A Low-Cost Robotic System for the Efficient Visual Inspection of Tunnels

S.A.I. Stent[a], C. Girerd[b], P.J.G. Long[a] and R. Cipolla[a,c]

[a] Department of Engineering, University of Cambridge, U.K.
[b] Department of Mechatronics, Institut Français de Mécanique Avancée, France
[c] Toshiba Research Europe Ltd., Cambridge, U.K.
Corresponding author e-mail: sistent@cantab.net

ABSTRACT

We introduce a low-cost robotic system designed to enable the safe, objective and efficient visual inspection of tunnels. The system captures high resolution images and processes them to produce maps of tunnel linings that are suitable for detailed inspection.

It is unique in that the total cost of hardware is an order of magnitude less than most existing systems while producing an equivalent or higher quality of output. The device makes use of consumer-grade digital cameras and high-power LEDs in a rotating rig, carried by a lightweight aluminium frame which is designed to reduce vibrations during data capture. It is portable and installable by hand and has a modular design, making it possible to adapt to different types of carriage units, tunnels and sensors.

Within the paper, we share insight into features of the device's design, including lessons learned from trials of earlier prototypes and comparisons with alternative systems. Using field data gathered from a 2km utility tunnel, we demonstrate the use of our system as a means of visualising tunnel conditions through image mosaicing, cataloguing tunnel segments using barcode detection and improving the objectivity of visual condition surveys over time by the detection of sub-mm crack growth.

We believe that our device is the first to provide comprehensive survey-quality data at such a low cost, making it very attractive as a tool for the improved visual monitoring of tunnels.

Keywords

Automation, Robotics, Tunnel Inspection, Tunnel Monitoring, Computer Vision, Image Processing

1 Introduction

The operation and maintenance of ageing civil infrastructure such at tunnels is one of the greatest challenges facing society today [1, 2]. Visual inspection, which allows the measurement and assessment of structural health over large spatial areas, is a key component. Timely and thorough inspection can result in the early identification and resolution of any problems.

Traditionally, the visual inspection of large-scale assets has been an expensive, labour-intensive process. While inspection methods across the industry are gradually being modernised with the help of automation and machine vision, many asset owners still rely on manual inspection by teams of trained inspectors. For tunnels, this is carried out by time-consuming manual walkthroughs, which can pose a health and safety risk as the environment is frequently dusty, dark and hazardous. Inspection records are often subjective, incomplete and prone to human error.

The benefits of automating the inspection process are well appreciated, and companies offering such solutions are now commonplace, particularly for road and rail. In tunnel inspection, numerous systems exist and many are in active research and development [e.g. 3-14].

In this work, we describe our efforts to develop a system to allow the autonomous capture of images suitable for the detailed visual inspection of tunnels. There are several features of our system which differentiate it from existing technology: the low material cost which is a fraction of existing systems; the high portability, modularity and capture autonomy; and the high quality of the data output, which allows both broad inspection using image mosaics and fine inspection of sub-0.3mm cracks. The main trade-off of our system is that its capture speed is relatively slow, limiting any very large scale application of the system to relatively low-traffic tunnels.

The article is organised as follows. Section 2 introduces the background of the project, certain case-specific design issues and a brief review of existing technology. Section 3 describes the key features of our robot design. Section 4 details field testing carried out in high-voltage cable tunnels in London, UK. Section 5 uses the field data gathered to illustrate some useful machine outputs for visual inspection. Section 6 concludes with a final evaluation of the present system and details about the future development of the project.

2 Background

The goal of this project was to develop an automated imaging system for a network of utility tunnels carrying high-voltage cables. Electrical power tunnels are expected to become increasingly common in cities in the future, as burying power cables beneath roads and open land becomes prohibitively expensive. Previous efforts to develop such a system highlighted a number of important design issues detailed in Section 2.1. Existing technology in this area is reviewed for comparison in Section 2.2.

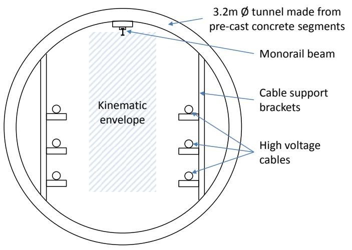

Figure 1. Typical cross-section of the completed tunnel showing monorail, cables and the allowable kinematic envelope in which the robot can operate.

