

A Traffic Volume Detection Algorithm for Overlapping Noise Identification: Postprint

Authors: Ma Qinglu, Zou Zheng

Date: 2019-01-28T00:00:00+00:00

Abstract

In traditional traffic volume detection, the short-time energy method and short-time average amplitude method suffer from the inability to identify overlapping vehicle noise segments. Extracting the spectrum view (SV) most correlated with overlapping signals as a feature for these segments helps address this issue. Based on the above conclusions, a traffic volume detection algorithm based on SV features is proposed. The algorithm first performs filtering and denoising on traffic noise containing overlapping vehicle noise segments; then extracts SV features from the traffic noise via fast Fourier transform (FFT) and applies smoothing to these features; finally, based on the SV features and employing the dual-threshold decision principle, performs endpoint detection on the traffic noise to separate the overlapping vehicle noise segments. Experiments conducted on a traffic noise dataset from single-lane, low-traffic road sections demonstrate a 20% improvement in accuracy compared to traditional methods, thereby validating the effectiveness of the proposed algorithm.

Full Text

Preamble

A Traffic Volume Detection Algorithm for Identifying Overlapping Noise

Ma Qinglu, Zou Zheng†

(School of Traffic & Transportation, Chongqing Jiaotong University, Chongqing 400074, China)

Abstract: Traditional traffic volume detection methods based on short-term energy and short-term average magnitude cannot effectively identify overlapping traffic noise segments. Extracting the spectrum view (SV), which is most

correlated with overlapping signals, as a feature of overlapping traffic noise segments can help address this limitation. Building on this insight, this paper proposes a traffic volume detection algorithm based on SV features. The algorithm first filters and denoises traffic noise containing overlapping segments. It then extracts SV features via Fast Fourier Transform (FFT) and smooths these features. Finally, employing dual-threshold decision principles based on SV features, the algorithm performs endpoint detection on traffic noise and separates overlapping segments. Experiments on a dataset of single-lane, low-traffic road noise demonstrate that the proposed algorithm achieves a 20% improvement in accuracy compared to traditional methods, thereby validating its effectiveness.

Keywords: traffic engineering; traffic volume detection; traffic noise; overlapping noise recognition; spectrum view; endpoint detection

0 Introduction

Urban roads generate massive amounts of traffic noise data daily, and noise pollution represents a significant external cost of urban transportation that causes discomfort, annoyance, and distress [1]. However, traffic noise also carries substantial useful information that can be extracted before isolation, shielding, or elimination measures are applied. Since speech processing techniques have matured considerably [2], they can be adapted for traffic noise analysis. Research on traffic noise is gradually emerging as a new direction in intelligent transportation information collection.

Most existing literature on traffic noise focuses on traffic assessment [3–8]. Joshi et al. [9] proposed a traffic density state classification method for chaotic traffic conditions using traffic noise measurements. Bhandarkar et al. [10] suggested training support vector machines and artificial neural networks to classify traffic density levels based on features extracted from vehicle noise signals: MFCC and energy. Borkar et al. [11] developed a neuro-fuzzy classifier based on MFCC representation to classify traffic flow into three density levels, achieving over 95% classification accuracy. Guan et al. [12] conducted noise testing and TNM analysis on a new LOS noise barrier experimental section of Interstate 465 in Indiana, USA, investigating the effects of pavement type, vehicle speed, traffic volume, noise reduction coefficient, barrier height, and barrier extension on LOS noise barrier performance. Their results indicated that LOS barriers should be designed with a noise reduction greater than 6.7 dBA, a minimum height exceeding that of large vehicles, a maximum height not exceeding 6.6 m, with optimal height determined through simulation analysis of noise reduction effects over the design life.

