Mechanical Systems and Signal Processing 95 (2017) 363–379

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

A data fusion approach for track monitoring from multiple in-service trains

George Lederman [a,b,*], Siheng Chen [b], James H. Garrett [a], Jelena Kovačević [b,c], Hae Young Noh [a], Jacobo Bielak [a]

[a] Civil and Environmental Engineering, Carnegie Mellon University, USA
[b] Electrical and Computer Engineering, Carnegie Mellon University, USA
[c] Biomedical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

ARTICLE INFO

Article history:
Received 24 October 2016
Received in revised form 12 March 2017
Accepted 17 March 2017
Available online 6 April 2017

Keywords:
Data fusion
Signal processing
Adaptive Kalman filter
Vehicle-based inspection
Inertial sensing

ABSTRACT

We present a data fusion approach for enabling data-driven rail-infrastructure monitoring from multiple in-service trains. A number of researchers have proposed using vibration data collected from in-service trains as a low-cost method to monitor track geometry. The majority of this work has focused on developing novel features to extract information about the tracks from data produced by individual sensors on individual trains. We extend this work by presenting a technique to combine extracted features from multiple passes over the tracks from multiple sensors aboard multiple vehicles. There are a number of challenges in combining multiple data sources, like different relative position coordinates depending on the location of the sensor within the train. Furthermore, as the number of sensors increases, the likelihood that some will malfunction also increases. We use a two-step approach that first minimizes position offset errors through data alignment, then fuses the data with a novel adaptive Kalman filter that weights data according to its estimated reliability. We show the efficacy of this approach both through simulations and on a data-set collected from two instrumented trains operating over a one-year period. Combining data from numerous in-service trains allows for more continuous and more reliable data-driven monitoring than analyzing data from any one train alone; as the number of instrumented trains increases, the proposed fusion approach could facilitate track monitoring of entire rail-networks.

1. Introduction

Monitoring the health of a track network is important both to ensure the safety of its users, and to reduce maintenance costs by early detection of faults [1,2]. Track operators want inspection techniques that are low-cost and reliable, and that can detect faults soon after they occur. Currently, the most widely used methods of track inspection are visual inspection and inspection with track geometry cars [1]. The former is subjective, so can have low reliability, but tends to be employed frequently to identify major faults. The latter is more reliable, but given the high cost of operating dedicated track geometry cars, tends to be performed infrequently.

* Corresponding author at: Civil and Environmental Engineering, Carnegie Mellon University, USA.
E-mail addresses: lederman@cmu.edu (G. Lederman), sihengc@andrew.cmu.edu (S. Chen), garrett@cmu.edu (J.H. Garrett), jelenak@cmu.edu (J. Kovačević), noh@cmu.edu (H.Y. Noh), jbielak@cmu.edu (J. Bielak).

http://dx.doi.org/10.1016/j.ymssp.2017.03.023
0888-3270/

A number of researchers have proposed inspecting track networks using sensors aboard in-service trains [3–17]. Such a technique could be a low-cost method for providing near continuous inspection, particularly if enough in-service trains were instrumented. Most researchers have studied how to increase inspection reliability by developing new features that could be extracted from the raw data of individual sensors, such as wavelets [3,4,18,19], the Short Time Fourier Transform [7], signal standard deviation [12], and signal energy [11]. Others have proposed techniques to estimate the track profile [5,8,14,16], where the track profile itself serves as a feature of track condition. Rather than propose a new feature, here we introduce a novel data fusion approach that can take as input the features currently used in the monitoring community. As will be shown, our fusion approach not only enables more continuous monitoring by leveraging data from multiple trains, but also, by combining multiple sensors, increases the overall reliability of inspection from in-service trains.