2.1 Case-Specific Design Issues

- **Monorail systems preferred.** Monorail was preferred over track or ground based systems due to the lower installation and maintenance cost (vs. track), the smaller kinematic envelope and the reduced risk of derailment, collisions and trip hazards. A typical tunnel cross-section and ideal kinematic envelope is shown in Figure 1.
- **Images affected by monorail smoothness**. The quality of the ride on the monorail, which consists of I-beam sections joined together (by electrically insulated connections, welding or pins), is poor. Previous efforts showed that continuous video data was inadequate for inspection purposes due to excessive monorail vibrations.
- **Size and maintenance**. Due to the difficulty of access, a lightweight (portable and installable by hand) and low maintenance system was essential.
- **Setup speed favoured over capture speed.** The system should be fast to set up and extract, keeping human tunnel time to minimum, since the tunnel environment is considered hazardous. In contrast, the data capture speed is of secondary importance provided it is unmanned.
- **Sensor type.** The initial interest was to gather high quality RGB images suitable for image-based reconstruction and the production of high-resolution image mosaics suitable for the detection of sub-mm features. The design should be able to integrate other sensors in the future.

2.2 Existing Technology

The spectrum of existing robotic tunnel inspection systems is very broad, and an up-to-date survey can be found in [1]. As described in [15], tunnel inspections vary in purpose, from routine visual inspections to more detailed, specific inspections. Routine inspection typically utilises purely visible and infra-red vision and/or geometry-based inspection systems [3-12] and operates as a kind of early warning system. However, many systems are developed for more detailed diagnosis, using additional sensors such as ultrasound, thermography, ground penetrating radar and mechanised hammer testing [e.g. 13-14].

Among the vision-based systems, the variety in approach is driven largely by the constraint of time available for asset inspection, which varies from case to case. Road and rail tunnels in particular have high traffic flows and therefore favour fast inspection systems in order to reduce disruption. Since road and rail can support fast locomotion, inspection systems can be mounted on manned vehicles and make use of sensors which are suitable for rapid capture. Examples include [3], which uses an array of narrow field-of-view laser-camera units mounted on a truck or rail, and [4-6], which combine arrays of line-scanning or high-frequency camera arrays and powerful lighting. These systems have fast capture times, but achieve them at the expense of higher equipment costs, longer set-up times and greater bulk to support the larger numbers of sensors and increased computing and illumination requirements.

In contrast, tunnels not intended for transport such as ours generally have little or no traffic and therefore less constrictive inspection windows. However, they are also more likely to be considered hazardous to human health, since they may not be designed to carry humans, and may be more variable in size. Hence lighter, more flexible systems such as [7, 8, 11, 12], and unmanned systems such as [9, 10] are more suitable.

The system of [9] uses a line-scanning camera for crack detection and relies on the robot maintaining a constant distance from the target to gather useful data. While line-scanning cameras can acquire data quickly, due to their limited field of view, image distortion can be

