A New Approach for Crack Detection and Sizing in Nuclear Reactor Cores

Michael G Devereux^{a,*}, Paul Murray^a, Graeme M. West^a

^a*Department of Electronic and Electrical Engineering, University of Strathclyde, 204 George Street, Glasgow, G1 1XW, U.K.*

Abstract

Remote Visual Inspection (RVI) of reactors in nuclear power plants allows station operators to assess the health and condition of their plant. In the UK, most nuclear stations are of the Advanced Gas-cooled Reactor (AGR) design. During planned periodic outages, a representative portion of each AGR core is inspected using specialist tools equipped with various sensors including a video camera for RVI. If cracks are observed in the core during data capture, a stitched image of the region needs to be created so that the crack can be analysed and sentenced (classifying the crack morphology, location, orientation and size) before the station is returned to service, provided return to service is justified. Currently, the crack analysis and sizing activities are conducted manually by expert analysts in a laborious process. In this paper, we present a new image processing approach capable of automating aspects of the crack analysis process. Specifically, we describe a set of techniques for quickly and accurately detecting the presence of cracks in AGR fuel channel inspection images. We also present a method for detecting circular channel features known as trepanned holes whose dimensions are known and can thus be used for scaling. The results of applying the proposed techniques are evaluated on image data from real AGR fuel channels and are shown to produce comparable results to those obtained manually. The advantage of the proposed approach is that it is fast, robust and more repeatable than the existing manual approach.

Keywords: Decision support systems, nuclear power generation safety, nuclear engineering, nuclear power plant

1. Introduction

Seven of the eight nuclear power plants in the UK, are of the Advanced Gas-cooled Reactor (AGR) design and these stations provide enough electrical energy to meet around 20% of the consumer demand each year. Many of the UK's (and the world's) Coble et al. (2012) nuclear power plants are approaching the end of their original design lifetime and a number of others have already exceeded theirs. This comes as a result of various plant lifetime extensions which have been granted based on extensive supporting evidence which demonstrates that the stations remain safe and fit for continued operation. Part of the evidence used to support the case for lifetime extension is gathered during plant inspection which occurs regularly during planned

periodic outages when the reactor is not at power. Plant inspection also allows station operators to regularly assess the condition of their reactor cores and to identify, locate and quantify any cracks or other defects in the fuel channels which form the core itself.

In the AGR plant design, the core is comprised of hundreds of channels constructed from cylindrical graphite bricks which provide housing for the fuel and allow the insertion of control rods to control the reaction. The inspection of selected fuel channels is conducted using specialist equipment and tools which house various sensors including a video camera. The camera is used to perform Remote Visual Inspection (RVI) and records videos of the fuel channel's inner wall at 6 different, but overlapping, orientations. In the event that any cracks are observed during the inspection process, stitched images (known as crack montages) of the affected region need to be produced from the video footage. Cracks present in this image are subsequently quan-

*Corresponding author. Tel.: 0141 444 7247

Email address: michael.devereux@strath.ac.uk
(Michael G Devereux)

tified and scaled using a conversion factor based on the size of known features in the images. This image stitching and analysis activity lies on the critical outage path as it must be completed before the station can be returned to power. From a commercial perspective, these inspection activities can be extremely costly.

Currently, the data analysis, montage generation and crack detection and sizing activities are conducted in a lengthy and onerous manual process. This is performed by specialists who view inspection videos, extract relevant frames and construct the crack montages. These specialists then analyse the stitched images and manually measure the length and width of cracks. Finally, the outcome of this initial analysis are provided to a Graphite Assessment Panel (GAP) who conduct a process known as crack sentencing (classifying the crack morphology, location, orientation and size) before the station is returned to service.

Given that the activities described lie on the critical path¹, it is highly desirable to minimize the time between gathering data and returning the station to power, especially if this can be achieved while making better use of the captured data which already exists. Furthermore, from an auditability perspective it is critical that any measurements made during the analysis are robust and repeatable. To this end, a new software tool known as ASIST (Automated Software Image Stitching Tool) (Murray et al. (2016b)) and (West et al. (2015a)), has recently been trialed by a team of specialists to assess the feasibility of using this tool to replace the existing manual image stitching method to generate crack montages. ASIST processes entire inspection videos to create full channel images known as chanoramas (fuel channel panoramas) which offer a 360° panoramic view of inner walls of the fuel channel from which crack montages can be easily extracted. An example chanorama is shown in Figure 1). ASIST has now been used to process over 500 different inspection videos (Murray et al. (2016a)) to generate chanoramas for over 500 different fuel channels from across the UK's nuclear fleet. Taking less than 30 minutes to generate a chanorama from a given inspection video, ASIST has the potential to significantly speed up the generation of crack montages. However, the subsequent crack detection and sizing activities are still performed manually. This

¹The critical path is a sequence of tasks that must be completed before the plant can be returned to service.

Figure 1: Chanoramas and crack montages (a) chanorama showing 8 brick layers (b) manually generated crack montage of layer 2nd from top (c) corresponding crack montage from chanorama

introduces subjectivity into the process.

In this paper we present a new, innovative, image processing system capable of automatically analysing chanoramas to detect the presence of cracks and other channel features known as trepanned holes. Trepanned holes can be seen in AGR core inspection videos of fuel channels from which small samples of graphite have been trepanned during inspection such that the irradiated graphite's properties can be measured. Since the size of the samples are known, the dimensions of the holes that remain in the fuel channels after the trepanning process can be used to give an indication of a scale for converting from pixel based measurements in the image data into real quantitative values. Automatic crack detection is performed using a bespoke Bag-of-Visual Words (BoVW) (Sivic et al. (2005)) object detection framework which is able to distinguish regions of a chanorama containing cracks from regions which are uncracked. Trepanned holes are detected using the Hough Transform (Duda and Hart (1972)) and measured using ellipse detection techniques.