Research on vehicle identification and traffic volume detection using traffic noise is relatively limited. Alexandre et al. [13] combined extreme learning machines with genetic algorithms to select relevant features from traffic noise for vehicle classification, improving correct classification rates from 74.83% (without

feature selection) to 93.74% (with feature selection). Lefebvre et al. [14] extracted MFCC feature vectors from multi-lane traffic noise and used support vector machines to estimate traffic volume over 11.5 days, with experimental results showing estimation errors around 85%. Mato-Mendez et al. [15] proposed a stereo recording-based system for detecting vehicle passages and identifying vehicle categories, classifying vehicles into motorcycles, cars, and trucks. Testing on approximately 761 signals yielded correct identification rates of about 95% per category. Torija et al. [16] used measured traffic noise spectral profiles as input data for real-time short-term urban road traffic flow estimation, achieving an average classification success rate of 96.1%.

Research on traffic volume detection when traffic noise signals overlap is even scarcer. Regarding overlapping sound signals, Dennis et al. [17] utilized local spectral features to recognize overlapping sound events and separate background noise, demonstrating effectiveness that could potentially be applied to overlapping traffic noise scenarios. He Xiaohua [18] proposed an endpoint detection method for traffic noise collected under single-lane, low-flow conditions based on short-term energy features, but did not consider overlapping traffic noise cases. Therefore, to address the problem of overlapping traffic noise, this paper proposes a traffic volume detection method based on spectrum view (SV) feature analysis of traffic noise, comparing it with traditional short-term energy and short-term average magnitude methods to validate the SV algorithm.

1 Basic Concepts

Traffic noise comprises a combination of vehicle noise and environmental noise. The definition of traffic noise is illustrated in Figure 1. Figure 1 shows a 35-second sample extracted from 126 seconds of traffic noise, containing overlapping vehicle noise segments (included for subsequent analysis and comparison of features between normal and overlapping segments). Overlapping vehicle noise segments occur when the time headway between adjacent vehicles is small. Traffic noise analysis involves preprocessing, median filtering, and normalization techniques, which are explained below.

Preprocessing: Before feature extraction, traffic noise must be preprocessed, including denoising, windowing, and framing. The LMS adaptive filter can automatically adjust its parameters without requiring prior knowledge [19], making it suitable for filtering and denoising traffic noise. Figure 1(a) shows the traffic noise waveform before LMS adaptive filtering, while Figure 1(b) shows the waveform after filtering. The denoised waveform is noticeably smoother, with higher distinguishability between vehicle noise and environmental noise.

Windowing: Among window functions, the Hamming window exhibits relatively small spectral leakage, making it the preferred choice. The Hamming window function is defined in Equation (1). The result of applying the window to framed signals is shown in Equation (2).

Smoothing and Normalization: After LMS adaptive filtering, traffic noise still has a low signal-to-noise ratio, causing significant fluctuations in feature curves within environmental noise segments. Therefore, further smoothing of feature parameter curves is necessary. Median filtering can remove isolated outliers without destroying step changes between smooth segments. Increasing the number of median filtering iterations can enhance smoothing effectiveness. Additionally, min-max normalization converts all data to values between $[0,1]$, eliminating magnitude differences across dimensions and preventing large computational errors due to significant scale variations. The min-max normalization calculation is shown in Equation (3).

2 Methodology

The fundamental principle of traditional traffic volume detection methods is to extract short-term energy (STE) or short-term average magnitude (STAM) features from traffic noise, then directly distinguish vehicle noise from environmental noise based on these features and count the number of vehicle noise segments, which represents the vehicle count in the traffic noise. However, these traditional algorithms do not consider overlapping vehicle noise segments, causing them to detect multiple overlapping segments as a single vehicle noise segment. To solve this problem, this paper proposes the SV algorithm. The SV algorithm's basic principle is: first, perform Fast Fourier Transform (FFT) on each frame of traffic noise and extract SV features; then distinguish vehicle noise from environmental noise based on SV features and separate overlapping vehicle noise segments; finally, count the number of vehicle noise segments as the vehicle count.