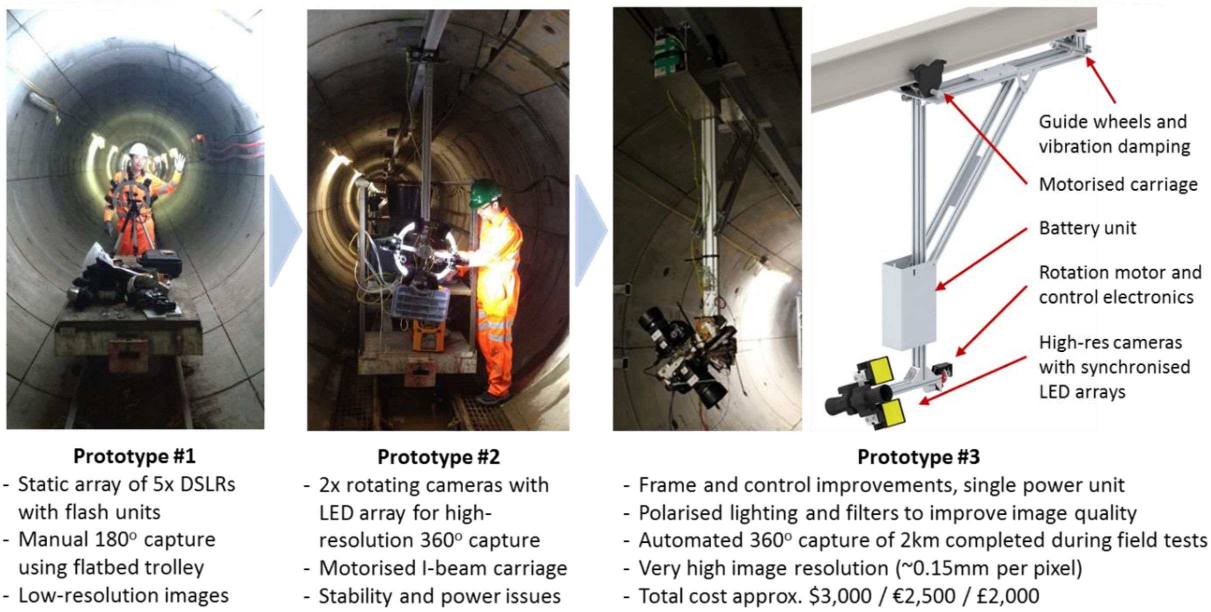

Figure 2. The evolution of our design over the life of the project. The initial proof of concept was gradually refined to become more automated and solve practical issues related to operating in the tunnel environment while maintaining a very low overall cost.

quite severe and hence the robot motion relative to the target must be carefully controlled. The system of [10] uses a fisheye camera with structured light both for inspection and to determine the robot's location via visual odometry. Fisheye systems are suitable for very low diameter tunnels or pipelines, but resolution and lighting become problematic at larger scales. The potential of a simple DSLR for accurate image mosaicing is shown in [11], but the technique is highly manual, requiring specified reference points or laser markers and assuming a precise geometry. A similar, more automated system is presented in [12] using a rig of DSLR cameras, but the capture process is manual.

Our contribution is a system which is several orders of magnitude cheaper than [3-8], more scalable than [9-12], and can be quickly adapted to provide images suitable for the detection of 0.3mm cracks in a range of tunnel scenarios.

3 Robot Design

3.1 Evolution of the Design

During testing, we evolved from a static array of cameras, similar to [3-6,12], to a rotating system. We found that the static array, despite allowing faster capture speeds, made the system much less flexible. For example, varying the resolution of capture to account for different tunnel dimensions required changing both the lens and orientation of each camera in the array as well as re-calibration of the new setup.

The rotating design provided much greater flexibility. Arbitrary resolution could be achieved by switching to more powerful lenses, and the number and orientation of images per revolution could be quickly adjusted programmatically to compensate. Reducing the number of cameras also reduced the number of lenses requiring calibration, cut down the weight and cost and resulted in more consistent lighting conditions in the images.

3.2 Features of Current Prototype

3.2.1 Portability and modularity

Figure 3 shows a schematic diagram which highlights the modular design of the current system. The sensor unit, control unit and motorised rotation unit combine into a single detachable and lightweight "capture unit", which is pictured in Figure 4. Modularity means that the capture unit can easily be switched from a monorail carriage to a ground or rail-based carriage if required for other tunnels. It also allows for easy installation: the monorail carriage can be permanently stationed at the bottom of tunnel shafts, while the more delicate capture unit can be stored off-site, away from the corrosive tunnel environment, and easily attached/detached by hand when required.

3.2.2 Image sensor and data capture strategy

Our choice of consumer DSLR cameras is unusual for this type of application, but important to maintain the low overall cost. We used 12MP Canon EOS 1100D, which are more than an order of magnitude cheaper than industrial cameras of comparable image quality. In addition, they have many convenient hardware features such as a jack connector and hot shoe connector to allow precise synchronisation with lighting, while being robustly designed for consumer use and compatible with a range of low-cost, high-quality lenses. The model is also supported by *gphoto2*, an open-source command-line software and library which we used for remote-controlled configuration and capture. The main constraint imposed by this choice of camera is the limited USB data transfer speed, which makes it impractical for applications requiring faster capture rates.

The dual camera setup allows for a number of modes of operation. A standard scan can capture a little over 180 degrees of data with each camera and the two image streams can be registered to cover 360 degrees. For more detailed scans, each camera can individually capture a 360 degree ring. In this setup, one camera was focused on a narrow field of view to acquire highly detailed images, sufficient for picking up hair-line cracks such as in Figure 5. The second camera had a wider field of view, with overlapping images to facilitate image registration, reconstruction and mosaicing. The two data streams could be registered to one another by taking into account the fixed relative pose between diametric pairs of images. This setup acquired twice as much data as the standard scan, taking nearly twice the time.

