A NEW PROBE CONCEPT FOR INTERNAL PIPEWORK INSPECTION

Rahul Summan, Neha Mathur, Gordon Dobie, Graeme West, Stephen Marshall, Charles Macleod, Carmelo Mineo, Gareth Pierce

University of Strathclyde,
Dept. of Electronic and Electrical Engineering
204 George Street,
Glasgow

rahul.summan@strath.ac.uk, neha.mathur@strath.ac.uk, gordon.dobie@strath.ac.uk, graeme.west@strath.ac.uk, stephen.marshall@strath.ac.uk, charles.macleod@strath.ac.uk carmelo.mineo@strath.ac.uk,s.g.pierce@strath.ac.uk

ABSTRACT

The interior visual inspection of nuclear pipework is a critical inspection activity required to ensure the continued safe, reliable operation of plant and thus avoid costly outages. Typically, the video output from a manually deployed probe is viewed by an operator online with the task of identifying and estimating the location of surface defects such as cracks, corrosion and pitting. However, it is very difficult to estimate the nature and spatial extent of defects from the often disorientating small field of view video of a relatively large structure. This work describes a new visual inspection system incorporating a wide field of view machine vision camera and additional sensors designed for inspecting 3 - 6 inch diameter pipes. The output of the system is a photorealistic model of the internal surface of the pipework. The generation of this model relies upon a core component of the system in the form of image feature extraction which estimates camera location. This paper considers the accuracy of this estimation as a function of the number and configuration of the extracted image features.

Key Words: Visual Inspection, Laser Profiler, Structure from Motion

1 INTRODUCTION

The internal visual inspection of pipework in the nuclear industry is a critical inspection activity that is carried out periodically to ensure the continued safe operation of plant. The technology used for such inspections are typically rudimentary and provide a live camera view of the internal surface which is monitored by an expert operator. If a surface defect is identified, the operator can attempt to manipulate the position and orientation of the probe to improve visibility, however, this can be challenging especially with distance. In this work, a new probe concept is proposed. Using bespoke hardware and an image feature based Structure from Motion (SFM) algorithm, the system is capable of generating a photorealistic, geometrically accurate surface model. This system is designed for deployment in empty pipes and although the application is nuclear focused, it is suitable for any pipework e.g. pressure vessels and heat exchanger tubing with enough surface texture to allow the probe localization system to operate.

The accuracy of the reconstructed surface model is a function of the accuracy of the estimated probe motion which is computed from 2D image features. In the target application of pipe mapping, such features tend to be low in number and non-uniformly distributed due to the underlying uniformity of the pipe's internal surface. Through simulation, this paper presents a systematic investigation of the error caused by a non-uniform 2D spatial distribution of features on the resultant estimate of camera motion. Accuracy is considered as a function of the number of correspondences and feature coverage pattern around a circular

image acquired by a fisheye lens. This paper initially provides an overview of the individual sub-systems of the probe and proceeds to detail the accuracy evaluation.

2 SYSTEM OVERVIEW

The probe consists of a machine vision (2MP) camera coupled with a fisheye lens. Thus in a single image, a 360° cross sectional view of the surface is captured where the lateral extent is a function of pipe diameter. At the cost of increased signal processing complexity, this configuration leads to quicker inspections [1], [2], [3], [4] in comparison to approaches which use mechanically swept narrow field of view cameras [5]. The current hardware prototype designed for 3 inch diameter pipes is shown in Figure 1 where each subsystem is highlighted. The system is driven in two distinct modes in alternation: colour image capture of the surface and cross sectional shape measurement capture. In the first mode, the internal volume of the pipe is illuminated with high power LEDs to acquire a colour image of the surface – these are subsequently used for SFM processing. Secondly, the LEDs are switched off and a laser ring profiler is activated which projects a red laser ring onto the pipe's surface and into the field of view of the camera. Through knowledge of the rigid body transform between the camera and profiler, triangulation is used to estimate the geometry of the illuminated cross section. Such data can be used to measure the dimensions of local surface defects such as cracks and pits and also global shape in the form of, for example, pipe ovality. An overlap of at least 50% between successive photographic image enables the cameras motion to be calculated which is not possible with laser images since only measurements along a line are acquired. In summary, the photographic images are used to calculate the probes motion while the

