Applied Acoustics 223 (2024) 110079

Contents lists available at ScienceDirect

Applied Acoustics

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

An introduction to acoustic location Template-based positioning model

Hongyu You [a,*], Ming Yang [a,*], Hanchao Li [b,*], Kuo-Ming Chao [c,*], Xiang Fei [a,*]

[a] Coventry University, Coventry CV1 5FB, UK
[b] Loughborough University, London E20 3BS, UK
[c] University of Roehampton, London SW15 5PH, UK

ARTICLE INFO

Keywords:
Acoustic positioning
Indoor localisation
Template pattern matching

ABSTRACT

Indoor location information, which is an indispensable label for linking the vast amount of data circulating in the world of IoT, 5G, and AI, is essential. This paper proposes a novel template pattern matching-based acoustic positioning method, Location Template-based Positioning Model, for indoor acoustic source localisation. This innovative technology, similar to the passive SONAR systems on submarines, uses audible acoustic waves to locate acoustic sources in complex indoor environments. The proposed LTPM is a statistical model, and it calculates the coordinates of acoustic sources by matching input acoustic signals with template acoustic signals that are pre-collected from preset locations. A comprehensive overview of the principles and implementations of LTPM is provided, and its feasibility is verified through preliminary positioning tests in this paper. It is hoped to provide research references for the advancement of pattern matching-based acoustic indoor positioning technologies.

1. Introduction

Indoor positioning technologies, a cornerstone of modern engineering, hold immense practical significance. Notable examples include ultrasonic positioning, Wi-Fi network-based location fingerprint positioning, and Ultra-Wideband (UWB) positioning. In real-world positioning tasks, ultrasonic positioning and UWB positioning face challenges in indoor environments due to base station installation complexities and specific requirements on propagation paths between base stations and target objects [1]. However, the Wi-Fi network-based location fingerprint positioning technology shines in its excellent adaptability to complex indoor environments since it relies on the pre-measurement of Received Signal Strength (RSS) at preset locations instead of measuring time differences. Drawing inspiration from the ultrasonic positioning technology and the Wi-Fi network-based location fingerprint positioning technology, a Template Pattern Matching (TPM)-based acoustic positioning method which utilises audible acoustic signals as the system input and adapts to complex indoor environments is proposed and tested preliminarily in this paper.

TPM was exploited in the EU project Tangible Acoustic Interfaces for Computer-Human Interaction (TAI-CHI) to improve the two-dimensional positioning accuracy [2]. However, existing positioning research tends to achieve signal source localisation with radio networks rather than acoustic waves. Therefore, a TPM-based acoustic positioning method, Location Template-based Positioning Model (LTPM), is introduced in this paper to fill the gap. LTPM combines TPM with audible acoustic waves to realise three-dimensional acoustic source localisation in indoor environments.

In static environments, acoustic waves emitted by acoustic sources with fixed directivity patterns at different locations form different acoustic patterns due to multipath propagations. A direct quantitative reflection of different acoustic patterns is the change of acoustic features in received acoustic signals. Acoustic patterns corresponding to different locations are distinguishable when multiple signal features are defined. Acoustic sources can be located with template acoustic signals collected from preset locations, signal processing and pattern matching techniques. The entire positioning procedure is similar to the passive SONAR positioning system on submarines. This positioning method does not rely on measuring the time difference of arrival; therefore, it has excellent adaptability to complex indoor environments compared to

* Corresponding authors at: Faculty of Engineering, Environment and Computing, Coventry University, Coventry, CV1 5FB, UK (Hongyu You, Ming Yang, Xiang Fei); Institute for Digital Technologies, Loughborough University (London Campus), London, E20 3BS, UK (Hanchao Li); School of Arts, Humanities, and Social Sciences, University of Roehampton, London, SW15 5PH, UK (Kuo-Ming Chao).

E-mail addresses: youh2@uni.coventry.ac.uk (H. You), ab2032@coventry.ac.uk (M. Yang), h.li6-22@student.lboro.ac.uk (H. Li), kuo-ming.chao@roehampton.ac.uk (K.-M. Chao), aa5861@coventry.ac.uk (X. Fei).

https://doi.org/10.1016/j.apacoust.2024.110079
Received 19 December 2023; Received in revised form 19 April 2024; Accepted 14 May 2024
Available online 1 June 2024

TDOA-based ultrasonic/electromagnetic indoor positioning systems.

The correlation between locations and acoustic features is critical since the TPM-based positioning method matched the input signal with pre-collected template signals according to the defined acoustic features. The acoustic multipath effect ensures that each preset location corresponds to a set of unique acoustic features. In indoor environments, acoustic waves propagate at a certain velocity in the air and interact with substances inside the environment. Physical effects such as attenuation, refraction, diffraction, reflection, path loss, etc., occur during the propagation of acoustic waves, as shown in Fig. 1. The acoustic signal received by the sensor at a fixed position is a mixed signal, which consists of a series of acoustic waves. This physical effect is the multipath effect or acoustic reverberation [3]. The multipath effect brings undesirable signal components to the received acoustic signal, as shown in Fig. 2. As a result, specific differences exist in the acoustic signals emitted by the same acoustic source at different locations. The proposed LTPM utilises this phenomenon to locate acoustic sources. As a comparison, in TDOA-based indoor positioning studies, signal components caused by the acoustic multipath effect in the received acoustic signal are always suppressed to ensure signal consistency for accurate measurement of the time difference of arrival.

Current electromagnetic signal feature-based TPM positioning technologies use the Received Signal Strength (RSS) as the primary signal feature. Such positioning systems can locate signal sources at 500 mm intervals [4]. However, improving the matching accuracy further is difficult because the RSS values of template signals fluctuate, and the RSS values collected at adjacent locations overlap severely [5]. Multiple dimensional and dimensionless features are defined in LTPM to prevent such a problem from happening. In this way, the floating influence of individual features is suppressed, and the complementarity of acoustic features among different signal groups is utilised.

The key contributions of this study are:

- A TPM-based acoustic indoor positioning technology, Location Template-based Positioning Model (LTPM), is proposed and tested.
- A matching method that combines signal processing and machine learning is developed to replace RSS-based similarity analysis and utilise complementarity between different acoustic features.
- Prerequisites of LTPM are summarised and a brief comparison between different positioning technologies that are applicable in indoor environments is provided. In addition, the application potential of LTPM is briefly analysed.
- This paper presents the first comprehensive introduction to the TPM-based acoustic positioning technology and provides research references for subsequent studies on acoustic indoor positioning.

The paper consists of five sections. Section 1 presents the background introduction. Section 2 introduces the LTPM. Section 3 introduces details on data processing and model training. Section 4 presents the preliminary positioning tests, while Section 5 presents the conclusion and future works.

2. LTPM and template signal collection

The Location Template-based Positioning Model (LTPM) is an acoustic positioning method that determines an acoustic source's location by matching the acoustic signal generated by the acoustic source with labelled template signals in the matching database. The positioning principle of LTPM is introduced in this section.

2.1. Location Template-based positioning model

LTPM achieves acoustic source localisation by matching the input acoustic signal with template signals collected from preset locations. The red dots in Fig. 3 are preset locations in an indoor environment. An acoustic source is repeatedly placed at all preset locations to generate template acoustic signals. Template signals at each preset location form a corresponding signal group, and all acoustic signal groups compose the template signal database for pattern matching. In positioning, the positioning system identifies the signal group with the highest probability of generating the features of the input signal. Then, the coordinates of the identified signal group are determined as the coordinates of the acoustic source.