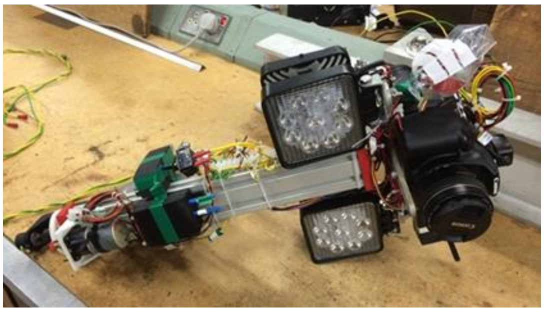

Figure 4. Detachable capture unit. From left to right: rapid-prototyped external gearbox and rotation motor; sealed control electronics with dials to control capture mode and parameters; rotating unit holding cameras and LED arrays.

Figure 5. Left: 0.3mm wide cracks are clearly visible and their evolution may be tracked. Right: a human hair for comparison (approx. 0.1mm). The documentation of cracks wider than 0.3mm is a standard requirement in tunnel construction.

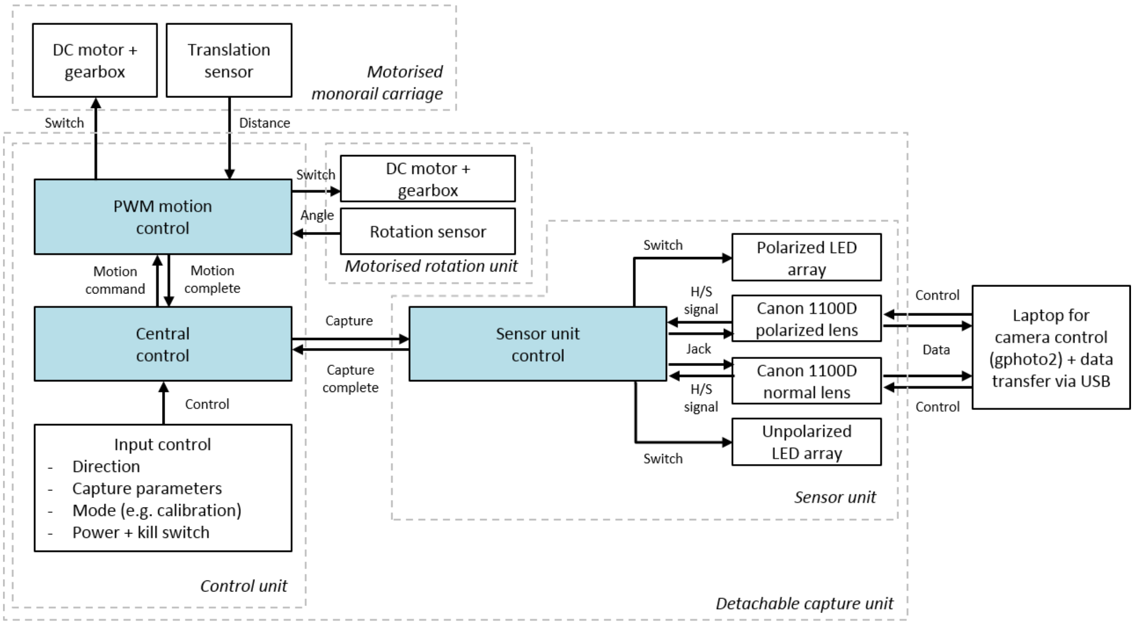

Figure 3. A schematic diagram showing the main system components and the simple interface between them. The modular design means that any one of the units can be upgraded or modified independently. The detachable capture unit can be easily removed from the motorised monorail carriage by hand, considerably simplifying setup and extraction of the system.

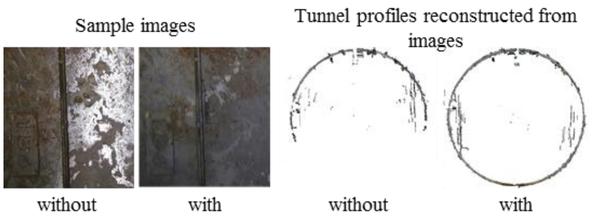

Figure 6. Left: identical images with and without light polarisation and filtering; right: the effect on tunnel profile reconstruction using images taken in wet conditions. Images without polarisation had too much specular variation in damp areas along the bottom of the tunnel and could not be registered, resulting in incomplete profiles.

