An Image-Based System for Change Detection on Tunnel Linings

Simon Stent*, Riccardo Gherardi[†], Björn Stenger[†], Kenichi Soga*, Roberto Cipolla*

* Department of Engineering,   [†] Toshiba Research Europe,
University of Cambridge, UK   Cambridge, UK
{sais2, ks207, cipolla}@cam.ac.uk   {firstname.lastname}@crl.toshiba.co.uk

Abstract

We present an automated system for detecting visual changes on tunnel linings. By registering new images to a three-dimensional tunnel surface model, recovered using Structure and Motion techniques, we are able to detect and localise changes accurately in order to assist visual inspection by a human expert. We formulate the problem of detecting changes probabilistically and exploit different feature maps and a novel geometric prior to achieve invariance to noise and nuisance sources such as parallax and lighting changes. System performance is assessed on a real data set labelled with ground-truth and collected with a prototype capture device. By localising, organising and classifying image data automatically, our system is a step towards higher frequency visual inspection at a reduced cost.

1 Introduction

Efficiently monitoring the structural health of large-scale tunnel infrastructure is a socially important challenge. Tunnel infrastructure is commonly located in urban areas and provides essential functions to society such as transport, electricity and communications. As urban populations grow, tunnels are often worked beyond their original design specifications for both function and lifespan, and hence face a growing risk of structural failure.

A critical requirement for performance failure prevention is the early detection of any visual changes in the tunnel surface, such as leakages, cracks and corrosion. Early detection allows early intervention, keeping the cost of remedial measures low and reducing the risk of unexpected failure. Detecting such changes is often the work of human inspectors, but given the sometimes adverse working environments and the extensive coverage areas, this is a costly and time-consuming process. More recently, digital camera and laser technology have been introduced [1, 2], but systems remain expensive and routine inspections are typically conducted only once every few years.

In this paper, we present a vision-based change detection system which is a step towards low-cost, high-frequency monitoring. The system automatically detects and localises changes in tunnel linings, reducing the workload and improving the effectiveness of expert human inspectors. It leverages techniques from Structure and Motion (SaM) and image localisation to be device-independent, and is demonstrated to work using inexpensive, off-the-shelf digital cameras.

2 Related work

Our two-part system consists of a preliminary data acquisition and reconstruction step followed by a capture and change detection stage.

Data acquisition and reconstruction. Existing approaches for tunnel lining inspection make use
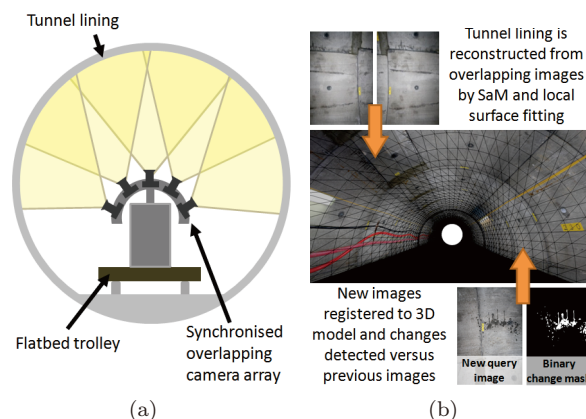
Figure 1: Overview of the system. *(a) Low-cost camera array for data capture; (b) Pipeline summary.*

of more expensive or bespoke visual capture systems such as laser scanners [1], line sensor cameras [2] or calibrated laser/camera hybrids [3]. Instead, we opt for a fixed but otherwise unconstrained array of synchronised, overlapping off-the-shelf digital cameras (Fig. 1(a)). The low cost of the capture device allows for the possibility of assigning one or several devices to monitor individual tunnels continuously for better risk evaluation, rather than using a single expensive device to monitor many tunnels sporadically as is common. From the captured image set, we use SaM techniques [4] to recover 3D geometry, and model the tunnel lining by locally fitting quadric surfaces to the resulting point cloud using robust non-linear optimisation.

As in [5], we find that in the tunnels we consider, a piecewise cylindrical representation is sufficient; though the system can be trivially extended to any shape. A 3D wireframe surface model, texture-mapped with capture images, is shown in Fig. 1(b).

