Marker-based Framework for Structural Health Monitoring of Civil Infrastructure

Milán Magdics[1,2,a], Rubén Jesús Garcia [1,b] and Mateu Sbert [1,c]

[1]Institute of Informatics and Applications, University of Girona, Girona, Spain

[2]Department of Control Engineering and Information Technology,

Budapest University of Technology and Economics, Budapest, Hungary

[a]magdics@imae.udg.edu, [b]rgarcia@imae.udg.edu, [c]mateu@imae.udg.edu

Keywords: Health monitoring, marker tracking, motion detection, civil infrastructure.

Abstract. As no civil infrastructure can escape aging and deterioration, health monitoring can prevent and report serious structural damage. With the rapid evolution of computer vision algorithms, optical-based systems become an increasingly feasible option for automatic monitoring. This paper proposes a cheap and flexible standalone system based on marker tracking to report deflection of structural elements of civil infrastructure. A single marker is placed on tracked objects, which allows unambiguous identification of objects and accurate movement tracking. Accuracy of the system is discussed by presenting a theoretical analysis of the translation error. Additionally, as a proof of concept we extend our work with a low-cost laboratory test implementation.

Introduction

Civil infrastructure must withstand infrequent but high forces coming from different sources and thus their deterioration is inevitable. To avoid serious damage, tracking the structural integrity and determining safety levels (*health monitoring*) becomes vital. An important component of structural health monitoring is following the trajectory of structural elements over time, which can indicate anomalous behavior. Despite its importance, measurement data is still often collected in the traditional way as on-site expert inspection, greatly limiting frequency and increasing costs. Therefore, there is a high desire for low-cost sensor systems providing automatic, remote data collection and evaluation. 3D object tracking sensors can be categorized into *inertial*, *magnetic*, *acoustic* and *optical* systems. Inertial systems suffer from the integration drift problem [1] making accurate 3D position estimation difficult, while magnetic and acoustic systems are less economic and have limited spatial and temporal precision. In contrast, computer vision algorithms have reached a level which allows high performance object tracking, turning *computer vision based optical systems* into a cheap, compact and robust option. Furthermore, since computer vision has numerous applications, there is an increasingly wide range of available hardware directly supporting many protocols such as of network communications and offering a number of performance and price tradeoffs.

In this paper we propose a camera-based sensor system for structural health monitoring. Markers are placed on objects of interest, which are tracked by cameras to obtain the 3D movement over time. The markers used by our method were successfully used in augmented reality systems [2], showing that a single camera per object is sufficient, and a camera may track multiple, uniquely identified objects at a time with real-time performance. Our paper is organized as follows. First, we review the background of vision-based systems designed for health monitoring. Next, we introduce our system, followed by error analysis. Finally, results and conclusions are presented.

Background

Most computer vision based tracking systems basically consist of three stages [3,4,5]. First, 3D information is obtained, usually as a point cloud, using stereoscopic vision theory, video frames or a-priori information. Then, 3D objects are identified with an object recognition method. Finally, object position is matched against stored values of older timeframes to obtain translation or rotation information. Although this pipeline is very flexible, the first two stages introduce the uncertainty and error of 3D reconstruction and object recognition.

This uncertainty and error can be reduced by the use of *fiducial markers* [6]. These are planar patterns that can be identified and located very reliably and accurately on an image or video stream. In fact, the assumption of a planar object allows the reduction of both detection and recognition tasks from 3D to 2D and thus, makes the use of 3D reconstruction (e.g. requiring multiple cameras) unnecessary. Additionally, tracking a well-characterized pattern instead of an arbitrary object increases robustness further. Fiducial markers are most commonly used in augmented reality systems [2,5] and provide high precision even at real-time rates. General-purpose marker tracking systems available on the market offer various time and spatial resolutions [7,8,9]. The idea of using markers was also applied for structural health monitoring [10], however, this method also relies on image registration techniques.

In this paper, we propose a standalone marker-based health monitoring system requiring neither image registration nor 3D reconstruction. As a proof of concept we show that with even low-cost hardware, high tracking accuracy can be achieved.

