Steering Measurement Decomposition for Vehicle Lane Keeping – A Study of Driver Behaviour

Yu Zhang*, Timothy Gordon, Miguel Martinez-Garcia, Chris Bingham

School of Engineering, University of Lincoln, Lincoln, LN6 7TS, U.K.

Abstract

Steering control for vehicle lane keeping has attracted significant attention from both automotive industries and researchers. To describe intermittent pulse-like qualities imparted by drivers that are seen in real-world steering measurements, a pulse control model (PCM) is presented for vehicle lane keeping. Inspired by the PCM, a steering angle measurement is decomposed into a combination of trend, integrated sine components (ISCs) and sine components (SCs), where trend corresponds to the path curvature, ISCs to the heading angles, and SCs to the lateral positions. Trends are extracted through the use of empirical mode decomposition (EMD) and principal component analysis (PCA), with singular spectral analysis (SSA) and Fourier curve-fitting (FCF) being employed to determine the ISCs and SCs in the main pulses. Through statistical pattern analysis on experimental measurements of drivers’ steering performance, it is revealed that (1) the pulse steering behaviour from real drivers shows the benefit of the proposed PCM for steering control during lane keeping, and (2) classification of pulse steering characteristics can be used for normal driver state identification and highlight abnormal driving behaviour, leading to the prospect of identifying driving characteristics typical of impaired concentration, substance misuse or tiredness, for instance.

Highlights

- The PCM can describe intermittent pulse-like qualities in real-world steering measurements.
- Steering angle measurements can be decomposed into a combination of trend, ISCs and SCs.
- EMD and PCA present appropriate tools for measurement trend extraction.
- SSA and FCF are appropriate techniques for the extraction of ISC and SC pulses.
- Patterns of pulse characteristics are useful for PCM validation and driver identification.

Keywords: Pulse control model, empirical model decomposition, principal component analysis, singular spectral analysis, Fourier curve-fitting, statistical pattern analysis.

*corresponding author: +44(0)1522 837938; yzhang@lincoln.ac.uk
1. Introduction

Intelligent vehicle technologies have many different functions, all aimed at improving comfort, efficiency and safety. The systems, sub-systems and algorithms that provide intelligent vehicle control are based on a range of disciplines: mathematics, computer science and engineering, which all contribute towards the goal of complementing or partially replacing the human driver in navigation and driving related behavioural tasks [1]. Of late, the intelligent vehicle concept has undergone a substantial increase in interest, as some of the required goals to achieve automated vehicles seem near accomplishment (examples include Volvo Steer Assist technologies, already successfully implemented, and the prototypical Mercedes self-driving trucks).

More recent information reveals that, with Volvo Pilot Assist, the driver can set a 'time gap' between their vehicle and the vehicle in front. The automation system can adjust the speed to maintain that 'time gap' according to the current speed. The system also helps the driver to maintain the vehicle inside the lane, but the driver can override the system by steering actively. Thus this is an example of shared control. Mercedes has an Adaptive Cruise Control system that performs a similar method for speed adjustment. The fact that, Level 5 automation – according to the SAE – does not seem to be happening soon due to technological constraints. Specifically, artificial intelligence does not seem to cope well in urban environments. On the other hand, shared control technologies seem to be more achievable. But in order to achieve effective transitions between the human and the vehicle, automated systems must be able to identify driver states. Drivers don't drive in the same mode at all times, e.g. sometimes more distracted or drowsy, and other times more alert, so an effective driver assistance technology should be able to identify these changes. In this paper, the attention is on identifying driver behaviour related to steering control.

