3-D ADVANCED GAS-COOLED NUCLEAR REACTOR FUEL CHANNEL RECONSTRUCTION USING STRUCTURE-FROM-MOTION

Kristofer Law, Graeme West, Paul Murray
Department of Electronic and Electrical Engineering
University of Strathclyde
Royal College Building, Glasgow, G1 1XW
kristofer.law;graeme.west;paul.murray@strath.ac.uk

Chris Lynch
EDF Energy Generation
Barnett Way, Barnwood, Gloucester, GL4 3RS
chris.j.lynch@edf-energy.com

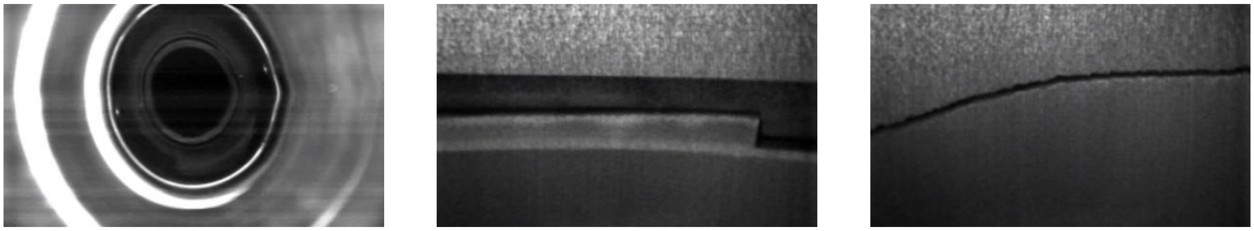
ABSTRACT

During planned, periodic outages, a selection of fuel channels within the UK fleet of Advanced Gas-cooled Reactor (AGR) cores are inspected using specialist tools which record video footage and other sensory data for each channel which undergoes inspection. Current visualization techniques comprise of manually produced montages by inspection engineers of points of interest (i.e. structural defects) and 2-D panoramic images of the fuel channels automatically produced using bespoke image stitching software. Both techniques however provide limited structural information due to the loss of depth data as a result of the image formation process. By recovering the depth information from the footage, a 3D model could be constructed and subsequently, allow for more accurate profiling of specific defects observed during inspection in addition to obtaining the fuel channels structure using existing footage. This work explores the preliminary application of a 3-D visualization technique known as Structure-from-Motion (SfM) which aims to obtain 3-D information by exploiting image correspondences across multiple viewpoints of the same scene in the RVI footage. This paper investigates the difficulties of applying state-of-the-art SfM to RVI footage and we present new techniques to improve feature correspondence searching in repetitive, non-descript environments.

Key Words: Visual inspection, Structure-from-Motion, Image processing, Feature detection

1 INTRODUCTION

Within the UK, there exists a fleet of seven Advanced Gas-cooled Reactors (AGR) which are approaching or have exceeded their original, conservative design lifetime estimations. As a result, it is important to rigorously inspect and monitor the condition of components which contribute to the design lifetime limitation. One of the major life-limiting components within the AGR design is the graphite reactor core. The graphite core is comprised of a sixteen sided, lattice structure of interconnected cylindrical graphite bricks which accommodate fuel and superheat high pressure carbon dioxide coolant gas using a sustained, controlled nuclear reaction. Gradually, fast neutron irradiation and radiolytic oxidation causes degradation of the graphite bricks resulting in physical and structural changes of the brick properties which can induce structural defects [1].



a) b) c)

Figure 1. Image stills of RVI footage. a) View obtained from the forward facing camera looking down the fuel channel, b) Brick interface between two vertically stacked graphite bricks after viewpoint is changed by mirror, c) Body of a graphite brick between brick layers which contains a structural defect. Note the imbalance of lighting conditions in b) and c) in addition to the repetitive texture of the graphite brick.

