Automated image stitching for enhanced visual inspections of nuclear power stations

Paul Murray, Graeme West, Stephen Marshall and Stephen McArthur
Department of Electronic and Electrical Engineering, University of Strathclyde
Glasgow, G1 1XW, Scotland
+44 (0) 141 548 3535

paul.murray@strath.ac.uk

Abstract

In the UK, visual inspection of the fuel channels of the Advanced Gas-cooled Reactor (AGR) nuclear power stations forms an integral part of understanding the health of the reactor cores. During a statutory outage, video footage of the inside of selected fuel channels is recorded. Features of interest and anomalies are manually identified by an expert who extracts frames from the video to create a composite image for the feature of interest. This is a laborious and time consuming process which can be costly to station operators who must produce these images before returning the station to service. This paper describes an automatic technique capable of generating a 2D image of the entire internal bore of the channel. The technique uses the position of the camera coupled with advanced image processing techniques to generate a high-resolution image of the whole channel. This allows surface details to be viewed in relation to each other, and the rest of the channel, while facilitating a direct comparison of any anomalies over time. In addition, the time taken by this automated technique to produce a full core image is a fraction of that taken to manually stitch an image for a much smaller area.

1 Introduction

Seven out of the eight nuclear power stations in operation in the UK are of the Advanced Gas-cooled Reactor (AGR) design. A major factor that determines the lifetime of the AGR stations is the condition of the graphite reactor core. The graphite core is constructed from a large number of moderator bricks, which are configured to form various channels to accommodate fuel, control rods and control the passage of the CO_2 gas coolant. As well as providing the physical structure, the core also acts as a neutron moderator. Under fast neutron irradiation and through the process of radiolytic oxidation, the physical properties, such as the strength, Young's modulus and dimensions all change. These changes are tracked through condition monitoring systems, which gather data during routine station operation, and through measurement and inspection of the core during statutory outages.

When the channels are refuelled during an outage, there is an opportunity to inspect the inside of the fuel channels to determine their condition. One inspection technique involves lowering a TV camera into a fuel channel and recording footage of the inner surface of the channel. By fixing the camera to point in a set orientation and raising it through the core, a single strip of the channel can be covered. Rotating the camera and repeating this process allows full coverage of the channel to be built up. When rotating the camera, the engineer manipulating the tool ensures some overlap between neighbouring scans so that the entire channel is inspected and that no part is unseen. Throughout this process, video data of each vertical scan is viewed and recorded for offline analysis and processing. If anything unusual is identified in the video, snapshots of the video are taken and manually assembled to generate a montage of the region of interest. This is a critical part of the inspection process and a statement detailing the results of the visual inspection of each fuel channel examined is required in order to return the station to service.

Currently, montage generation of regions of interest is performed by an expert who watches the video, extracts frames of interest and stitches them together. This is a costly process for station operators who cannot return the station to service without the stitched image. One drawback with the manual process is that the time taken for any montage to be produced cannot be quantified in advance. Furthermore, while the montages provide a visual representation of the region of interest, the majority of the data that is captured is not used. This means that the stitched images that are produced only provide a very small snapshot of the entire channel since it is not practical to manually stitch a full channel image.

An alternative to the manual approach described above is to use image stitching techniques to automatically produce full channel images. Developing methods for seamlessly stitching together images of a scene from different viewpoints has been of interest to the image processing community for many years [1-9]. Example applications include the creation of digital maps and satellite images, stabilisation of digital cameras and the creation of ultra wide panoramic images which are generated by stitching together images captured in quick succession by a moving camera. In this paper we present an algorithm that can be used to generate full channel images of the inside of fuel channels within AGR cores of nuclear power stations. It is shown in this paper that the proposed algorithm is capable of producing high resolution images of the entire channel in a fraction of the time taken with the existing approach to produce a single montage of just the region of interest.

The remainder of this paper provides some background information on image stitching and image alignment techniques before presenting, in detail, a novel automatic method for generating full channel images from video data captured from inside fuel channels. The results of this novel algorithm are compared with the existing manual approach in terms of efficiency and the extent of the channel coverage. It is shown that the proposed approach offers a significant improvement over the manual stitching method in both of these aspects. The automatic method also offers additional benefits in terms of performing direct comparisons of images generated from the same channel at different points in time. The paper also presents a number of opportunities for further work and research throughout.