3.2.3 Ensuring high data quality

To ensure the best possible image quality, we focused on stabilising the device during image capture. Instead of actioning the sensor unit from its central axis of rotation, a large external gear was used to allow a longer radial contact and an anti-backlash spring-loaded gearbox was designed and 3D-printed. Used with pulse width modulation (PWM) and a proportional corrector for DC motor control, the sensor unit could quickly rotate between positions in a smooth and stable fashion.

LEDs on the wide field of view side were fitted with linear polarising filters and the corresponding camera lens with a perpendicular polarising filter. This significantly reduced reflections from damp patches as shown in Figure 6. In the polarised image, the reduction in specularity means stable scene features become visible through water (left). This improves image registration leading to more complete reconstruction in damp conditions (right). Lights and cameras were also triggered sequentially in order to avoid interference between cameras and ensure the most consistent lighting conditions possible.

3.2.4 Low cost

Every component in the system was chosen to help maintain a low cost. The advantage of a low cost system is that many units can be bought and used, leading to better redundancy (if one system breaks down, it is easy to replace) and coverage (the same system can be replicated across many different sites in the network, rather than moved from site to site). Overall, the cost of the system components was under £2,000.

4 Field Testing

The robot was used to gather data in a 3.2m diameter high-voltage cable tunnel currently under construction in London, UK. In total, 700,000 images were taken using the two cameras and a total distance of 2.9km was covered by the robot. The longest stretch was 2.1km, where approximately five days of capture time were taken to complete a detailed scan.

5 Data Analysis

5.1 Reconstruction system

Our first task with the data was to register it into a common coordinate frame, which we achieved using a Structure-from-Motion (SfM) pipeline built around [16]. The image capture strategy was carefully calculated to provide sufficient image overlap and baseline for reconstruction. This produced sparse 3D point clouds as shown in Figure 7, as well as inferred camera poses for every image in the collection. The known linear structure of the data was used to make the process fast and robust.

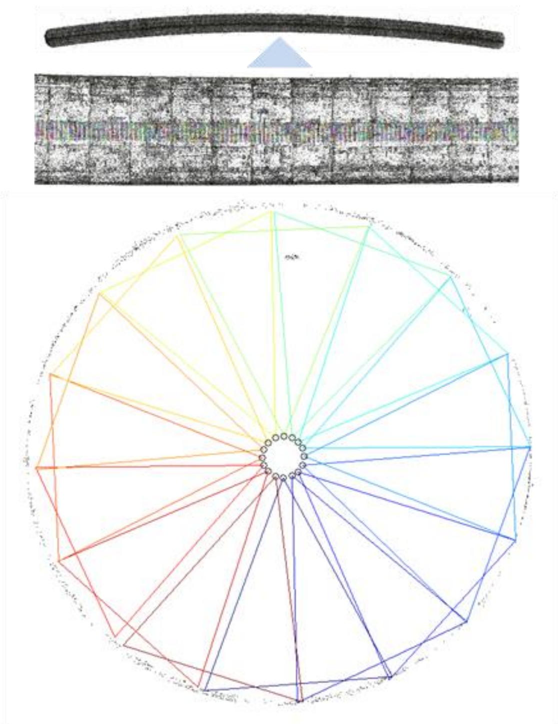

Figure 7. Sample of sparse reconstruction output. Top: a long reconstructed section; middle: close-up sideways orthographic view showing distinct tunnel segments; bottom: plot of a reconstructed tunnel slice showing camera centres (black circles in a central ring), their projected overlapping fields of view (~45º coloured triangles projecting out to the tunnel lining) and reconstructed cylindrical structure around the outside (black dots). The monorail is also visible towards the top of the cylinder as a thin row of points.

5.2 Image mosaicing for condition inspection

Having recovered camera poses and sparse scene structure, the data is suitable to produce large mosaic images of the tunnel lining such as Figure 8, allowing inspectors to get a rapid overview of the tunnel condition. We use a cylindrical projection, mapping 3D (X,Y,Z) co-ordinates from SfM to a 2D (t,θ) image mosaic plane, where t is the robot's translation along the tunnel and θ is the angle of rotation around the sensor axis.

To compile the mosaic, our approach selects central image quadrilaterals from individual images and pieces them together with planar perspective warping to produce a jigsaw-like image (Figure 10). Each quadrilateral is normalised to account for rig lighting, reducing visible seams in the final mosaic (Figure 11). Because of the dense camera sampling and the fact that local planarity holds true for most of the tunnel surface, the result is of suitably high quality for inspection, while also being relatively fast to compute and not requiring the manual identification of control points as in [11] or parametric surface modelling [12].