To visually assess the reactors' operational degradation, Remote Visual Inspection (RVI) collects video footage whilst the reactor is temporarily shut down for periodic examination. During inspection, the fuel is extracted from the channels and inspection engineers evaluate each pre-selected fuel channel using specialist equipment. The RVI process begins by first lowering the inspection tool in the fuel channel which faces forward as it descends. When the inspection tool reaches the bottom of the channel, a mirror is deployed to change the observable orientation of the camera to face the channel wall where successive scans of the channel wall is obtained by maneuvering the camera up and down at an orientation step size of $\pm 60°$. Due to the limited Field-of-View (FoV) of 70°, each scan has a small degree of overlap between recordings to ensure that full channel coverage is obtained [2]. Image stills of the different stages of this process can be viewed in Fig.1. Throughout the inspection, engineers will take note of regions of interest such as cracks within the examined fuel channel that will be further scrutinized. Once the inspection is complete, the RVI data is further analysed using image montages of the regions of interest. This is performed manually from additional footage which further observes defects often termed *crack following* footage. This is critical as the image montages provide valuable information about condition of the channel and highlight any structural defects and their properties before being able to return the station to service.

Historically, montage generation was a process undertaken by inspection engineers where frames of only the region of interest are selected and pieced together by hand. The process however takes time and the reactor cannnot be returned to power until the montages provided are quantified. Recently, research by Murray et al. [3] has allowed for autonomous generation of 360° 2-D panoramic images of an AGR fuel channel interior which accelerated this process dramatically whilst giving a full panoramic view of the fuel channels. In [2], the next logical step of visualization of the fuel channels was observed by inferring depth information through the use of a *pivot video* to illustrate the usefulness of 3-D visualization within the channel[1]. The work of [2,3] provided engineers the ability to examine structural defects and effectively pivot around features of interest within the channel. This highlighted the possibility of being able to deploy techniques which aim to facilitate the extraction of depth information directly from the RVI footage using a technique such as Structure-from-Motion (SfM) [4]. This process however is challenging due to the constrained and feature-poor environment observed within the AGR fuel channels.

This paper presents the preliminary results of the novel application of state-of-the-art SfM techniques to AGR inspection footage that aims to extract accurate 3-D depth information from captured 2-D RVI footage. The paper highlights the difficulties that must be overcome to facilitate the extraction of depth from historical RVI footage in Section 3 with emphasis on the challenges associated with feature correspondence searching. The underlying ambition of this work is to create a generalized 3-D reconstruction framework that can be extended across all domains of visual inspection and subsequently improve visualization techniques which in turn, improve human decision making.

[1] An example of the pivot video can be viewed at http://personal.strath.ac.uk/graeme.west/PBI.avi

2 STRUCTURE-FROM-MOTION

SfM is a well-established image processing technique which ascertains the 3-D geometry of a scene or object from a 2-D image sequence. It operates by identifying visual correspondences between multiple, overlapping viewpoints of the same object/scene then estimates the 3-D point locations using robust visual correspondences. Through triangulation of the location of multiple points in 3-D, a sparse representation of the scene (a.k.a. a 3-D point cloud) is formed and the intrinsic/extrinsic parameters of the camera from every viewpoint is calculated. This technique has been applied in a variety of fields such as robotics, augmented reality, architecture, archaeology and visual inspection [5]. SfM is performed in a multitude of steps in a quasi-parallel fashion which come together to form what is identified as a generic reconstruction pipeline. Generally, SfM approaches are composed of two key stages; correspondence searching and reconstruction. These are key steps in all SfM frameworks regardless of the reconstruction approach taken. An example of a typical SfM reconstruction framework is shown below in Fig.2.

