Onboard Estimation and Classification of Railroad Curvature

Jacob Trehag, Peter Händel, Senior Member, IEEE, and Maria Ögren.

Abstract—In this study, a method to classify and identify railroad plane geometry is developed. The aim is to determine where the transition curves, straight tracks and circular curves are located along the railroad. For this purpose, the railroad curvature measurements are used. The method is based on a Double Exponential Smoothing algorithm (DES). The DES algorithm is described with the aid of linear filters, where the outputs are the estimate of the curvature and its rate of change. These quantities are utilized to form a detector, where the time instants of a curvature change are determined. Owing to noisy measurement data, a FIR-median Hybrid Filter with predictive FIR substructures is implemented. This nonlinear filter improves the tracking ability in the DES-filter outputs.

Index Terms—least squares methods, median filters, nonlinear filters, piecewise linear approximation, rail transportation testing.

I. INTRODUCTION

Railroads have been the key infrastructural enablers for transportation of goods and passengers during the last 200 years. The development of contemporary train-sets requires instrumentation for test purposes based on digital signal processing of sensor data. Railroad plane geometry is generally described as straight tracks, circular curves with a constant curvature, and transition curves that connect a straight track segment to a constant curvature segment.

The formation of the transition curves can vary, e.g. in Sweden the transition curves are restricted to a clothoid formation, i.e., the radius varies linearly with its length, yielding a piecewise linear railroad curvature Fig. 1.

In Sweden, the map information is provided by Banverkets Information System (BIS), which is an office in the Swedish Rail Administration. The data provided by BIS contain information on the plane geometry of the Swedish railroads and is presented as a function of distance, i.e., the start point of where the construction of the railroad track began, and defines where the distance begins. Accordingly, based on the location of the train set, information on the plane geometry of the track is given by BIS.

Jacob Trehag and Maria Ögren are with Interfleet Technology AB, Stockholm, Sweden. Email: jacob.trehag@interfleet.se.


©2008 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

One of the approaches to determine the position of the train is to use the global positioning service (GPS), with a GPS receiver mounted on the vehicle. Standard GPS solutions are, however, not accurate enough to provide positioning information with sufficient precision and accuracy. An earlier study [1] presents such a system where GPS equipment is mounted on the railroad vehicle, in combination with local reference station providing GPS-correction data by radio. With the use of such a differential GPS, the vehicle position and accordingly the railroad plane geometry can be identified with sufficient accuracy. However, this study considers an alternative to the approach mentioned earlier [1], where the plane geometry is directly estimated from the measurements provided by onboard mounted sensors, yielding a more flexible and practical solution for the long-distance testing conditions. In particular, the method is independent of a database such as BIS. By detecting where the transition curves are located along the railroad, the extreme states in a railroad vehicle’s dynamical behavior can be extracted. The dynamical behavior depends on the length and radius of the circular curves, i.e. the radius impact the magnitude of the lateral forces [2]. The instrumentation presented in this study is based on onboard mounted sensors only, and thus, is suitable for long-term test scenarios on public railroads.

Estimation of the railroad plane geometry is considered based on the in-train measurements of speed and yaw rate, providing an assessment on the curvature of the railroad. Furthermore, a GPS-based preprocessing for debiasing the measurement is employed. The railroad curvature is modeled as a first-order, piecewise linear polynomial representing sections of straight tracks, transition curves and circular curves along the railroad. In short, the curvature describes how much a geometric figure deviates from being straight. The employed approach combines the model-based linear filtering incorporating local models of the railroad curvature with the
nonlinear filtering to suppress noise preserving the edges in the curvature transitions. The linear filtering is based on a polynomial model of the data and a fading memory least-squares problem is solved [3]. Here, the approach yields the estimates of the curvature and its rate of change. By implementing the asymptotic (as the number of data tends to infinity) least-squares solver, a simple yet effective solution is established with guaranteed stability, where the trade-off between noise suppression and transient behavior set by a single-user-chosen design parameter. In other contexts, the fading memory least-squares modeling methodology is applied to the design of the digital notch filters in [4].