2-1 is the positioning equation for TPM-based localisation. The location of a signal source is determined by comparing the input signal with template signals, which are pre-collected at preset physical locations [6]. Due to the positioning mechanism, LTPM can only locate one acoustic source at a time. However, it is feasible to simultaneously locate two or multiple different acoustic sources with different frequency components if specific signal filtering and separation algorithms are merged into the signal processing module. The received composite signal can be processed and decomposed into two corresponding acoustic signals with bandpass filters and then input into the positioning model sequentially for coordinates calculations. However, in this study, acoustic sources with single frequencies are utilised in the illustration and tests to facilitate the introduction of LTPM.

$$V_f = \underset{i\in(1,2\cdots n),n\in N}{argmin}\left(\left\{\left|f(S_{input}) - f(S_{V_i})\right|\right\}\right) \quad (2.1)$$

where $f(S_{input})$ is the feature extracted from the input signal, e.g. RSS. $f(S_V)$ stands for the feature extracted from a template signal. n is the number of template signals. The output V_f is the location label of the matched template signal.

The implementation of LTPM consists of three steps, as shown in Fig. 4. The first step is the collection of template acoustic signals. An acoustic source is placed at a preset location, and a microphone collects the acoustic wave generated by the acoustic source. The location template provides coordinates for template acoustic signals, as shown in Fig. 5. In the test, cubic location templates are designed according to a UR-10 robot's coordinate system. In this way, the robot can precisely transport the acoustic source to preset locations, and a template signal matching database is established by merging the collected template signals with preset coordinates.

In the second step, signals in the template signal matching database

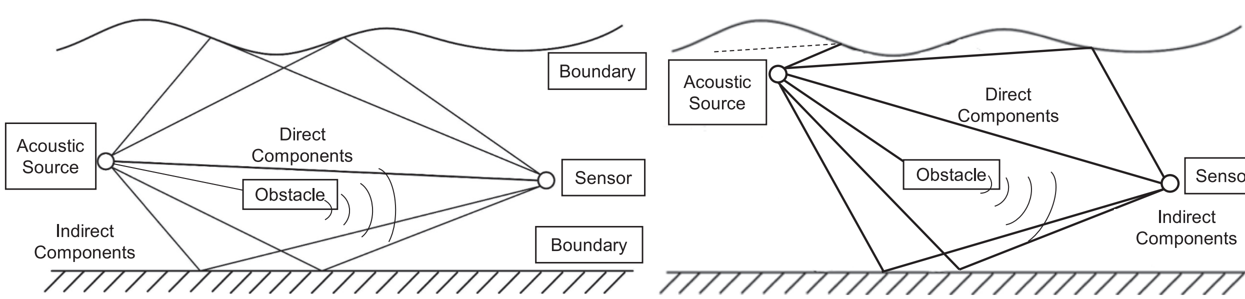

Fig. 1. The acoustic multipath effect.

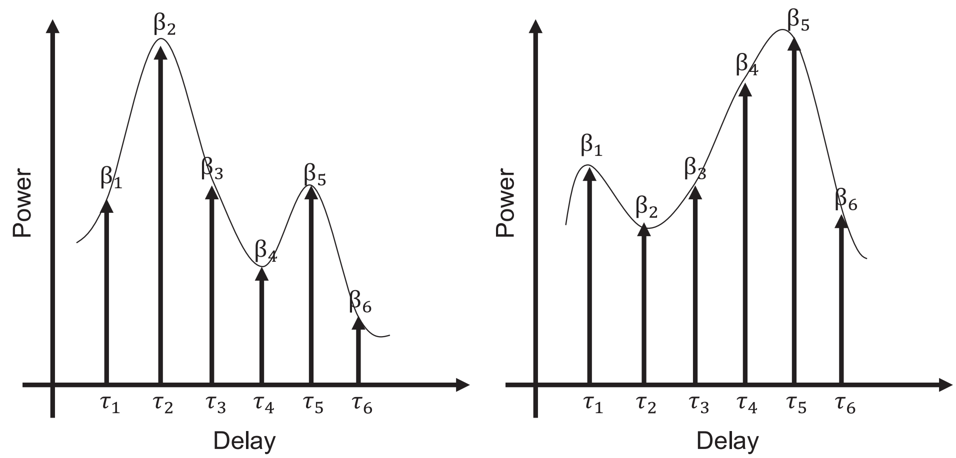

Fig. 2. The power distributions of two acoustic signals sampled by the sensor (corresponding to acoustic signals received by the sensor in Fig. 1).

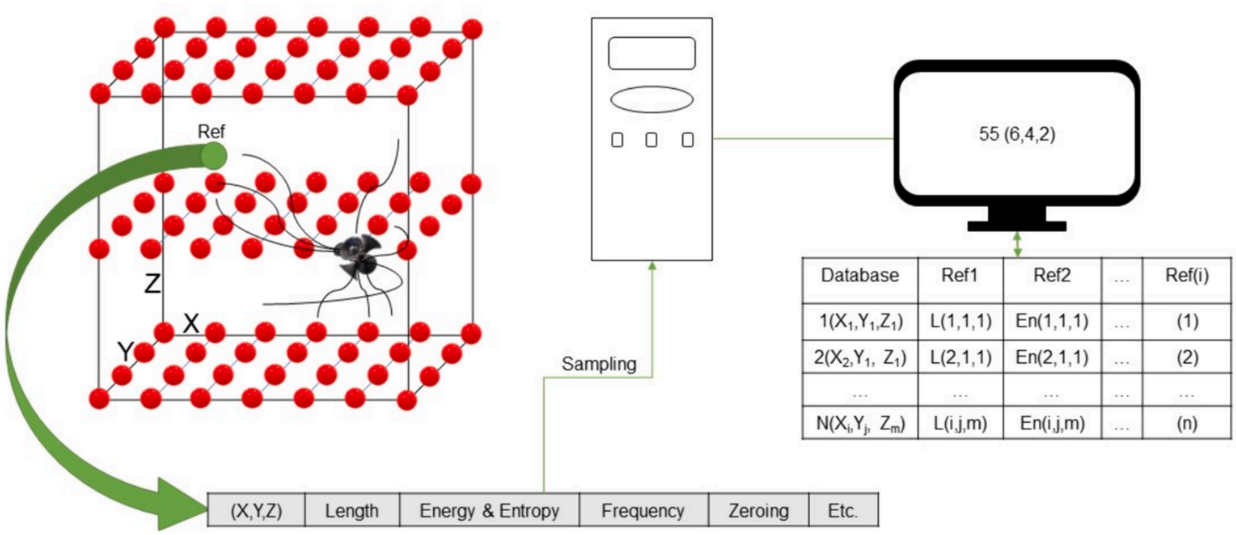

Fig. 3. The illustration of location templates and signal collection.

