**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

SOURCE LOCALIZATION ON RIGID
SURFACES FOR APPLICATIONS IN
HUMAN-COMPUTER TOUCH
INTERFACE

NGUYEN QUANG HANH

SCHOOL OF ELECTRICAL AND ELECTRONIC ENGINEERING

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

2017

Abstract

This thesis addresses the problem of impact localization on rigid surfaces using vibration signals. This allows conversion of daily objects, such as tabletops and glass panels, into human-computer touch interfaces using low-cost piezoelectric sensors. Determination of the impact location is achieved via location-template matching (LTM) and/or time-difference-of-arrival (TDOA)-based localization.

LTM algorithms utilize location-dependent features of the signal to perform localization. Extraction of such features, however, requires high signal-to-noise ratio (SNR) to avoid any noise-induced alteration to the features, which may affect the matching results. In this interdisciplinary research, we propose an LTM algorithm that is robust to band-limited noises. We first analyze the mechanical model for vibrations due to impacts on a plate surface and show that band-limited components (BLCs) of a signal can be employed as its location-dependent features. A measure for signal matching is subsequently proposed based on the cross-correlation of BLCs. The measure exploits velocity dispersion among the BLCs to discriminate the library data that matches the test signal from those that do not match. Experiment results show that under noisy environments the proposed BLC-LTM outperform existing LTM techniques when the environment noise is band-limited.

TDOA-based localization on rigid surfaces suffers from dispersion and multi-path. To mitigate the effect of these phenomena, the time of arrival (TOA) is first estimated for a component within a narrow frequency band of each sensor-

received signal. The TDOA between a sensor pair is subsequently computed as the difference in their corresponding TOAs. TOA estimation is a highly challenging task when the sensor is far from the impact location causing the signal to arrive at the sensor gradually. In this thesis, three algorithms are proposed to address the problem of gradual transition. In the proposed TOA-HAD algorithm, short-time-Fourier-transform (STFT) coefficients of the signal are first converted into Hermitian angle distributions (HADs). Although the resultant HADs vary with time, significant variation is exhibited only within the transition period. This period is subsequently determined using a proposed statistical measure, which serves to quantify the time variation of HADs. The obtained knowledge on the transition period facilitates TOA estimation even when the noise-to-signal transition is gradual. In addition, the proposed framework allows simultaneous TOA estimation across all the sensors to minimize TDOA estimation errors. Experiment results show that the proposed algorithm outperforms existing techniques for source localization on solid surfaces of different materials.

For longer propagation distances such as those often seen in large surfaces, the noise-to-signal transition occurs gradually over a relatively longer duration. Since the number of spikes introduced by noise increases with the duration of the transition period, estimation of TOA from such spurious transition is prone to error, resulting in poor subsequent TDOA estimates. In the proposed STFT-Logistic algorithm, a smooth parametrized function is employed in modeling the gradual noise-to-signal energy transition at each sensor. Specifically, the noise-to-signal transition is modeled by a four-parameter logistic function. The TDOA is subsequently estimated as the difference in time shifts of the functions fitted to the sensor data. While the TOAs estimated from existing techniques are severely affected by spurious spikes, the proposed fitting procedure involves received signal over the noise-to-signal transition period hence mitigating such detrimental effects. In the third and final algorithm, to avoid the uncertainty due to grad-

Abstract

ual noise-to-signal transition, its starting point is estimated as the arrival time. This is achieved by characterizing the arrival of each sensor-received signal by the arrival times of its frequency components. To estimate the arrival time of each frequency component, each component is modeled as a harmonic random process. A kurtosis sequence, which exhibits a sharp rising edge when the signal starts to deviate from the background noise, can then be obtained for each component. The starting point of the rising edge, in turn, is determined using the curvature of the computed kurtosis sequence. Experiment results show that the proposed algorithms significantly outperforms existing techniques which adopt the abrupt change model for TOA estimation on large surfaces.

Acknowledgements

I would like to acknowledge many people who have helped me along the course of my PhD program.

First and foremost, I would like to thank my advisor, Assoc. Prof. Andy W. H. Khong. My life as a Ph.D. candidate is full of struggles and I am grateful that Andy has always been there, pushing me through the hard times, even when I have little confidence in myself. His encouragement, guidance and support are invaluable and essential for me to reach this far.

Secondly, I would like to thank Dr. V. G. Reju for taking his valuable time to help me in doing research and proof read my works. I have learned a lot from his valuable advices and suggestions, and I am grateful for that.

I would like also to extend my thanks to our previous team member A. B. Sulaiman for his significant efforts spent on collecting the data that has been used in many of my experiments. I am also grateful to each and every member of our research team, especially Arun, Benxu, Divya, Huizhi, Rohith and Wu Kai for their open sharing of knowledge and experiences which led to finding answers to my many questions that came up during the working process.

I am especially grateful to all above mentioned for being the wonderful friends that they are.

Finally, I would like to acknowledge my family, who always have faith in me and provide me with unconditional love and support. None of my accomplishments would have been possible without them, and I love them very much.

To my family, who always believe in me

To my supervisor Andy W. H. Khong

Table of Contents

Abstract	iii
Acknowledgements	vii
Table of Contents	ix
List of Figures	xiii
Abbreviations	xvii
List of Notations	xviii
1 Introduction	1
1.1 Problem definition and motivation	2
1.2 Organization of the thesis	5
1.3 Contributions and related publications	7
I Algorithms Based on Location Template Matching	9
2 Literature Review of LTM Algorithms	10
2.1 Introduction to LTM	10
2.2 Time-reversal theory-based LTM	12
2.3 Classical plate theory-based LTM	14
2.3.1 All-pole filter model-based LTM (AP-LTM)	16

TABLE OF CONTENTS

2.3.2	Zak transform for time-frequency-based LTM (Z-LTM) . . .	18
2.4	Concluding Remark	20
3	Noise-Robust LTM Using Band-Limited Components	21
3.1	Flexural vibration in thin plates	22
3.2	The proposed BLC-LTM algorithm	23
3.3	Experiment Results	30
3.4	Concluding Remark	34
II	TDOA-Based Localization	35
4	Literature Review of TDOA-based Localization Algorithms	36
4.1	Introduction to TDOA-based localization	36
4.2	Generalized cross-correlation for direct TDOA estimation	39
4.3	Methods for indirect TDOA estimation	41
4.3.1	Wavelet analysis for TOA estimation	41
4.3.2	Kullback-Leibler discrimination information for phase tran- sit point detection	44
4.3.3	Hermitian angle-based onset detection	46
4.4	Concluding Remark	48
5	Hermitian Angle Distribution for TOA Estimation	50
5.1	Generalization of TiF-HA from a statistical perspective	51
5.2	The proposed TOA-HAD algorithm	54
5.2.1	Conversion of STFT coefficients to HADs	54
5.2.2	Change point detection	60
5.2.3	Estimation of noise and signal HADs	61
5.2.4	Noise and signal contributions in HAD	63
5.2.5	Determination of γ^t	65
5.2.6	Computational complexity	68

TABLE OF CONTENTS

5.3	Experiment results	69
5.3.1	Performance vs choice of frequency range	70
5.3.2	Performance comparison with existing methods	74
5.4	Concluding remark	75
6	Logistic Fitting for TOA Estimation	77
6.1	Time-frequency analysis and the use of a parametrized logistic function	78
6.2	Four-parameter logistic model fitting	82
6.2.1	Derivation of the cost function	82
6.2.2	Estimating the occurrence of the transition period	84
6.2.3	Estimation of γ and κ	85
6.3	Computational complexity	86
6.4	Experiment Results	87
6.5	Concluding remark	89
7	Harmonic-Process Modeling for TOA Estimation	90
7.1	Analysis of impact-induced signals	91
7.2	Kurtosis for TOA estimation	93
7.3	Arrival time estimation using curvature	97
7.4	TDOA estimation	98
7.5	Computational complexity	99
7.6	Experiment Results	100
7.7	Concluding Remark	101
8	Conclusions and Future Research Direction	102
8.1	Conclusions	102
8.2	Future Research Directions	105
A	Hermitian angle distribution estimation	107

TABLE OF CONTENTS

Author's Publication	110
References	111

List of Figures

1.1	(a) A setup for human-computer interface application where a transparent glass plate is placed above a flat-panel display. Sensors are mounted along the edges of the plate to pick up vibration signal and forward to computer for processing. (b) A low-cost Murata PKS1-4A10 piezoelectric shock sensor used in the setup.	2
1.2	Signal propagation model.	3
2.1	Block diagram of a typical LTM algorithm	11
2.2	A typical LTM setup for touch-interface applications.	12
2.3	The shaping function Λ_{lb} for the mode of vibration $(l, b) = (5, 5)$. .	15
2.4	Spectrum overlayed with all-pole model of testing vibration signals, for $\mathcal{N} = 3$	17
3.1	Proposed matching measure between the test signal $\mathbf{x}$ and a library signal $\mathbf{x}^{\mathcal{L}_q}$	24
3.2	Cross-correlations between the test signal $\mathbf{x}$ and the library signals $\mathbf{x}^{\mathcal{L}_{\tilde{q}}}$ and $\mathbf{x}^{\mathcal{L}_q}$. For $\mathbf{x}^{\mathcal{L}_{\tilde{q}}}$, which corresponds to the true source location, the peaks of the correlation functions are aligned whereas for $\mathbf{x}^{\mathcal{L}_q}$, which corresponds to a location other than that of the source, the peaks are not aligned.	26
3.3	The schematic of a typical LTM setup for touch-interface applications.	31

LIST OF FIGURES

3.4	Variation in performance of (a) BLC-LTM with change in the number of filters and (b) BLC-LTM and CC-LTM with change in the noise bandwidth.	32
3.5	Localization accuracy vs SNR on glass for (a) white noise and (b) band-limited noise, and on aluminum for (c) white noise and (d) band-limited noise.	33
4.1	A typical example of TDOA-based source localization. The impact location and sensor locations are respectively denoted by $\circ$ and $\bullet$	37
4.2	3D plot of the scalogram of the CWT in which the ridge is taken at the instantaneous frequency of 310 kHz.	43
4.3	Diagram representing difference between transit time, phase transit time and centroid transit time.	44
4.4	Left and right data frames at time index n	45
4.5	A typical sensor-received signal generated by a finger tap on a glass plate (a) in the TF domain via STFT, (b) the HAs computed from the STFT coefficients, and (c) the standard deviation of HAs across selected frequency bins.	47
5.1	The HA standard-deviation sequences, obtained via the TiF-HA method, with a zoom-in view on the transition period for signals received at sensors mounted (a) near and (b) far away from the impact location.	52
5.2	The variation of the histograms of $\{\theta(k, m) \mid k = k_b, \dots, k_e\}$ with respect to time for the signal received at a sensor mounted far away from the impact location.	53
5.3	(a) Evolution of HADs over time for a real signal generated by finger on a glass surface and (b) HADs at time frames before, during and after transition.	56

5.4	Normalized ℓ^2 -distance calculated for distribution pairs of (a) the proposed $\hat{F}_{\Theta}(\theta; m)$, (b) the samples $\{\theta(k, m) \mid k = k_b, \dots, k_e\}$ used in TiF-HA, and (c) the STFT-coefficient magnitude distributions. Each distribution pair correspond to a pair (m_1, m_2) of time-frame indices $0 \leq m_1, m_2 < M$	58
5.5	Block diagram of the proposed TOA-HAD algorithm.	60
5.6	Estimates of (a) the density $\hat{f}_{\Theta}^N(\theta)$ of noise HA and (b) the density $\hat{f}_{\Theta}^S(\theta)$ of signal HA, and their contributions in $\hat{f}_{\Theta}(\theta; m)$ at time frames (c) before transition and (d) after transition.	62
5.7	Variation in the signal contribution $\gamma^S(m)$ with respect to time. The TOA m^t is determined as when $\gamma^S(m)$ exceeds γ^t	66
5.8	Pseudo code for computing the sum in (5.6) with a computational complexity of $\mathcal{O}(\Delta k + N_{\theta})$	69
5.9	Localization of the impacts due to finger taps on glass surface using TOA-HAD. The sensor coordinates, actual impact locations and estimated locations are denoted by $\bullet$, $\circ$, and $\times$, respectively. . . .	70
5.10	Variation of RMSE and standard deviation of localization errors (STDE), measured in meter, on aluminum surface with respect to center frequency for different numbers of frequency bins. (a) RMSE when the impacts are due to metal stylus, (b) RMSE when the impacts are due to finger, (c) STDE when the impacts are due to metal stylus and (d) STDE when the impacts are due to finger.	72
5.11	Variation of RMSE and standard deviation of localization errors (STDE), measured in meter, on glass surface with respect to center frequency for different numbers of frequency bins. (a) RMSE when the impacts are due to metal stylus, (b) RMSE when the impacts are due to finger, (c) STDE when the impacts are due to metal stylus and (d) STDE when the impacts are due to finger.	73

LIST OF FIGURES

5.12	Performance comparison of the proposed TOA-HAD algorithm with that of TiF-HA using (a) RMSE and (b) STDE.	74
6.1	Signal received by a sensor placed at position (1.1 m, 0.9 m) due to an impact at location (0.3 m, 0.3 m) on a glass surface of dimension $1.2 \text{ m} \times 1.0 \text{ m} \times 5 \text{ mm}$ (a) in the time domain and (b) in the time-frequency domain.	79
6.2	Fitting of the logistic function $f(m)$ to $\chi(m)$	81
6.3	Estimating the occurrence of transition period. The points corresponding to the end-points of the transition period are denoted by $\bullet$, and the point corresponding to the minimum of the total variance $\sigma_{\chi}^2(m)$ is denoted by $\circ$	84
6.4	Performance comparison of the proposed algorithm with that of TiF-HA using (a) RMSE and (b) STDE.	88
7.1	Frequency components extracted from a typical vibration signal induced by a finger tap on glass surface.	91
7.2	Kurtosis sequences and respective curvatures	94
7.3	Frequency components extracted from a typical vibration signal induced by a finger tap on glass surface where white Gaussian noise is added to the signal at SNR of 10 dB.	96
7.4	Performance comparison of the proposed algorithm with that of TiF-HA and STFT-Logistic using (a) RMSE and (b) STDE. . . .	100
A.1	Distribution of complex uniform reference noise over the circle of radius ϑ centered at the origin on the complex plane.	108

Abbreviations

BLC	Band-limited component
CWT	Continuous wavelet transform
DZT	Discrete Zak transform
DFT	Discrete Fourier transform
FFT	Fast Fourier transform
GCC	Generalized cross-correlation
HCI	Human-computer interface
LTM	Location template matching
RMSE	Root-mean-square error
STDE	Standard deviation of localization errors
STFT	Short-time Fourier transform
TFR	Time-frequency representation
TDOA	Time difference of arrival
TOA	Time of arrival

List of Notations

Basic notation

$\ell^2(f, g)$	Euclidean distance between functions f and g
$\ z\ $	Magnitude of complex number z
z^*	Conjugate of complex number z
$*$	Convolution operator

Sets

$\mathbb{Z}^+$	Positive integers
$\mathbb{R}$	Real numbers
$\mathbb{C}$	Complex numbers

Notation for Probability and Statistics

X	Real random variable
$\underline{X}$	Complex random variable
F_X	Cumulative distribution function of X
f_X	Probability density function of X
$\Pr \{.\}$	Probability
$\mathbb{E} \{X\}$	Expectation
$\text{var} \{X\}$	Variance
$\text{median} \{X\}$	Median
$\text{kurt} \{X\}$	Kurtosis

Mechanical Properties

L_x, L_y, L_z	Dimensions of the touch interface
ρ	Mass density
$\dot{D}$	Flexural rigidity
E	Young's modulus of elasticity
ν	Poisson's ratio
μ	Absorption coefficient
$\tilde{\mu}$	Reduced absorption coefficient

Chapter 1

R	Number of sensors
N	Length of a received signal
K	Number of DFT frequency bins
(u_s, v_s)	Impact location
(u_i, v_i)	Location of the i th sensor
$s(t), s(n)$	Vibration signal generated at the impact location
$x_i(t), x_i(n)$	Vibration signal received at the i th sensor location
$h_i(t), h_i(n)$	Impulse response between (u_s, v_s) and (u_i, v_i)
$w_i(t), w_i(n)$	Background noise at the i th sensor
$d_i(u, v)$	Distance between (u, v) and (u_i, v_i)
$\Delta\varphi_i$	Phase lag between $x_i(t)$ and $s(t)$
α_i	Attenuation factor of $h_i(t)$

Chapters 2 and 3

(l, b)	Vibration mode indices
ω_{lb}	Natural frequency of vibration mode (l, b)
Q	Number of calibration points
$(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$	q th calibration point
(u_o, v_o)	Sensor location
$\delta(n)$	Dirac delta function
$x(n), x^{\mathcal{L}_q}(n)$	Test and q th library signals
$h(n), h^{\mathcal{L}_q}(n)$	Impulse responses for the test and q th library signals
$ \mathcal{P}_x(e^{j\omega}) ^2$	Spectral density of $x(n)$
$\Psi_{x, x^{\mathcal{L}_q}}(n)$	Cross-correlation between $x(n)$ and $x^{\mathcal{L}_q}(n)$.
$\mathcal{Z}_x(k, m)$	Discrete Zak transform coefficients for $x(n)$
$\sigma_x^2(k)$	Variance of $\mathcal{Z}_x(k, m)$ across time
$\mathbf{x}, \mathbf{x}^{\mathcal{L}_q}$	Vector form of test and q th library signals
$\mathbf{z}_k, \mathbf{z}_k^{\mathcal{L}_q}$	BLCs of $\mathbf{x}$ and $\mathbf{x}^{\mathcal{L}_q}$ extracted using the k th bandpass filter
$\mathbf{w}$	Background noise
$\mathbf{v}_k$	BLC of $\mathbf{w}$ corresponding to the k th bandpass filter

Chapters 4 and 5

$\tau_{i,j}$	TDOA between $x_i(n)$ and $x_j(n)$
$\mathcal{H}_{ij}$	Hyperbolic curve determined from $\tau_{i,j}$
$\Psi_{i,j}(n)$	Generalized cross-correlation between $x_i(n)$ and $x_j(n)$
$\underline{\lambda}(k)$	Generalized frequency weighting function
a, b	Dilation and translation parameters of a function
$\mathcal{W}_\psi(a, b)$	CWT of $x(t)$ with an analyzing wavelet $\psi(t)$
$\psi_G(t), \psi_M(t)$	Gabor and Morlet wavelets
$\mathbf{x}(m)$	Length- L data frame $[x(m), \dots, x(m + L - 1)]^T$
$\mathcal{P}_{\mathbf{x}(m)}(\omega)$	Periodogram of $\mathbf{x}(m)$
$\mathcal{D}_{\text{KL}}$	Kullback-Leibler information divergence
τ^{P}	Phase-transit time
$\mathcal{F}_K$	Fourier transform matrix
$\underline{x}(k, m)$	STFT of $x(n)$
$\underline{w}(k, m)$	STFT of synthetically generated reference noise
$k_{\text{b}}, k_{\text{e}}$	Starting and end bins of the selected frequency band
Δk	Number of selected frequency bins
$\Theta(m)$	Time-varying Hermitian angle random variable
$F_\Theta(\theta; m)$	Distribution function of $\Theta(m)$
$f_\Theta(\theta; m)$	Probability density function of $\Theta(m)$
$\theta(k, m)$	(Discrete) realization of $\Theta(m)$ corresponding to frequency bin k
$\hat{\mu}(m), \hat{\sigma}(m)$	Mean and standard deviation of $\{\theta(k_{\text{b}}, m), \dots, \theta(k_{\text{e}}, m)\}$

List of Notations

$\underline{W}$	Reference noise random variable on the TF domain
$\underline{X}(m)$	Signal random variable on the TF domain
$F_{\underline{X}}(x; m), F_{\underline{W}}(w)$	Distribution functions of $\underline{X}(m)$ and $\underline{W}$
$F_{\ \underline{X}\ }(z; m), F_{\ \underline{W}\ }(u)$	Magnitude distribution functions of $\underline{X}(m)$ and $\underline{W}$
$f_{\ \underline{X}\ }(z; m), f_{\ \underline{W}\ }(u)$	Magnitude probability density functions of $\underline{X}(m)$ and $\underline{W}$
τ^b, τ^e	Beginning and end points of the transition period
$F_{\Theta}^S(\theta), F_{\Theta}^N(\theta)$	HAD over the signal and noise periods
$\gamma^S(m), \gamma^N(m)$	Probability contribution of signal and noise to $f_{\Theta}(\theta; m)$
$\gamma_{-}^S, \gamma_{+}^S$	Signal contributions in $f_{\Theta}(\theta; m)$ over the noise and signal period
γ^t	Common threshold on signal contributions for TOA estimation
$\gamma_i^S(m)$	Signal contribution sequence for $x_i(n)$

Chapter 6

$\chi(m)$	Mean spectral sequence
$f(m)$	Logistic function fitted to $\chi(m)$
κ	Translation parameter of $f(m)$
γ	Parameter determining the slope of $f(m)$
A, B	Minimum and maximum values of $f(m)$
$\hat{A}^0, \hat{B}^0, \hat{\gamma}^0, \hat{\kappa}^0$	Initial estimates of A, B, γ and κ
$w(m)$	Weighting function for fitting $f(m)$ to $\chi(m)$
$\varepsilon(m)$	Residual/error in fitting $f(m)$ to $\chi(m)$

Chapter 7

K	Number of frequency components to be extracted
$z_k(n)$	Component corresponding to frequency f_k
$\underline{Z}_k(n)$	Harmonic random process modeling $z_k(n)$
$\underline{\mathbf{Z}}_k(n)$	$[\underline{Z}_k(n), \dots, \underline{Z}_k(n + L - 1)]^T$, a sliding window of size L on $\underline{Z}_k(n)$
$\kappa_2 \{\underline{\mathbf{Z}}_k(n)\}, \kappa_4 \{\underline{\mathbf{Z}}_k(n)\}$	Second and fourth cumulants of $\underline{\mathbf{Z}}_k(n)$
$\gamma_k(n)$	Kurtosis of $\underline{\mathbf{Z}}_k(n)$
$\xi_k(n)$	Curvature of $\gamma_k(n)$
τ_i^k	TOA of $z_i^k(n)$
$\tau_{i,j}^k$	TDOA between $z_i^k(n)$ and $z_j^k(n)$ for sensor pair (i, j)

Chapter 1

Introduction

Advancements in technology have been revolutionizing the way humans interact with computers. Traditional input devices such as keyboard and mouse are being replaced by new human-computer interfaces (HCIs). In addition, HCIs have become more natural towards human usage in a way where user inputs can be delivered via speech [1], facial expression [2,3], hand gesture [4,5], or finger tap [6,7]. Among existing human-centric HCIs, touch interfaces make up a significant portion of the market due to the advent of mobile devices such as smart-phones and tablets. In addition, touch interfaces nowadays are not limited to pre-designed products but extended to objects that are readily available in daily life [8] such as tabletops and glass panels. This thesis focuses on new algorithms for localization of impacts generated by fingers or styluses on the surfaces of such objects. Using low-cost sensors to capture surface vibrations, the proposed algorithms offer cost-effective and flexible means of converting large surfaces into touch interfaces.

