Data Denoising Based on Hadamard Matrix Transformation and Rayleigh Quotient Maximization: Application to GNSS Signal Classification

Jiang Yue, Bing Xu, *Member, IEEE*, and Li-Ta Hsu, *Member, IEEE*

***Abstract*— Global navigation satellite system (GNSS) signal type classification based on machine learning is an effective way to improve urban positioning performance. However, GNSS signal type features extracted are unrelated, and the number of features is limited, referred to as nonlocal- and few-feature issues, which limits the classification performance. This article presents a new data denoising theory to boost the classification performance based on concepts of Hadamard matrix transformation and Rayleigh quotient maximization. Hadamard matrix transformation increases the distance between different classes, i.e., interclass distance, by projecting the data into a new space, thereby increasing the classification performance. To improve the signal-to-noise ratio (SNR) of features, we maximize the Rayleigh quotient of the interclass distance. The proposed denoising approach is, in particular, effective for nonlocal- and few-feature signals. We applied the proposed data denoising theory to the GNSS signal type classification problem. Results indicate that GNSS signal type classification performance (microaveraging recall, i.e., Recall_μ) can be improved by about 5% ∼ 10% in a static test. For the dynamic test, about 1.5% ∼ 3.5% improvement is achieved.**

***Index Terms*— Classification, data denoising, global navigation satellite system (GNSS), multipath (MP), non-line-of-sight (NLOS) signal, reversible transformation.**

I. Introduction

THE global navigation satellite system (GNSS) positioning is challenging in urban canyons due to severe multipath (MP) interference and non-line-of-sight (NLOS) receptions [1], [2]. MP signals occur when the direct signal is reflected by nearby objects; thus, multiple signals are received, including the line-of-sight (LOS) signal. Different from MP, the NLOS receptions occur when the direct path is blocked, and only the reflected signals are received. As such, NLOS signals always result in positive-ranging errors that are independent of the receiver design [3]. A variety of techniques to tackle the MP/NLOS issue have been proposed, which can be broadly grouped into three categories, i.e., mitigation, detection, and correction, with their pros and cons. For instance, the mitigation of MP/NLOS signals is to remove or reduce their effects on the positioning. A typical example of this method is a choke-ring antenna that gives low gains to satellites of low elevation angles because low-elevated satellites are more likely to generate reflected signals [4]. Consistency checking is another way to find and remove MP/NLOS signals [5]. Detection tends to classify the received signals into different signal types, i.e., LOS/MP/NLOS, using some machine learning algorithms [6], [7]. Finally, correction is to apply corrections to measurements, so as to improve positioning accuracy. As an example, Xu *et al.* [8] proposed a vector tracking-based NLOS correction algorithm. Among these approaches, GNSS signal type classification is an effective and low-cost approach to improving urban GNSS positioning performance. The identified signal types are useful for both ranging-based least-squares GNSS positioning and 3-D mapping-aided GNSS positioning algorithms [9].

Machine learning algorithms have been widely applied to classification applications in the field of GNSS, e.g., the GNSS receiver context and channel classification [10], [11] and GNSS signal type classification [12]. Regarding GNSS signal type classification, researchers group the features into three categories, i.e., National Marine Electronics Association (NMEA), receiver independent exchange format (RINEX), and correlator levels based on the level at which the features are extracted [7]. A critical problem for classifying GNSS signals is that features extracted are linearly inseparable [13], as shown in Fig. 1, due to complicated propagations of rays in urban areas, as well as the dynamics of the receiver platform. Fig. 1 presents the point cloud of RINEX-level features of carrier-to-noise-density ratio (CNR), satellite elevation angle, and pseudorange rate residual for a low-cost commercial receiver. As can be seen, it is difficult for conventional machine learning algorithms to distinguish between different signal types due to noise interference.

Our previous study achieved an LOS/MP/NLOS classification rate of 75% using a support vector machine (SVM) classifier based on RINEX-level features [14]. An accuracy of 91.8% was reported in [12] using the algorithm called

Manuscript received 7 March 2022; revised 5 May 2022; accepted 31 May 2022. Date of publication 20 June 2022; date of current version 29 June 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 62103346 and Grant 61601225, and in part by the Guangdong Basic and Applied Basic Research Foundation under Project 2022A1515011680. The Associate Editor coordinating the review process was Yan Zhuang. *(Corresponding author: Bing Xu.)*

Jiang Yue is with the School of Science, Hohai University, Nanjing 211100, China (e-mail: hhu_yue@hhu.edu.cn).

Bing Xu is with the Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China and The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen 518057, China (e-mail: pbing.xu@polyu.edu.hk).

Li-Ta Hsu is with the Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China (e-mail: lt.hsu@polyu.edu.hk).

Digital Object Identifier 10.1109/TIM.2022.3184357

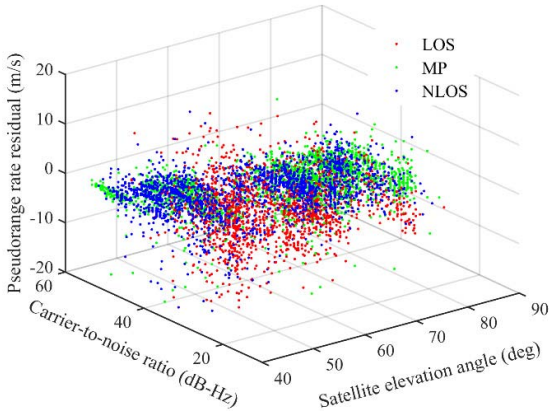


Fig. 1. Point cloud of RINEX-level features of GNSS signals.

the adaptive neurofuzzy inference system. Other than the classification rate, as introduced in [6], a confidence value of the classification is also provided using the robust decision tree-based LOS/NLOS classifier. This method, however, should make a tradeoff between high accuracy and low computational load. Suzuki *et al.* [15] propose to identify the NLOS reception by constructing an SVM classifier using features extracted from multiple autocorrelation functions (ACFs) based on a fact that the NLOS reception distorts the ACF of the direct LOS signal. It should be noted that, as of now, there is no benchmark data for evaluating the GNSS signal classification accuracy. Therefore, the accuracy reported in the existing literature is only valid for the specific dataset used. Xu *et al.* [7] made a fair comparison between the three-level classifiers using SVM and found that correlator-level performed best because of deeper-level measurements.

In addition to the commonly used SVM, decision-tree, and K-nearest-neighbor (KNN) based techniques, the neural network (NN) is also applied to GNSS signal type classification. In [16], an NN is constructed for LOS/NLOS classification with sequential ACF outputs as inputs. Features are extracted automatically within the network, instead of specifying features manually. Quan *et al.* [17] propose a convolutional NN (CNN)-based MP detection method, where time-series CNR and MP-induced pseudorange error are used as inputs to the network. In fact, NN is heavily dependent on several powerful nonlinear techniques, such as convolutional layers and pooling layers [18], [19]. The former is to detect local conjunctions of features from the previous layer, and the latter is to merge semantically similar features into a single one [20]. As such, a premise of the success of NN is that array data, such as images, have highly correlated features. However, there exist signals that have different physical meanings for different features, such as the GNSS signal. In general, GNSS signal type features are extracted from the observables and measurements, such as CNR, satellite elevation angle, and code pseudorange measurement, which have different meanings, referred to as nonlocal-feature signals in this article. Even worse, there are only a few features that can be extracted for these signals, referred to as a few-feature issue, causing a difficulty to build a reliable network. In view of this, this article proposes a data denoising technique to solve nonlocal- and few-feature issues.

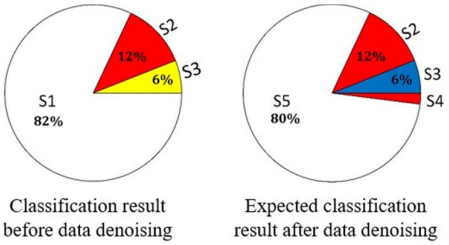


Fig. 2. Illustration of the idea of data denoising. S1 means samples correctly classified without denoising. S2 represents samples misclassified both with and without denoising. S3 represents samples misclassified without denoising but can be rectified with denoising. The noise of samples in S3 is comparable to interclass distances. S4 represents samples correctly classified without denoising but misclassified with denoising. S5 is a subset of S1, representing samples that can be correctly classified both with and without denoising.