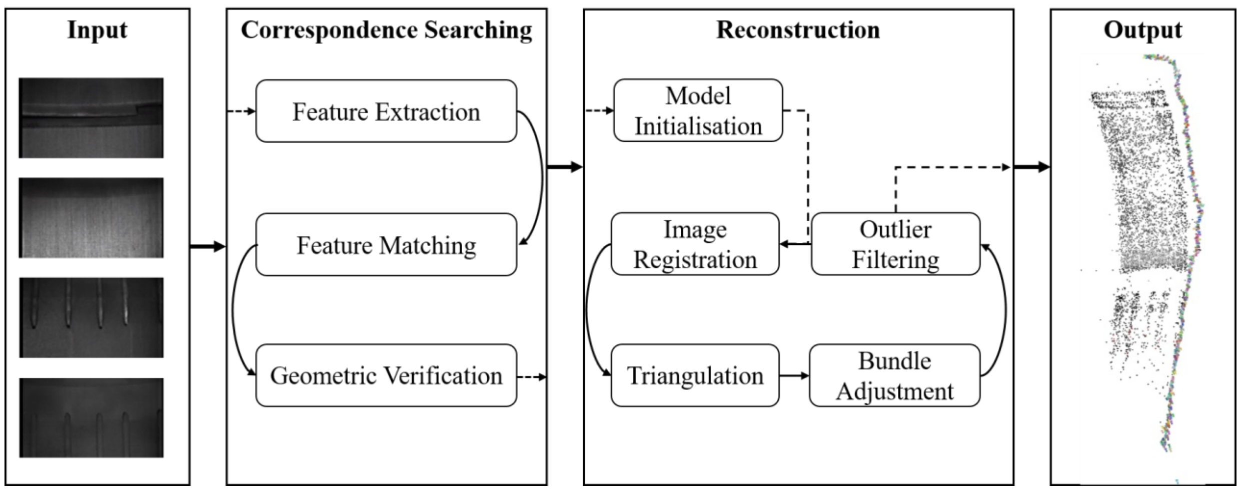


Figure 2. Flowchart showing a generic SfM pipeline for 3-D reconstruction and highlighting the key stages of each part of the pipeline: The input images, correspondence searching, reconstruction and the point cloud output [6].

2.1 Correspondence Searching

Correspondence searching determines overlapping viewpoints from the source input of images of the same object or scene from different vantage points. Direct approaches which operate on every pixel within all images are computationally complex therefore feature extraction is deployed using techniques such as the Scale Invariant Feature Transform (SIFT) [7] to produce an abstract set of discriminative points in the form of a feature descriptor. Each descriptor contains characteristics pertaining to the neighboring pixels around each point in the extracted set which are invariant to radiometric or geometric changes. The descriptors from each image are then matched using a metric to determine visual correspondence and the resulting geometric transformation between both viewpoints.

2.2 Reconstruction

The reconstruction stage of the SfM framework uses the matches determined by the correspondence searching process as an input. There are three different approaches to reconstruction; 1) Incremental which starts from a view pair which have a strong correspondence then iteratively builds the model; 2) Global which considers all correspondences during reconstruction; 3) Hierarchical which breaks the reconstruction problem into subsets which are then integrated into a singular, global model. All techniques simultaneously solve for a sparse 3-D representation of the observed object or scene as a set of

3-D points and the camera pose estimates. Afterwards, a refinement process known as Bundle Adjustment (BA) is applied to minimize the image distance between the projected 3-D point and the true projection in combination with a loss function to remove erroneous outliers for all pertinent images [6]. These refined sparse reconstructions can then be the basis of dense reconstructions using Multi-View Stereo (MVS), mesh reconstruction and texturing to generate photorealistic reconstructions [8]. For this paper, incremental reconstruction will be used as the basis of experimentation due to its prominence within SfM literature and the availability of open and closed source solutions for testing.

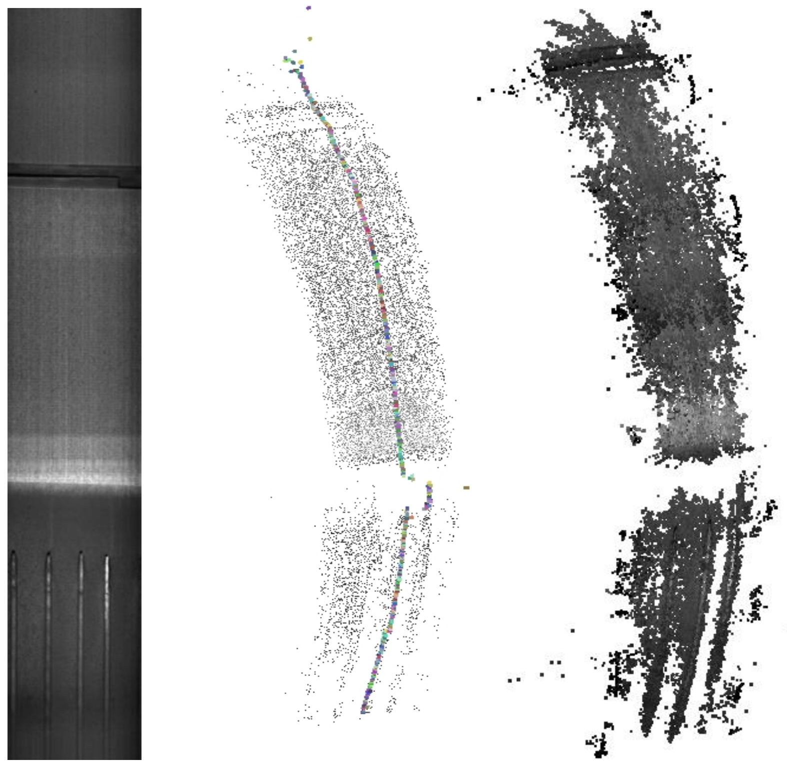
2.3 Application specific difficulties within the AGR