2 An overview of image alignment techniques

This section summarises various techniques which can be used to determine the alignment between images in order that they can be accurately stitched together. In [1] it is explained that alignment techniques are generally classified as direct methods or feature based methods. In this paper we also consider a third technique which is referred to as a knowledge based approach. A comprehensive tutorial on image stitching, which includes a detailed description of direct and feature based alignment techniques, can be found in [1].

2.1 A direct approach to image stitching

Given a pair of overlapping images, each of which contains a different view of some common scene, the aim is to estimate how one image can be geometrically warped such that the common points from both images can be used to accurately stitch the images together. The direct approach to image stitching performs a direct comparison of image pixels in both images to estimate the geometric transformation which must be applied to one image so that it aligns with another.

When designing any direct method for image stitching, suitable techniques for comparing the images based on pixel similarity must be identified. Error metrics like the Sum of Squared Differences (SSD) or the Sum of Absolute Differences (SAD) (see [1]) provide possible solutions. An alternative approach is to use cross correlation (see [10]) to compare the images to be stitched. When a suitable metric is established, the optimal shift/warping can be obtained by using methods to locate the minimum/maximum value in the output of the comparison. Direct comparisons can be computationally expensive; however, windowing techniques can be used to eliminate pixels from the calculation when they do not overlap with the image being stitched. Furthermore, hierarchical techniques and comparisons in the frequency domain can be used to improve computational efficiency [1].

2.2 A feature based approach to image stitching

A feature based approach to image stitching requires a set of features to be extracted from each of the images to be stitched. Features are then matched between images, and after so called outliers have been rejected, it is possible to use statistics of the position of the matched points to compute a geometric transform which can be used to align the images being stitched [1, 5].

In recent years, feature based approaches have dominated the image stitching literature due to a significant improvement in the reliability of methods for extracting and describing features [1, 5]. A large number of feature extractors have been published over the years including SIFT (Scale-Invariant Feature Transform) [3], SURF (Speeded Up Robust Features) [4, 7], and GLOH (Gradient Location-Orientation Histogram) [8], to name but a few. An extensive review of feature extractors is given in [6], and in [8], the authors compare 10 techniques including their own GLOH approach.

2.3 A knowledge based approach to image stitching

When there is knowledge about the position of the camera during imaging, there exists a trade-off between the required accuracy and complexity of image stitching routines and the accuracy of the positional estimation. If the positional estimate is very accurate there is no need to align and stitch images and vice versa. See Figure 1.

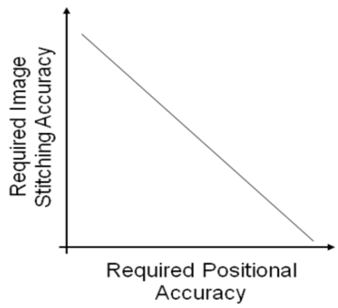

Figure 1 Trade-off between the accuracy of positional data and the accuracy of the stitching algorithm

In cases where the position, speed and orientation of the camera is known, it is possible to place each image from the video in the correct place in a global output image before simply blending them together (if necessary). This approach would not require overlapping images to be compared by a direct or feature based approach in order to find the transform required to align and position the images, that relationship would already be known. In such information rich cases, where the camera position and camera orientation estimations are consistently accurate, the knowledge-based approach can provide a simple yet effective solution with minimal computation and complexity.

3 2D image generation of the entire channel bore from videos

The video data captured during fuel channel inspections contains footage of vertical scans of the fuel channel with the camera facing six different orientations: North; North +60°; North +120°; North +180°; North +240° and North +300°. These orientations are chosen to ensure sufficient overlap of the scans in the spatial domain when capturing the data. For the data considered in this paper, the camera moves from the bottom of the channel to the top six times, and each time, the camera is rotated through 60° before a new scan is performed.

From an image stitching perspective, this poses two problems. The video sequence of each vertical scan must be accurately stitched to produce six vertical strips – one for each camera orientation (see Section 3.3). Then, these strips should be placed side-by-side and aligned vertically (see Section 3.4), before they are stitched together horizontally to produce the full channel image (see Section 3.5).