Figure 1.1(a) illustrates the conversion of an existing flat-panel display into an HCI by adding a transparent glass panel on top. In this setup, low-cost Murata PKS1-4A10 piezoelectric shock sensors, shown in Fig. 1.1(b), are mounted on the top and along the edges of the glass panel to capture any vibration due to user interaction on the surface. The captured signals are subsequently forwarded to a computer where a localization algorithm is applied to determine the interacting point.

Figure 1.1: (a) A setup for human-computer interface application where a transparent glass plate is placed above a flat-panel display. Sensors are mounted along the edges of the plate to pick up vibration signal and forward to computer for processing. (b) A low-cost Murata PKS1-4A10 piezoelectric shock sensor used in the setup.

1.1 Problem definition and motivation

In the problem of impact localization on solid surfaces, the signal received at each sensor is subjected to distortion due to dispersion - a phenomenon which refers to the dependence of wave propagation velocity on the frequency of vibration [9, 10]. The longer the propagating distance is, the more distorted the signal becomes. Another source of waveform distortion is from the interference of waves reflected from the medium boundary, a phenomenon known as multipath [11, 12]. The amount of dispersion is dependent on properties of the material while multipath interference is dependent on the geometry of the surface being considered.

In the scope of this thesis, the surfaces of interest are rectangular in shape and are made of materials that are isotropic and homogeneous. The impact localization problem is then formulated as follows: Assume that an impact is exerted at a specific location (u_s, v_s) on the surface and this results in a vibration signal $s(t)$ that emanates from the impact location in all directions. The vibration

1.1 Problem definition and motivation

Figure 1.2: Signal propagation model.

signal propagates with speed $c(\omega)$ as a function of frequency ω , and is received by R sensors positioned at (u_i, v_i) , $i = 1, \dots, R$, on the plate surface. Denoting the impulse response between the impact location and the i th sensor location by $h_i(t)$, a vibration signal $s(t) * h_i(t)$ is induced at the sensor location, where $*$ denotes the convolution operator. The signal received by the i th sensor $x_i(t)$ is then a combination of the induced vibration and the uncorrelated environment noise $w_i(t)$. Subsequently, the analog signal $x_i(t)$ is sampled to obtain the discrete signal $x_i(n)$, $n = 0, 1, \dots, N - 1$, where N is the length of the received data. Fig. 1.2 illustrates the signal propagation model in terms of the sampled versions of $s(t)$, $h_i(t)$, $w_i(t)$ and $x_i(t)$, which are respectively denoted by $s(n)$, $h_i(n)$, $w_i(n)$ and $x_i(n)$. Assuming that each sensor has uniform gain for simplicity, the signal received at the i th sensor is given as [11]

$$x_i(n) = s(n) * h_i(n) + w_i(n). \quad (1.1)$$

Note that in a dispersive environment, due to the dependence of velocity on frequency, signal components corresponding to different frequencies require different amounts of time to travel the same distance. Denoting the distance between the i th sensor and the impact source by $d_i = \sqrt{(u_i - u_s)^2 + (v_i - v_s)^2}$,

the component of $s(n)$ corresponding to a frequency ω requires a time $t_i(\omega) = d_i/c(\omega)$ to arrive at the i th sensor. As a consequence, for each frequency ω , there exists a lag in phase $\Delta\varphi_i(\omega) = \omega t_i(\omega)$ when the corresponding component of $s(n)$ arrives at the i th sensor. The effect of the impulse response $h_i(n)$ can therefore be expressed in the frequency domain as a combination of amplitude attenuation (due to propagation) and phase lag, i.e.,

$$\underline{h}_i(k) = \alpha_i(k)e^{j\Delta\varphi_i(\omega_k)}, \quad k = 0, \dots, K-1, \quad (1.2)$$

where $\underline{h}_i(k)$ is the discrete Fourier transform (DFT) of $h_i(n)$, K is the DFT length, $0 \leq \alpha_i(k) \leq 1$ is the attenuation factor and $\Delta\varphi_i(\omega_k)$ is the phase lag corresponding to the normalized frequency $\omega_k = 2\pi k/K$. The expression of the signals received at the i th sensor in the frequency domain is then

$$\underline{x}_i(k) = \alpha_i(k)\underline{s}(k)e^{j\Delta\varphi_i(\omega_k)} + \underline{w}_i(k), \quad (1.3)$$

where $\underline{s}(k)$, $\underline{x}_i(k)$ and $\underline{w}_i(k)$ are the DFTs of $s(n)$, $x_i(n)$ and $w_i(n)$, respectively.

Approaches for estimating (u_s, v_s) from $x_i(n)$, $i = 1, \dots, R$, include location template matching (LTM) and time-differences-of-arrival (TDOA) [13–15]. The LTM methods [16–19] exploit location-dependent features of the received signals. These methods include two phases: calibration and test. During calibration, impacts are exerted at pre-defined locations on the surface. Induced vibrations captured by a surface-mounted sensor are then analyzed to obtain location-dependent features which, in turn, are stored in a library. In the test phase, the same features are extracted from the received (test) signals. These features are subsequently compared with those stored in the library. By finding the closest match, an estimate of the source location is obtained. On the other hand, for the TDOA-based approach, the TDOA information is estimated for each pair of sensors using their respective received signals. Each TDOA determines a locus of possible source locations as a hyperbolic curve on the surface, with the corresponding sensors

1.2 Organization of the thesis

being the foci. The intersection of all such curves provides an estimate of the source location. Accuracy in TDOA estimation therefore plays a crucial role in determining the localization performance.

Both TDOA-based and LTM approaches have their advantages and disadvantages. While detectable locations are limited to only the calibrated points for LTM algorithms, TDOA-based approaches can localize impacts made at arbitrary points on the surface. On the other hand, in the presence of dispersion and multipath, distortions added to the received signals depend on the sensor location with respect to the impact location. LTM algorithms can therefore utilize the location-dependent distortion to effectively distinguish vibrations originated from different locations. Such distortion, however, adversely affects the direct estimation of TDOA when using the conventional generalized cross-correlation (GCC) method [20]. While dispersion can be circumvented to some extent [21–23] by isolating frequency components prior to the application of GCC, multipath distortion remains and results in erroneous TDOA estimation. An approach to mitigate the effect of multipath propagation is to perform time-of-arrival (TOA) estimation on the portion of the signal that contains its initial arrival before obtaining TDOAs as differences in the estimated TOAs.

1.2 Organization of the thesis

This thesis addresses the problem of impact localization on rigid surfaces using both the LTM and TDOA-based approaches. The thesis is divided into two parts, with the first part describing the LTM approach and the second part the TDOA-based approach.

In the first part, Chapter 2 reviews the background and existing LTM algorithms for HCI applications. A new LTM algorithm is proposed in Chapter 3. In this algorithm, by analyzing the mechanical model for vibrations due to impacts on plate surface, band-limited components (BLCs) of a signal are shown to be

capable of being employed as its location-dependent features. The characteristics of BLCs are also taken into account to derive a measure for signal matching based on the cross-correlation of BLCs.

The second part of this thesis is dedicated to TDOA-based localization algorithms. Existing methods including direct and indirect TDOA estimation are reviewed in Chapter 4. In Chapter 5, a new TDOA-based algorithm is proposed based on the conversion of the signal representation in the time-frequency domain into distributions of Hermitian angles (HADs). The signal-to-HAD conversion is achieved by combining the spectral density of the signal with the distribution of a reference noise. New measures are also proposed to compute the probability contributions of the background noise and vibration signal to each HAD. By tracking these time-varying contributions, the TOA can be determined as the instant when the signal contribution dominates that of the noise. The proposed framework also allows simultaneous TOA estimation across all the sensors to minimize TDOA estimation errors.

In Chapters 6 and 7 localization algorithms for larger surfaces are proposed. In Chapter 6, a smooth parametrized function is employed to model the gradual noise-to-signal energy transition at each sensor. Specifically, the noise-to-signal transition is modeled by a four-parameter logistic function. The TDOA is then estimated as the difference in time shifts of the functions fitted to the sensor data. In Chapter 7, each frequency component of the signal is modeled as a harmonic random process. A kurtosis sequence, which exhibits a sharp rising edge when the signal starts to deviate from the background noise, can then be obtained for each component. By accurately estimating the starting point of the rising edge as the arrival time, the proposed algorithm can avoid the uncertainty due to gradual noise-to-signal transition.

Finally, Chapter 8 concludes the thesis and discusses future research directions.

1.3 Contributions and related publications

This section summarizes the contributions made by the author in the LTM (Chapter 3) and TDOA-based (Chapters 5, 6 and 7) approaches.

The novelty of the proposed BLC-LTM algorithm in Chapter 3 is the utilization of propagation modes of the vibration signal in location matching. It will be shown in Chapter 3 that an arbitrary combination of these modes is unique to each location. Such a combination, which can be extracted from the signal as a band-limited component (BLC) using a filter, can therefore serve as a location-dependent feature and be employed for LTM. An additional contribution is the exploitation of the dispersiveness among these components to construct a cross-correlation-based matching measure. In the presence of band-limited noise, the use of multiple BLCs with the proposed measure allows robust location matching. The publication arising from this work [C1] was awarded the Best Paper Award at the 10th International Conference in Information, Communications and Signal Processing (ICICS 2015).

In Chapter 5, the contribution lies in the proposed framework for the conversion of signal to HADs. Although the resultant HADs vary with time, significant variation is exhibited only within the transition period. The variation of HADs, quantified by a proposed statistical measure, facilitates TOA estimation even when the noise-to-signal transition is gradual. In addition, the proposed framework allows *simultaneous* TOA estimation across all the sensors. This effectively reduces the variance of TDOA estimates compared to existing techniques where TOAs are estimated *independently*, since, for the latter, the variance of the TDOA estimate for each sensor pair equals the sum of the variances of the respective TOA estimates. This work has been submitted to IEEE Transactions on Multimedia and is currently under review [J1].

In Chapter 6, a novel approach for TDOA estimation is introduced, where TOA estimation is not required. The gradual noise-to-signal transition at each

sensor is modeled by a smooth parametrized function, and TDOAs are estimated as the differences in lateral shift of the fitted functions. In addition, the proposed fitting procedure involves the received signal over the entire transition period. This mitigates the detrimental effects of spurious peaks due to noise such as often seen in existing algorithms for TOA estimation. The work in this chapter was presented at the 41st IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016), and was published as [C2].

In Chapter 7, to circumvent the uncertainty incurred by the gradual transition, the proposed algorithm attempts to estimate the starting point of the transition period. Procedures to improve TOA/TDOA estimation accuracy in the presence of noise are also proposed in this chapter. This relies on constructing a statistical TOA profile for each sensor-received signal from its frequency components. The publication related to the work in this chapter is [C3], which was presented at the 42nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2017).

Part I

Algorithms Based on Location Template Matching

Chapter 2

Literature Review of LTM Algorithms

Location template matching (LTM), also known as *location pattern matching* [24], exploits dissimilarity in characteristic between source-sensor paths to distinguish signals originated from different locations. Since any signal distortion due to dispersion and/or multipath is highly dependent on the impact location, LTM is particularly suitable for localizing impacts in environments where these phenomena exist. Therefore, in the scope of applications involving rigid surfaces as touch interfaces, LTM is a popular approach for determination of locations where user interactions are made.

This chapter is organized as follows: Procedures for determination of an impact location using the LTM approach are discussed in Section 2.1. Sections 2.2 and 2.3 review key LTM algorithms that are deployed in HCI applications. Concluding remarks are given in Section 2.4

2.1 Introduction to LTM

Impact localization via LTM involves two stages, calibration and test, as illustrated in Fig. 2.1. In what follows, the superscript $\mathcal{L}_q$ is used to denote for the q -th library data/location to be defined during the calibration process. Let Q be the total number of calibration locations determined on the surface of interest. These locations are denoted by $(u^{\mathcal{L}_1}, v^{\mathcal{L}_1}), \dots, (u^{\mathcal{L}_Q}, v^{\mathcal{L}_Q})$. In a typical LTM

2.1 Introduction to LTM

Figure 2.1: Block diagram of a typical LTM algorithm

set-up, the determined set of calibration points forms a grid on the surface, as depicted in Fig. 2.2. During calibration, vibrations are generated at each of the grid points using either a finger or a stylus. The vibration signals $x^{\mathcal{L}_q}(n)$, $1 \leq n \leq N$, captured by a single sensor mounted at location (u_o, v_o) , are stored in a library $\{x^{\mathcal{L}_q}(n); q = 1, \dots, Q\}$. During the test phase, provided an impact is exerted at one of the calibration locations, $(u_s, v_s) \equiv (u^{\mathcal{L}_{\tilde{q}}}, v^{\mathcal{L}_{\tilde{q}}})$ for some $1 \leq \tilde{q} \leq Q$, the objective of an LTM algorithm is to determine the impact location given the received test signal $x(n)$.

Based on the assumption that an impact-induced signal preserves information about its source location, LTM-based algorithms focus on the extraction of location-specific information from the sensor-received signal. Such information encompasses the signature that distinguishes impacts at one location from those made at other locations. In the test phase, the signature of $x(n)$ is compared with that of each signal stored in the library $\{x^{\mathcal{L}_q}(n); q = 1, \dots, Q\}$. The *match* for the test signal is the library signal with the most similar signature. The source location of the test signal is subsequently determined as that of the matching library signal.

For an LTM algorithm, it all boils down to which location-dependent fea-

Figure 2.2: A typical LTM setup for touch-interface applications.

tures are extracted and how they are used in the matching process. In what follows, LTM algorithms and their considerations of which feature to exploit as the signature for the signal will be reviewed.

2.2 Time-reversal theory-based LTM

The time-reversal theory, introduced in [25], allows one to convert a *divergent* wave emanating from an acoustic source into a *convergent* wave focusing on the source. Application of this theory in impact localization on solid plates has been extensively studied in the literature [16, 24, 26–29]. As will be shown later in this section, time-reversal-based impact localization involves correlating a received test signal with those pre-collected and stored in a library. This effectively utilizes the signal itself as a location-dependent feature. Due to the need for computing the cross-correlation between the test and library signals, time-reversal-based LTM is also known as cross-correlation-based LTM (CC-LTM).

With reference to the time-reversal theory, an impact generated by a user

2.2 Time-reversal theory-based LTM

at an arbitrary point (u, v) is assumed to be sufficiently brief so that it can be approximated by the Dirac delta function, i.e., $s(n) \approx \delta(n)$. It follows that during calibration, where an impact is generated at each calibrating location $(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$, the signal $x^{\mathcal{L}_q}(n)$ received at the sensor is correspondingly given by

$$\begin{aligned} x^{\mathcal{L}_q}(n) &\approx \delta(n) * h^{\mathcal{L}_q}(n) \\ &= h^{\mathcal{L}_q}(n), \end{aligned} \quad (2.1)$$

where $*$ is the convolution operator and $h^{\mathcal{L}_q}(n)$ is the impulse response between the impact location $(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$ and the sensor location (u_o, v_o) . It is worth noting that $h^{\mathcal{L}_q}(n)$ is determined by the propagation path from $(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$ to (u_o, v_o) and is therefore unique for each $(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$.

During the test phase, an impact generated at one of the calibrated locations $(u_s, v_s) \equiv (u^{\mathcal{L}_{\tilde{q}}}, v^{\mathcal{L}_{\tilde{q}}})$ for some $1 \leq \tilde{q} \leq Q$ yields a sensor output $x(n)$, which approximates $h^{\mathcal{L}_{\tilde{q}}}(n)$. According to the time-reversal theory, if $x(n)$ is time-reversed and “emitted” from the sensor location (u_o, v_o) , the vibration signal at each calibrated location $(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$ will be given by

$$\begin{aligned} \Psi_{x, h^{\mathcal{L}_q}}(n) &= x(-n) * h^{\mathcal{L}_q}(n) \\ &\approx h^{\mathcal{L}_{\tilde{q}}}(-n) * h^{\mathcal{L}_q}(n). \end{aligned} \quad (2.2)$$

This, in fact, corresponds to the cross-correlation between $h^{\mathcal{L}_{\tilde{q}}}(n)$ and $h^{\mathcal{L}_q}(n)$, and can therefore be rewritten as

$$\Psi_{x, h^{\mathcal{L}_q}}(n) = \sum_{\ell=1}^N h^{\mathcal{L}_{\tilde{q}}}(n) h^{\mathcal{L}_q}(n + \ell). \quad (2.3)$$

It can easily be seen that $\Psi_{x, h^{\mathcal{L}_q}}(n)$ achieves its maximum if and only if $h^{\mathcal{L}_{\tilde{q}}}(n) = h^{\mathcal{L}_q}(n)$, which is equivalent to $(u^{\mathcal{L}_{\tilde{q}}}, v^{\mathcal{L}_{\tilde{q}}}) \equiv (u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$. Therefore, the impact location can be determined as

$$\tilde{q} = \operatorname{argmax}_{1 \leq q \leq Q} \{ \max \Psi_{x, h^{\mathcal{L}_q}}(n) \}. \quad (2.4)$$

Recall from (2.1) that $h^{\mathcal{L}_q}(n)$ has been approximated by $x^{\mathcal{L}_q}(n)$. The function $\Psi_{x,h^{\mathcal{L}_q}}(n)$ is therefore equivalent to $\Psi_{x,x^{\mathcal{L}_q}}(n)$, which quantifies the similarity/dissimilarity between $x(n)$ and $x^{\mathcal{L}_q}(n)$. Other similarity measure can be utilized in place of $\Psi_{x,x^{\mathcal{L}_q}}(n)$. One such measure is

$$\tilde{\Psi}_{x,x^{\mathcal{L}_q}}(n) = \frac{1}{\sigma_x \sigma_{x^{\mathcal{L}_q}}} \Psi_{x,x^{\mathcal{L}_q}}(n), \quad (2.5)$$

where the standard deviations of $x(n)$ and $x^{\mathcal{L}_q}(n)$, respectively denoted by σ_x and $\sigma_{x^{\mathcal{L}_q}}$ are introduced to compensate for the variation in signal energy for each user interaction [14]. Other methods to enhance the reliability of the similarity measure are discussed in [24].

2.3 Classical plate theory-based LTM

While time-reversal theory can be utilized in various applications, for the specific case of impact localization on a rigid plate, the performance can be improved by incorporating signal propagation models in algorithm development. A survey of theories on wave propagation on solid plates can be found in [30]. Within the scope of this thesis, however, only models pertaining to thin rectangular solid plate are of concern. Such models include the classical plate theory [31, 32], the Mindlin theory (which includes corrections for both rotary inertia and shear effects) [33], and Rayleigh-Lamb theory [34, 35]. Among these models, the classical plate theory is the simplest and can provide a sufficiently good approximation of the vertical displacement of the plate surface due to an impact as long as the vibration is contained within a sufficiently low-frequency band [30]. As reported in [19] and [27], for touch-interface applications, components of a vibration caused by finger or metal stylus are of frequencies less than 10 kHz, the classical plate theory can therefore be utilized.

According to the classical plate theory [31, 32], flexural vibration on a thin plate is propagated in modes identified by frequencies ω_{lb} where $l, b \in \mathbb{Z}^+$ are the

Figure 2.3: The shaping function Λ_{lb} for the mode of vibration $(l, b) = (5, 5)$.

mode indices. Given mechanical properties of the plate including the dimensions L_x , L_y , and L_z , density ρ , absorption coefficient μ , and Poisson's ratio ν , the mode frequency is given by

$$\omega_{lb} = (\beta_l^2 + \beta_b^2) \sqrt{\frac{\dot{D}}{\rho L_z}}, \quad (2.6)$$

where $\beta_l = \pi l / L_x$, $\beta_b = \pi b / L_y$, $\dot{D} = EL_z^3 / 12(1 - \nu^2)$ and E is the Young's modulus. The vertical displacement at the sensor location (u_o, v_o) due to an impact made at an arbitrary location (u_s, v_s) is then given as

$$x(u_s, v_s; u_o, v_o, t) = \frac{4}{\rho L_x L_y L_z} \sum_{l=1}^{\infty} \sum_{b=1}^{\infty} \frac{\Lambda_{lb}(u_s, v_s; u_o, v_o) e^{-0.5\tilde{\mu}t} \sin(\varpi_{lb}t)}{\varpi_{lb}}, \quad (2.7)$$

where $\tilde{\mu} = \mu / (\rho L_z)$ is the reduced absorption coefficient, $\varpi_{lb} = \sqrt{\omega_{lb}^2 - (\tilde{\mu}/2)^2}$ and

$$\Lambda_{lb}(u_s, v_s; u_o, v_o) = \sin(\beta_l u_s) \sin(\beta_b v_s) \sin(\beta_l u_o) \sin(\beta_b v_o). \quad (2.8)$$

Equation (2.8) implies that the sensor location imposes different scaling effect on Λ_{lb} corresponding to each mode (l, b) . The shape of Λ_{lb} as a function of (u_s, v_s) , however, is deterministic, as depicted in Fig. 2.3 where its maximum amplitude is normalized to 1.

In what follows, LTM techniques utilizing the model in (2.7) will be reviewed. Based on this model, these techniques rely on location-dependent features that can be extracted from the signal. Derivation of such features and the matching measures utilized will be discussed. Theoretical derivations are performed on the continuous-time version of the signals, while actual extractions and measure computations are performed on the digitized signals.

