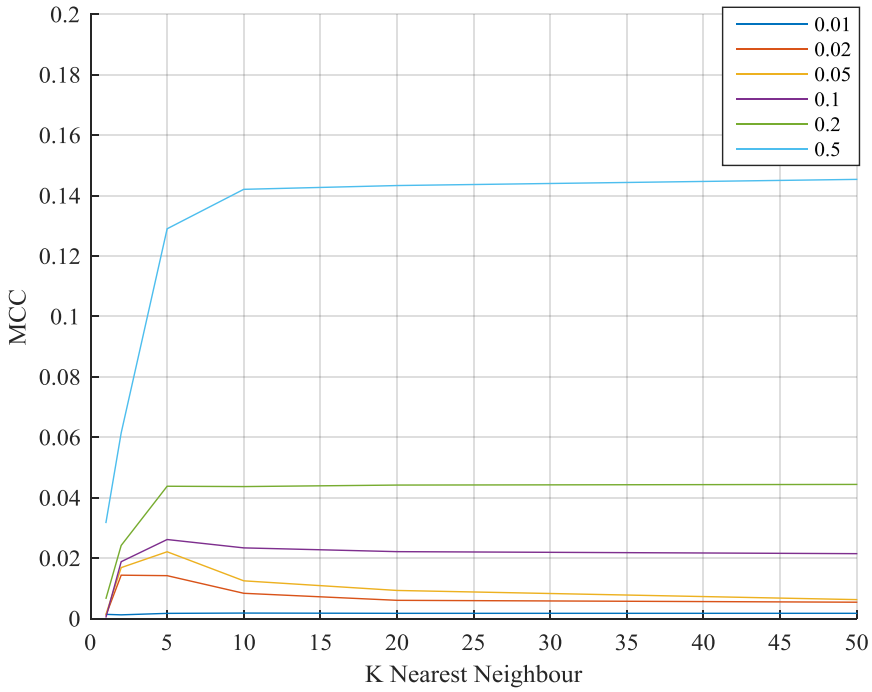


Figure 4-57- Dataset3 SHOT - Graph of MCC performance metric

Table 4-43 - Dataset3 SHOT - Table of MCC performance metric

MCC	Radius					
K Nearest Neighbour	0.01	0.02	0.05	0.1	0.2	0.5
1	0.0014	0.0010	0.0013	0.0004	0.0065	0.0316
2	0.0013	0.0144	0.0169	0.0188	0.0242	0.0615
5	0.0017	0.0142	0.0221	0.0262	0.0438	0.1290
10	0.0018	0.0084	0.0125	0.0234	0.0437	0.1421
20	0.0018	0.0061	0.0093	0.0222	0.0442	0.1433
50	0.0018	0.0054	0.0063	0.0215	0.0444	0.1454

Figure 4-58 shows a graph of the AuROC performance metric against the number of neighbours used to create a mean differencing across various radii. For lower radii the initial peak is seen when the k value is 5 and then decreases. The highest performance is seen at a radius of 0.5 of which does not appear to decrease a values of k are increased. Table 4-44 shows the corresponding data to this graph.

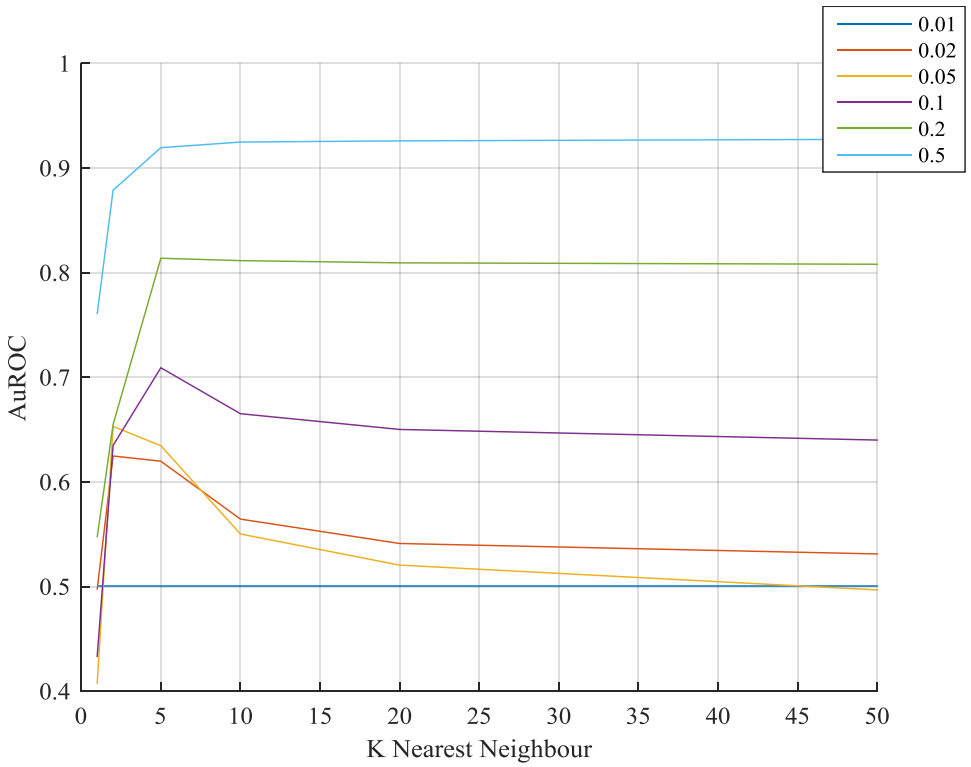


Figure 4-58- Dataset3 SHOT - Graph of AuROC performance metric

Table 4-44- Dataset3 SHOT - Table of AuROC performance metric

AuROC	Radius					
K Nearest Neighbour	0.01	0.02	0.05	0.1	0.2	0.5
1	0.5006	0.4970	0.4070	0.4327	0.5472	0.7603
2	0.5006	0.6248	0.6531	0.6350	0.6549	0.8789
5	0.5006	0.6199	0.6347	0.7092	0.8139	0.9195
10	0.5006	0.5646	0.5503	0.6653	0.8116	0.9249
20	0.5006	0.5412	0.5206	0.6502	0.8095	0.9261
50	0.5006	0.5312	0.4968	0.6401	0.8081	0.9275

Figure 4-59 shows a graph of the FPR@90%TPR performance metric against the number of neighbours used to create a mean differencing across various radii. For lower radii the initial peak performance is seen when the k value is 5 and then decreases. The highest performance is seen at a radius of 0.5 of which does not appear to decrease a values of k are increased. Table 4-45 shows the corresponding data to this graph.

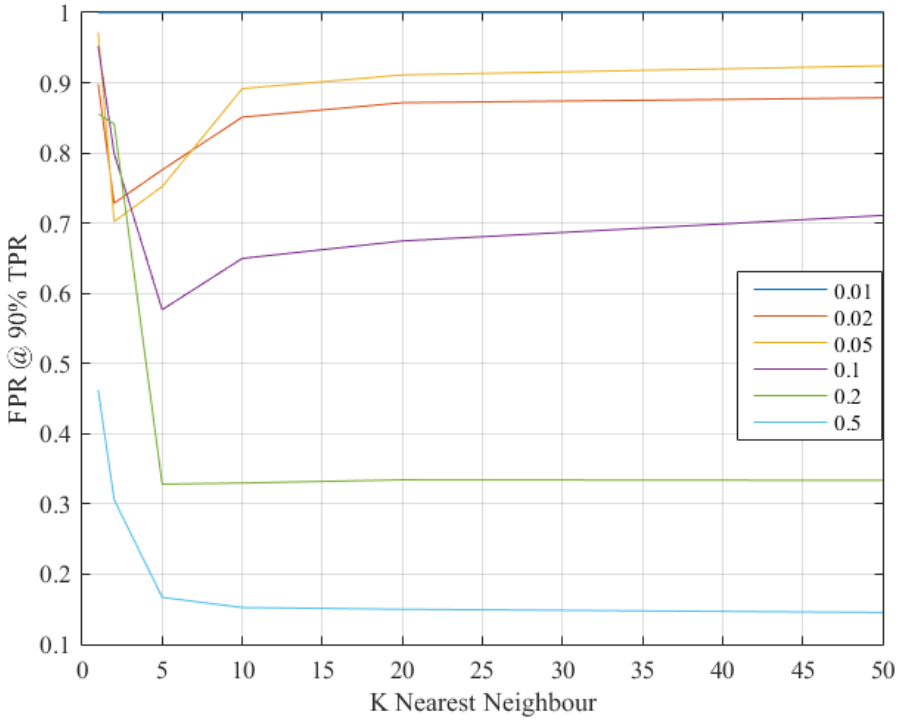


Figure 4-59 Dataset3 SHOT - Graph of FPR @ 90% TPR

Table 4-45 Dataset3 SHOT - Table of FPR @ 90% TRP

FPR @ 90% TPR (P_{fa} @ 90% P_d)	Radius					
k	0.01	0.02	0.05	0.1	0.2	0.5
1	1	0.898229	0.971949	0.952526	0.855582	0.462469
2	1	0.728602	0.702432	0.798905	0.841287	0.306041
5	1	0.775953	0.752093	0.576797	0.328229	0.166929
10	1	0.851007	0.891636	0.649819	0.329846	0.152497
20	1	0.871597	0.911171	0.674634	0.334353	0.150022
50	1	0.878556	0.924186	0.711023	0.333595	0.14547

4.5.5 SHOT Colour

This section looks at the results for using the SHOT Colour descriptor as a differentiator between the two epochs of data across Dataset 3. Figure 4-60 shows a graph of the FPR@90%TPR performance metric against the number of neighbours used to create a mean differencing with varied radii. Here, instability a low k numbers can be seen and then a levelling effect through the rest of the data. The best performance appears to be from a radius of 0.5 with decreasing performance as the radius is decreased. Table 4-46 shows the corresponding data to this graph.