The output is analysed to extract the major and minor axes (in pixels) of each trepanned hole for scaling. The key contributions of this paper are:

- a novel application of a BoVW approach to detect cracks in nuclear plant inspection images
- a new method for automatically sizing cracks AGR inspection images through automated identification and sizing of known reference features.

In the next section we review the state-of-the-art in RVI in the nuclear industry, in Section 3, we provide details on the AGR core along with a detailed description of the remote visual inspection process. In Section 4, we present our new image processing method for inspecting nuclear reactor cores. The results of applying our approach to real AGR inspection footage are discussed in Section 5 before we conclude in Section 6.

2. Related Research

The automated detection of visible structural defects is of interest in a range of problems ranging from crack detection on concrete bridges (Adhikari et al. (2014); Prasanna et al. (2014)) to nuclear reactor internals (Devereux et al. (2018)).

Jahanshahi and Masri (2012) propose a RVI framework for the condition assessment of structures using a robotic inspection system where 2D images are recorded and all quantitative measurements are read by reconstructing a 3D scene from the 2D images. This is a preferable alternative to the labor intensive case where an inspector would need to travel to each target structure. In some cases such as nuclear reactor internals inspection, RVI is the only possible option (Lee et al. (2006), Choi et al. (2005), Cho et al. (2004), Glass et al. (1999), West et al. (2015b)). While this paper is focused on automatically detecting specific physical features of interest in RVI images including cracks, there has already been some research published focusing on the automated detection of cracks in images from the nuclear industry. Schmugge et al. (2014) propose a framework for the automated detection of cracks on steel plates using segmentation and a regrowing approach. Schmugge et al. (2016) later propose a new approach for detecting both synthetic and real cracks on steel plates using advances in deep neural networks to classify image

tiles. Schmugge et al. (2017) further refine this approach to label individual pixels as cracked or uncracked by adapting and retraining Segnet (Badrinarayanan et al. (2015)), a deep learning semantic segmentation framework.

Chen et al. (2017) develop a method for detecting cracks on metallic surfaces in nuclear power plants using local binary patterns, support vector machines and data fusion to aggregate information obtained from different video frames to make the process robust. Chen and Jahanshahi (2017) further improve the detection rate of cracks on a synthetic dataset using a deep learning framework. This approach combines the the output of a convolutional neural network (CNN) and a Naive Bayes data fusion scheme to detect cracks with great accuracy. While most of this research focuses on metallic surfaces, this paper only considers crack detection in graphite bricks. The graphite surface does not suffer the same reflection problems from different views with different lighting that polished metallic surfaces typically exhibit.

3. AGR Core Inspection

AGR cores are constructed from layers of around 300 hollow, cylindrical graphite bricks which each stand approximately 0.85m - 0.9m tall. Individual bricks are keyed together to form layers which are stacked one on top of the other to create around 300 fuel channels, each with 11 brick layers (this varies from station to station), and this forms the nuclear reactor core. The fuel channels provide housing for the fuel and moderation for the reaction. Bricks in interstitial positions allow the insertion of control rods which are used to control and stop the reaction. The AGR core is recognized as one of the life-limiting component of this type of reactor (West et al. (2011)) and it is regularly inspected to ensure it remains safe and fit for continued operation. While various tools are used to measure a number of properties of the fuel channels selected for inspection, in this paper we describe only the RVI and trepanning processes as these are relevant to the work that is presented.

3.1. Remote Visual Inspection

Part of the inspection uses tools known as NICIE Mk2 (New In-Core Inspection Unit Mark 2) and CBIU (Channel Bore Inspection Unit) to gather data from inside a subset of fuel channels which

are selected for inspection based on various criteria (Cole-Baker and Reed (2007)). NICIE and CBIU are equipped with various sensors including a video camera which allows RVI of the inner channel walls. This tool is lowered into each fuel channel being inspected using a winch system. The camera and inspection tool are connected to a control console on the pilecap which is manually operated and controls the system while also recording the inspection footage. Live video footage is captured from inside each channel which undergoes inspection. Each video contains 2 sections if the channel is defect free and 3 sections otherwise. These can be summarized as:

- 1 Footage from a down channel camera as the inspection tool is lowered from the top of a fuel channel to the bottom.
- 2 6 overlapping scans of the side walls of the channel captured at different orientations with a 70° field of view.
- 3 Crack following footage (only if cracks are seen in 2) which requires lowering the inspection camera to the defect location and imaging it from starting point to end point.

The inspection camera is mounted so that the field of view is looking down the fuel channels but it is possible to record the inside wall of the fuel channel using a mirror that is part of the inspection tool. The position of this mirror is controlled by a motor so it is possible to view either down the fuel channel or the side walls as required. This mirror can be controlled remotely from the console so this allows RVI of the side walls of the channel after the camera has been inserted into the core. A more detailed description of the inspection process may be found in West et al. (2015b).

Historically, the crack following footage has been captured to facilitate the manual assembly of crack montages which could be analysed and used for crack sentencing prior to returning the station to service. The manual assembly of each image can take up to a full working day to complete depending on the extent and complexity of the crack. If multiple cracks are found in a channel this can impact on the time taken capture the additional data needed to build crack montages and can thus prolong the outage and increase its costs. Recently, crack montages extracted from chanoramas created using the techniques described in (Murray et al. (2016b)),

like the one shown in Figure 1, have been analysed in the crack quantification and sentencing process. Thus, there is a possibility that, in future, crack following footage may not need to be captured. The time associated with each channel inspection could therefore be reduced. While the crack montage generation process looks set to become automated, subsequent analysis of the stitched images produced remains a manual process.

