

Method For Visual Inspection And Apparatus Thereof

Field

5 The invention relates to a method for visual inspection and an apparatus thereof, in particular, the visual inspection involves detecting a defect on an object.

Background

10 Computer vision is a sub field of artificial intelligence that investigate how to make computer algorithms perceive and understand the world of a human through images. During an emulation, light is detected by cameras and/or sensors, while high-level of image understanding is processed by using computer algorithms. There are many applications where computer vision plays an important role. For instance, computer vision can be used for range sensing, industrial inspection of manufactured parts and/or parts that are currently in use and susceptible to wear and tear, as well as object recognition.

15 Visual inspection of surfaces for defects can cover both geometry (dimension) inspection and surface (texture) inspection of the surfaces. Dimension inspection is based on geometric shapes and typically captured using a positioning device holding a geometry sensor (a depth sensor, or a 3D scanner and etc.), while surface (texture) inspection is based on images. Sensor planning, i.e. identification of sensor locations to capture different parts of an object with required specifications is important for acquiring desired geometric shapes for dimension inspection. On the other hand, surface inspection relies on the imaging quality of camera(s) used. Different views of a target region may need to be acquired for analysis to identify the existence of a defect.

20 While some inspection systems can capture both 3D geometry and 2D images to enable the dimension inspection and surface inspection, the viewpoints for all these systems are predefined before the sensing, e.g. with active sensor planning. For example, in surface inspection of objects of the same type, manual selection is required to define viewpoints for image capturing, while in dimension inspection systems, CAD-based active sensing is adopted to derive viewpoints for capturing the 3D geometry. Multi-view based inspection systems may acquire inspection data by combining both shapes and images from different viewpoints.

30 In addition, computer vision is particularly valuable in 3D construction and/or reconstruction because computer vision can be used to obtain a three-dimensional position of an object point from its projective points on the image plane. A 3D construction and/or reconstruction is necessary in applications where accuracy of the 3D shape represents a fundamental factor to the performance of a computer vision device.

35 In conventional methods, high-end 3D scanners are required to achieve a good 3D reconstruction of a target object under inspection. There are also strict requirements to the positioning of the target object and the quality of images captured for the target object. This means that the set up of the scanner has to be precise and it will take a lot of time and effort to set up before different type of defects can be detected and/or a different type of object can be inspected.

Summary

40 According to an example of the present disclosure, there are provided a method and apparatus as claimed in the independent claims. Some optional features are defined in the dependent claims.

Brief Description of Drawings

45 In the drawings, the same reference numerals generally relate to the same parts throughout different views, unless otherwise specified. Embodiments of the invention will be better understood and readily apparent to one skilled in the art from the following written description, by way of example only and in conjunction with the drawings briefly described as follows.

Figure 1A shows how a crack is observed in a hash-outlined box and how a crack is occluded in a black-outlined box in one captured photograph.

50 Figure 1B shows how a crack is occluded in the hash-outlined box of Figure 1A in another captured photograph.

Figure 1C shows how a crack is observed in the black-outlined box of Figure 1A in another captured photograph.

Figure 2 shows a flow chart of how a method for performing single-view mobile inspection of defects can be provided according to one example of the present disclosure.

Figure 3 is a geometry illustration for detection of a hole crack according to one example of the present disclosure.

Figure 4A shows a test image comprising a crack region that is marked by a visual inspector.

Figure 4B is an exploded view of the crack region in Figure 4A.

5 Figure 4C shows a plurality of detected cracks in the test image according to one example of the present disclosure.

Figure 5 shows a flow chart of how classification-based single-view inspection is learnt from training images according to one example of the present disclosure.

10 Figure 6A shows how a plurality of image patches are obtained from a captured image according to an example of the present disclosure.

Figure 6B shows how erosion/corrosion regions can be indicated as on a captured image according to one example of the present disclosure.

Figure 7 shows a flow chart for single-view inspection trained with a 2D reference image according to one example of the present disclosure.

15 Figure 8 shows a flow chart for single-view inspection trained with a 3D reference model according to one example of the present disclosure.

Figure 9A shows a sensor planning system for computer vision according to one example of the present disclosure.

20 Figure 9B shows a comparative example of an existing model of sensor planning system for computer vision.

Figure 10 shows a work flow of a proposed Snap-to-inspect three Dimensional system (Snap2Inspect3D) according to one example of the present disclosure.

Figure 11 shows a hardware set up for capturing an image for visual inspection with active 3D measurement according to one example of the present disclosure.

25 Figure 12 shows how system calibration can be formulated according to one example of the present disclosure.

Figure 13A shows an example of a coded pattern according to one example of the present disclosure.

Figure 13B shows how a portion of the coded pattern in Figure 13A can be decoded.

Figure 13C shows how a cylinder is reconstructed after decoding the coded pattern in Figure 13A.

30 Figure 13D show an example of a checkerboard pattern that can be used for system calibration according to one example of the present disclosure.

Figure 13E show an example of asymmetrical circle patterns that can be used for system calibration according to one example of the present disclosure.

35 Figure 14 shows how a projection pattern can be generated for defect measurement according to one example of the present disclosure.

Figure 15 is a screenshot of system calibration with asymmetrical circle patterns according to one example of the present disclosure.

Figure 16 is an example of a system calibration result after system calibration in Figure 15.

40 Figure 17 is an example of height measurement of a rectangular object on a surface using Snap2Inspect3D.

Figure 18 is a flow chart on how Snap2Inspect3D operates according to one example of the present disclosure.

Figure 19 is a block diagram illustrating a system architecture according to an example of the present disclosure.

45

Detailed Description

Visual inspections expect quick and accurate evaluation of visual contents for significant abnormalities, such as defects, on a 3D object surface. Occlusions and/or shadings may hide defect features from adequate observations and detection. Figures 1A to 1C are images of a blade surface taken from different views, and these figures illustrate how occlusion and/or shading can hide defect features. For example, the crack in hashed-outlined box 101 can be observed in Figure 1A, but hardly seen in Figure 1B, and the crack in black-outlined box 102 can be clearly observed in Figure 1C, but is very weak in Figure 1A. Hence, it is important to use an appropriate view for optimal defect detection.

55 Mobile (portable or robotic) imaging devices enable flexible observation settings (angles, distances, etc.). A quick snap or image capture of the object surface without intentional planning using a mobile device is likely to result in a sub-optimal observation of defects. Handheld devices (such as a smart phone or an imaging device that can be held by a user for capturing images) are even more likely to generate sub-optimal observation due to inaccurate positioning. As a result, the picture(s) taken by a handheld device may not be at an optimal viewing angle and distance (or optimal

60

observation settings) to distinguish an abnormality from the background or other content on a target object. This technique imposes challenges to the demand of excellent feature representation and excellent detection algorithms which can detect abnormalities under various imaging and lighting conditions.

5 A further challenge is resolving the scale of the target object where visual inspection tasks require measurements in absolute scale.

A solution for mobile inspection of defects assisted by a variety (or different types) of visual knowledge is provided to address the challenges described above.

10 "Single view inspection" used herein refers to inspection using only a single view for capturing three dimensional (3D) geometry and/or two dimensional (2D) images to enable dimension inspection and surface texture inspection. This can be referred as "one-shot scan" in the present disclosure.

"Portable" used herein refers to a lighter and smaller version (than usual) of an object that allows it to be easily carried or moved or transported around from one place to another place by a user.

15 "Engine blade" used herein may refer to a turbine blade of a gas turbine or a steam turbine, or may refer to fan blades or propellers used in an automotive vehicle (such as a jet), industrial condensers, industrial fans and the like.

In one exemplary technique, there is provided a method and a device for single-view mobile inspection of defects assisted by different types of visual knowledge. In one aspect of the exemplary technique, visual knowledge about the defect characteristics, such as training images or reference image, is acquired as part of input of the defect detection algorithms. In another aspect of the exemplary technique, defects are detected from individual single-view images captured by an image capturing device, preferably hand held by a user conducting the inspection, and concurrently, a confidence score is calculated, and returned as feedback to the user. In another aspect of the exemplary technique, better inspection results are incrementally achieved when imaging parameters are tuned, preferably by the user.