The work proposed in this paper aims to automate the process of generating the full channel image from available video data. This requires key pieces of information to be known, such as the camera orientation, the frames of the video which contain scan data as well as the vertical and horizontal alignment parameters required for stitching the data together. To achieve automation of this process, modules must be implemented to:

- identify frames of interest in the video (not all video data is required)
- determine the camera orientation during each vertical scan
- stitch each vertical scan to produce a vertical image strip
- align these vertical image strips in a global image
- estimate the extent to which each strip overlaps another so that the images may be correctly stitched horizontally.

A block diagram demonstrating the modules that need to be implemented and their current status in terms of whether they are manual, automated, or semi-automated is shown in Figure 2.

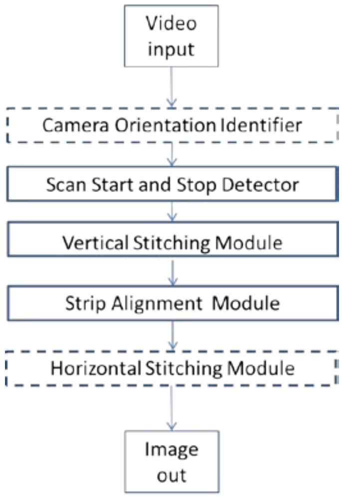

Figure 2 Overview of the process of converting a video sequence of scans from inside the fuel channels to a high resolution 2D image of the entire channel. Dashed boxes denote a manual input is required and solid boxes mean the module is either automatic or semi-automatic.

The remainder of this section describes the modules shown in Figure 2 and explains how automation either: is, or can be achieved.

3.1 Determining camera orientation

The camera orientation must be known so that the stitched vertical strips can be correctly related to each other prior to horizontal stitching. In addition to the image data that can be extracted from the video data captured during inspections, extra information, including the camera orientation, is contained within an overlay, which is present in every frame of the sequence. It is anticipated that optical character recognition (OCR) techniques [11] could be used to read the overlay in order to obtain this information automatically. For now, this information is manually entered into our system by a user who simply needs to glance at any frame in the video to obtain it. Techniques for automating this part of the process will be investigated in future.

3.2 Towards automatically identifying scan start and stop frames

In addition to scanning the channel in six different orientations during an inspection, a number of other features of the channel are inspected using the camera while the video is captured. As a result the video contains additional data that is not required for this process, so a means of identifying the start and end of the scans is required. A simple way to identify the frames which correspond to the start and stop of a scan is to watch the video and note the appropriate times before converting these times to frame numbers which can be entered into the system. However, this is a time consuming process when videos contain up to 2 hours of footage. A better approach would be to implement a method to support the identification of these start and end points without the requirement to view the entire video.

Certain properties of the fuel channel at each station are known. For example, the number of bricks used to construct the channel, and hence the number of brick interfaces (the join between two bricks) is known for each station. In the video data, the bricks generally appear as some light shade of grey in contrast to the brick interfaces which are generally dark. This is shown in the image in Figure 3(a) where the brick interface in the centre of the image is significantly darker than the imaged regions of the bricks above and below it. By plotting the average intensity of a 25 pixel high window located in the centre of every 5th frame (see Figure 3(b)), it is possible to identify frames which contain brick interfaces since the average intensity value within the window in these frames is lower than that of neighbouring frames. Using knowledge of the number of brick interfaces that are included in each scan (based on the number of bricks in the channel at that station), it is then possible to identify where a scan starts and ends by analysing a plot of the video intensity over time, see Figure 3(c).

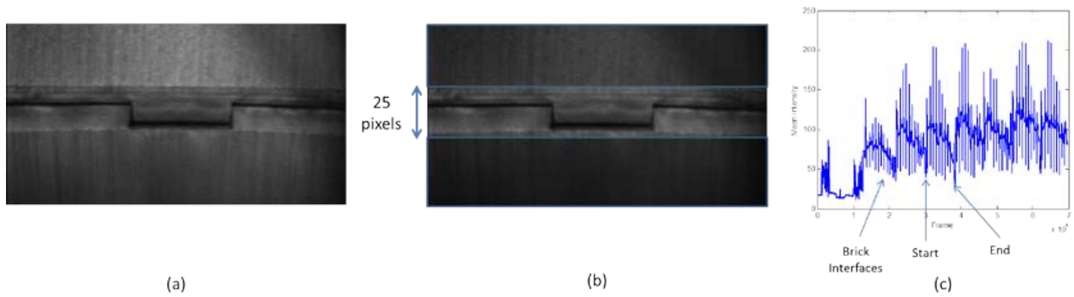

Figure 3 Video profile which can be analysed to identify frames which correspond to scans and their start and stop times. (a) Original image of a brick interface (b) Image and the window which is sampled and used to compute average for the video profile curve. (c) Video profile curve computed for an entire video sequence.