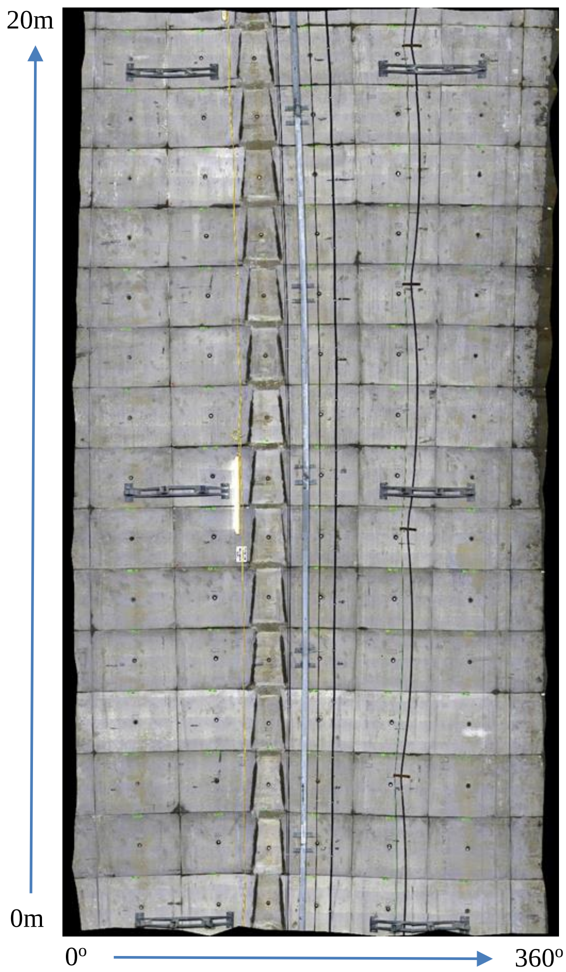

Figure 8. Sample mosaic output over 20m section.

5.3 Cataloguing tunnel segments using barcode detection

Another useful application of the data was the cataloguing of tunnel segments, a task which might usually require a manual walk-through and scanning of stickers attached to the segments. A simple image filtering pipeline was made to detect and retrieve stickers from the image database. The pipeline, illustrated in Figure 9, consisted of (i) thresholding in colour space to segment green patches, (ii) filtering to remove misshapen patches, (iii) description of each detected object using GIST features [17], and (iv) k-means clustering in the first two principal components of the GIST feature space to distinguish between types of sticker. The system successfully detected all stickers in a test image set of 9,600 images, thanks to their unique appearance, and 1,427/1,438 (99%) were correctly classified into barcode vs. non-barcode categories. Since the image locations were known from the 3D reconstruction, they could be pinpointed to a 3D position.

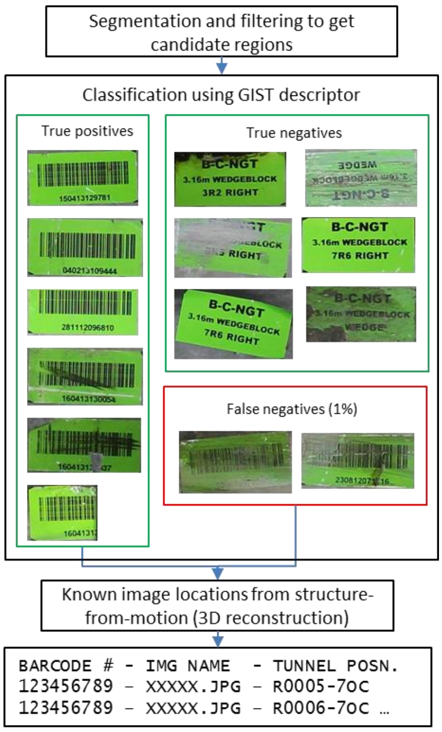

Figure 9. A simple image processing pipeline for barcode detection to catalogue tunnel segments.

This simple example highlights the effort that can be saved when such image datasets can be efficiently gathered and put to good use. A challenging, more useful problem for future work is how to tally and locate arbitrary objects within the dataset, without the need to design a separate detection system each time.