Change detection. Change detection in 2D images is a well-studied problem, particularly in the fields of remote sensing, video surveillance and medical imaging [6, 7]. Systems exist for similar applications to ours, such as pattern-matching for concrete crack detection or road surface condition monitoring [1, 2]. However, the detection of general anomalies on tunnel linings does not seem well explored, despite its importance. We identify the three main challenges and related work from the computer vision literature:

1. *Query image registration.* Accurate registration is an essential prerequisite for accurate change detection. In remote sensing, the standard means of registration is by coarse localisation via GPS, then feature-matching and homography estimation (assuming planar world geometry). In the tunnel environment, GPS is unavailable and 3D relief necessitates a geometric model. Recent voxel-based approaches [8, 9] show promise for multi-view change detection in cluttered scenes with many occlu-

Figure 2: Sources of nuisance variability. *From left to right: new query image with no relevant change; warped matching image from database; absolute difference image. The circled areas highlight significant nuisance variability.*

Table 1: Distance functions.

Feature	Dist.	Formula/Reference
RGB (no norm.)	L2	(R, G, B)
Chromaticity	L2	$(R, G, B)/R + G + B$
Gray-world	L2	$(R/\bar{R}, G/\bar{G}, B/\bar{B})$
Multi-scale Retinex	L2	See [11]
Dense SIFT	Angle	See [12, 13]
Grayscale	NCC	$\frac{1}{n}\sum_{x,y}\frac{(c_{x,y}^q - \bar{c^q})(c_{x,y}^m - \bar{c^m})}{\sigma_{c^q}\,\sigma_{c^m}}$

sions. The tunnel environment is in general un-cluttered and well-defined, and as we are specifically interested in its lining, we opt for a simpler, more scalable, local quadric surface model.

2. *Nuisance variability.* Fig. 2 shows some typical sources of nuisance variability in the tunnel environment. One source of false changes between the registered images is image parallax from un-modelled geometry. This can be avoided by explicitly modelling all geometry [9], however, as the tunnel surface is our main concern, we circumvent the problem using a nuisance mask, in the style of [10], which downweights regions of the change image depending on their adherence to the fitted surface. A further source of nuisance variability is illumination, amplified by the enclosed and poorly lit nature of the tunnel environment. In surveillance applications, background modelling is used to mitigate this variability, but is not feasible with limited temporal information. We investigate single image colour-normalisation and colour-constancy techniques such as Multi-Scale Retinex (MSR) [11], to counter both high and low frequency illumination variability.

3. *Many modes of relevant changes.* Many existing systems for large-scale infrastructure monitoring focus on detecting specific features such as cracks in concrete. Our main concern is to capture all statistical outliers which are not accounted for by understood modes of nuisance variability.

3 Theory

We denote the query image by I^q, a function $D^q \rightarrow \mathbf{R}^3$ from location $\mathbf{x}$ to an RGB value $I^q(\mathbf{x})$. The set of matching images is given by $\{I_i^m\}_{i=1}^M$. This is the set of images taken at a previous time instance with non-zero intersection of the image domains, registered to the query camera viewpoint.

We are interested in obtaining a change map $C : D^q \rightarrow \{0, 1\}$, which maps a location to 1 in case of a (relevant) change and 0 otherwise. The goal is to achieve invariance to nuisance variability as described above and return a probability of relevant change for pixel-level or image-level classification.

3.1 Change detection

We first consider the case of estimating a single change map C_i from the query image I^q and one of the pictures I_i^m matched to it from the database:

$$p(C_i \mid I^q, I_i^m) \propto p(I^q, I_i^m \mid C_i)\, p(C_i) . \qquad (1)$$

As observation we define a distance function d between the query and the matching image, so the likelihood term $p(I^q, I_i^m \mid C_i)$ is defined as the distribution of the values of d given that a change has occurred at a particular location:

$$p_d(I^q, I_i^m \mid C_i) = \begin{cases} \exp\left(-d(I^q, I_i^m)/\sigma^2\right) & \text{if } C_i = 0 \\ \mathcal{U}(d) & \text{if } C_i = 1 \end{cases}, \quad (2)$$

where $\mathcal{U}(d)$ is the uniform distribution over the range of values of d. The smoothing constant σ is set as the mean value of d over the whole query image set. Each matching image, I_i^m, provides information for changes to be identified in its areas of overlap with I^q.

The term $p(C_i)$ is the prior probability of a changed pixel, which we set to a constant value, but can be varied depending on the location of the image pixel within the tunnel to vary the change detection sensitivity.

The distance function, d, maps corresponding query and matching image pixels into a feature space, before returning a value, $d(I^q, I_i^m)$, using some distance metric. A good choice of distance function is any function that detects relevant changes, yet is invariant to changes due to nuisance variables.

