

Review of non-contact methods for automated aircraft inspections

Pasquale Lafiosca^{a*}, Dr Ip-Shing Fan^a

^aIntegrated Vehicle Health Management Centre, Cranfield University, United Kingdom

Abstract

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

Keywords: Aircraft inspection; Pattern recognition; Triangulation; Stereo vision; Structured light; Time of flight; Thermography.

1 Introduction

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

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

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

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

Today many tools are readily available for the inspection process, but the final assessment still relies on the inspector's experience. Drones equipped with a camera combined with pattern recognition algorithms are used to evaluate aircraft airworthiness after hail, lightning strike or other accidental damages. Better GVI automation would lower the downtime of the aircraft while improving safety and reliability, so that the inspector's expertise can be spent on evaluating the actual damage (detailed inspection) without having to search over the whole aircraft structure. The aim of this paper is to evaluate non-destructive inspection (NDI) technologies from the new perspective of their *capability to be automated*. In this review, only non-contact methods for external NDI of the aircraft surface are included, as they are more suitable for GVI automation in the medium term. Without contact, an end effector composed of multiple sensors could move fast and freely around the aircraft, completing a scan within hours and potentially delivering a report with sufficient level of certainty.

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

Pattern recognition is briefly illustrated in Sec. 2. However this technology can not be applied to defects like dents as they do not have sharp boundaries and require to catch the *third dimension*. Automated 3D measurement methods using passive and active triangulation are introduced in Sec. 3, while time of flight cameras are presented in Sec. 4. Different materials do not react equally to damage. A carbon fibre reinforced composite part, in fact, may show an almost invisible dent compared to an aluminium one exposed to the same impact and yet present a subsurface *delamination* [3], then, again, the judgment of the human inspector is critical. An attempt to detect subsurface damages can be pursued with thermography, as introduced in Sec. 5. Despite its pertinence to detect delamination, ultrasound scanning is not discussed here because it generally requires contact. As this review aims to address in situ aircraft inspection, methods that require

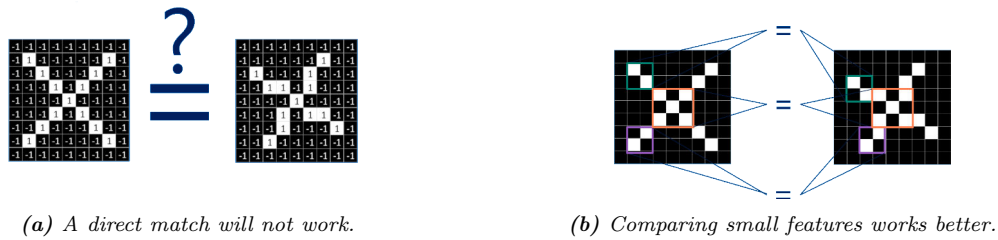
*Corresponding author: pasquale.lafiosca@cranfield.ac.uk

highly controlled environment or complex optical structures [4] are omitted. Sec. 6 concludes with the applicability of each technology to different damage types, and the challenges to automate these methods for aircraft inspection.

2 Pattern recognition

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

Pattern recognition has been widely applied in many fields, as well in aircraft inspections. Commercial drones equipped with a camera¹ are already available on the aircraft MRO market, thus allowing to scan the entire aircraft quickly. To some extent, the GVI task can be interpreted as a 2D image classification problem, looking for recurring image patterns typically found as positive flaws by the inspector, e.g. lightning strikes or scratches. Although there are several classification algorithms usable for this scope, the state-of-the-art performances are reached using convolutional neural networks (CNN) [5].