2.3.1 All-pole filter model-based LTM (AP-LTM)

The impulse response of the vibration sensor can be described by a second-order linear differential equation. By integrating this sensor model to the mechanical model of flexural vibration in thin plates, it has been shown in [17] that a sensor output is a function of frequency. In addition, its spectrum can be utilized in matching $x(n)$ to the library $x^{\mathcal{L}_q}(n)$ via determination of their dominant frequencies. Dominant-frequency extraction on the spectrum of a signal is achieved by using the all-pole model for infinite impulse response (IIR) filters to model the signal. Denoting the all-pole coefficients for $x(n)$ by α_p , $p = 1, \dots, \mathcal{N}$, where $\mathcal{N}$ is the prediction order, model fitting is achieved by minimizing the difference between $x(n)$ and its estimate $\hat{x}(n) = \sum_{p=1}^{\mathcal{N}} \alpha_p x(n-p)$, given by

$$\mathcal{J} = \sum_{n=1}^N (\varepsilon(n))^2, \quad (2.9)$$

where $\varepsilon(n) = x(n) - \hat{x}(n)$. The spectrum of $x(n)$ can now be estimated as

$$|\mathcal{P}_x(e^{j\omega})|^2 = \frac{\sigma_\varepsilon^2}{|1 + \alpha_1 e^{-j\omega} + \dots + \alpha_{\mathcal{N}} e^{-j\omega\mathcal{N}}|^2}, \quad (2.10)$$

where σ_ε^2 is the variance of $\varepsilon(n)$. Factorizing the denominator in (2.10), the spectrum is given as

$$|\mathcal{P}_x(e^{j\omega})|^2 = \frac{\sigma_\varepsilon^2}{|(e^{j\omega} - e^{j\omega_1}) \dots (e^{j\omega} - e^{j\omega_{\mathcal{N}}})|^2}, \quad (2.11)$$

2.3 Classical plate theory-based LTM

Figure 2.4: Spectrum overlayed with all-pole model of testing vibration signals, for $\mathcal{N} = 3$ [after [17]].

where the poles $\omega_1, \dots, \omega_{\mathcal{N}}$ correspond to dominant frequencies in the spectrum of $x(n)$. Estimation of dominant frequencies via all-pole modeling for several typical vibration signals are illustrated in Fig. 2.4. Here, IIR filter of order $\mathcal{N} = 3$ is utilized in modeling. The filter coefficients α_p are evaluated for signals received from impacts generated at locations $(0.41, 0.20)$ and $(0.42, 0.20)$, with the sensor mounted at $(0.05, 0.25)$ (all coordinates are in meter). As can be seen, each spectrum estimate (in red) peaks at the dominant frequencies of the corresponding signal. In addition, it is useful to note that even though the impact locations are spatially close, their dominant frequencies differ from each other.

Since the dominant frequencies are unique for each signal, they can be utilized as its signature for location matching. Equivalently, the coefficients can also be employed as the signal's signature. Denoting the coefficients for $x^{\mathcal{L}_q}$ by $\alpha_p^{\mathcal{L}_q}$, $p = 1, \dots, \mathcal{N}$, the dissimilarity measure between $x(n)$ and $x^{\mathcal{L}_q}(n)$ is given by

$$\mathcal{D}_{x, x^{\mathcal{L}_q}}^{\text{AP}} = \sum_{p=1}^{\mathcal{N}} (\alpha_p - \alpha_p^{\mathcal{L}_q})^2. \quad (2.12)$$

The source location of $x(n)$ is subsequently estimated as $(u^{\mathcal{L}_{\tilde{q}}}, v^{\mathcal{L}_{\tilde{q}}})$ that yields

the minimum dissimilarity, as given by

$$\tilde{q} = \underset{q}{\operatorname{argmin}} \mathcal{D}_{x, x^{\mathcal{L}_q}}^{\text{AP}}. \quad (2.13)$$

2.3.2 Zak transform for time-frequency-based LTM (Z-LTM)

According to the classical plate theory, the library signals $x^{\mathcal{L}_q}(n)$ can be expressed using (2.7) as given by

$$x^{\mathcal{L}_q}(t) = \frac{4}{\rho L_x L_y L_z} \sum_{l=1}^{\infty} \sum_{b=1}^{\infty} \frac{\Lambda_{lb}(u^{\mathcal{L}_q}, v^{\mathcal{L}_q}) e^{-0.5\tilde{\mu}t} \sin(\varpi_{lb}t)}{\varpi_{lb}}, \quad (2.14)$$

where

$$\Lambda_{lb}(u^{\mathcal{L}_q}, v^{\mathcal{L}_q}) = \sin(\beta_l u^{\mathcal{L}_q}) \sin(\beta_b v^{\mathcal{L}_q}) \sin(\beta_l u_o) \sin(\beta_b v_o). \quad (2.15)$$

Note that since only one sensor at (u_o, v_o) is utilized, the dependency of Λ_{lb} on (u_o, v_o) is implied without being explicitly shown in (2.14). In addition, $\Lambda_{lb}(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$ serves as the amplitude of the vibration mode (l, b) . Since Λ_{lb} is unique to the impact location $(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$, the variance across time for each mode is also dependent on $(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$. By defining

$$\vartheta_{lb}(t) = \frac{\Lambda_{lb}(u^{\mathcal{L}_q}, v^{\mathcal{L}_q}) e^{-0.5\tilde{\mu}t} \sin(\varpi_{lb}t)}{\varpi_{lb}}, \quad (2.16)$$

the variance across time for each mode of vibration (l, b) is given by

$$\begin{aligned} \sigma_{lb}^2(u^{\mathcal{L}_q}, v^{\mathcal{L}_q}) &= \mathbb{E}_{\vartheta_{lb}} \{t^2\} - [\mathbb{E}_{\vartheta_{lb}} \{t\}]^2 \\ &= \int_0^{\infty} t^2 \vartheta_{lb}(t) dt - \left[\int_0^{\infty} t \vartheta_{lb}(t) dt \right]^2 \\ &= \frac{\Lambda_{lb}(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})}{\varpi_{lb}} \left(\frac{\tilde{\mu}^2 \varpi_{lb}^2 (3 - \varpi_{lb} - \tilde{\mu}) - \varpi_{lb}^5 + 0.375\tilde{\mu}^4}{\omega_{lb}^2} \right), \end{aligned} \quad (2.17)$$

where $\mathbb{E}_{\vartheta_{lb}} \{\cdot\}$ is the expectation operator with density function $\vartheta_{lb}(t)$. Being a function of $(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$, the variance σ_{lb}^2 can be employed as a location-dependent feature of the signal.

To analyze the variance across time for each mode frequency, it is desirable to

2.3 Classical plate theory-based LTM

utilize a time-frequency representation (TFR) that can track the instantaneous change in time and frequency content of the signal. More specifically, the choice of TFR should allow one to obtain the instantaneous energy by summing its energy distribution across all frequencies, and the energy density spectrum by summing across time. These properties are referred to as the time- and frequency-marginal conditions [36]. The authors in [19] have chosen the discrete Zak transform (DZT) [37, 38] to represent the signal in the time-frequency domain since this transform satisfies the marginal conditions. The DZT of a received signal $x(n)$ is obtained by first shaping the length- N vector $\mathbf{x} = [x(1), \dots, x(N)]^T$ into a $L \times M$ matrix defined as

$$\mathbf{x}_{L \times M} = \begin{bmatrix} x(1) & x(2) & \cdots & x(M) \\ x(M+1) & x(M+2) & \cdots & x(2M) \\ \vdots & \vdots & \ddots & \vdots \\ x((L-1)M+1) & x((L-1)M+2) & \cdots & x(LM) \end{bmatrix}, \quad (2.18)$$

where $LM = N$, and subsequently performing DFT on each of the columns. The resultant DZT matrix $\mathbf{Z}_x = [\mathcal{Z}_x(k, m)]$ is of size $K \times M$ with K being the number of frequency bins. Each element of the DZT matrix is given by [39]

$$\mathcal{Z}_x(k, m) = \sum_{\ell=1}^L x(m + \ell M) e^{-j\ell\omega_k}, \quad (2.19)$$

where $\omega_k = 2\pi k/K$ is a normalized angular frequency, $1 \leq k \leq K$. The variance of $\mathbf{Z}_x$ across time for each frequency ω_k is then estimated as

$$\hat{\sigma}_x^2(k) = \frac{1}{M} \sum_{m=1}^M \left[|\mathcal{Z}_x(k, m)| - \frac{1}{M} \sum_{m'=1}^M |\mathcal{Z}_x(k, m')| \right]^2, \quad (2.20)$$

and the vector

$$\hat{\boldsymbol{\sigma}}_x^2 = \left[\hat{\sigma}_x^2(1), \dots, \hat{\sigma}_x^2\left(\frac{K}{2}\right) \right]^T \quad (2.21)$$

defines a location-dependent feature for impacts made at location $(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$. The variance $\hat{\sigma}_x^2(k)$ is computed only up to $k = K/2$ due to the symmetric property

of the DFT. A (dis)similarity measure between the test signal $x(n)$ and a library signal $x^{\mathcal{L}_q}(n)$ is subsequently given as

$$\mathcal{D}_{x,x^{\mathcal{L}_q}}^{\text{Zak}} = \sum_{k=1}^{K/2} \left(\hat{\sigma}_x^2(k) - \hat{\sigma}_{x^{\mathcal{L}_q}}^2(k) \right)^2, \quad (2.22)$$

and the index of the location where the test signal is generated is obtained as

$$\tilde{q} = \underset{1 \leq q \leq Q}{\operatorname{argmin}} \mathcal{D}_{x,x^{\mathcal{L}_q}}^{\text{Zak}}. \quad (2.23)$$

2.4 Concluding Remark

In this chapter, two approaches for LTM have been reviewed. In the time-reversal-theory-based approach, the signal itself is deployed as the signature for matching. On the other hand, the approach based on the classical model only extracts certain location-dependent features for utilization in matching. The latter approach requires less computation complexity since it is sufficient to store and process only partial information of each vibration signal. This however, results in algorithms being sensitive to any signal distortion that may corrupt the feature(s) of interest. To this extent, time-reversal-based LTM is more robust due to the inclusion of the whole signal in the matching process. Combining the advantages of both approaches, a noise robust LTM algorithm will be proposed in the next chapter.

Chapter 3

Noise-Robust LTM Using Band-Limited Components

The time-reversal theory and the classical plate theory have been utilized extensively in existing LTM algorithms. The integration of a model for vibration helps to determine features of the received signal that can be used for matching. The use of only certain features of the signal not only reduces computational complexity but also improves performance compared to time-reversal-based algorithms where the whole signal is utilized during the signature-matching process. However, this advantage is only achieved when the signal-to-noise ratio (SNR) is sufficiently high, since features extracted from the signal are, in general, prone to corruption by background noise.

Based on the classical plate theory, this chapter presents a new method that extracts other features of a vibration signal for location matching. The analysis is discussed in Section 3.1, and the BLC-LTM algorithm is subsequently proposed in Section 3.2. This algorithm utilizes a correlation-based measure to achieve robustness to band-limited noise, and therefore outperforms existing methods in such situations, as will be discussed in Section 3.3. Concluding remarks are given in Section 3.4.

3.1 Flexural vibration in thin plates

In this section, the classical plate theory will be revisited for the purpose of identifying new signal features that are suitable for location matching. Consider the flexural vibration on a rectangular solid plate of dimension $L_x \times L_y \times L_z$. An impact exerted at a source location (u_s, v_s) will induce a vibration signal $x(u, v, t)$ at location (u, v) on the plate surface at time instant t . Employing Kirchoff's assumptions for linear, elastic small-deflection theory of thin plates, the governing differential equation for $x(u, v, t)$ is given by [31]

$$\dot{D}\nabla^4 x(u, v, t) + \mu \frac{dx(u, v, t)}{dt} + \rho L_z \frac{d^2 x(u, v, t)}{dt^2} = p(u, v, t), \quad (3.1)$$

where $p(u, v, t)$ is the lateral load at location (u, v) , ρ the material density, μ the absorption coefficient, $\dot{D} = EL_z^3/12(1 - \nu^2)$ the plate's stiffness, E the Young's modulus, ν the Poisson's ratio, and the biharmonic operator $\nabla^4 = (\partial^4/\partial u^4) + 2(\partial^4/\partial u^2\partial v^2) + (\partial^4/\partial v^4)$.

By modeling the load as an impact at location (u_s, v_s) , i.e.,

$$p(u, v, t) = \alpha \delta(t) \delta(u - u_s) \delta(v - v_s), \quad (3.2)$$

where $\delta(\cdot)$ is the Dirac delta function, and α denotes the strength of the impact, the solution to (3.1) can be obtained for a simply supported plate with zero initial condition as [17]

$$x(u, v, t) = \frac{4\alpha}{\rho L_x L_y L_z} \sum_{l=1}^{\infty} \sum_{b=1}^{\infty} \frac{A_{lb} e^{-0.5\tilde{\mu}t} \sin(\varpi_{lb}t)}{\varpi_{lb}}, \quad (3.3)$$

where the vibration at (u, v) is expressed as an infinite sum of vibration modes (l, b) with $l, b \in \mathbb{Z}^+$ being the mode indices. Each of the modes is an exponentially decaying sinusoidal wave with reduced absorption coefficient $\tilde{\mu} = \mu/(\rho L_z)$, frequency $\varpi_{lb} = \sqrt{\omega_{lb}^2 - (\tilde{\mu}/2)^2}$, and

$$A_{lb} = \sin(\beta_l u_s) \sin(\beta_b v_s) \sin(\beta_l u) \sin(\beta_b v), \quad (3.4)$$

3.2 The proposed BLC-LTM algorithm

where $\omega_{lb} = (\beta_l^2 + \beta_b^2)\sqrt{\frac{D}{\rho L_z}}$, $\beta_l = \pi l/L_x$, and $\beta_b = \pi b/L_y$.

Since each mode can be identified by its frequency ϖ_{lb} and amplitude Λ_{lb}/ϖ_{lb} , it can be represented by the value pair $(\Lambda_{lb}, \varpi_{lb})$. The set of value pairs $\Omega = \{(\Lambda_{lb}, \varpi_{lb}) \mid l, b = 1, 2, \dots\}$ is therefore another representation of the vibration $x(u, v, t)$. Note from (3.4) that the amplitude Λ_{lb}/ϖ_{lb} of each mode is unique to the source and sensor locations. As a result, any non-empty subset of Ω can serve as a location-dependent feature of $x(u, v, t)$.

Consider a component of $x(u, v, t)$ which is formulated as a combination of modes whose frequencies are within some band $B_f = (f - \Delta f, f + \Delta f)$. This *band-limited component* (BLC) corresponds to the set $\Omega_f = \{(\Lambda_{lb}, \varpi_{lb}) \mid \varpi_{lb} \in B_f; l, b \in \mathbb{Z}^+\}$, which is a subset of Ω , and is given by

$$z_f(u, v, t) = \frac{4\alpha}{\rho L_x L_y L_z} \sum_{\varpi_{lb} \in B_f} \frac{\Lambda_{lb} e^{-0.5\tilde{\mu}t} \sin(\varpi_{lb}t)}{\varpi_{lb}}. \quad (3.5)$$

Due to the equivalence of $z_f(u, v, t)$ and Ω_f , the component can be employed as a signature for the source location (u_s, v_s) .

In the next section, an LTM algorithm will be proposed which utilizes BLCs as location-dependent features. The proposed BLC-LTM also employs a cross-correlation-based measure for evaluating the similarity between any two feature profiles. The proposed measure exploits the dispersion across the BLCs to improve the robustness of the algorithm to background noise.

3.2 The proposed BLC-LTM algorithm

In the context of solid plates, any signal distortion is attributed to the effects of multipath and dispersion. Multipath refers to the phenomenon where the signal is reflected at the medium boundary and arrives at the sensor via multiple paths. Such reflection causes signal distortion which is unique to the geometrical relation between the source, sensor and boundary. Dispersion, on the other hand, refers to the dependence of propagation velocity on the frequency of vibration;

Figure 3.1: Proposed matching measure between the test signal $\mathbf{x}$ and a library signal $\mathbf{x}^{\mathcal{L}_q}$.

the further a signal travels, the more its frequency components are separated in space. While CC-LTM exploits the uniqueness in the shape of the signal brought about by these two phenomena, the dispersion property is further exploited here to serve as a new matching measure for better localization performance.

Recall from Section 3.1 that BLCs of a signal can serve as its location-dependent features. Here, a new matching measure involving these features is proposed. Figure 3.1 shows the process to obtain the proposed measure. The BLCs are first extracted via signal filtering using K band-pass filters with frequency responses $H_k(f)$ where $k = 1, \dots, K$ is the band index. Inputs to the band-pass filters correspond to the test signal $\mathbf{x} = [x(1), \dots, x(N)]^T$ and the library data $\mathbf{x}^{\mathcal{L}_q} = [x^{\mathcal{L}_q}(1), \dots, x^{\mathcal{L}_q}(N)]^T$, $q = 1, \dots, Q$. Outputs of the filters (the BLCs) are respectively denoted by $\mathbf{z}_k$ and $\mathbf{z}_k^{\mathcal{L}_q}$. To evaluate the similarity between $\mathbf{x}$ and $\mathbf{x}^{\mathcal{L}_q}$, consider the correlation between the two corresponding BLCs, given by

$$\Psi_{\mathbf{z}_k, \mathbf{z}_k^{\mathcal{L}_q}}(\ell) = \sum_{n=1}^N z_k(\ell) z_k^{\mathcal{L}_q}(\ell + n), \quad (3.6)$$

3.2 The proposed BLC-LTM algorithm

and denote the lag corresponding to its maximum as

$$\ell_{z_k, z_k}^{\max, \mathcal{L}_q} = \arg\max_{\ell} \Psi_{\mathbf{z}_k, \mathbf{z}_k}^{\mathcal{L}_q}(\ell). \quad (3.7)$$

Utilizing the velocity dispersion property of the material, it will first be shown that the highest peaks of $\Psi_{\mathbf{z}_k, \mathbf{z}_k}^{\mathcal{L}_q}(\ell)$ align at the same lag *for all* band index k only for a certain set of library data. This set of $\mathbf{x}^{\mathcal{L}_q}$ corresponds to locations that are equidistant from the sensor where the distance is the same as that between the sensor and the true source location. In other words, it will be shown for all pairs $k \neq k'$, $1 \leq k, k' \leq K$, that

$$\begin{cases} \ell_{z_k, z_k}^{\max, \mathcal{L}_q} = \ell_{z_{k'}, z_{k'}}^{\max, \mathcal{L}_q}, & \text{if } d(u_s, v_s) = d(u^{\mathcal{L}_q}, v^{\mathcal{L}_q}), \end{cases} \quad (3.8a)$$

$$\begin{cases} \ell_{z_k, z_k}^{\max, \mathcal{L}_q} \neq \ell_{z_{k'}, z_{k'}}^{\max, \mathcal{L}_q}, & \text{if } d(u_s, v_s) \neq d(u^{\mathcal{L}_q}, v^{\mathcal{L}_q}), \end{cases} \quad (3.8b)$$

where $d(u', v') = \|(u', v') - (u, v)\|$ is the distance from an arbitrary point (u', v') to the sensor. This property is subsequently exploited for the proposed measure to match signals. This measure achieves a high value only when the test signal $\mathbf{x}$ matches with the library data from the same source location. A detailed proof of (3.8a) and (3.8b) is given below.

Proof. In the proposed method, the K non-overlapping narrow band-pass filters are chosen in such a way that, for each band, the difference in velocity among the frequency components constituting the respective BLC is negligible. The k -th BLC can therefore be considered as propagating with a constant velocity c_k that is determined by the center frequency f_k of the corresponding band-pass filter, i.e., $c_k = c(f_k)$. On the other hand, the choice of filters also guarantees velocity dispersion exist between different BLCs, i.e., $c_k \neq c_{k'}$ if $k \neq k'$, $1 \leq k, k' \leq K$.

Consider any two BLCs $\mathbf{z}_k$ and $\mathbf{z}_{k'}$ of $\mathbf{x}$, $k \neq k'$, $1 \leq k, k' \leq K$. When arriving at the sensor via direct path, the BLCs would have propagated a distance $d(u_s, v_s)$ with velocities c_k and $c_{k'}$. The arrival times of $\mathbf{z}_k$ and $\mathbf{z}_{k'}$ are therefore given by $\tau_k = d(u_s, v_s)/c_k$ and $\tau_{k'} = d(u_s, v_s)/c_{k'}$, respectively. The difference in arrival

Figure 3.2: Cross-correlations between the test signal $\mathbf{x}$ and the library signals $\mathbf{x}^{\mathcal{L}_{\bar{q}}}$ and $\mathbf{x}^{\mathcal{L}_q}$. For $\mathbf{x}^{\mathcal{L}_{\bar{q}}}$, which corresponds to the true source location, the peaks of the correlation functions are aligned whereas for $\mathbf{x}^{\mathcal{L}_q}$, which corresponds to a location other than that of the source, the peaks are not aligned.

times is then given by

$$\begin{aligned} \tau_{k,k'} &= \tau_k - \tau_{k'} \\ &= d(u_s, v_s) \left(\frac{1}{c_k} - \frac{1}{c_{k'}} \right). \end{aligned} \quad (3.9)$$

By similar arguments, the difference between arrival times for the BLCs $\mathbf{z}_k^{\mathcal{L}_q}$ and $\mathbf{z}_{k'}^{\mathcal{L}_q}$ of $\mathbf{x}^{\mathcal{L}_q}$ is given by

$$\tau_{k,k'}^{\mathcal{L}_q} = d(u^{\mathcal{L}_q}, v^{\mathcal{L}_q}) \left(\frac{1}{c_k} - \frac{1}{c_{k'}} \right). \quad (3.10)$$

Note that, as opposed to the conventional TDOA concept, which refers to the difference in arrival times between signals of *two* sensors, here $\tau_{k,k'}$ or $\tau_{k,k'}^{\mathcal{L}_q}$ refers to the difference in arrival times between BLCs of *one* sensor-received signal.

In theory, $\tau_{k,k'}$ can also be obtained as the difference between the silent-period lengths of the data frames $\mathbf{z}_k$ and $\mathbf{z}_{k'}$, and so as $\tau_{k,k'}^{\mathcal{L}_q}$ with respect to $\mathbf{z}_k^{\mathcal{L}_q}$ and $\mathbf{z}_{k'}^{\mathcal{L}_q}$.

3.2 The proposed BLC-LTM algorithm

It follows that

$$\begin{aligned} L_k - L_{k'} &= \tau_{k,k'} \\ &= d(u_s, v_s) \left(\frac{1}{c_k} - \frac{1}{c_{k'}} \right) \end{aligned} \quad (3.11)$$

and

$$\begin{aligned} L_k^{\mathcal{L}_q} - L_{k'}^{\mathcal{L}_q} &= \tau_{k,k'}^{\mathcal{L}_q} \\ &= d(u^{\mathcal{L}_q}, v^{\mathcal{L}_q}) \left(\frac{1}{c_k} - \frac{1}{c_{k'}} \right), \end{aligned} \quad (3.12)$$

where L_k , $L_{k'}$, $L_k^{\mathcal{L}_q}$ and $L_{k'}^{\mathcal{L}_q}$ are the lengths of the silent periods of $\mathbf{z}_k$, $\mathbf{z}_{k'}$, $\mathbf{z}_k^{\mathcal{L}_q}$, and $\mathbf{z}_{k'}^{\mathcal{L}_q}$, respectively. Note that even though z_k and $z_{k'}$ are extracted from the same signal, their silent periods L_k and $L_{k'}$ are different due to velocity dispersion. This is also the case for $L_k^{\mathcal{L}_q}$ and $L_{k'}^{\mathcal{L}_q}$.