2.1 Traditional Algorithm

The definitions of STE and STAM for each frame of traffic noise are given in Equations (4) and (5), respectively, where E_i represents the STE of the i -th frame and M_i represents the STAM of the i -th frame.

Let P denote either STE (E) or STAM (M), and P_i represent the feature value of the i -th frame calculated using Equation (4) or (5). Let P_1 be the higher threshold and P_2 the lower threshold for feature P ; let l_{noveh} be the minimum length of environmental noise segments and l_{veh} be the minimum length of vehicle noise segments, both measured in frames; let Q_0 be the true vehicle count and Q the current detected vehicle count during algorithm execution, with ε_r as the relative error limit for Q . The dual-threshold endpoint detection process based on feature P is as follows [20]:

- a) Initialize P_1 , P_2 , l_{veh} , l_{noveh} , and ε_r .
- b) Sequentially calculate feature values P_i for each frame. If $P_i \geq P_1$ and $P_r \geq P_2$ ($r = i + 1, i + 2, \dots, i + l_{veh} - 1$), then set $i_{start}^Q = i$, where i_{start}^Q is the starting endpoint of the Q -th vehicle noise segment. Continue

calculating P_i from $i = i + l_{veh}$ until encountering $P_i < P_2$ and $P_s < P_2$ ($s = i + 1, i + 2, \dots, i + l_{noveh} - 1$), then set $i_{end}^Q = i + l_{noveh} - 1$, where i_{end}^Q is the ending endpoint of the Q -th vehicle noise segment. The interval $[i_{start}^Q, i_{end}^Q]$ represents the Q -th vehicle noise segment. Repeat step b) until all frames are processed, then proceed to step c).

- c) Count the number of detected vehicle noise segments (i.e., detected vehicles) Q . If $|Q - Q_0|/Q_0 \leq \varepsilon_r$, save the current parameters $P_1, P_2, l_{veh}, l_{noveh}, \varepsilon_r$ and all endpoint detection results, then stop; otherwise, adjust parameters and restart from step a).

Applying the traditional algorithm to the sample traffic noise in Figure 1(b) yields the endpoint detection results (i.e., vehicle count detection) shown in Figure 2. Figures 2(b) and 2(c) show vehicle counts of 4 veh and 2 veh, respectively; Figures 2(e) and 2(f) also show counts of 4 veh and 2 veh.

Observation of Figures 2(b) and 2(e) reveals that the peak values of the first two non-overlapping vehicle noise segments are below or equal to the value at point V (collectively denoted as V for both V_E and V_M), while the third non-overlapping segment's peak exceeds V . In Figure 2(b), threshold E_1 is set below V_E , detecting all non-overlapping segments but incorrectly classifying the overlapping segment as non-overlapping. In Figure 2(c), threshold E_2 is set above V_E , detecting the overlapping segment but missing all non-overlapping segments. Note that the third vehicle noise segment remains undetected because its length fails to meet the l_{veh} requirement at the current threshold.

2.2 SV Algorithm

As shown in Figure 3 [Figure 3: see original paper], Figure 3(a) presents the extracted SV features of traffic noise, while Figure 3(b) provides a graphical analysis of these features. The yellow S-region indicates the noise area above the overlapping signal portion. Vehicle noise segments appear significantly darker than environmental noise segments. Environmental noise information primarily distributes in low-frequency regions, whereas vehicle noise information extends continuously from low to mid-high frequencies, gradually lightening in color with increasing frequency until approaching the color of environmental noise.

Graphical analysis of SV features reveals: (a) non-overlapping vehicle noise segments can be identified directly by their amplitude values exceeding those of environmental noise segments; (b) the S-region contains almost no information and its color depth nearly matches that of environmental noise. Based on these characteristics, overlapping vehicle noise segments can be effectively separated and then converted into non-overlapping segment identification problems.