Existing steering control models usually rely on two aspects: (i) a simplified models of visual information processing based on the numerical interpretation of perceived lane keeping error; (ii) a control law that models the manipulative control actions to keep the vehicle within the road boundaries and to track the road path. An example of such an approach is the well-known linear feedback law proposed by Salvucci and Gray [2] [3]:

$$\delta = k_1 \theta_n + k_2 \dot{\theta}_n + k_3 \dot{\theta}_f$$

where $\delta(t)$ is the steering angle, $\theta_n$ and $\dot{\theta}_n$ are the angle and angular rate to a near-point at a distance $d_1$ ahead of the vehicle (which a the lane centring abilities of the driver and the control of lateral velocity), and $\dot{\theta}_f$ defines the angular rate of the far-point at a distance
In [4], the adequacy of this model was tested with naturalistic driving data (NDD). It was seen that, although it is possible to fit the parameters of the control law (1) in a consistent manner, the fitted parameters lead to instability in closed-loop simulation. On the other hand, it has been verified experimentally that human drivers can achieve sufficient steering performance by using only the near and far information of the road [5]. Thus deficiencies in the Salvucci and Gray linear model to reproduce human steering behaviour arise from the control law itself. Although it is easy to optimize the parameters to produce high performance control, there is no evidence that the resulting control law is effective at representing human driving dynamics. Indeed, control law (1) is deficient in that it is purely reactive, there is no memory component in the model.

In [4], an alternative nonlinear control law was proposed—based on short pulse-like corrections, instead of a continuous steering wheel motion. In the 'pulse control model' (PCM), the driver is presumed to apply a ramp steer input to correct yaw-rate deviations. The ramp pulses are supplemented with further integrated sine component (ISC) pulses to ‘trim’ the yaw angle relative to the local direction of the lane, and sine component (SC) pulses to adjust lateral offsets. The pulse paradigm for steering control modelling can take account of attention switching and can be used to represent different driver states. Additionally, the intermittent rate of the steering pulses and their length can model distinct levels of fatigue and attention. Pulse modelling control works as an open-loop control action during the application of a pulse, as further error tracking or correction is stopped until the termination of the pulse. Thus the PCM can simulate pre-learned or pre-cognitive [6] steering control actions.

Inspired by PCM, this paper aims to capture short-pulse behaviour whilst recognising that there exists additional noise patterns that have no particular relevance, possibly as a consequence of vibration and other disturbances, and pre-emptive control where the driver sees changing curvature and makes anticipative larger-scale adjustments. This is achieved by searching for an optimal decomposition of the steering angle measurements with a global trend combined with elements of ISC and SC pulses, and accomplished using trend extraction, smoothing and curve fitting techniques.

The presented study is based on data obtained from tests with subjects in real driving settings, with the hypothesis of low-workload driving conditions, e.g. straight roads, light traffic and lack of strong disturbances such as crosswinds. The tests were executed in a fleet of Nissan Altima 3.5SE vehicles, specially instrumented to record several components of...
driver behaviour. The data is part of the road departure crash warning field operational tests [7] conducted in South-East Michigan. During the tests, all the recording sensors were set not to interfere with the driving task, therefore the data collected are considered 'naturalistic', which represent normal driving. The data was segmented into 60-second driving events which only include low workload steady highway driving – of radius of curvature never exceeding 500m. From the events lane changes and or other interventions were excluded. Each driving event has an initial speed in the range 28-32 ms$^{-1}$, overall maximum variation less than ±5 ms$^{-1}$. Fig. 1 shows an example steering time history.

![Steering angle measurement set](image)

**Figure 1.** Example steering angle measurement set

Trend extraction is typically used to highlight the long-term characteristics of measurements. An example tool for this that of empirical mode decomposition (EMD) [8] and using a subset of the resulting intrinsic mode functions (IMFs) to reconstruct an extracted trend within the underlying data based on the local characteristic time scale of the signal [9]. The effectiveness of EMD based technique for signal decomposition has been demonstrated in [10], where EMD is applied to extract higher order statistical parameters for gear fault diagnosis. [11] has also used modification of EMD to quantify the nonlinearity and non-stationarity of dynamic response of a vehicle–track coupling system.