The Proposed Framework

Overview. Fig. 1 shows the overview of our proposed system. Cameras are placed at the monitored infrastructure, collecting data regularly and/or on demand by making photos or videos of the object of interest. Depending on the availability of network connection, raw measurement data consisting of the images or videos are transferred via network or stored locally and transferred manually. Note that without network availability no systems could provide remote access, regardless the sensor type. Captured data is processed independently of data collection and can be either in real-time or offline. 3D position, pose and movement of the tracked object w.r.t. the camera is obtained by locating the marker attached to the object on the photos. Details of marker detection are explained in the following.

Marker Tracking. As stated in the previous section, markers are planar patterns that can be easily located on an image using computer vision algorithms. The location of the marker on the image in pixel coordinates unambiguously identifies the relative position of the marker w.r.t. camera in the 3D space and assuming that the size of the marker is known in some metrics, its position is given in absolute measures. Similarly, from the distortion of the pattern we can extract the relative orientation. Accurate results require lens distortion to be taken into account [6,10], therefore, our system supports camera calibration. Note that this step has to be performed only once for every camera type.

Object Recognition and Tracking Multiple Objects. Since tracked objects correspond to markers in our framework, the complex task of object recognition is simplified to marker identification, which greatly increases robustness and flexibility. Additionally, we designed our marker detection layer to support a large number of unique markers. Thus, attaching different markers to distinct objects allows unambiguous object identification without further processing, even when a single camera follows multiple targets.

Markers. Choosing the proper marker is crucial for performance. Based on former studies made by Rice et al. and Köhler et al. [11,12] and evaluation criteria of Fiala [6], we have chosen square markers encoding binary codes. Firstly, square markers have higher information payload than circular ones of the same size [11], which is important for unique identification of markers. Secondly, binary codes undoubtedly have the greatest performance in terms of false positive detection and intermarker confusion (i.e. reporting the wrong marker identifier) [6]. Markers are extracted from the images using the Aruco library [13] developed at the University of Córdoba in Spain and freely available with BSD license, which applies adaptive thresholding for increased robustness for changing lighting conditions and supports up to 1024 different markers, while offering real-time performance. It was successfully tested in augmented reality applications [2].

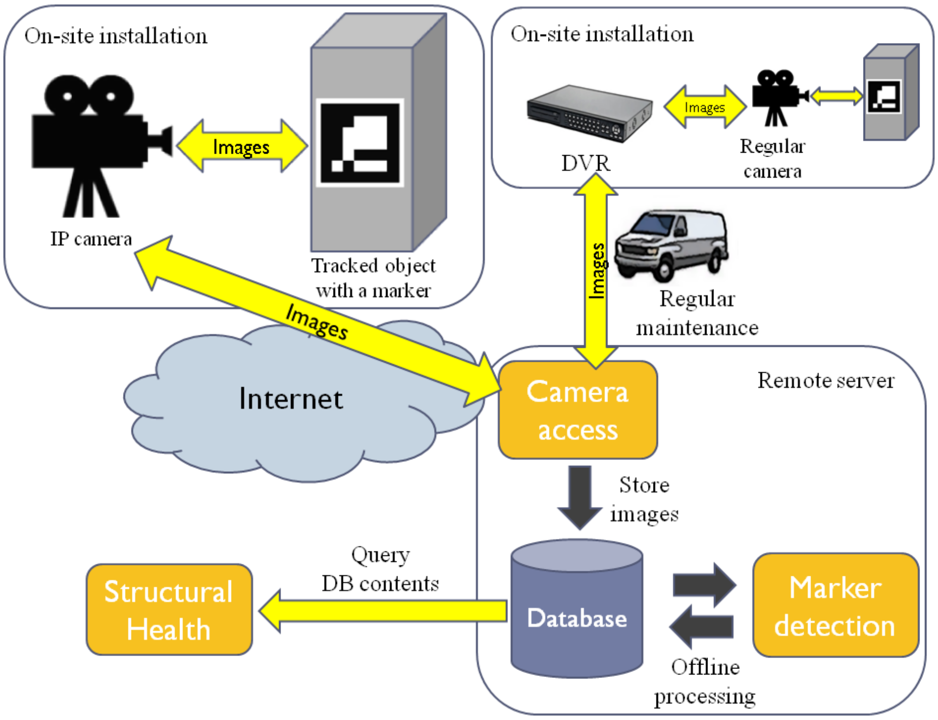

Figure 1: System overview

Limitations. Our system has the common limitations of vision-based approaches: the performance is affected by the viewing conditions such as bad weather or occlusion. Proper camera placement can deal with occlusion issues by guaranteeing that the tracked object is always visible to the camera and infrared-capability of cameras requires low additional cost, while allowing measurements in low lighting conditions. Additionally, our system interprets object movement as the movement of a planar marker, which has two main drawbacks. First, compound objects with higher degree of freedom require a marker for each joint. In addition, planar markers cannot be painted on objects with bent surfaces. However, this latter issue can be handled by attaching a planar plate containing the marker on the surface of the object. Finally, precision, optimal distance and the size of markers depend on the camera resolution; details are presented in the next section.