The FFT result $Y_i(k)$ for each framed signal $y_i(n)$ is given by Equation (6) [21]. Since $Y_i(k) = Y_i(N - k)$, only $[N/2]_{int} + 1$ sample points need processing. The frequency interval $\Delta f = f_s/N$, where f_s is the sampling frequency.

The SV feature calculated from Equation (6) can be represented by matrix A_{sv}

as shown in Equation (7). The frame time vector t and frequency scale vector f for SV are defined in Equation (8). In SV, the horizontal axis represents time, the vertical axis frequency, and coordinate values a_{ni} represent amplitude values after FFT. Each column of A_{sv} corresponds to frequency values in f , and each row corresponds to time values in t^T . Since 2D planes express 3D information, amplitude magnitude is indicated by color depth—darker colors represent larger amplitudes.

Based on Equations (6)-(8), the SV features of the sample traffic noise in Figure 1(b) are extracted. Let f_0 and f_1 be predefined upper and lower frequency limits, with corresponding row indices k_1 and k_0 in A_{sv} . Let t_{noveh} be the duration of leading environmental noise segments, N_{noveh} the number of samples within t_{noveh} , and β a tunable coefficient representing the amplitude upper limit of no-vehicle segments in SV. The calculations for λ , μ , and N_{noveh} are shown in Equation (9), where μ is the average amplitude within t_{noveh} at frequency f_1 , and λ is a scaling factor.

For more intuitive comparison with traditional algorithms, short-term average magnitude M'_i is defined based on SV features, analogous to Equation (5), as shown in Equation (10), where a_{ni} represents elements in A_{sv} . Endpoint detection results based on M'_i are shown in Figure 5 [Figure 5: see original paper], yielding 5 veh—identical to SV-based detection (provable through derivation). Thus, SV features can replace M'_i features.

The SV algorithm implementation proceeds as follows:

- a) Initialize f_0 , f_1 , l_{veh} , l_{noveh} , t_{noveh} , β , and ε_r , starting execution from row k_0 in A_{sv} .
- b) Traverse amplitude values a_{ni} in the current row sequentially. If $a_{ni} > \beta$ and $a_{nr} > \beta$ ($r = i + 1, i + 2, \dots, i + l_{veh} - 1$), set $i_{start}^Q = i$ as the starting endpoint of the Q -th vehicle noise segment. Continue traversing from $i = i + l_{veh}$ until encountering $a_{ni} < \beta$ and $a_{ns} < \beta$ ($s = i + 1, i + 2, \dots, i + l_{noveh} - 1$), then set $i_{end}^Q = i - 1$ as the ending endpoint, with interval $[i_{start}^Q, i_{end}^Q]$ representing the Q -th vehicle noise segment. The point $i_{end}^Q + 1$ marks the start of the Q -th environmental noise segment. Proceed to step c.
- c) Continue traversing from $i = i + l_{noveh}$ until finding $a_{ni} > \beta$ and $a_{nr} > \beta$ ($r = i + 1, i + 2, \dots, i + l_{veh} - 1$), then set $i_{start}^{Q+1} = i$ as the starting endpoint of the $(Q + 1)$ -th vehicle noise segment and $i_{end}^Q = i - 1$ as the ending endpoint of the Q -th environmental noise segment. The interval $[i_{start}^{Q+1}, i_{end}^Q]$ represents the Q -th environmental noise segment. Proceed to step d.
- d) Check whether interval $[i_{start}^{Q+1}, i_{end}^Q]$ contains any interval from set Φ . If not, add this interval to Φ and assign μ to all values in the submatrix formed by rows k_0 to k_1 and columns i_{start}^{Q+1} to i_{end}^Q in A_{sv} . If it does

contain an interval, decrement Q by 1. Repeat steps b-d until traversing all frames in row k_0 , then proceed to step e.

- e) Count the current number of vehicle noise segments Q . If $|Q - Q_0|/Q_0 \leq \varepsilon_r$, save all current parameters and detection results, then terminate; otherwise, adjust parameters and restart from step a.