However, IMFs after EMD do not quantify their relative contribution to the overall trend, therefore principal component analysis (PCA) is considered as a means of supplementary
selection of a minimal set of principal components (PCs) [12]. PCA can be used to synthesize and recompose the results of EMD. The concept is evaluated in [13] for non-invasive cardiovascular signals, where it is stated that PCA is optimal in generating a smaller set of orthogonal components and removing the correlations between the IMFs after EMD. PCA is also used in [14] to sift the extracted features after ensemble EMD in different scales of time domain and frequency domain. It is stated that the features of high priority through PCA present the important dynamical characteristics for gear fault diagnosis. In this paper, through the combination of EMD and PCA, the global trend of the steering measurement sets can be extracted and separated from the additional pulses.

Following trend extraction, the next task is to extract meaningful pulses from the detrended signals. Singular spectral analysis (SSA) is used extensively for feature extraction and predictive modelling to decompose time series into reconstructed subseries (RSs) that are nearly periodic [15]. For instance, [16] has applied multi-scale higher order SSA method to reveal features of the non-Gaussian and nonlinear characteristics in vibration signals for rotating machinery fault detection. Here, SSA is utilised for de-noising purpose, and it has shown to be beneficial preserving the nearly periodic pulses without compromising the sharp edges. Furthermore, multi-harmonic sine fitting [17] [18] is employed to extract the ISC and SC pulses implicit in the PCM through Fourier curve fitting (FCF). Finally, statistical patterns of the extracted pulse signals can be revealed.

The rest of the paper is organised as following: Section 2 explains the underlying principles of the applied techniques, i.e. PCM, EMD, PCA, SSA and FCF, with demonstrations of the measurement decomposition and pulse extraction from one example steering signal; Section 3 shows the generalization of experimental trials from 4 drivers and 20 steering signals each driver, with results of the statistical patterns of driver steering pulse behaviours; finally, Section 4 concludes the paper and suggests applications of the pulse extraction techniques for both PCM validation and driver identification, etc.

2. Underpinning Principles

Inspired by PCM, steering angle measurements are decomposed into their trend, ISC and SC constituents. Specifically, EMD and PCA extract the underlying trend and the dominant pulses in the steering angle measurements. The ISCs and SCs are obtained from the extracted main pulses using SSA and FCF. Additionally, by applying condition rules, short-pulse patterns are revealed that identify the normal driving behaviours of different drivers.
To provide a seed for what follows, the proposed methodology is depicted in Fig. 2.

![Outline Methodology](image)

**Figure 2.** Outline Methodology.

### 2.1 The Pulse Control Model

PCM is based on simple control logic, where yaw-rate deviations are presented by imparting a ramp steer input which is then augmented by a combination of ISC pulses corresponding to local angle adjustments and SC pulses corresponding to local lateral position adjustments. Prescribed component pulses, $f_a$, $f_0$ and $f_b$ are shown in Fig. 3, and simulation of vehicle motion responding to the above pulses is given in Fig. 4, where in this case, vehicle speed is 20 m/s, and the pulses are applied $t_0 = 2.0$ seconds. The peak angular velocity at the steering wheel is 40 deg/sec. Clearly, $f_a$ (ramp) imparts a long-term change in path curvature. $f_0$ (ISC) imparts a discrete change in direction and $f_b$ (SC) provides a small change of lateral displacement. Based on this description, is applied in response to lane boundary conflicts at the far point as defined by yaw rate errors (YREs) [19], while is applied to correct for near-point yaw angle errors (YAEs), and is applied for
near-point central lane keeping (CLK).

Figure 3. Steering pulses $f_a$, $f_0$ and $f_b$ used to control YRE, YAE and CLK respectively.

Figure 4. Vehicle responses for steering pulses $f_a$, $f_0$ and $f_b$ respectively.
2.2 Trend Extraction using Empirical Mode Decomposition

EMD is based on the direct detection of local signal extrema at a variety of intrinsic time scales to decompose the steering measurement sets. The resulting IMFs are considered as the most important characteristics of the data, and since the decomposition is based on localised timescales, is readily applicable to nonlinear and nonstationary signals. EMD is therefore characteristically different from more traditional wavelet denoising techniques [20], for instance.