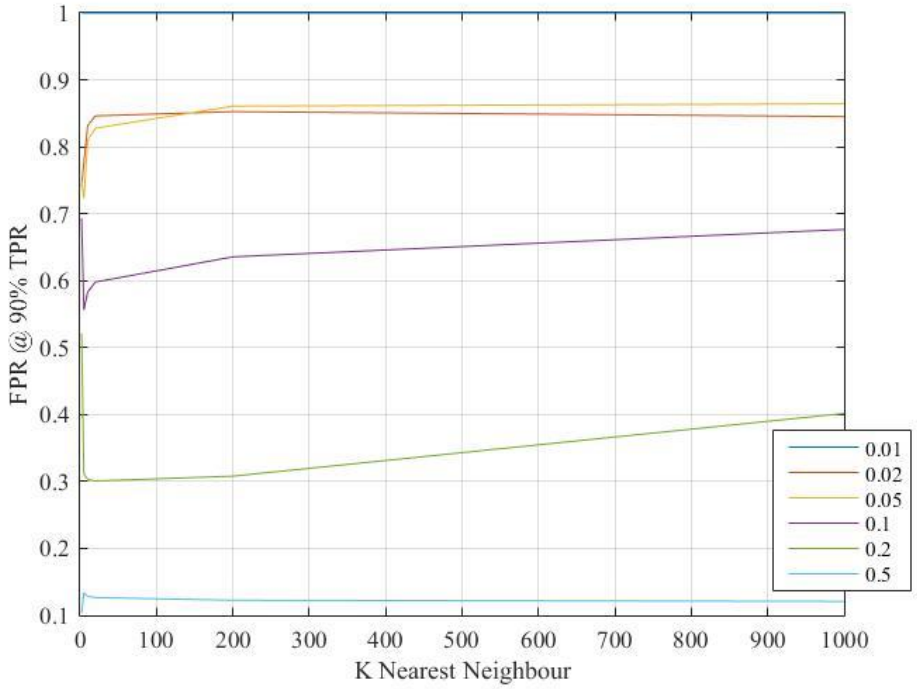


Figure 4-60 Dataset3 SHOT Colour - Graph of FPR @ 90% TPR

Table 4-46 Dataset3 SHOT Colour - Table of FPR @ 90% TRP

FPR @ 90% TPR (P_{fa} @ 90% P_d)	Radius					
k	0.01	0.02	0.05	0.1	0.2	0.5
2	1.0000	0.7409	0.7490	0.6932	0.5213	0.1030
5	1.0000	0.7758	0.7231	0.5560	0.3116	0.1337
10	1.0000	0.8315	0.8093	0.5830	0.3029	0.1284
20	1.0000	0.8461	0.8277	0.5979	0.3007	0.1265
200	1.0000	0.8525	0.8607	0.6355	0.3079	0.1226
1000	1.0000	0.8452	0.8642	0.6765	0.4016	0.1208

Figure 4-61 shows a graph of the TPR-FPR performance metric against the number of neighbours used to create a mean differencing with varied radii. Here, instability a low k numbers can be seen and then a levelling effect through the rest of the data. The best performance appears to be from a radius of 0.5 with level performance throughout variance of k values. Table 4-47 shows the corresponding data to this graph.

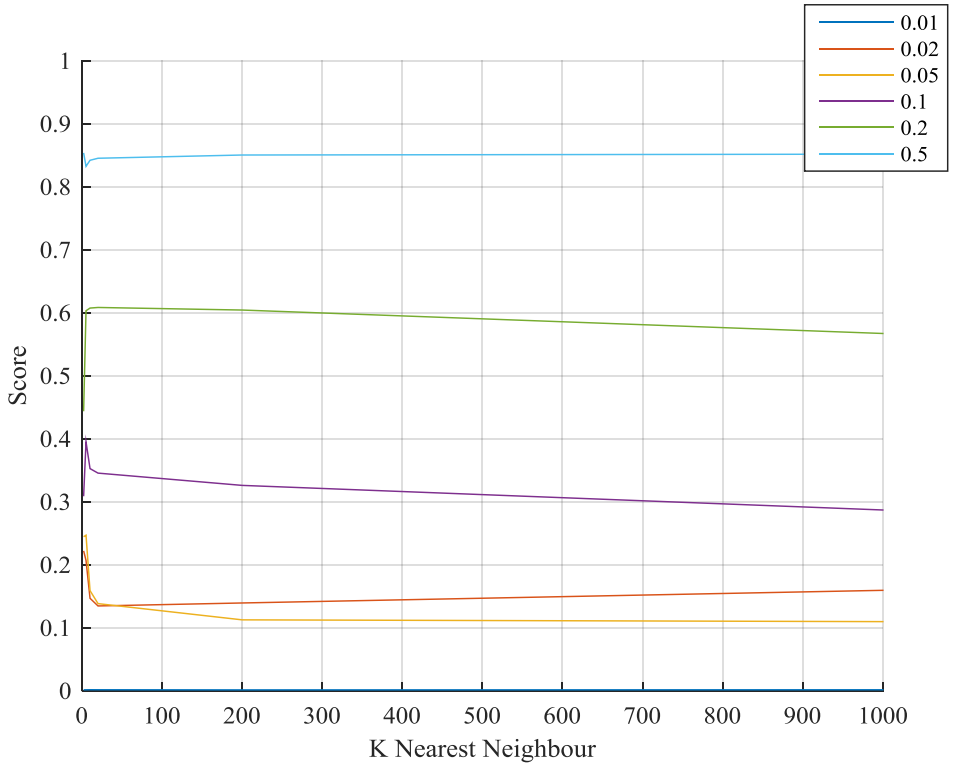


Figure 4-61- Dataset3 SHOT Colour - Graph of TPR - FPR Performance metric

Table 4-47 - Dataset3 SHOT Colour - Table of TPR - FPR performance metric

Max TP Rate - FP Rate	Radius					
K Nearest Neighbour	0.01	0.02	0.05	0.1	0.2	0.5
2	0.0014	0.2220	0.2445	0.3087	0.4436	0.8537
5	0.0015	0.2069	0.2469	0.3950	0.6031	0.8324
10	0.0016	0.1467	0.1589	0.3526	0.6074	0.8420
20	0.0015	0.1346	0.1383	0.3454	0.6083	0.8452
200	0.0013	0.1393	0.1125	0.3259	0.6042	0.8503
1000	0.0015	0.1594	0.1097	0.2869	0.5671	0.8517

Figure 4-62 shows a graph of the MCC performance metric against the number of neighbours used to create a mean differencing with varied radii. Here, instability a low k numbers can be seen and then a levelling effect through the rest of the data. The best performance appears to be from a radius of 0.5 with level performance throughout variance of k values. Table 4-48 shows the corresponding data to this graph.

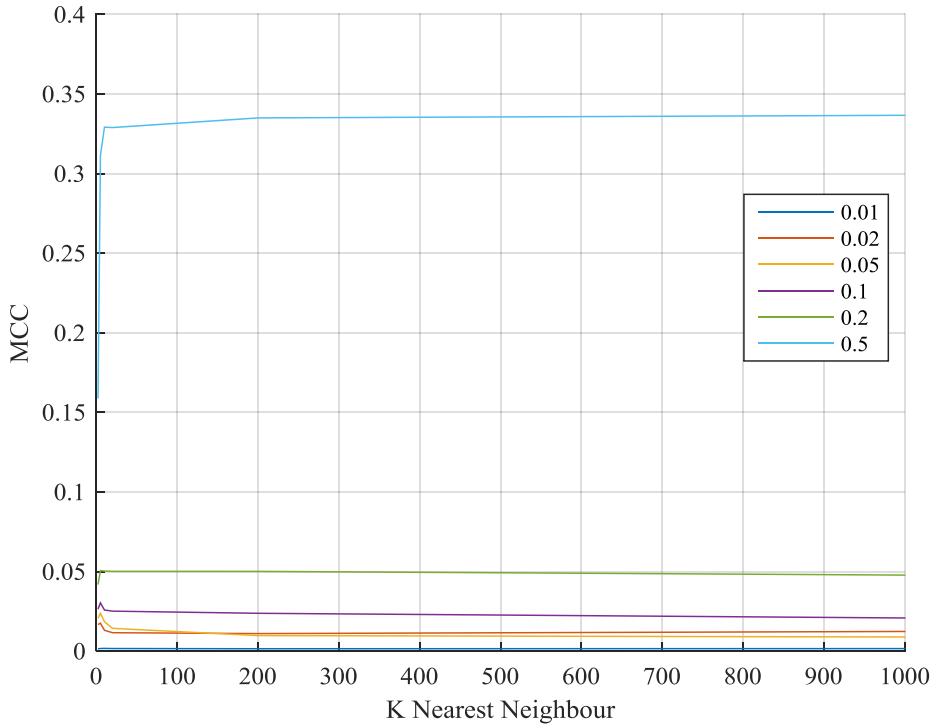


Figure 4-62- Dataset3 SHOT Colour - Graph of MCC performance metric

Table 4-48 - Dataset3 SHOT Colour - Table of MCC performance metric

MCC	Radius					
K Nearest Neighbour	0.01	0.02	0.05	0.1	0.2	0.5
2	0.0014	0.0167	0.0207	0.0263	0.0418	0.1587
5	0.0016	0.0175	0.0237	0.0303	0.0506	0.3114
10	0.0017	0.0131	0.0185	0.0258	0.0503	0.3291
20	0.0016	0.0116	0.0143	0.0251	0.0501	0.3289
200	0.0015	0.0110	0.0097	0.0238	0.0501	0.3349
1000	0.0016	0.0123	0.0089	0.0208	0.0477	0.3366

Figure 4-63 shows a graph of the AuROC performance metric against the number of neighbours used to create a mean differencing with varied radii. Here, instability a low k numbers can be seen and then a levelling effect through the rest of the data. The best performance appears to be from a radius of 0.5 with level performance throughout variance of k values. Table 4-49 shows the corresponding data to this graph.