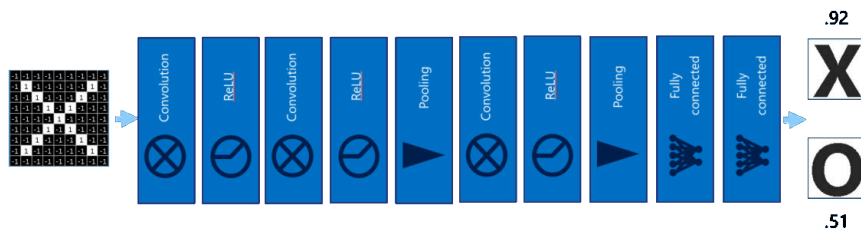


Source: Brandon Rohrer, CC0 (public domain).

Figure 1: Basic CNN approach.

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

The CNN scheme is particularly adapt to visual tasks since is unaffected by position shift and inspired by biological processes [6]. Its approach is to select *relevant* pieces from the original image called *features* (Fig. 1), which are smaller two-dimensional matrix that are most able to distinguish e.g. an “X” from a “O” [7]. The layers can be stacked multiple times in every order (Fig. 2) and the network may have very different configurations. As a rule of thumb, the more the layers (so that the network is *deep*), more sophisticated features of the image can be understood. CNN require a



Source: Brandon Rohrer, CC0 (public domain).

Figure 2: An example of complete CNN layout with two outputs.

significant amount of *training* data available, i.e. “damage” and “non-damage” labeled data so that the network can iteratively *learn* which features are more important to classify the input. This is done calculating the error at each iteration and doing the *backpropagation* to adapt weights accordingly.

2.1 Pros and cons

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

- It requires huge amounts of training data. In many situations such data is not easy to gather, for example because that kind of damage does not happen very often.
- It operates well only in diffused light conditions, while could easily fail in presence of non-uniform light.

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

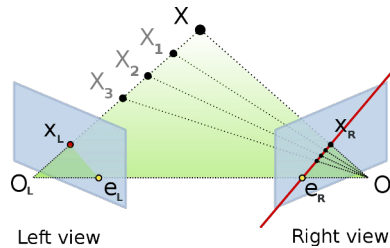
- Missing the third dimension, it is incapable to detect dents and other shape deformations, while it works good for scratches or lightning strikes.
- The output is only a prediction that a part of an image belongs to classes of flaws within a certain confidence. E.g. an image could have an output of 0.70 in the class “scratch” and, at the same time, 0.40 in the class “lightning strike”. From these values one may infer that the image contains a scratch with a probability of 70% and a lightning strike with a probability of 40%, but in any case this is giving no *quantitative* evaluation of the damage. As a consequence, the inspector would probably need to go through many false positives.

3 Triangulation

The basic principle to recover the depth z from images is to measure the triangle among the object located in an unknown point X and two known points. The two known points can be cameras or a camera and a light source. Using triangulation without manipulating light sources is called *passive* triangulation [8], *active* triangulation otherwise. The nature of the problem between the two is different, because in passive methods a correspondence needs to be established between features from two different camera images of the same scene, while in active methods the correspondence problem has already been solved using an artificial source of illumination [9].

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

In a given stereo system the so-called *fundamental matrix* establishes a relation between a point X_L in one image and an *epipolar line* onto the other one. Thanks to the epipolar geometry, the *stereo correspondence* is a one-dimension problem that consists to find the point X_R corresponding to X_L on that epipolar line only (Fig. 3) without having to span the whole image.



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

Figure 3: The search for point X_L on the right image is restricted to the epipolar line passing by e_R and X_R . If the image is rectified, this line becomes horizontal.

Since epipolar lines are generally oblique, after the acquisition of a image pair it is convenient to *rectify* the two images, i.e. to transform the two images as if they were taken with only a horizontal displacement and projected to a common image plane parallel to the baseline. This reduces the matching problem on the x -axis only as the algorithm searches for the corresponding points on a horizontal line [12]. Although rectification is not a necessary step, it is largely used to speed up the matching process.

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

3.1 Passive triangulation

Photogrammetry is a passive triangulation method that involves using many perspective views of the same scene. More than two points of views can be useful in some situations, e.g. where horizontal obstacles are present [13]. *Stereo vision* is the sub-class of photogrammetry that only involves two points of view to extract information (Fig. 4).