A. Related Work on Denoising

Denoising is a well-known problem in the signal processing community and continues to attract researchers' attention. Some commonly used denoising algorithms include local filters [21], [22], total variation [23], [24], sparsity prior [25], [26], global regularization [27], [28], and so on. Among these algorithms, a fundamental idea is to combine similar patches for effective denoising [29]. If the signal is contaminated by the noise with different frequencies, the Fourier transform and the wavelet transform can be used to remove the noise by converting the noisy signal into a new domain and then modifying the noisy coefficients according to certain rules [30]. One of the most well-known rules is soft thresholding introduced in [31]. However, wavelet coefficients are statistically dependent, thereby causing difficulties to find proper models for wavelet coefficients of natural signals [30], [32].

B. Our Approach and Contributions

In this article, we propose a new framework to improve the classification performance of nonlocal- and few-feature signals. We apply the proposed algorithm to GNSS signal type classification. More specifically, we improve GNSS LOS/MP/NLOS signal classification performance by rectifying misclassified samples. Fig. 2 is an illustration of the idea of this article. The white area (S1) denotes samples that are easily identified due to low noise interference, while the red area (S2) represents samples that are misclassified because of large noise disturbance. The yellow area (S3) denotes the case where noise is comparable to the interdistance, which poses a great difficulty for classification and is the target area of our proposed data denoising theory. As shown in Fig. 2, with data denoising, samples in the yellow area are expected to be rectified. Besides, the data denoising method may also have a slight negative effect, as indicated by S4 area in Fig. 2. Notice that the numbers in Fig. 2 are typical values for real-world GNSS signal type classification in our experiments (see Section IV). For other applications, these numbers might be different, but the idea is the same, i.e., rectifying samples that are misclassified without data denoising due to comparable interclass distance and noise interference.

To improve classification performance, we propose a data denoising algorithm based on concepts of the Hadamard

matrix transformation and the Rayleigh quotient maximization. More specifically, we first define a new metric, difference-distance, which is used to better identify whether the sample is misclassified or not. Then, we improve the classification performance from two perspectives. On one hand, we transform the data into a new space using Hadamard matrix transformation. In the new space, a sample that is misclassified in the original space can now be classified correctly. On the other hand, we enhance the signal-to-noise ratio (SNR) of data based on the concept of the Rayleigh quotient, considering the analogous form of SNR and Rayleigh quotient definitions.

The main contributions are summarized as follows.

1) A metric called "difference-distance" is proposed to describe the misclassification of samples. Difference-distance is the difference between two Euclidean distances, which is used to quantify the misclassification performance.

2) We prove that the difference-distance metric with Hadamard $\mathbf{S}^{-1}$ transformation can rectify samples that are easily misclassified in the original space. Detailed derivations are given.

3) We improve the SNR of data based on the concept of the Rayleigh quotient. An algorithm is developed to find the reversible transformation matrix.

4) We apply the proposed denoising theory to machine learning-based GNSS signal type classification. Real-world data show that SVM, KNN, and NN's performances are all improved with data denoising.

The remainder of this article is organized as follows. Section II introduces the background of the proposed data denoising theory, including the definitions of distance metrics and the Hadamard $\mathbf{S}^{-1}$ matrix. Section III presents the proposed data denoising theory in detail, including the performance of difference-distance metrics in the $\mathbf{S}^{-1}$ transformation space and how to improve the SNR of data. In Section IV, we apply the proposed approach to a practical application, i.e., GNSS signal type classification. Section V concludes this article and suggests future work.

II. Background

A. Distance Metric for Misclassification

The Euclidean distance is widely used to assess the similarity of samples. However, when the noise disturbance is comparable to the Euclidean distance, the Euclidean distance cannot be used for classifying these samples, as shown in Fig. 3. In Fig. 3, samples a and b with feature vectors $\mathbf{f}_a$ and $\mathbf{f}_b$, respectively, belong to Class 2, while sample c with feature vector $\mathbf{f}_c$ is from Class 1. Due to the comparable noise between samples a and b to the Euclidean distance between samples a and c, sample a is easily misclassified to Class 2. Hence, a new metric, difference-distance with two states (positive or negative), is proposed to describe whether the sample is misclassified or not.

Before that, we introduce several assumptions and notations. Let $\mathbf{f} \in \mathbf{R}^N$ be a vector representing the features of one sample in the target data, with N being the number of features. A basic assumption is that the noise in different features is independent of each other. The Euclidean distance is used to evaluate the distinction between two feature vectors $\mathbf{f}_a$ and $\mathbf{f}_b$

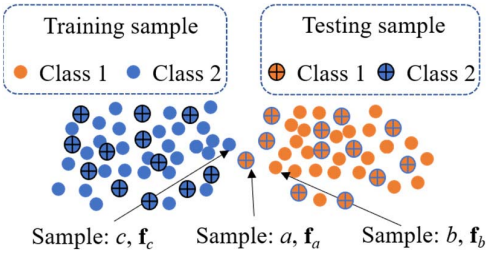


Fig. 3. Illustration of the definition of difference-distance. Assuming that samples b and c are labeled training data, sample a is a testing sample. If $D_{a-cb} \leq 0$, the sample a is likely to be misclassified to Class 2.

$$D_{ab} = \left|(\mathbf{f}_a - \mathbf{f}_b)^T(\mathbf{f}_a - \mathbf{f}_b)\right|^{1/2} = \left|\Delta\mathbf{f}_{ab}^T \Delta\mathbf{f}_{ab}\right|^{1/2} \tag{1}$$

where $\Delta\mathbf{f}_{ab} = \mathbf{f}_a - \mathbf{f}_b$ and the superscript T represents matrix transpose.

Definition 1 (Difference-Distance): The difference-distance is defined as the difference between two Euclidean distances. Specifically, letting D_{ab} be the Euclidean distance between samples a and b, and D_{ac} be the Euclidean distance between samples a and c, the difference-distance between D_{ab} and D_{ac} is defined as $D_{a-cb} = D_{ac} - D_{ab}$. If $D_{a-cb} \geq 0$, sample a is closer to sample b; that is, samples a and b are from the same class.

Now, we consider the noise vector $\mathbf{n}$. In addition, considering that samples a and b are from the same class, the difference in features $\Delta\mathbf{f}_{ab}$ is negligible. Alternatively, this feature difference residual can be absorbed in noise difference $\Delta\mathbf{n}_{ab}$. The difference-distance with noise disturbances between interclasses and intraclasses is

$$\begin{aligned} D_{a-cb} &= D_{ac} - D_{ab} \\ &= \left|(\Delta\mathbf{f}_{ac} + \Delta\mathbf{n}_{ac})^T(\Delta\mathbf{f}_{ac} + \Delta\mathbf{n}_{ac})\right|^{1/2} - \left|\Delta\mathbf{n}_{ab}^T\Delta\mathbf{n}_{ab}\right|^{1/2}. \end{aligned} \tag{2}$$

Considering the noise, the SNR is used as a metric for evaluating the disturbance in the Euclidean distance

$$\mathrm{SNR} = \left|\frac{\Delta\mathbf{f}^T\Delta\mathbf{f}}{\Delta\mathbf{n}^T\Delta\mathbf{n}}\right|^{1/2}. \tag{3}$$

In this article, the denoising algorithm is developed following two rules, i.e., rectifying the misclassified samples by correcting D_{a-cb} from negative to positive (referring to Fig. 3) and maximizing the SNR defined in (3).

B. Hadamard $\mathbf{S}^{-1}$ Matrix Transformation

Let $\mathbf{H}_m$ be a Hadamard matrix [33], which is an $m \times m$ matrix with entries ± 1 and satisfies $\mathbf{H}_m\mathbf{H}_m^T = m\mathbf{I}_m$. Its distinct row vectors are mutually orthogonal. Hadamard matrices have been used in a plenty of practical applications, such as coding theory and cryptology [34] and signal processing [35]. Wang *et al.* [36] and Yue *et al.* [37] applied it to the Hadamard transform spectrometry (HTS) where the denoising capability of Hadamard coding measurements with different reconstruction methods is analyzed. In fact, another Hadamard

matrix, called $\mathbf{S}$, is also widely used in spectroscopy [38]–[40], as defined in the following.

Definition 2 (Hadamard $\mathbf{S}$ *Matrix):* The first row and column of the Hadamard $\mathbf{H}_m$ matrix contain only $+1$ s. Deleting this row and column, we obtain an $(m-1)\times(m-1)$ matrix $\mathbf{G}_{m-1}$. Then, we change the $+1$ s in $\mathbf{G}_{m-1}$ to 0 s and -1 s to 1 s, so we have a matrix, $\mathbf{S}$ [38]. The matrix $\mathbf{S}^{-1}$ is the inverse matrix of the Hadamard $\mathbf{S}$ matrix. If $\mathbf{f}$ is a feature vector, we say that $\mathbf{S}^{-1}\mathbf{f}$ is an $\mathbf{S}^{-1}$ transformation of $\mathbf{f}$.