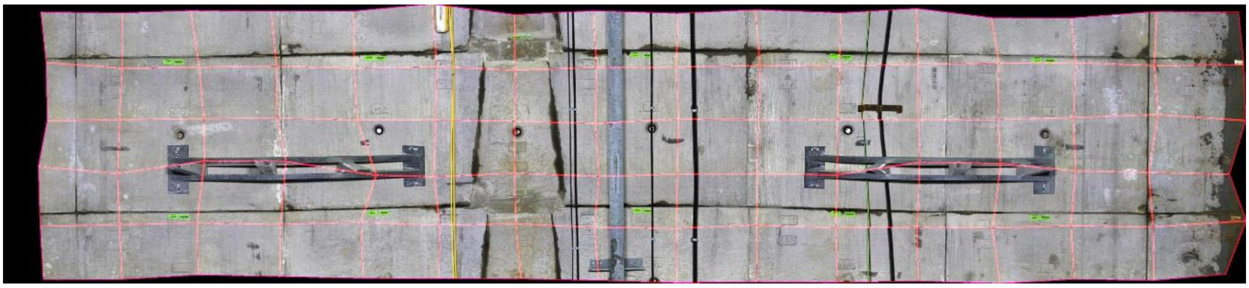

Figure 10. Sample mosaic highlighting individual image segments in red. The fit is found automatically, using 3D points and camera poses recovered from reconstruction. Best viewed in colour.

Figure 11. Above: a single mosaic strip without lighting correction. Seams between image segments are noticeable. Below: after normalising for rig lighting, the seams are reduced, producing a cleaner output image.

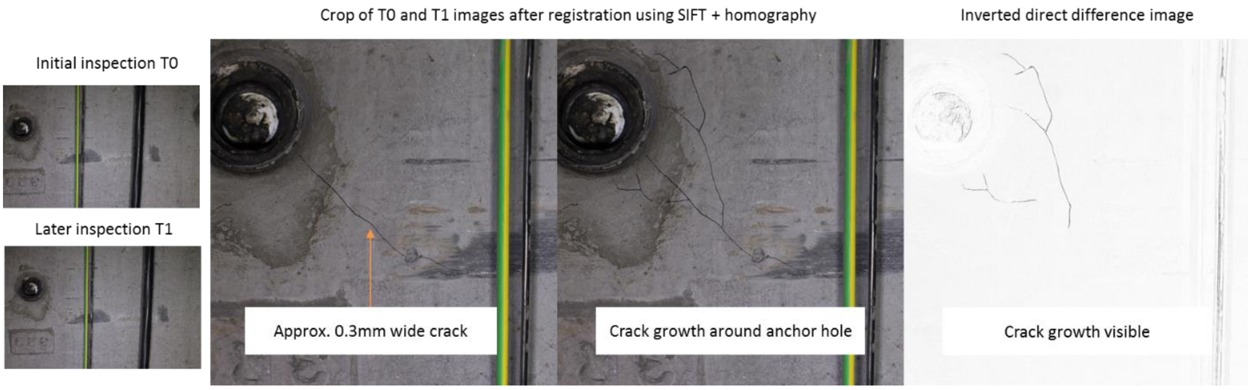

Figure 12. Crack change detection using high-resolution images taken some time apart. The close-up shows the growth of a ~0.3mm wide crack (simulated using a fine marker pen) and the resulting difference image.

5.4 Improving the objectivity of visual condition surveys over time

Finally, we demonstrate a simple example of change detection in tunnel condition over time. Figure 12 shows two images from separate datasets of an area scanned several days apart, where growth of a ~0.3mm width crack has been simulated. A simple image comparison which uses SIFT [18] feature matching to first localise and then register the images via a homography estimation illustrates the potential of this data to reveal changes that manual inspectors may not pick up. While this example does not account for the considerable variation in appearance that can happen over longer time periods, we intend to conduct a more detailed analysis of the problem using the data gathered with our system, in a manner similar to [12].

6 Conclusions

In this work we have described the development of a robot suitable for automated, large-scale, high-resolution image capture in tunnels. We believe our system to be unique for its favourable combination of price, flexibility and data quality. We briefly demonstrated some possible benefits of the field data gathered using the robot, such as image mosaicing for remote inspection, barcode detection for cataloguing tunnel segments and high resolution change detection.

There are numerous areas for further research. We intend to concentrate our immediate efforts on a more detailed analysis of the data gathered during field tests thus far. In particular, we wish to create an improved fault detection system and test it more extensively, comparing the inspection performance between manual, fully-automated and robot-assisted inspection.