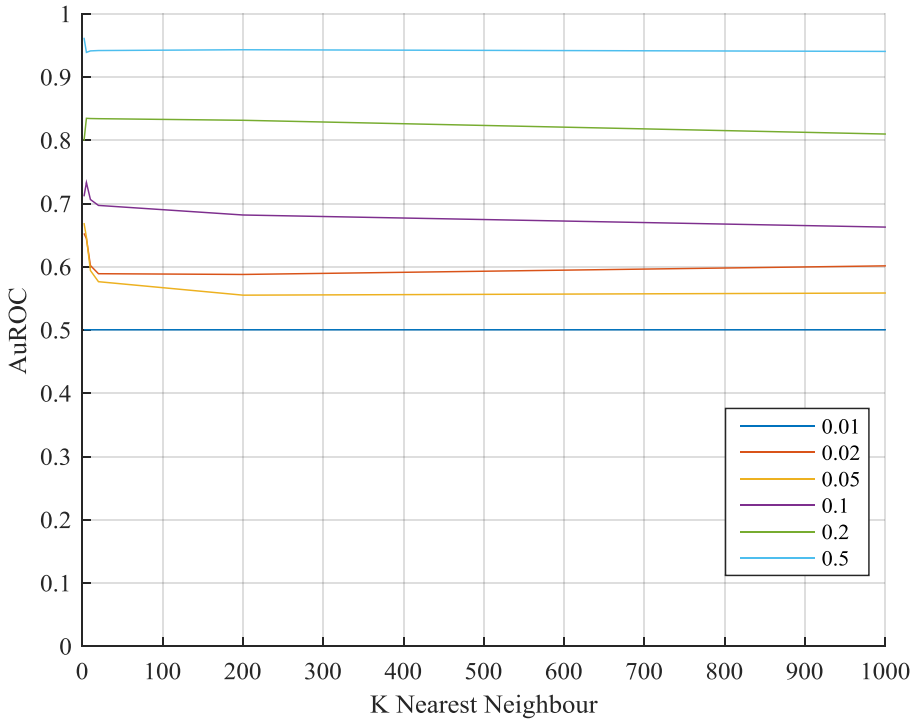


Figure 4-63- Dataset3 SHOT Colour - Graph of AuROC performance metric

Table 4-49 - Dataset3 SHOT Colour - Table of AuROC performance metrics

AuROC	Radius					
K Nearest Neighbour	0.01	0.02	0.05	0.1	0.2	0.5
2	0.5006	0.6530	0.6694	0.7116	0.7995	0.9623
5	0.5006	0.6448	0.6469	0.7333	0.8350	0.9392
10	0.5006	0.6022	0.5942	0.7065	0.8346	0.9415
20	0.5006	0.5892	0.5766	0.6972	0.8344	0.9421
200	0.5006	0.5881	0.5553	0.6821	0.8318	0.9434
1000	0.5006	0.6016	0.5587	0.6629	0.8101	0.9408

4.5.6 3DSAC

This section looks at the results for using the 3DSAC descriptor as a differentiator between the two epochs of data across Dataset 3.

Figure 4-65 visualises the change detection results across varied weighting of colour to structure.

Figure 4-65A shows 100% structural weighting and the changes can be seen clearly with little background noise surrounding it. As structural weighting is decreased and colour increased throughout the series, the size of the detected change appears to decrease. It is also observed there appears to be less noise as the structural weighting is decreased.

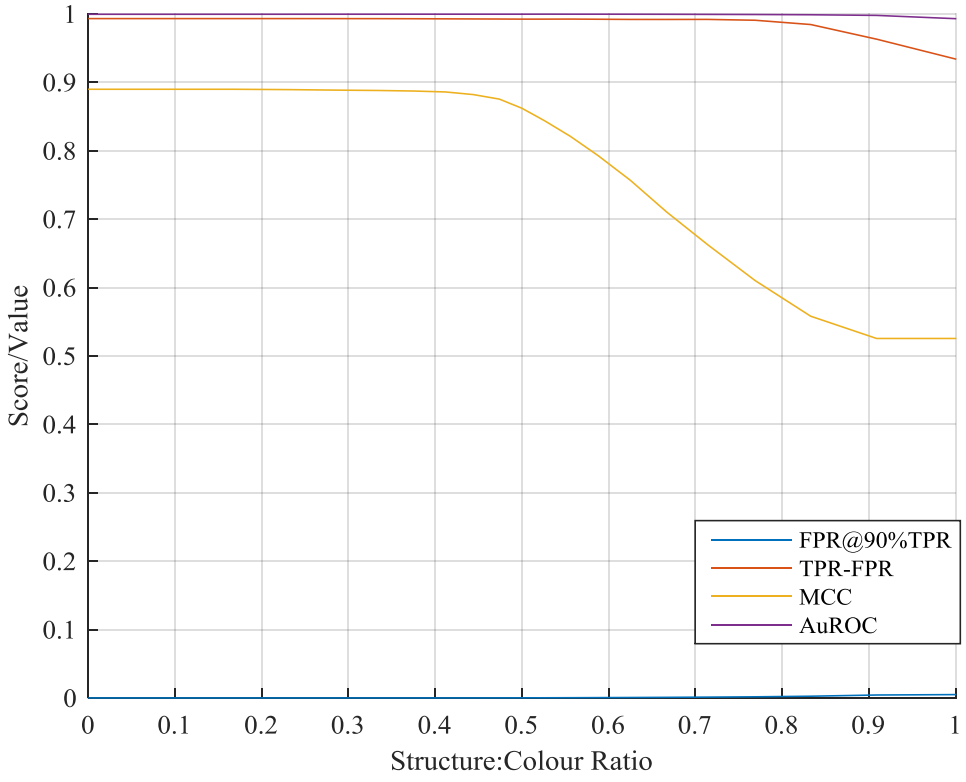


Figure 4-64 - Dataset3 3DSAC - Graph of performance metrics

Table 4-50 - Dataset3 3DSAC - Table of performance metrics

Ratio of colour to structure	FPR @ 90% TPR (P_{fa} @ 90%Pd)	Max TP Rate - FP Rate	Max MCC	Max AuROC Curve
0.0000	0.0002	0.9932	0.8899	0.9995
0.0909	0.0002	0.9932	0.8899	0.9995
0.1670	0.0002	0.9932	0.8899	0.9996
0.2310	0.0002	0.9932	0.8894	0.9996
0.2860	0.0002	0.9932	0.8887	0.9996
0.3330	0.0002	0.9932	0.8882	0.9996
0.3750	0.0002	0.9930	0.8873	0.9996
0.4120	0.0002	0.9929	0.8860	0.9997
0.4440	0.0002	0.9927	0.8822	0.9997
0.4740	0.0002	0.9926	0.8756	0.9997
0.5000	0.0003	0.9925	0.8622	0.9997
0.5260	0.0004	0.9925	0.8440	0.9997
0.5560	0.0005	0.9925	0.8210	0.9997
0.5880	0.0006	0.9923	0.7928	0.9996
0.6250	0.0008	0.9920	0.7569	0.9996
0.6670	0.0010	0.9920	0.7104	0.9995
0.7140	0.0013	0.9921	0.6629	0.9994
0.7690	0.0018	0.9908	0.6102	0.9992
0.8330	0.0027	0.9846	0.5582	0.9988
0.9090	0.0044	0.9632	0.5256	0.9979
1.0000	0.0050	0.9341	0.5256	0.9931

Figure 4-64 shows a graph of the performance metrics against the weighting used to create a differencing for 3DSAC. Here, the TPR-FPR, MCC and AuROC metrics start very high when maximum structural change weighting is defined. Only the MCC provides a discriminator whereby performance drops off as structure weighting is decreased in favour of colour. FPR@90TPR maintains a relatively constant value. Table 4-50 shows the corresponding data to this graph.