3.2. Trepanning

It is explained in Cole-Baker and Reed (2007) that weight loss in the graphite core occurs as a result of radiolytically-induced oxidation which increases the sensitivity of the core reactivity to certain fault scenarios. To assess weight loss in the core, a specialist inspection tool, equipped with angular-controlled, diamond-tipped, hollow cutting tool is used to trepan approximately 19 mm (previously 40mm) diameter and up to 70mm long samples from select channels during outages (Bradford (2015)). These samples are analysed offsite to assess and quantify weight loss. The remaining holes in the graphite bricks are visible in the inspection footage and are approximately 23mm in diameter. From a RVI perspective, this process provides an additional channel feature known as a trepanned hole (only in channels which have undergone this process) which can be seen in chanoramas. If these holes can be detected and sized in the chanoramas, they can be used for converting pixel based crack measurements into mm since the size of these features is approximately known.

4. A new decision support tool for remote visual inspection of Nuclear Reactors

The proposed new decision support tool for inspecting nuclear reactor cores is an image processing based solution capable of automatically analysing chanoramas to detect and highlight the presence of cracks and to automatically detect and scale trepanned holes. Figure 2 provides an example of a region of a chanorama which better illustrates what trepanned holes look like when viewed up-close.

4.1. Chanorama images and challenges for automated processing

Chanoramas are 8-bit greyscale images constructed from the red channel output of a radiation hardened RGB camera which captures 576 x

Figure 2: Example region of a single scan demonstrating the challenging textures and illumination of the imagery. Also shown are two trepanned holes.

720 pixel images and is embedded within each of the NICIE Mk2 and CBIU inspection tools (Murray et al. (2016b)). Each chanorama produced by ASIST represents approximately 30-45 minutes of video which is presented in the form of a single image. A typical chanorama consists of 45000-50000 row pixels and 4500-5000 column pixels which immediately presents a computational challenge for fast automated analysis. There also exists a significant variation in channel conditions which can affect the scene and make automated analysis challenging. An example of this is the varying bore diameter which changes the proximity of the channel bore to the sensor and illumination resulting in changes in brightness and contrast. All of this, as well as the fact that the image acquisition parameters such as illumination and zoom can be (and are regularly) varied by the inspection engineer make it very challenging to apply automated analysis techniques to detect cracks and other features directly from chanoramas. The section of chanorama shown in Figure 2 demonstrate some of the undesirable artefacts posing significant challenges for designing robust algorithms for automated analysis.

4.2. Automatic Crack Detection

The aim of our proposed, automatic crack detection system is to detect any cracks in images of each brick layer such that their presence can be highlighted to inspection engineers for further analysis. In this study we consider the binary classification problem of separating individual brick layers into two distinctive sets (cracked and uncracked) and use customized, supervised machine learning techniques to distinguish between the two. As with any supervised machine learning approach, a labelled data set is required and we partition this into a set

of training images and a set of test images. All training is performed on the training set and the test set is used solely and in isolation to evaluate performance and test the algorithm’s generalisation. It should be noted that the raw image data for the cracked bricks cannot be fed directly into the classifier as the individual images are too large and would provide a poor representation of the physical features of interest in each image. The reason for this is that despite the image showing a crack in the brick, by area, the majority of the image shows a brick which is not cracked.

The approach used for crack detection is a Bag-of-Visual-Words (BoVW) object detection framework (Sivic et al. (2005)) which we customize for our application. While a number of supervised machine learning techniques could have been adopted, BoVW models were chosen since they are robust, efficient and have been frequently used in computer vision for object detection. They are capable of exceptional performance in a wide range of applications ranging from classification of flowers to detecting objects in X-ray images captured by Baggage Security Screening Systems (Nilsback and Zisserman (2006); Kundegorski et al. (2016); Csúrka et al. (2004); Chatfield et al.). The BoVW model of feature-driven image understanding considers the task of classifying a given visual instance as an analogy to performing the classification of written articles (by type/genre) based on the statistical occurrence of different key words or word patterns. In this particular case, the “visual” words are common physical features found in each image of the training set. Each of these physical features can be statistically described using an appropriate feature descriptor. The feature descriptor is a (large) multi-dimensional vector of statistical image gra-

dient occurrences and, for this work, Speeded-up Robust Features (SURF) were used (Bay et al. (2006)). A number of other popular feature descriptors including ORB (Oriented FAST and rotated BRIEF²) and SIFT (Scale-invariant feature transform) were tested but SURF feature descriptors performed best. SURF is designed to be robust to appearance variations and is therefore invariant to illumination and orientation. The speed of travel of the inspection tool is constant during inspections and as we are applying the framework to tiles of a single composite images (chanormas) constructed from inspection frames, our framework is also invariant to inspection speed. As mentioned previously, reflections on the graphite surface do not pose the same challenges to an automated framework as a polished metal surface and it is possible to match SURF descriptors from multiple views. Physical features in the images such as cracks of varying widths actually aid the feature descriptors as these areas of the image vary significantly statistically which in turn aids in creating unique visual words which further aid in discriminating cracked versus non cracked tiles. Features of interest or keypoints may be found by a number of methods such as automated detection based on saliency (uniqueness) or by uniformly sampling the image and extracting features from each location on a regular grid. Due to the feature poor nature of the uncracked class in our data set, we use the latter method for selecting keypoint locations.

The visual words or vocabulary for each class are computed by performing multidimensional clustering over all of the feature point descriptors extracted from all the example imagery (for all object classes) to produce a set of feature clusters that characterise the overall feature space in a fixed dimension. Individual instances (each image in training and test sets) are then encoded by calculating the feature vector for each image in turn and matching the extracted features with the vocabulary to create a histogram of occurrences of each "visual" word. This fixed length "distribution of features" is used to differentiate between positive and negative classes and is required to train a classifier as well as to evaluate new instances once the classifier has been trained.

