Crack detection in inspection by a fusion of computer vision and machine learning methods

DNV GL, Høvik and Trondheim,
in cooperation with Prof. Rudolf Mester, IDI

Background

DNV GL has actively researched how to apply computer vision and machine learning techniques to automation of our remote inspection services.

Currently, we are running an NFR project, Autonomous Drone-based Surveys of Ships in Operation (ADRASSO), which aims to develop computer-vision-based framework to automate the DNV GL ship survey process. The application scenario is that an autonomous drone equipped with RGB cameras flies within a ship tank and inspects the condition of the tank by taking videos. The videos taken by the drone might be processed either on-board of the drone for real-time (defects detection) display and navigation guidance of drone flight path, or offline at the DNV GL office premise using a highly reliable and robust defect detection model.

To develop such a computer-vision-based pipeline, we have investigated several state-of-the-art object recognition models including object classification (image-level), object detection (bounding box + localization), and semantic segmentation (pixel-level) to detect the existence of a particular type of detects, i.e., cracks, in static images. Both object detection and semantic segmentation models have achieved very good results on our static ship survey images. However, we still face the challenges:

- The number of false positives ('false alarms' is to be kept as low as possible
- To achieve a very low false negatives rate (missed cracks) is very critical for the whole application.

Since the ultimate goal of the ADRASSO project is to analyse videos to reliably detect actual cracks while not generating too many false positives, we can utilize the temporal-spatial correlation of videos to combat the weakness of the object detection model.

This setting defines the environment for several research topics that can be addressed individually, or in a cooperative manner by several students.

(3) Detecting cracks from a stack of motion compensated images

This task combines modern method of video processing, in particular scene tracking an camera motion estimation with detection methods which today most often come from the domain of machine learning. The approach to the crack detection problem proposed here exploits that idea that cracks typically do not only have a two-dimensional appearance (for instance, like a line drawn on the material), but also have a three-dimensional structure, often a very tiny gap in the surface. It can be expected that a moving camera, in particular when it carries its own illumination, will show the crack not only by its two-dimensional appearance, but also by the change of this appearance, caused by illumination and image acquisition from different viewpoints.

In order to exploit this effect, a set of image frames from the video stream has to 'motion-compensated', that is: the position and orientation of a crack candidate is stabilized in the visual field for a short while, say, for about 3 seconds or 10-100 image frames. Here, temporal subsampling, that is: taking only each n-th frame may be permitted. In the stabilized image subsequence, it is expected that the crack varies in appearance in a typical way which is related to the change of illumination.

The motion-compensated image subsequence is then fed into a detection unit which does not work on a single RGB image, but on a stack of such images. A deep network trained on true (or synthetic) crack images with a CNN in the front end is a highly promising candidate for the detector architecture. Also for this task it will be a challenge to come up with training methods that allow to train the network with a limited set of training images.

A potential side aspect of this task which might turn out to be of key importance to the feasibility of the solution is the attempt to generated synthetic, but still realistic crack images, e.g. by combining real crack images which are artificially augmented by a hallucinated 3D surface structure and 're-illuminated' using advanced computer graphics methods. This approach is optional, but might develop into a rewarding endeavour for students which are also interest in image rendering.

Contact person: Prof. Rudolf Mester, IDI, Rudolf.Mester@ntnu.no