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

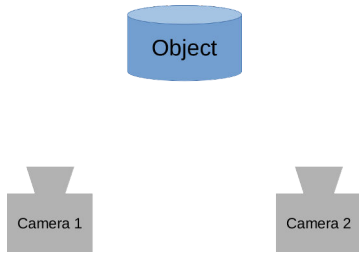


Figure 4: Passive stereo system elements.

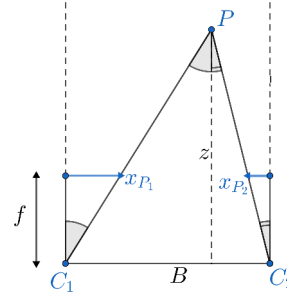


Figure 5: Recovering the depth information from disparity is based on similar triangles. The triangle PC_1C_2 is similar to the union of triangles built on the segments XP_1 and XP_2 .

12], achieving our final goal. The depth z of the corresponding pixel point P can be calculated with:

$$z = \frac{f \cdot B}{d_P} \quad (1)$$

where f is the camera's focal length² and B is the stereo baseline (Fig. 5). Note that in case of cameras with different focal lengths, one can be rescaled to match the other one.

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

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

$$d_P = \underset{0 \leq d \leq d_{max}}{\operatorname{argmin}} \sum_{x \in w} \operatorname{diff}(x, x - d) \quad (2)$$

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

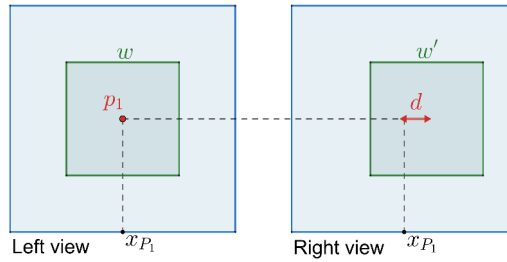


Figure 6: Local algorithms look for a window w shifted by d_P so that the overall colour differences between w and w' is minimised.

Many complex local methods have been developed replacing the window definition and the cost function of Eq. (2) to obtain more accuracy and computational efficiency. Most of the local approaches are based on:

- Adaptive windows: the principle is to keep windows large but without overlap disparity discontinuities, so to get clear edges.
- Weighted pixels: the key idea is to give a different weight to each pixel in influencing the matching process.
- Colour: takes into account adjacent colour areas which are likely to have the same disparity.

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

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

$$E(D) = E_{data}(D) + \lambda \cdot E_{smooth}(D) \quad (3)$$

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

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

$$E_{data}(D) = \sum_{p \in I_l} c(p, p - d_p) \quad (4)$$

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

$$E_{smooth}(D) = \sum_{p, q \in \eta} s(d_p, d_q) \quad (5)$$

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

The target of global methods is to find a disparity map that minimises the equation (3), but finding the global minimum is, in general, an np-complete problem [16]. One way to reduce the complexity of the energy equation is to map neighbouring pixels into a tree on the image grid, so that η does not contain cycles and an optimal solution can be computed efficiently via dynamic programming. However, global methods also find their limits because of the energy formulation, which does not efficiently represent the stereo matching problem [14], despite further developments have been made towards better converging algorithms.

3.1.1 Pros and cons

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

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

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

For a benchmark evaluation many of the passive stereo matching algorithms refer to a common evaluation dataset, e.g. KITTI⁴. At date the method with best performance is “LEAstereo”, achieving a 1.65% of outliers⁵ over the all ground truth pixels of the test set of the 2015 KITTI dataset.

3.2 Active triangulation

Active triangulation introduces the control of light sources. In general the term *active stereo* refers to a configuration with two cameras and one light source (Fig. 7a), while *structured light* refers to a configuration with one camera and one light source (Fig. 7b). In both cases light is projected with a known pattern onto the object, so that the way that shapes are observed becomes a new source of information.