Individual bricks layers are too large to directly apply the BoVW model so each brick is further sub-

²BRISK: Binary Robust Invariant Scalable Keypoints and FAST: Features from accelerated segment test

Figure 3: Crack likelihood map overlaid on the original image of a single brick layer. Red indicates high probability of crack and blue represents low probability of crack. Each square represents an evaluated image tile.

divided into smaller images. For the negative class, each brick layer image is simply divided into a 5 by 5 array of smaller tiles. For the positive class, cracks are manually cropped from the image and further cropped and segmented so that the size of the cracked tiles is similar to the size of the uncracked examples. The positive tiles are required to be a similar size to the negative tiles as unseen data is evaluated using a sliding window approach of overlapping tiles with fixed size. For this work, we considered 50% overlap between each window. The size of the tiles and the overlap between them determine the granularity of any resulting probability maps generated. Chen and Jahanshahi (2017) employs a similar strategy of scanning individual frames with a sliding window approach and use a 120×120 pixel window with a step size of 8 pixels. Schmugge et al. (2016) also use this approach and employ a 224×224 pixel window with 75% overlap to classify individual frames. Each of these approaches also classify the same features from different view points in different frames and aggregate the results to return more accurate classifications. The small step sizes and large overlaps across multiple frames result in quite fine classification maps which work well when considering low resolution frames. When considering large single

images such as chanoramas which sometimes exceed 180 Megapixels, considering a larger number of smaller tiles with larger overlap would be exceedingly computationally expensive. Other approaches to crack classification consider classifying every single pixel as cracked or not cracked such as the frameworks presented by Schmugge et al. (2014), Schmugge et al. (2017) and Oliveira and Correia (2014) but these approaches would be prohibitively computationally expensive for chanoramas.

A feature vector is calculated for each tile as described above and the resulting array of feature vectors serve as the instances for training a classifier. As the number of uncracked tiles vastly exceeds the number of cracked tiles (approximately 7000 uncracked tiles : 100 cracked tiles), a random subset of uncracked tiles was chosen for training. Again, due to the feature poor nature of the negative tiles, the number of negative instances was chosen to be 750 to enable the characteristics of the negative class to be sufficiently "learnt". It may seem counterintuitive to have a much larger number of negative image instances than positive image instances when creating a balanced data set but this is required as the BoVW model only considers strong features (detected SURF keypoints include a strength or saliency metric for every keypoint detected) therefore additional image instances are required to create a balanced array of keypoints suitable for clustering. The number of negative image instances was determined empirically to give the best results.

When the feature extraction process is complete, the next step is to classify the image as cracked or uncracked. The classifier of choice for this type of problem is a Support Vector Machine (SVM) due to its excellent performance in similar binary classification problems (Kundegorski et al. (2016)). SVMs are trained by determining the optimal separating hyperplane for the two classes in either instance space (linear SVMs) or in an inner product subspace of a high dimensional feature space (Radial Basis Function kernels, Polynomial kernels, Sigmoid kernels, etc.). New instances are classified by evaluating their relative position to this hyperplane. For a more detailed introduction to SVMs see Cristianini and Shawe-Taylor (2000).

To determine whether a full brick layer is a cracked or uncracked using this classifier, we consider a two stage approach. We first apply a sliding window approach and classify a series of overlapping regions or tiles in both the x and y directions

ensuring that all regions of the brick are classified. The result of this first stage is a coarse spatial crack probability map for the brick layer like the example shown in Figure 3 where red denotes the highest probability of a crack while blue represents the lowest probability. A coarse binary detection image for the positions of cracks in the bricks may be obtained by thresholding this probability map. To determine if there is a crack in the full brick, additional post-processing is required. Some bricks have a small number of individual tiles incorrectly labeled as being cracked when in fact they are not. These false positives are often caused by seams in the chanorama or in situations where features such as the brick interface are not being entirely cropped out. The BoVW only classifies individual tiles and does not consider their spatial relationship. To address this we consider the connectivity of the positive instances in the binary detection image. A brick is labelled as cracked if the number of connected positive instances in the binary detection mask exceeds a certain threshold. For this paper we focus on full length axial or circumferential cracks in the main brick area. A second specialist classifier could be trained to deal with smaller cracks and interfaces. These full length cracks will be a number of tiles in length and the tiles will all be connected. Thus, we impose a minimum threshold on the size of each cluster of positive detections and we apply a minimum threshold on the total number of detections in a single brick image. Suitable thresholds can be determined from the training data using an optimization method such as a Genetic Algorithm Beasley et al. (1993); Deb (1999) to maximise performance metrics when testing on a subset of the training data. The parameters of interest to optimise are the minimum size and connectivity of a cluster of pixels in order to maximise performance metrics for labelling an entire brick layer as cracked or not cracked.

An alternative decision tool to assist in the detection of cracks may be obtained by considering the probabilistic output of the SVM classifier for each individual tile classified during the sliding window phase and creating a heatmap of crack likelihood. This map can then be scaled to match the size of the original chanorama and using an appropriate transparent colourmap, overlaid on the chanorama image as shown in Figure 4. This provides an easy to use visual tool that immediately draws a user's attention to cracks or interfaces which appear like cracks. Bonnín-Pascual and Ortiz (2010) consider

Figure 4: Automated crack detection used to generate probability heat map indicating likelihood of a crack in the brick.

a similar approach using heatmapping to indicate the probability of successful classification of corrosion on the metal surfaces of a ship’s hull. The overlay in this case is opaque which obscures the inspection surface in contrast to our approach which allows analysis to still be performed while guided by the heatmap.