By observation of the video intensity profile shown in Figure 3(c), it is easy to see the frames which correspond to the brick interfaces – they appear as local minima along the profile. Currently, the video intensity profile is shown to the user who can manually interpret the plot and enter the frames of interest which correspond to each scan. For example, if the channel contains 10 bricks (9 interfaces), the user simply needs to identify the first brick interface of a scan, then find the next 8 interfaces and the corresponding frame number for each. The frame number which contains the end of the scan is the one containing the 9^{th} brick interface.

Techniques for automating the process further will be investigated in future. However, it should be noted that the time taken to generate and analyse a video intensity plot, like the one shown in Figure 3(c), is significantly less than the time it would take to watch up to 2 hours of video footage to obtain the same information.

3.3 Automatic generation of vertically stitched image strips

It was explained in Section 2.3 that if the relative position of the images is well known, then a knowledge based approach can be used to stitch the data without the need for sophisticated algorithms to compute the alignment between images. For the particular problem of generating vertical strips from each vertical scan of the fuel channel, the position of the camera in the channel, the speed at which it moves and its orientation, are either known or they can be estimated. Provided that the estimates of position, speed and orientation are accurate, then it is possible to create a large global output image for each vertical strip and place each image from the video in the correct place in the global image.

In practise, the data is acquired at a rate much greater than the speed at which the camera traverses the channel and hence there is significant overlap between successive frames. Therefore, only a small region from each frame is used when stacking the images one above the other in the output image. By manually designing the optimal window size, which is a window containing only the pixels that have changed between the current frame and the previous one, no blending is required. The images are simply abutted together, one on top of the other, in the global space. An example of this approach is shown in Figure 4.

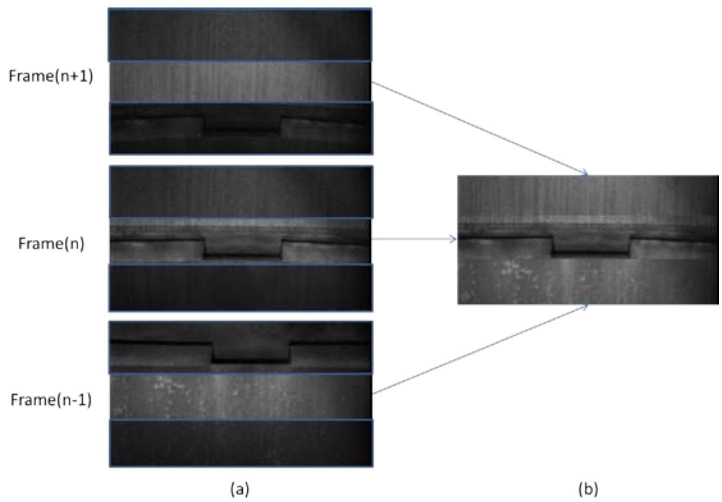

Figure 4 Example of generating a 2D image from 3 sequential frames captured during a bottom to top scan of the inside of the fuel channel. (a) Video sequence from which a small region of each frame is extracted and placed in the global output image. (b) Global output image created by stacking a cropped region from sequential frames in the video.

The knowledge based approach demonstrated in Figure 4 is used in our system to automatically stitch six vertical strips of the channel - one for each traversal: North; North +60°; North +120°; North +180°; North +240° and North +300°.

3.4 Automatic alignment of vertical strips in the global output image

When all six vertical strips have been stitched it is required that they are placed side by side and aligned vertically in a global image before the horizontal alignment and stitching is performed. This is not a trivial task since the vertical strips generally differ in height. For this reason, it is not sufficient to place the strips side by side without first computing the optimal vertical alignment between them. An example of two strips (rotated through 90° to save space) placed side by side before alignment is shown in Figure 5(a).