20 In another exemplary technique, there is provided a method, (vision) system and/or apparatus for performing single-view mobile inspection of defects that is guided by different types of visual knowledge. These method, system and/or apparatus are termed in the present disclosure as "snap-to-inspect", "snap2inspect2D" or "snap2inspect". An extension of these method, system and/or apparatus is termed as "snap2inspect3D". A flow chart illustrating the operation principles of the snap-to-inspect example is illustrated in Figure 2.

With reference to Figure 2, an image capturing unit (e.g. camera, sensor and the like) of a mobile device (such as smartphone) or an image capturing device is instructed to capture a first image 204 of a target object. For instance, such instruction is inputted by a user 202 or issued by a command received from a control unit of the mobile device. Defect inspection algorithms 208 are used to process the image 204 to detect whether the target object has a defect 210. These defect inspection algorithms are guided by different types of visual knowledge 206, which will be elaborated later. A confidence score 215 for the detection on whether the target object has the defect 210 is computed and fed back to the user 202 to assist the user 202 or the control unit to capture one or more images (or snapping one or more images using a camera) with more optimal imaging parameters (angle, distance, etc.) while performing different automatic inspection tasks (e.g. object recognition, scene reconstruction, feature detection and the like). This means that the image capturing device determines, based on the calculated confidence score 215, whether to generate better view recommendations 220 that may include improved imaging parameters with respect to translation and/or rotation movement of the imaging capturing device. The improved imaging parameters may be generated, if required, after considering the calculated confidence score. In one example, upon receiving an input, the image capturing device is instructed to capture, according to the improved imaging parameters, a second image of the object that results in calculation of an improved confidence score. With collation of confidence scores and observation of changing trends from different images captured by the image capturing unit on a common defect on the object, better imaging observations are incrementally achieved for each specific inspection task, thereby facilitating better detection performance. This means that the collated confidence scores can be used to fine tune detection of the common defect on the object or another object (of a same type as the object).
50 "Better imaging observations" can refer to better imaging quality achieved through application of improved imaging parameters with respect to translation and/or rotation movement of the imaging capturing device.

55 In another example, the calculated confidence score is generated as feedback to the user after comparing the calculated confidence score against a threshold, wherein the threshold is used to reduce a number of false detection. The improved imaging parameters can be generated after
60

comparing the calculated confidence score against the threshold. Thereafter, upon receiving an input, the image capturing device is instructed to capture, according to the improved imaging parameters, a second image of the object that results in calculation of an improved confidence score

5 A computer program product (i.e. application or program) may be provided to cause the execution of the above-described defect inspection algorithm on a program controlled entity. Such computer program product and its instructions thereof may be stored in a memory accessible to a processor and be executable by the processor. The computer program product may be provided or delivered as a recording means, including a non-transitory computer readable medium, such as a memory card, USB-stick, CD-ROM, DVD or also in form of a downloadable file from a server in a network. This can, for example, be achieved through a wireless communication network by the transmission of a respective file having the computer program product. The computer program product includes a mobile application tied to an operating system of a smartphone.

10 In one example, a decision-theory (also known as by-contrary decision) based single-view inspection without reference images is provided where an observed geometric pattern might occur in a random image, e.g. the detection of a rare event such as a crack on the inspected surface, is based on decision theory to produce a confidence score that represents defect pattern strength (e.g. crack strength). In this example, the inspection may exploit generic knowledge that is learned beforehand in the decision-theory based single-view inspection method. "Generic knowledge" used herein refers to factual information or concepts in computer visioning system for detection of defect(s) in an object.

20 In another example, a classification-based single-view inspection with training images is provided where visual features can be used to classify image patches into normal or abnormal (such as non-erosion or erosion regions), based on learned knowledge from training data to produce a confidence score that represents class membership probabilities.

25 In yet another example, a single-view inspection with 2D reference image is provided where missing materials or geometry defects are detected by comparing the real (detected) shape against the ideal shape mapped from a reference image that has no defect to produce a confidence score that represents registration quality (e.g. number of matched feature points after registration).

30 In a further example, a single-view inspection with 3D reference model is provided where missing materials/geometry defects are detected by comparing the real (detected) shape against the ideal shape registered/projected from a reference 3D model that has no defect to produce a confidence score that represents registration quality (e.g. number of matched feature points after registration).

35 For a given visual inspection task (based on a selected mode of operation), visual features are learned based on generic knowledge or from training data regarding defects. The visioning system according to snap2inspect comprises a plurality of operation modes for selection, wherein each operation mode is for detecting one type of defect, and the processing of a captured image of an object to detect whether the object has the defect is performed according to the selected operation mode. The plurality of operation modes may include at least two of the following operation modes:

- 40 • an operation mode to detect a crack on the object using a decision-theory based algorithm (Refer to "**1. Single-view inspection without training images**" below for details);
- an operation mode to detect erosion on the object using a classification algorithm that trains on a plurality of training images (Refer to "**2. Classification-based single-view inspection with training images**" below for details);
- 45 • an operation mode to detect missing material on the object using an algorithm that matches feature points of a reference 2D image with feature points of a captured image of the object (Refer to "**3. Single-view inspection with a 2D reference image**" below for details); and
- an operation mode to detect missing material on the object using an algorithm that matches features points of a 3D model with feature points of a captured image of the object (Refer to "**4. Single-view inspection with a 3D reference model**" below for details).

50 Several examples of defect detection by selecting one of the plurality of operation modes are shown in Table I.

Table I. Utilization of visual knowledge with respect to defect types

Visual knowledge	Type of defect	Detection method	Confidence score
Generic knowledge	Crack	by contrary decision	crack strength
Training data	Erosion	feature extraction and classification	class membership probabilities
2D reference image	missing material	2D image registration	registration quality, e.g., number of matched feature points
3D model	missing material	3D model projection to the respect image	registration quality, e.g., number of matched feature points

1. Single-view inspection without training images

5

In one example, there is provided a method and apparatus for single-view inspection without reference images (or training images). The method is conducted based on a “*a contrario*” statistical framework (also known as ‘by contrary decision’). In order to observe a meaningful (or significant) geometric event/pattern, the occurrence expectation should be small in the image. This means that a geometric pattern is considered meaningful in the “*a contrario*” framework if expectation of its occurrence would be very small in a random image. The lower the probability of such event, the more meaningful the geometric pattern is. This means that the expected occurrence of such event (represented by ϵ) can be set as one for a random image to reduce false detection (of defects) in the image. It should be appreciated that ϵ can be inputted by a user or predefined by the apparatus.

Mathematical expectation is the summation or integration of a possible value from a random variable. It is also known as the product of the probability of an event occurring, denoted as $P(x)$, and a value corresponding with an actual observed occurrence of the event. The expected value is a useful property of any random variable. The mathematical expectation is usually notated as $E(X)$ and such expected value can be computed by overall summation of distinct values that the random variable can take. The mathematical expectation will be given by mathematical formula $E(X) = \sum (x_i p_i, x_2 p_2, \dots, x_n p_n)$, where x is a random variable with a probability function, $f(x)$, p is the probability of the occurrence, and n is the number of all possible values.

Figure 3 illustrates an example of a crack on a target object (e.g. engine blade) represented by a discrete line segment $c: \{p, r, \alpha\}$, characterized by its starting point p (from hole), its length $r \in \mathbb{N}$ and its direction α .

In order to be robust to shading changes on the target object that is being captured by a portable imaging device (or handheld imaging device), the intensity image is locally normalized by dividing it by its median on a small neighborhood. That is, for $q = (x, y)$, a normalized intensity image is defined as

$$\hat{I}(q) = \frac{I(q)}{\text{median}(I(N(q)))}. \quad \text{- Equation (1)}$$

Consequently, strength of a crack, $t(c_p(r, \alpha))$, is defined as the negative sum of normalized intensity of pixels on the crack of the target object. The strength of the crack is represented by mathematical equation (2) below:

$$t(c_p(r, \alpha)) = - \sum_{q \in c_p(r, \alpha)} \hat{I}_p(q) \quad \text{- Equation (2)}$$

Let $\mathcal{C}$ be the set of all possible cracks in a discrete image patch I . For $\epsilon > 0$, a crack c is said to be ϵ -meaningful if its number of false alarm (NFA) under a hypothesis $\mathcal{H}_0$ satisfies

$$NFA(c) := \#\mathcal{C} \cdot F_c(t(c)) \leq \epsilon \quad \text{- Equation (3)}$$

where $F_c(t)$ is the probability that a random variable t is larger than a given threshold. The value ϵ corresponds to an expected number of false detections under the hypothesis $\mathcal{H}_0$, where $\mathcal{H}_0$ is a stochastic model for unstructured data. The smaller the quantity $NFA(c)$ is, the more meaningful the

crack c . Note that a detection method is said to produce ϵ -meaningful detections (events), when the expected number of false detections is bounded by ϵ . See definition of $\#C$ in Equation (4) below.