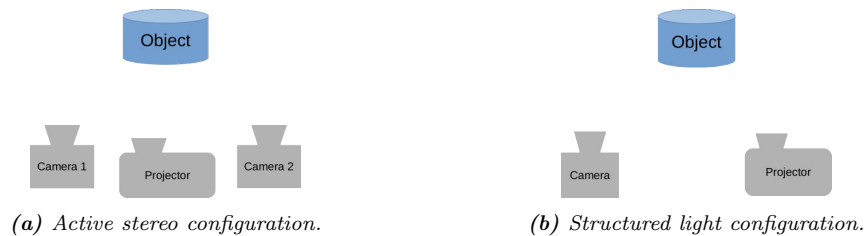


Figure 7: Common active triangulation configurations.

Active stereo is more common as it allows to do triangulation between the two cameras as in passive stereo, without the need to calibrate the projector. In fact, even if methods for projector calibration have been designed [17, 18], it

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

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

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

remains a laborious operation as the projector is not capable to capture images, thus leading to lower overall accuracy. Anyway the methods discussed below can be found in both configurations.

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

$$z = \frac{d}{\tan \theta} \quad (6)$$

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

The working principle of structured light is the same of the point gauge but, to match multiple points simultaneously, complex patterns are projected instead of a single spot of light. A naive way to speed up the process would be to project n stripes. That will divide the required time by n but it will cause ambiguity, since the same pattern is repeated multiple times (Fig. 9).

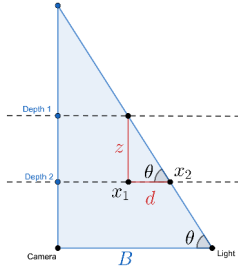


Figure 8: Also point gauges make one-dimension measures based on similar triangles.

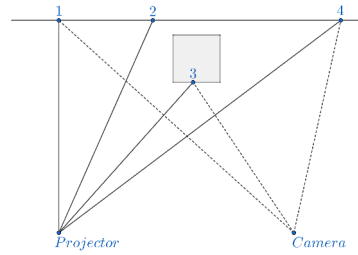
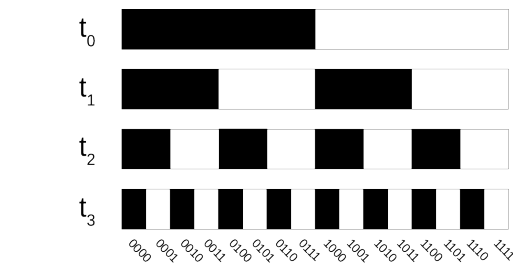


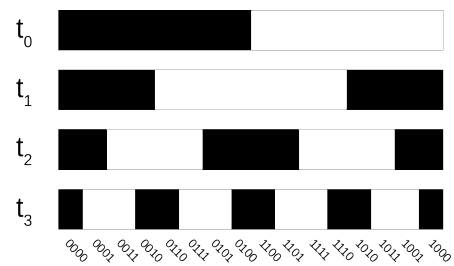
Figure 9: Projecting an identical stripe multiple times would cause ambiguity from the camera perspective.

Approaches to resolve ambiguity follow *temporally* or *spatially* coding. An effective temporally coded method is *binary coding* [19]: since the correspondence problem is one-dimensional and simplified after rectification, the aim is to represent each vertical stripe on the projector with a unique code. The code is realised by projecting a sequence of black and white patterns, so that the camera knows exactly which vertical stripe the pixel belongs to. In order to do this, binary encoded stripes are projected sequentially (Fig. 10a), such that the stripes have a crescent number from 0 to the maximum value. The resolution is limited to half the width of the smaller stripe.