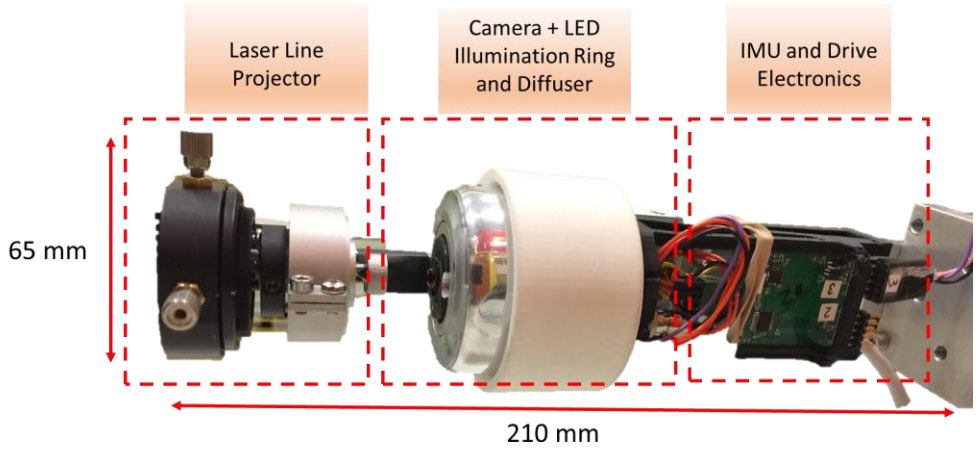


Figure 1 - Development Probe with sub-systems highlighted

laser is used to measure defects. Each image also has an associated timestamp and orientation measurement from an onboard inertial measurement unit (IMU) – this data can be used to assist the image processing in areas of low surface texture.

2.1 Laser Image Processing

A laser profiler image is shown in the image of Figure 2(a). Note that in the lower half of the image, part of the laser line is not visible due to a blind spot in the cameras field of view caused by a mechanical support arm for the profiler. In general this line will not be a circle, rather it will be a curved line which may contain discontinuities due to surface defects. In order to extract such a curve from the image, a radial line search in the red component is performed to determine the set of pixels corresponding to the laser. The

geometry of the laser and camera configuration is shown in Figure 2(b) where the pipe radius is denoted by, r, and the measured point on the surface is, p_w. This arrangement contains two geometric parameters, B and θ, which are needed for triangulation – these can be obtained through calibration prior to operation.

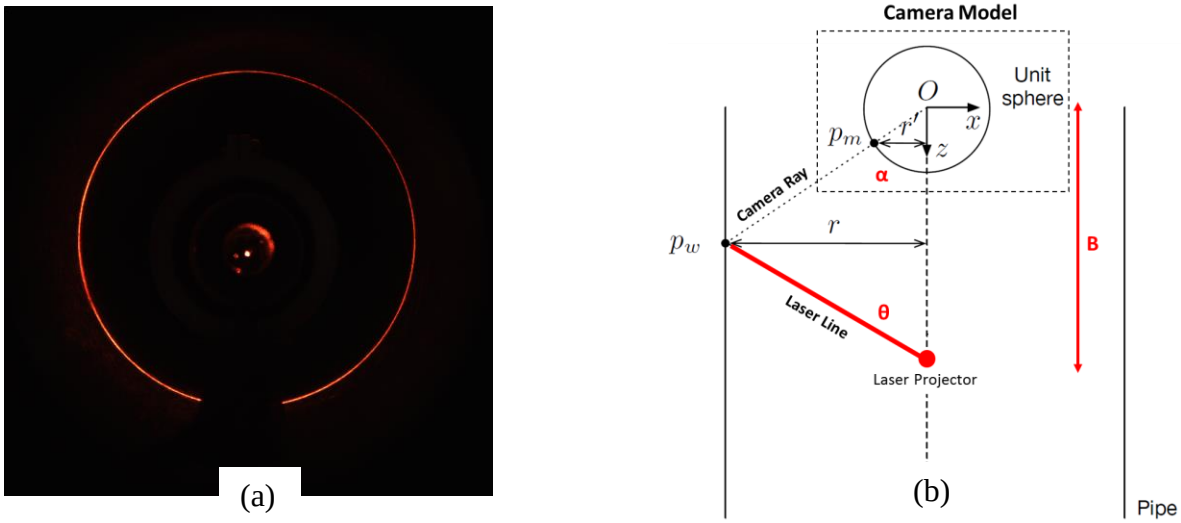


Figure 2 - Laser Profiler Image (a) Laser Image (b) Camera Laser Geometry

The laser is emitted orthogonally to the optical axis of the camera corresponding to an angle θ = 180°. Given knowledge of the baseline, B, and the direction vector corresponding to each laser pixel triangulation may be employed to estimate the 3D position of each these pixels. The vector corresponding to a pixel is estimated from the camera model - unit sphere in Figure 2(b) - described by Scaramuzza et al [6].