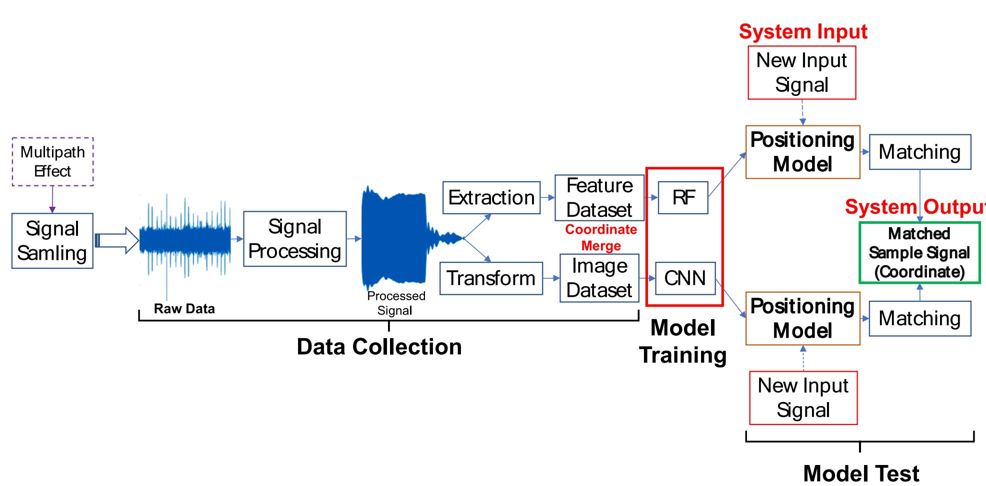

Fig. 4. The architecture of machine learning supported LTPM.

are processed to generate feature datasets and a positioning model is trained with the processed datasets. Acoustic source localisation is available after the positioning model training is completed. In the positioning test, acoustic sources are always estimated at one of the preset locations on the location template. Notably, different machine learning methods can be applied to train the positioning model

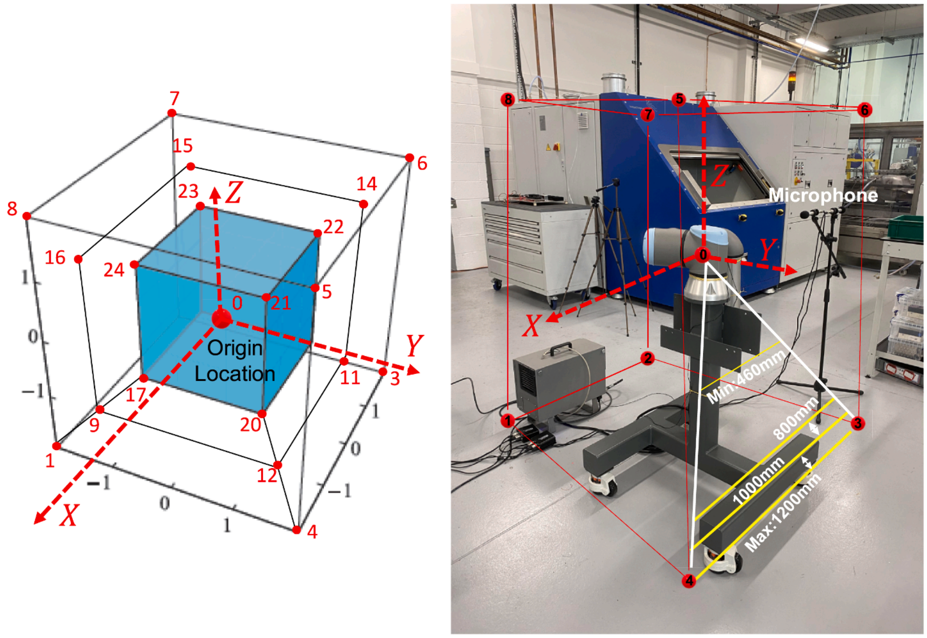

Fig. 5. The design of cubic location templates (left) and the physical setups in the positioning test (right).

according to the selected acoustic signal processing technique. For example, the random forest is used to classify acoustic features calculated with signal extraction functions. In contrast, the convolutional neural network processes signal images generated with signal transformation functions. Predictably, the positioning models trained with different machine learning algorithms have different positioning performances; thus, machine learning supported LTPM has great application potential.

The third step is to test the trained positioning model. Multiple positioning models are trained and tested with different training and test datasets to determine the positioning performance of LTPM. Positioning test results and summaries are presented in Sections 4 and 5.

2.2. Acoustic source and sampling system

In the indoor environment, the acoustic signal received by a microphone at p1 is a superimposed acoustic signal which is composed of direct signal components and multipath signal components, as shown in

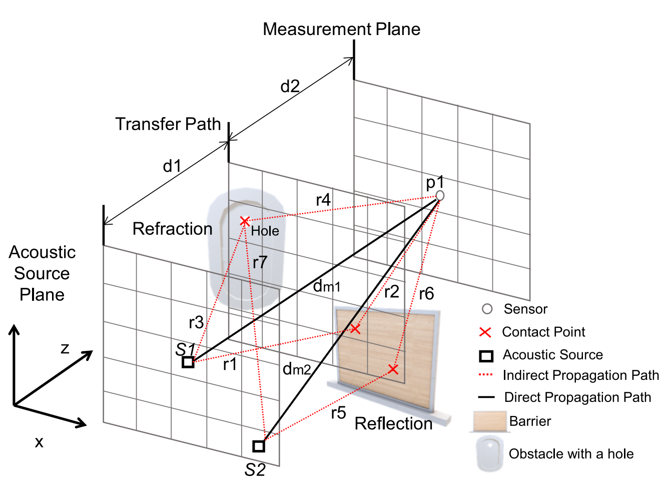

Fig. 6. The multipath propagation of acoustic waves in an indoor environment. Although the same acoustic .
Source is placed at S1 and S2, respectively, acoustic signals received at p1 and p2 differ due to the multipath propagation of acoustic waves

Fig. 6. For an acoustic source at S1, the acoustic signal received at p1 consists of three acoustic signals: Fig. 7.

$$Y_{p1}(x_1) = Y_{dm1}(x1) + Y_{r2}(x2) + Y_{r4}(x3) \tag{2.2}$$

For an acoustic source at S2, the propagation paths of acoustic waves are different. The acoustic signal received at p2 is:

$$Y_{p2}(x_2) = Y_{dm2}(x4) + Y_{r4}(x5) + Y_{r6}(x6) \tag{2.3}$$

Acoustic signals emitted by the same acoustic source at different positions differ, and the difference is utilised to match the input signal with template signals. $Y_{p1}(x_1)$ and $Y_{p2}(x_2)$ are composite acoustic signals, and they can be decomposed into plural sinusoidal waves with the Fourier Transform. On the contrary, a composite acoustic signal can be acquired by superimposing multiple acoustic signals with single frequencies and performing the Inverse Fourier Transform; therefore, if acoustic waves with single frequencies can be used for pattern matching, then acoustic waves with multi-frequencies can also be used to achieve pattern matching. Moreover, abundant feature variation patterns of acoustic waves with multi-frequencies are beneficial for improving matching accuracy. However, to focus on the introduction and implementation of LTPM, an acoustic source with a single resonate frequency is selected in this study.

The acoustic source used in the three-dimensional positioning test is a piezoelectric buzzer with a resonance frequency of 3100 Hz, and acoustic signals generated by the buzzer are isolated from the background noise with an FIR bandpass filter (2900 Hz – 3500 Hz).

The buzzer is powered by the peripheral extension interface on the robot module, and the robot transports the buzzer to preset spatial locations labelled by the location template. Meanwhile, a sampling module is deployed beside the robot module for signal collection. The sampling system consists of a GRAS AE146 microphone, an AA-12 power module, and a DAQ-2010 data acquisition card. The DAQ-2010 is configured with API functions in MATLAB. The sampling rate of the sampling system is set to 50,000 Hz.