SfM has several theoretical pre-requisites to be able to operate and reconstruct with precision: The image capture process requires consistency in illumination; substantial areas of unique; discriminative texture in addition to a high degree of overlap from differing viewpoints. The inspection footage obtained within the AGR fuel channels however presents a number of challenges:

- Camera calibration parameters are not able to be captured which complicates the reconstruction process as calibration describes the mapping between world, camera and image co-ordinate systems.
- Non-uniform illumination within most recordings due to inspection engineers engaging with the inspection tool. This changes the underlying physical properties of the features and makes them radiometrically dissimilar and therefore extremely difficult to consistently match.
- The texture of the AGR fuel channel interior is repetitive (as the same graphite brick design is vertically stacked to form the fuel channel) which produces erroneous matches and confuses the triangulation stage of the reconstruction.
- Low usable video resolution due to video overlay ($\sim 720 \times 400$).

The current RVI image capture process is appropriate for visual inspection but it was not designed with utilizing 3-D visualization techniques such as SfM in mind. Furthermore, developing and deploying new visualization hardware is expensive due to the hazardous nature of the environment in addition to the added cost of certification for use in a nuclear application. This highlights the need for a bespoke framework to circumvent issues introduced by the capture process which makes it unsuitable for 3-D reconstruction. To highlight the application specific challenges which are introduced when applying SfM to inspection footage, experimentation was performed using a number of state-of-the-art SfM frameworks such as *COLMAP*, *Bundler*, and *VisualSFM* [6,9,10] that provide the entire reconstruction functionality visualized in Fig.2 as a minimum requirement.

The experimentation process consisted of manually splitting the inspection video into subsets of images which correspond to each inspected orientation whilst disregarding forward facing camera and crack following footage. A typical full channel scan at each set orientation ($0°, 60°, 120°, 180°, 240°, 300°$) consists of approximately 8000 images which was then fed into the tested reconstruction frameworks. However, all reconstruction frameworks failed to produce any meaningful reconstructions when using images consisting of an entire scan at a set orientation. To investigate further, the channel footage was partitioned correspondingly to each AGR fuel channel brick layer comprising of approximately 700 images per layer; an evaluated region of the fuel channel which has been stitched together using ASIST (Automated Software Image Stitching Tool) software [2,3,4] can be seen in Fig.3a). With this approach, the only reconstruction technique which produced a tangible result using RVI footage was *VisualSFM*. To obtain this result, a dense variation of the SIFT feature descriptor was used [7] for feature extraction. The feature matching process was performed in a constrained, sequential manner where an image could only be matched to 15 subsequent frames to suppress the illumination difference in features between the top and bottom of the images as seen in Fig.1b) and Fig.1c). The resulting sparse and dense reconstructions produced by *VisualSFM* can be seen in Fig.3b) and Fig.3c) respectively.



a) b) c)

Figure 3. a) Image of the evaluated region of AGR fuel channel b) Sparse 3-D point cloud and the camera pose estimates *P*. c) Dense 3-D reconstruction. Note the bending of the sparse and dense reconstructions caused by the differences in illumination.