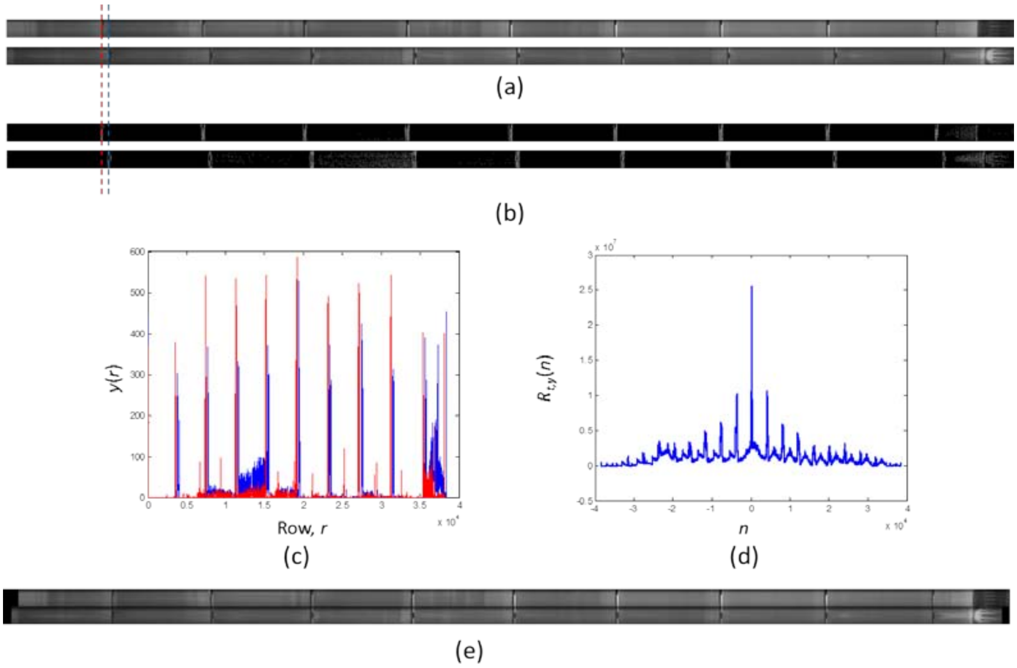

Figure 5 Example of computing vertical alignment of stitched strips. (a) Stitched strips of adjacent scans, N and N+60. (b) Stitched strips of N and N+60 shown in (a) after performing horizontal edge detection using a Sobel mask. (c) Sum of pixels in each row of N (red) and N+60 (blue) edge detected image. (d) Result of performing a cross correlation of the signals shown in (c) which are created by computing the sum of each row in the edge image. The position of the maximum value in the correlation result can be used to compute the offset between the signals. (e) N and N+60 strips aligned vertically in a global output image.

Ultimately, the vertical offset between the images is computed using cross correlation, see [10] for a full explanation. However, a series of pre-processing steps are performed first to improve the efficiency and accuracy of calculating the cross correlation to estimate the offset. It should be noted that all vertical strips are correlated with the same reference scan (North) in order to alleviate error drift. Each strip is also zero padded to the length of the longest strip to facilitate the computation of the cross correlation.

Figure 5 provides an illustrative example of the process used to compute the optimal offset between two strips. Firstly, a Sobel mask (see [10]) is used to locate horizontal edges in the original image (Figure 5(a)) for each vertical strip. This produces a binary edge image for each strip, like the ones shown in Figure 5(b). While the cross correlation can be computed using the original images, or the edge images, to estimate the offset which offers the best alignment, this process can be computationally intensive. For this reason, the two edge images are reduced to a 1D profile which is computed by summing each row in the edge image. Let, y denote a 1D vector and $y(r)$ denote a value in that vector which corresponds to row r in the edge image I. It is possible to compute a 1D profile y for any edge image I using:

$$y(r) = \sum_{c=0}^{W-1} I(r,c) \tag{1}$$

where c is the column index of image I of width W. Since black pixels are represented by a 0, and white pixels by a 1, the 1D profile has a high value in areas which have strong horizontal edges (the brick interfaces) and low values elsewhere – see Figure 5(c). Each edge image (one for each orientation) is reduced to a representative 1D signal using (1) before the cross correlation, $R_{t,y}$, of these 1D signals is computed,