3. Data processing and model training

3.1. Pre-processing and signal separation

Data collection and processing are completed in a closed test cell

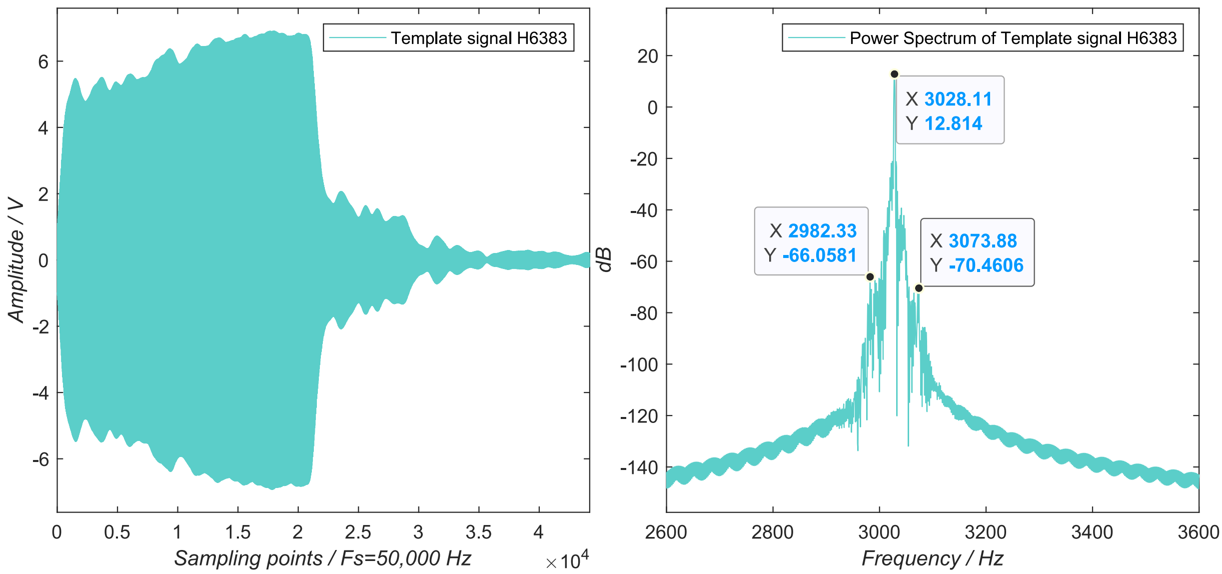

Fig. 7. A sampled signal in the time domain (left) and its power spectrum (right). The acoustic .
Source is a 3100 Hz buzzer, but it still consists of multiple signal components with different frequencies

with a noise level of 60–65 dB. The signal acquisition system consists of a measurement microphone, an amplification power supply module and a data acquisition board. Acoustic signals are sampled and transferred to the MATLAB workspace through Direct Memory Access (DMA). Digital filters, compensators, and energy-based signal separation algorithms are integrated into MATLAB for post-processing. The overall signal-to-noise ratio is maintained at 20 dB (power spectral density estimation). The acoustic environment contains realistic environmental factors, and the quality of sampled acoustic signals is sufficient to support pattern matching.

All template signals in the database and input signals need to be processed equally for feature extraction. The signal processing consists of three steps: pre-processing of signal sequences, signal separation and feature extraction, as shown in Fig. 8. The pre-processing of signal sequences aims to improve the signal-to-noise ratio of the sampled signal sequence (20 dB after compensation) and all sampled signal values are regulated within [-10,10]. The signal separation aims to separate acoustic signals from the processed signal sequences. The time-domain image of a separated signal is shown in the middle of Fig. 8.

Next, feature extraction functions extract features from the separated acoustic signals. 43 acoustic features are defined in the feature extraction function. The 43 acoustic features consist of statistical features and information theoretic features, as shown in Table 1. The statistical features are selected according to the physical properties of acoustic waves. For example, the propagation of acoustic waves is subject to the multipath effect; thus, acoustic signals received by the sensor at a fixed location have different lengths due to multipath propagations. Similarly, centroid frequency is selected because the midpoint frequencies of acoustic signals shift with different locations. Entropy features are selected because entropy represents the uncertainty of the timing signal series. For instance, the permutation entropy indicates the complexity of a time series [7] and the signal entropy represents the probability

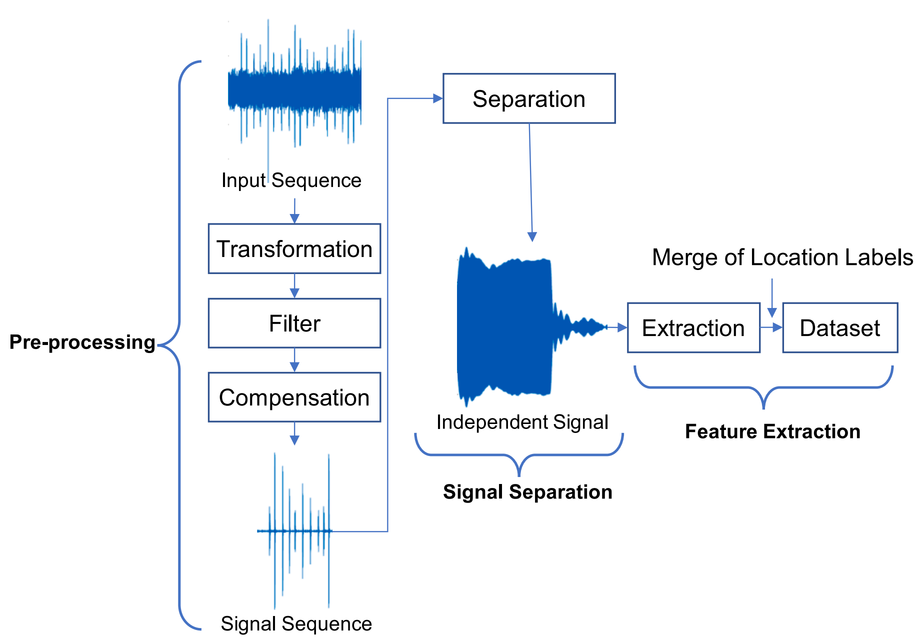

Fig. 8. Signal processing for LTPM.

Table 1
Acoustic features of LTPM.

Signal domains	Signal features
Time domain	Signal Length, Energy, Mean value, Absolute Mean value, Variance, Standard Deviation, Absolute sum, Peak value, Valley value, Valley-Peak, Median value, Kurtosis, Skew, Root Mean Square, Sin factor, Crest factor, Impulse factor, L factor, Zero Crossing Rate, Singular Value Decomposition
Frequency domain	Max Amplitude, Min Amplitude, Median Amplitude, Mean Amplitude, Valley-Peak, Centroid Frequency, Mean Square Frequency, Mean Square Root Frequency, Frequency Variance, Frequency Standard Deviation, Centroid Frequency of Kurtosis Diagram
Power Spectrum	Max Power, Min Power, Median Power, Mean Power, Centroid Frequency, Signal to Noise Ratio, Occupied Bandwidth, Pitch
Entropy	Signal Entropy, Spectrum Entropy, Sample Entropy, Permutation Entropy

distribution of the acoustic signal series [8].

Part of the features is introduced below:

A signal sequence received by the sensor is:

$$X(i) = s(i - t_d) + \eta(i) \tag{3.1}$$

whereas $s(i)$ is the source signal. t_d is the time delay. $\eta(i)$ is random Gaussian white noise. i is the length of the signal sequence.

The centroid frequency of an acoustic signal is defined as:

$$CentroidFrequency = \frac{\int_0^\infty fS(f)df}{\int_0^\infty S(f)df} \tag{3.2}$$

where $S(f)$ is the amplitude corresponding to the bin f in the Fast Fourier Transform (FFT) power spectrum. The centroid frequency describes the distribution of the signal spectrum [9].

The permutation entropy is defined as:

$$\mathrm{H_{pe}}(\mathrm{m}) = -\sum_{j=1}^{m} \mathrm{p}_j \mathrm{log} \mathrm{p}_j \tag{3.3}$$

where m is the embedded dimension. p_j is the probability function. The acoustic series is partitioned into vectors, and the permutation entropy is calculated by calculating the probability of each permutation of vectors. The permutation entropy indicates the complexity of signals [7].