4.3. Trepanned hole detection in chanoramas

In order to automatically detect trepanned holes in each brick layer of the chanorama, a technique known as the Hough Transform (HT) is used (Hough (2006)). The HT was originally proposed for detecting straight lines in images (Hough (2006)), was later extended in Duda and Hart (1972) for circle detection and detection of curves (Duda and Hart (1972); Merlin and Farber (1975)), and, in Ballard (1981), it was generalized for detecting any arbitrary shape. The standard HT for circle detection operates on each edge point in an image following the application of an edge detector such as Sobel, Canny, Roberts, Laplacian of Gaussian, or any suitable alternative, see Gonzalez and Woods (2006) for more examples. The HT treats each edge pixel as the centre of candidate circles of radius r and increments all pixels coinciding with the circumference of each candidate circle in the so-called Hough Parameter Space (HPS). If an edge pixel does in fact lie on a circle of radius r , then the corresponding pixel at the true circle’s centre will be incremented many times in this process and can thus be easily detected as a hot-spot in the HPS by thresholding or searching for local maxima. If we seek circles of a known radius, r , then the HPS is 2-dimensional. If circles of multiple (or unknown) radii are required to be detected, the HPS becomes 3-dimensional where each 2D accumulator array in the 3D space is used to test for a circle with some

unique radius r . For our application, we designed the Hough Transform to detect circles whose radii lie in the range $[40, 85]$ pixels which accounts for all dimensions of trepanned holes observed in our data.

While the Hough Transform works extremely well for detecting trepanned holes, some lens distortion and other factors affecting data acquisition can cause the sought features to appear elongated in some chanoramas. As a result, the output of the Hough Transform cannot be used directly to provide an accurate measurement of the dimensions of such trepanned holes. To obtain more accurate sizes for each detected trepanned hole we use a variation of the Hough Transform designed for ellipse detection Xie and Ji (2002). This approach fits an ellipse by examining all possible major axes (all pairs of points) before computing the minor axis using the Hough Transform. The ellipse detection approach is applied locally at the location of each detected trepanned hole. Of course, in theory, the ellipse detector could be applied from the outset without the need to first apply the Hough Transform for detection. However, since the ellipse detection algorithm is attempting to fit 5 parameters, it is computationally expensive making it unfeasible for application from the outset. Thus, we found that using the Hough Transform to detect the location of trepanned holes before measuring them using a localized ellipse detector to be the best solution in terms of computation efficiency and accuracy.

5. Results and Discussion

In this section we present a rigorous evaluation of the techniques proposed in Section 4 when used to detect the presence of cracks and trepanned holes in

Figure 5: Trepanned hole detection (a) original image (b) circles detected overlaid on original.

chanorama images created using real AGR in-core inspection videos. We also assess the performance of the automated system for measuring the dimensions of trepanned holes when compared with manually prepared ground truth measurements before concluding with a case study to demonstrate the application of the proposed techniques.

The detection of cracks and trepanned holes can each be considered in turn as a separate binary classification problem. Therefore, we can label the two classes as either positive i.e. crack/trepanned hole present and negative i.e. crack/trepanned hole absent. Popular choices of performance metrics for binary classification problems include true positive rate (recall), false positive rate (fall out), precision, accuracy, etc. Another suitable performance metric for algorithm evaluation is the F_β Score. This is the weighted harmonic mean of precision and recall and is defined as

$$F_\beta = (1 + \beta^2) \frac{\text{Precision} \times \text{Recall}}{(\beta^2 \times \text{Precision}) + \text{Recall}}$$

where Precision and Recall are defined by

$$\text{Precision} = \frac{\Sigma \text{True Positives}}{\Sigma \text{True Positives} + \Sigma \text{False Positives}}$$

and

$$\text{Recall} = \frac{\Sigma \text{True Positives}}{\Sigma \text{True Positives} + \Sigma \text{False Negatives}}$$

The F_β score can be adapted for a particular mode of operation by changing the dimensionless

parameter β to achieve either higher precision or recall performance at the cost of a decrease in performance of the other metric. When β is greater than 1, recall is weighted more heavily than precision and when β is less than 1 is, precision is weighted more heavily than recall. The most commonly used variation is the F_1 score in which precision and recall are weighted equally. Accuracy may also be used to evaluate the performance of the crack detection frameworks as we know the number of uncracked bricks. Accuracy is defined as

$$\text{Accuracy} = \frac{\Sigma \text{True Positives} + \Sigma \text{True Negatives}}{\Sigma \text{Positives} + \Sigma \text{Negatives}}$$

To evaluate crack detection performance, we considered a data set of 83 chanoramas. This contains 64 cracked bricks and 766 uncracked bricks. Unlike a typical machine learning problem, we did not apply the usual partition of 2/3 training data and 1/3 test data. We instead considered a test set of 542 bricks of which 43 were cracked and 499 were not. This was due to the scarcity of cracked bricks thus a large test set was needed to allow an accurate evaluation of the performance of the crack detection framework for full bricks and fuel channels. The remaining 21 cracked bricks and 267 uncracked bricks were used to construct the training set. Given the training data bricks were further subdivided into cracked/uncracked tiles and there remained sufficient data to train a robust BoVW and SVM classifier. The results of using the trained system to recognize and highlight cracked vs uncracked bricks are summarized in Table 1.

Accuracy was calculated to be 92.0% with a precision of 48.8% and recall of 90.7% yielding F_1 and

Figure 6: Examples of automated trepanned hole detection. Shown on the left of each example is the original image, in the centre the binary mask used by the Hough Transform for ellipse detection and on the right the detected ellipse is superimposed on the original image.