2.2 RGB Image Processing

Image feature extraction forms a core component of the system's operation. Due to the computational expense of operating upon the image directly, it is firstly transformed into a set of salient points which essentially represent gradient information in the intensity profile of the image. A feature consists of such a salient point and also a descriptor which encodes statistics computed from a window of pixels surrounding

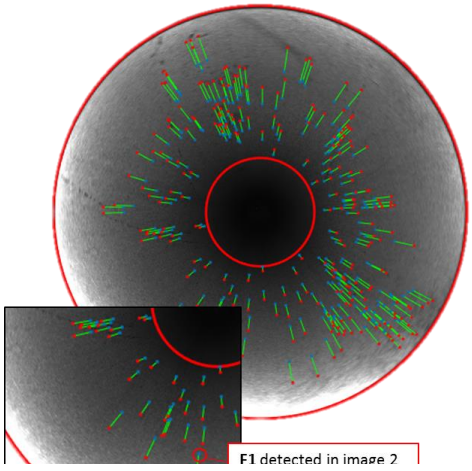


Figure 3 – Feature Extraction on stainless steel pipe. The features lie upon radial lines corresponding to linear motion

this point. The descriptor is used to identify the same feature in other images – through consideration of the 2D motion of features, the 3D motion of the probe may be inferred [7]. The feature extraction and matching process forms the fundamental input to the SFM processing pipeline. The widely used Scale Invariant Feature Transform (SIFT) [8] was selected for performing feature extraction in the visual front-end of the system. An example of feature extraction resulting from forward linear motion is shown in Figure 3, the red dots correspond to image extracted in the first image while the blue dots show where the features move to in the next image. Assuming that the camera model is known, (normalized) features $\tilde{x}_{i-1}$ and $\tilde{x}_i$, matched between images $i-1$ and i must satisfy the following constraint [7]:

$$\tilde{x}_{i-1} E \tilde{x}_i = 0 \tag{1}$$

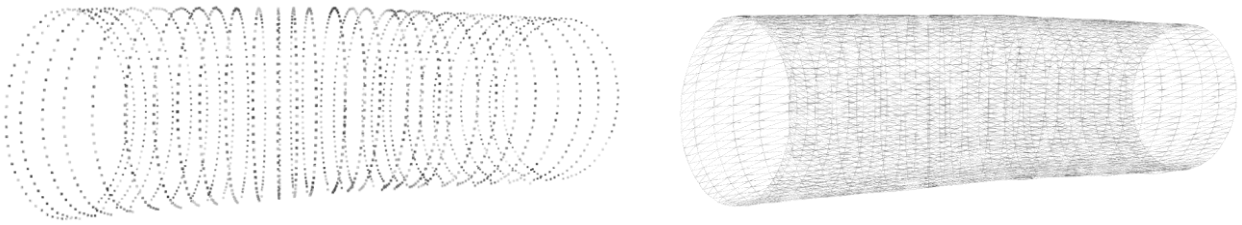
where, E, is the Essential matrix which encodes the rigid body transform [9]:

$${}^{i-1}T_i = \begin{bmatrix} R & \boldsymbol{t} \\ \mathbf{0} & 1 \end{bmatrix} \tag{2}$$

between the two images where $\boldsymbol{t} = [x, y, z]^T$ is 3D translation and $R \in SO(3)$ is 3D rotation. A minimum of 8 features matched are required to estimate, E, using the normalized 8 point algorithm [7]. These relative transforms are multiplied together to construct the absolute probe trajectory.