43 Feature extraction functions are compiled in MATLAB. In the feature extraction stage, these functions extract the acoustic features of each separated acoustic signal. Next, the extracted features are merged with coordinates to generate input data for LTPM. The data format is shown in 3–4.

$$S_{Ni} := \begin{bmatrix} s_{11}, s_{12}, s_{13} \cdots s_{143}, coordinate1 \\ s_{21}, s_{22}, s_{23} \cdots s_{243}, coordinate2 \\ \cdots \\ s_{N1}, s_{N2}, s_{N3} \cdots s_{N43}, coordinateN \end{bmatrix} \tag{3.4}$$

The weights of these features are automatically assigned and adjusted during the training of the positioning model and examined with the MDI (Mean Decrease Impurity) feature importance measuring method, as shown in Fig. 9.

Feature reduction tests on the 43-feature set are also performed to verify the positioning performance of LTPM after removing some features with low weights. Test results in Fig. 10 showed that the model maintains the same level of classification accuracy when the feature dimensionality decreases to 25. However, the computational complexity of the model is not improved significantly after removing some features; thus, all features are listed and calculated in this study.

The variation pattern of each feature set represents the inherent pattern of acoustic signals at the corresponding preset physical location. As a preliminary test, 7 location templates (56 locations) are established in this study, and the total number of collected signals is 123,500.

The signal database, which contains more than 32,000 signals, is a primary database. The numbers of signals of the three primary databases are presented in Table 2.

The signal database, which contains less than 3,200 signals, is a secondary database. The numbers of signals of the four secondary databases are presented in Table 3.

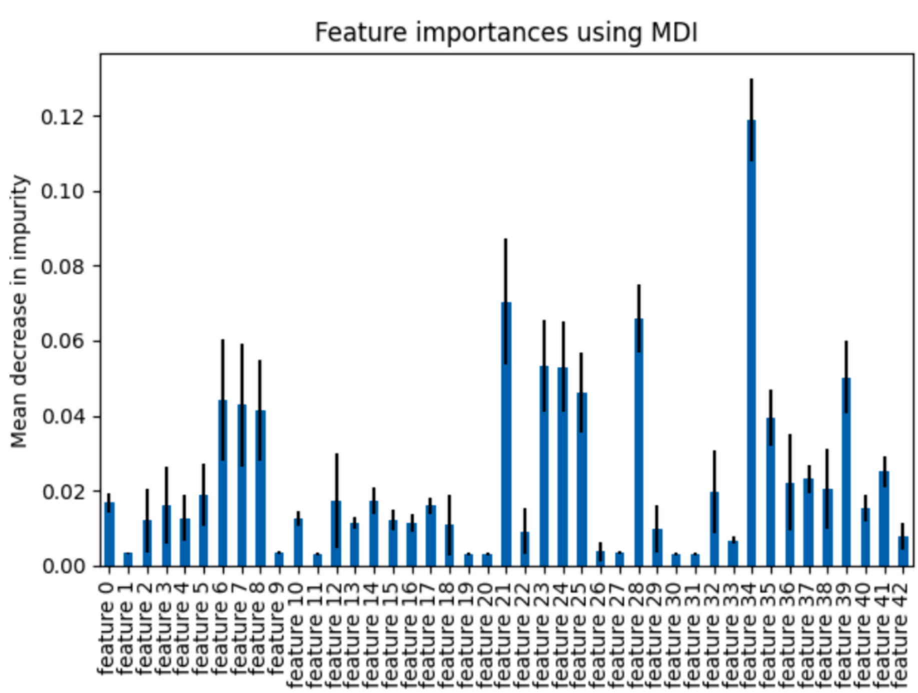

Fig. 9. The ranking of numerical features. The Mean Decrease Impurity (MDI) uses in-sample (IS) measurements to estimate feature importance for tree-based classifiers.

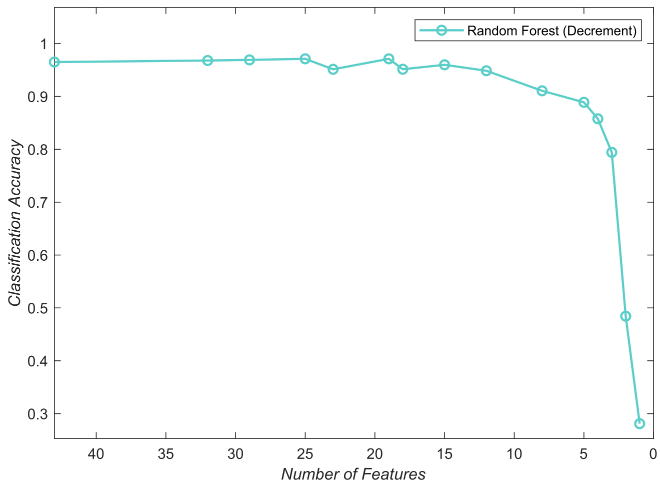

Fig. 10. Feature reduction test result. Although features with low importance have a limited impact on the overall classification accuracy, the classification accuracy begins to decrease significantly when the number of features reduces to 15.

Table 2
The three primary databases.

The side length of the location template	1200 mm	1000 mm	800 mm
Number of samples	48,000	40,000	32,000

Table 3
The four secondary databases.

The side length of the location template	460 mm	750 mm	900 mm	1050 mm
Number of samples	3,200	100	100	100

3.2. Positioning model

The second step of LTPM is to classify the data collected at preset locations with the machine learning algorithm. Random Forest (RF) was selected for the training of the acoustic positioning model since Random Forest and TPM have consistent requirements for the input data. That the input data should contain both features (acoustic features) and labels (coordinates). The RF-based positioning model is continuously optimised to improve the matching efficiency and accuracy between features and labels during the training so that the trained positioning model is able to predict the location label of the input signal.

When An input signal s_r is provided to the positioning model, features of the input signal are extracted as $F\{f_1, \cdots, f_{43}\}$. Then these features are sent to the trained positioning model as a system input as shown in 3–5. Next, the positioning model calculates a system output (a location label) according to the system input. This paper focuses on introducing the TPM-based positioning method rather than machine learning since RF is introduced as a classification tool; thus, the parameters and structures of machine learning will not be listed and discussed.

$$L_{predict} = C^{*}(F\{s_r\}) \tag{3.5}$$

where C^{*} is the trained classification model and L is the system output.

Training a Random Forest classification model is computationally intensive as it involves constructing multiple decision trees given the training data. In contrast, during inference, the positioning execution is less computationally intensive as the model makes predictions on new input data by simply passing it through each decision tree and then aggregating the results. Although training the model involves forward and backward passes through the network, the model only performs forward passes to predict locations during inference. In this study, the LTPM is trained offline; thus, the computational complexity of the positioning models for online inferencing is acceptable.

LTPM is essentially a statistical model rather than a metric location algorithm. The output of a trained positioning model is one of the preset spatial coordinates on the location template; thus, the output has only two conditions: true localisation or false localisation. In traditional positioning methods, positioning accuracy refers to the closeness between the measured location and the true or correct location of a point of interest. Still, the concept of positioning accuracy does not apply to LTPM.

In this paper, classification accuracy is used to reflect the positioning performance of LTPM. The output of LTPM is determined as a True Positive (TP) result if the coordinates of the test acoustic signal are calculated correctly by the positioning model. Otherwise, this localisation is a False Positive (FP) result. The classification accuracy is the proportion of the total number of predictions that were correct.

$$ClassificationAccuracy = \frac{TP}{TP + FP} * 100\% \tag{3.6}$$

Classification accuracy is one of the evaluation metrics of statistics. In this study, the classification accuracy of LTPM is jointly determined by the design of location templates (i.e. the number of preset locations in the space, the density of preset locations, the interval between location templates, etc.), data (features), data volume, and the classification model.