An intrinsic mode function (IMF) of a signal from EMD satisfies the following two conditions: (1) the number of extrema and the number of zero crossings are equal, or their difference is no more than 1, and (2) its local mean is zero. Specifically, given a signal \( x(t) \), the constituent IMFs, \( c_i(t) \), can be obtained and summed such that [21]:

\[
x(t) = \sum_{i=1}^{n_{IMF}} c_i(t).
\]

By way of example, the decomposition of the example steering angle measurement (in Fig. 1) is shown in Fig. 5, with the 11 associated IMFs—the final IMF is termed a residual vector.

![Figure 5. IMFs of the example steering angle measurement after EMD](image)
2.3 Selection of IMFs using Principal Component Analysis

PCA is applied to reduce large datasets whilst preserving ‘sufficient characteristics’ of the original measurements by determining the principal components (PCs) [22]. By way of example, after performing PCA on the matrix of the 11 decomposed IMFs from Fig. 5, the resulting 11 PCs are shown in Fig. 6(a). The cumulative sum of the percentage of the PC reconstructions is shown in Fig. 6(b). It can be seen that the first 2 PCs contribute more than 80% of the total sum and can be used to reconstruct the main features, or trend, of the original signal—see Fig. 7.

![Figure 6. PCA: (a) PCs of the IMFs; (b) Cumulative sum of the eigenvalues of the PCs](image-url)
2.4 Use of Singular Spectral Analysis for Smoothening

Dominant pulses are extracted by removing the underlying trend from the original signal, as shown in Fig. 7, and SSA is subsequently applied to the extracted pulses to decompose the time series into reconstructed subseries (RSs) that are nearly periodic [15]. The near-periodic components are useful for extraction of ISCs and SCs. SSA is also shown to robustly filter secondary small noise components whilst preserving meaningful high-bandwidth transients when performing measurement smoothening.

Consider $x_n$, $n = 1,\ldots, N$, to be a given time series with a window length $L$ ($1 < L < N$), and $K = N - L + 1$. The trajectory matrix of $X$ with a dimension of $L \times K$ is

$$X = \begin{bmatrix} x_1 & x_2 & \cdots & x_K \\ x_2 & x_3 & \cdots & x_{K+1} \\ \vdots & \vdots & & \vdots \\ x_L & x_{L+1} & \cdots & x_{N} \end{bmatrix}. \quad (3)$$

Let $S = XX^T$, with the constituent $\lambda_i$ ($i = 1,\ldots, L$) being the set of eigenvalues of $S$ in a decreasing order. Set $d = \max\{j, \lambda_j > 0\}$, with $U_j$ ($j = 1,\ldots, d$) being the associated eigenvectors. The factor vectors are defined as $V_j = X^T U_j / \sqrt{\lambda_j}$ ($j = 1,\ldots, d$). Donate
\( \tilde{X}_j = \sqrt{\lambda_j} U_j V_j^T \), the singular value decomposition (SVD) of the trajectory matrix is represented by

\[
X = \tilde{X}_1 + \cdots + \tilde{X}_d. \tag{4}
\]

The indices set \( \{1, \ldots, d\} \) are grouped into \( m \) disjoint subsets \( I_1, \ldots, I_m \), where each subset has a length \( l \). A new expansion of (4) is then:

\[
X = \tilde{X}_{I_1} + \cdots + \tilde{X}_{I_m}. \tag{5}
\]

Each matrix \( \tilde{X}_I \) is hankelized with the resulting Hankel matrix being transformed into a new series of length \( N \), \( \tilde{x}_n^{(k)} \). Finally, the initial time series \( x_n \) can be expressed as a combination of \( m \) RSs:

\[
x_n = \sum_{k=1}^{m} \tilde{x}_n^{(k)} \quad (n = 1, \ldots, N), \tag{6}
\]

where each subseries can equivalently be considered as frequency components [23].