Before describing our proposed method, we briefly discuss prior work on data fusion in the vehicle-based infrastructure monitoring space. To do so in a structured manner, we introduce the most prevalent data fusion model, the JDL Model (Joint Directors of Laboratories of the US Department of Defense), which has been developed and refined for defense applications since the mid-1980s [21–25]. In this model, data fusion is categorized into distinct "levels" within a data-processing pipeline. Level 0 Fusion is "Sub-object Data Association and Estimation," Level 1 Fusion is "Object Refinement," and Level 2 Fusion is "Situation Refinement." In our proposed data processing pipeline, shown in Fig. 1, Level 0 Fusion means combining raw data from different sensors, for example, fusing position data with data about the track condition. Level 1 Fusion means combining object level data, in this case, features extracted from multiple passes over the same track, where the track section is the "object" of interest. Level 2 Fusion means combining situation level information, in this case, the detection outcome from anomaly detection. In our proposed pipeline, all the data are fused at Level 1, so no Level 2 Fusion is required.

One of the earliest studies in train-based track monitoring was conducted by Bocciolone et al. [3]. They developed a system to detect track corrugations in the Milan Metro by looking at the wavelet transform of data acquired from accelerometers mounted on the axle of a passing train. In their study, they performed Level 0 Fusion by combining accelerometer and position data; they then analyzed the accelerometer data in the wavelength domain which is position dependent. Their study does not go beyond feature extraction; they do not combine data from different passes over the track, different accelerometers, or different trains.

The same year, Weston et al. [12] published a study on monitoring track alignment using sensors placed on a train in the Tyne and Wear Metro. They propose a technique to extract the standard deviation of the track profile, which is a common parameter of track geometry used in traditional track geometry car inspection. In terms of data fusion, Weston et al. go beyond the work by Bocciolone et al. by presenting results from multiple passes over the same track separated by several months. They note regions in which changes in the track appear to have occurred, but they do not present formal Level 1 Fusion approaches or anomaly detection techniques.

While higher-level data fusion approaches have not been studied for train-based infrastructure monitoring, they have been employed in related fields. Eriksson et al. [20] propose a method for pothole detection using vibration data collected from cell-phones in taxi-cabs. In their study, potholes are defined as events which exceed a predefined threshold; several detection events from individual vehicles must occur in the same vicinity before a pothole is detected by the overall system. Essentially, data from each individual pass is analyzed independently all the way through anomaly detection; if a pothole is detected, this detection serves as a "vote" that a pothole has occurred at that location. Voting represents the most common type of Level 2 Fusion.

Performing anomaly detection directly on data from a single sensor and a single pass requires that the effect of the damaged track on the signal is known *a priori*. This is the case in some train-based monitoring studies, for example, in the work of Molodova et al. [9]. They study rail-squats and define this type of damage as any point in the tracks where a particular feature exceeds a threshold set *a priori*.

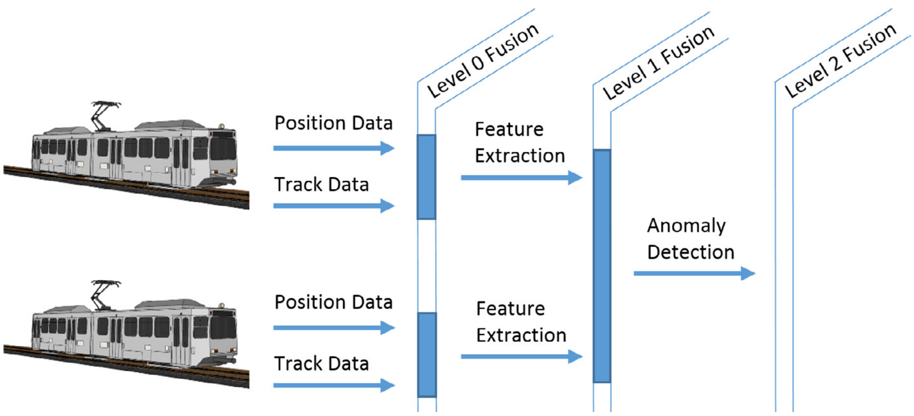

Fig. 1. Proposed processing pipeline. The proposed data fusion approach in this paper is a Level 1 Fusion method. Level 0 Fusion is the combination of raw data, Level 1, of features, and, Level 2, of decisions. Although Level 2 Fusion is not required for the proposed pipeline, it describes a type of fusion used in related studies [20], and is thus included for completeness.