The RSS and Wi-Fi network-based positioning technology is adaptable to complex indoor environments. The proposed LTPM inherits this advantage and has better data efficiency. The location label of the input signal is estimated with the data inference model in LTPM according to a series of features extracted from the input signal. In RSS and Wi-Fi network-based positioning, the matching algorithm searches for the most appropriate signal strength range that covers the signal strength of the received signal. Compared to the RSS and Wi-Fi-based positioning technology, the positioning model trained with acoustic data locates acoustic sources more accurately with audible acoustic waves, as shown in Table 4.

The RSS-based positioning technology has a high classification accuracy at long distances due to the severe signal attenuation. In 2003, Ahonen and Eskelinen proposed to locate mobile phones with 3G UMTS networks. A classification accuracy of 67 % was achieved by using RSS and 3G mobile signals at gap distances of 25 m between preset locations. The accuracy increased to 95 % when the interval between preset locations was set to 188 m [10]. Nowadays, the positioning accuracy of the RSS-based positioning technology has been improved to 500 mm – 1500 mm. The classification accuracy of the RF-based LTPM has dramatically improved compared to the RSS and Wi-Fi-based positioning technology, as shown in Table 4. This advantage enables LTPM-based positioning systems to locate acoustic sources at gaps of 173 mm.

The positioning accuracy is still available in RSS-based positioning systems since the relationship between RSS values and distances is linear [11]. The average RSS attenuation (in dB) at different distances from the signal source is proportional to the logarithm of the propagation distance. Therefore, the locations of signal sources not at preset locations can be roughly predicted according to the changes in RSS values. In this study, the same techniques have been applied to LTPM, but test results show that the changes in multiple feature patterns in the database are not proportional to distances; thus, acoustic sources that are not at preset locations can hardly be predicted. For more details, please refer to the test results in Section 4 and the conclusion in Section 5.

3.3. Prerequisites of LTPM

Overall, the prerequisites for the implementation of LTPM include location template design, data collection, data processing, and

Table 4
Comparison between different localisation technologies.

Positioning Technology	Method	Application Range	Transmitter Dependence	Medium	Positioning Accuracy	Classification Accuracy	Sensor	Number
Location Template-based Positioning Model (Random Forest)	Audible Acoustic Machine Learning-based Pattern Matching	3D Room Scale	No	Air	N/A	98.9 % at 173 mm interval	Microphone	1
RSS and Wi-Fi-based Location Template Matching	Electromagnetic Pattern Matching	3D Room Scale	Yes	Air	500 mm	70 % – 90 % at 500—1500 mm intervals	Wireless Switch	2 or more
3D Ultrasonic Positioning	Ultrasonic TOA/ TDOA	3D Room Scale	Yes	Air	20–50 mm	N/A	Transducer & Receiver	**Multiple**
TAI-CHI	Audible Acoustic TDOA	2D Surface	No	**Solid**	14 mm	N/A	Piezoceramic Sensor	4
UWB	Electromagnetic TDOA	3D Room Scale	Yes	Air	225 mm	N/A	Transceiver	**5 or more**

positioning model training. LTPM requires a certain amount of acoustic template signals from preset spatial locations for model training. To deliver the research objectives, an industrial robot and a high-performance computer are integrated into the positioning system to perform repetitive sampling and data processing in this study.

Besides, although obstacles between the signal source and the microphone will not affect the positioning performance of LTPM, test results indicate that the positioning performance of LTPM will decrease if objects with large surfaces, such as meeting tables/blackboards in the environment, are moved after the training completes. This is related to the disappearance of multipath components in the received signal. The positioning model and training datasets must be calibrated and re-collected to adapt to the changed environment. Therefore, a static environment is critical for LTPM to maintain a stable positioning performance.

4. Test design

4.1. Test objectives

Different types of localisation tests are designed according to the practical needs of indoor localisation, including:

i. Verify the feasibility of LTPM with indoor positioning tests and measure the classification accuracy of LTPM in situations that acoustic sources are at preset locations.
ii. Determine the positioning performance of LTPM in situations that acoustic sources are not at preset locations.

Predictably, the positioning performance of LTPM will decrease when acoustic sources deviate from preset locations. The linear distance between the acoustic source and the closest preset location is defined as the deviation distance, as shown in Fig. 11. To fulfil the second objective, two secondary test objectives are established:

i. Verify whether LTPM can locate acoustic sources to the closest preset locations.
ii. Determine the relationship between the classification accuracy of LTPM and the deviation distance.

4.2. The single location template positioning test

The Single Location Template (SLT) positioning test aims to evaluate the classification accuracy of LTPM in situations that acoustic sources are at preset locations, as shown in Fig. 12. This test consists of five independent tests. The training and test data in each independent test are collected from the same location template. The first four tests correspond to the 1200 mm, 1000 mm, 800 mm, and 460 mm datasets. For each test, the input data is divided into a training dataset and a test dataset with a training-to-test ratio of 4:1. Each model has eight outputs corresponding to the eight vertex locations. Test results are presented in Table 5.

In the fifth test, the 1200 mm, 1000 mm and 800 mm datasets are mixed. Overall, 120,000 samples from the three primary datasets are divided into a training dataset containing 96,000 samples and a test dataset containing 24,000 samples. Since all three location templates are used, the LTPM has 24 outputs. The test result is shown in the last

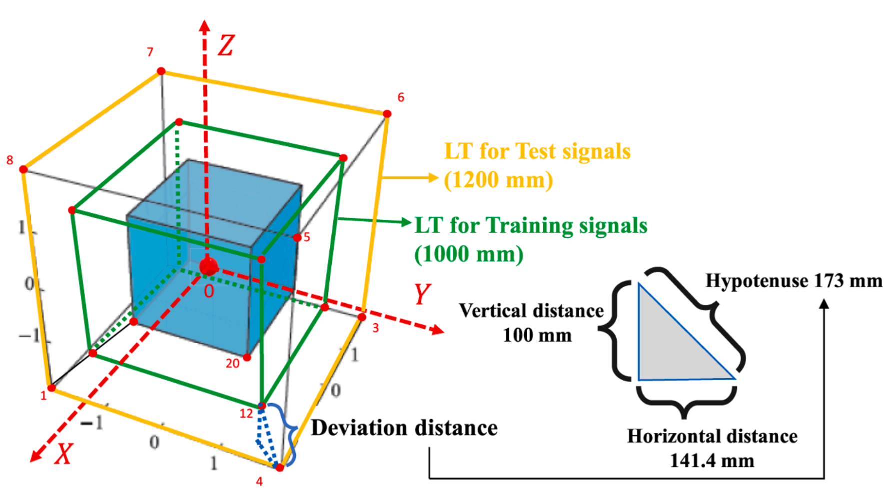

Fig. 11. The illustration of deviation distance.

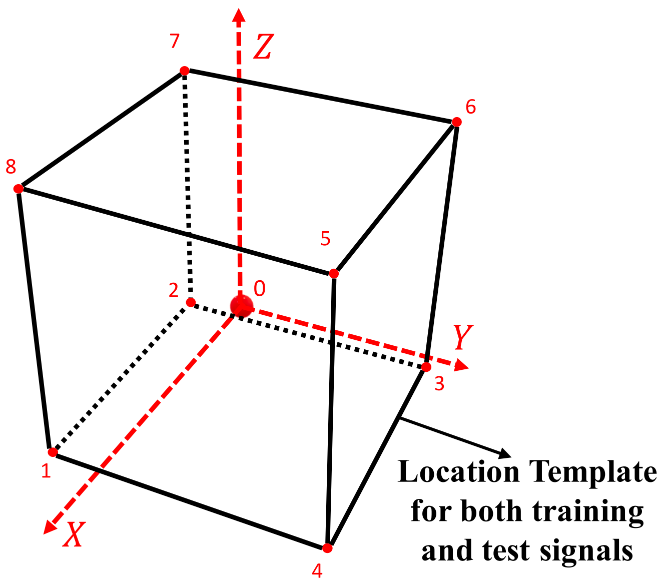

Fig. 12. The location template for the SLT test.