Owing to the noise and bias properties of the employed sensors, the curvature readings are preprocessed both in terms of debiasing and noise suppression. To preserve the edges, nonlinear filtering is utilized for the latter task. The nonlinear preprocessing employs the same polynomial model of the curvature readings as for the proceeding linear filtering, described earlier. The combination of linear predictors and median filters is known as the median filters with predictive substructures, introduced in the late 1980 [5]. The class of generalized nonlinear filtering is utilized for the latter task. The nonlinear filtering is found to be a versatile tool for a variety of applications, as introduced in the late 1980 [5].

In the worst-case, the computing time delay can be less than the sampling period. This can lead to a situation involving a cascade of recursive median filters. Subsequently, the set of feedback coefficients is optimized to minimize a desired cost function – resulting in a IIR-predictor has been designed in [7]. In this design [7], a feedback of the FIR forward predictors is introduced and subsequently, the set of feedback coefficients is optimized to minimize a desired cost function – resulting in a IIR-type predictor [7]. Furthermore, hybrid median filtering is also applied to determine the dynamic characteristics of the sensors by system identification in dynamic measurement [8]. The FIR-median hybrid filter has been utilized in combination with polynomial fitting [9] by combining the linear prediction and linear smoothing operators.

In a similar context, the use of recursive median filtering techniques have been employed to increase the performance of detecting the trend changes [10]. For this purpose, a method to diagnose jet turbine performance has been proposed, comprising of a cascade of recursive median filters. Subsequently, the detection is carried out with the aid of a gradient and laplacian edge detector. This kind of diagnostics require high accuracy of detecting the trend changes while keeping the rate of false alarms on a low level.

A block diagram of the proposed signal processing is shown in Fig. 2. The paper is organized as follows: in Section II, the problem is defined and the sensors and sensor outputs are presented along with a description on the debiasing of data. Section III introduces the edge-preserving nonlinear process-

![Fig. 2. Signal processing for estimation and classification of railroad curvature. The 1 Hz GPS readings are used for debiasing the 300 Hz yaw rate \( \dot{\phi}(n) \) and speed \( \nu(n) \) measurements with the aid of extended Kalman filtering (EKF). The raw curvature data \( \gamma_c(n) \) is downsampled with a factor 3 to reduce the computation time and subsequently preprocessed by an edge-preserving nonlinear preprocessor (FMH), producing \( \gamma_p(n) \). The smoothed curvature output \( \hat{\gamma}(n) \) and its rate of change \( k(n) \) are the outputs from the discounted least-squares filter (DES).]

II. MEASURING RAILROAD CURVATURE

A. Definition and modeling the curvature

The track curvature \( c(n) \) [m\(^{-1}\)] is the ratio between the railroad vehicle yaw rate \( \dot{\phi}(n) \) [rad/s] and the velocity in its direction of travel (i.e., the speed) \( \nu(n) \) [rad m/s], i.e.,

\[
c(n) = \frac{\dot{\phi}(n)}{\nu(n)}.
\] (1)

Here, and in the sequel, \( n \) denotes the discrete time in terms of integer sample instants corresponding to a sampling rate of \( F_S \) [s\(^{-1}\)].

The curvature, \( c(n) \), is derived from the arc length \( s(n) = \gamma(n) r(n) \), yielding

\[
c(n) = \frac{1}{r(n)} = \frac{\gamma(n)}{s(n)} = \frac{\dot{s}(n)}{s(n)}.
\] (2)

where \( \dot{s}(n) \) and \( s(n) \) are discrete time readings corresponding to the time derivative of the arc length and angle of the arc.