F_2 scores of 63.4% and 66% respectively. The precision and F Scores are relatively low for this algorithm but these metrics offer a poor representation of the algorithm’s overall capability. The low precision is due to the relatively large number of negative instances in the test set resulting in a proportionately larger number of false positives. This is a common problem for datasets with a class imbalance Japkowicz (2000). In a binary classification problem like this, the trade off for an improvement in precision is a decrease in recall. The system was designed to achieve a very high recall by tuning the parameters to maximize F_2 score. This was considered more important than high precision as

		Actual	
		Positive	Negative
Prediction	Positive	TP = 39	FP = 41
	Negative	FN = 4	TN = 458

Table 1: Crack Detection Analysis.

this particular application requires a very low false negative rate. If accuracy alone was considered as a measurement of this system, one would conclude excellent performance but this is due again to the class imbalance and this also demonstrates why multiple performance metrics need to be considered for a problem like this and more importance assigned to one or two for the desired mode of operation.

Furthermore, the system yields a false positive rate of 8.2% which is sufficiently low to make the proposed approach a useful tool for decision support. In context, of the 542 bricks evaluated in this study, only 80 bricks (39 true positives and 41 false positives) would need to be passed to an operator for further analysis which would offer a sufficient time saving compared to manually evaluating the full set of bricks. If the system highlights a large number of brick layers for further analysis and they do not contain cracks, the level of decision support offered by the system would be unreliable and operators would not be likely to trust or adopt it. The framework therefore remains very useful as the large majority of both cracked and uncracked bricks are correctly labeled.

The output of the BoVW classifier for each overlapping tile can also be used to generate crack likelihood heatmaps and an example of this is given in Figure 4.

An example of the output of the trepanned hole detection method is demonstrated in Figure 5. The original image (Figure 5(a)) used for testing contains two trepanned holes, one above and one below a brick interface. The proposed method is clearly able to detect both trepanned holes accurately as demonstrated in Figure 5(b) where the result of overlaying the resultant binary detection image on top of the original is shown. The algorithm was applied to the same data set of 83 chanoramas with a total of 45 trepanned holes and the results are summarized in Table 2. Precision was calculated to be 93.33% with a recall of 93.33% yielding an F_1 score of 93.33%. The number of true negatives is the number of not a trepanned holes in bricks and is not a meaningful metric. It should be noted that some trepanned holes are detected multiple times when sweeping through the range of circle sizes being sought. In these situations, only one detection is counted during the evaluation. In general, these initial detection results are promising and the small number of false positives and false negatives could be easily corrected through a small amount of manual intervention.

		Actual	
		Positive	Negative
Prediction	Positive	TP = 42	FP = 3
	Negative	FN = 3	TN: NA

Table 2: Trepanned Hole Detection Analysis.

In addition to evaluating the ability of the HT based trepanned hole detection algorithm to reliably recognize these features only when they are present, we have also manually measured the major and minor axis of each trepanned hole in the dataset and compared these dimensions with those obtained by the proposed automated localized ellipse detection approach. The result of this analysis is shown in Table 3 where TH1 denotes trepanned hole 1 and so forth. To ease readability, Table 3 shows the error between width and height measurements made by the proposed system and the manually produced ground truth. It is clear by reference to Table 3 that in many cases the automated approach accurately measures the dimensions of trepanned holes seen in the data to within a few pixels (≤ 5) error when compared with the manually produced ground truth measurements. In fact, the average error was calculated to be 5 pixels for the entire data set considered. It should also be noted that the largest errors appear in Trepanned Holes 40 and 42. The reason is that due to saturation and non-uniform illumination, the ellipse detector has misidentified other features (i.e. not the correctly identified trepanned hole) in the images and has measured these in error. An example of this issue is shown in Figure 6(c) alongside examples of the techniques working well in Figure 6(a) and Figure 6(b) for reference.

Having managed to accurately detect and measure cracks and trepanned holes in the chanoramas tested, we show in Figure 7 an example of how these objects can be accurately scaled using the proposed techniques. The image in Figure 7 shows an individual brick layer containing a single crack. Six trepanned holes were also detected in this chanorama and these have been cropped and placed below the brick layer image as a reference. These correspond to TH 15 - TH 20 in Table 3. The crack in the image shown in Figure 7 has been measured to be 4510 pixels long and 12 pixels wide at its widest point. The average width of the trepanned holes detected in this chanorama is 95 pixels and

Figure 7: Image of brick layer showing brick layer dimensions as well as a crack and six trepanned holes cropped from other brick layers in the same chanorama.

is known to correspond to approximately 23mm in real metric units. Using this known feature for scale, we can say that 1 image pixel equates to 0.242mm. Thus, we can determine that the crack in the image is 2.9 mm wide at its widest point. The average height of the trepanned holes detected in this chanorama is 135 pixels. Using a similar approach we calculate that 1 image pixel equates to 0.171mm. Using this scaling factor for height we calculate true metric height of the crack to be 778mm. There will be uncertainty in the physical dimensions of the trepanned holes which will also introduce some uncertainty into the scaling factors calculated based off these dimensions. This uncertainty has been studied and it has been found to have no significant impact on the calculated widths of cracks.

6. Conclusion

In this paper we have presented a new, automatic method for analysing video footage captured during routine inspection of the AGR cores in the UK's fleet of nuclear power plants. The techniques provide new decision support tools for remote visual inspection of these critical assets. The proposed