To search for a candidate of crack, the starting point $\mathbf{p}$ is set in the $w \times w$ window (depicted as square in Figure 3) and centered at a lower vertex $\mathbf{o}$ on a major axis 304 of an ellipse 302 representing a hole. The $w \times w$ window represents an image patch I of a captured image (e.g. 204 of Figure 2). Details on how an image patch of a captured image can be obtained will be discussed later. O_1 to O_5 are points on a boundary of the ellipse 302. In one example, each of these points O_1 to O_5 can be used as an origin to generate fan-shape scanning regions for crack detection.

The length of crack r ranges discretely from $r_{\min}$ to $r_{\max}$: $r \in [r_{\min}, r_{\max}]$. The crack direction α is within a $\pi/2$ range from the ellipse orientation: $\alpha \in [\varphi - \frac{\pi}{4}, \varphi + \frac{\pi}{4}]$. The number of discrete direction at a given crack length r is set to be $K(r) = \frac{\pi}{2}r$. Therefore, the number of all possible cracks in the image patch I is

$$\#C = \#\{\mathbf{p} \in \mathcal{N}_w(\mathbf{o})\} \sum_{r=r_{\min}}^{r_{\max}} K(r) = w^2 \sum_{r=r_{\min}}^{r_{\max}} \frac{\pi}{2}r \quad \text{- Equation (4),}$$

where $\#C$ is the number of all possible cracks in an image patch I , $\#\{\mathbf{p} \in \mathcal{N}_w(\mathbf{o})\}$ is the number of a probability that starting point $\mathbf{p}$ lies on the lower vertex $\mathbf{o}$ of the ellipse, and $\sum_{r=r_{\min}}^{r_{\max}} K(r)$ is the summation of the discrete directions.

To compute the distribution of the strength $t(c)$ of a crack, empirical distribution of $\hat{I}(q)$ is estimated from its histogram. Thereafter, the law of the strength of a crack $t(c_p(r, \alpha))$, i.e. sum of the r pixel contributions, is obtained by convolving r copies of the distribution μ .

$$F_c(t) := \mathbb{P}_{\mathcal{H}_0}[t(c) \geq t] = \int_t^{+\infty} d \left(\begin{matrix} r \\ \star \\ j=1 \end{matrix} \mu \right) \quad \text{- Equation (5),}$$

where $\star$ denotes the convolution operator.

In this exemplary technique, a crack is detected as ϵ -meaningful in the image patch I if its strength $t(c)$ is larger than the threshold, $\hat{t}(r, \epsilon)$, defined in Equation 6.

$$\hat{t}(r, \epsilon) := \min \left\{ t; F_c(t) \leq \frac{\epsilon}{\#C} \right\} \quad \text{- Equation (6),}$$

The value of ϵ can be selected as 1 and a different threshold on $t(c)$ can be yielded for each value of the crack length r . Among all the ϵ -meaningful cracks (if there are any) in the image patch I , the one with the highest strength is considered as a detected crack. An example of crack detection result is shown in Figures 4A to 4C. Figure 4A is a test image for detecting cracks according to one example of the present disclosure. On region 400 is marked out as an example. Figure 4B is an exploded view of the region 400 of Figure 4A that is marked by an inspector to indicate a crack in several apertures 402 and it matches with a detection result in Figure 4C that is indicative of whether the test image has a defect. Figure 4C shows computed crack lines 404 in the several apertures 402.

2. Classification-based single-view inspection with training images

In the following paragraphs, a classification-based single-view inspection of defects with training images is discussed. Colour, texture, and Histogram of Gradient features are calculated from the sample images and later classified by using Support Vector Machines (SVM). A SVM is a discriminative classifier formally defined by a separating hyperplane. In other words, given labelled training data (supervised learning), the algorithm outputs an optimal hyperplane which categorizes the training data according to their labels. A SVM model is a representation of the examples (training data) as points in space, mapped so that the examples (training data) of the separate categories are divided by a clear gap that is as wide as possible. New examples (test data) are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall.

Some examples of the feature vectors that can be extracted from sample image patches are listed in Table II.

Table II. Features extracted from sample images patches

Feature	Dimension
HSV color histogram	32
Color auto correlogram	64
Color moments	6
Gabor wavelet	48
Wavelet moments	40
Histogram of Gradient	81
Total	271

Using the SVM trained by the predetermined visual features of sample images, image patches are classified into normal/abnormal (e.g., erosion or non-erosion) regions based on learned knowledge. In this case, confidence score is given by class membership probabilities. An image patch described in the present example refers to a portion (usually a rectangle window) of a training image or a test image that is being processed for detection of a defect. The patches are automatically cropped from the original training image or the original test image with some overlaps between each patch in a sliding-window way. As illustrated with reference to Figure 6A, a plurality of image patches (e.g. 601, 602, 605, 606 and 607) are cropped from a test image 610 such that each image patch has an overlapping portion with at least two image patches proximate to it. As an example, there is an overlap 603a between image patch 601 and image patch 602, and an overlap 604a between image patch 601 and image patch 605. Likewise, there is also an overlap 603b between image patches 602 and 606 and an overlap 603c between image patches 606 and 607. If an image patch contains a detected defect, the image patch may be highlighted (for example, displaying as a black-outlined box or a hashed-outlined box) on the original test image for viewing.

The flowchart of classification-based single-view inspection with training images is shown in Figure 5. Figure 5 indicates that training data are provided to an image capturing device (visioning system) at a step S502, and the predetermined features (such as colour, texture, and histogram of gradient) are extracted out of each training image obtained from the training data at a step S504. The extracted features of the respective training images are sent to a classifier (e.g. a support vector machine) to learn before grouping each training image accordingly at a step S506. The image capturing device is configured to output and monitor a classification result indicating whether a sample object comprises a defect and/or a number of false detections at a step S512. The classification result may be used to fine-tune the rules for classifying the training images and/or test images. Preferably, the image capturing device is configured to compute a confidence score 215 from the classification result as a feedback to the image capturing device and/or a user.

The image capturing device processes the testing data in a similar fashion as the training images. For instance, testing data are provided to the image capturing device at a step S508, and the predetermined features (such as colour, texture, and histogram of gradient) are extracted out of each captured image obtained from the testing data at a step S510. The image capturing device is also configured to compute a classification result indicating whether the test object comprises a defect at the step S506. The image capturing device is configured to output and monitor a classification result indicating whether the test object comprises a defect and/or a number of false detections at a step S512, and to fine-tune the rules to classify the training images and/or test images. Preferably, the image capturing device is configured to compute a confidence score 215 from the classification result as a feedback to the image capturing device and/or a user. An example of erosion detection result is shown in Figure 6B. Image patches indicated on a test image 610 as a plurality of black-outlined boxes in Figure 6B are erosion regions 620.

3. Single-view inspection with a 2D reference image

Geometry defects on a surface of a material, or missing portions from a surface of a material can be detected quickly from a comparison between a 2D image of the surface of the material with a single 2D reference image that has no geometry defect. This is referred to as Single-view inspection with 2D reference image in the present disclosure. In the present example pertaining to Single-view

inspection with 2D reference image, the discussed technique applies to a new image and a reference image that is mono-chromatic or grayscale (instead of a full colour image).

The flowchart of single-view inspection with 2D reference image is shown in Figure 7. Firstly, a rigid transform between a reference image 702 with ideal geometry and a new image 704 is determined by registration of their corresponding key point positions (e.g. edges, corners and the like). Note that image registration is a process of transforming different sets of data into one coordinate system (space). Data may be multiple images, or images captured from different times, viewpoints or sensor. The ideal geometry 706 in the reference image 702 is transformed to a coordinate space of the real geometry 708 and missing materials/geometry defects is detected by comparing features at the corresponding positions.

