Measurement 83 (2016) 44–56

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Measuring human-induced vibrations of civil engineering structures via vision-based motion tracking

Feng Zheng [a], Ling Shao [b], Vitomir Racic [c], James Brownjohn [d,*]

[a] Department of Electronic and Electrical Engineering, University of Sheffield, United Kingdom
[b] Department of Computer Science and Digital Technologies, Northumbria University, United Kingdom
[c] Department of Civil and Structural Engineering, University of Sheffield, United Kingdom
[d] College of Engineering, Mathematics and Physical Sciences, University of Exeter, United Kingdom

ARTICLE INFO

Article history:
Received 12 July 2015
Accepted 7 January 2016
Available online 27 January 2016

Keywords:
Object tracking
Human induced vibration
Ensemble learning
Online learning

ABSTRACT

We present a novel framework for measuring the body motion of multiple individuals in a group or crowd via a vision-based tracking algorithm, thus to enable studies of human-induced vibrations of civil engineering structures, such as floors and grandstands. To overcome the difficulties typically observed in this scenario, such as illumination change and object deformation, an online ensemble learning algorithm, which is adaptive to the non-stationary environment, is adopted. Incorporated with an easily carried and installed hardware, the system can capture the characteristics of displacements or accelerations for multiple individuals in a group of various sizes and in a real-world setting. To demonstrate the efficacy of the proposed system, measured displacements and calculated accelerations are compared to the simultaneous measurements obtained by two widely used motion tracking systems. Extensive experiments illustrate that the proposed system achieves equivalent performance as popular wireless inertial sensors and a marker-based optical system, but without limitations commonly associated with such traditional systems. The comparable experiments can also be used to guide the application of our proposed system.

1. Introduction

In civil engineering dynamics, there have been many problems related to vibrations of floors[1], footbridges [2], assembly structures (grandstands, spectator galleries, etc.), due to crowds or groups of human occupants walking, running, dancing and jumping. For example, the London Millennium Footbridge [3] opened on 10 June 2000 was closed almost immediately for nearly two years because of the unexpected movements occurred when a large crowd of pedestrians crossed the bridge. Just a year before, a similar vibration serviceability problem was observed on the newly built Solferino footbridge in Paris [4]. Also, in 2000 during a concert event, the cantilevers of the Cardiff Millennium stadium experienced excessive vibration amplitudes caused by people jumping so that the concert had to be stopped. Moreover, the modern structures have become more flexible and prone to human induced vibrations. Consequently, extensive research into the human-structure dynamic interaction phenomenon was launched. The research results were incorporated into two key design guidelines relevant to crowd loading of footbridges (France) [5] and grandstands (UK) [6] for civil engineers.

Human motion and the induced force have drawn much attentions of researchers from different areas for many years [7,8]. Several reliable force models for active individuals [9,10] are available. However, there is a lack of models describing dynamic loading of structures due to groups or

* Corresponding author. Tel.: +44 1392 723698.
E-mail address: J.Brownjohn@exeter.ac.uk (J. Brownjohn).

http://dx.doi.org/10.1016/j.measurement.2016.01.015
0263-2241/© 2016 The Authors. Published by Elsevier Ltd.

crowds of people. How to use a model of individual loading to generate load models of multiple people still remains a challenge. It is unknown how people interact in groups of various sizes and what the level of synchronisation is between individuals under different circumstances, such as various visual and tactile stimuli. The main difficulty is to collect simultaneously the body motion data for multiple people in groups or crowds on real structures. Therefore, the key aim of this paper is to develop a new vision-based system which will enable robust collection of fundamental body data.

Although vision-based methods for human motion analysis have caught much attention of researchers and practitioners involved in gaming, security and other related applications, the robustness of the systems is far from ideal. The key reasons for this are difficulties in setting up tracking targets and the environmental conditions. At present, these challenges can be partially solved using the robust object descriptors and adaptive appearance models [11–14]. These methods can work well on data sets recorded under controlled conditions. However, due to the unpredictability of environmental changes, most existing methods cannot be applied directly in a real-world situation. In addition, they are usually unable to cope with the challenges appearing in a video sequence simultaneously. Thus, in this paper, a real-time system which contains a vision-based multiple object tracking algorithm [15] and a set of carefully selected hardware components is constructed to deal with the weaknesses of current systems.