On the hardware side, numerous improvements are possible. While the speed of the system was sacrificed in favour of cost, many inspection scenarios require higher speed. This could be achieved by switching to a higher frame-rate camera and using more powerful illumination – both of which are possible without major modification to the existing design. In addition, the rotating sensor unit could be adjusted to include other potentially useful sensors such as a thermal imaging camera or a range scanner.

Acknowledgements

We thank Mark Farmer at National Grid, David Naylor at Mott MacDonald, Riccardo Gherardi and Björn Stenger at Toshiba Europe Computer Research Laboratory and Kenichi Soga at the University of Cambridge for valuable discussions and for their help and cooperation in the project.

References

[1] Balaguer C., Montero, R., Victores, J.G., Martínez, S. and Jardón, A. “Towards Fully Automated Tunnel Inspection: A Survey and Future Trends”, *31st International Symposium on Automation and Robotics in Construction & Mining (ISARC)*, 2014.

[2] National Academy of Engineering. “Grand Challenges for Engineering”, on-line: http://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=02152008, 2008

[3] Gavilan M., Sanchez F., Ramos J. and Marcos O. “Mobile Inspection System for High-Resolution Assessment of Tunnels”, *The 6th International Conference on Structural Health Monitoring of Intelligent Infrastructure*, Hong Kong, 2013.

[4] Shen B., Zhang W., Qi D. and Wu X. “Wireless Multimedia Sensor Network Based Subway Tunnel Crack Detection Method”, *International Journal of Distributed Sensor Networks*, 2014.

[5] Muller U., Kuhn P. and Kadner G. “Crack Detection - Potential Savings in Motion”, *World Tunnelling Magazine (March 2012)*: 25-27, 2012.

[6] Ukai M. “Advanced Inspection System of Tunnel Wall Deformation using Image Processing”, *Quarterly Report of RTRI*, Vol. 48, No. 2, 2007.

[7] Sandrone F. and Wissler R. “Laser Scanning Images Analysis for Tunnel Inspection”, *Harmonising Rock Engineering and the Environment*, 454, 2012.

[8] Paar G. and Kontrus H. “Three-Dimensional Tunnel Reconstruction Using Photogrammetry and Lasers Scanning”, *Proceedings of the 3rd Nordost, 9. Anwendungsbezogener Workshop zur Erfassung, Modellierung, Verarbeitung und Auswertung von 3D-Daten,* 2006.

[9] Yu S., Jang J. and Han C. “Auto Inspection System Using a Mobile Robot for Detecting Concrete Cracks in a Tunnel”, *Automation in Construction,* 16 (3), pages 255–261, 2007.

[10] Hansen P., Alismail H., Rander P. and Browning B. “Visual Mapping for Natural Gas Pipe Inspection”, *International Journal of Robotics Research (online)*, 2014.

[11] Lee C., Chiu Y., Wang T. and Huang T. “Application and Validation of Simple Image-Mosaic Technology for Interpreting Cracks on Tunnel Lining”, *Tunnelling and Underground Space Technology,* 34 61-72, 2013.

[12] Stent S., Gherardi, R., Stenger, B., Soga, K. and Cipolla, R. “Visual change detection on tunnel linings”, *Machine Vision and Applications (online)*, 2014.

[13] Loupos, K. et al. “Robotic Intelligent Vision and Control for Tunnel Inspection and Evaluation – The ROBINSPECT EC Project”, *IEEE International Symposium on Robotic and Sensors Environments (ROSE)*, 2014.

[14] Victores, J.G., Martínez, S., Jardón, A. and Balaguer C. “Robot-Aided Tunnel Inspection and Maintenance System by Vision and Proximity Sensor Integration”, *Automation in Construction*, 20 629–63, 2011.

[15] Malva R. “The Inspection, Monitoring and Diagnosis of Tunnels”, *Proceedings of the World Tunnel Congress 2014 – Tunnels for a better Life,* Foz do Iguaçu, Brazil, 2014.

[16] Wu C. “Towards Linear-time Incremental Structure From Motion”, *3DV*, 2013.

[17] Oliva A. and Torralba A. “Modeling the Shape of the Scene: a Holistic Representation of the Spatial Envelope”, *International Journal of Computer Vision*, Vol. 42(3): 145-175, 2001.

[18] Lowe, D. "Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision,* 60.2: 91-110, 2004