Monitoring of Bridges using Computer Vision Methods

ABSTRACT: This paper is an overview of the development and application of Computer Vision for the Structural Health Monitoring (SHM) of Bridges. A brief explanation of SHM is provided, followed by a breakdown of the stages of computer vision techniques separated into laboratory and field trials. Qualitative evaluations and comparison of these methods have been provided along with the proposal of guidelines for new vision-based SHM systems.

KEY WORDS: Computer Vision; Bridge Monitoring, Structural Health Monitoring.

1 INTRODUCTION

Existing Civil infrastructure is under an increasing level of stress from loading/environmental effects. These effects can be detrimental to the integrity of the bridges, and must be monitored in order to avoid dangerous incidents and insure public safety.

Visual inspections remain to be the most common method of bridge inspection worldwide. This method is used as a means of detecting obvious damage to structures such as cracks/ shifting of components and is carried out by following a set of established guidelines according to bridge type. This method has many limitations which affect its reliability and is extremely sensitive to human error, particularly since a visual inspection is rarely carried out by a senior engineer. A survey of the reliability of visual inspections has detailed the high level of variability in this assessment method [1].

In recent years Structural Health Monitoring (SHM) systems have been developed to try to overcome these limitations. SHM can provide an unbiased means of determining the true state of our aging infrastructure. Sensor systems are used to monitor bridge deterioration and provide real information on the capacity of individual structures, hence extending bridge life and improving safety. Changes in stiffness is usually measured using strain sensors, but recent research has indicated that measuring displacement changes from calibrating vehicles can be used as a method of detecting bridge condition[2]. The deflection readings are gathered with respect to vehicles of known weight passing over the structure. If pre-weighed calibration trucks are not available for testing, vehicle weights can be either gathered from a database in order to gain approximate readings, or by using a weigh-in-motion system to gain precise information on vehicle weights.

Traditionally displacement is measured using transducers or accelerometers which are attached to fixed points on the structure. The transducers, such as linear variable differential transformers (LVDT), give a direct reading of displacement but generally require an independent frame for mounting, which for most bridges make them impractical for use in the field. On the other hand, accelerometers can be attached directly to the structures (i.e. do not need a fixed reference) but post processing of the data is required to convert acceleration to displacement, as shown in Figure 1.

Figure 1. Integration Algorithm for Converting Accelerometer Data to Displacement [3]

While these readings can, under certain circumstances, provide a reasonably accurate estimate of deflection, there are several related problems with the method, such as:

1) Durability as the sensors can be damaged easily.
2) Equipment is expensive to purchase and time consuming to set up on site.
3) Measurement noise in the acceleration signal can result in errors in the calculated displacement signal.

In recent years, new methods have been investigated to address these issues, such as GPS readers, Laser Vibrometers and Computer vision techniques. GPS readers have been proved to be a reliable method of detecting deflection but
require satellite connectivity to function, which may not be feasible at all sites, and they are susceptible to high electromagnetic noise[4] they are also not ideal for short span bridges where movement ranges are modest. Laser vibrometers are comparable in effectiveness with transducers [5], but they are costly and difficult to set up/operate. On the contrary, video cameras are cheap, reasonably easy to use and can inspect a structure from a distance. They also mimic human visual inspection. The main drawback in their use as part of an SHM system is that the captured images require complex and intelligent data processing and analysis through computer vision algorithms.

This paper aims to provide a background on the evolution, application and practicality of Computer Vision techniques in SHM along with recommendations for future work in this area.

2 COMPUTER VISION IN SHM EVOLUTION

Computer vision is becoming a more widespread method of SHM. It operates by recording motion pictures of a target area on a civil structure. Early applications of Computer Vision to SHM involved the development of a hybrid system composed of traditional sensors paired with cameras in order to produce the output (Figure 2). Zuarin et al. [6] proposed a system where the cameras would replace the systems used for attaining vehicle weights, and be paired with sensors for measuring displacement/acceleration. A camera monitors traffic on a bridge in order to determine the amount of traffic passing over it, while a transducer is affixed to the bottom of the bridge to measure deflection.

More recently computer vision has been also applied as a replacement of transducers and accelerometers on measuring the bridge displacement. In these systems, the basic principle involves taking an image of the bridge to establish reference points, this is then compared to subsequent images to determine displacement. This process is known as registration [6]. There are several methods of processing the images gathered, which will be discussed in the following section.