	Manual		Automatic		Error			Manual		Automatic		Error	
	Width (pixels)	Height (pixels)	Width (pixels)	Height (pixels)	Width (pixels)	Height (pixels)		Width (pixels)	Height (pixels)	Width (pixels)	Height (pixels)	Width (pixels)	Height (pixels)
TH 1	106	131	102	131	-4	0	TH 22	116	107	112	118	-4	11
TH 2	98	135	96	134	-2	-1	TH 23	70	91	94	103	24	12
TH 3	110	129	106	129	-4	0	TH 24	86	125	82	125	-4	0
TH 4	98	136	80	131	-18	-5	TH 25	95	124	88	128	-7	4
TH 5	105	127	114	135	9	8	TH 26	86	128	92	130	6	2
TH 6	96	138	94	135	-2	-3	TH 27	88	128	78	124	-10	-4
TH 7	105	133	104	135	-1	2	TH 28	85	130	74	127	-11	-3
TH 8	96	140	90	137	-6	-3	TH 29	85	129	100	125	15	-4
TH 9	106	140	92	131	-14	-9	TH 30	101	135	102	136	1	1
TH 10	104	137	106	135	2	-2	TH 31	84	129	68	128	-16	-1
TH 11	98	139	94	136	-4	-3	TH 32	87	135	76	133	-11	-2
TH 12	103	139	102	132	-1	-7	TH 33	100	130	98	129	-2	-1
TH 13	104	132	104	137	0	5	TH 34	86	134	90	134	4	0
TH 14	95	138	102	139	7	1	TH 35	145	155	140	144	-5	-11
TH 15	98	130	100	131	2	1	TH 36	127	159	118	157	-9	-2
TH 16	97	133	92	132	-5	-1	TH 37	148	159	148	156	0	-3
TH 17	96	135	96	134	0	-1	TH 38	127	159	104	155	-23	-4
TH 18	95	140	92	136	-3	-4	TH 39	145	158	148	159	3	1
TH 19	97	141	94	138	-3	-3	TH 40	145	160	58	176	-87	16
TH 20	100	142	96	138	-4	-4	TH 41	149	158	146	159	-3	1
TH 21	118	106	112	124	-6	18	TH 42	142	159	58	177	-84	18

Table 3: Trepanned Hole Detection Analysis.

system uses advanced image processing techniques known as a Bag of Visual Words which we have designed and customised to automatically detect the presence of cracks in AGR fuel channels in a fast and effective way. We have also presented a new approach for detecting channel features known as trepanned holes whose metric dimensions are also known and can thus be used to convert and scale pixel based crack measurements into metric values. The results presented in this paper demonstrate the potential of the proposed techniques when applied to images captured from within real nuclear reactor cores whose appearance and contents are highly variable and present a challenge for designing any automated processing system. The proposed system has the advantage that it operates in a robust and repeatable way which removes the subjectivity associated with manually analysing image data. The techniques described in this paper will initially be trialed by the same team of specialists who currently analyse all image and video data captured during inspection of the UKs fleet of nuclear reactors. In parallel we will continue to develop the tools described and will work to elicit expert knowledge in order that we can design suitable algorithms to detect other channel features for scaling and to automatically categorize cracks based on the outcomes of our automated crack detection and image analysis routines.

Acknowledgement

The authors would like to thank EDF Energy for providing the data used in this paper.

References

- Adhikari R, Moselhi O, Bagchi A. Image-based retrieval of concrete crack properties for bridge inspection. *Automation in construction* 2014;39:180–94.
- Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv preprint arXiv:1511.00561* 2015;.
- Ballard DH. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition* 1981;13(2):111–22. doi:[http://dx.doi.org/10.1016/0031-3203\(81\)90009-1](http://dx.doi.org/10.1016/0031-3203(81)90009-1).
- Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features. *Computer vision ECCV 2006* 2006;:404–17.
- Beasley D, Bull DR, Martin RR. An overview of genetic algorithms: Part 1, fundamentals. *University computing* 1993;15(2):56–69.
- Bonnín-Pascual F, Ortiz A. Detection of cracks and corrosion for automated vessels visual inspection. In: *CCIA*. 2010. p. 111–20.
- Bradford MR. Reactor inspection and monitoring and its role in the reactor safety case. In: *The 4th EDF Energy Nuclear Graphite Symposium*. 2015. p. 331–40.
- Chatfield K, Lempitsky VS, Vedaldi A, Zisserman A. The devil is in the details: an evaluation of recent feature encoding methods. In: *BMVC*. volume 2. p. 8.
- Chen FC, Jahanshahi MR. Nb-cnn: deep learning-based crack detection using convolutional neural network and naive bayes data fusion. *IEEE Transactions on Industrial Electronics* 2017;65(5):4392–400.
- Chen FC, Jahanshahi MR, Wu RT, Joffe C. A texture-based video processing methodology using bayesian data