Within the sparse reconstruction pictured in Fig.3b), the point cloud is densely populated with many of the channel features being visible such as the methane holes at the bottom of the layer and the brick interface between Layer 1 and 2 at the top. It is clear that the feature correspondence searching operates well only within a region which represents approximately 3% of the entire inspection video because of the abundance of discriminative features. However, this does not translate to other brick layers where it is repetitive and feature-poor and as a result, causes not only *VisualSFM* but all other tested frameworks at the feature matching to fail due to the lack of robust feature matches. At the reconstruction stage within the region shown by Fig.3a), the resulting reconstruction is erroneously bending backwards and the reason for this is due to the change in illumination as the camera enters the first brick layer of the channel (the bright horizontal line which can be seen in Fig.3a)). In turn, this introduces an error in the triangulation stage. Due to the incremental nature of the reconstruction, the error then cumulatively builds and results in the triangulated points to continue bending until the entire assessed region is processed. A similar observation by El Kahi et al. [11] brings more attention to this issue and highlights the importance of uniform illumination during the image capture process. Another issue through the deployment of incremental reconstruction is when the footage from multiple bricks are included, the reconstruction process will identify them as the same recurring brick instead of unique, individual layers which are stacked on top of each other and this highlights the incompatibility of incremental reconstruction. The preliminary reconstruction results illustrate that SfM is a plausible method for the AGR fuel channels however, to utilize SfM effectively, the SfM framework archetype illustrated in Fig.2 must be modified to improve correspondence searching and reconstruction within feature-poor and non-uniform environments.

3 CORRESPONDENCE SEARCHING

As highlighted with the experimentation using tested frameworks, one of the key issues of applying SfM to the RVI footage is the feature matching step within the correspondence searching stage. Due to the predefined challenges such as the inherently repetitive nature of the channel interior coupled with low resolution footage, many feature descriptors are erroneously matched together due to the uniformity of the detected features. Therefore, three different approaches which could be used to improve or assist the correspondence search using prior knowledge or through the combination of feature and direct-based approaches are discussed within this section.

3.1 Knowledge-based approach

A knowledge-based approach is inherently the simplest approach out of the three which incorporates expert knowledge about how the camera moves from frame to frame and forces the matching process to search within a defined region. With this approach, it would provide a simple solution whilst minimizing the computational complexity. During inspection, the camera undergoes vertical translation which equates to a change of approximately 5 pixels per frame of vertical displacement. Ideally, if the camera moved in a precise, exact motion between frames or if the trajectory of the camera was known through the use of an Inertial Measurement Unit (IMU) then the matching process would be simplified. However, during inspection, the camera does not move in a clean, vertical motion and often suffers from undesirable rotation/twisting as it is raised and lowered within the channel in addition to abrupt, transient motion caused by the wheels of the inspection tool entering brick interfaces or structural defects. Additionally, the camera may be stopped for further inspection of a defect which further disrupts the knowledge-based approach. By deploying a constrained matching process which only searches within a pre-defined window, the feature matching process will suffer from an accumulative drift over time and result in mismatching between uniform feature descriptors due to the inconsistency of rigid-body motion.

3.2 Optical Flow

Due to the uniformity of the majority of the channel interior and the irregular movement of the camera, feature-based approaches alone are unsuitable. To try and improve feature matching, the use of the SIFT feature descriptor [7] is hybridized with a motion estimation technique called optical flow to assist the feature extraction process [12]. Optical flow is an image processing technique which aims to characterise the dynamics of the perceived scene between contiguously captured image frames and gives an estimation of the motion of every pixel in a direct fashion between subsequent image frames in the form of a displacement field. Since the motion of the inspection tool is relatively inconsistent and therefore unknown between each frame, optical flow proves advantageous since it produces a velocity value for each pixel which allows it to capture not only basic rigid translation but also affine motion caused by the camera twisting. To perform optical flow, the Wedel et al. implementation of the Dual Total-Variation L^1 norm (Dual TV-L^1) algorithm [12] was selected particularly for its ability to handle changes in illumination between frames in addition to the use of additional filtering to suppress noise. This comes at the expense of computation with it taking approximately 2 seconds for every frame using an Intel i7 processor and 16GB of RAM but with an GPU implementation, this process is accelerated dramatically to near real-time performance.

After the Dual TV-L^1 optical flow algorithm is applied, the statistical mode, mean and median of the estimated displacement within each image is acquired then accumulated to demonstrate the camera motion within a 2-D space at each frame as shown in Fig.4. As expected, estimated motion of the camera in each frame does not follow a smooth, vertical translation but demonstrates clear, interspersed quivering during the inspection which in turn invalidates the usage of the pre-defined search window utilized within the knowledge-based approach. Using 2-D panoramic images produced using ASIST software [2,3,4], in the evaluated area where motion estimation was applied, there is a vertical shift of roughly 3700 pixels and a horizontal shift of approximately -50 pixels relative to the starting frame. Considering vertical displacement, the mode is very accurate with the cumulative vertical displacement of 3700 pixels exactly. Conversely, the median of the estimated cumulative horizontal displacement produces the most accurate result with a horizontal shift of -40 pixels with the modal horizontal displacement result being very inaccurate. With an approximate motion between frames known, the feature matching process could then be constrained using the vertical displacement value from the mode and the horizontal displacement using the median value.

