Advanced Engineering Informatics 29 (2015) 867–877

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Construction equipment activity recognition for simulation input modeling using mobile sensors and machine learning classifiers

Reza Akhavian [1], Amir H. Behzadan *

Department of Civil, Environmental, and Construction Engineering, University of Central Florida, 4000 Central Florida Blvd, Orlando, FL 32816-2450, USA

ARTICLE INFO

Article history:
Received 22 May 2014
Received in revised form 28 February 2015
Accepted 10 March 2015
Available online 1 April 2015 xxxx

Keywords:
Construction equipment action recognition
Smartphone sensors
Accelerometer
Data-driven simulation
Supervised machine learning
Big data analytics

ABSTRACT

Although activity recognition is an emerging general area of research in computer science, its potential in construction engineering and management (CEM) domain has not yet been fully investigated. Due to the complex and dynamic nature of many construction and infrastructure projects, the ability to detect and classify key activities performed in the field by various equipment and human crew can improve the quality and reliability of project decision-making and control. In particular to simulation modeling, process-level knowledge obtained as a result of activity recognition can help verify and update the input parameters of simulation models. Such input parameters include but are not limited to activity durations and precedence, resource flows, and site layout. The goal of this research is to investigate the prospect of using built-in smartphone sensors as ubiquitous multi-modal data collection and transmission nodes in order to detect detailed construction equipment activities which can ultimately contribute to the process of simulation input modeling. A case study of front-end loader activity recognition is presented to describe the methodology for action recognition and evaluate the performance of the developed system. In the designed methodology, certain key features are extracted from the collected data using accelerometer and gyroscope sensors, and a subset of the extracted features is used to train supervised machine learning classifiers. In doing so, several important technical details such as selection of discriminating features to extract, sensitivity analysis of data segmentation window size, and choice of the classifier to be trained are investigated. It is shown that the choice of the level of detail (LoD) in describing equipment actions (classes) is an important factor with major impact on the classification performance. Results also indicate that although decreasing the number of classes generally improves the classification output, considering other factors such as actions to be combined as a single activity, methodologies to extract knowledge from classified activities, computational efficiency, and end use of the classification process may as well influence one's decision in selecting an optimal LoD in describing equipment activities (classes).

1. Introduction

According to the United States Department of Commerce, construction and infrastructure projects comprise a trillion dollar industry with a continuous annual increase in pace [1]. Although there have been many efforts to increase the productivity of construction and infrastructure projects in recent years, the industry is still suffering from low productivity growth [2–5]. There are several key factors that can influence productivity in construction and infrastructure industry, including the uncertain, dynamic, and transient nature of most construction projects. During the pre-construction phase, and due to the lack of data, it is customary to make engineering assumptions about the availability of tools, resources, information, materials, equipment, construction methods, and flow of activities [3]. Although a level of versatility is often considered for such assumptions, the dynamics involved in most projects as they enter the construction phase, makes it necessary to revise initial project plans and decisions, which may in turn result in potential delays and rework [3,6,7].

As infrastructure projects increasingly become larger and more complex in nature, traditional manual quantitative analysis methods mostly fail to effectively and accurately capture key project productivity performance indicators [8]. Therefore, computer simulation models capable of modeling uncertainties and stochastic events have become more relevant to the decision-making process especially when real world evaluation is difficult, expensive, or time-consuming. To achieve the best results, a simulation model

* Corresponding author. Tel.: +1 407 823 2480; fax: +1 407 823 3315.
E-mail addresses: reza@knights.ucf.edu (R. Akhavian), amir.behzadan@ucf.edu (A.H. Behzadan).
[1] Tel.: +1 407 823 2480; fax: +1 407 823 3315.

http://dx.doi.org/10.1016/j.aei.2015.03.001
1474-0346/

should accurately represent the real engineering system through the integration of data that describe the real world resources and processes [5]. It is imperative that manual data collection techniques such as direct observations and field surveys are not efficient ways to obtain large volumes of high quality data in a timely manner [9]. Thus, automated data collection using sensors, vision-based systems, and laser scanners have gained credibility in quantitative analysis of construction activities.

Process-level data collection deals with data from construction resources (i.e. equipment, labor, material). Detailed resource activity recognition using these data has a great potential in discovering knowledge about activity durations and precedence, resource flows, and site layout. Among different types of process-level knowledge, activity duration is undoubtedly one of the most influential factors as there is always an uncertainty component to duration values that can propagate in time and/or space and consequently affect the outcome of the decision-making process [10,11]. Therefore, a systematic approach for action recognition that leads to precise activity duration extraction can boost the accuracy of decision-making tools such as simulation models. It has been widely discussed that inaccurate and unrealistic simulation models with static input data built upon expert judgments, secondary data (from past projects), and assumptions made on the basis of available resources and information during the pre-construction phase are major impediments that prohibit the widespread use of simulation models within the construction industry [8,12].