Higher the number of unique codes, higher the accuracy of the calculated depth, up to physical limits. In fact, one of the problems is that transition pixels create uncertainty areas for the camera: stripes can not be too thin as the camera would perceive blurred borders. A clever solution is then to minimise transitions in projected patterns, so the *Gray encoding* was developed (Fig. 10b). Gray code is less prone to errors due to the lower number of transition lines, still maintaining the same number of projected images and corresponding unique codes generated. Moreover, the width of the stripes is bigger for equal coding capacity, facilitating analysis on oblique surfaces and, since each neighbour code has only 1 bit variation, any wrong decoding introduces a misplacement of one unit only.



(a) A typical binary encoding that assigns a unique code to each vertical stripe.



(b) Gray encoding minimises transitions, thus is less prone to errors.

Figure 10: Stripes encoding commonly used.

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

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

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

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

Pure spatially coding requires larger alphabet, so that a single projected pattern is enough to encode the image. A good example of it is the *De Bruijn sequence* method [21, 22]. Apart from using colors to increase k , it introduces the concept of using neighbour stripes to uniquely determine the associated code, so the algorithm is good at identifying each stripe uniquely, e.g. the couple red-green occurs only once (Fig. 11).



Figure 11: An example of De Bruijn sequence of order 2 over an alphabet of 4 symbols mapped to 4 colours with no repetitions. A sequence of 2 colors uniquely identifies a position.

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

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

Consider a single sinusoidal fringe pattern represented as:

$$I(x, y) = I'(x, y) + I''(x, y) \cdot \cos[\phi(x, y)] = I'(x, y) + I''(x, y) \cdot \frac{[e^{j\phi(x, y)} + e^{-j\phi(x, y)}]}{2} \quad (7)$$

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

$$\tilde{I}(x, y) = I''(x, y) \cdot \frac{e^{j\phi(x, y)}}{2} \quad (8)$$

From which the phase can be calculated as:

$$\phi(x, y) = \arctan \left\{ \frac{\text{Im}[\tilde{I}(x, y)]}{\text{Re}[\tilde{I}(x, y)]} \right\} \quad (9)$$

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

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

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

Line laser scan technology involves building up an image one line at a time. The line sensor is usually composed by a planar laser and a camera. It slowly moves over the object (or, viceversa, the object moves under the sensor) following known positions on the scanning line. For each position the triangulation is trivial as there will be only one point (along the baseline direction) corresponding to the projected laser line. So, if time is not an issue and the system is well calibrated, *spacetime* analysis [24] can be used. Often, as in many laser optics applications, the laser beam is assumed to have an irradiance profile that follows an ideal Gaussian distribution⁶ [25]. Given the above assumption, for each image the camera will perceive only a few points of a Gaussian curve. With these points the position of the middle point is exactly assigned and used for triangulation. The advantage of this method over the previous ones is that it leads to *sub-pixel* accuracy, however it requires a substantial amount of time during which the object must stay completely still [26].

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

3.2.1 Pros and cons

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

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

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

Classic scanning systems oriented to manufacturing and reverse engineering usually operate in a controlled environment, with the object staying still and having relatively small dimensions. On the contrary, aircraft inspections are carried on in unconstrained environments and over large surfaces, so the compromise between accuracy and speed may be difficult to achieve.

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

4 Time of flight

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

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

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

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

Continuous modulation is the basis of most high accuracy TOF cameras currently on the market. It uses a sinusoidal modulated light source and measures the phase shift of the reflected wave, since the speed in a medium is known very precisely in different conditions and the phase shift directly relates with time shift [27, 28]. The emitted signal frequency f to set depends on the maximum distance to measure λ , also considering the return journey [29]:

$$v = \lambda \cdot f \quad \Rightarrow \quad f = \frac{v}{\lambda} \quad (10)$$

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

The third method is to provide on/off loops transmitting a 50% duty cycle square wave (stroboscopic light source) and record the amount of returned light in a specific interval using CMOS imaging sensors, called *smart pixels*. These

are particular LEDs having two capacitors per pixel [28, 30]. The capacitors are alternately connected to the sensor, using the same clock source of the stroboscopic light. This arrangement allows to directly relate the differential charge in the capacitors to the phase offset, then to the distance as seen above.