Applying the SV algorithm to the sample traffic noise in Figure 1(b) yields endpoint detection results shown in Figure 4 [Figure 4: see original paper], detecting 5 veh—matching the ground truth. Figure 4(a) shows successful separation of overlapping vehicle noise segments, with endpoints displayed in Figure 4(b).

3 Experimental Results and Analysis

For a single-lane road section, audio detectors collected traffic noise for 126 seconds at an 8 kHz sampling rate. Experimental tests determined that a frame duration of 25 ms (200 samples) with a 12.5 ms frame shift was appropriate for analyzing vehicle noise. Figure 6 [Figure 6: see original paper] shows waveforms before and after filtering. The ground truth vehicle count is 25, equivalent to an hourly traffic volume of 714 veh/h. There are 7 overlapping noise segments, and the signal-to-noise ratio improves significantly after LMS adaptive filtering.

First, basic parameters for both traditional and SV algorithms were determined from Figures 2 and 4. Vehicle detection results based on STE, STAM, and SV algorithms are shown in Figures 7 Figure 7: see original paper-(c). STE and STAM both detected 18 vehicles, while SV detected 23 vehicles against a ground truth of 25. Neither STE nor STAM identified overlapping segments, whereas SV detected 5 overlapping segments but missed 2.

Following the analytical approach in Figure 3(b), SV should theoretically detect all 7 overlapping segments in Figure 6. However, close examination of the 3rd and 5th undetected overlapping segments in Figure 7(c) reveals: the 3rd overlapping segment lacks the S-region described in Figure 3(b), while the 5th overlapping segment has an S-region that is too narrow (width smaller than l_{noveh}) and located inside the overlapping portion, making detection impossible through parameter adjustment alone. Other overlapping segments were accurately detected because they exhibited S-regions similar to Figure 3(b).

The hourly traffic volumes detected by STE, STAM, and SV algorithms within 126 seconds are 514 veh/h, 514 veh/h, and 657 veh/h, respectively. Absolute errors, relative errors, and accuracy rates are calculated and summarized in Table 1 .

As shown in Table 1, both STE and STAM algorithms achieve 72% accuracy, while the SV algorithm reaches 92% accuracy—a 20% improvement over traditional methods.

4 Conclusion

Addressing the inability of short-term energy and short-term average magnitude features to effectively separate overlapping vehicle noise signals, this paper proposes a novel traffic volume detection method using traffic noise—the SV algorithm—and compares it with traditional STE and STAM algorithms. Key conclusions are:

- a) Traditional and SV algorithms share the commonality of utilizing dual-threshold decision principles and effectively detecting non-overlapping vehicle noise segments. The crucial difference is that traditional algorithms cannot simultaneously detect overlapping segments while maintaining detection of non-overlapping segments, regardless of parameter tuning, whereas the SV algorithm can.
- b) In terms of traffic volume detection accuracy, the SV algorithm outperforms traditional methods by 20%, primarily by successfully detecting additional overlapping vehicle noise segments.
- c) The proposed SV algorithm cannot achieve 100% detection of all overlapping segments because not all overlapping segments exhibit the S-region characteristics described in Figure 3(b). Only those meeting these characteristics can be accurately detected.
- d) Future research will focus on improving the SV algorithm to address current limitations and expanding the experimental dataset to further validate and enhance the algorithm's accuracy and stability.

References

- [1] Basner M, Babisch W, Davis A, et al. Auditory and non-auditory effects of noise on health [J]. *The Lancet*, 2014, 383(9925): 1325-1332.
- [2] Khademian M, Homayounpour M M. Monaural multi-talker speech recognition using factorial speech processing models [J]. *Speech Communication*, 2018, 98: 1-16.
- [3] Kumar P, Nigam S P, Kumar N. Vehicular traffic noise modeling using artificial neural network approach [J]. *Transportation Research Part C: Emerging Technologies*, 2014, 40: 111-122.
- [4] Morel J, Marquis-Favre C, Gille L A. Noise annoyance assessment of various urban road vehicle pass-by noises in isolation and combined with industrial noise: a laboratory study [J]. *Applied Acoustics*, 2016, 101: 47-57.
- [5] Coensel B D, Brown A L, Tomerini D. A road traffic noise pattern simulation model that includes distributions of vehicle sound power levels [J]. *Applied Acoustics*, 2016, 111: 170-178.