$$R_{t,y}(n) = \sum_{m=-\infty}^{\infty} y[m]t[n+m] \tag{2}$$

In this work, t is fixed as the 1D profile for the North scan strip and y varies for all other scans which are compared to this one. This approach removes potential for error drift which could occur if each strip is compared to the one next to it. By computing the position of the maximum value in $R_{t,y}$, (see Figure 5(d)) the best alignment, a, may be easily computed,

$$a = \arg\max\left(R_{t,y}\right) \tag{3}$$

This estimate is used to align each image vertically in the global output image to produce a result similar to the one shown in Figure 5(e) prior to horizontal stitching.

3.5 Stitching the aligned vertical strips in the horizontal direction

When all six vertical strips have been automatically stitched together, placed side by side, and aligned vertically (see Figure 5), all that is left to do is identify the extent of the horizontal overlap between the vertical strips and to shift the images left or right until they overlap correctly. This process is currently achieved by manually shifting the vertical strip images using appropriate image editing software. While it is anticipated that a direct/feature based approach could be used to compute the alignment between images, this has not yet been investigated since the manual process is generally completed in less than 5 minutes.

3.6 Summary

This section has presented a method for converting lengthy video sequences captured during fuel channel inspections into high resolution 2D images of the entire channel bore. Techniques for robustly extracting the required frames from the video and accurately stitching them together have been explained in detail. The major benefit of this approach is that it allows the entire channel to be visually inspected in a single static image where all features are relative to each other. This offers a significant improvement over the current approach which requires watching long videos where it is difficult to visualise what is seen in one frame relative to a different frame 10 seconds later in the sequence. There is also less chance of details being missed when inspecting a static image than when watching a video captured using a moving camera.

4 Results and discussion

The method described in Section 3 has been tested on various videos from a number of fuel channels to generate high resolution 2D stitched images of the entire channel bore. The results are extremely promising and some of these are shown in Figure 6. It should be noted that only one full channel image has been shown here (Figure 6 (a)) to demonstrate the level of coverage that is achieved. The results shown in (Figure 6 (b), (c) and (d)) have been cropped from their respective full channel images to demonstrate the quality of the stitched image that is produced using our approach.

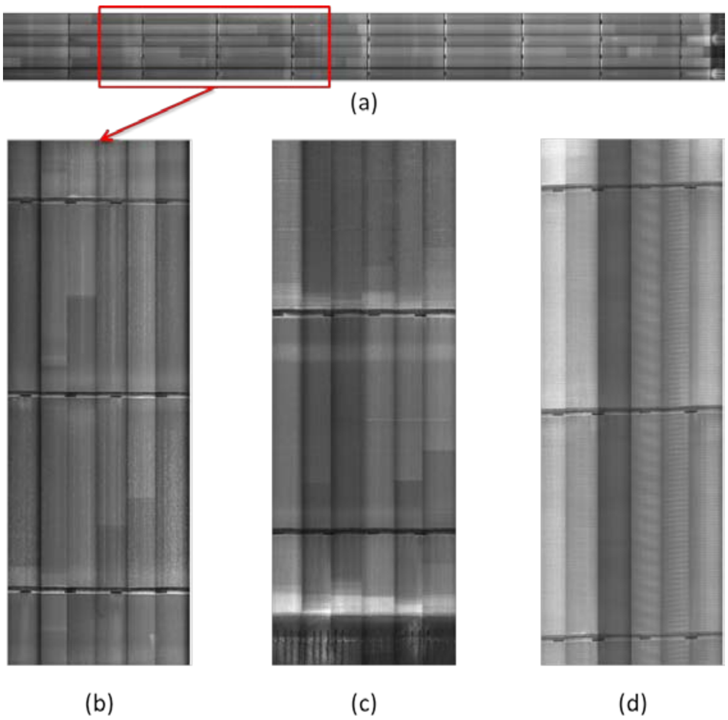

Figure 6 Results of using the approach described in this paper to generate full channel images for 3 different fuel channels at various stations in the UK. (a) full channel image, (b) region cropped from (a). (c) Cropped region from a full channel image showing methane holes (bottom) and other bricks. (d) Crop showing 2 full bricks from a full channel image.