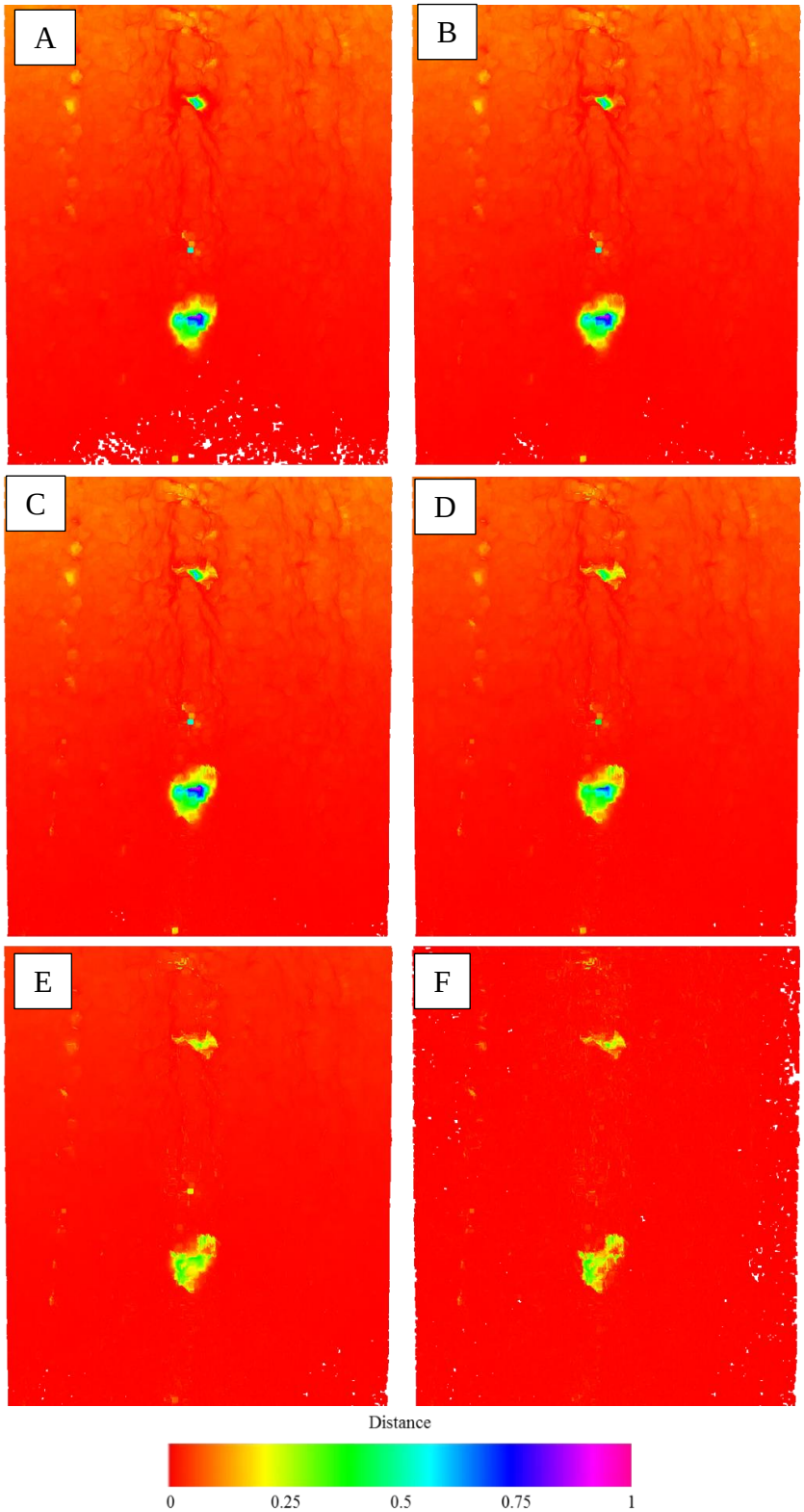


Figure 4-65 - Dataset3 3DSAC - Weighting Changes

A) 100% Structure 0% Colour B) 100% Structure 40% Colour C) 100% Structure 80% Colour D) 80% Structure 100% Colour E) 40% Structure 100% Colour F) 0% Structure 100% Colour

4.6 Summary of Results

Using the performance assessment techniques from Chapter 3, a single value of performance is derived from an analysis of measuring descriptor differences. Each of the descriptors is listed with its corresponding performance value. Where appropriate, further definition of the descriptor parameters are listed next to the descriptor.

4.6.1 Dataset 1

Below is a summary of Dataset 1 results (Table 4-51). Each of the descriptors full results are found in Section 4.3. A comprehensive analysis is detailed in Chapter 5.

Table 4-51 - Dataset 1 summary of performance results

Dataset 1-Descriptors		FPR @ 90% TPR	Max TPR-FPR	Max MCC	Max AuROC Curve
Colour	XYZ	0.4331	0.4878	0.1814	0.8260
	LAB	0.6958	0.4468	0.1694	0.7756
PFH	k = 200	0.7179	0.3766	0.2122	0.7451
FPFH	k = 20000	0.0042	0.9495	0.8301	0.9957
PCE	k = 1000	0.8731	0.0425	0.0119	0.5233
SHOT	k=2 radius = 0.5	0.1193	0.8513	0.2908	0.9126
SHOT Colour	k=2 radius = 0.5	0.1483	0.7652	0.2509	0.9091
Euclidean	k = 40	0.0011	0.9774	0.9158	0.9994
3DSAC	0% colour	0.0010	0.9866	0.9202	0.9995

The highlighted yellow cell for SHOT Colour Max AuROC curve has one slightly higher performance value when using a k value of 50. This gives a max AuROC curve of 0.9169, a 0.8% increase. It can be seen from these results that the most performant descriptor in this scenario is the 3DSAC technique (highlighted green). Across all metrics this offers the highest performance. Conversely the PCE technique offers the lowest performance across all the metrics.

4.6.2 Dataset 2

Below is a summary of Dataset 2 results (Table 4-52). Each of the descriptors full results are found in Section 4.4

Table 4-52 - Dataset 2 summary of performance results

Dataset 2- Descriptors		FPR @ 90% TPR	Max TPR-FPR	Max MCC	Max AuROC Curve
Colour	XYZ	0.3318	0.6288	0.194	0.9002
	LAB	0.6789	0.3977	0.0581	0.7559
FPFH	k = 1000	0.0205	0.9034	0.686	0.99
SHOT	k=200 radius =0.5	0.137	0.8035	0.1078	0.9071
SHOT Colour	k=200 radius =0.5	0.141	0.7993	0.0878	0.9255
Euclidean	k = 20	0.0139	0.9054	0.7032	0.9799
3DSAC	0% colour	0.0173	0.9007	0.7022	0.9770

Here it can be seen that the Euclidean technique to change detection is most performant in three of the four metrics in this scenario. Only FPFH exceeds this in the AuROC metric but is noticeably less performant in the MCC metric. The 3DSAC method closely matches the Euclidean metric while using 0% colour. Further analysis and discussion of this is detailed in Chapter 5.

4.6.3 Dataset 3

Below is a summary of Dataset 3 results. Each of the descriptors full results are found in Section 4.5

Table 4-53 - Dataset 3 summary of performance results

Dataset 3- Descriptors		FPR @ 90% TPR	Max TPR-FPR	Max MCC	Max AuROC Curve
Colour	XYZ	0.0066	0.9142	0.5544	0.9872
	LAB	0.005	0.9341	0.5256	0.9931
FPFH	k = 20000	0.4238	0.99	0.641	0.9992
SHOT	k=50 radius= 0.5	0.14547	0.8198	0.1454	0.9275
SHOT Colour	k=2 radius= 0.5	0.103	0.8537	0.1587	0.9623
Euclidean	k=20	0.0002	0.9933	0.8905	0.9995
3DSAC	0.167	0.0002	0.9932	0.8899	0.9996

Note: Highlighted cell for SHOT Colour Max MCC in results shows a significant outlier. It is noted performance for this single k nearest neighbour value is substantially lower. Where k = 5, Max MCC = 0.3114. Figure 4-62 shows this outlier graphically.

The most performant descriptor in this scenario is through a Euclidean methodology (highlighted green). It achieves the best performance across three of the four metrics (one joint with 3DSAC). And hence the 3DSAC method offers the best performance when measured across two of the four metrics. Further analysis and discussion of this is detailed in Chapter 5.

Chapter 5

Analysis and Discussion

This chapter analyses the results described in Chapter 4 to give a fuller understanding of their compositions and discusses their relevance.

First, each dataset is discussed and how this links back to the original aims in creating a scenario. An exhaustive list of each of the descriptors performance is analysed in isolation and any observations made. The analysis takes into consideration both the quantitative and qualitative assessments of the data.

Following this the results are globally discussed in full with a critical analysis and detail what has been learned regarding the technique proposed.

5.1 Analysis

The analysis section of this chapter reviews each of the descriptors in isolation to make observations through quantitative analysis and makes a qualitative analysis of their visualisation. The multi-coloured map is used for the primary analysis and will form the main point for qualitative discussion.

5.1.1 Dataset 1

Dataset 1 contains a scene with a circa 1.5% change introduced. Through a robust and rigorous truth derivation, and error to the real truth was estimated to be 1.48%. It is noted this is an over estimation of the magnitude of change, hence would result in the slight underperformance of a change detection measured against this. This noted, it will be consistent across all cases of change detection technique. This is deemed as acceptable to enable a valid performance assessment of any of those techniques and to enable a comparative analysis.

The change is split across two distinct and separated articles, a non-uniform hole and a partially buried spade. The hole presents a highly challenging scenario due to the lack of change in colour that would traditionally be the main source to discriminate between images. This presents a good scenario to display why 3D change detection is highly relevant in aerial imagery. The dominant change in this image is structural change. This can be seen in the colour results whereby the statistics of performance measurement are less than favourable.

This scenario fulfils the first scenario aim of:

- Scene changes with complex and non-uniform structural change and very little change in colour.

By looking at the summary of results in Table 4-51, the single best change detection technique by quantitative analysis is the 3DSAC algorithm closely followed by Euclidean. The 3DSAC method outperformed all other methods across all the metrics considered for performance assessment. The difference in using a Euclidean Squared

distance measurement over a Euclidean distance offers a performance increase in this scenario and by visual inspection there is a lot less noisy small magnitude change scattered around the scene. The only other descriptors suitable for detecting change in this scenario is the FPFH algorithm and the Colour based description, both performing well enough for positive visual change detection and reasonable statistical performance measures.

The summary of results show that the MCC metric provides a single performance value of 0.8301 for FPFH, 0.9158 for Euclidean and 0.9202 for 3DSAC. The area under the ROC curves provides 0.9957, 0.9994 and 0.9995 respectively. TPR-FPR provides 0.9495, 0.9774 and 0.9866 respectively. Significantly, the MCC metric provides a key discriminator in the performance of the methodologies, where the other metrics provide a validation of this. Their marginal difference between Euclidean, 3DSAC and the FPFH values does not discriminate such a large difference in performance.

Colour

This section looks at the results for using colour only as a differentiator between the two epochs of data across Dataset 1. Here the three descriptors LAB, RGB and XYZ are compared with the performance metrics for each of these listed in the results section.