Error Analysis

The movement of the marker will be detected by the pixel position of its corners. Since the camera has a finite resolution, an error is introduced: the distance between the exact marker position and the center of the pixel reported by the camera. Assuming properly removed lens distortion, under ideal conditions this error is at most half a pixel in magnitude; however since errors in all four corners may accumulate, in practice we may obtain a slightly higher error. With black and white markers and cameras which can capture many grey levels, objects can be tracked to subpixel accuracy. Since this depends on the exact camera type, here we present a general theoretical upper bound.

The size of a pixel in world space depends on the distance to the object and the resolution of the camera (Fig. 2 left). In the figure, a represents the distance to the marker, $2b$ is the height of the image, and c is the distance to the border of the image. θ is the vertical field of view (*fov*) of the camera. The situation is identical for the horizontal *fov*. The equations relating a, b, c and θ are:

$$a = c \cos(\theta/2); \qquad b = c \sin(\theta/2) = a \tan(\theta/2) \qquad (1)$$

Fov	Resolution	Distance (1cm error)	Error (1m distance)
60º	640x480	8.31 m	1.20 mm
60º	2592x1944	33.67 m	0.30 mm
5.5º	640x480	99.93 m	100 µm
5.5º	2592x1944	404.72 m	25 µm

Table 1: Maximum distance for 1 cm error and maximum error from 1m distance

Fov	Resolution	Distance		Marker Size	
		10% fov	50% fov	10% fov	50% fov
60º	640x480	96 cm	4.8 m	11.08 cm	2.77 m
60º	2592x1944	3.89 m	19.44 m	44.91 cm	11.22 m
5.5º	640x480	96 cm	4.8 m	9.22 mm	23 cm
5.5º	2592x1944	3.89 m	19.44 m	3.73 cm	93 cm

Table 2: Maximum distance for 1 cm error, maker covering 10% and 50% of camera fov

Fov	Resolution	Error		Marker Size	
		10% fov	50% fov	10% fov	50% fov
60º	640x480	10 mm	2 mm	11.54 cm	58 cm
60º	2592x1944	2 mm	514 µm	11.54 cm	58 cm
5.5º	640x480	10 mm	2 mm	9.60 mm	4.8 cm
5.5º	2592x1944	2 mm	514 µm	9.60 mm	4.8 cm

Table 3: Maximum error for 1m distance, marker covering 10% and 50 % of camera fov

The length of a pixel side in world coordinates is $2b/py$, where py is the vertical resolution. Since maximum horizontal and vertical error is proportional to this length, the relationship between error and distance is linear, and between error and resolution, inversely linear. The maximum error for a distance a for a fov θ and vertical resolution py is:

$$E = a \tan(\theta/2)/\, py; \qquad a = E\, py / \tan(\theta/2) \tag{2}$$

The horizontal and vertical errors are relatively independent of size of the marker; as long as the marker is large enough (more than 30 or 40 pixels), corners can be located with relative ease. Larger sizes, however, are more robust to eventual stains which hinder recognition. Table 1 contains the maximum distance for typical camera resolutions for a fixed 1 cm error, along with the error expected when the marker is located at 1 m from the camera. We can see that the error of the recognition is small enough that the quality of the marker printing becomes the bottleneck. Fig. 2 (right) shows how errors in recognition (Δb) relate to errors in depth (Δa) by the formulas:

$$\Delta a = a\, \Delta b / b; \qquad a = b\, \Delta a / \Delta b \tag{3}$$

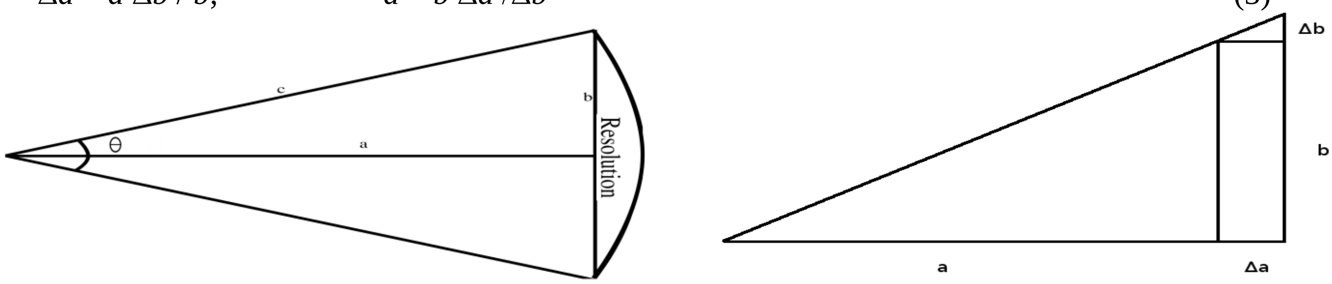

Figure 2: Left: Field of view of a camera (1D vertical slice). Right: Error in depth

In this case, the error in depth is inversely proportional to marker size and directly proportional to distance. Markers should therefore be as large as possible, while always completely visible by the camera. Table 2 shows the maximum distance for a 1 cm error and the needed marker size to cover the given percentages of the camera. Table 3 shows the depth error for a fixed distance of 1 m. We can see that the errors are larger than with horizontal and vertical movements, so if this is an issue, two perpendicular markers may be used. The use of large fovs makes the error in depth increase, so there is a tradeoff between precision and the number of simultaneously tracked markers.

Test implementation

We tested our concept using a low-cost (125$ in december 2012) IP-camera with infrared vision and wireless connection from Apexis [14], having 640×480 resolution and 60º horizontal *fov*. Images captured by the camera were transferred automatically via network connection and inserted into a remote database. Marker localization on the images was performed on the remote server using the Aruco library and the obtained 3D pose and position were inserted into the database. We also built a web-based interface where the user can dynamically query the pose, position or movement of the marker filtered by capture time, using a web-browser. Fig. 4 shows a typical graphical result of the query of movement in a given time period, we used basic augmented reality methods to depict the marker movement. We show a measurement taken at night using the infrared mode of the camera in Fig. 5 (right). Image quality is not considerably affected comparing to the images made at daylight.

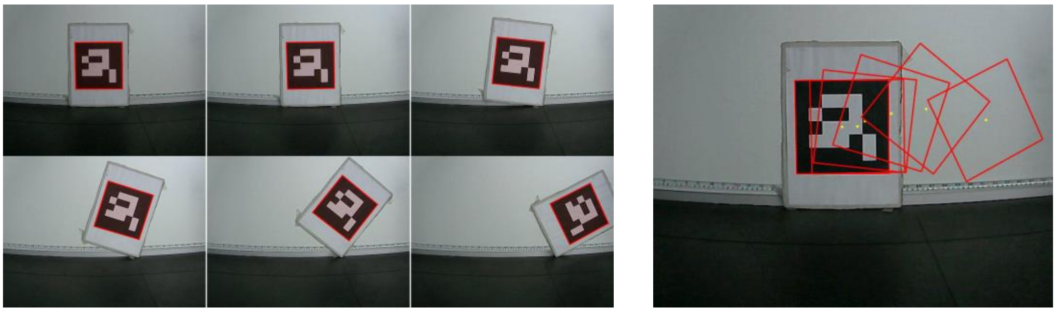

Figure 4: Visualization of captured data (left) using augmented reality (right)

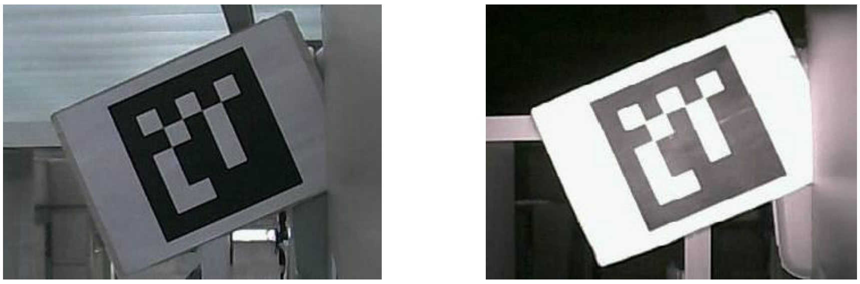

Figure 5: Images taken at daylight (left) and at night (right)