Typically, the speed \( \nu(n) = \nu_0 + \Delta \nu(n) \) for some positive speed \( \nu_0 \) and speed variation \( \Delta \nu(n) \) ensuring a well defined quotient in (1). The curve radii of interest are in the range from 250 [m] to \( \infty \) [2]. This implies that \( c(n) \) is in the range \( \pm C \) [m\(^{-1}\)], where \( C \in [1/\infty, 1/250] \) [m\(^{-1}\)]. The curvature’s rate of change depends on the length of the curve and the curve radius, i.e.,

\[
k(t) = \lim_{\Delta t \to 0} \frac{c(t + \Delta t) - c(t)}{\Delta t}.
\] (3)

In discrete time domain (3) can be expressed as

\[
k(n) = c(n) - c(n - 1).
\] (4)

As illustrated in Fig. 1, a local (i.e., within a segment of the track) model of the curvature is

\[
c(n) = k n + c_0,
\] (5)

where \( k \) is the rate of change and \( c_0 \) is a real-valued constant. With reference to (5) and the illustration presented in Fig. 1,
during a circular curve $k \approx 0$ and $c_0$ is a constant. During a transition curve, $k = \pm k_0$ for some positive $k_0$ determining the constant rate of change. For a straight track, $k \approx 0$ and $c_0 \approx 0$ holds true. The output of the proposed filtering algorithm is estimates at some time instant $n$ of the curvature as well as its rate of change.

**B. Sensor outputs and debiasing**

The onboard sensors provide sampled measurements of the yaw rate and the speed, and the measurements are denoted by $\dot{\varphi}_m(n)$ and $v_m(n)$, respectively. The yaw-rate data $\dot{\varphi}_m(n)$ is provided by a fiber-optic gyroscope of model SAAB rate-gyro GR-G5 8438009-107 with a sensitivity of $285 \cdot 10^{-3} \text{ rad/s}$. The speed $v_m(n)$ is measured with a DRS05 Doppler radar that has a standard deviation of 0.65 km/h after calibration in the vehicle environment. The Doppler radar measurement of speed has a high update rate and reliability, compared with the speed measurement by GPS. In general, the velocity can be measured with any arbitrary measurement method that is available. The only consideration is to measure the velocity in the vehicle’s direction of travel. The data acquisition and synchronization of data is done with a MCGplus data acquisition system, at a sampling rate of $F_S = 300 \text{ [s}^{-1}]$. Since the system collects several other quantities that urge a sampling rate of $300 \text{ [s}^{-1}]$, while the developed application can consider a lower sampling rate, the sensor outputs are downsampled with a factor 3, to reduce the computation time.

However, the sensor outputs are contaminated with sensor noise and systematic bias errors. The main error source is the yaw-rate gyroscope, both with respect to measurement noise and bias term. Accordingly, the measured yaw rate approximately fulfills

$$\dot{\varphi}_m(n) = \dot{\varphi}(n) + B + w'(n),$$

where $B$ is a bias term and $w'(n)$ is a zero mean-additive noise term. In practice, the bias $B$ slowly varies with time, (compared with the update rate given by $F_S$), urging for more advanced signal processing than a standard differentiation of data.

Let $\dot{\varphi}_r(n)$ and $v_r(n)$ denote the output from an initial processing stage driven by the measurements $\dot{\varphi}_m(n)$ and $v_m(n)$, respectively, where the bias terms are removed from the sensor signals. To remove the bias, a GPS-aided dead reckoning navigation system is employed to preprocess the yaw rate and velocity signals prior to forming the curvature reading [11]. In short, the preprocessing comprises an extended Kalman filter (EKF) to estimate the bias in the yaw rate and velocity signals. The EKF utilizes the GPS position and a trajectory, calculated from the yaw angle and velocity, to estimate the bias.

**C. Curvature observations and their processing**

Based on $\dot{\varphi}_r(n)$ and $v_r(n)$, the observed raw curvature $c_r(n)$ is formed by (1) by replacing $\dot{\varphi}(n)$ and $v(n)$ by the preprocessed measurements. The influence of the preprocessing stage on the curvature readings is illustrated in Fig. 3, where it can be noted that the biased quotients in (1), especially a biased yaw-rate signal, have an impact on the curvature readings. Assuming that the railroad vehicle is running on a straight track and the yaw-rate signal is biased towards some non-zero value instead of being zero, the curvature output will indicate a curve when the vehicle either decelerates or accelerates. A biased velocity signal only affects the curvature amplitude.