By visual inspection though the visualisation of the point cloud (

Figure 4-17), the hole appears to be less prominent in all 3 of the colour spaces. This is expected due to the lack of change in colour. One observation here is that the overturned and slightly damper sand that has split over the edge of the spade whilst moving the sand is what is mainly detected, other than the spade itself. This is particularly interesting as it further points that the lack of colour change in this section of the imagery. The spade object is well detected across all three colour spaces. Thresholding of the change down to the largest 15,000 points shows a difference in the density of the points attributed to being change. Here, the RGB colour space functions the worst across all but one statistic, the FPR @ 90% TPR. The best performer across all categories is the XYZ colour space, contrary to previous research detailed in a review of literature, but with a lack of colour change in the scene this should be negated.

Euclidean

The Euclidean difference of the two epochs provides a very high performance technique to measure the change in this scenario. The highest performing metrics of the algorithms are split across values of 10 to 50 nearest neighbours, with 40 neighbours achieving the best performance. The MCC gives the discriminator with highest variance to provide a performance measurement.

PFH

The performance statistics from PFH are less than favourable. It was found when processing, very large computation times limited the ability to process the data, taking 3 days to complete the calculations when using 200 nearest neighbours. While the performance can be seen to have a positive trend as more neighbours are used, the returns from taking this long to compute are minimal. The further use of the PFH algorithm was abandoned at this stage in favour of the FPFH algorithm.

FPFH

The FPFH algorithm picked up where the PFH algorithm left off. Its optimised computational efficiency allowed for many more neighbours to be considered, and in fact its performance at the same levels of neighbours outperformed the PFH algorithm at all values above a *k* value of 5. This further validates the decision to revoke the PFH descriptor for change detection in other datasets. Above *k* of 200 the algorithm started to perform much better, with the FRP @ 90% TRP dropping off allow more of the noise to be filtered out, and hence the other metrics showing it performing more favourably. The best performance was found at very large *k* values. One visual observation is that the change detected appears to be focussed around changing contours in the image, and areas where there is a higher surface roughness.

PCE

The PCE algorithm produced a poor discriminator for change in this scenario. A consistent high probability of false alarm with low probability of detection produced an almost random ability to detect change. The MCC score hover around the 0 mark, and the AuROC curve around 0.5, showing that it is no better at finding change than it is at not

finding it. Computation time was also very expensive for this algorithm, and by visual analysis of Figure 4-21 the chance can be seen to be grouped around the larger contours of the dataset. The algorithm is designed to describe principal curvatures, and hence this holds true. From this first dataset, it is evident that this would not be a good candidate for change detection in the application of aerial imagery. For this reason, this was not taken forward for the other datasets.

SHOT

Performance of the SHOT algorithm as a change detector was poor in this scenario. Two variables were varied to try and push performance out of the algorithm, and a matrix of these changes produced. The input parameters to the algorithm allow the radius of the sphere of comprehension to be specified, and this was varied with the number of nearest neighbours included in this computation. Initial experimentation found a radius of about 0.5 became impracticable for computation times, taking up to 2 days per scenario. Fortunately by this size, and while there is a positive trend in performance across all metrics with increase in radius, the performance appeared to have settled with little more to gain from the algorithm. The best performance was found at a radius of 0.5. Visualisation of the change in Figure 4-27 shows an interesting distribution of detected change. When filtered to the largest change, it appears to have mainly detected this around contours of the large true change. This is a feature rich area following the implemented change in comparison to the flat relatively featureless sand.

SHOT Colour

The SHOT Colour algorithm performs very similarly to the SHOT algorithm, with a slightly decreased performance. It is believed this is due to the nature of the change introduced having limited colour change.

3DSAC

At its extremes of either being only structure or colour, the 3DSAC algorithm performs very similarly to either the structure or colour algorithms on their own as expected. As the ratio of colour is increased and structure decreased, there is a gradual reduction in performance as per the disparity between the two individual algorithms. More

interestingly, by visual inspection of the change visualisation, both changes of the hole and the spade can distinctly be seen, with best results around 20%-40% colour. While this cannot be recognised statistically, by visual inspection this appears to offer the best chance of an analyst detecting changes. Figure 4-33A provides a visual representation of the change without any thresholding. Here, it can be observed that the detected change is focused on the hole, which subsequently is the largest structural change from the scene. The much larger change in this appears to flood the change from the spade and any minimal spillage of sand in the transfer between the hole dug and the mound created.

5.1.2 Dataset 2

Dataset 2 contains a scene with a circa 0.1% change introduced. Through a robust and rigorous truth derivation, and error to the real truth was estimated to be 0.87%. It is noted this is an under estimation of the magnitude of change, hence would result in the slight over performance of a change detection measured against this. That being said, this is equivalent across all the technique deployed in this scenario.

The change is split across two distinct and separated articles, the movement of a chalk rock amongst other chalk rocks, and the additional placement of a red strap. The rock presents a highly challenging scenario with minimal change in colour to the scene, it would be almost impossible to detect this by normal visual inspection. The introduction of the red strap shows a distinct colour change, while representing little structural change.

This scenario fulfils the original aim of:

- Scene changes with combinations of structural changes and colour changes.

The single best change detection technique by quantitative analysis in this dataset is using a standard Euclidean distance with 3DSAC closely following second (with 0% colour weighting). Again, the only other descriptor suitable for detecting change in this scenario is the FPFH algorithm, scoring acceptably well as a change detection method with all the performance metrics.

The summary of results show that the MCC metric provides a single performance value of 0.686 for FPFH, 0.7032 for Euclidean and 0.7022 for 3DSAC. The area under the ROC curves provides 0.99, 0.9799 and 0.9770 respectively. TPR-FPR provides 0.9034, 0.9054 and 0.9007 respectively. Again, the MCC metric provides a key and significant discriminator in the performance of the methodologies.

Colour

Within the colour techniques, the LAB colour space functions the worst across all but one statistic, the FPR @ 90% TPR. Across the other three performance metrics, LAB performs considerably worse. The best performer across all categories again is the XYZ colour space. By visual inspection, all seem to pick up the structural change and colour change, with CIE and RGB appearing to detect the red colour change better than XYZ.

Euclidean

The Euclidean difference of the two epochs again provides a very high performance technique to measure the change in this scenario. The highest performing metrics of the algorithms are split across the k = 2 to 20 of nearest neighbours, with a linear decease in performance after this. The MCC appears to give the best discriminator to provide a performance measurement, again with the best performer being k=2 nearest neighbours. Figure 4-51 provides a visual representation of the change without any thresholding. Here, it can be observed that the detected change is focused on the moved chalk rock, which subsequently is the largest structural change from the scene. The minimal structural change in the additional object is all but negated showing the limitations of using just structure for change detection.

FPFH

In dataset 2, the FPFH algorithm presents an acceptable change detection technique statistically. By visual analysis it can be seen to discriminate change clearly, although many other background features are also considered to be small but significant change. Visual analysis in Figure 4-41 of k=200 and k=1000 values show an interesting change in detecting the two objects well at k=200 but seems to get seduced onto the larger structural change in k=1000. As this descriptor takes only structure into consideration,

this follows the trend of the small structural change that this object represents, and hence the data is skewed towards that of the larger structural change.

SHOT

In dataset 2, the SHOT algorithm again offers a poor performance at detecting true change.

SHOT Colour

SHOT Colour presents a poor technique for detecting change in this scenario, and interestingly offers a worse performance than normal SHOT without colour consideration. Visual inspection does show some grouping of larger change value around the true change, but there is so much other change of similar magnitude that it cannot discriminate this.

3DSAC

In dataset 2 the 3DSAC algorithm performed best with using maximum (100%) weighting toward only structure (i.e. negating colour). It can be seen by visual inspection of the change visualisation that the colour change from the strap object is not detected very prominently when only Euclidean differencing is taken (i.e. 0 colour) and as the colour ratio is increased it becomes more visible. At its extreme where only colour is present in the algorithm, the structural change is difficult to detect. Again, a change value of around 20-40% offers an additional visually interpreted advantage over pure Euclidean change. After this, the performance metrics can be seen to slope off and reduce to a less favourable method. Its overall performance is very slightly less than that of Euclidean.

5.1.3 Dataset 3

Dataset 3 contains a scene with a circa 1.3% change introduced. Through a robust and rigorous truth derivation, and error to the real truth was estimated to be 1.06%. It is noted that this is an under estimation of the number of points and hence the estimated truth model is estimated to have less points that that of the real change. As the case in Dataset 2, this would result in the over performance of any change detection technique that is applied to this scenario. This is consistent across all the implementation and hence would present equal opportunities for performance analysis.

In this scenario, a single change in made whereby a plastic liner is moved within the ditch. This offers a higher colour change in comparison to its background with the plastic liner being white. The challenge here is that there is minimal structural change, only the movement of a thin sheet of which is nestled within the ditch; also some photos gave an occluded view of this due to the contour of the ground further reducing the structural change ability.

This scenario fulfils the original aim of:

- Scene changes with very little structural change but changes in colour.

The best change detection technique by quantitative analysis is split between standard Euclidean method and 3DSAC (16% Colour) techniques. With 3SAC having a higher AuROC metric but lower in MCC. Notably the colour descriptors perform much better in this scenario and offer an acceptable visual inspection method for change detection. FPFH again scores the next best after the Euclidean and 3DSAC methods.

The summary of results show that the MCC metric provides a single performance value of 0.641 for FPFH, 0.8905 for Euclidean and 0.8899 for 3DSAC. The area under the ROC curves provides 0.9992, 0.9995 and 0.9996 respectively. TPR-FPR provides 0.99, 0.9933 and 0.9932 respectively. Again, the MCC metric provides a key and significant discriminator in the performance of the methodologies.

Colour

In dataset 3, the Colour algorithm performs much better than in the other two scenarios. This can be attributed to the nature of the change that was introduced as colour is the dominant change in the scene. By visual analysis, not much difference can be seen by eye between Colour and the Euclidean based structural descriptors.

Euclidean

The Euclidean algorithm performs very favourably, keeping false positive low and detecting most if not all of the change most of the time. The highest performance is found with a *k* value of 20 and is the best performing method when MCC is taken into consideration.