Equations (3.11) and (3.12) imply that when $\mathbf{x}$ and $\mathbf{x}^{\mathcal{L}_q}$ correspond to locations that are equidistant from the sensor, i.e., $d(u_s, v_s) = d(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$, the above differences in arrival times are the same, i.e.,

$$L_k - L_{k'} = L_k^{\mathcal{L}_q} - L_{k'}^{\mathcal{L}_q}, \quad 1 \leq k, k' \leq K, \quad (3.13)$$

which can be rewritten as

$$L_k - L_k^{\mathcal{L}_q} = L_{k'} - L_{k'}^{\mathcal{L}_q}, \quad 1 \leq k, k' \leq K. \quad (3.14)$$

Here it is worth noting that since the multipath effect is neglected and only dispersion has been taken into account, the cross-correlation between $\mathbf{z}_k$ and $\mathbf{z}_k^{\mathcal{L}_q}$ achieves its maximum at the lag that compensates for the difference between their arrival times, i.e.,

$$\ell_{z_k, z_k^{\mathcal{L}_q}}^{\max} = L_k - L_k^{\mathcal{L}_q}. \quad (3.15)$$

The result in (3.14) therefore implies (3.8a). On the contrary, when $\mathbf{x}^{\mathcal{L}_q}$ is from a location where $d(u_s, v_s) \neq d(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$, the equalities in (3.13) and (3.14) no longer holds, which leads to (3.8b). The above implication is depicted in Fig. 3.2, where $\mathbf{x}^{\mathcal{L}_{\bar{q}}}$ and $\mathbf{x}$ are from locations that are equidistant from the sensor, whereas $\mathbf{x}^{\mathcal{L}_q}$ is from a location of a different distance from the sensor.

Note that there may exist multiple library data corresponding to locations which lie on the circle centered at the sensor with radius $d(u_s, v_s)$. While the BLCs' correlation peaks are aligned for all such locations, if multipath is taken into account, these peaks are low in amplitude except only for the library data collected at the true source location. This is due to the uniqueness of reflections from the boundary with respect to impact location. $\square$

It is important to note that (3.8b) does *not* apply for the case of non-dispersive signals. In the absence of velocity dispersion, it follows that $1/c_k - 1/c_{k'} = 0$, $1 \leq k, k' \leq K$. As a result, (3.11) and (3.12) always imply $L_k - L_{k'} = 0$ and $L_k^{\mathcal{L}_q} - L_{k'}^{\mathcal{L}_q} = 0$. These can be rewritten as $L_k - L_k^{\mathcal{L}_q} = L_{k'} - L_{k'}^{\mathcal{L}_q}$, or equivalently

$$\ell_{z_k, z_k}^{\max, \mathcal{L}_q} = \ell_{z_{k'}, z_{k'}}^{\max, \mathcal{L}_q}, \quad \forall 1 \leq q \leq K. \quad (3.16)$$

This implies that the correlations $\Psi_{\mathbf{z}_k, \mathbf{z}_k^{\mathcal{L}_q}}(\ell)$, $k = 1, \dots, K$, always have their maximum peaks aligned at the same lag, irrespective of the library data vector used for comparison.

With the above, a measure for signal matching is proposed as

$$\mathcal{Y}_q(\ell) = \sum_{k=1}^K \frac{1}{\sigma_{\mathbf{z}_k} \sigma_{\mathbf{z}_k^{\mathcal{L}_q}}} \Psi_{\mathbf{z}_k, \mathbf{z}_k^{\mathcal{L}_q}}(\ell), \quad (3.17)$$

where $\sigma_{\mathbf{z}_k}$ and $\sigma_{\mathbf{z}_k^{\mathcal{L}_q}}$ are the standard deviations of $\mathbf{z}_k$ and $\mathbf{z}_k^{\mathcal{L}_q}$, respectively. The proposed measure makes use of the fact that the maxima of $\Psi_{\mathbf{z}_k, \mathbf{z}_k^{\mathcal{L}_q}}(\ell)$, $k = 1, \dots, K$, are aligned at the same lag when $(u_s, v_s) = (u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$, resulting in a dominant maximum of $\mathcal{Y}_q(\ell)$. On the other hand, when $(u_s, v_s) \neq (u^{\mathcal{L}_q}, v^{\mathcal{L}_q})$,

3.2 The proposed BLC-LTM algorithm

the maxima of $\Psi_{\mathbf{z}_k, \mathbf{z}_k^{\mathcal{L}_q}}$ are smaller and misaligned. As a result, $\Upsilon_q(\ell)$ exhibits single/multiple peak(s) with magnitude(s) significantly lower than that of the dominant peak when there is a match. The source location is therefore identified as $(u_s, v_s) = (u^{\mathcal{L}_{\tilde{q}}}, v^{\mathcal{L}_{\tilde{q}}})$, where

$$\tilde{q} = \underset{1 \leq q \leq Q}{\operatorname{argmax}} \{ \max \Upsilon_q(\ell) \}. \quad (3.18)$$

The proposed correlation-based measure is depicted in Fig. 3.1.

In what follows, the behavior of $\Psi_{\mathbf{z}_k, \mathbf{z}_k^{\mathcal{L}_q}}(\ell)$ in the presence of noise is to be investigated. For simplicity, assume that calibration is conducted in a noise-free environment while noise may be present during the test phase. The library and test signals can then be modeled as

$$\mathbf{x}^{\mathcal{L}_q} = \mathbf{s}^{\mathcal{L}_q}, \quad \mathbf{x} = \mathbf{s} + \mathbf{w}, \quad (3.19)$$

where $\mathbf{s}^{\mathcal{L}_q} = [s^{\mathcal{L}_q}(1), \dots, s^{\mathcal{L}_q}(N)]^T$, $q = 1, \dots, Q$, and $\mathbf{s} = [s(1), \dots, s(N)]^T$ are the signals due to vibration and $\mathbf{w} = [w(1), \dots, w(N)]^T$ is the background noise. With $\mathbf{r}_k^{\mathcal{L}_q}$, $\mathbf{r}_k$, and $\mathbf{v}_k$ being the respective BLCs of $\mathbf{s}^{\mathcal{L}_q}$, $\mathbf{s}$, and $\mathbf{w}$, the BLCs of $\mathbf{x}^{\mathcal{L}_q}$ and $\mathbf{x}$ can then be expressed as

$$\mathbf{z}_k^{\mathcal{L}_q} = \mathbf{r}_k^{\mathcal{L}_q}, \quad \mathbf{z}_k = \mathbf{r}_k + \mathbf{v}_k. \quad (3.20)$$

In reality, either indoor or outdoor HCI applications suffer from band-limited noise. Such noise can arise from artificial sources such as road vehicles and air movement machinery including fans and ventilation or air-conditioning units [40]. The case where the noise $\mathbf{w}$ is band-limited is therefore considered here so that there exists a set of indices I_0 where the BLCs $\mathbf{r}_k$ for $k \in I_0$ have sufficiently high local SNRs, i.e., $\mathbf{v}_k \approx \mathbf{0}$. Here $\mathbf{0}$ is a length- N vector of zeros. The correlation in (3.6) can then be expressed as

$$\Psi_{\mathbf{z}_k, \mathbf{z}_k^{\mathcal{L}_q}}(\ell) = \begin{cases} \Psi_{\mathbf{r}_k, \mathbf{r}_k^{\mathcal{L}_q}}(\ell) & , \forall k \in I_0 \\ \Psi_{\mathbf{r}_k, \mathbf{r}_k^{\mathcal{L}_q}}(\ell) + \Psi_{\mathbf{v}_k, \mathbf{r}_k^{\mathcal{L}_q}}(\ell) & , \forall k \notin I_0 \end{cases}. \quad (3.21)$$

When the noise bandwidth is narrow, a few BLCs are contaminated with noise. Since each of the correlation functions are normalized before they are summed in (3.17), when the BLCs with high SNR outnumber the noisy BLCs, the contribution of the noise-free terms $\Psi_{\mathbf{r}_k, \mathbf{r}_k^{\mathcal{L}_q}}(\ell)$ dominates that of the noisy terms $\Psi_{\mathbf{v}_k, \mathbf{r}_k^{\mathcal{L}_q}}(\ell)$. As a result, $\mathcal{T}_q(\ell)$ is robust to noise even if the SNR of the sensor signal is low, provided the noise is sufficiently band-limited. This is not true for conventional CC-LTM described in Section 2.2.

3.3 Experiment Results

In this section, experiment results are presented to verify the performance of the proposed algorithm in the context of a touch interface. The experiments are conducted on rectangular plates made of aluminum and glass, with dimensions $0.6 \text{ m} \times 0.6 \text{ m} \times 5.0 \text{ mm}$ and $0.6 \text{ m} \times 0.6 \text{ m} \times 4.0 \text{ mm}$, respectively. A single Murata PKS1-4A1 shock sensor is mounted on each of the plates. Sensor-received signals are digitized using a PreSonus FireStudio at a sampling frequency of 96 kHz to achieve high resolution in time, although for Murata sensors, signals of frequency 8 kHz and above are attenuated and are negligibly small. The coordinate system is chosen on each plate so that the sensor is located at the origin, i.e., $(u, v) = (0, 0)$. A set of 25 calibration points is arranged in a 5×5 grid with $(u^{\mathcal{L}_1}, v^{\mathcal{L}_1}) = (10, 10) \text{ cm}$ lying at the lower left corner. The distances between the columns of the grid from left to right (and rows from bottom to top) are respectively 1, 3, 6, and 10 cm. The schematic of the experiment setup is depicted in Fig. 3.3. During calibration, five impacts are generated at each of the pre-defined locations, resulting in a total of 125 signals. In the test phase, each of the signal takes turn being the test signal, with the remaining 124 signals being the library. An accurate match is obtained when a test signal is matched with one of the other four that correspond to the same source location. For all experiments,

3.3 Experiment Results

Figure 3.3: The schematic of a typical LTM setup for touch-interface applications.

the classification performance of algorithms is quantified by accuracy defined as

$$\mathcal{C} = \frac{N_c}{N_t}, \quad (3.22)$$

where N_c is the number of accurate matches and $N_t = 125$ is the total number of tests conducted.

The effect of the choice of filters to the performance of the proposed BLC-LTM is first examined in what follows. Assume that sensor-received signals are band-limited with maximum frequency of f_B due to the sensor's frequency response. This band is covered by the passbands of the K filters by choosing the passband of the k -th filter to be $[f_B(k-1)/K, f_B k/K]$. Figure 3.4(a) shows the variation in performance of BLC-LTM for different values of K with SNR -10 dB and the noise being limited to the band $[3, 5]$ kHz. As discussed previously, as K increases, there are more BLCs that are free of noise, resulting in the dominance of the noise-free terms over the noisy terms in (3.17). The localization accuracy is therefore improved as K increases, as can be seen from the figure. In the following experiments, the value of K is fixed to be equal to 8, which corresponds to the choice of filters with passbands of $[0, 1]$ kHz, $[1, 2]$ kHz, $[2, 3]$ kHz, $[3, 4]$ kHz, $[4, 5]$ kHz, $[5, 6]$ kHz, $[6, 7]$ kHz, and $[7, 8]$ kHz.

Figure 3.4: Variation in performance of (a) BLC-LTM with change in the number of filters and (b) BLC-LTM and CC-LTM with change in the noise bandwidth.

The performance of the proposed BLC-LTM is next compared with that of the conventional CC-LTM [14] for various noise bandwidths. Test signals are to be added with synthetically generated noise signals within frequency bands of $[0, 1]$ kHz, $[0, 2]$ kHz, $[0, 3]$ kHz, $[0, 4]$ kHz, $[0, 5]$ kHz, $[0, 6]$ kHz, $[0, 7]$ kHz and $[0, 8]$ kHz. The result is shown in Fig. 3.4(b) where the SNR is maintained at -10 dB. It can be seen that the accuracies of both the algorithms are low when the noise bandwidth is large (above 4 kHz, which is half of the bandwidth of the signal received from the Murata sensor). However, unlike the CC-LTM method, the performance of BLC-LTM increases as the noise bandwidth decreases. It is apparent that the conventional cross-correlation does not provide any means to suppress the effect of noise regardless of bandwidth. On the contrary, the measure proposed in (3.17) makes use of the dominance of noise-free BLCs to achieve robustness of the proposed algorithm to band-limited noise.

Finally, BLC-LTM and CC-LTM are compared with the Zak transform-based LTM (Z-LTM) [19] discussed in Section 2.3, which is a non-correlation-based method. The performance of the algorithms is evaluated for two cases, white noise and band-limited noise of passband $[3, 4]$ kHz, with varying SNR. Figures 3.5(a) and 3.5(b) show the classification accuracies of the three algorithms on the glass plate for different values of SNR. Respective experiment results on the aluminum

3.3 Experiment Results

Figure 3.5: Localization accuracy vs SNR on glass for (a) white noise and (b) band-limited noise, and on aluminum for (c) white noise and (d) band-limited noise.

plate are shown in Figs. 3.5(c) and 3.5(d). It can be seen from the figures that the three algorithms perform equally well under high SNR conditions (above 0 dB) for all the cases. BLC-LTM and CC-LTM, however, perform better when the SNR is low. This is due to the robustness of cross-correlation-based methods to uncorrelated noise. On the contrary, non-correlation-based methods such as Z-LTM do not possess this property and have their extracted feature severely distorted when the noise energy begins to dominate the signal energy, leading to poor localization performance. While the proposed BLC-LTM can achieve the same performance as that of CC-LTM in lower SNR (up to -15 dB) for narrow-band noise, the latter modestly outperforms the former when the noise is white. This agrees with the previous discussion on BLC-LTM's performance versus noise bandwidth.

3.4 Concluding Remark

By analyzing the mechanical model for vibration due to impacts on plate surfaces, it has been shown that a signal's BLCs can be employed as location-dependent features. A measure for signal matching has also been proposed based on the cross-correlations between the respective BLCs of the test and library signals. The proposed measure exploits velocity dispersion among BLCs to exhibit clear distinction in value between the matched and non-matched cases in template matching. In addition, the use of multiple BLCs significantly improves the performance of the proposed BLC-LTM against existing methods when the background noise is band-limited.

However, one of the disadvantages of the LTM methods is that they can only localize impacts made at points belonging to a discrete set of predefined locations. Therefore, to achieve high resolution in impact localization, it is necessary to include as many points as possible in the predefined set. Given that features have to be extracted for every possible tap location made by the user, this incurs extensive calibration and intensive computation. In addition, the localization performance depends on the assumption of similarity between taps at the same location. This assumption, however, is often violated in reality due to inconsistencies in the style of tapping or any deviation of the test condition from that of the calibration phase. The TDOA-based approach for impact localization, which is free from such constraints, will be discussed in Part II of this thesis.

Part II

TDOA-Based Localization

Chapter 4

Literature Review of TDOA-based Localization Algorithms

While the localization ability of LTM methods is limited to a set of predefined discrete locations, TDOA-based methods do not suffer from such a constraint. In addition, TDOA-based localization utilizes more than one sensor, and the performance tends to improve with an increasing number of sensors. The concept of TDOA refers to the difference in propagation time of a signal from its source to each sensor of a given pair. By estimating and combining the TDOAs of all the sensor pairs, the source location can be determined. The main goal of algorithms in this family is therefore to obtain an accurate TDOA estimate for each sensor pair.

In this chapter, the steps in TDOA-based localization will first be discussed in Section 4.1, followed by a review of methods for direct and indirect TDOA estimation in Sections 4.2 and 4.3, respectively.

4.1 Introduction to TDOA-based localization

Consider a plate surface mounted with R sensors where the i th sensor is mounted at location (u_i, v_i) . Supposed that an impact is exerted at location (u_s, v_s) , generating the sensor outputs $x_i(n)$. For each sensor pair (i, j) , the TDOA $\tau_{i,j}$, in

Figure 4.1: A typical example of TDOA-based source localization. The impact location and sensor locations are respectively denoted by $\circ$ and $\bullet$.

samples, is subsequently estimated using the respective received signals $x_i(n)$ and $x_j(n)$. Each $\tau_{i,j}$ determines a locus of possible source locations as a hyperbolic curve on the surface, with (u_i, v_i) and (u_j, v_j) being the foci. Such a hyperbola, denoted by $\mathcal{H}_{ij}$, consists of all the point (u, v) satisfying the condition

$$\frac{d_i(u, v) - d_j(u, v)}{c} = \frac{\tau_{i,j}}{f_s}, \quad (4.1)$$

where $d_i(u, v) = \sqrt{(u - u_i)^2 + (v - v_i)^2}$ is the Euclidean distance from (u, v) to the i th sensor positioned at (u_i, v_i) , c is the wave propagation velocity, and f_s is the sampling frequency. Provided the value of c is known *a priori*, an estimate of the source location can be obtained as the intersection of all $\mathcal{H}_{ij}$ by solving the optimization problem

$$(\hat{u}_s, \hat{v}_s) = \underset{(u,v)}{\operatorname{argmin}} \sum_{i,j} \left(\frac{d_i(u, v) - d_j(u, v)}{c} - \hat{\tau}_{i,j} \right)^2, \quad (4.2)$$

A typical example of TDOA-based localization using three sensors is illustrated in Fig. 4.1 with the impact location denoted by $\circ$ and the sensor locations denoted by $\bullet$.

The propagation velocity c in (4.2) is usually an unknown parameter and

requires prior estimation. During calibration, impacts are generated at a set of Q pre-determined locations $(u^{\mathcal{L}_1}, v^{\mathcal{L}_1}), \dots, (u^{\mathcal{L}_Q}, v^{\mathcal{L}_Q})$. The TDOAs $\hat{\tau}_{i,j}^{\mathcal{L}_q}$ estimated for each impact are subsequently utilized to estimate c as

$$\hat{c} = \underset{c}{\operatorname{argmin}} \sum_q \sum_{i,j} \left(\frac{d_i(u^{\mathcal{L}_q}, v^{\mathcal{L}_q}) - d_j(u^{\mathcal{L}_q}, v^{\mathcal{L}_q})}{c} - \hat{\tau}_{i,j}^{\mathcal{L}_q} \right)^2. \quad (4.3)$$

It is worth noting that since velocity calibration depends on the underlying TDOA estimation algorithm, different algorithms may use different calibrated values for velocity.

Within the scope of this thesis, the human-computer touch interfaces of interest are made of solid materials such as glass and aluminum. In such environments, wave propagation is dispersive, i.e. the speed of propagation is dependent on the frequency of vibration. This causes distortion to the signal waveform and renders TDOA estimation using conventional techniques such as the generalized cross-correlation (GCC) method prone to errors. To mitigate the effect of dispersion, pre-processing of the sensor outputs is required [21–23]. Such pre-processing techniques often involve isolating signal components within a chosen narrow frequency band. As a result, the concepts of TDOA and propagation velocity become specific to the signal components that have their frequency contents encompassed within the selected band.

In addition to dispersion, waveform distortion due to multipath should also be taken into account. Multipath distortion are due to the interferences of waves reflected from the medium boundaries before reaching the sensors. This results in spurious peaks in the GCC functions, hence degradation in performance of the algorithm. To circumvent such problem, other techniques such as presented in [41–44] employs *indirect* approaches for TDOA estimation.

4.2 Generalized cross-correlation for direct TDOA estimation

The GCC method [20] assumes that the signals received at the sensors are scaled and delayed versions of each other, as given by

$$x_i(n) = \alpha_{i,j} x_j(n - \tau_{i,j}), \quad (4.4)$$

where $\alpha_{i,j}$ is the inter-sensor gain, and $\tau_{i,j}$ is the TDOA (in terms of number of samples) between the (i, j) sensor pair. Based on this assumption, the TDOA $\tau_{i,j}$ is estimated as

$$\hat{\tau}_{i,j} = \operatorname{argmax} \Psi_{i,j}(n), \quad (4.5)$$

where

$$\Psi_{i,j}(n) = \sum_{k=0}^{K-1} \underline{\lambda}(k) \underline{x}_i(k) \underline{x}_j^*(k) e^{j2\pi nk/K} \quad (4.6)$$

is the generalized cross-correlation function between $x_i(n)$ and $x_j(n)$. Here, the summation is performed in the frequency domain with K being the total number of frequency bins and k the frequency-bin index. The functions $\underline{x}_i(k)$ and $\underline{x}_j(k)$ are the DFTs of the signals received at the i th and j th sensors with $(\cdot)^*$ denoting the complex conjugate and $\underline{\lambda}(k)$ being the generalized frequency weighting. There are several member algorithms in the GCC family depending on how the weight $\underline{\lambda}(k)$ is defined. Among the commonly used weighting functions such as the smoothed coherence transform (SCOT) [45] and the Hannan-Thomson (HT) processor [46], the phase transform (PHAT) [20] has shown to achieve good performance in presence of low noise and multipath [47].

In dispersive environments, however, the assumption (4.4) does not hold in general because the delay $\tau_{i,j}$ is a function of frequency, as explained via the propagation model in Section 1.1. Direct application of the GCC method on dispersed signals is therefore prone to error. A method utilizing GCC for source localization in dispersive media has been proposed in [21]. To mitigate the effect of

dispersion, the algorithm isolates a single frequency of the signal before applying the GCC method. The isolation is achieved by correlating the signal with a Gaussian-modulated cosine pulse of a frequency ω' . The obtained component corresponding to the frequency ω' is given by

$$g(t) = e^{-[(t-t_c)^2/\sigma_c^2]} \cos(\omega' t), \quad (4.7)$$

where t_c and σ_c^2 are the center and variance in time of the Gaussian pulse. The use of the Gaussian pulse is motivated by its ability to taper the cosine function, thereby narrowing the frequency content of the signal before the application of GCC.

The idea of frequency isolation is revisited in [22], where a Butterworth filter is utilized to filter the signal component within an appropriate band. In [22], different frequency bands are examined so that the optimal band, which gives the minimal root-mean-square error (RMSE) in localization, can be selected. A different approach to signal component isolation using wavelet transform (WT) is also proposed in [23]. The WT decomposes the signal into frequency bands corresponding to different scales. Each scale provides an estimate of TDOA for each sensor pair by cross-correlating the respective signal components. The value leading to the maximum cross-correlation coefficient is subsequently chosen as the best estimate of the true TDOA.

While the above methods yield reasonable location estimates in the presence of dispersion, their performance is limited when multiple reflections from the medium boundaries generate spurious peaks in the cross-correlation coefficients. The localization performance is further reduced when the sensors are placed close to the edges. To circumvent the multipath phenomenon, many methods employ a different approach by estimating the time-of-arrival (TOA) of the signal to each sensor. The TDOA can then be obtained *indirectly* as the difference between the estimated TOAs. Different methods using this approach will be discussed in the

next section.

4.3 Methods for indirect TDOA estimation

Methods for indirect TDOA estimation focus on the portion corresponding to the initial arrival of the signal at each sensor in order to mitigate waveform distortion due to multipath. In what follows, since the signal received at each sensor is to be treated individually, the subscript i denoting the sensor index will be omitted for brevity, i.e., the sensor signal will be denoted by either $x(t)$ or $x(n)$ for its time-continuous or digitized version, respectively.