Thus, an additive noise model is appropriate, i.e., observed curvature $c_r(n)$ is modeled as

$$c_r(n) = c(n) + w(n),$$

where $w(n)$ is a zero mean additive noise term. In the subsequent sections of the paper, the bias-compensated curvature observation $\hat{c}_r(n)$ is further processed in two stages, where the first stage is a nonlinear finite-impulse response median hybrid-filter structure with the aim to reduce the additive noise preserving the edges in the curvature readings. The second stage is a model-based infinite-impulse response linear-filter producing smoothed estimates of the curvature and its rate of change, i.e., $\hat{c}(n)$ and $\hat{k}(n)$, respectively. The outputs from the second stage are transformed to form two detection variables to obtain the transitions between the different curve types and straight tracks.

**III. EDGE-PRESERVING NONLINEAR FILTERING**

Denoising of a corrupt direct current (DC) level with edge-preserving properties may be performed by a median filter, for which the output is the median of the measurements in a rectangular window [12]. A generalization of median filtering to ramp-like signals is provided by FIR-median hybrid filters with predictive FIR substructures [5]. A curvature trend consisting of constant levels and ramps makes it possible to use a zero- and first-order predictor. The zero order, denoted with the order $\ell = 0$, is for level prediction, while first order, denoted with $\ell = 1$, is for ramp prediction.

In short, the predictive substructures result in a set of prediction candidates based on the prior notation that the curvature is either constant in straight tracks and circular curves, or varies linearly in transition curves. The nonlinear median operation makes a proper selection to form the output, $c_{\ell}(n)$, of the FIR-median hybrid filter.

Consider a rectangular data window of length $M = 2L + 1$ and linear predictor outputs

![Fig. 3. An illustration of the bias reduction done in the preprocessing stage. In the marked area, the biased-curvature trend indicates a curve while the unbiased-curvature trend does not.](image-url)
\[ e_P^0(n) = \sum_{i=1}^{L} h_F^0(i) c_r(n - i) \]  
and
\[ e_P^1(n) = \sum_{i=1}^{L} h_F^1(i) c_r(n - i), \]
where \( h_F^0(i) \) is the forward FIR filter of length \( L \), for zero-order prediction defined as [5]
\[ h_F^0(i) = \frac{1}{L}, \quad i = 1, \ldots, L. \]  
Here, (10) is derived from the constraint to minimize the output noise power at zero frequency, yielding FIR filters that are pure averagers [5]. The forward FIR filter of length \( L \) for the first-order prediction is defined as [5]
\[ h_F^1(i) = \frac{4L - 6i + 2}{L(L - 1)}, \quad i = 1, \ldots, L. \]  
Here, (11) is derived to minimize the noise power in each sample under the assumption that the noise components in each sample are independent of each other [5]. The filter coefficients for backward prediction are defined by (10) and (11) in a time-reversed order. In Fig. 4, the step response of each predictor is shown with the intention to describe their different characteristics.

The filter output is formed as
\[ c_p(n) = \text{MEDIAN}[c_F^1(n), c_F^0(n), c_r(n), c_B^0(n), c_B^1(n)]. \]  
(12)

A. **Edge-Preserving Noise Attenuation by Iterative Filtering of a Sequence of Data**

The basic principles of hybrid median filtering with predictive substructures have been outlined earlier. Furthermore, it has been shown that if the median filtering is iterated, where the length of the predictive FIR filters is increased with the iterations, the output noise will decrease [5]. However, the length of the filters is constrained to ensure root-signal convergence [13]. A root signal is a signal which is invariant to the filter [5]. This implies that in an iterative filtering scheme, a root-signal convergence is ensured by setting the length of the filters to [5],
\[ L_r \leq 2L, \]
(13)
where \( L_r \) is the length of the root signal and \( L \) is the number of filter coefficients in each of the predictors. A sufficient condition for a signal to be a root signal is that it should only consist of constant neighborhoods and edges [5].