4.1 Pros and cons

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

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

5 Thermography

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

$$\lambda_{max} = \frac{b}{T} \quad (11)$$

where $b = 2.897\,771\,955 \dots \times 10^{-3}$ m K is the Wien's displacement constant and T is the temperature of the black body. For example, a black body at a temperature of 300 K has its peak of emitted energy at $\sim 9.6\,\mu\text{m}$, while a temperature of 400 K corresponds to $\sim 7.2\,\mu\text{m}$.

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

The standard sensing structure is made by Focal Plane Arrays (FPA) that measure the excitation generated by incident photons as change of electrical conductivity or by measuring a generated voltage. Some accurate but expensive technologies require cooling, while uncooled ones are more cheaper and compact, hence more popular.

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

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

5.1 Pros and cons

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

- Difficulties to obtain a uniform thermal stimulation over large surface. Defects seen by thermal images usually have high variability [38] due to different environmental temperature (e.g. outdoor), materials, perspective, etc.
- Effect of thermal losses which induces unwanted contrasts (noise).
- Ability to inspect materials of a limited thickness.
- In general, limited to defects that cause a measurable change of thermal properties (e.g. cracks are detected only if they induce thermal resistance) [34].

6 Conclusions

The aim of this paper is to consider the *suitability for automation* of aircraft inspections. The previous sections introduced the working principles of the major non-contact NDI methods. Their performance as reported in literature cannot be directly compared because of different equipment, conditions and inspected materials. A *qualitative* comparison is summed up (Fig. 12) with respect to speed, cost-effectiveness and ease of use and best use-case accuracy⁷.

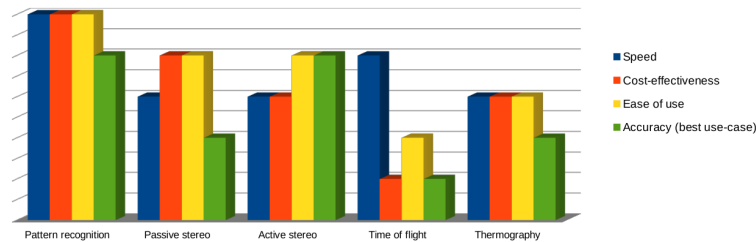


Figure 12: Indicative features of the technologies discussed.

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

Different NDI technologies work differently with respect to main damage types (Fig. 13).

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

Figure 13: Comparison of technologies with respect to types of damage.