4.3.1 Wavelet analysis for TOA estimation

The continuous wavelet transform (CWT) of an arbitrary analog signal $x(t)$ is defined as [48]

$$\mathcal{W}_\psi(a, b) = \frac{1}{a} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt, \quad (4.8)$$

where a and b denote, respectively, the dilation and translation in time of a function, and $\psi(t)$ is the analyzing wavelet function which satisfies the admissibility condition given by [49]

$$\int_{-\infty}^{\infty} \frac{|\underline{\psi}(\omega)|^2}{|\omega|} d\omega < \infty, \quad (4.9)$$

where $\underline{\psi}(\omega)$ is the Fourier transform of $\psi(t)$. Using Parseval's theorem, (4.8) can be rearranged to obtain an expression of the CWT via the Fourier transform $\underline{x}(\omega)$ of the signal $x(t)$ as

$$\mathcal{W}_\psi(a, b) = \frac{\sqrt{a}}{2\pi} \int_{-\infty}^{\infty} \underline{x}(\omega) e^{jb\omega} \underline{\psi}^*(a\omega) d\omega. \quad (4.10)$$

The discrete form in (4.10) leads to the wavelet domain of the sampled signal $x(n)$.

From (4.8) and (4.10), we note that the analyzing wavelet function $\psi(t)$ may be considered as a window both in the time and frequency domains, and the choice of this window distinguishes an algorithm from others in the WT family.

It is also worth mentioning that any windowing function is subjected to the uncertainty principle [50], which states that no windowing function can have a time-bandwidth product of less than 1/2.

Among all kinds of window functions, the Gabor function with its Gaussian shape provides a small (concentrated) window in the time-frequency domain [49], and has been proven to achieve the lower bound of the uncertainty principle [51]. For this reason, the authors in [41] employed the Gabor function as the analyzing wavelet. The Gabor function and its Fourier transform are respectively given by

$$\psi_G(t) = \frac{1}{\sqrt[4]{\pi}} \sqrt{\frac{\omega_0}{\gamma}} \exp \left[-\frac{(\omega_0/\gamma)^2}{2} t^2 + j\omega_0 t \right], \quad (4.11)$$

and

$$\underline{\psi}_G(\omega) = \frac{2\pi}{\sqrt[4]{\pi}} \sqrt{\frac{\gamma}{\omega_0}} \exp \left[-\frac{(\gamma/\omega_0)^2}{2} (\omega - \omega_0)^2 \right], \quad (4.12)$$

where ω_0 and γ are positive constants. The authors in [49] propose to set $\gamma = \pi\sqrt{2/\ln 2}$ so that the Gabor function approximately satisfies the admissibility condition (4.9) and $\omega_0 = 2\pi$ such that the dilation a is equal to the period of vibration.

The authors of [41] performed wavelet analysis on each sensor-received signal and stated that the arrival time of an incident wave can be estimated from the maximum of the magnitude of the wavelet transform. With the knowledge of the arrival time at each sensor, an optimization problem on a system of nonlinear equations was then formulated to jointly estimate the source location and the propagation velocity.

Another wavelet-based algorithm has been proposed in [42], but with the complex Morlet wavelet employed as the analyzing function. The Morlet wavelet is given by

$$\psi_M(t) = \frac{1}{\sqrt{\pi} f_b} e^{j\omega_c t} e^{-t^2/f_b}, \quad (4.13)$$

where ω_c is the central angular frequency, and f_b is the wavelet bandwidth functioning as a shape control parameter. This wavelet has been chosen for its ability

4.3 Methods for indirect TDOA estimation

Figure 4.2: 3D plot of the scalogram of the CWT in which the ridge is taken at the instantaneous frequency of 310 kHz [after [42]].

to separate amplitude and phase, which enables the measurement of instantaneous frequencies and their temporal evolution [52]. At each time instant, the instantaneous frequency is identified as the dominant frequency where the signal energy is most concentrated. This energy concentration corresponds to the CWT coefficient of highest modulus at the time being considered. The collection of such CWT coefficients in time determines *ridges* on the scalogram [52, 53]. Fig. 4.2 shows an example of a ridge on the scalogram of a signal in an experiment carried out in [42], where impacts were generated by dropping a 9 mm diameter steel ball on the surface of a $0.52 \text{ m} \times 0.41 \text{ m} \times 0.97 \text{ mm}$ plate. It can be seen that the ridge corresponds to a peak in the magnitude of CWT coefficient. Projecting the ridge location on the time and frequency domains, a time instant and the instantaneous frequency at that time instant can be obtained. The authors in [42] proposed to estimate the TOA by the time projection of the location corresponding to the maximal ridge.

Although wavelet transform is a popular choice for the analysis of nonstationary signals, the use of a single ridge of the scalogram (hence, a single frequency component) for arrival detection is not robust to environment noise. It has been reported in [44] that the wavelet-based method proposed in [41] did not provide

Figure 4.3: Diagram representing difference between transit time, phase transit time and centroid transit time [after [43]].

satisfactory performance with the average localization error of about 0.1 m on glass surfaces of dimension 0.6 m $\times$ 0.6 m for impacts generated by a finger or a metal stylus. The wavelet-based method is outperformed by the KLID [43] and TiF-HA [44] methods, both of which utilize a sufficiently wide frequency range in TOA estimation. These two methods will be reviewed in the subsequent sections.

4.3.2 Kullback-Leibler discrimination information for phase transit point detection

The authors in [43] associate signal arrival with the *phase transit point*, which corresponds to the maximum change in the signal spectrum. Figure 4.3 illustrates the phase transit time of a signal in comparison with other concepts including the *transit time* and the *centroid transit time* [43]. In order to determine the phase transit point, it is necessary to monitor the variation in signal spectrum. Consider the data frames obtained by applying a sliding window of length L on $x(n)$. Each data frame is given by

$$\mathbf{x}(m) = [x(m), \dots, x(m + L - 1)], \quad m = 0, 1, \dots, N - L. \quad (4.14)$$

4.3 Methods for indirect TDOA estimation

Figure 4.4: Left and right data frames at time index n .

Note that the frame index m is also the time index of the first element of a frame. An estimate of the spectral density (or frequency distribution) of $\mathbf{x}(m)$ can be given by its periodogram as

$$\hat{\mathcal{P}}_{\mathbf{x}(m)}(\omega) = \frac{1}{2\pi L} \left| \sum_{n=m}^{m+L-1} x(n) e^{-jn\omega} \right|^2. \quad (4.15)$$

Post-processing techniques to reduce frequency leakage such as those proposed in [54] and [55] can also be applied on the periodograms to obtain better spectral estimates.

To determine the point where maximum spectral variation occurs, for each time index n , $L-1 \leq n \leq N-L$, the two frames $\mathbf{x}(n+1-L)$ and $\mathbf{x}(n)$ adjacent respectively to the left and right of $x(n)$, as illustrated in Fig. 4.4, are considered. Denoting the estimated spectral density of the left frame by $\hat{\mathcal{P}}_L(n, \omega)$ and that of the right frame by $\hat{\mathcal{P}}_R(n, \omega)$, where $-\pi \leq \omega \leq \pi$ is the normalized frequency, it follows that

$$\hat{\mathcal{P}}_L(n, \omega) = \hat{\mathcal{P}}_{\mathbf{x}(n+1-L)}(\omega) \quad \text{and} \quad \hat{\mathcal{P}}_R(n, \omega) = \hat{\mathcal{P}}_{\mathbf{x}(n)}. \quad (4.16)$$

The change in spectral density at each time index n is then estimated as the discrepancy between $\hat{\mathcal{P}}_L(n, \omega)$ and $\hat{\mathcal{P}}_R(n, \omega)$, which is quantified using the Kullback Lebler information divergence as [56]

$$\mathcal{D}_{\text{KL}}(n) = \frac{1}{K} \sum_{k=0}^{K-1} \left(\frac{\hat{\mathcal{P}}_L(n, \omega_k)}{\hat{\mathcal{P}}_R(n, \omega_k)} + \frac{\hat{\mathcal{P}}_R(n, \omega_k)}{\hat{\mathcal{P}}_L(n, \omega_k)} \right), \quad (4.17)$$

where ω_k with $k = 0, 1, \dots, K-1$ are the normalized frequency points at which the densities are computed, and K is the total number of frequency bins. The phase transit point can subsequently be estimated as the instant where $\mathcal{D}_{\text{KL}}(n)$

achieves its maximum value, as given by

$$\hat{\tau}^p = \operatorname{argmax} \hat{\mathcal{D}}_{\text{KL}}(n), \quad L - 1 \leq n \leq N - L. \quad (4.18)$$

It is worth noting that KLID is not a distance but a divergence measure and therefore, while a large $\mathcal{D}_{\text{KL}}(n)$ tends to imply large discrepancy between density functions, the maximum of $\mathcal{D}_{\text{KL}}(n)$ does not necessarily correspond to maximum variation in spectral density.

4.3.3 Hermitian angle-based onset detection

The TiF-HA method [44] determines signal arrival as its onset point, and detects this point based on the idea of converting the short-time Fourier transform (STFT) coefficients of the received signal into Hermitian angles. Each length- L data frame $\mathbf{x}(m)$ is first transformed into the time-frequency (TF) domain using the $K \times K$ discrete Fourier transform matrix $\mathcal{F}_K$. The STFT obtained is given by

$$\begin{aligned} \underline{\mathbf{x}}(m) &= \mathcal{F}_K \mathbf{x}(m) \\ &= [\underline{x}(0, m), \dots, \underline{x}(k, m), \dots, \underline{x}(K - 1, m)]^T. \end{aligned} \quad (4.19)$$

To mitigate the effect of dispersion, only $\underline{x}(k, m)$ within a sufficiently narrow frequency band is considered. Within the band, consisting of frequency bins $k_b, k_b + 1, \dots, k_e$, variation in propagation velocity with frequency can be assumed negligibly small.

The magnitude of the STFT coefficients are shown in Fig. 4.5(a). Since $\|\underline{x}(k, m)\|$ may vary without any theoretical upper bound, it is challenging to determine an adaptive threshold to distinguish vibration signal from background noise. Therefore, each $\underline{x}(k, m)$ is being mapped to an angle to facilitate onset detection. Since $\underline{x}(k, m)$ is complex, the concept of Hermitian angle (HA) in the complex vector space is utilized. The advantage of HA over $\|\underline{x}(k, m)\|$ is that

4.3 Methods for indirect TDOA estimation

Figure 4.5: A typical sensor-received signal generated by a finger tap on a glass plate (a) in the TF domain via STFT, (b) the HAs computed from the STFT coefficients, and (c) the standard deviation of HAs across selected frequency bins.

the angle is bounded by $[0, \pi/2]$ and hence any thresholding process involved is simplified.

The resultant HA for each TF point (k, m) within the region limited by the chosen frequency range is given by

$$\theta(k, m) = \cos^{-1} \left| \frac{\|\underline{x}(k, m)\|}{\sqrt{\|\underline{x}(k, m)\|^2 + \|\underline{w}(k, m)\|^2}} \right|, \quad (4.20)$$

where $\|\underline{w}(k, m)\|$ is the reference noise that is randomly generated for each TF point. Figure 4.5(b) shows the HAs obtained from the STFT in Fig. 4.5(a).

It has been observed that when $\|\underline{x}(k, m)\|$ and $\|\underline{w}(k, m)\|$ vary within the same range, the distribution of $\theta(k, m)$ is widely spread over $[0, \pi/2]$ compared to when one varies significantly with respect to the other. Therefore, by choosing $\underline{w}(k, m)$ to have approximately the same power as that of the background noise, the resultant $\theta(k, m)$ exhibits high variation during the noise period. On the other hand, over the signal period, since $\|\underline{x}(k, m)\| \gg \|\underline{w}(k, m)\|$, $\theta(k, m)$ varies modestly within a neighborhood of 0. Therefore, the distribution of $\theta(k, m)$ is more widely spread over the noise period compared to that over the signal period.

As a result, by quantifying the distribution spread by the standard deviation across frequency bins $k_b, \dots, k_e$, the resultant standard-deviation sequence $\hat{\sigma} = [\hat{\sigma}(0), \dots, \hat{\sigma}(N - M + 1)]^T$ would exhibit a pronounced transition, as illustrated in Fig. 4.5(c). The onset point can then be determined by thresholding $\hat{\sigma}$, where each sequence element is given by

$$\hat{\sigma}(m) = \sqrt{\frac{1}{\Delta k} \sum_{k=k_b}^{k_e} [\theta(k, m) - \hat{\mu}(m)]^2}, \quad (4.21)$$

given that

$$\hat{\mu}(m) = \frac{1}{\Delta k} \sum_{k=k_b}^{k_e} \theta(k, m) \quad (4.22)$$

is the mean and $\Delta k = k_e - k_b + 1$. Here, both $\hat{\mu}(m)$ and $\hat{\sigma}(m)$ are bounded since $0 \leq \theta(k, m) \leq \pi/2$.

It is worth noting that an ideal threshold for $\hat{\sigma}$ may not always exist, and even when it does, there may not be a unique threshold. As a consequence, estimation of the onset point, which highly depends on the choice of threshold, is not consistent and may be prone to errors, leading to unsatisfactory localization performance.

4.4 Concluding Remark

In this chapter, existing methods for direct and indirect TDOA estimation have been reviewed. The advantage of indirect TDOA estimation is that waveform distortion due to multipath can be avoided by focusing on the signal's arrival. However, determining the TOA of the signal is a challenge itself. Existing techniques assume an abrupt noise-to-signal transition at each sensor output so that the TOA can be estimated as the instant where the transition occurs. This assumption is not always valid, especially when the distance from the impact location to the sensor is large. With the noise-to-signal transition occurring gradually over a relatively long duration, the TOA is effectively a point within the

4.4 Concluding Remark

transition period. In subsequent chapters, different algorithms will be proposed in attempt to mitigate the uncertainty in TOA estimation that is imposed by gradual transitions.

Chapter 5

Hermitian Angle Distribution for TOA Estimation

As discussed in Chapter 4, existing methods for TOA estimation such as TiF-HA or those based on wavelet or KLID rely on the assumption that the noise-to-signal transition is abrupt. In situations where the impact is further away from the sensor, this transition is gradual, imposing high uncertainty on TOA estimates obtained via such methods. This chapter addresses the problem of gradual transition by proposing a new a statistical framework where the time-domain sensor output is converted to probability distributions that vary with time. The algorithm proposed in this chapter can be considered as a generalization of TiF-HA to handle gradual noise-to-signal transitions.

As opposed to the TiF-HA method, where STFT coefficients of the received signal are converted into Hermitian angles, the TOA-HAD algorithm proposed here converts the STFT coefficients into *Hermitian angle distributions* (HADs). While the resultant HADs vary with time, significant variation is exhibited only within the transition period. Therefore, by proposing a statistical measure to quantify the time variation of HADs, the transition period can be determined. This facilitates TOA estimation even when the noise-to-signal transition is gradual. In addition, the proposed HAD framework enables *simultaneous* TOA estimation across all the sensors. This effectively reduces the variance of TDOA estimates compared to existing techniques where TOAs are estimated *independently*.

5.1 Generalization of TiF-HA from a statistical perspective

dently, since, for the latter, the variance of the TDOA estimate for each sensor pair equals the sum of the variances of the respective TOA estimates.

This chapter is organized as follows: In Section 5.1, the TiF-HA method will be analyzed and generalized from a statistical perspective. The construction of HADs from STFT coefficients of a received signal is subsequently proposed in Section 5.2. The concepts of signal and noise contributions in HADs will also be introduced in this section. These concepts quantify HAD variation over the noise-to-signal transition, and are utilized to obtain optimal TOA and TDOA estimates. The performance of the proposed algorithm is evaluated against existing techniques in Section 5.3 using taps generated by finger and stylus on aluminum and glass surfaces.

5.1 Generalization of TiF-HA from a statistical perspective

In statistics, a *sample* drawn from a distribution may refer to one or a set of realizations of the random variable governed by that distribution [57, 58]. In this chapter, a sample is referred to as a set of realizations. With that in mind, for each time-frame index m , the set $\{\theta(k, m) \mid k = k_b, \dots, k_e\}$ can be considered as a sample of size Δk drawn from a distribution $F_\Theta(\theta; m)$. In other words, each $\theta(k, m)$ is a realization of a time-varying random variable $\Theta(m)$ which is governed by $F_\Theta(\theta; m)$. The notation $F_\Theta(\theta; m)$ is termed the Hermitian angle distribution (HAD) at time frame m . Effectively, $\hat{\sigma}(m)$ in Section 4.3.3 is an approximate measure of the statistical spread of $F_\Theta(\theta; m)$. In practice, however, $\hat{\sigma}(m)$ does not always exhibit an abrupt high-to-low transition, especially for a signal received by a sensor far away from the impact location, as shown in Fig. 5.1. In addition, as seen in Figs. 5.1(a) and 5.1(b), $\hat{\sigma}(m)$ can achieve high values over the signal period, resulting in spurious peaks that adversely affect the determination of an optimal threshold for $\hat{\sigma}$. Since a modest change in the choice of threshold may

Figure 5.1: The HA standard-deviation sequences, obtained via the TiF-HA method, with a zoom-in view on the transition period for signals received at sensors mounted (a) near and (b) far away from the impact location.

cause a significant change in the TOA estimate when the transition is not abrupt, TOA estimation is prone to error and localization performance is degraded. It is therefore desirable to quantify the variation in HAD with a statistical measure that exhibits clear distinction in values between the noise and signal periods.

Although $F_{\Theta}(\theta; m)$ is not available in a practical scenario, the histogram of its sample $\{\theta(k, m) \mid k = k_b, \dots, k_e\}$ can approximately describe the probability density function (PDF) $f_{\Theta}(\theta; m)$, which is the derivative of $F_{\Theta}(\theta; m)$. For the signal received at a sensor located far away with $\hat{\sigma}(m)$ shown in Fig. 5.1(b), the histograms of $\{\theta(k, m) \mid k = k_b, \dots, k_e\}$ are plotted for each time frame m in Fig. 5.2. This series of histograms is illustrated over the interval $[0, \pi/2]$ on the angle axis, which has been evenly divided into twenty sub-intervals. Subsequently, for each time frame m , the number of elements of $\{\theta(k, m) \mid k = k_b, \dots, k_e\}$ that are included in each of the sub-intervals is plotted. From a statistical perspective, if the realizations $\theta(k, m)$ over the noise and signal periods are, respectively, generated according to a noise and a signal process, $\theta(k, m)$ over the transition period will be generated according to a combination of the two processes such

Figure 5.2: The variation of the histograms of $\{\theta(k, m) \mid k = k_b, \dots, k_e\}$ with respect to time for the signal received at a sensor mounted far away from the impact location.

that the contribution of the signal process becomes increasingly dominant with time. As demonstrated in Fig. 5.2, while the noise process contributes $\theta(k, m)$ of all values within $[0, \pi/2]$ with relatively even probability, the signal process contributes mostly $\theta(k, m)$ of small values within a neighborhood of 0.

Comparing Figs. 5.1(b) and 5.2, although $\hat{\sigma}(m)$ exhibits high value within the signal period intermittently, the signal contribution remains dominant due to $\theta(k, m)$ being concentrated at small values. As a result, compared to $\hat{\sigma}(m)$ shown in Fig. 5.1(b), a statistical measure that quantifies the probabilistic contribution of signal in each $f_{\Theta}(\theta; m)$ will exhibit clear distinction in value between the noise and signal periods. Utilization of such a statistic for TOA estimation will be proposed later. Since the probabilistic contribution takes into account all the HA values within $[0, \pi/2]$, it is desirable to approximate the true $f_{\Theta}(\theta; m)$ by a function that is continuous with respect to θ . Such an approximation for $f_{\Theta}(\theta; m)$ will also be proposed in place of the histogram of $\{\theta(k, m) \mid k = k_b, \dots, k_e\}$.

5.2 The proposed TOA-HAD algorithm

5.2.1 Conversion of STFT coefficients to HADs

In what follows, the estimates of $F_{\Theta}(\theta; m)$ and $f_{\Theta}(\theta; m)$ which are continuous with respect to θ are derived. These functions serve to track the noise-to-signal transition at each sensor in this proposed framework. As described in Section 5.1, $\{\theta(k, m) \mid k = k_b, \dots, k_e\}$ is a sample drawn from $F_{\Theta}(\theta; m)$. In a similar manner, the sets $\{\underline{w}(k, m) \mid k = k_b, \dots, k_e\}$ and $\{\underline{x}(k, m) \mid k = k_b, \dots, k_e\}$ in Section 4.3.3 can also be considered as samples drawn respectively from a reference noise distribution $F_{\underline{W}}(w)$ and an STFT-coefficient distribution $F_{\underline{X}}(x; m)$. Here, $F_{\underline{W}}(w)$ and $F_{\underline{X}}(x; m)$ are distributions over the complex plane and correspond to random variables $\underline{W}$ and $\underline{X}(m)$, respectively. The variable $\underline{W}$ is not dependent on m since the same reference noise distribution has been used for all m . On the contrary, $\underline{X}(m)$ and $\Theta(m)$ are random variables with time-varying distributions. Note that while $F_{\underline{X}}(x; m)$ is attributed to the received signal $x(n)$, $F_{\underline{W}}(w)$ is to be chosen such that the resultant $F_{\Theta}(\theta; m)$ exhibits desirable properties.

Before estimating $F_{\Theta}(\theta; m)$, its expression will first be derived by assuming that $F_{\underline{W}}(w)$ and $F_{\underline{X}}(x; m)$ are known. Similar to (4.20), the HA random variable at time frame m can be obtained from $\underline{W}$ and $\underline{X}(m)$ as

$$\Theta(m) = \cos^{-1} \left| \frac{\|\underline{X}(m)\|}{\sqrt{\|\underline{X}(m)\|^2 + \|\underline{W}\|^2}} \right|. \quad (5.1)$$

Here, $\|\underline{X}(m)\|$ and $\|\underline{W}\|$, representing the magnitudes of $\underline{X}(m)$ and $\underline{W}$, respectively, are also random variables. Their respective distribution functions are given as $F_{\|\underline{X}\|}(z; m) = \int_{\|x\| \leq z} f_{\underline{X}}(x; m) dx$ and $F_{\|\underline{W}\|}(u) = \int_{\|w\| \leq u} f_{\underline{W}}(w) dw$, where $f_{\underline{X}}(x; m)$ and $f_{\underline{W}}(w)$ are density functions corresponding to $F_{\underline{X}}(x; m)$ and $F_{\underline{W}}(w)$, respectively. Here, $z, u \in \mathbb{R}$ and hence $F_{\|\underline{X}\|}(z; m)$ and $F_{\|\underline{W}\|}(u)$ are distributions over the real line.