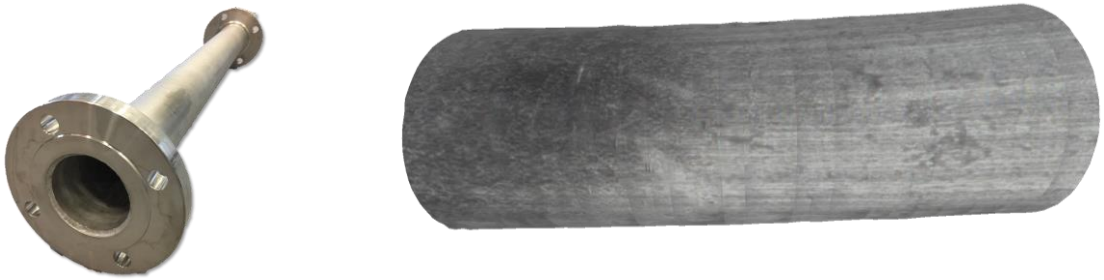
2.3 System Output

An example output of the system is shown in Figure 4 and Figure 5. The system firstly generates a point cloud, Figure 4(a), from the laser where each laser ring is placed in 3D space according to the camera positions/orientations calculated from the SFM processing. A meshing algorithm is subsequently applied to transform this point cloud into a triangular mesh as shown Figure 4(b). Image data is then mapped back onto the mesh through a texturing procedure to form the photorealistic model. This consists of projecting the vertices corresponding to each mesh triangle into the images to generate a set of candidate image patches. The patch which is most perpendicular with the view axis of the camera is chosen for mapping. The final output corresponding to a linear segment of 3 inch 304L stainless steel pipe in Figure 5(a), is displayed in Figure 5(b). The texture from the inside of the pipe is visible on the outer surface to provide a single global view of the data and thus ease of visualization. Ultimately, it is envisaged that the 3D model will be divided into linear segments where each segment is transformed into a planar surface for analysis.



(a) (b)

Figure 4 - (a) Laser point cloud (b) Meshed point cloud



(a) (b)

Figure 5 - 3D Model (a) 3 inch stainless steel pipe (b) Photorealistic model

3 FEATURE EXTRACTION PERFORMANCE

The feature extraction process is affected by both the camera's optics and the material under inspection. The matching process relies upon the patch of pixels corresponding to the image feature descriptor having a similar appearance in different images. However, the fisheye lens introduces a non-linear warp of these image patches as they move across the lens – this can be observed in Figure 6 with changing spatial step between successive images. The effect of the fisheye lens upon matching performance was evaluated for different inter-image spacing using, 5 mm, 10 mm and 15 mm steps. It is immediately clear that with greater step size the average number of feature matches reduced markedly.

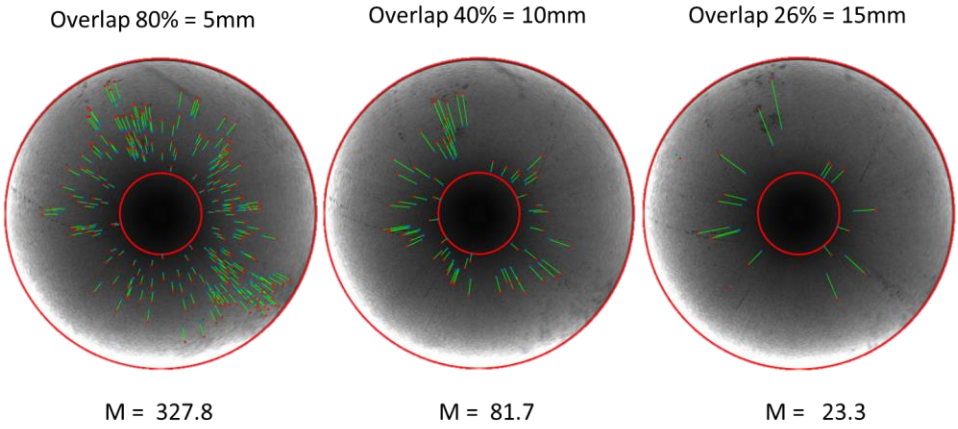


Figure 6 – Feature extraction as a function of image spatial separation where M is the mean number of matches

Since features are ultimately derived from the available texture in the images, the surface finish of the underlying material influences the performance of feature extraction [10]. In cases where the material has a homogenous appearance and thus a lack of gradient information, feature extraction performs poorly, this is clearly shown in Figure 7 where the matches are mostly clustered in one quadrant of the image. This clustering of features has the potential introduce error in the estimated pose, ${}^{i-1}T_i$, computed in the SFM

pipeline. Many systems that have been reported in the literature do not have such a marked issue because they are deployed into feature rich environments. However, it is recognized as a potential problem and often alleviated using a technique known as bucketing [11]. This technique attempts to uniformly sample from the available feature matches by superimposing a grid over the image then sampling from each grid cell. This technique can be extended to circular images through the use of sectors. However, in the current application of pipe mapping, an abundance of features may be unavailable thus forcing the use of a biased configuration of feature matches. The purpose of the investigation herein is to consider the effect on the estimated camera pose when clustered features are used.