Figure 2 Hybrid Camera-Sensor System proposed by Zuarin et al[7]

2.1 Traffic detection

A method for detecting and tracking traffic is detailed in [8]. The model uses a background subtraction method, wherein a reference image of the empty bridge is taken, and any images featuring traffic is compared to this image, and any new objects are assigned to the foreground. These objects are treated as “blobs” tracked through the video. While this method tracks vehicles well in optimum conditions, it is not capable of handling invariant natural conditions such as low light/ rain and fog which makes it impractical for long-term monitoring on site. In addition, this approach is not suitable in congested traffic, as multiple small vehicles could be misread as one large object or large vehicles could block smaller ones from the camera. Those are two common problems in the current state of the art in computer vision.

A superior method is laid out in [9] where the authors extract vehicle features and use these features to track the vehicle, as shown in Figure 3. The system is not sensitive to the effects of changing lighting due to weather because only the most prominent features are tracked through frames and then grouped by motion to give a reading of vehicle location. This method is less susceptible to the occlusion problems that occurred in the previous approach because the system groups features that move together, so the only time that multiple vehicles could be grouped together is if they were extremely close to each other in the same lane and moving at identical speeds, which would be unlikely due to safety constraints.

Figure 3 Features of Vehicles tracked [10]

A comprehensive survey of video processing for traffic applications has been carried out by Kastrinaki et al[11].

2.2 Hybrid Camera-Sensor Approach

Once the traffic can be tracked and classified, the gross vehicle weights can be determined from a reference database. This information can then be used to determine the expected deflection that will be read by the attached sensors in a Hybrid approach.

2.2.1 Laboratory Work

A hybrid sensor camera system for classifying vehicles was developed in [12]. In this paper, the authors laid out a system for grouping vehicles into 7 different classes depending on the readings from strain gauges that were time synchronised with video images of the vehicles passing over a test bed setup. A
neural network was developed to classify the vehicles based on the application of a Bayesian filter to the collected strain gauge readings. The main purpose of the video images was to establish the location of the vehicles on the deck. However it was suggested that they could also be used to give gross weights of vehicles by assigning weights based on classes that were determined by an image based neural network.

A lab trial of another camera-sensor method was carried out in [7]. The computer vision algorithms were primarily used to determine the type of vehicle crossing their model bridge. Additionally the position of the vehicle at certain times was logged in order to build the Unit Influence Line (UIL) using the data obtained from the transducers placed on the underside of the model bridge at the times determined by the video. The proposed system provided promising results, particularly in detecting changes to deflection based on various damage scenarios, but it could not be immediately transposed to the field because it does not deal with an inherent issue in sensor-camera systems: time synchronisation. The data logger and USB camera in this study are linked to the same computer and can be synchronised to the same clock, but this approach would not be feasible in the field as a USB camera would not have the required pixel resolution to detect deflection from a distance. Moreover, the vision based system can only differentiate between 3 types of pre-determined vehicles, while a larger database is needed in order to create a viable real solution.

2.2.2 Field Work

Fraser et al detailed a system for combining camera input with sensor data in [13]. This system involved the combination of accelerometers and strain gauges with cameras for detecting and classifying vehicles. The data acquisition for the sensors/cameras was managed by the use of a wireless cloud which accessed a wired network in a nearby building. While this method is suitable in a built up area, it may not be possible in rural locations. The LabVIEW code used to synchronise the sensors also controlled the traffic-facing camera and captured images at a rate of 3 frames per second and a resolution of 640x480 pixels. It has not been specified if the code is efficient enough to improve this rate of capture and resolution, which will hamper its application to real time measurement of displacement as the scanning rate is insufficient for real time dynamic analysis. Critically a method for correcting camera movement due to wind/vibration has not been specified, an issue detailed in [14].

Another hybrid-camera sensor method trialled in the field was [15]. This method delivers an improved reading of bridge condition due to the correlation of the class-based vehicle loads with recorded bridge responses. This allows the system to give an approximate indication of where damage has occurred on the bridge. These systems have served as a precursor to camera-camera setups as they have proved the viability of replacing sensors for determining vehicle loads on bridges. In order to achieve a camera-only SHM approach, the next step is the replacement of sensors with cameras for reading displacement on bridges, which is analysed in section 2.3.