From (5.1), the (cumulative) distribution function of $\Theta(m)$ is, by definition,

5.2 The proposed TOA-HAD algorithm

given as

$$\begin{aligned}
 F_{\Theta}(\theta; m) &= \Pr \left\{ \cos^{-1} \frac{\|\underline{X}(m)\|}{\sqrt{\|\underline{X}(m)\|^2 + \|\underline{W}\|^2}} \leq \theta \right\} \\
 &= \Pr \{ \|\underline{W}\| \leq \|\underline{X}(m)\| \tan \theta \} \\
 &= \int_{z=0}^{+\infty} \int_{u=0}^{z \tan \theta} f_{\|\underline{X}\| \|\underline{W}\|}(z, u; m) du dz, \tag{5.2}
 \end{aligned}$$

where $f_{\|\underline{X}\| \|\underline{W}\|}(z, u; m) = f_{\|\underline{X}\|}(z; m) f_{\|\underline{W}\|}(u)$ is the joint PDF of $\|\underline{X}\|$ and $\|\underline{W}\|$. This joint PDF is given as the product of their PDFs since the two random variables are independent.

Utilizing the empirical probability of its sample $\{\|\underline{x}(k, m)\| \mid k = k_b, \dots, k_e\}$ to approximate $f_{\|\underline{X}\|}(z; m)$ for each time frame m , $F_{\Theta}(\theta; m)$ is estimated from (5.2) as

$$\hat{F}_{\Theta}(\theta; m) = \frac{1}{\Delta k} \sum_{k=k_b}^{k_e} \int_0^{\|\underline{x}(k, m)\| \tan \theta} f_{\|\underline{W}\|}(u) du \tag{5.3}$$

from which its PDF is correspondingly estimated as

$$\begin{aligned}
 \hat{f}_{\Theta}(\theta; m) &= \frac{d\hat{F}_{\Theta}}{d\theta}(\theta; m) \\
 &= \frac{\sec^2 \theta}{\Delta k} \sum_{k=k_b}^{k_e} \|\underline{x}(k, m)\| f_{\|\underline{W}\|}(\|\underline{x}(k, m)\| \tan \theta), \tag{5.4}
 \end{aligned}$$

where $\Delta k = k_e - k_b + 1$. From (5.3) and (5.4), the resultant HADs are dependent on the magnitude distribution of the chosen reference noise. This fact is utilized to determine $f_{\underline{W}}(w)$ so that the HADs exhibit distinctive properties between noise and signal periods.

Assuming that $\underline{x}(k, m)$ exhibits higher power over the signal period, by selecting $f_{\underline{W}}(w)$ with a cut-off magnitude ϑ_0 , i.e., $f_{\underline{W}}(w) = 0 \ \forall w : \|w\| > \vartheta_0$, so that its magnitude distribution satisfies

$$f_{\|\underline{W}\|}(u) = 0 \quad \forall u > \vartheta_0, \tag{5.5}$$

the distinction between *noise HADs* and *signal HADs* can be achieved. More

Figure 5.3: (a) Evolution of HADs over time for a real signal generated by finger on a glass surface and (b) HADs at time frames before, during and after transition.

specifically, at any time frame m within the signal period, for any choice of ϑ_0 , there exists θ_0 such that $\|\underline{x}(k, m)\| \tan \theta_0 > \vartheta_0 \forall k \in \{k_b, \dots, k_e\}$, which, according to (5.4) and (5.5), results in $\hat{f}_\Theta(\theta; m) = 0 \forall \theta > \theta_0$. This implies that θ_0 serves as a cut-off angle of $\hat{f}_\Theta(\theta; m)$. Given a fixed choice of ϑ_0 , the resultant θ_0 is high when $\|\underline{x}(k, m)\|$ are low, and vice versa. Therefore, since $\|\underline{x}(k, m)\|$ exhibit lower values during the noise period compared to the signal period, as illustrated in Fig. 4.5(a), $\hat{f}_\Theta(\theta; m)$ within the noise period are more widely spread than those within the signal period. This is consistent with the observation made in [44], where $f_{\underline{W}}(w)$ is also chosen to have a cut-off magnitude. It is worth noting that $f_{\underline{W}}(w)$ is not mentioned explicitly in [44] but implied via drawing $\underline{w}(k, m)$ from a uniform distribution over a square region centered at the origin of the complex plane.

To investigate the behavior of HADs during the noise-to-signal transition, the variation of $\hat{f}_\Theta(\theta; m)$ with time is depicted in Fig. 5.3(a), while Fig. 5.3(b)

5.2 The proposed TOA-HAD algorithm

illustrates the profile of $\hat{f}_\Theta(\theta; m)$ at time frames before, during, and after the transition. For the computation of $\hat{F}_\Theta(\theta; m)$ and $\hat{f}_\Theta(\theta; m)$, there exist many possible choices of $f_{\underline{W}}(w)$ with cut-off magnitude ϑ_0 . However, as shown in the Appendix, a non-circular support on the complex plane for $f_{\underline{W}}(w)$ results in $\hat{F}_\Theta(\theta; m)$ and $\hat{f}_\Theta(\theta; m)$ being computationally expensive. The density $f_{\underline{W}}(w)$ is therefore proposed to be uniform over a circular region of radius ϑ_0 centered at the origin of the complex plane. This choice, as shown in the Appendix, simplifies the expressions for $\hat{F}_\Theta(\theta; m)$ and $\hat{f}_\Theta(\theta; m)$, and therefore reduces the computational complexity compared to that of a non-circular support. The simplified estimates are given by

$$\hat{F}_\Theta(\theta; m) = \frac{\tan^2 \theta}{\vartheta_0^2 \Delta k} \sum_{k \in D_\theta} \|\underline{x}(k, m)\|^2 + 1 - \frac{|D_\theta|}{\Delta k}, \quad (5.6)$$

and

$$\hat{f}_\Theta(\theta; m) = \frac{2 \tan \theta \sec^2 \theta}{\vartheta_0^2 \Delta k} \sum_{k \in D_\theta} \|\underline{x}(k, m)\|^2, \quad (5.7)$$

where $D_\theta = \{k \mid k \in [k_b, k_e]; 0 \leq \|\underline{x}(k, m)\| \leq \vartheta_0 \cot \theta\}$, and $|D_\theta|$ is the cardinality of D_θ .

With reference to (5.7), insights can be gained with regard to the choice of ϑ_0 that distinguishes HADs between the noise and signal periods. When the signal power is higher than that of the reference noise, i.e., $\|\underline{x}(k, m)\| > \vartheta_0$ for almost all $k \in \{k_b, \dots, k_e\}$, the condition $\|\underline{x}(k, m)\| \leq \vartheta_0 \cot \theta$ is satisfied only for sufficiently large values of $\cot \theta$ (which is equivalent to sufficiently small θ). For large θ , D_θ is almost empty, causing $\hat{f}_\Theta(\theta; m) \approx 0$. As a result, $\hat{f}_\Theta(\theta; m)$ within the signal period exhibits high concentration over the neighborhood $\theta \approx 0$, as illustrated in Fig. 5.3(b). To guarantee such behavior of $\hat{f}_\Theta(\theta; m)$ over the signal period, it is important to choose ϑ_0 sufficiently small. However, the reference noise power, which for our choice of $f_{\underline{W}}(w)$ is given by $\vartheta_0^2/2$, should not be less than the background noise power. This is to prevent $\hat{f}_\Theta(\theta; m)$ within the noise period from being concentrated over $\theta \approx 0$, which results in no distinction between noise and

Figure 5.4: Normalized ℓ^2 -distance calculated for distribution pairs of (a) the proposed $\hat{F}_\Theta(\theta; m)$ (computed using (5.6)), (b) the samples $\{\theta(k, m) \mid k = k_b, \dots, k_e\}$ used in TiF-HA [44], and (c) the STFT-coefficient magnitude distributions. Each distribution pair correspond to a pair (m_1, m_2) of time-frame indices $0 \leq m_1, m_2 < M$.

signal HADs. With that in mind, $\vartheta_0^2/2$ is chosen equal to the background noise power, which can be estimated over the silent period of the received signal. For TiF-HA [44], while the choice of $f_{\underline{W}}(w)$, implicitly made via $\underline{w}(k, m)$, is different from the choice made here, the reference noise power is made equal to that of the background noise for the same reason.

To illustrate the inter-dissimilarity among HADs of different time frames, and contrast with that among STFT-coefficient magnitude distributions, the ℓ^2 -

5.2 The proposed TOA-HAD algorithm

distance between distributions of same type is computed for all time-frame pairs (m_1, m_2) , $0 \leq m_1, m_2 < M$. Here, the ℓ^2 -distance between two distribution functions F_1 and F_2 , corresponding respectively to time frames m_1 and m_2 , is given by

$$\ell^2(F_1, F_2) = \sqrt{\int_{-\infty}^{+\infty} |F_1(x) - F_2(x)|^2 dx} . \quad (5.8)$$

For discrete samples, the empirical cumulative distribution functions are deployed in (5.8). The inter-dissimilarity between distributions is quantified by their ℓ^2 -distance normalized such that the maximum distance equals unity, and the results are shown in Fig. 5.4. For these plots, the dark regions represent time-frame pairs where $\ell^2(F_1, F_2) \approx 0$ corresponding to low dissimilarity.

The inter-dissimilarity among $\hat{F}_\Theta(\theta; m)$ computed from (5.6) is illustrated in Fig. 5.4(a). It can be seen that $\hat{F}_\Theta(\theta; m)$ from the noise period exhibit high dissimilarity with those from the signal period, while those from the same period exhibit low dissimilarity. Fig. 5.4(b) shows the inter-dissimilarity among the samples $\{\theta(k, m) \mid k = k_b, \dots, k_e\}$. Note that both Figs. 5.4(a) and 5.4(b) illustrate the dissimilarity between HADs, but using different approximations for the true $F_\Theta(\theta; m)$, namely $\hat{F}_\Theta(\theta; m)$ and $\{\theta(k, m) \mid k = k_b, \dots, k_e\}$. Comparing Figs. 5.4(a) and 5.4(b), it can be seen that, over both the noise and signal periods, the HADs approximated by the proposed $\hat{F}_\Theta(\theta; m)$ exhibit lower dissimilarity, represented by the darker and more uniformly shaded areas. This is due to the fact that $\theta(k, m)$ in TiF-HA [44] are obtained using the discrete realizations $\underline{w}(k, m)$ drawn from the distribution $F_{\underline{W}}(w)$, while, for the proposed approach, $\hat{F}_\Theta(\theta; m)$ are directly computed from $F_{\underline{W}}(w)$ itself, which is continuous. The latter is therefore a better approximation to the true HAD.

The inter-dissimilarity among the samples $\{\|\underline{x}(k, m)\| \mid k = k_b, \dots, k_e\}$ is illustrated in Fig. 5.4(c). Comparing Figs. 5.4(a) and 5.4(c), it can be seen that the STFT-coefficient magnitude distributions exhibit high dissimilarity over the signal period, as opposed to HADs. This is due to the non-stationary nature of

Figure 5.5: Block diagram of the proposed TOA-HAD algorithm.

vibration signals. It is therefore highly desirable to determine the signal period after converting STFT coefficients to HADs.

5.2.2 Change point detection

As described in Section 5.1, the TOA is determined as the change point between noise and signal processes. In the presence of gradual noise-to-signal transition, this change point is to be determined as the time instant when the signal process dominates the background noise process. This is achieved by developing an algorithm based on the concepts of noise and signal contributions in each HAD as

5.2 The proposed TOA-HAD algorithm

illustrated in Fig. 5.5.

Denoting the beginning and end points of the transition period by τ^b and τ^e , respectively, the length of the transition period is given by $\tau^e - \tau^b + 1$. This length determines whether the change occurs abruptly or gradually, with $\tau^b = \tau^e$ implying an abrupt change, where the transition period reduces to a single point. To analyze the signal, the sensor output is partitioned into three segments, the noise period $0 \leq m < \tau^b$, the transition period $\tau^b \leq m \leq \tau^e$, and the signal period $\tau^e < m < M$. Since $F_\Theta(\theta; m)$ exhibits low dissimilarity within the noise period as well as the signal period, it can therefore be assumed that, over the noise period, $\Theta(m)$ is governed by a common background noise HAD $F_\Theta^N(\theta)$, which implies

$$F_\Theta(\theta; m) = F_\Theta^N(\theta), \quad m = 0, \dots, \tau^b - 1. \quad (5.9)$$

Similarly, over the signal period, $\Theta(m)$ is governed by a common signal HAD, denoted by $F_\Theta^S(\theta)$, i.e.,

$$F_\Theta(\theta; m) = F_\Theta^S(\theta), \quad m = \tau^e + 1, \dots, M - 1. \quad (5.10)$$

Over the noise-to-signal transition period, $\Theta(m)$ is governed by both $F_\Theta^N(\theta)$ and $F_\Theta^S(\theta)$ where the latter dominates with time. In what follows, to quantify the dominance of one process over the other, the contributions of the noise and signal processes in each $F_\Theta(\theta; m)$ are to be defined. It is useful to note that since $F_\Theta(\theta; m)$ is dependent on time, the noise and signal contributions also vary with time. Before constructing the concepts of noise and signal contributions, the estimation of $F_\Theta^N(\theta)$ and $F_\Theta^S(\theta)$ from $\hat{F}_\Theta(\theta; m)$ will first be discussed in the next section.

5.2.3 Estimation of noise and signal HADs

In a real-application scenario, $F_\Theta^N(\theta)$ and $F_\Theta^S(\theta)$ can be estimated from $\hat{F}_\Theta(\theta; m)$ in (5.6) via two steps: determining the noise period and the signal period, followed

Figure 5.6: Estimates of (a) the density $\hat{f}_{\Theta}^N(\theta)$ of noise HA and (b) the density $\hat{f}_{\Theta}^S(\theta)$ of signal HA, and their contributions in $\hat{f}_{\Theta}(\theta; m)$ at time frames (c) before transition and (d) after transition.

by estimating $F_{\Theta}^N(\theta)$ and $F_{\Theta}^S(\theta)$ from $\hat{F}_{\Theta}(\theta; m)$ computed over these periods.

To determine the noise and signal periods, it suffices to estimate the transition period $[\tau^b, \tau^e]$, as described in [59]. The noise and signal periods are then respectively determined as $[\tau^b - L_N, \tau^b - 1]$ and $[\tau^e + 1, \tau^e + L_S]$, where L_N and L_S are their respective lengths, which can be determined empirically. These lengths should be sufficiently large so that a sufficient number of HADs are involved in the estimation of $F_{\Theta}^N(\theta)$ and $F_{\Theta}^S(\theta)$ to achieve high confidence in the resultant estimates. However, since our focus is on the noise-to-signal transition and data frames close to the transition period will provide more useful information, the lengths are limited to $L_N = L_S = 50$ samples. Such small values of L_N and L_S will also help reduce computational cost. The distributions $F_{\Theta}^N(\theta)$ and $F_{\Theta}^S(\theta)$ are then estimated as

$$\hat{F}_{\Theta}^N(\theta) = \sum_{m=\tau^b-L_N}^{\tau^b-1} \hat{F}_{\Theta}(\theta; m) \quad \text{and} \quad \hat{F}_{\Theta}^S(\theta) = \sum_{m=\tau^e+1}^{\tau^e+L_S} \hat{F}_{\Theta}(\theta; m). \quad (5.11)$$

5.2 The proposed TOA-HAD algorithm

Their density function estimates are respectively given by

$$\hat{f}_{\Theta}^N(\theta) = \sum_{m=\tau^b-L_N}^{\tau^b-1} \hat{f}_{\Theta}(\theta; m) \quad \text{and} \quad \hat{f}_{\Theta}^S(\theta) = \sum_{m=\tau^e+1}^{\tau^e+L_S} \hat{f}_{\Theta}(\theta; m). \quad (5.12)$$

Figs. 5.6(a) and 5.6(b), respectively, illustrate $\hat{f}_{\Theta}^N(\theta)$ and $\hat{f}_{\Theta}^S(\theta)$ which have been calculated from $\hat{f}_{\Theta}(\theta; m)$ shown in Fig. 5.3(a). It can be seen that $\hat{f}_{\Theta}^N(\theta)$ is relatively flat with a wide support similar to $\hat{f}_{\Theta}(\theta; m)$ before transition, while $\hat{f}_{\Theta}^S(\theta)$ exhibits a steep slope and a narrow support similar to $\hat{f}_{\Theta}(\theta; m)$ after transition. This is as expected since these estimates are obtained via averaging functions of similar shapes.

5.2.4 Noise and signal contributions in HAD

Having $F_{\Theta}^N(\theta)$ and $F_{\Theta}^S(\theta)$ estimated via (5.11), to calculate their respective probability contribution to $F_{\Theta}(\theta; m)$, we first compute their contributions for each possible value of $\Theta(m)$. Provided that a realization of $\Theta(m)$ has been observed with value $\theta \in [0, \pi/2]$, it is desirable to compute the probability that $\Theta(m)$ is governed by the noise/signal process. Since $F_{\Theta}^N(\theta)$ and $F_{\Theta}^S(\theta)$ are continuous with respect to θ , $\Pr \{ \Theta(m) \sim F_{\Theta}^N(\theta) \mid \Theta(m) = \theta \} = 0$ and $\Pr \{ \Theta(m) \sim F_{\Theta}^S(\theta) \mid \Theta(m) = \theta \} = 0$. The probability contributions of $F_{\Theta}^N(\theta)$ and $F_{\Theta}^S(\theta)$ will therefore be compared within an interval $(\theta - \Delta\theta, \theta + \Delta\theta)$, for some $\Delta\theta > 0$, rather than at a single value θ . If $\Theta(m)$ is governed by $F_{\Theta}^N(\theta)$, then the probability of $\Theta(m)$ falling within $(\theta - \Delta\theta, \theta + \Delta\theta)$ is given by

$$\Delta F_{\Theta}^N = F_{\Theta}^N(\theta + \Delta\theta) - F_{\Theta}^N(\theta - \Delta\theta). \quad (5.13)$$

On the other hand, if $\Theta(m)$ is governed by $F_{\Theta}^S(\theta)$, the same is given by

$$\Delta F_{\Theta}^S = F_{\Theta}^S(\theta + \Delta\theta) - F_{\Theta}^S(\theta - \Delta\theta). \quad (5.14)$$

The contribution of either $F_{\Theta}^N(\theta)$ or $F_{\Theta}^S(\theta)$ to each HA value $\theta \in [0, \pi/2]$ is then obtained as

$$\begin{aligned}
p_N(\theta) &= \Pr \left\{ \Theta(m) \sim F_{\Theta}^N(\theta) \mid \Theta(m) \underset{\Delta\theta \rightarrow 0}{\in} (\theta - \Delta\theta, \theta + \Delta\theta) \right\} \\
&= \lim_{\Delta\theta \rightarrow 0} \frac{\Delta F_{\Theta}^N}{\Delta F_{\Theta}^N + \Delta F_{\Theta}^S} \\
&= \frac{f_{\Theta}^N(\theta)}{f_{\Theta}^N(\theta) + f_{\Theta}^S(\theta)},
\end{aligned} \tag{5.15}$$

and

$$\begin{aligned}
p_S(\theta) &= \Pr \left\{ \Theta(m) \sim F_{\Theta}^S(\theta) \mid \Theta(m) \underset{\Delta\theta \rightarrow 0}{\in} (\theta - \Delta\theta, \theta + \Delta\theta) \right\} \\
&= \lim_{\Delta\theta \rightarrow 0} \frac{\delta F_{\Theta}^S(\theta)}{\delta F_{\Theta}^N(\theta) + \delta F_{\Theta}^S(\theta)} \\
&= \frac{f_{\Theta}^S(\theta)}{f_{\Theta}^N(\theta) + f_{\Theta}^S(\theta)},
\end{aligned} \tag{5.16}$$

where $f_{\Theta}^N(\theta)$ and $f_{\Theta}^S(\theta)$ are density functions corresponding to the distributions $F_{\Theta}^N(\theta)$ and $F_{\Theta}^S(\theta)$, respectively. It is useful to note from (5.15) and (5.16) that $p_N(\theta) + p_S(\theta) = 1$.

The noise contribution in each $F_{\Theta}(\theta; m)$ is then obtained by integrating $p_N(\theta)$ over all possible values of $\theta \in [0, \pi/2]$, i.e.,

$$\begin{aligned}
\gamma^N(m) &= \Pr\{\Theta(m) \sim F_{\Theta}^N(\theta)\} \\
&= \int_0^{\pi/2} p_N(\theta) f_{\Theta}(\theta; m) d\theta,
\end{aligned} \tag{5.17}$$

where $f_{\Theta}(\theta; m)$ is the PDF of $\Theta(m)$. The signal contribution in $F_{\Theta}(\theta; m)$ is correspondingly given as

$$\begin{aligned}
\gamma^S(m) &= \Pr\{\Theta(m) \sim F_{\Theta}^S(\theta)\} \\
&= \int_0^{\pi/2} p_S(\theta) f_{\Theta}(\theta; m) d\theta.
\end{aligned} \tag{5.18}$$

Figs. 5.6(c) and 5.6(d), respectively, illustrate contributions of noise and signal processes for $\hat{f}_{\Theta}(\theta; m)$ estimated at time frames before and after transition. In

5.2 The proposed TOA-HAD algorithm

these plots, the darker region, with the area of $\gamma^S(m)$, is stacked over the lighter region, which corresponds to $\gamma^N(m)$. It can be seen that, the noise contribution dominates for $f_\Theta(\theta; m)$ before transition, as opposed to the dominance of the signal contribution for $f_\Theta(\theta; m)$ after transition.

Since $p_N(\theta) + p_S(\theta) = 1$, it follows from (5.17) and (5.18) that $\gamma^N(m) + \gamma^S(m) = 1$, $0 \leq m < M$. It therefore suffices to analyze the noise-to-signal transition using either $\gamma^N(m)$ or $\gamma^S(m)$. Fig. 5.7 shows the signal contributions computed for all the time-frame indices m . It can be seen that $\gamma^S(m)$ increases with time, which also implies that $\gamma^N(m)$ decreases with time. In addition, $\gamma^S(m)$ is consistently low during the noise period and consistently high after signal arrival. This is due to the fact that the respective HADs vary insignificantly over the corresponding noise/signal period, which has been utilized for (5.9) and (5.10). Substituting (5.9) into (5.18), the signal contribution in each HAD over the noise period is given by

$$\gamma_-^S = \int_0^{\pi/2} p_S(\theta) f_\Theta^N(\theta) d\theta. \quad (5.19)$$

Similarly, substituting (5.10) into (5.18), the signal contribution in each signal HAD is given by

$$\gamma_+^S = \int_0^{\pi/2} p_S(\theta) f_\Theta^S(\theta) d\theta. \quad (5.20)$$

The signal contribution in HAD therefore increases from γ_-^S to γ_+^S during the transition period, as illustrated in Fig. 5.7. The TOA is then estimated as the instant when $\gamma^S(m)$ exceeds a threshold $\gamma_-^S < \gamma^t < \gamma_+^S$, which signifies the dominance of $\gamma^S(m)$ over $\gamma^N(m)$. It is worth noting that, since $\gamma^N(m) + \gamma^S(m) = 1$, the statistic measure $\gamma^S(m)$ has effectively been normalized and therefore, γ^t is common for all the sensor data.