FPFH

FPFH statistically proves an acceptable method for detecting change, but at its worst across the three scenarios. By visual analysis, it is possible to discriminate the change but many artefacts are present and is not immediately clear.

SHOT

The SHOT descriptor again performs poorly, not presenting a suitable method to detect change. This can be also said for visual analysis using this technique.

SHOT Colour

The SHOT Colour descriptor offers a poor performance in dataset3 is very similar to that seen in dataset1, and outperforming dataset2.

3DSAC

This is the first scenario whereby the 3DSAC performs its best using a colour weighting. This runs true with the scenario aim, and validates this property of the descriptor quantitatively. The maximum AuROC is marginally higher than Euclidean distance, and by visual inspection the difference between 3DSAC and Euclidean is indistinguishable.

5.2 Discussion

One of the first things noticed from visual inspection of the colour change imagery in dataset 1 is that the overturned wet sand (and hence change in colour) was detected. This detected change was not accounted for in the creation of truth dataset and presents a flaw in the performance analysis metric, adversely effecting any performance figures. While the surface disruption was minimal, this was still enough to effect the colour results in a negative way. The primary aim for this dataset was to investigate a scene whereby there was minimal colour change with a dominant structural change. It is most likely why the 3DSAC algorithms best performance did not introduce any colour at all in this scenario, utilising only its structural change capability. While the colour descriptors with colour elements to their calculation (All 3 colour spaces, 3DSAC colour weighting and SHOT Colour) would have been adversely effected, the difference in Euclidean and Euclidean squared performance was found. It is evident that the Euclidean squared method within the 3DSAC change description technique provides a better solution, both by visual analysis of Figure 4-13 and by the all the performance metrics. The 3DSAC algorithm out performed all the other change detection techniques, with a marginal but measurable performance increase over Euclidean change in this scenario.

Between dataset1 and dataset2, there was an inherent expectation to have been an increase in colour change detection performance due to the scenario aim. In reality, performance increase was marginal, with relatively very low performance in comparison to using a Euclidean structural based descriptor. It is believed the more dominant structural change in this scene could have adversely affected the performance measurement of colour within all the colour utilising descriptors. The performance of using colour is much better in dataset3 and validates that colour is reactive to the scenario aims of introducing colour change. Colour on its own still represents a poor method for discriminating change in 3D point clouds, but in conjunction with structural change measured with Euclidean distance as per 3DSAC, a performance advantage can be found in both qualitative visual inspection and by statistical quantitative analysis.

Euclidean distance offers an excellent method for detecting and discriminating change, and performs consistently across all three scenarios. Its performance can be increased by varying the number of nearest neighbours and taking the mean distance. This acts to smooth the effect of outliers and surface variation brought around by point clouds in the SfM process.

A notable point in the performance analysis of 3DSAC is that only the nearest neighbour was considered. This is a flaw in the experimentation, and should be perused to enhance performance as validated by ever other descriptor throughout this research. Where the 3DSAC method was outperformed by standard Euclidean distance, the difference is highly marginal and it believed these would be exceeded by the 3DSAC method.

Statistically the descriptors calculated many millions of points and their performances were consistent. Across the datasets there is little variation, and where there is variation is explained by the aims of the research with the exception of colour performing worse once more colour change was introduced in dataset 2.

On an end to end statistical analysis, only 3 datasets were evaluated and this is a limitation in this research. The planning of flights to capture data and then process it is considerable, this and the prohibitively expensive processing times of the SHOT descriptors made even more difficult to end to more scenarios. That being said, the consistency of the descriptors give a good evaluation of the methods presented.

MCC offers best the discriminator in the performance analysis of chance detection, and statistics alone not enough to prove that one is better than the other, a visual analysis element is also required.

It is evident through these results that the large computational costs in describing points for the purpose of change detection are impractical and also do not present any performance advantage.

Chapter 6

Conclusions and Further Work

This section provides a review of the objectives set out in the introduction of this research and presents conclusions regarding these and any other notable points discovered through the research. Following this, the limitations are discussed and additional follow on work is discussed.

6.1 Conclusions

Chapter 1 defined a set of seven objectives and these will now be discussed.

Some of these are simple in nature and require little discussion, while some of the more key ones that formed the contribution to research (these being Objective 5, 6 and 7).

Objective 1: Develop realistic scenarios of challenging dynamic environments.

This research presented three realistic scenarios as follows:

- Scene changes with complex and non-uniform structural change and very little change in colour.
- Scene changes with combinations of structural changes and colour changes.
- Scene changes with very little structural change but changes in colour.

Through careful planning, each of these scenarios successfully presented challenging environments to detect change. The scenario development was a strong element of this research as it was crucial to validate performance assessment and provided an effective discriminator. Some additional considerations of this are discussed for further work.

Objective 2: Process 2D captured data into accurate 3D data that can be used efficiently to allow accurate change detection.

2D imagery was processed into 3D date in form of point clouds. As this is a current research topic in its own right and would require vast resource to develop a comparable technique, this was performed using COTS software. Photoscan was down selected by analysis of literature and was a suitable tool to achieve this. The models produced by this were of good quality and to a high fidelity.

Objective 3: Accurately define ground truth models to enable performance assessments.

Ground truth models were created by developing a robust and meticulous method to derive them. Through this it was possible to estimate the truth and the error in this. This provided credence to any of the experimentation and the errors measured were sufficiently low enough and understood. In retrospect, it could be possible to use a Lidar scanner to gather this information. This would present the object of change at a higher fidelity than the model but would most likely face similar challenges to separate this from the background scene.

Objective 4: Develop or propose a change detection process and framework.

For this research, a change detection process was developed that enabled the descriptors to be modularly inserted to perform their activities. It was also possible to use a synthetic dataset for development of the framework and pipelines which was also modular and swappable with real world data.

Objective 5: Develop a novel change detection technique that utilises 3D structure and Colour.

A simple mean Euclidean distance spread over between 20-50 nearest neighbours sometimes offers the best change detection technique when purely discussing structural change, but in reality this is not the case as colour change is always present. The tuneability of the 3DSAC algorithm allows the variability to introduce more or less colour change weighting. It is seen in every example that the 3DSAC algorithm offered an extra level of detection awareness over only using the Euclidean method. It offers the end user the ability to tune this to tailor suit the scenario or scene. This increase in performance over just structural change detection, or just colour change detection, however small,

could provide the difference when detecting critical changes such as when on search and rescue missions or detecting nefarious activities.

Objective 6: Propose a technique to quantitatively and qualitatively detect change

The quantitative detection of change is inherent of the processing framework whereby a simple thresholding technique to discriminate between change and non-change was implemented. This enabled the techniques quantify change and present this as a detection or to be discarded. Further, a visualisation technique was presented which allowed qualitative assessment. This proved very useful and helped understand the algorithms far beyond that of pure quantitative assessments.

Objective 7: Compare, contrast and critique the novel technique to other state of the art descriptors

This research provides an evaluation of some of the most prominent 3D descriptors across three challenging scenarios. It proposes the use of colour as well as structure to achieve 3D change detection from 2D aerial imagery. The research uses a multitude of techniques to analyse and measure the performance of each descriptor with the proposed 3DSAC algorithm providing the best solution for visual inspection while also maintaining a high statistical probability for detecting though the use of only filters or thresholding.

While many of the other 3D descriptors achieve a sometimes reasonable performance and can detect the change, they are not suited over change detection. PFH, FPFH, PCE, SHOT and SHOT Colour are not suitable for use as change descriptors in the case of aerial imagery. Changes in dynamic environments are best measured using structural differencing techniques, and colour can be used to enhance this.

6.2 Limitations

As raised in the discussion, a limitation of this research is that only three datasets have been implemented. These scenarios are quite open rural environments with limited

occlusion to objects. It's likely that if this was in an urban environment there would be much more occlusion. This is a limitation mainly around the data gathering, whereby if access was provided to view an area, the 3D model of that could be produced.

The scenes these show short term temporal differences in a controlled environment. It is likely that in a real world environment, there could be many more changes and could flood the change detection process. This research has been limited to those controlled situations where only relatively minimal change has been introduced. The methodology should show the relative change and hence small change in the presence of much large changing scenes could be lost among the noise.

6.3 Further Work

The work in this paper could be continued to by extending the 3DSAC technique to use *k* nearest neighbours to achieve the performance enhancements shown using other descriptors in this research. It is strongly anticipated this would give 3DSAC a hands down advantage over all the other change detection techniques in the application of point to point change comparison in aerial imagery.

Furthermore, more experimentation should be continued to give a second validation between only Euclidean, 3DSAC and FPFH methods.

Testing in different environments such as urban spaces would present a different use case, but limitations in data gathering under CAP722 make this difficult. With the more uniform shapes in urban environments, it is expected that structural change detection to be more efficient where suitable models can be created.

Scenes with more change in the number of objects and the disparity between object sizes would test the different techniques ability to discriminate between these and the background noise.

Also testing with larger seasonal change would present an interesting research topic and would raise new and different considerations using the methods discussed in this paper.

The change of vegetation and lighting conditions would add many more variations, and would be more challenging for detecting a salient change.

Currently there is a human in the loop to look at the change detected to understand this and to make decisions upon that information. In an age of increasing AI complexity, an understanding of the change could be explored, to identify and recognise the objects that have changed. Further to this, the system would be able to resolve if the change is of interest and make any future decisions on it, such as supplementary analysis or by performing an action.

Machine learning would also be an interesting follow on work whereby the algorithms could be optimised to find their absolute ultimate performances and hence further understand their performance envelopes.

References

Abeles, P (2012) 'Resolving implementation ambiguity and improving surf.' CoRR Vol.1202.0492.