The iterative edge preserving nonlinear filtering process is initialized with a length of the FIR predictors set to \( L^0 \). The root-signal length is set to be one-third of the amount of samples corresponding to a normal transition curve where a typical length of a normal transition curve varies from 70 to 150 m. The coefficients \( \beta \) and \( \tau \) are constant parameters in the iterative filtering process. The number of updates are set by \( \tau \) and is related to the iterator \( i \) as
\[ f = i \mod \tau. \]  
(14)
The length of the predictors are updated every time when \( f = 0 \) as
\[ L^i = L^{i-1} + \beta. \]  
(15)
The parameter \( \beta \) is related to constraint (13) and \( \tau \) as
\[ \beta = \left\lfloor \frac{L_r - L^0}{\tau} \right\rfloor, \]
(16)
where \( \lfloor \cdot \rfloor \) is the floor function.

The values of \( \beta \) and \( \tau \) have been empirically determined from experiments. Setting the root-signal length \( L_r \approx 300 \) samples, initializing the length of the predictors, \( L^0 = 25 \), iterating \( i = 100 \) times and updating the length of the predictors, \( L \), around 10 times \( (\tau = 10 \text{ and } \beta = 27) \) has provided satisfying performance in detecting trend changes in the curvature data.

An example of the iterative edge-preserving nonlinear filtering process is given in Fig. 5.

**IV. Discounted Least Squares Estimation**

Based on the debiased and enhanced curvature data, efficient decision variables can be obtained by finite memory least-squares modeling.

Based on the preprocessed curvature data \( c_p(n) \), the aim is to form decision variables based on estimates of the curvature and its rate of change. We model the input as
\[ c_p(n) = c(n) + w(n). \]  
(17)
This section summarizes the derivation of the discounted least squares estimation, i.e., \( \hat{c}(n) \) of \( c(n) \) and \( k(n) \) of \( k \). The quantity \( w(n) \) is \( \sigma_w^2 \), modeling the measurement noise and other model errors. Owing the nonlinear preprocessing, Gaussianity of the data cannot be expected. Thus, we employed a least-squares approach to extract estimates of the curvature, and its rate of change. Let \( \theta(n) \) be the time-varying state variable
\[ \theta(n) = \begin{pmatrix} c(n) \\ k \end{pmatrix}. \]  
(18)
Then, $\theta(n)$ obeys the recursion $\theta(n) = F \theta(n-1)$, where
\[
F = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.
\] (19)

Consider the prediction error $\varepsilon(n) = c_p(n) - \hat{c}(n)$, where $\hat{c}(n)$ is the one-step ahead prediction based on the data up to and including the time instant $n-1$. Such a prediction is given by
\[
\hat{c}(n) = (10)^T F \theta(n-1).
\] (20)

The weighted least-squares criterion is given as
\[
\sum_{n=0}^{N} \lambda^{(N-n)} \varepsilon^2(n).
\] (21)

Here, $0 \leq \lambda \leq 1$ is the forgetting factor. The memory length is given, by the rule-of-thumb as $2/(1 - \lambda)$. By neglecting the transient response of the recursion, (21) can be described with the help of the linear filters as
\[
x(n) = 2\lambda x(n-1) - \lambda^2 x(n-2) + (1 - \lambda)c_p(n),
\] (22)
\[
\hat{c}(n) = (1 + \lambda)x(n) - 2\lambda x(n - 1)
\] (23)
\[
\hat{k}(n) = (1 - \lambda)[x(n) - x(n - 1)].
\] (24)

Here, $\hat{c}(n)$ and $\hat{k}(n)$ denotes the least-squares solution, i.e., the filter outputs. A block diagram is depicted in Fig. 6.

The choice of $\lambda$ is a compromise between noise rejection and tracking ability. A large value will render in a more robust filter, but with a decreased tracking ability. On the other hand, a low value will increases the tracking ability, but will render in a less robust filter.

The influence of the preprocessing of data on the estimated rate of change is illustrated in Fig 7.