5.2.5 Determination of γ^t

As opposed to existing TOA-based localization algorithms which focus on TOA estimation at each sensor individually [41, 43, 44], the common threshold γ^t allows

Figure 5.7: Variation in the signal contribution $\gamma^S(m)$ with respect to time. The TOA m^t is determined as when $\gamma^S(m)$ exceeds γ^t .

TOAs to be estimated simultaneously across all the sensors. The TOA estimates are optimized for TDOA estimation error minimization, hence results in good localization performance.

Let $\gamma_i^S(m)$ denote the signal contribution sequence corresponding to the i th sensor, $i = 1, \dots, R$, where R is the number of sensors. While the noise-to-signal transition may not be abrupt, for simplicity, assume that $\gamma_i^S(m)$ increases linearly with slope η_i . Although this assumption may not be entirely valid, the algorithm achieves good results as will be shown in Section 5.3. The variation of $\gamma_i^S(m)$ over the transition period is then given as

$$\gamma_i^S(m) = \eta_i(m - \tau_i^b), \quad m \in [\tau_i^b, \tau_i^e], \quad (5.21)$$

where $[\tau_i^b, \tau_i^e]$ is the transition period at the i -th sensor. The arrival time of the signal, determined as when $\gamma_i^S(m)$ exceeds γ^t , is then given by

$$\tau_i = \frac{\gamma^t}{\eta_i} + \tau_i^b. \quad (5.22)$$

For an estimate $\hat{\gamma}^t$ of the threshold, the TOA is estimated as

$$\hat{\tau}_i = \frac{\hat{\gamma}^t}{\eta_i} + \tau_i^b. \quad (5.23)$$

The estimation error for the TDOA between a sensor pair (i, j) , $1 \leq i, j \leq R$, is

5.2 The proposed TOA-HAD algorithm

then given by

$$\begin{aligned}\Delta\tau_{i,j} &= (\hat{\tau}_i - \hat{\tau}_j) - (\tau_i - \tau_j) \\ &= \left(\frac{1}{\eta_i} - \frac{1}{\eta_j}\right) (\hat{\gamma}^t - \gamma^t).\end{aligned}\quad (5.24)$$

Since γ^t is common across all sensors, it implies that $\gamma_{-,i}^S \leq \gamma^t \leq \gamma_{+,i}^S$, $i = 1, \dots, R$, where $\gamma_{-,i}^S$ and $\gamma_{+,i}^S$ are, respectively, the signal contributions in noise HAD and signal HAD at the i th sensor. It follows that

$$\gamma_{-, \max}^S \leq \gamma^t \leq \gamma_{+, \min}^S, \quad (5.25)$$

where $\gamma_{-, \max}^S = \max\{\gamma_{-,i}^S \mid i = 1, \dots, R\}$, and $\gamma_{+, \min}^S = \min\{\gamma_{+,i}^S \mid i = 1, \dots, R\}$. Knowing only the bounds $\gamma_{-, \max}^S$ and $\gamma_{+, \min}^S$, it is equally likely for γ^t to take any value within the interval $[\gamma_{-, \max}^S, \gamma_{+, \min}^S]$. Therefore, γ^t can be assumed to be uniformly distributed over $[\gamma_{-, \max}^S, \gamma_{+, \min}^S]$ with probability density $1/(\gamma_{+, \min}^S - \gamma_{-, \max}^S)$.

The estimate $\hat{\gamma}^t$ is then chosen such that it minimizes the sum of squared TDOA errors between all sensor pairs and across all possible values of γ^t , i.e., to minimize

$$\begin{aligned}\mathcal{J} &= \int_{\gamma_{-, \max}^S}^{\gamma_{+, \min}^S} \sum_{1 \leq i, j \leq R} (\Delta\tau_{i,j})^2 \frac{1}{\gamma_{+, \min}^S - \gamma_{-, \max}^S} d\gamma^t \\ &= \int_{\gamma_{-, \max}^S}^{\gamma_{+, \min}^S} \sum_{1 \leq i, j \leq R} \left(\frac{1}{\eta_i} - \frac{1}{\eta_j}\right)^2 \frac{(\hat{\gamma}^t - \gamma^t)^2}{\gamma_{+, \min}^S - \gamma_{-, \max}^S} d\gamma^t.\end{aligned}\quad (5.26)$$

Taking the derivative of $\mathcal{J}$ with respect to γ^t ,

$$\frac{d\mathcal{J}}{d\gamma^t} = \sum_{1 \leq i, j \leq R} \left(\frac{1}{\eta_i} - \frac{1}{\eta_j}\right)^2 \int_{\gamma_{-, \max}^S}^{\gamma_{+, \min}^S} \frac{2(\hat{\gamma}^t - \gamma^t)}{\gamma_{+, \min}^S - \gamma_{-, \max}^S} d\gamma^t, \quad (5.27)$$

from which, by setting $d\mathcal{J}/d\gamma^t = 0$, we obtain

$$\hat{\gamma}^t = \frac{\gamma_{+, \min}^S - \gamma_{-, \max}^S}{2}. \quad (5.28)$$

The threshold in (5.28) can then be applied to the sequence $\gamma_i^S(m)$ at each sensor

to obtain the optimal TOA estimates.

5.2.6 Computational complexity

The computational load for TOA-HAD is mainly contributed by the following stages: STFT computation, STFT-to-HAD conversion, and signal contribution computation. The complexity of a K -point DFT using a fast Fourier transform algorithm is $\mathcal{O}(K \log K)$. This results in a complexity of $\mathcal{O}(MK \log K)$ for computing a K -by- M STFT coefficient matrix.

In estimating the HAD for each time-frame index m according to (5.6) and (5.7), $\hat{F}_\Theta(\theta; m)$ is first computed at N_θ values $\theta_q = \pi/2 \times q/(N_\theta - 1)$, $q = 0, \dots, N_\theta - 1$. A naive implementation of (5.6) would require $\mathcal{O}(\Delta k N_\theta)$ computations. However, a more efficient implementation can be achieved by first sorting $\underline{x}(k_b, m), \dots, \underline{x}(k_e, m)$ in order of ascending magnitude, which is of $\mathcal{O}(\Delta k \log \Delta k)$ complexity. The sum in (5.6) can subsequently be computed in $\mathcal{O}(\Delta k + N_\theta)$ via the procedure described in Fig. 5.8. This results in an implementation of (5.6) with the complexity of $\mathcal{O}(\Delta k \log \Delta k + N_\theta)$. Since $\hat{f}_\Theta(\theta; m)$ is the derivative of $\hat{F}_\Theta(\theta; m)$, the values $\hat{f}_\Theta(\theta_0; m), \dots, \hat{f}_\Theta(\theta_{N_\theta-1}; m)$ can be directly computed as the finite differences of $\hat{F}_\Theta(\theta_0; m), \dots, \hat{F}_\Theta(\theta_{N_\theta-1}; m)$ in $\mathcal{O}(N_\theta)$. The complexity for each HAD estimation is therefore $\mathcal{O}(\Delta k \log \Delta k + N_\theta)$. HAD estimation is also limited for only m within the interval $[\tau^b - L_N, \tau^e + L_S]$, where $L_{NS} = L_N + L_S + \tau^e - \tau^b \ll M$, to reduce unnecessary computations. This results in a complexity of $\mathcal{O}((\Delta k \log \Delta k + N_\theta)L_{NS})$ for the stage of STFT-to-HAD conversion.

For estimating $\gamma^S(m)$, (5.18) is approximated by

$$\hat{\gamma}^S(m) = \sum_{q=0}^{N_\theta-1} p_S(\theta_q) \hat{f}_\Theta^S(\theta_q), \quad (5.29)$$

which incurs $\mathcal{O}(N_\theta)$ complexity for each time-frame index m . The overall complexity for this stage is therefore $\mathcal{O}(N_\theta L_{NS})$.

5.3 Experiment results

Input: $\underline{x}(k_b, m), \dots, \underline{x}(k_e, m)$ sorted in order of ascending magnitude

Output: $\hat{F}_\Theta(\theta; m)$ evaluated at $\theta_0, \dots, \theta_{N_\theta-1}$

```

1:  $k := 0$ 
2:  $S := 0$ 
3: for  $q = 0$  to  $N_\theta - 1$  do
4:   while  $k < \Delta k$  and  $\|\underline{x}(k, m)\| \leq \vartheta_0 \cot \theta_q$  do
5:      $S := S + \|\underline{x}(k, m)\|^2$ 
6:      $k := k + 1$ 
7:   end while
8:    $\hat{F}_\Theta(\theta_q; m) := (\tan^2 \theta_q \times S + \Delta k - k) / \Delta k$ 
9: end for
10: return  $\hat{F}_\Theta(\theta_0; m), \dots, \hat{F}_\Theta(\theta_{N_\theta-1}; m)$ 

```

Figure 5.8: Pseudo code for computing the sum in (5.6) with a computational complexity of $\mathcal{O}(\Delta k + N_\theta)$.

In summary, the complexity of the proposed TOA-HAD is $\mathcal{O}(MK \log K + (\Delta k \log \Delta k + N_\theta)L_{\text{NS}})$, which is equivalent to $\mathcal{O}(MK \log K + N_\theta L_{\text{NS}})$ since $L_{\text{NS}} \ll M$ and $\Delta k < K$. Implementation of TOA-HAD on the ADSP-21469 EZ-Kit Lite evaluation hardware requires approximately 300 ms to localize a generated impact in real-time operation.

5.3 Experiment results

The performance of the proposed TOA-HAD algorithm is evaluated using different sets of real data collected on aluminum and glass plates. The aluminum plate is of dimension 0.6 m $\times$ 0.6 m $\times$ 2.5 mm and the glass plate 0.6 m $\times$ 0.6 m $\times$ 5.0 mm. Similar to [44], impacts are generated on the surfaces of these plates by either a finger or a metal stylus. The generated impacts are captured by eight Murata PKS1-4A10 piezoelectric shock sensors mounted on the surface of each plate, 0.1 m away from the edges, as denoted by $\bullet$ in Fig. 5.9. The received analog signals are then simultaneously sampled at a sampling frequency $f_s = 96$ kHz using an 8-channel 24-bit data acquisition system.

Vibration data are collected from different experiment setups to investigate the localization performance of the proposed algorithm. Two types of impacts

Figure 5.9: Localization of the impacts due to finger taps on glass surface using TOA-HAD. The sensor coordinates, actual impact locations and estimated locations are denoted by $\bullet$, $\circ$, and $\times$, respectively.

(metal stylus and finger) are generated on the aluminum and glass plates, making up four experiment setups in total. For each of the four setups, nine impact locations are arranged on the plate surface in a 3×3 grid array. At each location, as denoted by $\circ$ in Fig. 5.9, a set of five taps are generated, comprising a total of forty-five test cases. In the following experiments, data frame length $M = 256$ and DFT size $K = 1024$ is employed for STFT computation. The overall performance of each algorithm is quantified by the root-mean-square localization error (RMSE) and the standard deviation of localization errors (STDE), measured in meter.

5.3.1 Performance vs choice of frequency range

Since the propagation speed varies with frequency in a dispersive medium, effects of the choice of frequency range on the localization performance will first be investigated. A choice of $B = [k_b, k_e]$ can be characterized by its number of frequency bins $\Delta k = k_e - k_b + 1$ and center frequency $f_c = f_s(k_b + k_e)/(2K)$. While it is important to select a sufficiently narrow bandwidth to minimize the

5.3 Experiment results

effect of dispersion, Δk should be sufficiently large to achieve good estimates of HADs in (5.3) and (5.4). There is therefore a trade-off between minimizing the effect of dispersion and the quality of HAD estimates.

The choice of the center frequency f_c can be explained via the Lamb wave theory. This theory categorizes the wave propagating in a thin plate into two types of modes: symmetric and anti-symmetric [34, 60]. The former is termed as ‘compressional’ to describe the expansion and contraction of the plate thickness. Anti-symmetric wave propagation, on the other hand, exhibits constant-thickness flexing and is known as ‘flexural’. Existing works [60, 61] have shown that at low frequencies, most of the wave energy is contained in low-order flexural modes, while compressional modes are only visible at high frequencies. The frequency range considered in this work belongs to the region where only the zero-order antisymmetric mode a_0 is present. According to [34], the propagation velocity of the a_0 mode, as a function of frequency f , is proportional to the square root of f , i.e.,

$$c(f) \propto \sqrt{f}. \quad (5.30)$$

Differentiating both sides of (5.30) with respect to f gives

$$\frac{dc}{df} \propto -\frac{1}{\sqrt{f}}, \quad (5.31)$$

which implies that the variation in velocity between high-frequency components is smaller than that between low-frequency components. As a consequence, for the same number of frequency bins, bands centered at higher frequencies exhibit lower amounts of dispersion. This implies that operating at higher frequencies allows one to select a larger number of bins while maintaining the same amount of dispersion. Therefore, a choice of sufficiently high center frequency is desirable in order to achieve the low-dispersion and many-bin criteria. There is, however, an upper limit for the choice of frequency range. This is due to the frequency response of the Murata piezo-electric sensors rolling off beyond 10 kHz.

Figure 5.10: Variation of RMSE and standard deviation of localization errors (STDE), measured in meter, on aluminum surface with respect to center frequency for different numbers of frequency bins. (a) RMSE when the impacts are due to metal stylus, (b) RMSE when the impacts are due to finger, (c) STDE when the impacts are due to metal stylus and (d) STDE when the impacts are due to finger.

In our experiments, the performance of TOA-HAD is evaluated for different choices of frequency band B on the four experiment setups, namely metal stylus on aluminum, finger on aluminum, metal stylus on glass, and finger on glass. With each frequency bin corresponding to $f_s/K = 93.75$ Hz, the frequency band $B = [k_b, k_e]$ has its center frequency at $f_c = [(k_e - k_b)/2] \times 93.75$ Hz. Figs. 5.10 and 5.11 illustrate the performance of TOA-HAD for the aluminum and the glass plates, respectively. For each of the two figures, subplots (a) and (b) illustrate the RMSE for impacts generated by a metal stylus and finger, respectively, while subplots (c) and (d) illustrate the corresponding STDE. In each of the subplots, the quantified performance (RMSE or STDE) is plotted against f_c for different numbers of frequency bins $\Delta k = 20, 40$ and 60 .

Note that the variation of STDE and that of RMSE are similar. This reflects the consistency in quantifying the algorithm performance by these two measures.

5.3 Experiment results

Figure 5.11: Variation of RMSE and standard deviation of localization errors (STDE), measured in meter, on glass surface with respect to center frequency for different numbers of frequency bins. (a) RMSE when the impacts are due to metal stylus, (b) RMSE when the impacts are due to finger, (c) STDE when the impacts are due to metal stylus and (d) STDE when the impacts are due to finger.

It can also be seen that both the RMSE and STDE are low for B centered at some mid-range frequency, which to be referred to as the *operating band*. It is clear that the aluminum and the glass plates have different operating bands. For aluminum, as can be seen from Fig. 5.10, the operating band is centered at $f_c \approx 8$ kHz. The operating band for glass, as observed from Fig. 5.11, has a lower center frequency $f_c \approx 6$ kHz.

The selected bandwidth also affects the localization performance. However, the effect is only evident for narrow bands such as where $\Delta k = 20$. The performance is relatively less sensitive to wider bands, e.g. $\Delta k = 40$ and 60 . Based on these results, $\Delta k = 50$, $f_c = 6$ kHz will be used for both the aluminum and glass plates in the subsequent experiments.

Figure 5.12: Performance comparison of the proposed TOA-HAD algorithm with that of TiF-HA [44] using (a) RMSE and (b) standard deviation of the localization errors (STDE), both measured in meter.

5.3.2 Performance comparison with existing methods

The TiF-HA algorithm [44] has been shown to outperform other techniques including the KLID algorithm [43] and the wavelet method [41]. The proposed TOA-HAD method will therefore be compared with the TiF-HA method.

Figure 5.12 shows the localization performance of the two algorithms on all the four data sets. The RMSEs of the algorithms for each data set are compared in Fig. 5.12(a) whereas Fig. 5.12(b) compares the STDEs. Note that TOA-HAD outperforms TiF-HA in terms of both RMSE and STDE. For the case of metal stylus on glass plate, TiF-HA suffers from high localization errors with RMSE exceeding 2.5 cm and STDE of more than 1.0 cm, while TOA-HAD exhibits fairly good performance with RMSE of 1.25 cm and STDE of 0.56 cm. These results translate to an improvement in performance of more than 50% for TOA-HAD over TiF-HA. In addition, for finger taps on glass plate, while TiF-HA exhibits

5.4 Concluding remark

fairly good performance with RMSE and STDE of approximately 1.3 cm and 0.7 cm, respectively, TOA-HAD achieves sub-centimeter accuracy with RMSE of 0.73 cm and STDE 0.3 cm. Overall, TOA-HAD exhibits consistency across any combination of tapping style and plate material with RMSE of 1.0 ± 0.3 cm and STDE 0.4 ± 0.1 cm. This justifies the validity of the statistical approach employed by TOA-HAD.

To further illustrate its potential application for human-computer interface, the estimated impact locations for TOA-HAD method on the glass surface are depicted in Fig. 5.9. The result for finger taps on the glass surface is shown for its popularity in touch-screen applications. As can be seen, most of the estimates lie within the neighborhood of the impact locations.

5.4 Concluding remark

In this chapter, a statistical approach for source localization on solid surfaces based on the concept of HADs has been proposed. The HADs computed from the sensor output exhibit high similarity over both the noise and signal periods. This property is utilized to compute the probability contributions of the noise and signal processes to each HAD within the transition period. The arrival of the signal at each sensor is then estimated as when the corresponding signal contribution exceeds a threshold that is common for all sensors. By deriving the optimal threshold, it is possible to perform simultaneous TOA estimation across all the sensors. The resultant TOA estimates minimize the total TDOA estimation error and therefore yield optimal localization performance. Experiment results on aluminum and glass plates show that the proposed TOA-HAD outperforms existing algorithms such as TiF-HA by significant margins.

While the proposed HAD framework allows simultaneous TOA estimation, hence reducing variance in TDOA estimates, long transition periods still impose significant uncertainty on TOA estimation. In addition, even though improve-

ment in performance over existing methods is observed, modeling of the transition using a linear function is not entirely valid. In the next chapter, noise-to-signal transition modeling will be discussed. As will be shown, with an appropriate model, better estimation of TDOA can be achieved for larger surfaces.

Chapter 6

Logistic Fitting for TOA Estimation

Existing techniques for TOA estimation assume abrupt noise-to-signal transition [43, 44]. The proposed TOA-HAD achieves significant improvement against such methods by taking into account that the transition may occur over a relatively long duration. However, in TOA-HAD, the transition is modeled as a linear function, which is not entirely valid. To better model the noise-to-signal transition, the use of a smooth parametrized function is proposed in this chapter.

In what follows, four-parameter logistic functions are employed in transition modeling. A function is uniquely determined by a set of parameters, which include a parameter specifying the lateral shift of the function. Fitting a function to the transition data is essentially equivalent to determining its parameters. In this work, the TDOA between a sensor pair is estimated as the difference in lateral shift of the two respective fitted functions. Since the proposed fitting procedure involves the signal data over a relatively long period that includes the transition, the effects of any spurious spikes within the transition period is mitigated while estimating the lateral-shift parameter. This is not the case for techniques such as TiF-HA or TOA-HAD, where the TOA is estimated via thresholding, which is sensitive to any noise within the transition period. As a result, the proposed method can achieve higher accuracy in TDOA estimation, resulting in better localization of the impact source for HCI applications.

This chapter is organized as follows: In Section 6.1 the vibration signal will be analyzed in the time-frequency domain followed by the proposed use of logistic function for transition modeling. The STFT-Logistic algorithm for TDOA estimation is proposed in Section 6.2, which involves deriving a cost function for estimating the parameters of the logistic function to optimally model the signal. The performance of the proposed algorithm is evaluated in Section 6.4 using taps generated by finger and stylus on large glass surfaces.

6.1 Time-frequency analysis and the use of a parametrized logistic function

To illustrate how a parameterized function is used to model the noise-to-signal transition, vibration signals are first analyzed. A typical time-domain signal $x(n)$ of length N received by one of the sensors due to a finger tap on a large glass surface is shown in Fig. 6.1(a). The signal is captured using surface-mounted shock sensors on a glass plate of dimension $1.2 \text{ m} \times 1.0 \text{ m} \times 5 \text{ mm}$. The arrival of the vibration at the sensor is associated with a rise in energy of the sensor output to above the background-noise level. This change in energy can be observed clearly in the time-frequency domain via computing the STFT of the signal given by

$$\begin{aligned}\underline{\mathbf{x}}(m) &= \mathcal{F}_K \mathbf{x}(m) \\ &= [\underline{x}(0, m), \dots, \underline{x}(K-1, m)]^T,\end{aligned}\tag{6.1}$$

where $\mathbf{x}(m) = [x(m), \dots, x(m+M-1)]^T$ is a length- M signal frame, $m = 0, 1, \dots, N-M+1$ is the frame index, and $(.)^T$ denotes the transposition operator. The variable K is the number of frequency bins, and $\mathcal{F}_K$ is the $K \times K$ discrete Fourier transform (DFT) matrix. The spectrogram of $x(n)$, estimated using the magnitude of the computed STFT coefficients, is shown in Fig. 6.1(b) in logarithmic scale. Notice from the figure an increase in energy occurring as

6.1 Time-frequency analysis and the use of a parametrized logistic function

Figure 6.1: Signal received by a sensor placed at position (1.1 m, 0.9 m) due to an impact at location (0.3 m, 0.3 m) on a glass surface of dimension 1.2 m $\times$ 1.0 m $\times$ 5 mm (a) in the time domain and (b) in the time-frequency domain.

the signal arrives at the sensor. In this section, the objective is to estimate the TDOAs by analyzing such energy transition at each sensor. It is, however, worth noting that, due to velocity dispersion, energy transition varies with frequency. The sensor-received signal will therefore be processed within a sufficiently narrow frequency band so that any dispersive effect within the band is negligible. The energy transitions across frequencies belonging to the same band can then be considered to occur at the same time.