C. Euclidean Distance in the $\mathbf{S}^{-1}$ Transformation Space

The $\mathbf{S}^{-1}$ matrix is a basic transformation matrix to transform the data into a new space in the view of linear transformation. In the new space, the Euclidean distance between two features in (1) becomes

$$\begin{aligned} D_{ab}^{\mathbf{S}^{-1}} &= \left|\left(\mathbf{S}^{-1}\Delta\mathbf{f}\right)^T\left(\mathbf{S}^{-1}\Delta\mathbf{f}\right)\right|^{1/2} \\ &= \left|\Delta\mathbf{f}^T\left(\mathbf{S}^{-1}\right)^T\left(\mathbf{S}^{-1}\right)\Delta\mathbf{f}\right|^{1/2}. \end{aligned} \tag{4}$$

Letting $\mathbf{S}$ be a Hadamard S-matrix, $\mathbf{S}^{-1}$ can be written as [38]

$$\mathbf{S}^{-1} = \frac{2}{m+1}\left(2\mathbf{S}^T - \mathbf{J}_m\right) \tag{5}$$

where $\mathbf{J}_m$ is an $m\times m$ matrix with all elements being 1.

Substituting (5) into (4), we have [38]

$$\begin{aligned} D_{ab}^{\mathbf{S}^{-1}} &= \left|\Delta\mathbf{f}^T\left(\mathbf{S}^{-1}\right)^T\left(\mathbf{S}^{-1}\right)\Delta\mathbf{f}\right|^{1/2} \\ &= \left|\Delta\mathbf{f}^T\frac{4}{(m+1)^2}\left(2\mathbf{S}^T-\mathbf{J}_m\right)^T\left(2\mathbf{S}^T-\mathbf{J}_m\right)\Delta\mathbf{f}\right|^{1/2} \\ &= \left|\Delta\mathbf{f}^T\frac{4}{(m+1)^2}(2\mathbf{S}-\mathbf{J}_m)\left(2\mathbf{S}^T-\mathbf{J}_m\right)\Delta\mathbf{f}\right|^{1/2}. \end{aligned} \tag{6}$$

In addition, we have [33]

$$\mathbf{S}^T\mathbf{S} = \mathbf{S}\mathbf{S}^T = \frac{1}{4}(m+1)(\mathbf{I}+\mathbf{J}_m) \tag{7}$$

and [33]

$$\mathbf{S}\mathbf{J}_m = \mathbf{J}_m\mathbf{S} = \frac{1}{2}(m+1)\mathbf{J}_m. \tag{8}$$

With (7) and (8), the distance in the new space can be written as

$$\begin{aligned} D_{ab}^{\mathbf{S}^{-1}} &= \big|\Delta\mathbf{f}^T[(m+1)(\mathbf{I}+\mathbf{J}_m) \\ &\quad -2(m+1)\mathbf{J}_m + m\mathbf{J}_m]\Delta\mathbf{f}\big|^{1/2} \\ &= \left|\Delta\mathbf{f}^T[(m+1)\mathbf{I}-\mathbf{J}_m]\Delta\mathbf{f}\right|^{1/2}. \end{aligned} \tag{9}$$

III. Proposed Data Denoising Theory

For improving the classification performance, there are two steps implemented in this article, i.e., rectifying the difference-distance of samples from negative to positive and maximizing the SNR. This section describes the proposed data denoising approach in detail, including the Hadamard $\mathbf{S}^{-1}$ matrix transformation and its benefits and one self-defined matrix $\mathbf{A}$ that can improve the SNR of the target data.

A. Difference-Distance Metric in the $\mathbf{S}^{-1}$ Transformation Space

Specifically, as we mentioned earlier in Section II, if the difference-distance of a sample in the original space is negative, i.e., $D_{a-cb} < 0$, the sample would possibly be misclassified. Inversely, if the difference-distance of the sample in a new space satisfies $D_{a-cb}^{\mathbf{S}^{-1}} > 0$, it can be rectified. On the other hand, in the new space, we maximize the SNR of features to gain further classification rate improvement.

Theorem 1: Assume that the sample a with feature vector $\mathbf{f}_a$ and the sample b with feature vector $\mathbf{f}_b$ belong to Class 1. The sample c with feature vector $\mathbf{f}_c$ belongs to Class 2, as shown in Fig. 3. Based on the definition of difference-distance in (2), we have

$$P(D_{a-cb} > 0) < P\left(D_{a-cb}^{\mathbf{S}^{-1}} > 0\right) \tag{10}$$

where $P(\cdot)$ denotes the probability.

Proof: According to the definition of difference-distance in (2), if $D_{a-cb} = D_{ac} - D_{ab} \geq 0$, we have $(D_{ac})^2 - (D_{ab})^2 \geq 0$, and $D_{ac} + D_{ab} \geq 0$.

By applying the Hadamard matrix $\mathbf{S}^{-1}$ transformation to the Euclidean distance $\Delta\mathbf{f}^T\Delta\mathbf{f}$, we have

$$\begin{aligned} \left(\mathbf{S}^{-1}\Delta\mathbf{f}\right)^T\left(\mathbf{S}^{-1}\Delta\mathbf{f}\right) &= \frac{4\Delta\mathbf{f}^T\Delta\mathbf{f}}{(N+1)} - \frac{4\Delta\mathbf{f}^T\mathbf{J}_N\Delta\mathbf{f}}{(N+1)^2} \\ &= \frac{4}{(N+1)^2}\left\{\begin{array}{l}(N+1)\Delta\mathbf{f}^T\Delta\mathbf{f} \\ -\Delta\mathbf{f}^T\mathbf{J}_N\Delta\mathbf{f}\end{array}\right\} \end{aligned} \tag{11}$$

where $\mathbf{J}_N$ is an $N\times N$ matrix with all entries being 1. The terms $\Delta\mathbf{f}^T\Delta\mathbf{f}$ and $\Delta\mathbf{f}^T\mathbf{J}_N\Delta\mathbf{f}$ in (11) can be calculated, respectively, as

$$\Delta\mathbf{f}^T\Delta\mathbf{f} = \sum_{i=1}^{N}\left(\Delta\mathbf{f}^2(i)\right) = N\times E\left[\Delta\mathbf{f}^2(i)\right] \tag{12}$$

and

$$\begin{aligned} \Delta\mathbf{f}^T\mathbf{J}_N\Delta\mathbf{f} &= \Delta\mathbf{f}^T\begin{bmatrix}1&\cdots&1\\ \vdots&\cdots&\vdots\\ 1&\cdots&1\end{bmatrix}_{N\times N}\Delta\mathbf{f} \\ &= \left(\sum_{i=1}^{N}\Delta\mathbf{f}(i)\right)^2 \\ &= N^2\times(E[\Delta\mathbf{f}])^2 \end{aligned} \tag{13}$$

where $E[\cdot]$ is the expectation operator.

Substituting (12) and (13) into (11), we have

$$\begin{aligned} \left(\mathbf{S}^{-1}\Delta\mathbf{f}\right)^T\left(\mathbf{S}^{-1}\Delta\mathbf{f}\right) &= \frac{4}{(N+1)^2}\left\{\begin{array}{l}(N+1)\Delta\mathbf{f}^T\Delta\mathbf{f} \\ -\Delta\mathbf{f}^T\mathbf{J}_N\Delta\mathbf{f}\end{array}\right\} \\ &= \frac{4}{(N+1)^2}\left\{\begin{array}{l}\Delta\mathbf{f}^T\Delta\mathbf{f}+N^2E\left[(\Delta\mathbf{f})^2\right] \\ -N^2\times(E[\Delta\mathbf{f}])^2\end{array}\right\} \\ &= \frac{4}{(N+1)^2}\left\{\Delta\mathbf{f}^T\Delta\mathbf{f}+N^2\times\mathrm{Var}(\Delta\mathbf{f})\right\} \end{aligned} \tag{14}$$

where $\mathrm{Var}(\Delta\mathbf{f}) = \mathrm{E}[\Delta\mathbf{f}^2] - (\mathrm{E}[\Delta\mathbf{f}])^2$ is the variance of $\Delta\mathbf{f}$. According to (14), the difference-distance in the new space