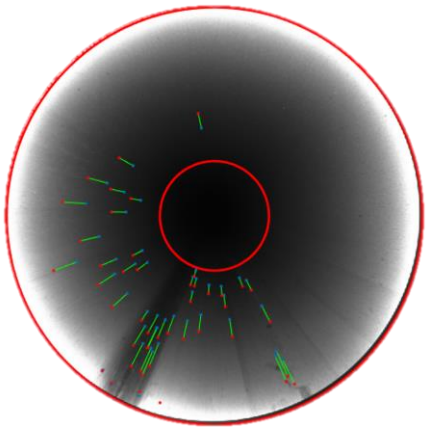


Figure 7 - Feature extraction on uniform surface leading to clustered features

A simulation environment was developed to systematically consider the effect of the spatial distribution and number of features upon the resultant pose estimate. Given camera poses, the images corresponding to 3D points belonging to a cylindrical surface could be generated through this environment. Two feature patterns, as shown in Figure 8, were chosen to capture the clustered nature of features. The pattern in Figure 8(b) captures the single global cluster configuration demonstrated in Figure 7 while Figure 8(a) captures local clusters of points which are spatially separated by an equal angle.

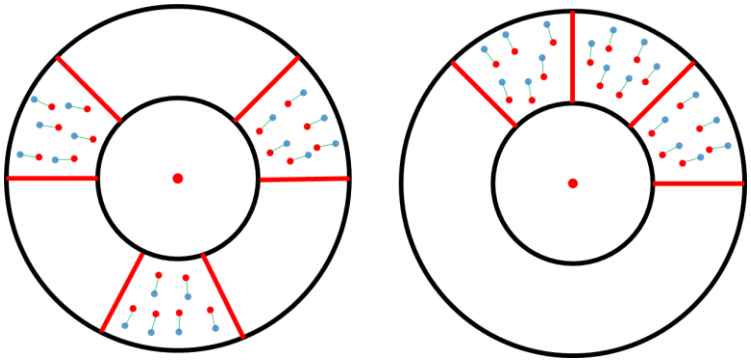


Figure 8 – Image Feature Patterns (a) Local cluster (b) Global cluster

Pose error was evaluated by successively adding sectors to an initially empty image circle according to the above patterns. Starting from a minimum of 8, 8 matches were successively added up to a maximum of 96

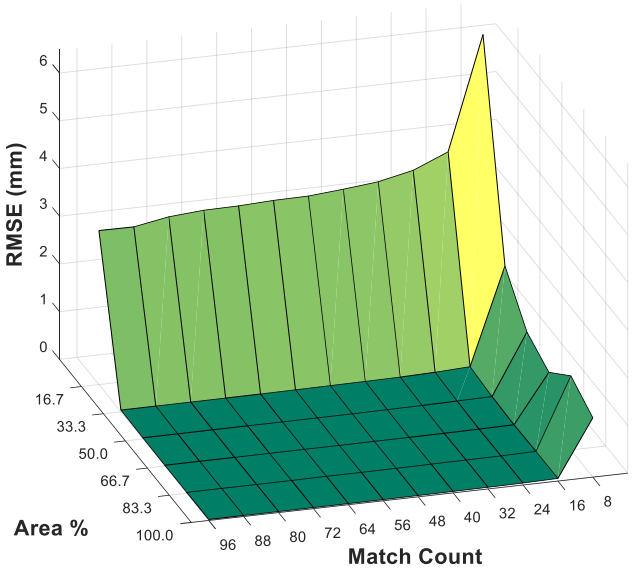


Figure 9 - Local cluster

matches, wherein each iteration they were equally (as near as possible) distributed amongst all available sectors. The relative pose between two images captured inside the cylinder of known radius was then computed at each iteration. This data allowed pose error to be considered as a function of both percentage area of the image occupied and the number of feature matches. Pose error expressed in terms of translation is shown in Figure 9 and Figure 10 for local and global clusters respectively. Each data point is the Root Mean Square Error (RMSE) of 500 trials.