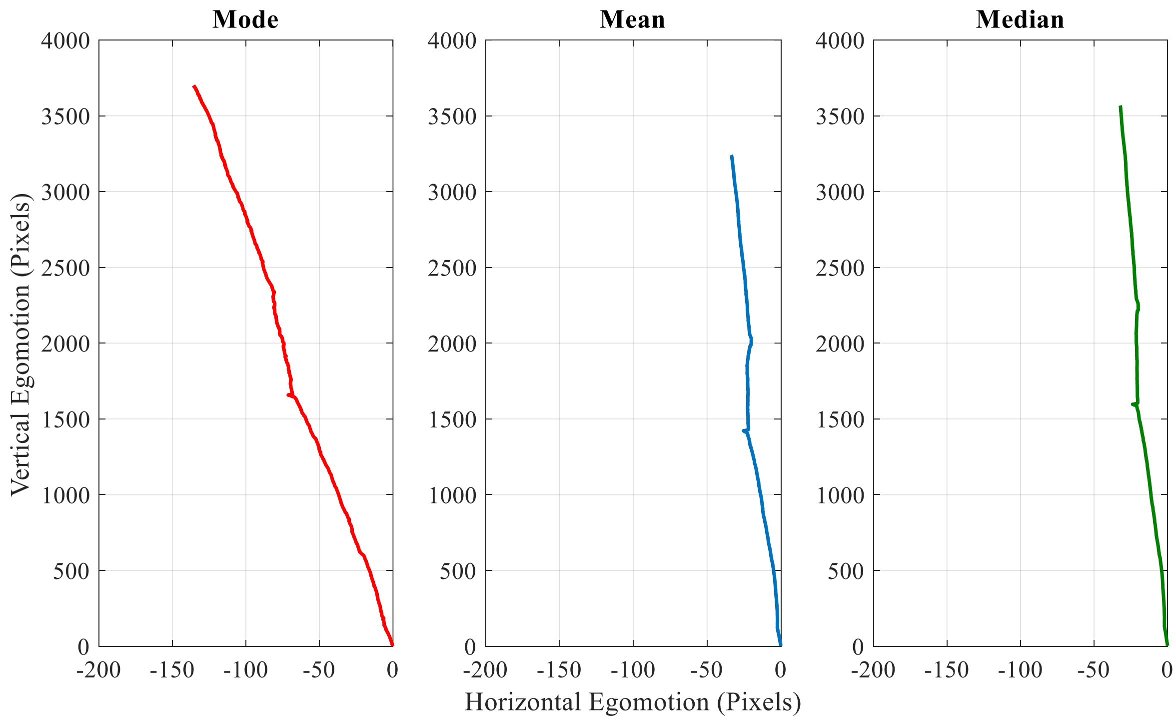


Figure 4. Estimated cumulative modal, mean and median displacement using the Dual TV-L^1 optical flow algorithm in each assessed frame.

3.3 Similarity

The use of similarity methods is another approach of aiding the feature matching process as it can be used to determine the similarity between two image regions. There are many forms of similarity methods such as Sum of Squared Differences (SSD), Sum of Absolute Differences (SAD), Census and Zero-mean Normalized Cross Correlation (ZNCC). SAD and ZNCC are regarded as the best methods when dealing with feature poor and feature rich environments respectively [8]. Conversely, the fundamental issue with SAD is the inability to cope with changes in illumination which make it inapplicable for feature matching using AGR inspection footage. Therefore, ZNCC is going to be focus of experimentation. The first stage of implementation is to apply the SIFT feature descriptor to the image to get a set of salient feature points. The second stage is used to apply filtering to the image and this is driven by the fact that the use of similarity measures is often hampered by noise. To filter out noise and suppress the changes in illumination, an 11×11 averaging filter is first applied to both compared images. Afterwards, ZNCC is applied using the following equation:

$$ZNCC(u,v) = \frac{\sum_{x,y}\left[I(x,y) - \bar{I}_{u,v}\right]\left[t(x-u,y-v) - \bar{t}\right]}{\left\{\sum_{x,y}\left[I(x,y) - \bar{I}_{u,v}\right]^2 \sum_{x,y}\left[t(x-u,y-v) - \bar{t}\right]^2\right\}^{0.5}} \in [-1,1] \tag{1}$$

where t is the template which will be searched in the image I and $\overline{I_{u,v}}$ is the mean of the region assessed by the template. The template is formed by truncating a region around each detected SIFT feature location. Two different forms of experimentation were carried out using ZNCC to determine the similarity. Primarily, the most common form of ZNCC which is using the template t to search the entirety of I_i, and secondly, a constrained radial search was implemented around the estimated feature location acquired by SIFT in the previous image with a pre-defined radius with a shift estimated using the Dual TV-L^1 algorithm. By doing this, we effectively combine the use of motion estimation and similarity to further strengthen the feature matching process. The results of this can be seen in Fig.5a) and Fig.5b) respectively.

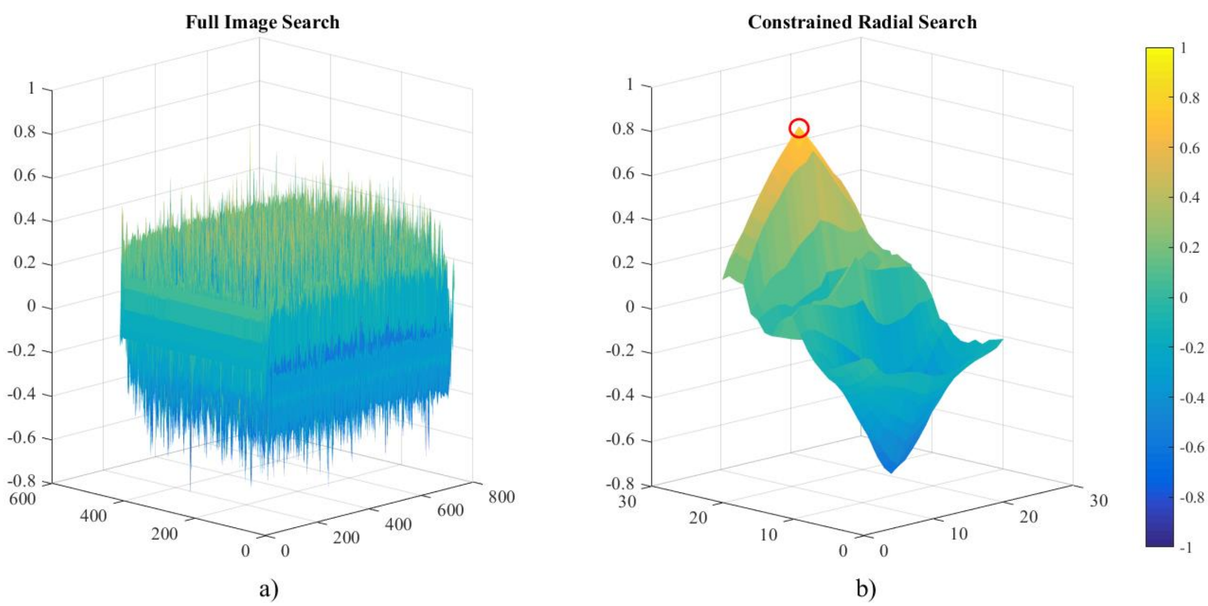


Figure 5. a) ZNCC performed using a full-image search in regards to the template b) Constrained radial ZNCC. Note the multiple peaks in a) in regards to b) where there is a singular maximum.

From Fig.5a), it is clear that due to the uniformity of the channel interior, there exists a large amount of erroneous maxima as a direct result of the repetitiveness of texture. Therefore, through a brute force matching procedure, it could identify any of the numerous maxima as a valid match which results in more matches that are essentially false positives. By applying a global search across an entire image using the similarity approach, it highlights the fundamental issue of generic feature matching in areas of repetitive texture regardless of the technique used. Therefore, by constraining the search within a radius defined using motion estimation, a maximum is more likely to be obtained using similarity methods such as ZNCC as seen in Fig.5b) and provides a more accurate means of matching since the search regions are localized. By deploying a constrained approach which combines multiple methods, the number of matches are vastly reduced as a result of localizing the feature matching process however the accuracy of the matches is substantially increased making the correspondence search process significantly more robust through reduction of false positive matches.