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

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

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

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

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

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

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

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

References

- [1] European Union Aviation Safety Agency. *Clarification of Definitions for General Visual (GVI), Detailed (DET), and Special Detailed (SDI) Inspections*. 2012. URL: <https://www.easa.europa.eu/sites/default/files/dfu/IP122%20Rev%202.pdf> (visited on 2019-10-22).
- [2] Lawrence Cook. "Visual inspection reliability for composite aircraft structures". In: (2009).
- [3] L. Cook et al. "Reliability of Damage Detection in Advanced Composite Aircraft Structures". In: (2013).
- [4] Feng Gao. "Interferometry for Online/In-Process Surface Inspection". In: 2017. ISBN: 978-953-51-2955-4. DOI: 10.5772/66530.
- [5] Touba Malekzadeh et al. *Aircraft Fuselage Defect Detection using Deep Neural Networks*. 2017. URL: <https://arxiv.org/abs/1712.09213> (visited on 2019-10-22).
- [6] Kuniyiko Fukushima. "Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position". In: *Biological Cybernetics* (1980).
- [7] Brandon Rohrer. *How do Convolutional Neural Networks work?* Creative Commons CC0 1.0 Universal Licence. 2016. URL: http://brohrer.github.io/how_convolutional_neural_networks_work.html (visited on 2019-10-23).
- [8] Kevin G. Harding. *Practical Optical Dimensional Metrology*. Bellingham, Washington: SPIE, 2019.
- [9] R. Umesh Dhond and J. K. Aggarwal. "Structure from Stereo - A Review". In: *Systems, Man, and Cybernetics, IEEE Transactions on* (1989).
- [10] Zhengyou Zhang. "A Flexible New Technique for Camera Calibration". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 22 (2000-12), pp. 1330–1334. DOI: 10.1109/34.888718.
- [11] M. Pilu. "Uncalibrated stereo correspondence by singular value decomposition". In: *HP Laboratories Technical Report* (1997-01), pp. 3–11.
- [12] L. Nalpantidis et al. "Efficient hierarchical matching algorithm for processing uncalibrated stereo vision images and its hardware architecture". In: *IET Image Processing* (2011). URL: <https://doi.org/10.1049/iet-ipr.2009.0262>.
- [13] Jan Kallwies and Hans-Joachim Wuensche. "Effective Combination of Vertical and Horizontal Stereo Vision". In: 2018-03, pp. 1992–2000. DOI: 10.1109/WACV.2018.00220.
- [14] Michael Bleyer and Christian Breiteneder. "Stereo Matching—State-of-the-Art and Research Challenges". In: *Advanced Topics in Computer Vision*. Ed. by Giovanni Maria Farinella, Sebastiano Battiato, and Roberto Cipolla. London: Springer London, 2013, pp. 143–179. ISBN: 978-1-4471-5520-1. DOI: 10.1007/978-1-4471-5520-1_6. URL: https://doi.org/10.1007/978-1-4471-5520-1_6.
- [15] Daniel Scharstein and R. Szeliski. "A Taxonomy And Evaluation Of Dense Two-Frame Stereo Correspondence Algorithms". In: *International Journal of Computer Vision - IJCV* 47 (2000-01), pp. 7–42.
- [16] Yuri Boykov, Olga Veksler, and Ramin Zabih. "Fast Approximate Energy Minimization via Graph Cuts". In: *IEEE Trans. Pattern Anal. Mach. Intell.* (2001), pp. 1222–1239. ISSN: 0162-8828. URL: <https://doi.org/10.1109/34.969114>.
- [17] Song Zhang and Peisen Huang. "Novel method for structured light system calibration". In: *Optical Engineering - OPT ENG* 45 (2006-08). DOI: 10.1117/1.2336196.
- [18] Xiaobo Chen et al. "Accurate calibration for a camera–projector measurement system based on structured light projection". In: *Optics and Lasers in Engineering* 47 (2009-03), pp. 310–319. DOI: 10.1016/j.optlaseng.2007.12.001.
- [19] Song Zang. "High-speed 3D shape measurement with structured light methods: A review". In: *Optics and Lasers in Engineering* (2018). URL: <https://doi.org/10.1016/j.optlaseng.2018.02.017>.
- [20] Daniel Scharstein and Richard Szeliski. "High-Accuracy Stereo Depth Maps Using Structured Light". In: *Proceedings of the CVPR 2003* (2003). URL: <https://doi.org/10.1109/CVPR.2003.1211354>.
- [21] Jordi Pages, Joaquim Salvi, and Josep Forest. "A new optimised De Bruijn coding strategy for structured light patterns". In: (2004-08). DOI: 10.1109/ICPR.2004.70.
- [22] Joaquim Salvi et al. "A state of the art in structured light patterns for surface profilometry". In: *Pattern Recognition* 43.8 (2010). ISSN: 0031-3203. URL: <https://doi.org/10.1016/j.patcog.2010.03.004>.
- [23] Song Zhang. "Absolute phase retrieval methods for digital fringe projection profilometry: A review". In: *Optics and Lasers in Engineering* 107 (2018).
- [24] Brian Curless and Marc Levoy. "Better Optical Triangulation through Spacetime Analysis". In: *ICCV*. 1995.