The method described in this paper has two major advantages over the existing techniques used to analyse and stitch data from videos captured during fuel channel inspection.

1. The proposed approach provides full channel coverage such that data can be viewed and analysed in a single snapshot
2. The algorithm described here is much more efficient than the existing manual techniques for montage generation.

Currently, only regions in the video which are deemed to be of interest are stitched since it is not practical to manually produce a full channel image. The major advantage of viewing the data as a single snapshot (produced by the proposed method) is that any regions of interest can be viewed in the context of the entire channel. This allows, for example, the extent of any anomalies to be easily viewed and sized in relation to the height of the bricks that form a particular channel. Furthermore, the image can be easily inspected multiple times to ensure that all regions of interest are analysed and that nothing is missed. In future, it would be possible to generate a new video from inside the same channel, and from this, create a full channel montage which can be overlaid on historic data to perform a direct comparison.

The other major advantage of the proposed approach lies in the efficiency with which the images are generated when compared with the existing manual approach. Due to the nature of the manual approach, it is difficult to quantify exactly how long it will take to manually generate the images required by station operators to return the station to service. This varies from channel to channel and is dependent on the experience of the person tasked with generating the montage. The proposed approach, however, has been demonstrated to generate full channel images which represent 100% of the channel in less than 3 hours.

5 Conclusions

This paper presents a technique which can be used to convert lengthy video sequences captured during routine inspections of fuel channels into high resolution 2D images for visual inspection. The method relies on a combination of user inputs and sophisticated image processing techniques to produce a high resolution image representing 100% of the channel bore in less than 3 hours. This allows features of the channel to be seen relative to each other and provides an easy way to visualise the data and interactively make measurements if required. Furthermore, by performing the image stitching automatically, the proposed approach is less prone to human error and it minimises the possibility that any regions of interest would be missed on viewing. The images produced by the method described in our work also facilitate comparisons of inspected channels with historic images generated in previous inspections. The next steps in this work are to automate the horizontal alignment and stitching of the vertical strips. When this is completed, techniques for fully automating the process of obtaining the camera orientation using OCR will be investigated and a method for automatically analysing the video intensity profile will be developed.

Acknowledgements

The authors would like to thank EDF Energy for funding this work and for allowing us to present the paper. Thanks also to Rahul Summan and Gordon Dobie, University of Strathclyde, for a number of interesting discussions about image stitching.

References

1. R Szeliski, 'Image alignment and stitching a tutorial', Technical Report MSR-TR-2004-92, Microsoft Research, 2006.
2. M Brown, and D G Lowe, 'Recognising panoramas', Proceedings of 9th IEEE International Conference on Computer Vision, Nice, France, Vol.2, pp 1218 - 1225, October, 2003.
3. D G Lowe, 'Distinctive Image Features from Scale-Invariant Keypoints', International Journal of Computer Vision, Vol 60, no. 2, pp 91-110, 2004.
4. H Bay, T Tuytelaars and L Van Gool, 'SURF: Speeded Up Robust Features' Lecture notes in computer science, pp 404-417, 2006.
5. R Summan, 'Positioning for Mobile NDE Inspection Platforms', PhD Thesis, in Dept. Electronic and Electrical Engineering, University of Strathclyde, Glasgow, 2013.
6. T Tuytelaars and K Mikolajczyk, 'Local invariant feature detectors: a survey', Foundations and Trends in Computer Graphics and Vision, Vol 3, No 3, pp 177-280, 2008.
7. H Bay, A Ess, T Tuytelaars and L Van Gool, 'Speeded-Up Robust Features (SURF)' Computer Vision and Image Understanding, Vol 110, No 3, pp 346-359, 2008.
8. K Mikolajczyk and C Schmid, 'A performance evaluation of local descriptors', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 27, No. 10, pp 1615-1630, 2005.
9. G Dobie, R Summan, C MacLeod, S G Pierce, 'Visual odometry and image mosaicing for NDE' NDT & E International, (in press).
10. R C Gonzalez and R E Woods, 'Digital image processing', Upper Saddle River, NJ: Pearson Education, 2008.
11. Y Yang, X Lijia and C Chen, 'English Character Recognition Based on Feature Combination' Procedia Engineering,. Vol 24, No 0, pp 159-164, 2011.