The remainder of this paper is organised as follows. The background of measuring dynamic load and the contributions are given in Section 2. In Section 3, we describe the framework of the adopted object tracking algorithm. How to align the signals generated by different sensors is detailed in Section 4. Extensive experiments in comparison to classical sensors are presented in Section 5. We conclude this paper and discuss future work in Section 6.

2. Background and contributions

2.1. Measuring dynamic load

Several researchers tried to adopt different systems to monitor activities of individual people and investigate the synchronisation phenomenon of groups or crowds. Early attempts to measure human induced loading [16–18] were based on direct force identification using force plates and instrumented treadmills. However, their size places restrictions on studies of loading induced by multiple people [18]. An alternative approach is to measure the loading indirectly. According to [19], if the accelerations of body motion are known or measured, the ground reaction force (GRF) $\mathcal{F}$ [8] can be calculated indirectly using the basic principles of Newtonian mechanics, i.e., force is equal to mass times acceleration. Therefore, using the acceleration and mass of the individual, the GRF generated by a crowd can be computed by

$$\mathcal{F} = \sum_i m_i a_i - g \sum_i m_i, \qquad (1)$$

where g is the static acceleration due to gravity, m_i is the body mass of the ith test subject and a_i is the dynamic acceleration due to body motion. Generally, the body mass is supposed to be known, while acceleration of the body needs to be experimentally measured or estimated.

Experimental characterisation of the body motion is possible using optical marker-based motion tracking [10], wireless inertial sensors [20], video-based monitoring [21] or multichannel interacting model [22]. In [10], the accelerations of body segments were measured by tracking optical markers (Codamotion) stuck to the surface of the human body, and then used to generate force signals. However, due to interaction with daylight and the limitation of the number of markers, marker-based optical tracking systems are usually constrained to artificial laboratory environments. Alternative wireless inertial sensors [23] can be used in outdoor environments but are expensive and typically suffer from synchronising individual units in a wireless network. Moreover, the number of units within a wireless network is limited, which in turn restricts the number of monitored individuals within a crowd.

To overcome the limitations of conventional motion tracking sensors, a vision-based method can be considered. Video data captured by a camera (CCD or CMOS sensor) are becoming increasingly discussed as an innovative tool for measuring the motion of humans, structures or animals. Combined with the right video analysis algorithms used to detect the motion trajectory in the image space, vision-based methods have the potential to save time and money over conventional sensors. Research in vision-based motion tracking methods is topical [24], with a wide range of applications, such as surveillance [25], augmented reality, robotics and human–computer interaction. Compared with the conventional systems, vision-based methods for measuring human motion have the following advantages: (1) It is possible to measure people in outdoor environments rather than laboratory setting. This is because the system is less sensitive to illumination changes than marker-based sensors. (2) The number of tracking individuals is not limited. Due to the entire scenario being captured and no special tracking target (such as a Codamotion marker) being predefined, it is easy to track much more targets in the view at all times. (3) People are not aware of being recorded. No markers or inertial sensors need to be worn by participants. This will save time for preparation and lead to more natural captured body movement of test subjects, although there are ethical considerations to be addressed. (4) It is a cheap, remote and long-term monitoring system. The available commercial marker-based or wireless inertial systems are typically expensive and require external power.

Some research on digital image correlation (DIC) [26] methods to track the movement of crowds does exist [21]. However, the suggested methods are built based on a strong assumption that the motion of each individual in a crowd is similar to the motion of surrounding people, i.e. everybody moves in the same direction. In reality, even when test subjects follow the same music, directions of their motion can be opposite. Moreover, each test subject has their own motion style or pattern, such as waving