Table 1 details a subset of the functions that we examined: chromaticity and gray-world (colour-normalisation techniques); MSR (a colour-constancy technique); dense SIFT (a spatial histogram of gradients technique); and grayscale normalised cross-correlation (NCC) (measuring textural similarity).

In many cases the query image contains regions which are visible in multiple matching images. In these areas, we combine the result of the individual change maps:

$$p(C(\mathbf{x})) = 1 - \prod_{\{j:\mathbf{x} \in D_j^m\}} (1 - p(C_j(\mathbf{x}))) , \qquad (3)$$

where D_j^m is the domain of the matching image I_j^m.

3.2 Geometric prior

We use the information available to us from our SaM reconstruction to form a geometric prior, $p(C|\mathcal{G})$, included as follows:

$$p(C_i \mid I^q, I_i^m, \mathcal{G}) = p(C_i \mid I^q, I_i^m)\, p(C|\mathcal{G}) . \qquad (4)$$

The prior makes use of the recovered scene geometry, $\mathcal{G}$, which maps image locations to corresponding 3D points: $D^m \subset D^q \rightarrow \mathbf{R}^3$. The objective of the prior is to mask out nuisance changes caused by geometry or poorly reconstructed features. It can thus also be thought of as an inverse 'nuisance map' [10].

To construct the prior, we first group the 2D SIFT features into an inlier (on-surface) and outlier (off-surface) set, based on the distance of their corresponding 3D points to their closest point on the locally fitted surface. Given the relatively sparse nature of $\mathcal{G}$

| (a) | (b) | (c) | (d) |
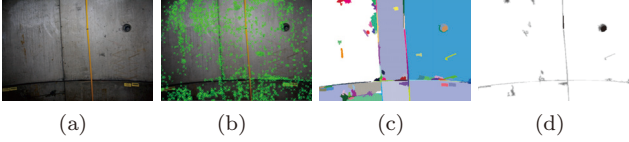
Figure 3: Geometric prior. *(a) Query image (b) Distribution of reconstructed SIFT features (green) (c) Mean-shift segmented image with colour-coded segments (d) Final geometric prior (black areas indicate off-surface or uncertain geometry).*

(Fig. 3(b)), we next apply mean-shift segmentation to the query image [14]. This delineates the image into pixel groups of similar colour and texture (Fig. 3(c)). Inliers and outliers contained within a pixel group vote towards its overall classification. Pixel groups containing only outliers are classified as off-surface and assigned a prior probability of zero, i.e. changes in those regions are considered to be nuisance variability and are ignored. Pixel groups containing more inliers than outliers are classified as on-surface and assigned a prior probability of one. For pixels lying in groups which contain no points, or fewer inliers than outliers, the prior depends on the distance of the pixel to the nearest inlier. The prior is therefore expressed:

$$p(C(\mathbf{x})|\mathcal{G}) = \begin{cases} 1, & \text{for on-surface groups.} \\ 0, & \text{for off-surface groups.} \\ \exp\left(\frac{-||\mathbf{x}-\mathbf{x}_{in}||}{\sigma_{\mathcal{G}}^2}\right), & \text{otherwise.} \end{cases}$$

(5)

where $\sigma_{\mathcal{G}}$ controls smoothness in uncertain regions and $\mathbf{x}_{in}$ signifies the nearest inlying 2D SIFT feature. The geometric prior thus downweights changes where the geometry is either known to be off-surface and known to be unreliable. The latter point is important in the tunnel environment, where off-surface features such as cables and boxes tend to have matte, featureless surfaces and are therefore reconstructed poorly using feature-based SaM. Fig. 3(d) shows an example of a geometric prior mask, downweighting changes along the yellow cable and in the panel anchor (surface hole).

4 Experiments

We used a prototype capture system consisting of five synchronised cameras with flash units, arranged in a semi-circular array as shown in Fig. 1(a). We captured data covering 180° of a 3m diameter tunnel section of 100m length. This comprised 1,000 images at a resolution of $3,888 \times 2,592$ pixels.

Next, artificial changes were made to the concrete tunnel lining to simulate leaking, cracking, spalling and other visual changes. A query set of 232 images was taken, of which 131 contained relevant changes and were hand-labelled with ground truth.

The texture of the tunnel lining was found to be sufficiently unique to accurately register query images within a 50m section using robust 2D-2D SIFT matching. Query image cameras were resectioned with radial distortion estimation. Query image matches were then back-projected into the recovered geometry, $\mathcal{G}$, and reprojected into the query image for change detection.