A confidence score 215 is obtained based on registration quality (e.g., number of matched feature points) as a feedback. Thereafter, better observation (e.g., translation/rotation of hand-held imaging devices) recommendation is given according to the obtained scores and/or score changes over time as more images are captured for capturing a defect.

4. Single-view inspection with a 3D reference model

Geometry defects on a surface of a material, or missing portions from a surface of a material can also be detected by the vision system with a 3D reference model 802 that has no geometry defect. The flowchart of single-view inspection with 3D reference model is shown in Figure 8.

With reference to Figure 8, a new 2D image 804 is registered to the 3D model 802 by registration of their corresponding key point positions (typically represented as n). With sufficient number of corresponding points (i.e. $n \geq 3$), the optimal projective transformation, associated viewing angles and other registration parameters can be derived. This means that the registration will use optimal viewing parameters derived from the corresponded key points. Then, with the optimal registration parameters, the 3D model is projected to a (2D) projection view 806 that corresponds to the respective 2D view of the new image 804. Common techniques known to a skilled person, such as Efficient Perspective- n - Point (EPnP) method developed by Lepetit et al. in their 2008 International journal paper can be used to estimate the 2D projection view 806 that corresponds to the respective 2D view of the new image.

Missing materials and/or geometry defects can be detected by comparing the corresponding positions between a real geometry 808 detected from the new 2D image 804 and an obtained ideal geometry of the projection view 806 in a coordinate space determined by transformation.

A confidence score 215 is calculated based on registration quality (e.g., number of matched feature points). Thereafter, better observation recommendation (e.g. translation and/or rotation movement of an imaging capturing unit, a sensor, a camera or the like) is calculated and generated according to the confidence scores/score changes.

5. Experimental results

The techniques discussed above are implemented (demonstrated) in an application of defect detection of engine blades. Out of the 56 hole crack regions marked out by a human inspector, the crack detection algorithm (without reference images) found 96.4% of them. The detection rate of missing material (with 2D reference image) and corrosion/erosion (with training images) are both 100%.

Table III. Part inspection results

S/N	Defect type	Detection rate	Number of samples
1	Hole crack	96.4%	56 regions in 51 blades
2	Missing material	100%	17 blades
3	Corrosion/erosion	100%	19 blades

The techniques disclosed herein can also be used in imaging sensor planning in robotic vision. The objective of imaging sensor planning in robotic vision can be summarized as follows:

With reference to Figure 9A, if information about the environment (e.g., the object 902 under observation, the available sensors 904) as well as information about the task(s) 906 that a vision system is to accomplish (i.e., detection of certain object features, object recognition, scene reconstruction, object manipulation) is provided during sensor planning, a sensor planning system 900 is able to develop strategies to automatically determine sensor parameter values 910 that can achieve a task with a certain degree of satisfaction. With such strategies, sensor parameters values

910 can be selected and can be purposefully changed in order to effectively perform the task at hand. Sensory systems are then able to operate more flexibly, autonomously, and reliably. The process of sensor planning system is plotted in Figure. 9A.

One existing model of sensor planning that is discussed in International Patent Application No. PCT/SG2017/050175 is illustrated in Figure 9B. The model in Figure 9B comprises three parts: system setup 952 for different types of objects, task specification 954 for objects of the same type and online inspection 956 for each object. The individual steps for inspection 956 are not relevant to the present disclosure and are not elaborated herein. It should be appreciated that various sensing and inspection steps can be used for the online inspection 956 in the existing method of sensor planning.

The system setup 952 consists of the following two parts:

- Sensor calibration: Calibrate different sensors (texture sensor, geometry sensor) into a same coordinate system, i.e. the robotic base coordinate system.
- System configuration: Configure device specifications, which will be considered in a viewpoint calculation. Specification includes:
 1. Geometry sensor: Field of view, clear distance, depth of field, precision for each measurement point, point density at different sensing distance.
 2. Texture sensor: Field of view, Depth of field
 3. Robot: Position accuracy.

The task specification 954 seeks to specify global parameters 960 for objects of the same type, which includes the desired geometry and texture quality. An object model 958 for the objects to be inspected has to be prepared. If sensors and the robot (the actuator) are fixed for inspecting different types of defects on different objects (like in examples of the present disclosure), system configuration remains unchanged, and thus, the sensors and the robot only need to be calibrated once. For inspecting objects of a same type (like in examples of the present disclosure), the task specification can be performed once. If new training data is available, prior knowledge training can be updated to include the new training data.

The differences between Snap2Inspect, which refers to techniques proposed in examples of the present disclosure, and existing sensor planning methods are summarized in Table IV below.

Table IV. Comparison between with Snap2Inspect and existing sensor planning methods.

	Existing sensor planning methods	Snap2Inspect
Image capturing mode	Robotic capturing (require accurate positioning of objects or imaging devices)	Mobile capturing (using a portable, preferably a, handheld device) (only require approximate positioning of objects or imaging devices)
No. of sensors	Multi-view or multi-sensor	Single view or single sensor
Input	Require both sensor model and object model (CAD model).	Inspection decision based on generic knowledge, training image data with or without reference models
Criteria for optimal sensor set-up	Geometry-oriented: Unobstructed view, in focus, properly magnified, well-centered within the field-of-view, imaged with sufficient contrast, optimal target tracking, and other set-ups for obtaining a good image	Defect-oriented: Confidence score of defect inspection (depends on the specific inspection task)
Tasks	Object-level: Feature detection, object recognition, scene reconstruction	Defect-level: Defect inspection

There are two main differences of the techniques disclosed in examples in the present disclosure and the existing sensor planning methods:

1) The criteria for optimal sensor setup in existing sensor planning methods is geometry-oriented (which aims at optimal output geometry quality). In contrast, the criterion for the techniques disclosed herein is defect-oriented which aims at better detection rate.

2) Although the vision tasks to be achieved by existing sensor planning include object recognition and feature detection (which is similar to (at least) one of the tasks implemented through the techniques disclosed herein), existing sensor planning is reliant on low level sensing parameters such as resolution during the actual planning process. In contrast, the techniques disclosed herein focus on defect-level pattern recognition tasks and the corresponding detected result is actually returned as a feedback during the imaging process.

The key features of a proposed solution based on the above exemplary techniques of examples of the present disclosure are as follows

- An automatic and efficient method to present a quick "evaluation" to the inspector (or user) when mobile imaging device is positioned quickly. The 'by contrary decision' approach is extended to the detection of any parameterizable geometric structures. The threshold is automatically determined so that parameter setting is minimal.
- In different operation modes, the input to the visual system can be generic knowledge, training image data or reference images. These different operation modes are non-model based (i.e. no sensor model and/or object model are used), and such non-model-based approaches are used to determine the best next view and sensor settings to incrementally acquire the object information. This is different from conventional sensor planning methods, because these methods use model-based approaches that require sensor model and object model to determine the optimal sensor placements and a shortest path through these viewpoints. Defect-oriented confidence scores are used to guide the inspector (user) to find the optimal imaging parameters and it is applicable to generic visual inspection of objects, structures & architectures. There is no requirement for the sensor set-up to be optimized before detection of defects. The defect-oriented confidence scores are designed for defect-level pattern recognition tasks. However, existing sensor planning methods are geometry-oriented and have to ensure that captured images have unobstructed view, are in focus, are properly magnified, are well-centered within a field-of-view, and imaged with sufficient contrast.

The advantages of the proposed solution are as follows:

- Applicable to generic visual inspection of objects or structures;
- Different operation modes: with or without reference image/model, with or without training data, etc.; and
- Automatic and efficient.

A toolbox based on the proposed solution, is useful for industries that require defect inspection with mobile (especially hand-held) imaging devices. The applications may include:

- Application on mobile imaging device for abnormality and/or defect inspection;
- Assistance to inspector to achieve better inspection performance with interactions; and
- Adaptation to task specific inspection where users can further develop based on teachings of the present disclosure. For example, classification-based single-view inspection with training images can be adapted for detecting other surface defects, such as, discoloration and burning, by using a classifier that is trained by a new set of training images of relevant defects.