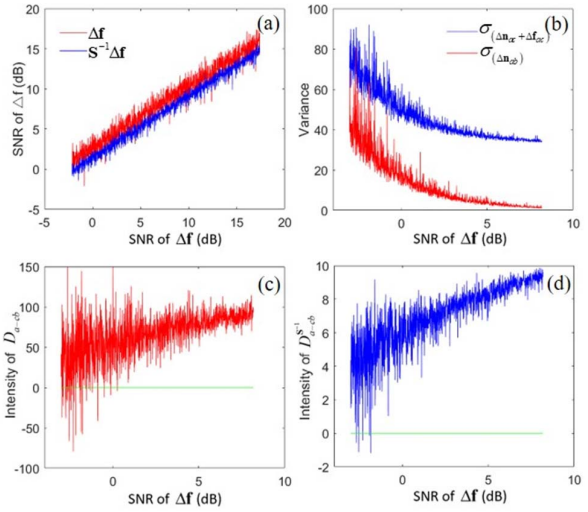


Fig. 4. Illustration of difference-distance metric in the $\mathbf{S}^{-1}$ transformation space. (a) Degradation of SNR caused by the matrix $\mathbf{S}^{-1}$. (b) Comparison of $\sigma_{(\Delta\mathbf{n}_{ac}+\Delta\mathbf{f}_{ac})}$ and $\sigma_{\Delta\mathbf{n}_{ab}}$. (c) Intensity of D_{a-cb}. (d) Intensity of $D^{\mathbf{S}^{-1}}_{a-cb}$.

can be derived to be

$$\begin{aligned} D^{\mathbf{S}^{-1}}_{a-cb} &= D^{\mathbf{S}^{-1}}_{ac} - D^{\mathbf{S}^{-1}}_{ab} \\ &= \frac{\left(D^{\mathbf{S}^{-1}}_{ac}\right)^2 - \left(D^{\mathbf{S}^{-1}}_{ab}\right)^2}{D^{\mathbf{S}^{-1}}_{ac} + D^{\mathbf{S}^{-1}}_{ab}} \\ &= \frac{4\left\{\begin{array}{l}(N+1)D_{a-cb}(D_{ac}+D_{ab})+\\ N^2\left(\sigma_{(\Delta\mathbf{n}_{ac}+\Delta\mathbf{f}_{ac})}-\sigma_{\Delta\mathbf{n}_{ab}}\right)\end{array}\right\}}{\left(D^{\mathbf{S}^{-1}}_{ac}+D^{\mathbf{S}^{-1}}_{ab}\right)(N+1)^2} \end{aligned} \tag{15}$$

where $\sigma_{(\Delta\mathbf{n}_{ac}+\Delta\mathbf{f}_{ac})}$ and $\sigma_{\Delta\mathbf{n}_{ab}}$ represent the variance of $(\Delta\mathbf{n}_{ac}+\Delta\mathbf{f}_{ac})$ and $\Delta\mathbf{n}_{ab}$, respectively. $\Delta\mathbf{f}_{ac} = \mathbf{f}_a - \mathbf{f}_c$, and $\Delta\mathbf{n}_{ac}$ means noise associated with $\Delta\mathbf{f}_{ac}$. According to (15), assuming the independence of $\Delta\mathbf{f}_{ac}$ and $\Delta\mathbf{n}_{ac}$, we have $\sigma_{(\Delta\mathbf{n}_{ac}+\Delta\mathbf{f}_{ac})} = \sigma_{(\Delta\mathbf{n}_{ac})} + \sigma_{(\Delta\mathbf{f}_{ac})}$. Given that $\Delta\mathbf{f}$ and $\Delta\mathbf{n}$ are comparable, the second term in the numerator of (15) can be considered positive. As a result, $D^{\mathbf{S}^{-1}}_{a-cb}$ is more likely to be positive compared with D_{a-cb}, which is in the first term of the numerator of (15). To illustrate the difference-distance metric in the $\mathbf{S}^{-1}$ transformation space, a simulation is performed with results shown in Fig. 4. As shown in Fig. 4(b), it is clear that the term $(\sigma_{(\Delta\mathbf{n}_{ac}+\Delta\mathbf{f}_{ac})} - \sigma_{\Delta\mathbf{n}_{ab}})$ is positive, and the intensity of $D^{\mathbf{S}^{-1}}_{a-cb}$ is more likely to be positive compared to D_{a-cb}. As a result, the classification performance is expected to improve. Theorem 1, therefore, is proofed.

Notice that the SNR of data may decrease due to the $\mathbf{S}^{-1}$ transformation, as shown in Fig. 4(a), thus degrading classification performance. To solve this issue, we introduce another matrix to improve the SNR by maximizing the Rayleigh quotient of the interclass distance, as introduced in Section III-B.

B. SNR Improvement via Rayleigh Quotient Maximization

In order to enhance the SNR of the target data, we transform (3) into a new space by designing a matrix $\mathbf{A}$ so that the SNR in the new space becomes

$$SNR' = \left|\frac{(\mathbf{A}\Delta\mathbf{f})^T(\mathbf{A}\Delta\mathbf{f})}{(\mathbf{A}\Delta\mathbf{n})^T(\mathbf{A}\Delta\mathbf{n})}\right|^{1/2} = \left|\frac{\Delta\mathbf{f}^T\left(\mathbf{A}^T\mathbf{A}\right)\Delta\mathbf{f}}{\Delta\mathbf{n}^T\left(\mathbf{A}^T\mathbf{A}\right)\Delta\mathbf{n}}\right|^{1/2}. \tag{16}$$

The question then becomes how to find the matrix $\mathbf{A}$ that satisfies

$$SNR' = \left|\frac{\Delta\mathbf{f}^T\left(\mathbf{A}^T\mathbf{A}\right)\Delta\mathbf{f}}{\Delta\mathbf{n}^T\left(\mathbf{A}^T\mathbf{A}\right)\Delta\mathbf{n}}\right|^{1/2} \geq \left|\frac{\Delta\mathbf{f}^T\Delta\mathbf{f}}{\Delta\mathbf{n}^T\Delta\mathbf{n}}\right|^{1/2} = \text{SNR}. \tag{17}$$

The concept of the Rayleigh quotient is introduced here considering that (17) has an analogous form to the Rayleigh quotient definition [41]

$$R(\mathbf{M}, \mathbf{x}) = \frac{\mathbf{x}^T\mathbf{M}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \tag{18}$$

where $\mathbf{M}$ is an $N \times N$ real and symmetric matrix. Letting $\mathbf{x} = \Delta\mathbf{f}$, we have $R(\mathbf{M}, \Delta\mathbf{f}) = ((\Delta\mathbf{f})^T\mathbf{M}\Delta\mathbf{f}/(\Delta\mathbf{f})^T\Delta\mathbf{f})$. As such, to find the solution to (17), we consider the maximization of the Rayleigh quotient, i.e.,

$$\arg\max(R(\mathbf{M}, \Delta\mathbf{f})). \tag{19}$$

The Rayleigh quotient has an interesting characteristic that it is invariable along the scale of $\Delta\mathbf{f}$, i.e.,

$$R(\mathbf{M}, \Delta\mathbf{f}') = \frac{c\Delta\mathbf{f}^T\mathbf{M}c\Delta\mathbf{f}}{c\Delta\mathbf{f}^Tc\Delta\mathbf{f}} = \frac{\Delta\mathbf{f}^T\mathbf{M}\Delta\mathbf{f}}{\Delta\mathbf{f}^T\Delta\mathbf{f}} = R(\mathbf{M}, \Delta\mathbf{f}) \tag{20}$$

where $\Delta\mathbf{f}' = c\Delta\mathbf{f}$ with c being a nonzero constant. Therefore, we can make a fine assumption that $\Delta\mathbf{f}^T\Delta\mathbf{f} = 1$. Now, (19) becomes an extremum problem with subsidiary conditions

$$\arg\max(R(\mathbf{M}, \Delta\mathbf{f})) \;\; s.t. \;\; \Delta\mathbf{f}^T\Delta\mathbf{f} = 1. \tag{21}$$

Using the Lagrange multiplier method, we obtain

$$\mathbf{M}\Delta\mathbf{f} = \lambda\Delta\mathbf{f} \tag{22}$$

where λ is a Lagrange multiplier. Therefore, the matrix $\mathbf{M}$ that maximizes the Rayleigh quotient has an eigenvector equal to $\Delta\mathbf{f}$, with λ being the corresponding eigenvalue. The Rayleigh quotient has a property that $\lambda_{\min} \leq R(\mathbf{M}, \Delta\mathbf{f}) \leq \lambda_{\max}$, where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalues of $\mathbf{M}$, respectively.