Agarwal, Sameer, et al. (2011): 'Building rome in a day.' Communications of the ACM 54.10

AgiSoft, L. L. C. (2017). 'Agisoft photoscan.' Professional Edition

Alexandre, L.A., (2012, October). '3D descriptors for object and category recognition: a comparative evaluation'. In Workshop on Color-Depth Camera Fusion in Robotics at the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vilamoura, Portugal (Vol. 1, No. 3, p. 7).

Allaire, S et al. (2008). 'Full orientation invariance and improved feature selectivity of 3D SIFT with application to medical image analysis.' CVPRW'08. IEEE Computer Society Conference

Allaire, S., Kim, J.J., Breen, S.L., Jaffray, D.A. and Pekar, V., (2008, June). Full orientation invariance and improved feature selectivity of 3D SIFT with application to medical image analysis. In 2008 IEEE computer society conference on computer vision and pattern recognition workshops (pp. 1-8). IEEE.

Arya, S., Mount, D. M., Netanyahu, N. S., Silverman, R., & Wu, A. Y. (1998). An optimal algorithm for approximate nearest neighbor searching fixed dimensions. Journal of the ACM (JACM), 45(6), 891-923.

Autodesk, Inc. (2016). '123D Catch'

Barber, D. M., D. Holland, and J. P. Mills. (2008). 'Change detection for topographic mapping using three-dimensional data structures.' International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 37.B4

Barnhart, Theodore B., and Benjamin T. Crosby. (2013). 'Comparing two methods of surface change detection on an evolving thermokarst using high-temporal-frequency terrestrial laser scanning, Selawik River, Alaska.' Remote Sensing5.6 : 2813-2837.

Bay, H., Tuytelaars, T. and Van Gool, L., (2006, May). Surf: Speeded up robust features. In European conference on computer vision (pp. 404-417). Springer, Berlin, Heidelberg.

Bianco, S., Ciocca, G. and Marelli, D. (2018) 'Evaluating the Performance of Structure from Motion Pipelines', Journal of Imaging, 4(8), p. 98. doi: 10.3390/jimaging4080098.

Black, JW. (1860). 'Balloon View of Boston'. Photograph by J.W. Black, October 13, 1860. Boston Pictorial Archive.

Blauensteiner, P., Wildenauer, H., Hanbury, A., & Kampel, M. (2006). 'On colour spaces for change detection and shadow suppression.'

Bolles R. C., H. H. Baker, and D. H. Marimont. (1987). Epipolar-plane image analysis: An approach to determining structure from motion. International Journal of Computer Vision, 1:7–55"

Boughorbel, S., Jarray, F. and El-Anbari, M. (2017) 'Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric', PLOS ONE. Edited by Q. Zou, 12(6), p. e0177678. doi: 10.1371/journal.pone.0177678.

Boyle, W. S. and Smith, G. E. (1970) 'Charge Coupled Semiconductor Devices', Bell System Technical Journal, 49(4), pp. 587–593. doi: 10.1002/j.1538-7305.1970.tb01790.x.

Bradski, G (2000). 'The OpenCV Library.' Dr. Dobb's Journal: Software Tools for the Professional Programmer 25.11 : 120-123.

Brutto, M. Lo, and Paola Meli. (2012).'Computer vision tools for 3D modelling in archaeology.' International Journal of Heritage in the Digital Era 1.1_suppl: 1-6.

Champion, Nicolas. (2007). '2D building change detection from high resolution aerial images and correlation digital surface models.' International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences36.3/W49A: 197-202.

Chicco, D. (2017) 'Ten quick tips for machine learning in computational biology', BioData Mining, 10(1), p. 35. doi: 10.1186/s13040-017-0155-3.

Choi, Sunglok, Taemin Kim, and Wonpil Yu. (1997). 'Performance evaluation of RANSAC family.' Journal of Computer Vision 24.3: 271-300.

Chum, Ondrej, and Jırı Matas. (2002). 'Randomized RANSAC with Td, d test.' Proc. British Machine Vision Conference. Vol. 2.

CIE. (1931). 'Commission internationale de l'Eclairage proceedings' 1931', Cambridge University Press.

Civil Aviation Authority. (2010).'CAP 722 Unmanned Aircraft System Operations in UK Airspace—Guidance.' Directorate of Airspace Policy.

de Jong, K. L. and Bosman, A. S. (2018) 'Unsupervised Change Detection in Satellite Images Using Convolutional Neural Networks'. Available at: http://arxiv.org/abs/1812.05815.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). 'Imagenet: A large-scale hierarchical image database'. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee.

Faugeras, O. D., Luong, Q. T., & Maybank, S. J. (1992, May). Camera self-calibration: Theory and experiments. In European conference on computer vision (pp. 321-334). Springer, Berlin, Heidelberg.

Fawcett, Tom. (2006). 'An introduction to ROC analysis.' Pattern recognition letters 27.8: 861-874.

Filipe, S. and Alexandre, L.A., (2014, January). A comparative evaluation of 3D keypoint detectors in a RGB-D object dataset. In 2014 International Conference on Computer Vision Theory and Applications (VISAPP) (Vol. 1, pp. 476-483). IEEE.

Fischler, M. A., & Bolles, R. C. (1981). 'Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography.' Communications of the ACM, 24(6), 381-395.

Flitton, G.T., Breckon, T.P. and Bouallagu, N.M., (2010, September). 'Object Recognition using 3D SIFT in Complex CT Volumes'. In BMVC (No. 1, pp. 1-12).

Forsyth, D. A., & Ponce, J. (2002). 'Computer vision: a modern approach'. Prentice Hall Professional Technical Reference.

Frahm, J.M., Fite-Georgel, P., Gallup, D., Johnson, T., Raguram, R., Wu, C., Jen, Y.H., Dunn, E., Clipp, B., Lazebnik, S. and Pollefeys, M., (2010, September). Building rome

on a cloudless day. In European Conference on Computer Vision (pp. 368-381). Springer, Berlin, Heidelberg.

Furukawa Y and Ponce J, (2010). 'Accurate, Dense, and Robust Multi-View Stereopsis', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 32, Issue 8, Pages 1362-1376.

Ganesan, P., V. Rajini, and R. Immanuvel Rajkumar. (2010). 'Segmentation and edge detection of color images using CIELAB color space and edge detectors.' Emerging Trends in Robotics and Communication Technologies (INTERACT), 2010 International Conference on. IEEE.

Girardeau-Montaut, D., Roux, M., Marc, R., & Thibault, G. (2005). 'Change detection on points cloud data acquired with a ground laser scanner'. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(part 3), W19.

Gómez-Gutiérrez, Álvaro, et al. (2015). 'Does HDR pre-processing improve the accuracy of 3D models obtained by means of two conventional SfM-MVS software packages? The case of the Corral del Veleta Rock Glacier.' Remote Sensing 7.8: 10269-10294.

Guo, Y., Bennamoun, M., Sohel, F., Lu, M., Wan, J., & Kwok, N. M. (2016). 'A comprehensive performance evaluation of 3D local feature descriptors'. International Journal of Computer Vision, 116(1), 66-89.

Hana, X. F., Jin, J. S., Xie, J., Wang, M. J., & Jiang, W. (2018). 'A comprehensive review of 3d point cloud descriptors'. arXiv preprint arXiv:1802.02297.

Hanley, James A., and Barbara J. McNeil. (1982). 'The meaning and use of the area under a receiver operating characteristic (ROC) curve.' Radiology 143.1: 29-36.

Hartley, R. I. (1992, May). 'Estimation of relative camera positions for uncalibrated cameras'. In European conference on computer vision (pp. 579-587). Springer, Berlin, Heidelberg.

Hartley, R. I. (1997). 'In defense of the eight-point algorithm'. IEEE Transactions on pattern analysis and machine intelligence, 19(6), 580-593.

Hartley, R., & Zisserman, A. (2003). 'Multiple view geometry in computer vision'. Cambridge university press.

Harwin, S. and Lucieer, A. (2012) ‘An accuracy assessment of georeferenced point clouds produced via multi-view stereo techniques applied to imagery acquired via unmanned aerial vehicle’, ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIX-B7, pp. 475–480. doi: 10.5194/isprsarchives-XXXIX-B7-475-2012.

Harwin, Steve, and Arko Lucieer. (2012). 'Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from unmanned aerial vehicle (UAV) imagery.' Remote Sensing 4.6: 1573-1599.

Horn, B. K. P. and Schunck, B. G. (1981) ‘Determining optical flow’, Artificial Intelligence, 17(1–3), pp. 185–203. doi: 10.1016/0004-3702(81)90024-2.

Hussain, Masroor, et al. (2013). 'Change detection from remotely sensed images: From pixel-based to object-based approaches.' ISPRS Journal of Photogrammetry and Remote Sensing 80 : 91-106.

Ioannidou, A., Chatzilari, E., Nikolopoulos, S., & Kompatsiaris, I. (2017). Deep learning advances in computer vision with 3d data: A survey. ACM Computing Surveys (CSUR), 50(2), 20.

Juan, Gwun. (2009). 'A comparison of SIFT, PCA-SIFT and SURF', International Journal of Image Processing (IJIP), Volume ,3 Issue 4, p143 - 152.

Ke, Y. and Sukthankar, R., (2004). PCA-SIFT: A more distinctive representation for local image descriptors. CVPR (2), 4, pp.506-513.

Kersten, Thomas P., and Maren Lindstaedt. (2012). 'Image-based low-cost systems for automatic 3D recording and modelling of archaeological finds and objects.' Euro-Mediterranean Conference. Springer Berlin Heidelberg.

Krig, S. (2016) ‘Interest Point Detector and Feature Descriptor Survey’, in Computer Vision Metrics. Cham: Springer International Publishing, pp. 187–246. doi: 10.1007/978-3-319-33762-3_6.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

Kromer, R., Abellán, A., Hutchinson, D., Lato, M., Edwards, T., & Jaboyedoff, M. (2015). A 4D filtering and calibration technique for small-scale point cloud change detection with a terrestrial laser scanner. Remote Sensing, 7(10), 13029-13052.