- [25] Rüdiger Paschotta. *Encyclopedia of Laser Physics and Technology*. Wiley, 2008. ISBN: 978-3-527-40828-3.
- [26] Marc Levoy et al. “The Digital Michelangelo Project: 3D Scanning of Large Statues”. In: *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '00. USA: ACM Press/Addison-Wesley Publishing Co., 2000, pp. 131–144. ISBN: 1581132085. DOI: 10.1145/344779.344849. URL: <https://doi.org/10.1145/344779.344849>.
- [27] Robert Lange and Peter Seitz. “Solid-State Time-of-Flight Range Camera”. In: *IEEE Journal of Quantum Electronics* 37 (2001).
- [28] Bernhard Buetten et al. “CCD/CMOS lock-in pixel for range imaging: Challenges, limitations and state-of-the-art”. In: (2005-01).
- [29] A Dorrington, Andrew Payne, and Michael Cree. “An evaluation of time-of-flight range cameras for close range metrology applications”. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII* (2010).
- [30] Mark Hughes. *How Do Time of Flight Sensors (ToF) Work?* 2019. URL: <https://www.allaboutcircuits.com/technical-articles/how-do-time-of-flight-sensors-work-pmdtechnologies-tof-3d-camera/> (visited on 2019-11-27).
- [31] S.A. Guomundsson, Henrik Aanæs, and Rasmus Larsen. “Environmental Effects on Measurement Uncertainties of Time-of-Flight Cameras”. In: 2007-08. DOI: 10.1109/ISSCS.2007.4292664.
- [32] Dragos Falie and Vasile Buzuloiu. “Noise Characteristics of 3D Time-of-Flight Cameras”. In: 1 (2007-06). DOI: 10.1109/ISSCS.2007.4292693.
- [33] Xavier P. Maldague. *Theory and Practice of Infrared Technology for Nondestructive Testing*. John-Wiley & Sons, 2001.
- [34] Xavier P. Maldague. “Introduction to NDT by active infrared thermography”. In: *Materials Evaluation* 60 (2002-09).
- [35] D. Wu et al. “Inspection of aircraft structural components using lockin-thermography”. In: *Quantitative Infrared Thermography* (1997-01).
- [36] D. Wu and G. Busse. “Lock-in thermography for nondestructive evaluation of materials”. In: *International Journal of Thermal Sciences* (1998).
- [37] Pierre Servais. “Development of a New NDT Method Using Thermography for Composite Inspection on Aircraft with Portable Military Thermal Imager”. In: 2006.
- [38] Robert J. Ducar. “Pulsed thermographic inspection and application in commercial aircraft repair”. In: *Thermosense XXI*. Ed. by Dennis H. LeMieux and John R. Snell Jr. Vol. 3700. International Society for Optics and Photonics. SPIE, 1999, pp. 77–83. DOI: 10.1117/12.342277.
- [39] Yang Zhongyuan et al. “Detection of morphology defects in pipeline based on 3D active stereo omnidirectional vision sensor”. In: *IET Image Processing* (2018). DOI: 10.1049/iet-ipr.2017.0616.
- [40] Priyan Gunatilake et al. “Image enhancement and understanding for remote visual inspection of aircraft surface”. In: *Nondestructive Evaluation of Aging Aircraft, Airports, and Aerospace Hardware*. Ed. by Raymond D. Rempt and Alfred L. Broz. Vol. 2945. International Society for Optics and Photonics. SPIE, 1996, pp. 416–427. DOI: 10.1117/12.259117. URL: <https://doi.org/10.1117/12.259117>.



DSSPACE

<https://dspace.org/>

Review of non-contact methods for automated aircraft inspections

Lafiosca, Pasquale; Fan, Ip-Shing

2020-12-01

British Institute of Non-destructive Testing

<https://doi.org/10.1784/insi.2020.62.12.692>

Lafiosca P, Fan IS. (2020) Review of non-contact methods for automated aircraft inspections. Insight: Non-Destructive Testing & Condition Monitoring, Volume 62, Issue 12, December 2020, pp. 691-701