6. Snap2Inspect3D

There is also provided a method, apparatus and/or a (visual or vision) system for visual inspection with active 3D measurement. Such method, apparatus and/or a (visual or vision) system is referred herein as Snap2Inspect3D. Once defects in 2D images are detected, for instance, through the snap2inspect or snap2inspect2D techniques disclosed previously, the targeted defect regions can be measured actively with an adaptive illumination pattern. This exemplary technique provides key 3D measurements such as geometry of the inspected surface in absolute scale and profile data of defects (length, width and depth) while keeping the number of image acquisitions low or within to a few snaps. It should be appreciated that this exemplary technique provides a lower cost, and yet faster method to obtain accurate 3D measurement of a detected defect as compared to existing methods.

One limitation in 3D reconstruction is known as correspondence problem, particularly when observing non-textured objects. In order to cope with the correspondence problem, methods based on structured light can be used to create correspondences and give specific code word to every position on an image.

A structured light pattern is projected onto a scene and imposes illusion of texture on an object, thereby increasing the number of correspondence on non-textured objects. Therefore, surface reconstruction is possible when looking for differences between projected and recorded patterns. Coded structured light is preferred because each element of the pattern can be unambiguously identified on the images. The aim of coded structured light is to robustly obtain a large set of correspondences per image independently of the appearance of the object being illuminated and the ambient lighting conditions.

A set of patterns may be successively projected onto a surface. In such an example, the codification is temporal. Such kind of patterns can achieve high accuracy in measurements, but the ability to measure moving surfaces is restricted to one-shot scans (i.e. not robust to motion acquisition). While classical-based multiplexing techniques like M-array based techniques perform one-shot 3-D reconstruction with good accuracy results, these techniques produce sparse (feature wise) reconstructions.

The following example demonstrates that by targeting at defect regions, it is possible to perform motion-robust and accurate 3D measurement of targeted defect using a hand-held device.

The workflow of the visual inspection with active 3D measurement, or Snap2Inspect3D is shown in Figure 10. With reference to Figure 10, first, an image is captured (S1) using an image capturing unit of a portable imaging device (e.g. a camera on a smartphone/tablet). Once a potential defect is identified by real-time image analysis, the surface geometry within the field of view is acquired in order to perform accurate 3D measurements. To acquire surface geometry using a portable imaging device (and preferably a hand-held imaging device), a spatial coded pattern is used to reconstruct (S2) sparse 3D points of an underlying surface of a detected defect region within a single snap (or by capturing a single image of the defect region). Based on the underlying surface geometry, the image captured by the camera can be overlaid with the surface as texture and the absolute scale of the defect on the image can be inferred. Guided by defect detection (S3), the system projects an adaptive pattern to measure the target region of the defect (i.e. a defect region) for dense and accurate 3D profiling (S4).

Hence, in the workflow, upon detection that an object has a defect, for instance, by the snap2inspect3D techniques, a light projector is instructed to project a coded structured pattern on the object. Thereafter, the image capturing device is instructed to capture a two dimensional image of the object comprising the projected coded structured pattern. A three dimensional (3D) model of the object is constructed from the captured two dimensional image to obtain surface geometry of the object to facilitate 3D measurements of the defect. Optionally, the coded structured pattern is a monochromatic spatially coded pattern.

To achieve optimized measurement over the defect region, we can adapt a coded projection pattern based on a viewpoint between defect and image capturing device, tune the colours of the pattern with regard to the one of the defect regions and increase spatial sample resolution for the targeted defect region. It provides key 3D measurements such as geometry of the inspected surface in absolute scale and the profile of defects (length, width and depth) while keeping the scanning within a few shoots. Finally, task-specified dimensional criteria (S5), such as the maximum acceptable width of cracks or the depth of a hole, can be taken into account in defect verification and decision-making. The minimum number of shoots required to perform 3D measurement is dependent on complexity of the obtained surface geometry of the target object. For example, one shoot is sufficient for a target object with spherical surfaces.

A hardware set-up of a proposed vision system 1100 is shown in Figure 11. It comprises of an image capturing device 1105, such as in the present example, a camera of a smart phone, and a portable projector 1110. To set up the vision system 1100, the image capturing device 1105 is mounted on an adjustable clamp (not shown in Figure 11) together with the projector 1110 at a certain angle θ between the imaging capturing device 1105 and a projected pattern 1115 based on a target object 1120 to be measured. A cable may be used to connect between an input/output port (e.g. USB slot) on the image capturing device 1105 and an input/output port (e.g. HDMI slot) on the projector 1110 to enable data transfer between them. For applications using coded structured light, the active device is typically a digital light projector and is modelled as the inverse of a camera. When a camera is used, light from the environment is focused on an image plane and captured. Such process reduces the dimensions of the data taken in by the camera from three dimensions to two dimensions (i.e. light from a 3D scene is stored on a 2D image), and hence a system calibration is required before dimensions of a defect on a 3D object can be measured from the camera images.

Accuracy and simplicity of such system calibration is a key challenge in building a 3D measurement system 1100. Figure 12 shows coordinate systems to be considered for the projector 1110 and the camera 1105 for system calibration. As illustrated in the example of Figure 12, the

dimensions of a three-dimensional (3D) point $P(X_w, Y_w, Z_w)$ taken in by a camera 1105 is reduced from three dimensions to two dimensions, and represented as $p(u_c, v_c)$. The coordinates of the point P is computed with respect to a 3D coordinate space 1202 formed by three axes X_w, Y_w, Z_w , which intersect at an origin O . The three axes X_w, Y_w, Z_w are orthogonal to one another. The coordinates of $p(u_c, v_c)$ are formed with respect to a 2D camera coordinate space formed by two axes u and v , which intersect at an origin o . The origin o of the 2D camera coordinate space corresponds to an origin O_c of a camera coordinate frame formed by three axes X_c, Y_c, Z_c . The three axes X_c, Y_c, Z_c are orthogonal to one another. The 3D point P has a corresponding projecting point $p'(u_p, v_p)$ at the projector 1110. The coordinates of $p'(u_p, v_p)$ are formed with respect to a 2D projector coordinate space formed by two axes u and v , which intersect at an origin o . The origin o of the 2D projector coordinate space corresponds to an origin O_p of a projector coordinate frame formed by three axes X_p, Y_p, Z_p . The three axes X_p, Y_p, Z_p are orthogonal to one another.

During calibration of the 3D measurement system 1100 of Figure 11 as shown in Figure 12, intrinsic parameters of the camera 1105 and the projector model 1110, and extrinsic parameters between the projector 1110 and the camera 1105 are determined. The terms "camera" and "camera model" and the terms "projector" and "projector model" are used interchangeably in the present example. The intrinsic parameters of the camera model 1105 include a focal length and co-ordinates of a principal point, which describes a relationship between a 3-D point $P(X_w, Y_w, Z_w)$ and a corresponding projection onto an image plane. The extrinsic parameters include rotation and translation between the coordinate systems of the camera and the projector, which describes the relative position and orientation between them. In the example of Figure 12, calibration can be performed by using a black image with an illuminated pixel as a pattern. In the example, only one point can be reconstructed through triangulation using pixel coordinates of the illuminated pixel in the pattern and the corresponding coordinates in the camera image.

The intrinsic parameters of the camera model 1105 can be estimated by widely used calibration processes, such as Direct linear transformation (DLT) method, Zhang's method and the like. For the extrinsic calibration, the system may use a method of projecting a calibration pattern such as through the use of a planar checkerboard to minimize calibration complexity and cost under a fixed setting with the camera 1105 and the projector 1110 fixed. An example of the planar checkerboard is shown in Figure 13d. In another example, in order to adapt the calibration to a mobile setting such as a hand-held image capturing device, the planar checkerboard as shown in Figure 13d can be replaced with asymmetrical circle patterns as shown in Figure 13e for the ease of real-time detection.

The proposed system 1100 and the corresponding method discussed above consider the projector as an inverse camera which maps 2D image intensities into 3D rays, and thus make the calibration of a projector-camera system a similar procedure as that of stereo cameras. In this way, by having the 2D projected points and its 3D correspondences, the system 1100 can be calibrated using any standard stereo camera calibration method.