**A. Detection and Segmentation**

The detection scheme is based on handling segments of transition curves and thereafter detecting the start and end time in each transition curve. In order to detect the end and start time of a transition curve, two detection variables are defined as
\[
\tau_c(n) = \frac{\hat{c}(n)}{c_{MAX}}
\] (25)

and,
\[
\tau_k(n) = \frac{\hat{k}(n)}{k_{MAX}}.
\] (26)

The quantities $\tau_c$ and $\tau_k$ are used for detection of the start and end time, respectively, and $c_{MAX}$ and $k_{MAX}$ are the corresponding maximum values of $\hat{c}(n)$ and $\hat{k}(n)$, respectively, in a segment.

The segmentation is done by locating the data indexes where the curvature’s rate of change is higher than the threshold value $\delta_c$. To carry out this, a segmentation variable is introduced and defined as
\[
\tau_{segm}(n) = \frac{|k(n)|}{k_g}.
\] (27)

where $k_g$ is a generic numerical value of the curvature’s rate of change in a transition curve.

In Fig. 9, the start and end time of a transition curve is presented. Here, a data segment starts at time $n_{segm}$. The start
time is detected when \( \tau_k \) deviates from a threshold value \( \delta_d \) (ideally \( \delta_d = 0 \)), and if this occurs at time \( n_s \) then the start time is expressed as

\[
t_{\text{start}} = n_{\text{segm}} + n_s.
\]

Furthermore, the end time is detected where \( \tau_c \) reaches its peak level and if this occurs at time \( n_e \) then the end time is expressed as

\[
t_{\text{end}} = n_{\text{segm}} + n_e - n_{\text{corr}}.
\]

The term \( n_{\text{corr}} \) is a correction term of the detected end-time, defined in Fig. 8. Note that the time correction is dependent on \( \lambda \).

The procedure is also applicable on an incoming transition curve (see Section I, Fig. 1). For an outgoing transition curve the data is reversed, thus the end and start time are determined in a similar manner.

Based on the detected transitions between the different types of curves, the curvature within a track segment can be subsequently estimated by a least-squares fit of a zero- or first-order linear regression, respectively.

V. RESULTS

In this study, the detected time instants of a trend change in curvature data have been compared with the curvature data from BIS. The data from BIS is a compilation of the intended location of a railroad segment, based on the information from construction phase of the railroad. This implies that the location can vary in reality, e.g., the topography may have an impact on the intended location of a railroad segment, yielding an uncertainty in the BIS data.

Furthermore, in the verification procedure, an integration of the speed has been done to simplify the comparison with BIS. With known distance locations in both the BIS and measurement data, the result from the developed application has been verified.

The measurement data have been acquired during the “Gröna Täget” [14] survey performed in the summer of 2007. In Fig. 11, a short segment of railroad curvature somewhere in between Ånge and Sundsvall, Sweden, has been processed and the detected trend changes in curvature data are presented. Similar segments of railroad curvature data have been investigated and verified against BIS data and the error of the detected trend changes where found to be within the range of \( \epsilon \in [-7,7] \) [m].

Furthermore, though there were error sources that influenced the verification, such as the uncertainty in BIS data and the
integration of the railroad vehicle’s speed, the result obtained were still considered to be satisfactory.

![Fig. 11. Detected trend changes in curvature measurements. The curvature data have been downsampled with a factor 3 to reduce the iterative FMH-filtering time.](image)

### VI. Discussions and Conclusions

The development of contemporary and future trains is important for sustainable transportation of humans and goods. The areas require novel measurement techniques for dynamic testing of train sets on public railroads. Key information include the instantaneous curvature of the track, which is the main parameter of interest in this study. Earlier methods require equipment installed along the track. However, in this study, this is not required, as our method is based on onboard mounted sensors only. A model-based signal processing comprising nonlinear and linear filtering blocks has been derived and its performance has been investigated based on the field-test data.

The application can be improved further by implementing a cancellation of the lateral twitches that impact the yaw-rate readings. Lateral twitches can be found in trains where the comfort is not considered, i.e., in cargo trains where the suspension does not interfere the lateral twitches yielding an influence in the yaw-rate readings. However, it is common to measure the lateral accelerations in both ends of the train when testing the running dynamics of a rail vehicle [2], and thus, a noise cancellation can be considered where the dynamics of the lateral twitches can be found from the lateral acceleration measurements.

### References