In an effort to address this challenge, the authors have been investigating the applicability of data-driven simulation for construction operations analysis [13,14]. In the authors' previous studies, a wireless network of sensors attached to different articulated parts of construction equipment was designed and implemented [13,14]. However, due to technical and practical difficulties associated with mounting sensors on construction equipment body parts (e.g. attachment and detachment of different sensors for every data collection session, construction site dust and noise) and data storage issues, a more pervasive data collection scheme is used in this study. This paper presents the latest findings on a critical component of an ongoing research, a ubiquitous data sensing and analysis system that captures multi-modal process data from construction equipment using mobile sensor nodes, and employs data mining and process reasoning methods to transform raw data into meaningful knowledge that can be ultimately incorporated into data-driven simulation models. In this paper, first, a comprehensive literature review is conducted to help identify the gaps in knowledge and practice, and put the presented work within proper context. Next, the requirements and necessary level of detail (LoD) and resolution in activity recognition is discussed, and the designed methodology is described. Finally, the experimental results of the developed methodology are presented and further discussion about the results is provided.

2. Previous work

The framework presented in this research consists of (a) an activity recognition architecture using built-in smartphone accelerometer, gyroscope, and positional sensors that is used to (b) detect distinct activities performed by construction equipment for (c) construction simulation input modeling. Therefore, this section provides a comprehensive literature review in each of these three domains.

2.1. Action recognition using accelerometer and gyroscope data

A three-dimensional (3D) accelerometer is a sensor that returns values of acceleration, and a 3D gyroscope is a sensor that returns the angular velocity about x, y, and z axes [15]. The idea of recognizing activities using accelerometers have been around since the 1990s where researchers leveraged wearable devices to report instantaneous and sudden vibrations of human targets [16–18]. More recently, the use of gyroscope for the same purpose has also attracted the attention of researchers [15,19]. In particular, the adoption of such sensors in smartphones has facilitated the emergence of more context-aware applications.

Several fields including but not limited to computer sciences, healthcare, and sports have benefited from these Micro-Electro-Mechanical Systems (MEMS) inertial sensors [15,19–22]. For example, wireless accelerometers were used for the analysis of soccer players' movement patterns [23]. Using both accelerometer and gyroscope, Li et al. [15] presented a fall detection algorithm capable of detecting static postures and dynamic transitions. However, they stated that more environmental and physiological information is needed to distinguish between more complex actions. In a similar study, identification of physical human activities using mobile accelerometer sensors was evaluated [24]. Motoi et al. [25] proposed a human posture and walking monitoring system that works based on the speed of the ambulatory subjects.

Despite the prevalent use of such context-aware systems in non-engineering domains, research on their applications in engineering fields has been relatively limited. For instance, in a driving safety application, Johnson and Trivedi [26] used accelerometers to detect, recognize, and record driving styles. In an industrial setting, Lukowicz et al. [27] developed a system for segmenting and recognizing typical user gestures in a wood workshop using body-worn microphones and accelerometers. In a prototype experiment that was conducted in a laboratory setting, they simulated the assembly of a simple wooden object to recognize specifically-designed activities. As discussed in more detail in the next Subsection, construction jobsites have unique characteristics that may prohibit the wide application of such pervasive mobile data collection techniques. Challenges include but are not limited to the unstructured arrangement of resources (i.e. equipment, labor, material) that creates technical and practical problems for installing and calibrating sensors, as well as storage of non-structured or semi-structured data. Moreover, unexpected and intermittent events such as equipment breakdowns, adverse weather, and human crew motion irregularities can also add to the difficulty of interpreting sensory data collected from construction jobsites.

2.2. Construction resource action recognition

Object recognition and tracking has been a major research direction of several ongoing efforts in the field of computer vision [28–30]. Unlike computer vision where almost all such studies target human action recognition and pose analysis, researchers in construction engineering and management (CEM) domain have applied similar algorithms mostly for vision-based construction resource recognition and tracking. For example, Brilakis et al. [31] proposed a framework for vision-based tracking of construction entities. Their methodology requires calibration of two cameras, recognition of construction resources and identification of the corresponding regions, matching the entities identified in different cameras, two-dimensional (2D) tracking of the matched entities, and finally calculation of 3D coordinates. This and similar vision-based approaches, although provide promising results for recognition and tracking of construction equipment, still require much computation in each one of the aforementioned steps. In another study, an image processing methodology was adopted for idle time quantification of hydraulic excavators [32]. The LoD of the framework, however, was limited to detection of only idle and busy states of a hydraulic excavator. For the purpose of learning and classification of labor and equipment actions, the concept