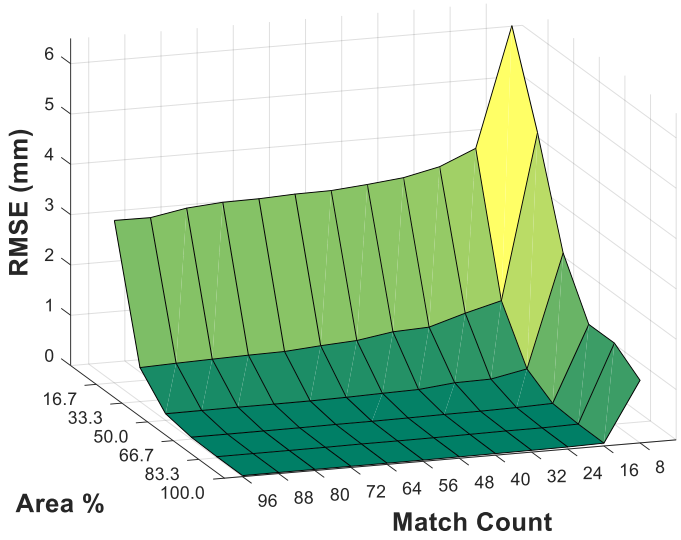


Figure 10 - Global clusters

It is readily apparent that the local cluster case converges more rapidly than the global case. At around 33% area the local clusters converge to the final error value while for the global case this happens at approximately 66% area. In addition to this, beyond 16 feature matches, there is no substantial improvement in the final RMSE error. When the number of features is greater than 16 matches, the area occupied plays a more important role than the number of matches. This simulation has shown in a well-founded manner that matches which are spread across an image yield smaller error in comparison to the same number of clustered matches.

4 CONCLUSIONS

In conclusion, this paper has presented an overview of a unique video probe system for conducting the internal visual inspection of pipelines in the nuclear industry. The system makes use of a feature based approach for estimating the trajectory of the probe. High accuracy measurements of cross sectional geometry are then projected into this path to form a skeleton structure representing surface geometry. The final output of the system is a photorealistic model which is created by meshing the laser point cloud and back-projecting image data onto the mesh. This form of output is far easier for the end user to interpret than the raw video data. In the target application of pipe mapping, features can appear clustered due to a uniform surface finish of the material from which the pipe is constructed. This introduces error into the estimated camera path which directly propagates into the 3D model at the output of the system. An investigation into the error caused by global and local clusters of feature matches was conducted. It was found that the local case lead to faster convergence to the minimum error and also that beyond 16 feature matches there was not substantial improvement in error.

5 ACKNOWLEDGMENTS

This work was funded by Innovate UK in the “Mosaicing for Automated Pipe Scanning” project. The authors would like to thank the project partners: National Nuclear Laboratory, Wideblue, Inspectahire and Sellafield Ltd.

6 REFERENCES

1. J. Kannala, S. S. Brandt, and J. Heikkila, “Measuring and Modelling Sewer Pipes from Video,” *Machine Vision and Applications*, **vol. 19**, no. 2, pp. 73–83, 2008.
2. P. Hansen, H. Alismail, P. Rander, and B. Browning, “Pipe Mapping with Monocular Fisheye Imagery,” in *Intelligent Robots and Systems* (IROS), 2013 IEEE/RSJ International Conference on, pp. 5180–5185, IEEE, 2013.
3. K. Matsui, Y. Atsushi, and K. Toru, "3-D Shape Measurement of Pipe by Range Finder Constructed with Omni-directional Laser and Omni-directional Camera." *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010.
4. H. Song, G. Kunshan, Q. Di, W. Huapu and Y. Jing. "Design of In-Pipe Robot Based on Inertial Positioning and Visual Detection." *Advances in Mechanical Engineering* 8.9 2016
5. P. Murray, G. West, S. Marshall, S. McArthur, "Automated image stitching for enhanced visual inspections of nuclear power stations." *10th International Conference on Condition Monitoring and Machine Failure Prevention Technologies*. 2013.

6. D. Scaramuzza, A. Martinelli, and R. Siegwart. "A Flexible Technique for Accurate Omnidirectional Camera Calibration and Structure from Motion." *Computer Vision Systems, 2006 ICVS'06. IEEE International Conference on*. IEEE, 2006.
7. R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision. Cambridge university press, 2003.
8. M. Brown, and D. G. Lowe. "Automatic Panoramic Image Stitching Using Invariant Features." *International journal of computer vision,* pp. 59-73, 2007.
9. P. Corke, Robotics, Vision and Control: Fundamental algorithms in MATLAB. Springer, **Vol. 73**. 2011.
10. G. Dobie, R. Summan, C. MacLeod, and S. G. Pierce, "Visual Odometry and Image Mosaicing for NDE," *NDT & E International*, **vol. 57**, pp. 17–25, 2013.
11. Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong, "A Robust Technique for Matching two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry," *Artificial intelligence*, **vol. 78,** no. 1,pp. 87–119, 1995.