- fusion for autonomous crack detection on metallic surfaces. *Computer-Aided Civil and Infrastructure Engineering* 2017;32(4):271–87.
- Cho BH, Byun SH, Shin CH, Yang JB, Song SI, Oh JM. Ke-provt: underwater robotic system for visual inspection of nuclear reactor internals. *Nuclear engineering and design* 2004;231(3):327–35.
- Choi CH, Kim SH, Kim CH, Seo YC, Shin HC, Lee SU, Jung KM, Jung SH, Choi YS. The robots for nuclear power plants. In: *Conference on Korea Nuclear Society* 2005. 2005. .
- Coble JB, Ramuhalli P, Bond LJ, Hines J, Upadhyaya B. Prognostics and health management in nuclear power plants: a review of technologies and applications. *National Technical Information Service* 2012;.
- Cole-Baker A, Reed J. Measurement of AGR graphite fuel brick shrinkage and channel distortion; *Royal Society of Chemistry*; volume 309. p. 201–8.
- Cristianini N, Shawe-Taylor J. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.
- Csurka G, Dance C, Fan L, Willamowski J, Bray C. Visual categorization with bags of keypoints. In: *Workshop on statistical learning in computer vision, ECCV*. Prague; volume 1; 2004. p. 1–2.
- Deb K. *An introduction to genetic algorithms*. Sadhana 1999;24(4-5):293–315.
- Devereux M, Murray P, West G, Buckley-Mellor S, Cocks G, Lynch C, Fletcher A. Automated analysis of AGR fuel channel inspection videos. In: *6th EDF Energy Nuclear Graphite Conference*. 2018. .
- Duda RO, Hart PE. Use of the hough transformation to detect lines and curves in pictures. *Commun ACM* 1972;15(1):11–5. doi:10.1145/361237.361242.
- Glass S, Levesque M, Engels G, Klahn F, Fairbrother D. Underwater robotic tools for nuclear vessel and pipe examination. In: *Proceedings ANS 8th International Topical Meeting on Robotics and Remote Systems*. 1999. p. 25–9.
- Gonzalez RC, Woods RE. *Digital Image Processing* (3rd Edition). Prentice-Hall, Inc., 2006.
- Hough PVC. Machine analysis of bubble chamber pictures. In: *International Conference on High Energy Accelerators and Instrumentation*. 2006. doi:citeulike-article-id:476172.
- Jahanshahi MR, Masri SF. Adaptive vision-based crack detection using 3d scene reconstruction for condition assessment of structures. *Automation in Construction* 2012;22:567–76.
- Japkowicz N. The class imbalance problem: Significance and strategies. In: *Proc. of the International Conf. on Artificial Intelligence*. 2000. .
- Kundegorski ME, Akçay S, Devereux M, Mouton A, Breckon TP. On using feature descriptors as visual words for object detection within x-ray baggage security screening. In: *7th International Conference on Imaging for Crime Detection and Prevention (ICDP 2016)*. IET; 2016. .
- Lee SU, Choi YS, Jeong KM, Jung S. Development of a tele-operated underwater robotic system for maintaining a light-water type power reactor. In: *2006 SICE-ICASE International Joint Conference*. IEEE; 2006. p. 3017–21.
- Merlin PM, Farber DJ. A parallel mechanism for detecting curves in pictures. *IEEE Transactions on Computers* 1975;C-24(1):96–8. doi:10.1109/T-C.1975.224087.
- Murray P, West G, Law K, Buckley-Mellor S, Cocks G, Lynch C. Automated video processing and image analysis software to support visual inspection of AGR cores. 2016a. .
- Murray P, West G, Marshall S, McArthur S. Automated in-core image generation from video to aid visual inspection of nuclear power plant cores. *Nuclear Engineering and Design* 2016b;300:57–66.
- Nilsback ME, Zisserman A. A visual vocabulary for flower classification. In: *Computer Vision and Pattern Recognition*. IEEE; volume 2; 2006. p. 1447–54.
- Oliveira H, Correia PL. Crackit - an image processing toolbox for crack detection and characterization. In: *2014 IEEE international conference on image processing (ICIP)*. IEEE; 2014. p. 798–802.
- Prasanna P, Dana KJ, Gucunski N, Basily BB, La HM, Lim JS, Parvardeh H. Automated crack detection on concrete bridges. *IEEE Transactions on automation science and engineering* 2014;13(2):591–9.
- Schmugge SJ, Nguyen NR, Thao C, Lindberg J, Grizzi R, Joffe C, Shin MC. Automatic detection of cracks during power plant inspection. In: *Proceedings of the 2014 3rd International Conference on Applied Robotics for the Power Industry*. IEEE; 2014. p. 1–5.
- Schmugge SJ, Rice L, Lindberg J, Grizzi R, Joffey C, Shin MC. Crack segmentation by leveraging multiple frames of varying illumination. In: *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE; 2017. p. 1045–53.
- Schmugge SJ, Rice L, Nguyen NR, Lindberg J, Grizzi R, Joffe C, Shin MC. Detection of cracks in nuclear power plant using spatial-temporal grouping of local patches. In: *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE; 2016. p. 1–7.
- Sivic J, Russell BC, Efros AA, Zisserman A, Freeman WT. Discovering objects and their location in images. In: *Proceedings of the IEEE International Conference on Computer Vision*. volume I; 2005. p. 370–7.
- West G, Murray P, Marshall S, McArthur S. Improved visual inspection of advanced gas-cooled reactor fuel channels 2015a;6(Special Issue - Nuclear Energy).
- West G, Murray P, Marshall S, McArthur S. Improved visual inspection of advanced gas-cooled reactor fuel channels. *International Journal of Prognostics and Health Management* 2015b;6(Special Issue).
- West GM, McArthur SDJ, Towle D. Knowledge-directed characterization of nuclear power plant reactor core parameters. *Nuclear Engineering and Design* 2011;241(9):4013–25. doi:http://dx.doi.org/10.1016/j.nucengdes.2011.08.013.
- Xie Y, Ji Q. A new efficient ellipse detection method. In: *Pattern Recognition, 2002. Proceedings. 16th International Conference on*. volume 2; 2002. p. 957–60.

Dr. Michael Devereux received a BSc in applied physics and Ph.D. in applied mathematics from the University of Limerick in Ireland, in 2007 and 2012, respectively. He has held postdoctoral posts in Durham University, University of Bristol, Heriot-Watt University in the fields of signal and image processing, machine learning and computer vision.

Dr. Paul Murray received the MEng and PhD degrees in Electronic and Electrical Engineering

from the University of Strathclyde, Glasgow, UK, in 2008 and 2012, respectively. He is currently a Lecturer in the Department of Electronic and Electrical Engineering at the University of Strathclyde. His research interests include signal and image processing, image stitching, mathematical morphology, and machine learning for image processing in a diverse range of applications ranging from agri-tech to nuclear plant inspection.

Dr. Graeme West received the B.Eng. (Hons.) and Ph.D. degrees from the University of Strathclyde, Glasgow, U.K., in 1998 and 2002, respectively. He is a Senior Lecturer in the Department of Electronic and Electrical Engineering at the University of Strathclyde. His research interests include intelligent system applications in condition monitoring, diagnostics and prognostics, particularly applied to the nuclear power generation industry.