With regard to light coding, for motion-robust acquisition of surface geometry under a mobile setup, a monochromatic spatially coded pattern is preferred. Colored patterns can also be used although colored patterns are generally less robust than monochromatic patterns. For instance, in one example, the coded pattern can be adapted based on an M-array based theory, with uniqueness of the code for each 3 by 3 window. An algorithm such as that is proposed by Albiter et al; IEEE int. conf. on image processing, pages 529-532, 2007 (hereinafter "Albiter et al") may be used. An alphabet of three symbols that is associated to geometrical shapes (disc, circle, and stripe) may be used. The length of the code associated to each primitive is 9 since a 3x3 window centred on this primitive is considered in this example. In one example, an obtained matrix of dimensions 27x29, which means 783 primitives, verifying first three desired criteria and using three symbols is obtained through the algorithm applied (see Figure 13a). The first three desired criteria refers to the three desired criteria used in the algorithm proposed by Albiter et al to ensure that the system decodes the coded patterns correctly are central symmetry, uniqueness of the code of each 3x3 window, and hamming distance being more than 3 (i.e. each code word is different from each other by at least three symbols). Projected pattern needs to be segmented and decoded for 3D reconstruction (see Figures 13b and 13c). First, the contours are segmented and then classified into three pattern primitives. Once each primitive together with its neighbors is detected, its code is determined. Finally, the 3D position of each decoded point can be calculated through triangulation between the observed pattern and the original one.

According to the light coding description above, the projector 1110 can be configured to project a structured coded pattern on a 2D image captured by the camera 1105 for defect detection. Thereafter, a 3D model is reconstructed by decoding the projected coded pattern as discussed to obtain 3D coordinates to detect a defect on the 2D image. Once a defect is identified on the 2D image

with its underlying surface parameters estimated using the method described above, the 3D coordinates of any points in the defect region can be calculated from the 2D image.

Thereafter, the system 1100 is configured to generate a defect-specified pattern to measure the defect automatically. Figure 14 shows essentially the same elements as Figure 12 but in addition, there is shown a cube 1404 provided for illustrative purposes and it represents the 3D object to be scanned for defects. As shown in Figure 14, the system 1100 is configured to assume a point, p , as a centre of defect in a 2D image, with its 3D point, $P (X_w, Y_w, Z_w)$, in a camera frame 1402 (i.e. video frame captured by the camera 1105). Based on the camera-projector calibration,

$$p' = K_{proj}^{-1} T P, \quad \text{Equation (7),}$$

where p' is the corresponding point for the projector, K is the intrinsic parameters for the projector and T is the extrinsic parameter between camera frame 1402 and projector frame 1406.

Various pattern codification strategies can be applied. For example, with reference to Figure 17, to measure a depth of a dent on a surface, a pattern of four-points are required, where three points form a reference plane with the remaining one point as target point. Thus, the depth of the dent can be calculated as distance between the target point and the reference plane.

In order to obtain accurate 3D measurements, system calibration should be performed. As discussed with reference to Figures 11 and 12, to perform the calibration between the camera 1105 and projector 1110, a user (investigator) may point both the camera 1105 and projector 1110 toward a calibration board with asymmetrical circle patterns (e.g. as shown in Figure 13e). The detected patterns are automatically marked on a screen 1502 by the system 1100 is illustrated in Figure 15, and recorded for correlation. The calibration is repeated at least a further two times, with the camera 1105 and the projector (of the system 1100) 1110 tilted at a different angle θ toward the calibration board with asymmetrical circle patterns to collect at least three sets of data. The system calibration result with 10 sets of data is illustrated in Figure 16.

In other words, with reference to Figure 15, generally, the image capturing device 1105 and the light projector 1110 are directed towards a calibration board comprising an area 1510 with calibration pattern 1506 such that the light projector 1110 projects asymmetrical circle patterns 1504 outside the area 1510 with calibration pattern 1506. The calibration pattern 1506 and the projected asymmetrical circle patterns 1504 are detected by capturing, using the image capturing device 1105, an image of the calibration board comprising the calibration pattern 1506 and the projected asymmetrical circle patterns 1504 at a first instance. The detection of the calibration pattern 1506 and the projected asymmetrical circle patterns 1504 is graphically represented as lines 1512 and 1508 respectively. In this example, the line 1508 is formed by joining each center point of each asymmetrical circle of the pattern 1504 together, whereas the line 1512 is formed by joining each center point of each asymmetrical circle of the pattern 1506 together. Thereafter, the image capturing device 1105 and the light projector 1110 are positioned or tilted at a different angle θ towards the calibration board for at least two more times. It should be appreciated that when the image capturing device 1105 and the light projector 1110 are positioned or tilted differently, the calibration pattern 1506 and the projected asymmetrical circle patterns 1504 are detected again for each additional image captured at the different angle θ by the image capturing device 1105. After capturing at least three images at respective different angles, a calibration result is obtained based on collated data of the detected calibration pattern and detected projected asymmetrical circle patterns. The calibration pattern 1506 may be different or the same as the projected asymmetrical circle patterns 1504. For example, the calibration pattern 1506 can be asymmetrical circle patterns or planar checkerboard pattern. In another example, the calibration pattern 1506 in the area 1510 may be overlaid with the structured coded pattern projected by the light projector 1110 during calibration.

Following the completion of the calibration procedure (and taking into consideration the calibration result), the system 1100 can perform dimensional measurement using stereo-triangulation on the selected object. For example, a height of an object on a surface, or a depth of a cavity in a surface can be measured. Figure 17 illustrates an example of height measurement of a rectangular object 1704 on a surface 1706. Four illuminated dots (or points) 1702 are projected onto the object 1704 placed on the surface 1706, and their 3D coordinates 1708 with respect to a centre of the camera 1105 is calculated in real-time based on the triangulation. Height or depth on the upper left corner, for example, (x_1, y_1, z_1) , is calculated by the distance from a centre point 1710 to the plane formed by the rest of three points, for example, (x_0, y_0, z_0) , (x_2, y_2, z_2) , and (x_3, y_3, z_3) . In other words, generally, the four illuminated dots 1702 are projected onto a surface containing an elevated or a dented point for measuring respective height or depth of the point relative to the surface, wherein three of the four illuminated dots 1702 for forming a reference plane are projected on the surface and the remaining one illuminated dot (e.g. the centre dot or point 1710 in Figure 17) is projected on the point. After obtaining 3D coordinates of each of the four illuminated dots 1702 with respect to a centre

of the image capturing device 1105 through triangulation, the respective height or depth of the point relative to the surface based on distance between the point and the reference plane can be calculated.

In the aforementioned examples relating to a toolbox (an application that may include a mobile application of a smartphone) for performing single-view mobile inspection of defects which is guided by different types of visual knowledge (Snap2Inspect), a confidence score is provided to assist an inspector to capture images using optimal imaging parameters (angle, distance, etc.) when performing different automatic inspection tasks. With the scores and their changing trends, better imaging observations are incrementally achieved for the specific inspection task, facilitating better detection performance.

Monocular vision might not be able to resolve the scale of the object. This presents a challenge while performing visual inspection tasks at various times, because detailed defect analysis may require measurement in absolute scale. The proposed method and the corresponding system 1100 resolves the scale of the inspected object by stereo-triangulation with an additional projector. A flow chart 1800 in Figure 18 describes how the proposed method performed through the corresponding system 1100 fits into the framework of Snap2Inspect3D.