2.3 Video Registration for SHM

2.3.1 Target Based

The initial applications for Computer Vision as registration for SHM involved the use of target-based approaches. This involved affixing premade targets or markers such as LEDs, speckle or other randomised patterns to the bridge which are used as stable easily identifiable features to be tracked through a video, as shown in Figure 4.

![Figure 4 Speckled pattern applied for field test][16]

2.3.1a Laboratory Work

The readings obtained from camera-only systems are compared to readings from traditional sensors in order to determine their suitability for replacing these sensors in the future. The study detailed in [17] involves applying a Digital Image Correlation (DIC) technique to video of a shake table test carried out in the lab. The DIC readings were compared to verified measurements from accelerometers, which were attached to the test specimen. The method used a correlation approach [18], where an image was divided into sub-images, and the position of sub-image A in the reference image is compared to its position in further images in order to plot a translation matrix and determine displacement of the target point. The method has comparable results to accelerometer readings, but again does not cater for camera movement/vibration or differing light levels since it is run in a lab controlled environment. No information has been provided to clarify how the pixel units were converted to engineering ones for comparison with accelerometer readings, therefore the accuracy of their results cannot be verified.

A lab trial carried out in [19] describes an accurate (within 0.09mm of LVDT at monitoring distance of 2m) method for calculating displacement based on movement of white points on a black background inside a region of interest. This point tracking method is similar to an optical flow methodology proposed by [20], where key-point features are selected based on relative light intensity to the neighbouring pixels [21][22]. The features are then tracked through the subsequent frames and a plot of their movement can be created. The trial also used a region of interest in the image based on expected displacement and camera zoom capabilities, which reduces the image size that needs to be processed and enhances the processing time of the algorithm to within 1/30s and makes real-time measurement possible. This system also incorporates a method for synchronization of multiple cameras using a master-slave setup where PCs are linked over a wireless network in order to track multiple points of displacement simultaneously. This method proved to be quite effective in solving the issue of time synchronization in the lab, the issue of setting up a wireless network in the field could still be a factor however, as explained before. There is also no consideration given to overcoming the difficulties of
movement of the camera or environmental light changes during monitoring. The calibration of the camera to ensure accurate pixel-engineering units is an important step, and a valid method is laid out in [23], where the authors use the extrinsic parameters of the camera to obtain the conversion factor for that particular camera, which removes the need for a calibration target to be attached to the structure and made visible in the images recorded.

Correlation and optical flow approaches are merged in [24] in an attempt to gain sub-pixel accuracy of object displacement. The authors used computer-generated random patterned images and applied rotation and transformation effects to them. Their method did deliver the desired accuracy (maximum mean bias error of 0.03 pixels for 0.5 pixel displacement) and could potentially be modified for use on real images as it is accurate and computationally efficient.

2.3.1.b Field work

Recently, the replacement of sensors by camera systems has been trialed in the field. In [16] a pattern is applied to a bridge and displacement readings from a pre-weighted truck performing passes is verified against LVDT readings and a predictive model developed in the lab. The results are not used as a method of detecting damage, merely to verify the accuracy of their predictive model. The accuracy obtained is reasonable (+0.15mm compared to the model), but the monitoring difficulties of environmental change/vibration of camera in the field are briefly discussed without any proposed solution.

Further work was carried out by Lee et al in [25] where the feasibility of a vision system for use on a long-span bridge was discussed. A key point raised from this paper is the treatment of the angular orientation of the camera with respect to the target; this issue is particularly relevant to bridges as the ideal scenario of the camera being placed 90° from the target is not always possible. This paper also discusses the possibility of monitoring the bridge during the night hours using artificial light to ensure constant light levels, a promising solution to the problem of differing environmental light. However it must be considered that traffic patterns could vary largely at night and so the data collected may not be representative of the daily conditions.

2.3.2 Feature Based

The application of targets to the bridge is a limiting factor in camera based monitoring due to the location of some monitoring points on long-span or otherwise remote locations. To address this issue, the next generation in SHM for bridges involves the use of non-target “contactless” approaches that only use the natural patterns created by concrete pour, bolts or environmental effects as points of reference to extract and track features.