Table 5
The SLT test results.

	LTPM-8 (1200)	LTPM-8 (1000)	LTPM-8 (800)	LTPM-8 (460)	LTPM-24
Minimum Interval between Acoustic Sources	1200 mm	1000 mm	800 mm	460 mm	173 mm
Classification Accuracy	98.21 %	96.47 %	98.90 %	88.66 %	98.90 %

column of Table 5.

The RF-based LTPM maintains a classification accuracy of over 96 % in the first three SLT localisation tests. In the fifth test, the RF-based LTPM reaches a classification accuracy of 98.9 %. These test results indicate that LTPM locates acoustic sources accurately in the situation that acoustic sources are at preset locations. In contrast, the classification accuracy of the LTPM trained with the 460 mm dataset decreases to 88.66 % since the 460 mm dataset only contains 3,200 signals, indicating that the classification accuracy of LTPM varies with the data volume.

The test results preliminarily verify the feasibility of LTPM and that LTPM accurately locates acoustic sources at a three-dimensional interval of 173 mm.

4.3. The multiple location template (MLT) positioning test

The MLT localisation test consists of two secondary tests. The purpose of the first secondary test is to verify the positioning performance of LTPM in the situation that acoustic sources are not at preset locations, as shown in Fig. 13. The second secondary test aims to determine the classification accuracy of LTPM at different deviation distances.

In the first secondary test, one of the three primary datasets is used as the test dataset to test the LTPM trained with the other two datasets. The LTPM has 16 outputs since the training data are from two location templates. In the first subtest, a positioning model is trained with the 1000 mm and 1200 mm datasets then tested with the 800 mm dataset. In the second subtest, a positioning model is trained with the 800 mm and 1200 mm datasets then tested with the 1000 mm dataset. The second subtest is unique since it aims to determine the classification accuracy of LTPM when an acoustic source is equidistant from two preset locations. In the third subtest, the positioning model is trained with the 800 mm

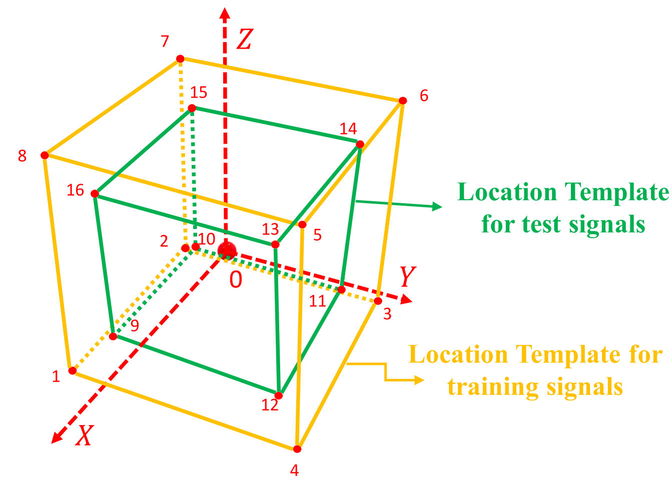

Fig. 13. Location templates for MLT tests.

and 1000 mm datasets and then tested with the 1200 mm dataset. The test results are shown in Table 6.

In MLT tests, the training data and the test data are collected separately from different location templates. The matching accuracy of the positioning model decreases inevitably because the trained positioning model does not contain the mapping relationship between the test data and preset locations. Since the 1000 mm location template is equidistant from the 800 mm location template and the 1200 mm location template, the classification accuracy increases slightly from 4.74 % to 8.67 %. But the overall low classification accuracy of the RF-based LTPM indicates that the RF-based LTPM cannot accurately locate acoustic sources to the nearest preset location.

The second secondary test consists of five subtests. The 1000 mm dataset is used as the training data. The trained LTPM is then tested with five test datasets. In the first test, the trained LTPM is tested with the 460 mm dataset. In the second test, the same LTPM is tested with the 750 mm dataset. The same LTPM is tested with the 800 mm, 900 mm, and 1050 mm datasets in the third, fourth, and fifth tests. The LTPM used in the tests has eight outputs corresponding to the eight cubic vertices of the 1000 mm location template, and test results are shown in Table 7 in descending order of the deviation distance.

Similar to the previous test results, the classification accuracy of LTPM is 15.65 % when the deviation distance is 468 mm. The classification accuracy further decreases to 5 % as the deviation distance decreases to 217 mm. Then the classification accuracy of LTPM gradually increases as the deviation distance decreases and stabilises at 12.5 %. These results further verify that patterns of acoustic signals emitted by an acoustic source at distant locations differ. Millimetre-level deviations will cause corresponding changes in acoustic features, as shown in Fig. 14; thus, LTPM cannot accurately match the input acoustic signals since the training database excludes acoustic signals collected from the deviated locations.

5. Conclusion

This paper gives a comprehensive introduction to LTPM. The

Table 6
Test results of the first secondary test.

	16 locations 800 −(1000 + 1200)	16 locations 1000 − (800 + 1200)	16 locations 1200 − (800 + 1000)
Deviation Distance	173 mm / 346 mm	173 mm	173 / 346 mm
Classification Accuracy	4.74 %	8.67 %	0.00 %

Table 7
The results of the second secondary test.

	8 locations 460 mm – 1000 mm	8 locations 750 mm – 1000 mm	8 locations 800 mm – 1000 mm	8 locations 900 mm – 1000 mm	8 locations 1050 mm – 1000 mm
Deviation Distance	468 mm	217 mm	173 mm	87 mm	43 mm
Classification Accuracy	15.65 %	5.00 %	5.01 %	12.50 %	12.50 %

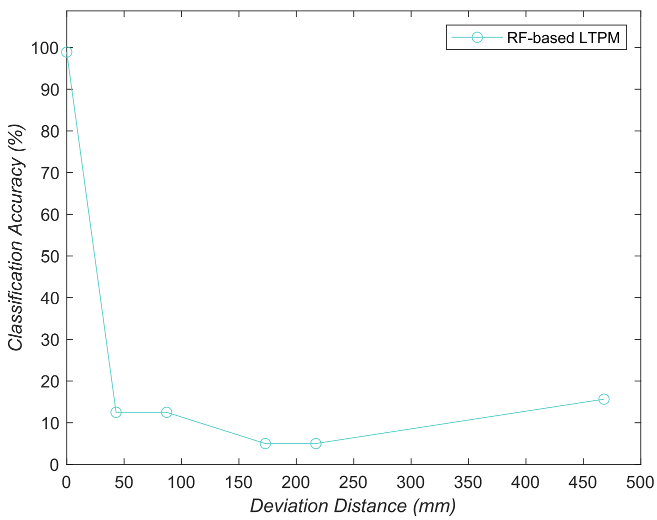

Fig. 14. The relationship between classification accuracy and deviation distance.

implementation of the proposed acoustic positioning method is divided into three steps: the data acquisition from preset locations, positioning model training, and positioning model test. The proposed LTPM is a statistical model that achieves acoustic source localisation by matching the input acoustic signal with template signals collected from preset locations; thus, the classification accuracy reflects the positioning performance of LTPM.