- [6] Can A, Leclercq L, Lelong J, et al. Accounting for traffic dynamics improves noise assessment: experimental evidence [J]. *Applied Acoustics*, 2009, 70(6): 821-829.
- [7] Can A, Leclercq L, Lelong J, et al. Traffic noise spectrum analysis: dynamic modeling vs. experimental observations [J]. *Applied Acoustics*, 2010, 71(8): 764-770.
- [8] Tyagi V, Kalyanaraman S, Krishnapuram R. Vehicular traffic density state estimation based on cumulative road acoustics [J]. *IEEE Trans on Intelligent Transportation Systems*, 2012, 13(3): 1156-1168.
- [9] Joshi V, Rajamani N, Prathapaneni N, et al. Traffic density state estimation based on acoustic fusion [C]//Proc of IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver, BC: IEEE Press, 2013: 478-482.
- [10] Bhandarkar M, Waykole T. Vehicular mechanical condition determination and on road traffic density estimation using audio signals [C]//Proc of CICN International Conference on Computational Intelligence and Communication Networks. New York: CICN Press, 2014: 395-401.
- [11] Borkar P, Sarode M, Malik L. Employing speeded scaled conjugate gradient algorithm for multiple contiguous feature vector frames: an approach for traffic density state estimation [J]. *Procedia Computer Science*, 2016, 78: 740-747.
- [12] Guan Bowen, Li Shuo, Jiang Yi, et al. Noise reduction effect evaluation of LOS noise barrier [J]. *Journal of Traffic and Transportation Engineering*, 2015, 15(4): 26-33.
- [13] Alexandre E, Cuadra L, Salcedo-Sanz S, et al. Hybridizing extreme learning machines and genetic algorithms to select acoustic features in vehicle classification applications [J]. *Neurocomputing*, 2015, 152: 13-22.
- [14] Lefebvre N, Chen Xiaodong, Beuseroy P, et al. Traffic flow estimation using acoustic signal [J]. *Engineering Applications of Artificial Intelligence*, 2017, 64: 164-171.
- [15] Mato-Mndez F J, Sobreira-Seoane M A. Blind separation to improve classification of traffic noise [J]. *Applied Acoustics*, 2011, 72(8): 690-698.
- [16] Torija J A, Ruiz P D. Using recorded sound spectra profile as input data for real-time short-term urban road-traffic-flow estimation [J]. *Science of the Total Environment*, 2012, 435: 270-279.
- [17] Dennis J, Tran H D, Chng E S. Overlapping sound event recognition using local spectrogram features and the generalised hough transform [J]. *Pattern Recognition Letters*, 2013, 34: 1085-1093.
- [18] He Xiaohua. Research on online traffic volume statistics algorithm based on vehicle running noise [D]. Changchun: Northeast Normal University, 2008.

- [19] Gao Ye, Ni Jing' en, Chen Jie, et al. Steady-state and stability analyses of diffusion sign-error LMS algorithm [J]. *Signal Processing*, 2018, 149: 242-251.
- [20] Onumanyi A J, Abu-Mahfouz A M, Hancke G P. A comparative analysis of local and global adaptive threshold estimation techniques for energy detection in cognitive radio [J]. *Physical Communication*, 2018, 29: 224-238.
- [21] Jodra L J, Gurrutxaga I, Muguerza J, et al. Solving Poisson' s equation using FFT in a GPU cluster [J]. *Journal of Parallel and Distributed Computing*, 2017, 102: 28-36.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv –Machine translation. Verify with original.