2.3.2.a Laboratory Work

Many of the studies detailed previously used either a correlation approach, an optical flow or a combination of both these approaches. The study carried out in [26] proposed a new method of tracking displacement where images are subject to down-sampling and intensity interpolation in order to generate moiré fringes, which are then put through a phase-shifting process to determine displacement distribution. It is stated in this paper that sub-millimetre deflection readings can be obtained, but they do not publish a full set of results for verification. Running times for the algorithm have also not been detailed in the paper so it cannot be compared to that of existing algorithms, but it is declared in that the comparison is favourable.

The lab study carried out in [27] did not use intensity of pixels for determining features to be tracked, but rather a process known as orientation-code matching (OCM), which performs well against effects such as lighting changes [28] and incorporates a subpixel method to reduce measurement errors (maximum error vs LVDT of feature tracking was 4.82% at a monitoring distance of 9.15m). This error percentage increases in proportion to monitoring distance, which is a common occurrence in field work with computer vision in SHM. Another system which catered for lighting changes was detailed in [29]. This method proposed the usage of a normalized correlation metric in order to account for changing lighting conditions, the results are shown in figure 5. This system had issues with very small displacements, which would make it unsuitable for use on short span bridges.

![Figure 5 Results of Correlation metric applied to subsets of an image with changes of light](image)

2.3.2.b Field Work

A successful study into measuring displacements of a bridge was carried out by Feng et al in [31]. Their method involved determining an approximate value for displacement using a correlation/template matching approach to find the correlation peak of the sub-image. This sub-image is then upsampled in order to find sub-pixel displacement of the target area. This method worked well compared to LVDT readings, but two challenges which were not overcome were the effect of camera vibration and a heat haze which occurred when the air was non-uniformly heated by ambient temperatures. The heat haze problem is an issue in locations with high average temperatures, therefore it has not been mentioned in other publications but would definitely need to be overcome if computer vision is to be a viable solution worldwide.

A multi-point optical flow system was trialled by Kim et al in [32] with satisfactory results (less than 2% error due to displacement responses), but they had issues with image noise from smoke affecting accuracy of their captured images and low pixel resolution of the camera used in the test was also a factor in obtaining less than perfect accuracy in their readings.
The use of a higher resolution camera could make their solution more practical for widespread use in the field.

An advanced method for detecting displacements and vibrations in a structure is outlined in [33], as shown in Figure 6. A new framework for detecting and tracking key points (FREAK) is developed and integrated with a calibration method for converting pixel-based units to engineering units. This and similar methods differ from other optical flow or correlation-based methods on the use of a small set of sparse key-points rather than a dense correspondence of pixels.

Finally, a system which combines the traffic detection methods detailed above with camera-based displacement detection is [2]. This system measures axle spacing and position of vehicles crossing the bridge and synchronises their position with deflection readings taken by a second camera placed at a perpendicular angle to the bridge soffit. The videos are synchronised by the use of two LEDs which simultaneously flash through usage of an interval timer. The authors catered for the issue of camera vibration by creating a custom tripod system to smooth out vibrations and provide a sturdy base for the camera.

![Camera setup from [2]](image)

Figure 7 Camera setup from [2]

The accuracy of this system did not extend to sub-millimetre levels, primarily due to the resolution of the images provided by the camera. With a superior camera or usage of post-processing algorithms, it is possible that sub-pixel accuracy could be obtained.

CONCLUSIONS

The area of Computer vision in SHM is still a relatively new one, but the methods detailed above represent significant progress in recent times. It is believed by the authors that the methods proposed in [26], [27], [30], [33] present viable solutions for detecting bridge displacement or strains if adopted properly. If some/all of these methods could be combined with a calibration process similar to [34], a method for compensating for camera movement detailed in [35] and time-synchronization of cameras using a method akin to the one detailed in [36] it is possible that an efficient and accurate solution could be developed and put into practice in the near future and the authors are currently working on a solution that incorporates these elements.

ACKNOWLEDGMENTS

The Department for Employment and Learning Northern Ireland, Invest Northern Ireland, the National Science Foundation, and Science Foundation Ireland are gratefully acknowledged for their financial support.

REFERENCES