Specifically, with reference to Figure 18, an inspector 1802 obtains an image of a target object and performs a one shot-scan for 3D surface reconstruction. Thereafter, Snap2Inspect3D technique 1804 described above is applied to detect a defect of the target object using visual knowledge and surface geometry involving 3D measurement (1806). That is, Snap2Inspect3D 1804 involves active sensing for surface geometry and 3D defect measurement (1808). In one example, four points may be illuminated on a detected defect region, and a 3D measurement of a defect can be subsequently made if positions of the projector 1110 of Figure 11 and camera 1105 of Figure 11 are fixed. A confidence score 1810 for the detection on whether the object has a defect is computed and fed back to an inspector 1802 to assist the inspector 1802 to capture one or more images with more optimal imaging parameters (angle, distance, etc.). After the confidence score 1810 is computed, documentation of the defect finding takes place (1812). The system 1100 then determines, based on the calculated confidence score 1810, whether to generate better view recommendations 1814. For example, better view recommendations 1814 may include improved imaging parameters with respect to translation and/or rotation movement of the camera 1105 and generates the improved imaging parameters if required. Upon receiving an input, the camera 1105 may be instructed to apply the improved imaging parameters to capture a second image of the object that results in calculation of an improved confidence score. With the collated confidence scores and the observed changing trends, better imaging observations can be incrementally achieved for each specific inspection task, thereby facilitating better detection performance.

Traditional precision industries measure 3D profile of a surface using high-end 3D scanners, which require to be mounted on a flexible holder, or a hand-held scanner, which tracks the position of the scanner based on the texture on the object surface. However, acquisition of accurate 3D surface profile takes a long scan time with those high-cost 3D scanners, and additional efforts are required to extract and analyze the regions of 3D surface profile related to defects by manually drawing boundary of regions or reference points. To address this problem, a 3D surface inspection tool was introduced. Such 3D surface inspection tool can scan the surface and analyze three defined types of defects: dents, rivet flushness-levels and gaps. The detection result is subsequently displayed on the measured surfaces. However, such a system is designed to measure specifically the above-mentioned three types of defects. The conventional system is also bulky because a pole is required to fix the inspection system against the surface during scanning, and cannot be hand held.

The proposed method and corresponding system 1100 of Figure 11 for visual inspection with active 3D measurement offers active defect measurement guided by defect detection, which could be readily implemented using existing hardware, such as smartphone or tablet with image capturing capabilities, and a hand-held projector into a low-cost palm-sized platform for on-site visual inspection. The proposed method and corresponding system 1100 further allows active 3D measurement with adaptive illumination pattern guided by the image-based defect detection. It provides key 3D measurements such as geometry of the inspected surface in absolute scale and profile of defects (length, width and depth) while keeping the scanning time within a few snaps.

The proposed method and corresponding system 1100 advantageously provides intelligent defect detection with absolute scale, where the image captured by the camera 1105 can be overlaid with the surface geometry as texture. Based on the underlying geometry, the absolute scale of the detected area on the image can be inferred. Thus, task-specified dimensional criteria such as the maximum acceptable width of crack can be taken into consideration during defect detection and decision making process.

The image capturing device 1105 in Figure 11 may be a computing or mobile device, for example, smart phones, tablet devices, and other handheld devices. The image capturing device 1105 in Figure 11 and the projector 1110 in Figure 11 may be able to communicate through other communications network, such as, wired network, but these are omitted from Figure 11 for the sake of clarity. There may be provided software, such as one or more computer programs being executed within the image capturing device 1105, and instructing the image capturing device 1105 to connect and communicate with the projector 1110 according to the operations described with reference to the earlier figures. Likewise, there may be provided software, such as one or more computer programs being executed within the projector 1110, and instructing the projector 1110 to connect and communicate with the image capturing device 1105 according to the operations described with reference to the earlier figures.

The architecture of a system or an apparatus proposed in an example of the present disclosure may be an apparatus 1902 in Figure 19 and comprises a number of individual components including, but not limited to, processing unit 1916 (or processor), a memory 1918 (e.g. a volatile memory such as a Random Access Memory (RAM) for the loading of executable instructions 1920, the executable instructions defining the functionality the apparatus 1902 carries out under control of the processing unit 1916. The apparatus 1902 also comprises a network module 1925 allowing the apparatus 1902 to communicate over the communications network 1908 (for example the internet, Bluetooth network). User interface 1924 is provided for user interaction and may comprise, for example, conventional computing peripheral devices such as display monitors, computer keyboards and the like. The apparatus 1902 may also comprise a database 1926 to store defect detection results (and corresponding confidence scores and/or system settings of the image capturing device 1905 as well as the projector 1110). The database 1926 may also be configured to store a coded projection pattern for projecting on a defect region of an object. It should also be appreciated that the database 1926 may not be local to the apparatus 1902. The database 1926 may be a cloud database.

The processing unit 1916 is connected to input/output devices such as a computer mouse, keyboard/keypad, a display, headphones or microphones, a video camera and the like (not illustrated in Figure) via Input/Output (I/O) interfaces 1922. The components of the processing unit 1916 typically communicate via an interconnected bus (not illustrated in Figure 19) and in a manner known to the person skilled in the relevant art.

The processing unit 1916 may be connected to the network 1908, for instance, the Internet, via a suitable transceiver device (i.e. a network interface) or a suitable wireless transceiver, to enable access to e.g. the Internet or other network systems such as a wired Local Area Network (LAN) or Wide Area Network (WAN). The processing unit 1916 of the apparatus 1902 may also be connected to one or more external wireless communication enabled device 1904 through the respective communication links 1910 and 1912 via the suitable wireless transceiver device e.g. a WiFi transceiver, Bluetooth module, Mobile telecommunication transceiver suitable for Global System for Mobile Communication (GSM), 3G, 3.5G, 4G telecommunication systems, or the like.

In another example, the architecture of a system or an apparatus proposed in an example of the present disclosure may be a device 1904 in Figure 19. The device 1904 may comprise a number of individual components including, but not limited to, microprocessor 1928 (or processor), a memory 1930 (e.g. a volatile memory such as a RAM) for the loading of executable instructions 1932, the executable instructions defining the functionality the device 1904 carries out under control of the microprocessor 1928. The device 1904 also comprises a network module (not illustrated in Figure) allowing the device 1904 to communicate over the communications network 1908. User interface 1936 is provided for user interaction and control that may be in the form of a touch panel display and presence of a keypad as is prevalent in many smart phone and other handheld devices. The device 1904 may also comprise a database (not illustrated in Figure), which may not be local to the device 1904 but a cloud database. The device 1904 may include a number of other Input/Output (I/O) interfaces 1934 as well but they may be for connection with headphones or microphones or projectors, Subscriber identity module (SIM) card, flash memory card, USB based device, and the like, which are more for mobile device usage.

The software and one or more computer programs stored may include, for example, applications that may include one or more applications for e.g. internet accessibility, operating the device 1904 and apparatus 1902 (i.e. operating system), network security, file accessibility, database management, which are applications typically equipped on a desktop or portable (mobile) device. The software and one or more computer programs may be supplied to a user of the device 1904 and/or the apparatus 1902 encoded on a data storage medium such as a CD-ROM, on a flash memory carrier or a Hard Disk Drive, and are to be read using a corresponding data storage medium drive for instance, a data storage device (not illustrated in Figure 19). Such application programs may also be

downloaded from the network 1908. The application programs are read and controlled in its execution by the processing unit 1916 or microprocessor 1928. Intermediate storage of program data may be accomplished using RAM 1920 or 1930.

Furthermore, one or more of the steps of the computer programs or software may be performed in parallel rather than sequentially. One or more of the computer programs may be stored on any machine or computer readable medium that may be non-transitory in nature. The computer readable medium may include storage devices such as magnetic or optical disks, memory chips, or other storage devices suitable for interfacing with a general purpose computer or mobile device. The machine or computer readable medium may also include a hard-wired medium such as exemplified in the Internet system, or wireless medium such as exemplified in the Wireless LAN (WLAN) system. The computer program when loaded and executed on such a general-purpose computer effectively results in an apparatus (e.g. 1105 and/or 1110 in the respective Figures) that implements the steps of the computing methods in examples herein described.

In summary, examples of the present disclosure may have the following features.
A method for detecting a defect on an object, the method comprising:
instructing an image capturing device to capture a first image of an object;
processing the captured first image to detect whether the object has a defect;
calculating a confidence score for the detection on whether the object has the defect;
generating the calculated confidence score as a feedback to a user; and
determining based on the calculated confidence score whether to generate improved imaging parameters with respect to translation and/or rotation movement of the image capturing device;
generating the improved imaging parameters if required after considering the calculated confidence score; and
receiving an input to instruct the image capturing device to capture, according to the improved imaging parameters, a second image of the object that results in calculation of an improved confidence score.