Thus, if we let $\mathbf{A}^T_{\max}\mathbf{A}_{\max} = \mathbf{M}_{\max}$ (i.e., $\mathbf{A}^T_{\max}\mathbf{A}_{\max}$ has an eigenvalue $\lambda_{\max}$ and the corresponding eigenvector $\Delta\mathbf{f}$), we have

$$\frac{\Delta\mathbf{f}^T\left(\mathbf{A}^T_{\max}\mathbf{A}_{\max}\right)\Delta\mathbf{f}}{\Delta\mathbf{f}^T\Delta\mathbf{f}} = \lambda_{\max} \tag{23}$$

$$\frac{\Delta\mathbf{n}^T\left(\mathbf{A}^T_{\max}\mathbf{A}_{\max}\right)\Delta\mathbf{n}}{\Delta\mathbf{n}^T\Delta\mathbf{n}} = \eta \tag{24}$$

where $\lambda_{\min} \leq \eta \leq \lambda_{\max}$. Therefore,

$$\begin{aligned} SNR' &= \left|\frac{\Delta\mathbf{f}^T\left(\mathbf{A}^T_{\max}\mathbf{A}_{\max}\right)\Delta\mathbf{f}}{\Delta\mathbf{n}^T\left(\mathbf{A}^T_{\max}\mathbf{A}_{\max}\right)\Delta\mathbf{n}}\right| \\ &= \left|\frac{\lambda_{\max}}{\eta}\frac{\Delta\mathbf{f}^T\Delta\mathbf{f}}{\Delta\mathbf{n}^T\Delta\mathbf{n}}\right|^{1/2} \\ &= \left|\frac{\lambda_{\max}}{\eta}\right|^{1/2}\text{SNR} \\ &\geq \text{SNR} \end{aligned} \tag{25}$$

where the equality holds if and only if $\Delta\mathbf{f} = \Delta\mathbf{n}$.

Algorithm 1 Construction of Transformation Matrix $\mathbf{A}$ That Maximizes SNR' in (17)

Input: Training data $\mathbf{x}$ with different Classes $\mathbf{X}_1$ and $\mathbf{X}_2$

Initialization:

- Feature difference $\Delta\mathbf{f} = \frac{1}{n_1}\sum_{\mathbf{x}\in\mathbf{X}_1}^{\mathbf{x}} - \frac{1}{n_2}\sum_{\mathbf{x}\in\mathbf{X}_2}^{\mathbf{x}}$, where n_1 and n_2 are sample numbers in Classes $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively.
- Two uniformly distributed random $N \times N$ matrices $\mathbf{B}_1$ and $\mathbf{B}_2$
- Matrix $\Sigma = \text{diag}(\lambda_N, \lambda_{N-1}, \ldots, \lambda_1)$, $1 \le \lambda_1 \le \lambda_2 \le \cdots \le \lambda_N$

Steps:

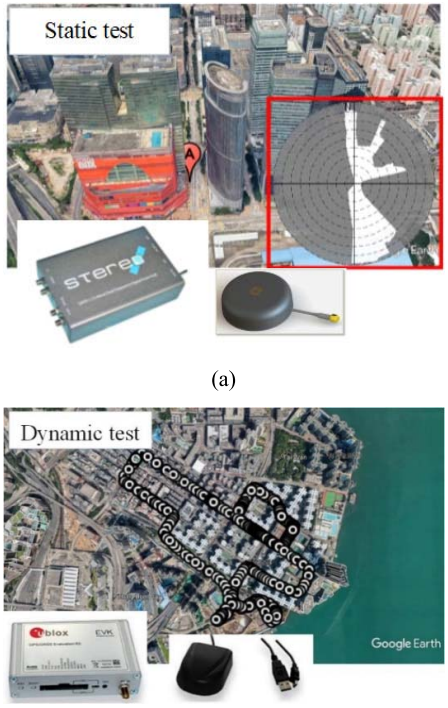
1: **for** $k = 1, 2, \ldots, M$ where M is the search number
2: Apply QR decomposition to $\mathbf{B}_1$ and $\mathbf{B}_2$:
$\mathbf{B}_1 = \mathbf{Q}_1\mathbf{R}_1$, $\mathbf{B}_2 = \mathbf{Q}_2\mathbf{R}_2$ where $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are two orthogonal matrices that have column partitions $\mathbf{Q}_1 = [\mathbf{q}_{11}, \mathbf{q}_{12}, \ldots, \mathbf{q}_{1N}]$ and $\mathbf{Q}_2 = [\mathbf{q}_{21}, \mathbf{q}_{22}, \ldots, \mathbf{q}_{2N}]$, respectively.$\mathbf{R}_1$ and $\mathbf{R}_2$ are upper triangular matrices.
3: Let $\mathbf{q}_{11} = \Delta\mathbf{f}$.
4: Apply Schmidt normalization to $\mathbf{Q}_1$:
$\rightarrow \quad \boldsymbol{\beta}_1 = \mathbf{q}_{11},$
$\rightarrow \quad \boldsymbol{\eta}_1 = \frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|} = \frac{\Delta\mathbf{f}}{\|\boldsymbol{\beta}_1\|}$
$\rightarrow \quad \boldsymbol{\beta}_2 = \mathbf{q}_{12} - \langle\mathbf{q}_{12}, \boldsymbol{\eta}_1\rangle\boldsymbol{\eta}_1$
$\rightarrow \quad \boldsymbol{\eta}_2 = \frac{\boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|} = \frac{\mathbf{q}_{12} - \langle\mathbf{q}_{12}, \boldsymbol{\eta}_1\rangle\boldsymbol{\eta}_1}{\|\boldsymbol{\beta}_2\|}$
$\rightarrow \quad \ldots$
$\rightarrow \quad \boldsymbol{\beta}_N = \mathbf{q}_{1N} - \sum_{i=1}^{N-1}\langle\mathbf{q}_{1N}, \boldsymbol{\eta}_i\rangle\boldsymbol{\eta}_i$
$\rightarrow \quad \boldsymbol{\eta}_N = \frac{\boldsymbol{\beta}_N}{\|\boldsymbol{\beta}_N\|} = \frac{\mathbf{q}_{1N} - \sum_{i=1}^{N-1}\langle\mathbf{q}_{1N}, \boldsymbol{\eta}_i\rangle\boldsymbol{\eta}_i}{\|\boldsymbol{\beta}_N\|}$
$\rightarrow \quad \mathbf{Q}_1 = [\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_N]$
5: Let $\mathbf{A} = \mathbf{Q}_2\Sigma\mathbf{Q}_1^T$, and transform $\mathbf{x}$ into a new space by $\mathbf{x}' = \mathbf{A}\mathbf{x}$.
6: Do classification on the transformed data $\mathbf{x}'$ with 2-fold using KNN.
7: **end for**

As of now, we have proven that the SNR of the target data can be improved with the transformation matrix $\mathbf{A}$, for which $\mathbf{A}^T\mathbf{A}$ has an eigenvector $\Delta\mathbf{f}$. Considering (22), we construct the matrix $\mathbf{A}$ with one eigenvector being $c\Delta\mathbf{f}$ based on the singular value decomposition (SVD) and QR decomposition methods. Detailed steps are listed in Algorithm 1.

Suppose there are n samples in the training dataset with N dimensions. The computational complexity of step 2 is $O(4 \times N^3)$. Step 4 is to apply the Schmidt normalization to $\mathbf{Q}_1$. Its computational complexity is $O(2 \times N^3)$. In the last step, the KNN is employed to benchmark the data denoising, which has a computational complexity of $O(1/2 \times n \times N)$. Since the training runs for M times, the overall computation complexity of Algorithm 1 is $O(M \times (6 \times N^3 + 1/2 \times n \times N))$.

IV. Application to GNSS Signal Classification

For urban GNSS positioning, the classification of GNSS signal types is highly desirable. Unlike features of images, human faces, and so on, the GNSS signal type features are independent of each other, causing limited performance using common techniques, such as KNN, Gauss process maximum likelihood, and SVM. This section presents an application of the proposed data denoising theory to the GNSS signal classification problem. First, features of the GNSS signal are introduced, followed by the experimental setup, data collection, and labeling. Then, we present the classification result and analysis.



(a)

(b)

Fig. 5. Location and equipment for (a) static test (Location A is one of the several locations in the test) and (b) dynamic test.

A. Experimental Setup, Data Collection, and Labeling

We conducted both static and dynamic tests. In the static test, more than 30-min intermediate frequency (IF) data were collected using the GNSS radio frequency (RF) front end at several different urban areas in Hong Kong with one example location shown in Fig. 5(a). The sky mask in the lower right corner of Fig. 5(a) indicates a very limited sky area in urban areas. The data collected were postprocessed using an open-source GPS software receiver [42], which allows extractions of correlator-level features. Table I lists the parameters of the experimental configuration. In the vehicular dynamic test, as shown in Fig. 5(b), a commercial receiver, Ublox M8T with a patch antenna, was used.