For acoustic sources at preset locations: The LTPM trained with 96,000 signals locates acoustic sources at 173 mm intervals with 98.9 % classification accuracy. Since test acoustic sources are placed at preset locations, the LTPM has a positioning error of 0 mm if the output of LTPM is true positive.

For acoustic sources that deviate from preset locations, The classification accuracy of LTPM is severely degraded. The classification accuracy decreases to 12.5 % when the acoustic source is 43 mm from the nearest preset location. The primary causes of this phenomenon are:

i. The feature pattern of the input signal at a non-preset location is not included in the training data. Thus, the input signal can only be randomly matched to a signal group according to features with high weights, resulting in a decreased classification accuracy.
ii. The classification accuracy is subjected to the classification model, defined acoustic features and data volume. In the test, 43 features are defined as the basis for matching, and the theoretical minimum distinguishable distance of the LTPM trained with 40,000 template signals is less than 43 mm. In other words, the trained model is overfitting; it can accurately locate acoustic sources at preset locations but cannot locate acoustic sources to the nearest preset location.

The minimum distinguishable gap distance is set to 43 mm in this study because the distance is subject to the test equipment, especially the robot. The acoustic source is transported to preset locations by the robot, and the acoustic wave from the source interacts with the robot during its propagation; thereby, the proportion of multipath components in the received signal is increased, resulting in improved classification accuracy. The robot arm used in the test has a 45 mm radius. Thus, the minimum gap distance between location templates is set to 50 mm to avoid overlapping positions of the robot arm so that the impact of the robot on the test results is minimised.

5.1. Technical advantages of LTPM

The acoustic positioning technology developed by TAI-CHI transforms solid substances such as tables, walls and glasses into two-dimensional tangible interfaces [12]. In comparison, this study aims to achieve three-dimensional localisation with acoustic waves and TPM. Current positioning technologies (both ultrasonic and electromagnetic) are based on TDOA and Time of Arrival (TOA) algorithms. These technologies have certain limitations in practical use and always require the deployment of multiple transceivers [13]. Besides, the system response speed of TDOA-based positioning systems is more than twice that of LTPM since the acoustic waves emitted by transducers need to reach the target and then return.

LTPM, conversely, does not require the participation of signal transmitters since the positioning model directly infers the location label with audible acoustic signals from the acoustic source. The positioning performance of LTPM relies on the design of location templates, data, data volume and the classification algorithm. The advantages of LTPM are:

i. LTPM performs error-free positioning for acoustic sources at preset locations.
ii. LTPM achieves acoustic source localisation with audible acoustic waves. The deployment of signal transmitters is unnecessary.
iii. LTPM has low hardware costs and excellent environmental adaptability as it requires only one microphone and does not rely on measuring time differences.

In addition, the disadvantage of LTPM is that it cannot predict acoustic sources at non-preset locations, but this defect can be compensated to a certain extent by replacing or combining the random forest with other machine learning algorithms.

Unlike the TDOA-based positioning technology, LTPM is an acoustic positioning-oriented enabling technology that does not require direct paths between signal sources and sensors. This characteristic of LTPM covers potential three-dimensional interactive demands in Virtual Reality technology. In addition, the application scenarios of LTPM are not limited to indoor environments. Similar positioning methods have been applied to acoustic source localisation on battlefields. Microflown AVISA has developed a 3D acoustic situational platform to detect audible battlefield threats, and the TPM-based acoustic positioning is used to detect the incoming direction of bullets [14].

5.2. Future works

The test result shows that a classification accuracy of 98.9 % is achieved by using LTPM at gap distances of 173 mm between locations. The minimum gap distance is set to 43 mm in the positioning test. But this result is subject to the testing equipment. Further studies on the minimum distinguishable gap distance will be conducted with a smaller robot system.

On the other hand, research on the classification algorithm of the LTPM will be conducted continuously. Integrating different machine learning algorithms may enable fuzzy localisation so acoustic sources at non-preset locations can be located at the nearest preset location. The random forest algorithm used in this study will be replaced with

Convolutional Neural Network (CNN) and the positioning performance of the CNN-based LTPM will be tested and analysed in future works.

Subsequent research includes the impact of different directivity patterns of acoustic sources on the positioning performance and the synchronous localisation of two or multiple acoustic sources to improve the practicality of LTPM.

CRediT authorship contribution statement

Hongyu You: Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis. **Ming Yang:** Writing – review & editing, Supervision, Project administration. **Hanchao Li:** Software. **Kuo-Ming Chao:** Writing – review & editing, Supervision, Project administration, Formal analysis. **Xiang Fei:** Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] S. Holm, "Ultrasound positioning based on time-of-flight and signal strength", presented at the *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sydney, NSW, Australia, 2012, pp. 1-6, doi: 10.1109/IPIN.2012.6418728.

[2] S. -P. Kuo, B. -J. Wu, W. -C. Peng and Y. -C. Tseng, "Cluster-Enhanced Techniques for Pattern-Matching Localization Systems", in *Proc. MASS*, Pisa, Italy, 2007, pp. 1-9.9.

[3] Czapiewska A, Luksza A, Studanski R, Zak A. Reduction of the multipath propagation effect in a hydroacoustic channel using filtration in cepstrum. Sensors (Basel) 2020;20(3):751–73. https://doi.org/10.3390/s20030751.

[4] P. Li, C. Han, A. Khan, U. Raza, R. Piechocki, A. Doufexi, and T. Farnham, "Deep transfer learning for WiFi localization", in *Proc. RadarConf21*, 2001, pp. 1-5.

[5] N. Swangmuang and P. Krishnamurthy, "Location Fingerprint Analyses Toward Efficient Indoor Positioning", in *Proc. PerCom*, Hong Kong, China, 2008, pp. 100-109, doi: 10.1109/PERCOM.2008.33.

[6] Sun G, Chen J, Guo W, Liu KJR. Signal processing techniques in network-aided positioning: a survey of state-of-the-art positioning designs. IEEE Signal Processing 2005;22(4):12–23. https://doi.org/10.1109/MSP.2005.1458273.

[7] Cao WW, Tung JB, Gao VAP, Hively LM. Detecting dynamical changes in time series using the permutation entropy. Physical Review E 2004;70(4):046217. https://doi.org/10.1103/PhysRevE.70.046217.

[8] Lin J. Divergence measures based on the Shannon entropy. IEEE Trans Information Theory 1991;37(1):145–51. https://doi.org/10.1109/18.61115.

[9] Quan Y, Harris JM, Chen X. Acoustic attenuation logging using centroid frequency shift and amplitude ratio methods: A numerical study. In SEG Technical Program Expanded Abstracts 1994:8–11.

[10] Ahonen S, Eskelinen P. Mobile terminal location for UMTS. IEEE Aerospace and Electronic Systems 2003;18(2):23–7. https://doi.org/10.1109/MAES.2003.1183866.

[11] Perahia E, Stacey R. Next Generation Wireless LANs-Throughput. Robustness, and Reliability. 1st ed. Cambridge University Press; 2008.

[12] Rolshofen W, Dietz P, Schafer G. TAI-CHI: tangible acoustic interfaces for computer-human interaction. Fortschritte Der Akustik 2005;31(2):773.

[13] F. Ijaz, Hee Kwon Yang, A. W. Ahmad and C. Lee, "Indoor positioning: A review of indoor ultrasonic positioning systems", presented at the 2013 15th International Conference on Advanced Communications Technology (ICACT), PyeongChang, Korea (South), 2013, pp. 1146-1150.

[14] L. Solomon, W. Wang, and M. Häge, "Battlefield Acoustic Sensing, Multimodal Sensing, and Networked Sensing for Intelligence, Surveillance, and Reconnaissance (ISR) Applications", 2015.