We also measured precision by setting the distance to the camera to 1 m and using a marker of size 17.7^2 cm^2. As Table 4 shows, even with this relatively small marker size and low-cost hardware, less than 1 cm precision could be achieved. Due to the use of fiducial markers, we have not met any case so far when the system failed to locate or identify an object visible to the camera.

Translation (cm)	**Error (cm, horizontal)**	**Error (cm, vertical)**	**Error (cm, depth)**
10	0.3	0.4	0.3
20	0.4	0.3	0.5
30	0.4	0.5	0.8
40	0.3	0.3	0.9

Table 4: Average measurement errors of the system along different axis

Summary

The paper presented a low-cost, flexible system based on marker tracking, for monitoring structural health of civil infrastructure. Both theoretical studies and laboratory tests showed that even with a cheap camera, we are able to track moving objects with high precision remotely and automatically, and regardless of the lighting conditions. As future work we will install the system at the bridge over river Arousa in Spain to track the moving parts of the elastomeric support.

Acknowledgements

This work has been supported by the research projects coded TIN2010-21089-C03-01, IPT-2011-1516-370000 and IPT-2011-0885-430000 (Spanish Commission for Science and Technology), and by grant 2009SGR643 (Catalan Government).

References

[1] C Bouten, K Koekkoek, M Verduin, R Kodde, and J Janssen. *A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity*. IEEE Transactions on Biomedical Engineering, 44, 3, pp 136 –147, DOI=10.1109/10.554760 (2010)

[2] *OpenSpace3D*, http://www.openspace3d.com/ [March 20 2013]

[3] I Brilakis, F Dai, and S-C Radopoulou. *Achievements and challenges in recognizing and reconstructing civil infrastructure*. In Proceedings of the 15th international conference on Theoretical Foundations of Computer Vision. Springer-Verlag, Berlin, Heidelberg, pp 151-176 (2011).

[4] J Teizer, F Bosche, C.H. Caldas, and C.T. Haas, 2006. *Real-time Spatial Detection and Tracking of Resources in a Construction Environment*, International Conference on Computing and Decision Making in Civil and Building Engineering, Montreal, Canada, pp 494-502.

[5] H Fathi,, I. Brilakis, and P Vela. 2011. *Automated 3D Structure Inference of Civil Infrastructure Using a Stereo Camera Set*. Computing in Civil Engineering (2011): pp 118-125.

[6] M Fiala. *Designing Highly Reliable Fiducial Markers*. IEEE Trans. Pattern Anal. Mach. Intell. 32, 7, pp 1317-1324 (2010).

[7] GOM, http://www.gom.com/metrology-systems/3d-motion-analysis.html [March 20 2013]

[8] *Correlated Solutions*, http://www.correlatedsolutions.com/ [March 20 2013]

[9] *Limess*, http://www.limess.com/ [March 20 2013]

[10] P Kohut, K Holak, T Uhl. *Prototype of the vision system for deflection measurements*, Diagnostyka / Polskie Towarzystwo Diagnostyki Technicznej, 4, pp 3-11 (2011)

[11] A C. Rice, A R. Beresford, and R K. Harle. 2006. *Cantag: an open source software toolkit for designing and deploying marker-based vision systems*. In Proceedings of the Fourth Annual IEEE International Conference on Pervasive Computing and Communications pp 12-21 (2006).

[12] J Köhler, A Pagani, D Stricker. *Detection and Identification Techniques for Markers Used in Computer Vision*. In Visualization of Large and Unstructured Data Sets - Applications in Geospatial Planning, Modeling and Engineering. Volume 19 of OASICS, pp 36-44 (2010)

[13] *ArUco: a minimal library for Augmented Reality applications based on OpenCv*, 2012. Available: http://www.uco.es/investiga/grupos/ava/node/26 [March 20 2013].

[14] Apexis, *APM-J0233-WS-IRC IR IP*, http://www.apexis.com.cn/en/productsdetails_187.html [March 20 2013]