SSA is applied to the main pulses in order to reduce the effect of secondary noise components whilst preserving the near-periodic cyclic behaviour—see Fig. 8. The frequency component is selected based on the least mean square prediction error. It is notable that, unlike alternative smoothing techniques, the sharp (high bandwidth) edges are largely preserved.

![Figure 8. Reconstructed primary pulses using SSA](image-url)
2.5 Application of Fourier Curve-fitting

Subsequent to processing using SSA to denoise the measurement sets, Fig. 8, Fourier curve-fitting [18] is applied to the data over 2 second intervals (i.e. 20 sample points), since it is expected that the pulses are presented with durations of < 2 seconds [4], using

\[ f(x) = a_0 - \sum_{i=1}^{6} a_i \cos(\omega_i x) + \sum_{i=1}^{6} b_i \sin(\omega_i x), \tag{7} \]

where \( \omega_1 \) to \( \omega_6 \) are selected such that the durations of the Fourier sine and cosine waves span 0.4 to 1.4 seconds in general (averagely 1.4, 1.2, 1.0, 0.8, 0.6 and 0.4 seconds for indices 1 to 6).

Specifically, the curve is fitted using a ‘Bisquare’ method, which minimizes a weighted sum of the squared errors. The weight is allocated to each data point according to the distance of the point from the fitted curve, e.g. points at the fitted curve get full weight, whereas points farther get reduced weight – extreme outliers may get zero weight. This ‘Bisquare’ method is chosen over the conventional least-squares approach, because it reduces the effect of outliers, which the real data contain considerably.

Consequently, coefficients \( a_1 \) to \( a_6 \) and \( b_1 \) to \( b_6 \) are calculated, in Fig. 9, and indicate the amplitudes of the corresponding ISC and SC pulses, as discussed in Section 2.1.

![Figure 9. FCF to SSA reconstructed pulses](image-url)
Considering the primary pulses that are extracted in Fig. 8, the amplitudes of the ISC and SC pulses are plotted in Fig. 10. (Note: For improved clarity of presentation, pulse amplitudes less than 0.2 degrees are omitted in Figs. 10 to 12.) The diffusion pattern of Fig. 10 reveals that for shorter time intervals the local corrections are more stable. The outer bound indicates that there is a maximum angular rate and that this is independent of the size of the pulses. It also reveals that maximum pulse amplitudes decrease when the duration is greater than 1.1 ~ 1.2 seconds.

3. Experimental Trials

Pulse investigation is important since on straight or gently curving roads, lane keeping can be maintained by small incremental adjustments, ramps (the trend) for changing path curvature, short high-bandwidth ISC pulses to change direction, and SC pulses to adjust lateral offset. However, during real-world driving conditions, drivers will often use alternative behaviour such as low frequency feedback (slow ramps). The aim is therefore to look for pulses occurring intermittently, and to characterize them.

Following the pulse extraction procedure of Section 2, total 20 steering angle measurement sets obtained from the same driver (Driver 1, for Fig. 1) are investigated, and
the amplitudes of pulses vs. time intervals are revealed (procedure as described by Fig. 10), with the results shown in Fig. 11(a) and (b), for ISCs and SCs. The positive and negative envelopes are drawn based on exponential functions excluding outlier data points. The ISCs and SCs possess similar characteristic envelopes.

Figure 11. Driver 1—Fourier fitting coefficients: (a) ISCs; (b) SCs
Further, from Fig. 12, which considers results from 4 different drivers (with 20 steering signals each), the envelopes are also similar, indicating that the maximum angular rate is largely independent of the size of the pulses and the driver.