Lague, Dimitri, Nicolas Brodu, and Jérôme Leroux.(2013) 'Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (NZ).' ISPRS Journal of Photogrammetry and Remote Sensing 82 (2013): 10-26.

Lawrence, G. (1906). Photograph of San Francisco in ruins from Lawrence Captive Airship Geo. R. Lawrence Co.

Levenberg, Kenneth. (1944). 'A method for the solution of certain non-linear problems in least squares.' Quarterly of applied mathematics 2.2: 164-168.

Longuet-Higgins, H. C. (1981). A computer algorithm for reconstructing a scene from two projections. Nature, 293(5828), 133-135.

Lourakis, Manolis IA, and Antonis A. Argyros. (2009). 'SBA: A software package for generic sparse bundle adjustment.' ACM Transactions on Mathematical Software (TOMS) 36.1: 2.

Lowe, D G. (1999). 'Object recognition from local scale-invariant features.' Computer vision, 1999. The proceedings of the seventh IEEE international conference on. Vol. 2. Ieee.

Lu, Dengsheng, P. Mausel, E. Brondizio, and Emilio Moran. (2004). 'Change detection techniques.' International journal of remote sensing 25, no. 12 : 2365-2401.

Lucas, B D., and Kanade, T. (1981). 'An iterative image registration technique with an application to stereo vision': 674-679.

Luong, Q. T., Deriche, R., Faugeras, O., & Papadopoulo, T. (1993). On determining the fundamental matrix: Analysis of different methods and experimental results.

Ma, Y., Huang, K., Vidal, R., Košecká, J., & Sastry, S. (2004). Rank conditions on the multiple-view matrix. International Journal of Computer Vision, 59(2), 115-137.

Marquardt, Donald W. (1963). 'An algorithm for least-squares estimation of nonlinear parameters.' Journal of the society for Industrial and Applied Mathematics11.2 : 431-441.

Matthews, Brian W. (1975). 'Comparison of the predicted and observed secondary structure of T4 phage lysozyme.' Biochimica et Biophysica Acta (BBA)-Protein Structure 405.2: 442-451.

Meer,Subbarao. (2006). 'Beyond RANSAC: User independent robust regression' Computer, Vision and Pattern Recognition Workshop, 2006. CVPRW '06. Conference on 17-22, June (2006)"

Mikolajczyk, K.; Schmid, C. (2005). 'A performance evaluation of local descriptors', Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.27, no.10, pp.1615-1630

Muja M, Lowe, D. (2009). 'Fast approximate nearest neighbors with automatic algorithm configuration', In VISAPP International Conference on Computer Vision Theory and Applications."

Nistér, David. (2005). 'Preemptive RANSAC for live structure and motion estimation.' Machine Vision and Applications 16.5: 321-329.

Oliensis, J. (2000). 'A critique of structure-from-motion algorithms.' Computer Vision and Image Understanding 80.2: 172-214.

Olsen, Michael J., Shawn Butcher, and Evon P. Silvia. (2012). 'Real-time change and damage detection of landslides and other earth movements threatening public infrastructure.'.

Ostrowski, S., et al. (2014). 'Analysis of point cloud generation from UAS images.' ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2.1: 45.

Peterson, Eric & Klein, Matthew & L. Stewart, Robert. (2015). Whitepaper on Structure from Motion (SfM) Photogrammetry: Constructing Three Dimensional Models from Photography, Report number: ST-2015-3835-1, Affiliation: U.S. Bureau of Reclamation.

Powers, David Martin. (2011). 'Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.'

Raguram, R., Frahm, J. M., & Pollefeys, M. (2008, October). A comparative analysis of RANSAC techniques leading to adaptive real-time random sample consensus. In European Conference on Computer Vision (pp. 500-513). Springer, Berlin, Heidelberg.

Rusu, R. B., & Cousins, S. (2011, May). 3d is here: Point cloud library (pcl). In 2011 IEEE international conference on robotics and automation (pp. 1-4). IEEE.

Rusu, R. B., Blodow, N., & Beetz, M. (2009, May). Fast point feature histograms (FPFH) for 3D registration. In 2009 IEEE International Conference on Robotics and Automation (pp. 3212-3217). IEEE.

Rusu, R. B., Marton, Z. C., Blodow, N., & Beetz, M. (2008). Persistent point feature histograms for 3D point clouds. In Proc 10th Int Conf Intel Autonomous Syst (IAS-10), Baden-Baden, Germany (pp. 119-128).

Rusu, Radu Bogdan, et al. (2008). 'Persistent point feature histograms for 3D point clouds.' Proc 10th Int Conf Intel Autonomous Syst (IAS-10), Baden-Baden, Germany.

Rusu, Radu Bogdan, Nico Blodow, and Michael Beetz. (2009). 'Fast point feature histograms (FPFH) for 3D registration.' Robotics and Automation, 2009. ICRA'09. IEEE International Conference on. IEEE.

Rusu, Radu Bogdan. (2010). 'Semantic 3d object maps for everyday manipulation in human living environments.' KI-Künstliche Intelligenz 24.4: 345-348.

Salti, S., Tombari, F., & Di Stefano, L. (2014). SHOT: Unique signatures of histograms for surface and texture description. Computer Vision and Image Understanding, 125, 251-264.

Salti, S., Tombari, F., Spezialetti, R. and Di Stefano, L., (2015). Learning a descriptor-specific 3D keypoint detector. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2318-2326).

Salvi, J., Armangué, X., & Batlle, J. (2002). A comparative review of camera calibrating methods with accuracy evaluation. Pattern recognition, 35(7), 1617-1635.

Schonberger, J. L. and Frahm, J. M. (2016) ‘Structure-from-Motion Revisited’, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem, pp. 4104–4113. doi: 10.1109/CVPR.2016.445.

Scovanner, P., Ali, S. and Shah, M. (2007) ‘A 3-dimensional sift descriptor and its application to action recognition’, Proceedings of the ACM International Multimedia Conference and Exhibition, (c), pp. 357–360. doi: 10.1145/1291233.1291311.

Shi, J (1994). 'Good features to track.' Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on. IEEE.

Shi, Juan, Jinling Wang, and Yaming Xu. (2011)'Object-based change detection using georeferenced UAV images.' Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci 38: 177-182.

Singh, Ashbindu. (1989). 'Review article digital change detection techniques using remotely-sensed data.' International journal of remote sensing 10.6: 989-1003.

Smith, M. W., J. L. Carrivick, and D. J. Quincey. (2016). 'Structure from motion photogrammetry in physical geography.' Progress in Physical Geography40.2: 247-275.

Szeliski, R. (2010). Computer vision: algorithms and applications. Springer Science & Business Media.

Tombari, F., Salti, S. and Di Stefano, L. (2013) 'Performance Evaluation of 3D Keypoint Detectors', International Journal of Computer Vision, 102(1–3), pp. 198–220. doi: 10.1007/s11263-012-0545-4.

Tombari, Federico, Samuele Salti, and Luigi Di Stefano. (2010). 'Unique signatures of histograms for local surface description.' European conference on computer vision. Springer, Berlin, Heidelberg.

Tombari, Federico, Samuele Salti, and Luigi Di Stefano. (2011). 'A combined texture-shape descriptor for enhanced 3D feature matching.' Image Processing (ICIP), 2011 18th IEEE International Conference on. IEEE.

Torr, P. H. S., & Davidson, C. (2003). IMPSAC: Synthesis of importance sampling and random sample consensus. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(3), 354-364.

Torr, P. H., & Murray, D. W. (1997). The development and comparison of robust methods for estimating the fundamental matrix. International journal of computer vision, 24(3), 271-300.

Torr, P. H., & Zisserman, A. (2000). MLESAC: A new robust estimator with application to estimating image geometry. Computer vision and image understanding, 78(1), 138-156.

Tsai, R Y., and Huang, T (1984). 'Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces.' IEEE Transactions on pattern analysis and machine intelligence 1: 13-27.

Tuytelaars T, Mikolajczyk K. (2008). 'Local invariant feature detectors: a survey', Foundations and Trends in Computer Graphics and Vision archive Volume 3 Issue 3, 2008

Ullman, S. (1979). The interpretation of structure from motion. Proceedings of the Royal Society of London. Series B. Biological Sciences, 203(1153), 405-426.

Varghese, A., Gubbi, J., Ramaswamy, A., & Balamuralidhar, P. (2018). ChangeNet: a deep learning architecture for visual change detection. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 0-0).

Verhoeven, Geert. (2011). 'Taking computer vision aloft–archaeological three-dimensional reconstructions from aerial photographs with photoscan.' Archaeological Prospection 18.1: 67-73.

Wallach, H., & O'connell, D. N. (1953). The kinetic depth effect. Journal of experimental psychology, 45(4), 205.

Waltz, E and Llinas J. (1990). 'Multisensor data fusion.' Vol. 685. Boston: Artech house.

Westoby, M. J., et al. (2012). '‘Structure-from-Motion’photogrammetry: A low-cost, effective tool for geoscience applications.' Geomorphology 179: 300-314.

Xiao, W. (2012). Detecting changes in trees using multi-temporal airborne LiDAR point clouds. University of Twente Faculty of Geo-Information and Earth Observation (ITC).

Zavodny, Alexandri Gregor. (2012).Change detection in lidar scans of urban environments. University of Notre Dame.

Zhang, Z., Vosselman, G., Gerke, M., Tuia, D., & Yang, M. Y. (2018). Change detection between multimodal remote sensing data using siamese CNN. arXiv preprint arXiv:1807.09562.

Zhong, Y. (2009) ‘Intrinsic shape signatures: A shape descriptor for 3D object recognition’, in 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops. IEEE, pp. 689–696. doi: 10.1109/ICCVW.2009.5457637.