4.1 Qualitative results

Fig. 4 illustrates various distance functions, the geometric prior mask and a final change detection image for three sample queries. Relevant changes in

sequences 1 and 2 include leaking, fine chalk markings, discolouration and new items attached to the surface. The three illustrated distance functions, gray-world, MSR and NCC, pick up details with different degrees of success. Gray-world tends to amplify changes, with good resolution, but the model we use is a global one and hence illumination effects are also undesirably amplified. MSR removes a smoothed range of low-frequency nuisance variation, retaining a high-frequency, detailed mask. However, grayscale changes are not as prominent as colour changes. NCC highlights even very fine changes by taking into account both intensity and spatial information, but at the cost of reduced resolution. All methods falsely detect changes from the lighting units and cabling. This is even more evident in sequence 3, with strong parallax and specularity in the scene.

The geometric prior in all three cases correctly identifies and removes most of the nuisance change caused by off-surface features. The final column shows a probabilistic output change mask, formed by a combination of gray-world, MSR and NCC features, factored by the geometric prior as per (4). In sequences 1 and 2, its performance is close to ground truth. Sequence 3 illustrates a failure case, caused by the unusual presence of some thread on the normally featureless red cable.

### 4.2 Quantitative feature comparison

The image-level detection performance of various combinations of features is compared in Fig. 5(a). Distance functions taking into account local spatial information (NCC and DSIFT) performed better than methods which do not take advantage of dependencies between neighbouring pixels, such as MSR and gray-world. NCC proved best at detecting all-round changes, picking up almost all of the positive change cases at about 50% false positive rate. A combination of NCC with both MSR and gray-world approaches was examined. Qualitatively, results on a pixel level were improved, though image-level performance with this combination of pixel and region-based methods was found to be unchanged.

4.3 Effect of geometric prior

We compared the image-level detection performance of two feature combinations, gray-world and gray-world NCC, in three scenarios: without a prior; using the mean-shift based geometric prior described in Section 3.2; and one using SLIC superpixel segmentation instead of mean-shift [15]. ROC curves are shown in Fig. 5(b). Classification performance after the introduction of the geometric priors increases substantially. With no prior, gray-world is initially far more discriminative than NCC, which is more prone to detect changes across nuisance areas of the image space. When a prior is introduced however, nuisance regions are masked out and NCC can safely exploit local spatial information solely on the area of interest (tunnel surface) and outperform gray-world. Finally, the performance of our proposed mean-shift prior is shown to improve on the more local, SLIC-based prior.

5 Conclusions

We have presented a system which is suitable for automated monitoring and detection of general visual changes on smooth, unpainted, concrete tunnel linings.
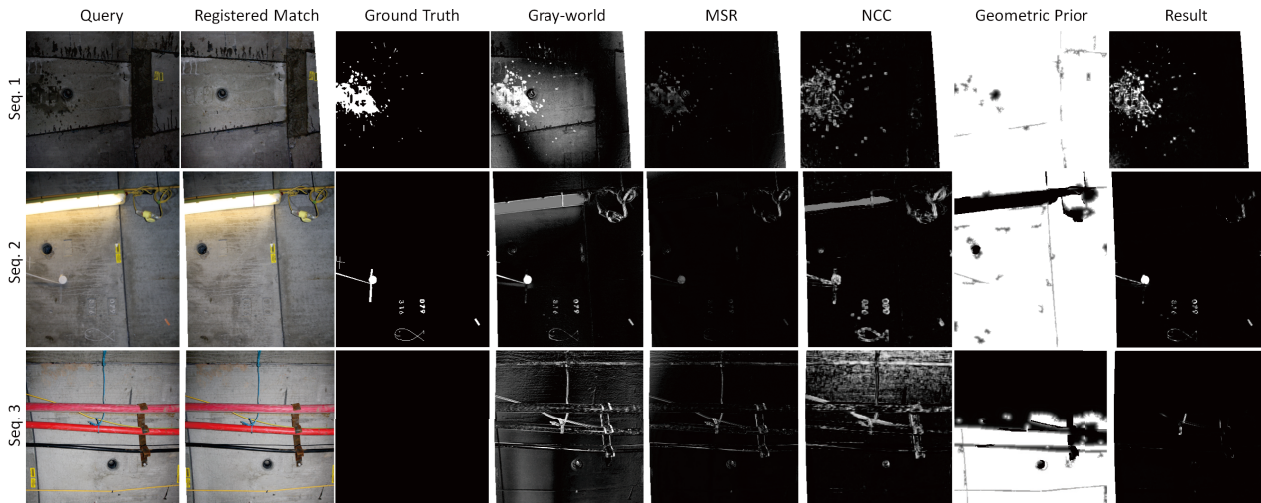
Figure 4: Illustrative results for three cases. *Two sequences (1 and 2) feature relevant change: water leakage, chalk marks, added features; all three sequences feature nuisance change: lighting change (1), cable and fixtures (2 and 3).*