To differentiate the pulse behaviour of different drivers, histograms of the 12 Fourier fitting coefficients are plotted in Fig. 13 (a)—(d) for Drivers 1—4, respectively. It is shown, for instance that the results for Driver 1 exhibits greater pulse behaviour (i.e. more instances of large amplitude pulses), whilst Driver 4 imparts fewer high-bandwidth pulses than other drivers. By neglecting secondary noise components (i.e. pulses with amplitude <0.2 degrees), it is clear that ISC and SC pulses with durations of 1.0–1.2 seconds ($a_2$, $a_3$ and $b_2$, $b_3$) are a prevalent choice in driving behaviour.

To provide more clarity of the salient features, cumulative probability curves of the most meaningful fitting parameters $a_2$, $a_3$ and $b_2$, $b_3$ are also given in Fig. 14(a) and (b), respectively. From these it is clear that all drivers impart similar pulse driving characteristics in all cases. It is also notable that individual drivers show a greater propensity to using a particular pulse characteristic. For instance, Driver 1 shows a greater propensity to using positive SC pulses (corresponding to $b_2$ and $b_3$, as mentioned, to correct small lateral
offsets) with duration 1.0~1.2 seconds, as shown in Fig. 14(b).
Figure 13. Histograms of the Fourier fitting coefficients for 4 drivers: (a) Driver 1; (b) Driver 2; (c) Driver 3 and (d) Driver 4.
In this paper, we have characterized normal driving based on realistic driving data. Thus
a system could work by identifying that the driver is in 'normal model' and not in need of steering assistance through the identified parameters. Drivers present curves or 'valid parameters' according to their driving state [24]. When the driver is outside this range a warning system could alert the driver without the need of cameras to monitor the driver. The identified similarities in driver pulse behaviour demonstrates the relevance of the PCM concept, whilst the differences in imparted pulse behaviour from different drivers can be employed for driver identification purposes.

The pulse control scheme was also validated in [25], where a similar driver modelling approach was proposed based on the primitive steering pulses. However, its model has some limitations, as the parameters were fitted based on Salvucci and Gray model, which responds simultaneously to near and far information – nevertheless, the assumption is not valid in real life driving.

Among other shared control technologies, the framework of optimal control is often considered in the literature [26][27], but their approach has the inconvenience that it is very difficult to establish what is the 'real' objective function that the human driver uses. In this paper, we seek to describe the human from the point of view of extracting its steering primitives. It is shown that the steering signals can be reconstructed through the measurement decomposition and data analysis techniques, so this establishes a methodology for biofidelic driver model representation.

4. Conclusion

This paper has focused on the application of signal decomposition of steering behaviour measurements for vehicle lane keeping. Based on PCM, steering angle measurement can be considered as a combination of trend, which contributes to global steering direction, ISC pulses, which represent local heading corrections, SC pulses that adjust vehicle lateral offsets, and secondary noise components. In this paper, EMD and PCA are used to extract the global trend, while SSA and FCF are the applied to extract the ISC and SC pulses. The similarity of the statistical actuation patterns for different drivers’ pulse behaviour has demonstrated the validity of using the PCM for vehicle lane keeping, which establishes the benchmarking for normal driver behaviours; whilst the differences pulse behaviour can be used for driver identification or to highlight deviations from normal driving characteristics—the latter therefore providing a mechanism to inform on impaired concentration through distraction, drug misuse or tiredness, for instance.
**Abbreviations**

PCM — Pulse Control Model  
NDD — Naturalistic Driving Data  
ISC — Integrated Sine Component  
SC — Sine Component  
EMD — Empirical Mode Decomposition  
IMF — Intrinsic Mode Function  
PCA — Principal Component Analysis  
PC — Principal Component  
SSA — Singular Spectral Analysis  
RS — Reconstructed Subseries  
FCF — Fourier Curve-Fitting  
YRE — Yaw Rate Error  
YAE — Yaw Angle Error  
SVD — Singular Value Decomposition  
CLK — Central Lane Keeping

**Acknowledgements**

The authors would like to thank the University of Michigan Transportation Research Institute for access to the driving data used in this analysis.

**References**


M. Martinez-Garcia, Y. Zhang, T. Gordon, “Modelling lane keeping by a hybrid open-closed-loop pulse
