Detecting Change for Multi-View, Long-Term Surface Inspection

Simon Stent[1]
sistent@cantab.net
Riccardo Gherardi[2]
riccardo.gherardi@crl.toshiba.co.uk
Björn Stenger[2]
bjorn.stenger@crl.toshiba.co.uk
Roberto Cipolla[1]
rc10001@cam.ac.uk

[1] Machine Intelligence Lab
Department of Engineering
University of Cambridge
Cambridge, U.K.

[2] Toshiba Research Europe Ltd.
Cambridge, U.K.

We describe a system for the detection of changes in multiple views of a textured surface taken at different times by a moving camera. Our motivation is the development of a non-contact inspection system, summarised in fig. 1, to be used for detecting anomalous visual changes on surfaces - in this case on concrete tunnel linings. This application is of increasing social importance as tunnels and other large-scale infrastructure age and more efficient methods for structural inspection are required to allow their continued safe operation.

The problem is challenging for several reasons: (i) Size and nature of changes. Changes of interest are often small and subtle – e.g. a fattening in the width of a hairline crack or a patch of discolouration caused by organic growth or surface damage. (ii) Nuisance factors. A sizeable proportion of the observed change over time is caused by nuisance factors, either internal to the acquisition system (such as different image sensors, capture settings or lighting setup) or due to external causes (for example, seasonal changes of temperature and humidity). (iii) Registration error. Achieving the pixel-accurate registration typically required for change detection is challenging because neither the sensor positions nor the tunnel geometry can be reliably determined. Parallax errors are common.

We address these challenges by first using a structure-from-motion pipeline to approximately register images from our robotic inspection rig to the reconstructed surface of interest. Given a pair of registered image patches from different times, our main contribution is a novel approach to detect changes between the patches using a two-channel convolutional neural network (CNN). CNNs have recently been shown to be very effective at learning invariance to certain modes of image variability, but require large amounts of labelled image data to train. We create an unlimited source of negative pairs (i.e. patches where no abnormal change has occurred) by taking registered viewpoints from different cameras from the same time. We supplement this with a smaller dataset of negative pairs across the different test times from regions where no changes of interest have occurred. This requires a limited effort in coarsely labelling a small subset of the test data. Together, these negative pairs capture much of the natural variance from nuisance factors and registration error. For the positive (changed) pair generation, we provide randomly sampled pairs as well as synthetically generated changes using a crack model (fig. 2). The homogeneity of the tunnel environment allows a network to generalize well from a manageable amount of labelled ground-truth data.

Our approach is similar to [1], who learn a two-channel CNN for the inverse problem of similarity measurement between image patch pairs. A key difference is that we train directly on a mixture of synthetic data and task data generated by our own pipeline, allowing us to learn task-specific invariances for the improved detection of changes.

We evaluate our system using three datasets captured from a live tunnel over two months (fig. 3). A trained inspector was tasked with simulating real changes in the tunnel between captures and a set of ground truth change images were generated for testing. We compare our method against our existing probabilistic change detection approach [2] and against the results of a manual inspection carried out by a second trained inspector in the field (table. 1). This comparison is of particular importance to industry, since manual inspection is still in common practice. Finally, we visualise our system's output using the Google Maps API (fig. 4).

[1] S. Zagoruyko and N. Komodakis. Learning to compare image patches via convolutional neural networks. In *CVPR*, 2015.

[2] S. Stent, R. Gherardi, B. Stenger, K. Soga, and R. Cipolla. Visual change detection on tunnel linings. *Machine Vision and Applications*, pages 1–12, 2014.

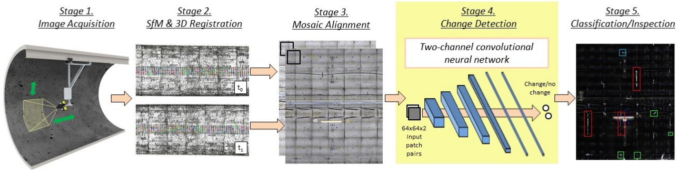
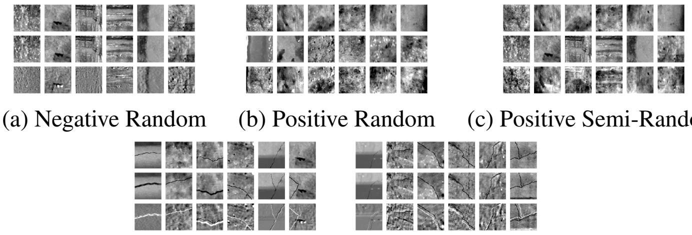


Figure 1: System overview. The main novelty is in stage 4, in which changes are detected between registered sets of image mosaics captured at different times. We propose and evaluate a new approach using a two-channel CNN. The network learns a model for normal modes of image variation, so as to detect abnormal changes with fewer false positives.



(a) Negative Random (b) Positive Random (c) Positive Semi-Random

(d) Positive Crack (e) Negative Crack

Figure 2: Sample training pairs (rows 1+2) and their difference images (row 3) from different training sets: (a) negative (unchanged) pairs; (b) positive (changed) random pairs, with both members chosen randomly; (c) semi-random positive pairs, combining (a) and (b); (d) positive crack pairs, including crack appearance/disappearance, extension and widening; (e) negative crack pairs.

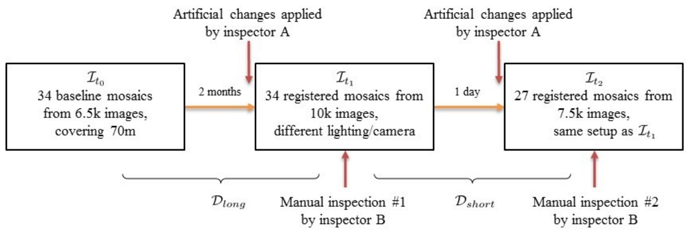


Figure 3: Timeline and datasets gathered for the inspection experiments.

Inspection dataset:	$\mathcal{D}_{short}$ (1 day)			$\mathcal{D}_{long}$ (2 months)		
FPR per pixel:	0.01	0.05	0.10	0.01	0.05	0.10
Manual inspection	0.29	0.29	0.29	**0.58**	0.58	0.58
Prior method [2]	0.20	0.55	0.64	0.00	0.00	0.32
Proposed method	**0.73**	**0.87**	**0.87**	0.26	**0.65**	**0.84**

Table 1: Percentage of artificial changes detected by the compared systems at different false positive rates. Changes are considered detected if they are greater than >50% positively labelled.

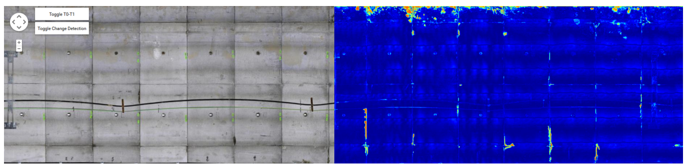


Figure 4: Visualisation using Google Maps API with maximum resolution of 0.3mm/px. Left: mosaic image of an 8m section; right: change mask.