4 FUTURE WORK

Current open and closed source solutions cannot reconstruct fuel channels using full orientation scans of the fuel channel or be able to combine them without a priori knowledge due to the inherent similarity between brick layers. Therefore, work will be instigated into performing sparse and dense reconstructions using a bespoke approach. The feature extraction and matching mechanisms explained in this paper will be deployed within the reconstruction framework to improve the triangulation of 3-D point locations accurately and with precision from the 2-D correspondences. Other future work will be the integration of piecewise reconstruction framework which will allow fusion between sparse and dense reconstructions obtained using multiple orientation scans to form a 3-D cylindrical model of the AGR fuel channels for improved visualization.

5 CONCLUSIONS

This paper describes the preliminary work in developing a bespoke SfM pipeline to reconstruct AGR fuel channels in 3-D from RVI footage with particular emphasis on improving the robustness of feature correspondence matching within constrained, non-descript environments. Experimental work was presented and the different methods of assisting the feature correspondence search was proposed. The initial results are promising and improve the robustness of the feature matching process using a combination of indirect methods and direct methods at the expense of more computation. In the future, the reconstruction stage will be built from the ground up to account for the issues introduced by incremental reconstruction in a repetitive environment as well as methods to suppress non-uniform illumination within the RVI footage.

6 ACKNOWLEDGMENTS

This work was funded by EDF Energy. The views presented in this work do not necessarily represent the views of EDF Energy.

7 REFERENCES

1. C.J.Wallace, G.M.West, G.J.Jahn, S.D.J.McArthur, D.Towle, G.Buckley, "Control Rod Monitoring of Advanced Gas-Cooled Reactors", *Seventh American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies NPIC&HMIT 2010*, Las Vegas, Nevada, November 7-11th, 2010 (2010)
2. G.M.West, P. Murray, S.Marshall, S.McArthur, "Improved visual inspection of advanced gas-cooled reactor fuel channels", *Nuclear Engineering and Design*, **Volume 6** (2015)
3. P. Murray, G.M.West, S.Marshall, S.McArthur, "Automated in-core image generation from video to aid visual inspection of nuclear power plant cores", *International Journal of Prognostics and Health Management*, **Volume 300**, pp. 57-66 (2016)
4. P. Murray, G.M.West, K.Law, S.Buckley-Mellor, G.Cocks, C.Lynch, "Automated video processing and image analysis software to support visual inspection of AGR cores", *5th EDF Energy Generation Ltd Nuclear Graphite Conference*, Southampton, UK, May 9-12th 2016 (2016)
5. P. Hansen, H.Alismail, P.Rander, B.Browning, "Visual mapping for natural gas pipe inspection", *International Journal of Robotics Research*, **Volume 34**, pp. 532-558 (2015)
6. J.L. Schönberger, J-M.Frahm, "Structure-from-Motion Revisited", 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, June 26th – July 1st 2016, pp. 4104-4113 (2016)

7. D.G.Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision*, **Volume 60**, pp. 91-110 (2004)

8. Y.Furukawa, C.Hernández, "Multi-View Stereo: A Tutorial", *Foundations and Trends® in Computer Graphics and Vision*, **Volume 9**, pp.1-148 (2015)

9. N.Snavely, S.M.Seitz, R.Szeliski, "Modeling the World from Internet Photo Collections", International Journal of Computer Vision (IJCV), **Volume 80**, Issue 2, pp. 189-210 (2008)

10. C.Wu, "Towards Linear-Time Incremental Structure from Motion", *Proceedings of the 2013 International Conference on 3D Vision (3DV)*, Seattle, USA. 2013, pp. 127-134 (2013)

11. S. E. Kahi, D. Asmar, A. Fakih, J. Nieto and E. Nebot, "A vison-based system for mapping the inside of a pipe," 2011 IEEE International Conference on Robotics and Biomimetics, Karon Beach, Phuket, 2011, pp. 2605-2611 (2011)

12. A.Wedel, T.Pock, C.Zach, H.Bischof, D.Cremers, "An Improved Algorithm for TV-L^1 Optical Flow", *Statistical and Geometrical Approaches to Visual Motion Analysis*, Dagstuhl Castle, Germany, July 13-18th 2008 pp. 23-45 (2009)