(a)

(b)
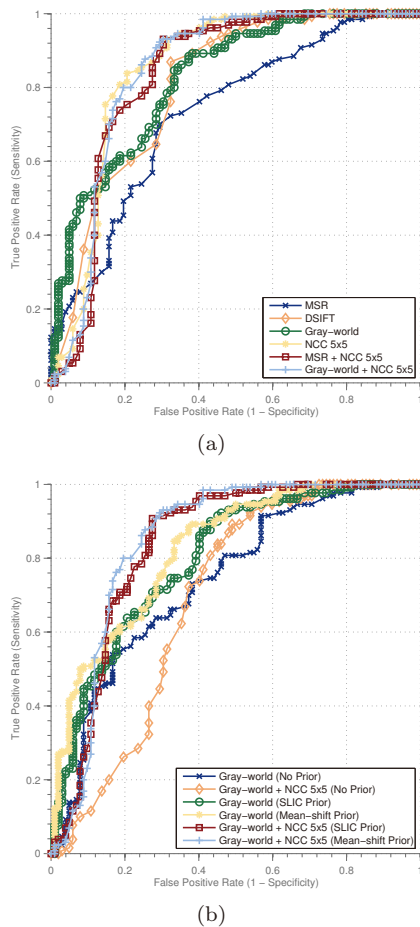
Figure 5: ROC performance comparison. *(a) Image-level performance of various features; (b) Image-level performance with and without geometric prior. Area-based methods with a discriminative geometric prior were found to perform best.*

Our system is inexpensive to implement and can reduce the workload for visual inspection significantly, facilitating high frequency, effective tunnel inspections. In future, we plan to test our system in an active tunnel environment to detect real changes. We are developing our capture device to acquire larger volumes of data automatically and will further explore nuisance invariant features. We also wish to extend the approach to more complex tunnel geometries.

References

[1] SPACETEC, "http://www.spacetec.de," 2012.

[2] M. Ukai, "Advanced Inspection System of Tunnel Wall Deformation using Image Processing," *Quarterly Report of RTRI*, vol. 48, no. 2, pp. 94–98, 2007.

[3] T. Ortner, G. Paar, G. Hesina, R. F. Tobler, and B. Nauschnegg, "Towards True Underground Infrastructure Surface Documentation," *Real Corp Proceedings 2010, Vienna*, pp. 18–20, 2010.

[4] N. Snavely, S. M. Seitz, and R. Szeliski, "Modeling the world from internet photo collections," *Int. Journal of Computer Vision*, vol. 80, no. 2, pp. 189–210, 2008.

[5] K. Chaiyasarn, T.-K. Kim, F. Viola, R. Cipolla, and K. Soga, "Image mosaicing via quadric surface estimation with priors for tunnel inspection," in *ICIP*, pp. 537–540, Nov. 2009.

[6] D. Lu, P. Mausel, E. Brondízio, and E. Moran, "Change detection techniques," *International Journal of Remote Sensing*, vol. 25, pp. 2365–2407, June 2004.

[7] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, 2005.

[8] T. Pollard and J. L. Mundy, "Change Detection in a 3-d World," in *CVPR*, June 2007.

[9] A. Taneja, L. Ballan, and M. Pollefeys, "Image based detection of geometric changes in urban environments," in *ICCV*, pp. 2336–2343, 2011.

[10] A. Ravichandran and S. Soatto, "Long-range spatio-temporal modeling of video with application to fire detection," *ECCV*, pp. 329–342, 2012.

[11] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes.," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 965–76, 1997.

[12] D. Lowe, "Distinctive image features from scale-invariant keypoints," *IJCV*, vol. 60, no. 2, pp. 91–110, 2004.

[13] A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms," 2008.

[14] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *Trans. PAMI*, vol. 24, no. 5, pp. 603–619, 2002.

[15] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels," *EPFL Tech. Report No. 149300*, 2010.