B. Features of GNSS Signal

In general, features for GNSS signal type classification are extracted from the NMEA or RINEX file that records measurements, such as pseudorange, Doppler, and the CNR [12], [14]. With the availability of GNSS raw measurements in mass-market devices, e.g., tablets and smartphones with the Android 7 operating system [43], deeper level GNSS

TABLE I
EXPERIMENTAL CONFIGURATIONS

Equipment	Parameter	Value
Antenna	Polarization	RHCP
	Low noise amplifier gain	27 dB
	Noise figure	≤2 dB
RF front end	GNSS signal	GPS L1 C/A
	Sampling rate	26 MHz
	IF	0 MHz
	Double sided bandwidth	8 MHz
	Noise figure	8 dB
	RF gain	10 dB
GPS software receiver	Correlator numbers	25
	Correlator spacing	0.05 chip
	Coherent integration time	1 millisecond

TABLE II
FEATURES USED FOR GNSS SIGNAL CLASSIFICATION

Level	Feature
RINEX	Satellite elevation angle
	Carrier-to-noise ratio
	Pseudorange residual
	Pseudorange rate residual
Correlator	Ratio between the measured maximum correlation value and the standard value
	Mean of the multicorrelator peak delay
	Variance of the multicorrelator peak delay
	Mean of the code discriminator output over 1 s
	Variance of the multicorrelator peak delay
	Early late phase

measurements are accessible, such as navigation message bits and correlation results of each channel [44]. Our previous works have explored the performance improvement of GPS signal classification using these deeper level features over the RINEX-level classifiers based on a GPS software-defined receiver (SDR) [7]. Features extracted at different levels are shown in Table II. Readers can refer to [7] for a detailed description of these features. In this article, for the static test, we use a hybrid feature vector consisting of the correlator-level features in Table II and one RINEX-level feature, i.e., satellite elevation angle. Since we have no access to the correlators within a Ublox receiver, we use RINEX-level features in the dynamic test.

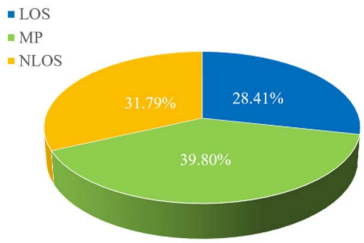


Fig. 6. Distribution of each signal type.

The signal types are labeled using an algorithm called code pseudorange double difference (DD) proposed in [7]. The code pseudorange DD observable can be expressed as [7]

$$\mathrm{DD}_{AB}^{12} = \nabla\Delta r_{AB}^{12} - \gamma_{B,\mathrm{MP}}^{1} - \gamma_{B,\mathrm{NLOS}}^{1} + \nabla\Delta e_{AB}^{12} \tag{26}$$

where $\nabla\Delta r_{AB}^{12}$ is the DD between geometric satellite-to-user range for the reference receiver, A, and rover receiver, B. The superscripts 1 and 2, respectively, denote the target satellite to be labeled and the master satellite, which has a high elevation angle and is assumed to be LOS. $\gamma_{B,\mathrm{MP}}^{1}$ and $\gamma_{B,\mathrm{NLOS}}^{1}$ are pseudorange errors caused by MP and NLOS, respectively, for satellite 1. $\nabla\Delta e_{AB}^{12}$ is the receiver noise-induced pseudorange error. A robust algorithm based on the code pseudorange DD observable is developed to label signal type in [7]. The proportion of each signal type is shown in Fig. 6, showing a relatively even distribution of different signal types. We can also see that about 70% of the measurements are contaminated by MP or NLOS receptions, which reflects the importance of signal type classification for urban positioning applications.

C. Classification Result and Analysis

1) Static Test: In this article, SVM, KNN, and NN are employed as machine learning tools. Without loss of generality, half of the samples are randomly selected as training data, and the rest half is for testing. Since the purpose of the proposed denoising algorithm is to rectify misclassified samples, we choose recall as the performance measure, which measures a model's ability to detect positive samples. In addition, considering the multiclass classification in this application, i.e., LOS/MP/NLOS classification, the microaveraging recall is selected to be the performance measure for the overall effectiveness, which is defined as [45]

$$\mathrm{Recall}_{\mu} = \sum_{i=1}^{l} \mathrm{TP}_i \Big/ \sum_{i=1}^{l} (\mathrm{TP}_i + \mathrm{FN}_i) \tag{27}$$

where l is the number of classes, TP_i denotes the number of samples that belong to Class c_i and are correctly identified, and FN_i is the number of samples that belong to Class c_i but are misclassified to other classes. For the per-class performance, the recall is used [45]

$$\text{Recall for class } c_i = \frac{\mathrm{TP}_i}{\mathrm{TP}_i + \mathrm{FN}_i}. \tag{28}$$

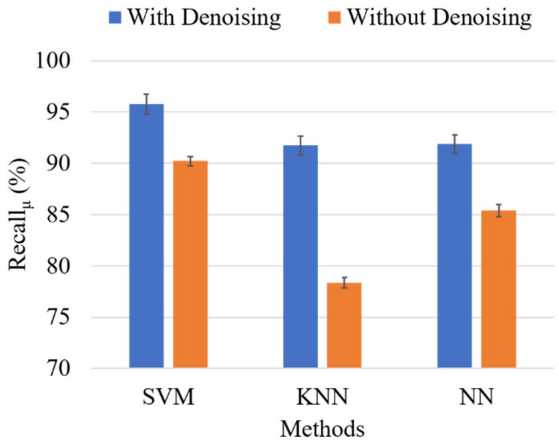


Fig. 7. Overall effectiveness of classification with and without denoising in the static test.

TABLE III
CONFUSION MATRIX FOR SVM

Signal type (number of samples)		Prediction without denoising			Prediction with denoising		
		LOS	MP	NLOS	LOS	MP	NLOS
Label	LOS	1185	108	24	1275	35	7
	MP	124	1567	154	35	1737	73
	NLOS	19	73	1382	6	4	1464

The computer program is executed ten times. The mean value of the overall accuracy with uncertainty is shown in Fig. 7. With data denoising, the overall accuracy for SVM and KNN is improved by more than 5% and 10%, respectively, with very small uncertainty. The explanation for the bigger improvement for KNN than that for SVM is that, on the one hand, KNN is a type of instance-based learning, and samples from different classes are more separable in the transformed domain in this article. On the other hand, the basic idea for an SVM learning classifier is to find the optimum hyperplane in the high-dimensional feature space that maximizes the margin between different classes and minimizes the error. The result verifies the effectiveness of the proposed data denoising approach. Compared to SVM, NN has a lower performance because of the nonlocal- and few-feature issues, as introduced in Section I. Notice that NN's performance is also improved with the proposed data denoising approach, which verifies the effectiveness of the proposed approach.

Tables III–V show the confusion matrices for SVM, KNN, and NN, respectively, from one of the ten tests. Fig. 8 is the per-class classification performance in terms of recall. The proposed data denoising approach is feasible for all signal types. A closer look at the result shows that MP has the biggest improvement over the other two types for SVM and KNN. The reason is that MP is the combination of both direct and reflected signals. As a result, the MP signal has a complicated and time-varying propagation, and its property can be similar to both LOS and NLOS signals. As can be seen in Fig. 8,

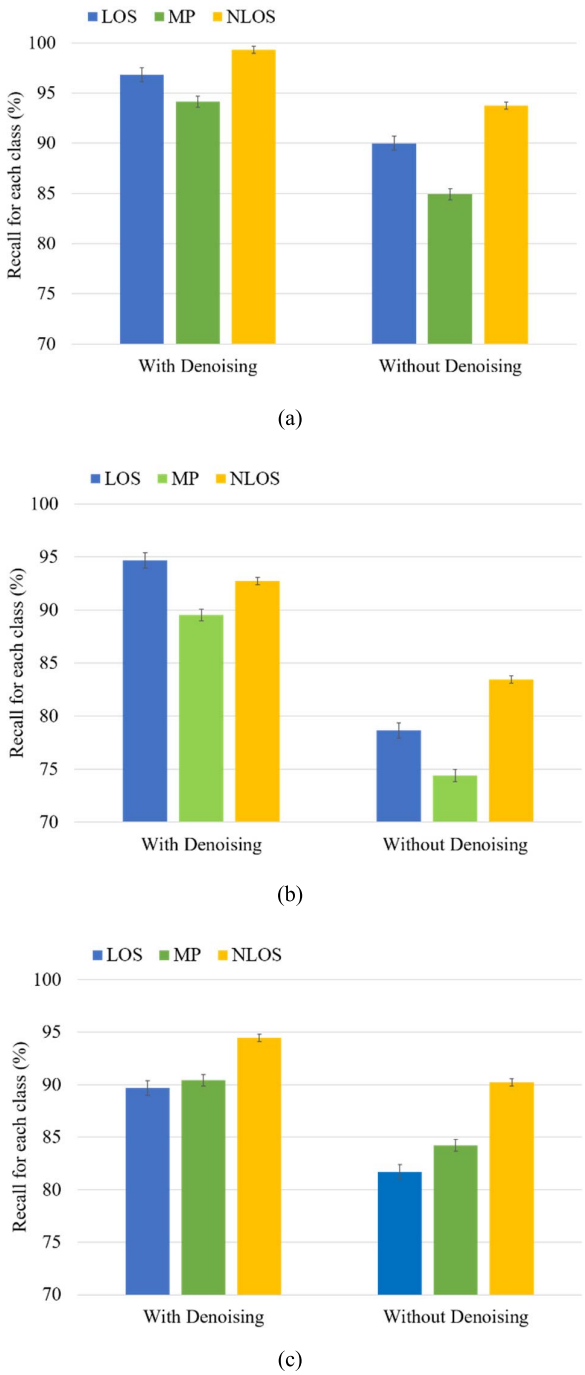


Fig. 8. Per-class classification performance using (a) SVM, (b) KNN, and (c) NN.

for both SVM and KNN, the MP is the most difficult to distinguish from the other two types. With data denoising, those misclassified MP samples are more likely to be rectified.