The method may further comprise:
collating confidence scores obtained from different images captured by the image capturing device for a common defect on the object; and
using the collated confidence scores to fine tune detection of the common defect on the object or another object.

The method may further comprise:
selecting an operation mode from a plurality of operation modes so that the processing of the captured first image to detect whether the object has the defect is performed according to the selected operation mode,
wherein each operation mode is for detecting one type of defect, wherein the plurality of operation modes includes at least two of the following operation modes:
an operation mode to detect a crack on the object using a decision-theory based algorithm;
an operation mode to detect erosion on the object using a classification algorithm that trains on a plurality of training images;
an operation mode to detect missing material on the object using an algorithm that matches features points of a reference two dimensional (2D) image with feature points of a captured image of the object; and
an operation mode to detect missing material on the object using an algorithm that matches features points of a three dimensional (3D) model with feature points of a captured image of the object.

The calculated confidence score may generated as feedback to the user after comparing the calculated confidence score against a threshold, wherein the threshold is used to reduce a number of false detection.

The method may further comprise:
upon the detection that the object has the defect, instructing a light projector to project a coded structured pattern on the object;
instructing the image capturing device to capture a two dimensional image of the object comprising the projected coded structured pattern;
constructing a three dimensional (3D) model of the object from the captured two dimensional image by decoding the projected coded structured pattern to obtain surface geometry of the object; and

performing 3D measurements of the defect using the surface geometry obtained.
The coded structured pattern may be a monochromatic spatially coded pattern.
The method may further comprise:

- adapting the coded structured pattern based on viewpoint between a defect and the image capturing device; and/or
 tuning colours of the coded structured pattern in regard to a specific defect region; and/or
 increasing spatial sample resolution for a target defect region considered.
- 5 The method may further comprise:
 directing the image capturing device and the light projector towards a calibration board comprising an area with calibration pattern such that the light projector projects asymmetrical circle patterns outside the area with calibration pattern;
 detecting the calibration pattern and the projected asymmetrical circle patterns by capturing
 10 using the image capturing device an image of the calibration board comprising the calibration pattern and the projected asymmetrical circle patterns;
 tilting the image capturing device and the light projector at a different angle towards the calibration board for at least two more times;
 detecting the calibration pattern and the projected asymmetrical circle patterns for additional
 15 images captured at the different angles by the image capturing device; and
 obtaining a calibration result based on collated data of the detected calibration pattern and detected projected asymmetrical circle patterns.
 The calibration pattern may be a plurality of asymmetrical circle patterns.
 The method may further comprise:
 20 taking into consideration the calibration result, performing measurement using stereo-triangulation on an image of the object captured by the image capturing device.
 The method may further comprise:
 projecting four illuminated dots onto a surface containing an elevated or a dented point for
 measuring respective height or depth of the point relative to the surface, wherein three of the four
 25 illuminated dots for forming a reference plane are projected on the surface and the remaining one illuminated dot is projected on the point;
 calculating 3D coordinates of the four illuminated dots with respect to a centre of the image capturing device based on triangulation; and
 after obtaining the 3D coordinates of the four illuminated dots, calculating height or depth of
 30 the point relative to the surface based on distance between the point and the reference plane.
 An apparatus for detecting a defect on an object, the apparatus comprises:
 a processor configured to execute instructions in a memory to control the apparatus to:
 instruct an image capturing device (e.g. 1105 of Figure 11) to capture a first image (e.g. 204 of
 Figure 2) of an object (e.g. 1704 of Figure 17);
 35 process the captured first image to detect whether the object has a defect (e.g. 210 of Figure 2);
 calculate a confidence score (e.g. 215 of Figure 2) for the detection on whether the object has the defect;
 generate the calculated confidence score as a feedback to a user (e.g. 202 of Figure 2);
 40 determine based on the calculated confidence score whether to generate improved imaging parameters (e.g. 220 of Figure 2) with respect to translation and/or rotation movement of the image capturing device;
 generate the improved imaging parameters if required after considering the calculated confidence score; and
 45 receive an input to instruct the image capturing device to capture, according to the improved imaging parameters, a second image of the object that results in calculation of an improved confidence score.
 The apparatus may be controllable to:
 collate confidence scores obtained from different images captured by the image capturing
 50 device for a common defect on the object; and
 use the collated confidence scores to fine tune detection of the common defect on the object or another object.
 The apparatus may be controllable to select an operation mode from a plurality of operation modes so that the processing of the captured first image to detect whether the object has the defect is
 55 performed according to the selected operation mode,
 wherein each operation mode is for detecting one type of defect, wherein the plurality of operation modes includes at least two of the following operation modes:
 an operation mode to detect a crack on the object using a decision-theory based algorithm;
 an operation mode to detect erosion on the object using a classification algorithm that trains
 60 on a plurality of training images

an operation mode to detect missing material on the object using an algorithm that matches features points of a reference two dimensional (2D) image with feature points of a captured image of the object; and

- 5 features points of a three dimensional (3D) model with feature points of a captured image of the object.

The calculated confidence score may be generated as feedback to the user after comparing the calculated confidence score against a threshold, wherein the threshold is used to reduce a number of false detection.

- 10 The apparatus may be controllable to:

upon the detection that the object has the defect, instruct a light projector (e.g. 1110 of Figure 11) to project a coded structured pattern on the object;

instruct the image capturing device to capture a two dimensional image of the object comprising the projected coded structured pattern;

- 15 construct a three dimensional (3D) model of the object from the captured two dimensional image by decoding the projected coded structured pattern to obtain surface geometry of the object; and

perform 3D measurements of the defect using the surface geometry obtained.

- 20 13A). The coded structured pattern may be a monochromatic spatially coded pattern (e.g. Figure

The apparatus may be controllable to:

adapt the coded structured pattern based on viewpoint between a defect and the image capturing device; and/or

- 25 tune colours of the coded structured pattern in regard to a specific defect region; and/or increase spatial sample resolution for a target defect region considered.

After directing the image capturing device and the light projector towards a calibration board comprising an area (e.g. 1510 of Figure 15) with calibration pattern (e.g. 1506 of Figure 15) such that the light projector projects asymmetrical circle patterns (e.g. 1504 of Figure 15) outside the area with calibration pattern, the apparatus may be controllable to:

- 30 detect the calibration pattern and the projected asymmetrical circle patterns from an image captured using the image capturing device, wherein the image comprises the calibration board comprising the calibration pattern and the projected asymmetrical circle patterns; and

after tilting the image capturing device and the light projector at a different angle towards the calibration board for at least two more times, the apparatus may be controllable to:

- 35 detect the calibration pattern and the projected asymmetrical circle patterns for additional images captured at the different angles by the image capturing device; and obtain a calibration result based on collated data of the detected calibration pattern and detected projected asymmetrical circle patterns.

- 40 The calibration pattern may be a plurality of asymmetrical circle patterns.

The apparatus may be controllable to:

take into consideration the calibration result to perform measurement using stereo-triangulation on an image of the object that is captured by the image capturing device.

- 45 After projecting four illuminated dots (e.g. 1702 of Figure 17) onto a surface (e.g. 1706 of Figure 17) containing an elevated or a dented point (e.g. 1710 of Figure 17) for measuring respective height or depth of the point relative to the surface, wherein three of the four illuminated dots for forming a reference plane are projected on the surface and the remaining one illuminated dot is projected on the point, the apparatus may be controllable to:

- 50 calculate 3D coordinates (e.g. 1708 of Figure 17) of the four illuminated dots with respect to a centre of the image capturing device based on triangulation from an image captured by the image capturing device, wherein the image comprises the four illuminated dots, the surface and the point; and

after obtaining the 3D coordinates of the four illuminated dots, the apparatus is controllable to:

- 55 calculate height or depth of the point relative to the surface based on distance between the point and the reference plane.

Throughout this specification and claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated integer or group of integers or steps but not the exclusion of any other integer or group of integers.

5 While the invention has been described in the present disclosure in connection with a number of embodiments and implementations, the invention is not so limited but covers various obvious modifications and equivalent arrangements, which fall within the purview of the appended claims. Although features of the invention are expressed in certain combinations among the claims, it is contemplated that these features can be arranged in any combination and order.