A deeper analysis is illustrated in Fig. 9, similar to a polar plot, where samples at the same azimuth are exactly the same used in different tests. The benefit of presenting classification results using this plot is twofold, i.e., 1) one can clearly see the classification result for a single sample in different tests and 2) one can easily find the distribution of samples that are rectified with the proposed denoising algorithm. All tests

TABLE IV
CONFUSION MATRIX FOR KNN

Signal type (number of samples)		Prediction without denoising			Prediction with denoising		
		LOS	MP	NLOS	LOS	MP	NLOS
Label	LOS	1036	267	14	1247	62	8
	MP	256	1373	216	85	1652	108
	NLOS	69	175	1230	24	83	1367

TABLE V
CONFUSION MATRIX FOR NN

Signal type (number of samples)		Prediction without denoising			Prediction with denoising		
		LOS	MP	NLOS	LOS	MP	NLOS
Label	LOS	1076	188	53	1181	115	21
	MP	163	1554	128	87	1668	90
	NLOS	41	103	1330	19	63	1392

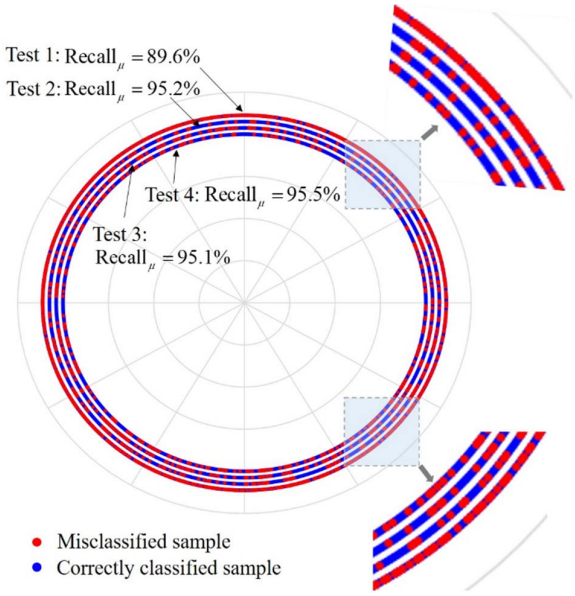


Fig. 9. Illustration of samples that are rectified with denoising for an SVM classifier. Test 1 is not using denoising, while Tests 2–4 are using denoising with a different matrix **A**.

in Fig. 9 use the SVM classifier and have the same training data. Test 1 refers to the classification without denoising, while Tests 2–4 are with denoising, and they have different values for matrix **A** due to different $\mathbf{Q}_1$'s used in Algorithm I. Notice that Fig. 9 excludes the samples that are correctly classified in all four tests; hence, one can clearly see samples that are misclassified without denoising and are now rectified with denoising. As shown in Fig. 9, the microaveraging recall is improved to about 95% in Tests 2–4 with a different matrix **A**. Taking a closer look at the result, as shown in the two zoomed-in parts in Fig. 9, we have two interesting findings. On the one hand, samples that are rectified with denoising

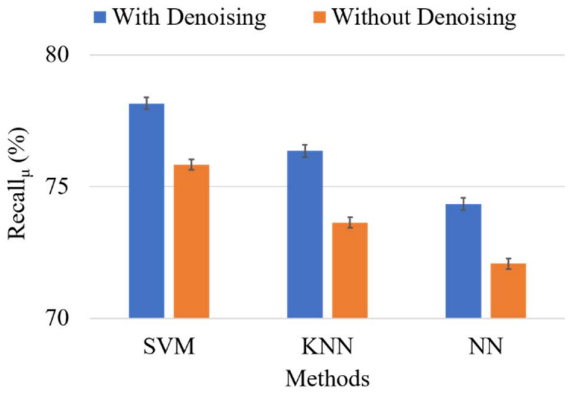


Fig. 10. Overall effectiveness of classification with and without denoising in the dynamic test.

are evenly distributed. On the other hand, for different **A**'s, different samples might be rectified. The reason is that, with different transformations, samples have different distributions in the new space. These results indicate that the proposed approach has a good generalization performance.

2) Dynamic Test: Fig. 10 shows the overall accuracy of classification results with and without denoising results in the dynamic test. Compared to the static test, the improvement of overall accuracy with data denoising over that without denoising is lower, ranging from 1.5% to 3.5%. A potential reason is that the dynamics make the GNSS measurements more complicated, which takes on a prominent position over the noise. Our proposed approach, however, is aimed to deal with the noise issue. Another reason might be that, according to the findings in [7], the correlator-level classifier outperforms the RINEX-level classifier. In the dynamic test, the commercial receiver, Ublox M8T, can only output RINEX-level measurements, whereas we have access to the correlator-level measurement in the static test. Notice that both SVM and KNN outperform NN in this test, which also verifies that the proposed data denoising is a feasible solution to nonlocal- and few-feature signal classification problems.

V. CONCLUSION AND FUTURE WORK

In this article, a new data denoising approach has been proposed by means of a series of matrix transformations. The core idea behind the proposed approach is to reduce the noise effect by projecting the target data into a new domain. With this transformation, samples for which interclass distance is comparable to noise in the original space can now be correctly classified. Rayleigh quotient is introduced to compensate for the SNR decrease due to the domain transformation. This approach is useful to solve the classification problem, in particular, for few-feature and nonlocal-feature signals, such as the GNSS signal. We conducted both static and dynamic tests using different GNSS receivers and extracted features at different levels. The results fully verify the effectiveness of the proposed method. Since our algorithm has general and simple assumptions making it available for most of the datasets, lane detection, and face recognition, implementing our algorithm in more applications would be our another future

work, not limited to machine learning but also deep learning, such as CNN.

REFERENCES

[1] L.-T. Hsu, "Analysis and modeling GPS NLOS effect in highly urbanized area," *GPS Solutions*, vol. 22, no. 1, p. 7, Jan. 2018.

[2] L. Cheng, K. Wang, M. Ren, and G. Yan, "Adaptive filter approach for GPS multipath estimation under correntropy criterion in dynamic multipath environment," *IEEE Trans. Signal Process.*, vol. 67, no. 22, pp. 5798–5810, Nov. 2019.

[3] G. A. McGraw, P. D. Groves, and B. W. Ashman, "Robust positioning in the presence of multipath and NLOS GNSS signals," in *Position, Navigation, and Timing Technologies in the 21st Century: Integrated Satellite Navigation, Sensor Systems, and Civil Applications*, vol. 1, Y. J. Morton, F. van Diggelen, J. J. Spilker, Jr., B. W. Parkinson, S. Lo, and G. Gao, Eds. Hoboken, NJ, USA: Wiley, 2021, ch. 22.

[4] M. S. Braasch, "Multipath effects," in *Global Positioning System: Theory and Applications* (Progress in Astronautics and Aeronautics), vol. 1, B. Parkinson, J. J. Spilker, Jr., P. Axelrad, and P. Enge, Eds. Washington, DC, USA: American Institute of Aeronautics and Astronautics, 1996, pp. 547–566.

[5] P. D. Groves and Z. Jiang, "Height aiding, C/N0 weighting and consistency checking for GNSS NLOS and multipath mitigation in urban areas," *J. Navigat.*, vol. 66, no. 5, pp. 653–669, Sep. 2013.

[6] R. Yozevitch, B. B. Moshe, and A. Weissman, "A robust GNSS LOS/NLOS signal classifier," *Navigation*, vol. 63, no. 4, pp. 429–442, Dec. 2016.

[7] Xu, Jia, Luo, and Hsu, "Intelligent GPS L1 LOS/multipath/NLOS classifiers based on correlator-, RINEX- and NMEA-level measurements," *Remote Sens.*, vol. 11, no. 16, p. 1851, Aug. 2019, doi: 10.3390/rs11161851.

[8] B. Xu, Q. Jia, and L.-T. Hsu, "Vector tracking loop-based GNSS NLOS detection and correction: Algorithm design and performance analysis," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 4604–4619, Jul. 2020, doi: 10.1109/TIM.2019.2950578.

[9] P. D. Groves and M. Adjrad, "Performance assessment of 3D-mapping–aided GNSS—Part 1: Algorithms, user equipment, and review," *Navigation*, vol. 66, no. 2, pp. 341–362, Jun. 2019, doi: 10.1002/navi.288.

[10] T. Ren and M. G. Petovello, "A stand-alone approach for high-sensitivity GNSS receivers in signal-challenged environment," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 5, pp. 2438–2448, Oct. 2017, doi: 10.1109/TAES.2017.2699539.

[11] N. Sokhandan, N. Ziedan, A. Broumandan, and G. Lachapelle, "Context-aware adaptive multipath compensation based on channel pattern recognition for GNSS receivers," *J. Navigat.*, vol. 70, no. 5, pp. 944–962, Sep. 2017, doi: 10.1017/s0373463317000121.

[12] R. Sun, L.-T. Hsu, D. Xue, G. Zhang, and W. Y. Ochieng, "GPS signal reception classification using adaptive neuro-fuzzy inference system," *J. Navigat.*, vol. 72, no. 3, pp. 685–701, vol. 2018, doi: 10.1017/s0373463318000899.

[13] Q. Liu, Z. Huang, and J. Wang, "Indoor non-line-of-sight and multipath detection using deep learning approach," *GPS Solutions*, vol. 23, no. 3, p. 75, Jul. 2019.

[14] L.-T. Hsu, "GNSS multipath detection using a machine learning approach," in *Proc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2017, pp. 1–6.

[15] T. Suzuki, Y. Nakano, and Y. Amano, "NLOS multipath detection by using machine learning in urban environments," in *Proc. 30th Int. Tech. Meeting Satell. Division Inst. Navigat. (ION GNSS+)*, Nov. 2017, pp. 3958–3967.

[16] T. Suzuki and Y. Amano, "NLOS multipath classification of GNSS signal correlation output using machine learning," *Sensors*, vol. 21, no. 7, p. 2503, Apr. 2021, doi: 10.3390/s21072503.

[17] Y. Quan, L. Lau, G. W. Roberts, X. Meng, and C. Zhang, "Convolutional neural network based multipath detection method for static and kinematic GPS high precision positioning," *Remote Sens.*, vol. 10, no. 12, p. 2052, Dec. 2018, doi: 10.3390/rs10122052.

[18] M. A. Nielsen, *Neural Networks and Deep Learning*. San Francisco, CA, USA: Determination Press, 2015.

[19] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Oct. 2015.

[20] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[21] P. Chatterjee and P. Milanfar, "Patch-based near-optimal image denoising," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1635–1649, Apr. 2012.

[22] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[23] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.

[24] G. Chantas, N. P. Galatsanos, R. Molina, and A. K. Katsaggelos, "Variational Bayesian image restoration with a product of spatially weighted total variation image priors," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 351–362, Feb. 2010.

[25] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[26] J. Han, J. Yue, Y. Zhang, and L. Bai, "Local sparse structure denoising for low-light-level image," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5177–5192, Dec. 2015.

[27] A. Kheradmand and P. Milanfar, "A general framework for regularized, similarity-based image restoration," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5136–5151, Dec. 2014.

[28] J. Pang and G. Cheung, "Graph Laplacian regularization for image denoising: Analysis in the continuous domain," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1770–1785, Apr. 2017.

[29] P. Chatterjee and P. Milanfar, "Is denoising dead?" *IEEE Trans. Image Process.*, vol. 19, no. 4, pp. 895–911, Apr. 2010, doi: 10.1109/TIP.2009.2037087.

[30] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2744–2756, Nov. 2002.

[31] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[32] E. P. Simoncelli, "Bayesian denoising of visual images in the wavelet domain," in *Bayesian Inference Wavelet-Based Models*. Cham, Switzerland: Springer, 1999, pp. 291–308.

[33] M. Harwit, *Hadamard Transform Optics*. Amsterdam, The Netherlands: Elsevier, 2012.

[34] C. Ding, T. Helleseth, T. Kløve, and X. Wang, "A generic construction of Cartesian authentication codes," *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 2229–2235, Jun. 2007, doi: 10.1109/TIT.2007.896872.

[35] K. R. Rao, M. A. Narasimhan, and K. Revuluri, "Image data processing by Hadamard–Haar transform," *IEEE Trans. Comput.*, vol. C-24, no. 9, pp. 888–896, Sep. 1975.

[36] Z. Wang *et al.*, "High-SNR spectrum measurement based on Hadamard encoding and sparse reconstruction," *Appl. Phys. B, Lasers Opt.*, vol. 123, no. 12, p. 277, Dec. 2017.

[37] J. Yue, J. Han, L. Li, and L.-F. Bai, "Denoising analysis of spatial pixel multiplex coded spectrometer with Hadamard H-matrix," *Opt. Commun.*, vol. 407, pp. 355–360, Jan. 2018.

[38] M. Harwit and N. J. A. Sloane, *Hadamard Transform Optics*. New York, NY, USA: Academic Press, 1979.

[39] A. Mrozack, D. L. Marks, and D. J. Brady, "Coded aperture spectroscopy with denoising through sparsity," *Opt. Exp.*, vol. 20, no. 3, pp. 2297–2309, 2012.

[40] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "Multiplexing for optimal lighting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 8, pp. 1339–1354, Aug. 2007.

[41] G. Strang, *Linear Algebra and its Applications*, 3rd ed. New York, NY, USA: Brooks/Cole, 1988.

[42] B. Xu and L.-T. Hsu, "Open-source MATLAB code for GPS vector tracking on a software-defined receiver," *GPS Solutions*, vol. 23, no. 2, pp. 1–9, Apr. 2019, doi: 10.1007/s10291-019-0839-x.

[43] S. Malkos, "Google to provide raw GNSS measurements," *GPS World*, vol. 27, no. 7, p. 36, Jul. 2016.

[44] P.-H. Jau *et al.*, "Adopting machine learning to GNSS positioning on MediaTek P60 platform," in *Proc. 31st Int. Tech. Meeting Satell. Division Inst. Navigat. (ION GNSS+)*, Oct. 2018, pp. 538–553.

[45] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manage.*, vol. 45, no. 4, pp. 427–437, Jul. 2009, doi: 10.1016/j.ipm.2009.03.002.

Jiang Yue received the B.S. and Ph.D. degrees from the Nanjing University of Science and Technology (NJUST), Nanjing, China, in 2008 and 2014, respectively.

He was a Lecturer with NJUST from November 2014 to August 2021. He was a Post-Doctoral Fellow with the Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University, Hong Kong, from October 2019 to August 2020. He is currently an Associate Professor with the College of Science, Hohai University, Nanjing. His research interests are signal processing and target detection.

Bing Xu (Member, IEEE) received the B.Eng. and Ph.D. degrees in network engineering and navigation guidance and control from the Nanjing University of Science and Technology, Nanjing, China, in 2012 and 2018, respectively.

He was a Post-Doctoral Fellow with the Department of Aeronautical and Aviation Engineering (AAE), The Hong Kong Polytechnic University (PolyU), Hong Kong, SAR, China, from August 2018 to April 2020. He is currently a Research Assistant Professor with the Department of AAE, PolyU. His current research interest is signal processing with applications to satellite and cellular positioning and wireless communications.

Li-Ta Hsu (Member, IEEE) received the B.S. and Ph.D. degrees in aeronautics and astronautics from National Cheng Kung University, Tainan, Taiwan, in 2007 and 2013, respectively.

He was a Post-Doctoral Researcher with the Institute of Industrial Science, The University of Tokyo, Tokyo, Japan. In 2012, he was a Visiting Scholar with the Faculty of Engineering, University College London, London, U.K. He is currently an Associate Head and an Associate Professor with the Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University, Hong Kong, SAR, China. His current research interests include global navigation satellite system positioning in challenging environments and localization for autonomous driving vehicles and unmanned aerial vehicles.

Dr. Hsu has been an Associate Fellow of the Royal Institute of Navigation (RIN) and a Council Member of the Institute of Navigation (ION) since 2019.