Suppose a sufficiently narrow frequency band that has been selected covers frequency bins from k_b to k_e where $1 \leq k_b < k_e \leq K$. The mean spectral value across the selected frequency band at each time-frame index m is then given by

$$\chi(m) = \frac{1}{k_e - k_b + 1} \sum_{k=k_b}^{k_e} \|\underline{x}(k, m)\|. \quad (6.2)$$

Notice in Fig. 6.2 that $\chi(m)$ increases from low values (corresponding to noise) to high values (corresponding to signal). Denoting the time interval corresponding to the noise-to-signal transition edge of $\chi(m)$ as the transition period, it can be seen that the variation in energy is modest before and after the transition period. The general waveform of $\chi(m)$ shown in Fig. 6.2 can therefore be modeled by a four-parameter logistic growth function [62,63]

$$f(m) = A + \frac{B - A}{1 + e^{-\gamma(m-\kappa)}}, \quad (6.3)$$

where A , B , γ , κ are the parameters that determine the shape of $f(m)$. The parameters

$$A = \lim_{m \rightarrow -\infty} f(m) \quad \text{and} \quad B = \lim_{m \rightarrow +\infty} f(m), \quad (6.4)$$

are, respectively, the minimum and maximum values of $f(m)$. When $f(m)$ is fitted to the sequence $\chi(m)$, as shown in Fig. 6.2, these parameters correspond, respectively, to the noise and signal energy levels. The parameter κ is the translation parameter that determines the location of $f(m)$ along the time axis such that when $m = \kappa$,

$$f(\kappa) = \frac{A + B}{2}, \quad (6.5)$$

which corresponds to the midpoint of the transition edge of $f(m)$. The point $m = \kappa$ therefore always lies within the transition period. The slope of $f(m)$ at $m = \kappa$ is given by the derivative

$$\begin{aligned} f'(\kappa) &= \left. \frac{(B - A)\gamma e^{-\gamma(m-\kappa)}}{(1 + e^{-\gamma(m-\kappa)})^2} \right|_{m=\kappa} \\ &= \frac{(B - A)\gamma}{4}, \end{aligned} \quad (6.6)$$

which implies

$$\gamma = \frac{4f'(\kappa)}{B - A}. \quad (6.7)$$

The parameter γ is therefore proportional to the slope of $f(m)$ at $m = \kappa$ and reflects how steep the change in $\chi(m)$ occurs when $f(m)$ is correctly fitted. It

6.1 Time-frequency analysis and the use of a parametrized logistic function

Figure 6.2: Fitting of the logistic function $f(m)$ to $\chi(m)$.

is worth noting that an abrupt transition is a special case of gradual transitions where $\gamma \rightarrow \infty$.

Before formulating the estimation of A , B , γ and κ in the next section, it is necessary to first describe how $f(m)$ is deployed for TDOA estimation and why it is expected to outperform existing algorithms. Suppose that $f_i(m)$, with translation parameter κ_i , is the function fitted to $\chi_i(m)$ of the i th sensor. The TDOA between a sensor pair (i, j) is then obtained as the difference between κ_i and κ_j . It is worth noting that, as opposed to conventional methods such as TiF-HA and KLID-based where the TDOAs are estimated via explicit estimation of TOAs, the proposed approach utilizes the time translation of the fitting functions to estimate the TDOAs.

To understand why the proposed algorithm can achieve better localization performance, it is necessary to note that existing algorithms estimate the TOA by applying a carefully determined threshold on statistics derived from the sensor output. When the transition edge is not abrupt such as the case for relatively long propagation distances in large surfaces, localization performance is highly dependent on the choice of threshold. In addition, the signal becomes more spurious due to noise over the long transition period, making the thresholding process more erroneous. Therefore, as the slope of the noise-to-signal transition edge decreases, the error in TOA estimation increases, resulting in poor TDOA estimates. On the contrary, fitting a function to the sensor data which focuses on

the transition period mitigates such error. With good fitting, the time-translation error of the fitted function is averaged across all the points along the transition edge. The translation-parameter estimates are therefore robust to the slope of the transition edges. As a result, TDOAs computed via translation parameters are less error-prone compared to those obtained via threshold-based TOA estimates employed in existing techniques.

6.2 Four-parameter logistic model fitting

In estimating A , B , γ , and κ from $\chi(m)$ it is worth noting that the transition period, which encapsulates the arrival of the signal, should be the focus when fitting $f(m)$ to $\chi(m)$. Therefore, as opposed to [64], where a function is fitted across the entire data range, a weighting sequence $w(m)$ is introduced here so as to give higher priority to data over the transition period. The proposed algorithm involves deriving a cost function which is independent of A and B , followed by a non-linear optimization process for the estimation of γ and κ . The estimated translation parameter $\hat{\kappa}$ is subsequently used in TDOA estimation between sensor outputs.

6.2.1 Derivation of the cost function

Let $\varepsilon(m)$ be the residual/error, modeled as a normal random variable, while fitting $f(m)$ to $\chi(m)$. With reference to (6.3), it is given that

$$\varepsilon(m) = \chi(m) - [A + (B - A)h(m)], \quad (6.8)$$

where $h(m) = [1 + e^{-\gamma(m-\kappa)}]^{-1}$. With $w(m)$ properly chosen to give higher priority to data over the transition period, the weighted sum of squared fitting errors is

6.2 Four-parameter logistic model fitting

given by

$$\begin{aligned}\mathcal{J}(A, B, \gamma, \kappa) &= \sum_{m=1}^M w(m) \varepsilon^2(m) \\ &= \tilde{\boldsymbol{\chi}}^T \boldsymbol{\chi} + (B - A)^2 \tilde{\mathbf{h}}^T \mathbf{h} + A^2 S_w - 2A \mathbf{w}^T \boldsymbol{\chi} \\ &\quad - 2(B - A) \tilde{\mathbf{h}}^T \boldsymbol{\chi} + 2A(B - A) \mathbf{w}^T \mathbf{h},\end{aligned}\quad (6.9)$$

where $\boldsymbol{\chi} = [\chi(1), \dots, \chi(M)]^T$, $\mathbf{w} = [w(1), \dots, w(M)]^T$, $S_w = \sum_{m=1}^M w(m)$, $\tilde{\boldsymbol{\chi}} = [\tilde{\chi}(1), \dots, \tilde{\chi}(M)]^T$, $\tilde{\mathbf{h}} = [\tilde{h}(1), \dots, \tilde{h}(M)]^T$, $\tilde{\chi}(m) = w(m)\chi(m)$ and $\tilde{h}(m) = w(m)h(m)$. Differentiating (6.9) with respect to A and B , we obtain

$$\frac{\partial \mathcal{J}}{\partial A} = 2[AS_w + (A - B)\tilde{\mathbf{h}}^T \mathbf{h} + \tilde{\mathbf{h}}^T \boldsymbol{\chi} - \mathbf{w}^T \boldsymbol{\chi} + (B - 2A)\mathbf{w}^T \mathbf{h}] \quad (6.10)$$

and

$$\frac{\partial \mathcal{J}}{\partial B} = 2[(B - A)\tilde{\mathbf{h}}^T \mathbf{h} - \tilde{\mathbf{h}}^T \boldsymbol{\chi} + A\mathbf{w}^T \mathbf{h}]. \quad (6.11)$$

Since the weighted sum $\mathcal{J}$ is minimum when $\partial \mathcal{J} / \partial A = 0$ and $\partial \mathcal{J} / \partial B = 0$, the values of A and B can be obtained from (6.10) and (6.11), respectively, as

$$A = \frac{(\mathbf{w}^T \boldsymbol{\chi})(\tilde{\mathbf{h}}^T \mathbf{h}) - (\mathbf{w}^T \mathbf{h})(\tilde{\mathbf{h}}^T \boldsymbol{\chi})}{S_w \tilde{\mathbf{h}}^T \mathbf{h} - (\mathbf{w}^T \mathbf{h})^2} \quad (6.12)$$

and

$$B = \frac{(\tilde{\mathbf{h}}^T \boldsymbol{\chi})(S_w - \mathbf{w}^T \mathbf{h}) - (\mathbf{w}^T \boldsymbol{\chi})(\mathbf{w}^T \mathbf{h} - \tilde{\mathbf{h}}^T \mathbf{h})}{S_w \tilde{\mathbf{h}}^T \mathbf{h} - (\mathbf{w}^T \mathbf{h})^2}. \quad (6.13)$$

Substituting (6.12) and (6.13) into (6.9), the cost function is now simplified to be independent of A and B , as given by

$$\mathcal{J}(\gamma, \kappa) = \tilde{\boldsymbol{\chi}}^T \boldsymbol{\chi} - \frac{(\mathbf{w}^T \boldsymbol{\chi})^2 (\mathbf{h}^T \mathbf{h}) + S_w (\tilde{\mathbf{h}}^T \boldsymbol{\chi})^2}{S_w \tilde{\mathbf{h}}^T \mathbf{h} - (\mathbf{w}^T \mathbf{h})^2} - \frac{2(\mathbf{w}^T \mathbf{h})(\mathbf{w}^T \boldsymbol{\chi})(\tilde{\mathbf{h}}^T \boldsymbol{\chi})}{S_w \tilde{\mathbf{h}}^T \mathbf{h} - (\mathbf{w}^T \mathbf{h})^2}. \quad (6.14)$$

With $\mathcal{J}(\gamma, \kappa)$ now only dependent on γ and κ , optimization methods such as that described in [65, 66] can be utilized to estimate the pair (γ, κ) that minimizes the cost function in (6.14). To facilitate such nonlinear optimization procedure, it is important to obtain a sufficiently good initial estimate of (γ, κ) . Therefore, it will be discussed in the next section the estimation of when the transition period occurs. This will serve to provide an initial estimate of (γ, κ) . It is worth noting

Figure 6.3: Estimating the occurrence of transition period. The points corresponding to the end-points of the transition period are denoted by $\bullet$, and the point corresponding to the minimum of the total variance $\sigma_\chi^2(m)$ is denoted by $\circ$.

that with the transition period estimated as $[\hat{\tau}_b, \hat{\tau}_e]$, a weighting sequence given by

$$w(m) = \begin{cases} 1, & \hat{\tau}_b \leq m \leq \hat{\tau}_e, \\ 0, & \text{otherwise,} \end{cases} \quad (6.15)$$

can be employed to give higher priority to the signal data over the transition period when minimizing $\mathcal{J}(\gamma, \kappa)$ for the purpose of fitting.

6.2.2 Estimating the occurrence of the transition period

Denoting the transition period by $[\tau_b, \tau_e]$, the procedure to obtain its estimate $[\hat{\tau}_b, \hat{\tau}_e]$ will now be discussed. Consider a frame index m that divides $\chi = [\chi(1), \dots, \chi(M)]^T$ into left and right partitions $\chi_m^\ell = [\chi(1), \dots, \chi(m-1)]^T$ and $\chi_m^r = [\chi(m+1), \dots, \chi(M)]^T$. The relative position of the partition point m with respect to τ_b and τ_e affects the total variance of χ_m^ℓ and χ_m^r . If $m < \tau_b$, χ_m^ℓ contains only noise energy, and therefore exhibits low variance. However, χ_m^r has high variance due to the inclusion of transition period in $\chi(m+1), \dots, \chi(\tau_e)$. Similarly, if $m > \tau_e$, χ_m^ℓ exhibits high variance while χ_m^r shows low variance. On the contrary, when $m \in [\tau_b, \tau_e]$, the number of outliers is reduced for both χ_m^ℓ and χ_m^r , resulting in smaller variances for both the sequences. The total variance

6.2 Four-parameter logistic model fitting

of the two partitions, given by

$$\sigma_\chi^2(m) = \text{var}\{\mathbf{x}_m^\ell\} + \text{var}\{\mathbf{x}_m^r\}, \quad (6.16)$$

can therefore reflect the position of the partitioning point with respect to the transition period, with small $\sigma_\chi^2(m)$ corresponding to $m \in [\tau_b, \tau_e]$. As illustrated in Fig. 6.3, the point

$$\varrho = \underset{m}{\text{argmin}} \sigma_\chi^2(m) \quad (6.17)$$

is therefore guaranteed to lie within the transition period.

Note that since $\chi(m)$ values are highly concentrated around A (defined in (6.4)) for $m \leq \tau_e$, $\text{median}\{\mathbf{x}_{\tau_e}^\ell\}$ is a sufficiently good estimate of A , where $\mathbf{x}_{\tau_e}^\ell = [\chi(1), \dots, \chi(\tau_e - 1)]^T$. Exploiting the fact that the transition period is comparatively short with respect to the data length, it follows that $\text{median}\{\mathbf{x}_\varrho^\ell\} \approx \text{median}\{\mathbf{x}_{\tau_e}^\ell\}$. An initial estimate of A can therefore be obtained as $\hat{A}^0 = \text{median}\{\mathbf{x}_\varrho^\ell\}$. Similarly, $\hat{B}^0 = \text{median}\{\mathbf{x}_\varrho^r\}$. The beginning of the transition period is then defined as the last time instant that $\sigma_\chi^2(m)$ decreases below the threshold $\hat{A}^0$, i.e.,

$$\hat{\tau}_b = \max\{m \mid 1 \leq m < \varrho; \sigma_\chi^2(m-1) < \hat{A}^0 \leq \sigma_\chi^2(m)\}. \quad (6.18)$$

Similarly, the end of the transition period is estimated as the first time instant that $\sigma_\chi^2(m)$ exceeds $\hat{B}^0$, i.e.,

$$\hat{\tau}_e = \min\{m \mid \varrho < m \leq M; \sigma_\chi^2(m) \leq \hat{B}^0 < \sigma_\chi^2(m+1)\}. \quad (6.19)$$

Figure 6.3 illustrates the estimated $\hat{\tau}_b$ and $\hat{\tau}_e$ using the above process.

6.2.3 Estimation of γ and κ

With the identification of the transition period, an approximation scheme can now be adopted where the transition edge, i.e., the portion of $f(m)$ over the period $[\hat{\tau}_b, \hat{\tau}_e]$, is approximated by a line that passes through the two points $(\hat{\tau}_b, \hat{A}^0)$ and

$(\hat{\tau}_e, \hat{B}^0)$. In other words, as can be seen from Fig. 6.3,

$$f(m) \approx \alpha m + \beta, \quad \forall m \in [\hat{\tau}_b, \hat{\tau}_e], \quad (6.20)$$

where the coefficients α and β are given by

$$\alpha = \frac{\hat{B}^0 - \hat{A}^0}{\hat{\tau}_e - \hat{\tau}_b}, \quad \text{and} \quad \beta = \hat{A}^0 - (\hat{B}^0 - \hat{A}^0) \frac{\hat{\tau}_b}{\hat{\tau}_e - \hat{\tau}_b}. \quad (6.21)$$

Recall that $\kappa \in [\tau_b, \tau_e]$, from (6.5) and (6.20), an initial guess for κ is the value $\hat{\kappa}^0 \in [\hat{\tau}_b, \hat{\tau}_e]$ that satisfies $\alpha \hat{\kappa}^0 + \beta = (\hat{A}^0 + \hat{B}^0)/2$, giving an initial value

$$\hat{\kappa}^0 = \frac{\hat{\tau}_b + \hat{\tau}_e}{2}. \quad (6.22)$$

Note that (6.20) also implies $f'(m) \approx \alpha, \forall m \in [\hat{\tau}_b, \hat{\tau}_e]$, which, together with (6.7) and (6.21), provides an initial estimate for γ given by

$$\hat{\gamma}^0 = \frac{4\alpha}{\hat{B}^0 - \hat{A}^0} = \frac{4}{\hat{\tau}_e - \hat{\tau}_b}. \quad (6.23)$$

With the initial values $(\hat{\gamma}^0, \hat{\kappa}^0)$, the solution $(\hat{\gamma}, \hat{\kappa})$ that minimizes the cost function in (6.14) can be obtained via iterative optimization algorithms such as the Nelder-Mead algorithm [65, 66].

6.3 Computational complexity

The two main stages of STFT-Logistic include computations of the mean spectral sequence and fitting of the sequence to a logistic function. As discussed in Section 5.2.6, the complexity of STFT computation is $\mathcal{O}(MK \log K)$, where M is the number of windowed data frames and K is the number of frequency bins. The sequence $\chi(m)$ can subsequently be obtained from $\underline{x}(k, m)$ via (6.2) in $\mathcal{O}(M\Delta k)$, resulting the overall complexity of $\mathcal{O}(MK \log K)$ for the first stage.

In the logistic fitting stage, the main computational load is due to the iterative optimization of the cost function $\mathcal{J}(\gamma, \kappa)$. For each iteration, depending on the method being used, it is necessary to perform several computations of

6.4 Experiment Results

$\mathcal{J}(\gamma, \kappa)$ and its derivatives. Here, the Nelder-Mead simplex method is used for optimization, and during each iteration, the Nelder-Mead simplex may change in five different ways: reflection, expansion, outside contraction, inside contraction, and shrink [67]. This corresponds to as many computations of $\mathcal{J}(\gamma, \kappa)$ using (6.14), each of which is of $\mathcal{O}(M)$ complexity. The number of iterations required to minimize $\mathcal{J}(\gamma, \kappa)$, however, depends on how fast the method can converge if at all and therefore is nondeterministic. In our particular application, the Nelder-Mead algorithm often converges within 50 iterations in average, which results in a negligible complexity for the fitting stage as compared to the mean-spectral-sequence computation stage. It can therefore be concluded that, empirically, the STFT-Logistic algorithm is of complexity $\mathcal{O}(MK \log K)$.

6.4 Experiment Results

The performance of the proposed STFT-Logistic algorithm is evaluated using different sets of real data collected on two glass plates of different dimensions. The small plate is of dimension $0.6 \text{ m} \times 0.6 \text{ m} \times 5.0 \text{ mm}$ while the large plate is of dimension $1.2 \text{ m} \times 1.0 \text{ m} \times 5.0 \text{ mm}$. On the surface of each of these plates, impacts are generated by either a finger or a metal stylus, making a total of four experiment setups. Eight surface-mounted Murata PKS1-4A10 piezoelectric shock sensors are deployed to capture impact-induced vibrations where the sensor outputs are digitized using $f_s = 96 \text{ kHz}$. The sensors are mounted at the corners and the midpoints of the edges, 0.1 m away from the plate boundaries. For the small plate, impacts are generated at locations $(0.2, 0.2)$, $(0.2, 0.3)$, $(0.2, 0.4)$, $(0.3, 0.2)$, $(0.3, 0.3)$, $(0.3, 0.4)$, $(0.4, 0.2)$, $(0.4, 0.3)$, and $(0.4, 0.4)$. For the large plate, the impact locations are $(0.3, 0.3)$, $(0.3, 0.5)$, $(0.3, 0.7)$, $(0.5, 0.3)$, $(0.6, 0.5)$, $(0.6, 0.7)$, $(0.9, 0.3)$, $(0.9, 0.5)$, and $(0.9, 0.7)$ (all the dimensions are in meters). At each location, a set of five impacts are generated, and hence a total of forty-five test cases are performed for each experiment setup. In order to quantify

Figure 6.4: Performance comparison of the proposed algorithm with that of TiF-HA [44] using (a) RMSE and (b) standard deviation of the localization errors (STDE), both measured in meter.

the overall performance of each algorithm on a set of data, the root-mean-square error (RMSE) and the standard deviation (STDE) of the localization errors for all impacts in the set, measured in meter, are employed.

Since the TiF-HA outperforms the KLID-based method, we only compare the performance of our proposed algorithm with that of TiF-HA [44]. It can be seen from Fig. 6.4 that the proposed STFT-Logistic outperforms the TiF-HA significantly albeit a modest improvement in performance for the finger tap on the large plate. These results illustrate the advantage of the adoption of gradual-change model over abrupt-change model. The modest improvement in localization performance of the proposed STFT-Logistic for finger taps on large surface may be explained as follows. Unlike a metal stylus which is rigid, a finger is soft and generates multiple vibrations due to flesh, nail and bone. These vibrations are of different types and propagate with different velocities. Over

6.5 Concluding remark

the small surface, due to the short propagation distance, the difference in arrival times at the sensor of these vibrations are negligibly small. On the contrary, for a large surface, their separation in arrival times affects the energy transition of the sensor output. This energy transition now corresponds to multiple vibrations while the transition modeling was intended for the arrival of a single vibration. Such model mismatch causes the performance of STFT-Logistic to degrade in this case and only to outperform that of TiF-HA modestly.

6.5 Concluding remark

In this chapter, the STFT-Logistic algorithm has been proposed, which employs four-parameter logistic functions to model gradual energy transitions in sensor-received signals. After fitting functions to the sensor data, TDOAs are estimated as the time-translational difference between the fitted functions. This effectively avoids direct TOA estimation, which is prone to error when the transition period is relatively long, and therefore outperforms existing techniques which adopt the abrupt change model for TOA estimation.

Energy transition modeling using logistic functions results in good localization performance in most cases. However, when the variation of signal within the transition period is too complicated for modeling, such as the case of finger impacts on large surface, the proposed STFT-Logistic exhibits limited performance. In the next chapter, to avoid the complicity of the noise-to-signal transition, another approach for TOA estimation will be proposed.

Chapter 7

Harmonic-Process Modeling for TOA Estimation

The STFT-Logistic algorithm proposed in Chapter 6 is an attempt to deal with gradual noise-to-signal transition often seen in larger surfaces. However, while the use of translation parameter partially compensates the uncertainty in TDOA estimation due to relatively long transition periods, it does not directly address the problem of gradual transition. The method begins to show limitations when the transition is too complicated for modeling using logistic functions. In this chapter, to circumvent the complicity of noise-to-signal transition, a statistical method is proposed.

The proposed algorithm first extracts signal components of different frequencies and subsequently estimates the TOA of each component as the starting point of its transition period. While the use of translation parameter in [59] inherits the uncertainty of a long transition period, the proposed approach of estimating the starting point of the transition period does not suffer from this uncertainty. It however requires a sufficiently high signal-to-noise ratio (SNR) to achieve good estimation. In this chapter, procedures to improve TOA/TDOA estimation accuracy in the presence of noise will also be proposed, which relies on constructing a statistical TOA profile for each sensor-received signal from its frequency components. It will be shown later in the experiment results that the proposed method of TOA Profiling on Harmonic Processes (TOA-HPP) can achieve better localiza-

7.1 Analysis of impact-induced signals

Figure 7.1: Frequency components extracted from a typical vibration signal induced by a finger tap on glass surface. The black solid lines represent the real parts of the frequency components while the dashed lines represent their respective magnitudes. The thick gray lines represent the kurtosis sequences $\gamma_k(n)$ obtained from (7.9). The black dots represent the TOAs estimated using $\gamma_k(n)$.

tion of the impact source on large surface compared to STFT-Logistic, TOA-HAD or other existing techniques.

This chapter is organized as follows: In Section 7.1 components of different frequencies of the vibration signal will be analyzed in the time domain for TOA profiling. The use of kurtosis and the corresponding curvature in TOA estimation for each frequency component is proposed in Sections 7.2 and 7.3. TDOA estimation using the derived TOA profiles is discussed in Section 7.4. Section 7.6 evaluates the performance of the proposed algorithm against other methods on large glass surfaces via experiments with taps generated by finger and stylus.

7.1 Analysis of impact-induced signals

Due to velocity dispersion, for a sensor-received signal $x(n)$, different frequency components arrive at the sensor at different time instants. Therefore, the TOA

can be determined only for each frequency component, rather than for the whole signal. By extracting components of $x(n)$ corresponding to a selected set of frequencies f_k , $k = 1, \dots, K$, the arrival of $x(n)$ can be characterized by the set that includes the TOA of each component.

Extraction of the component $\underline{z}_k(n)$ corresponding to frequency f_k can be achieved using a narrow bandpass filter centered at f_k . However, the narrower the passband is, the more widely spread is the impulse response, which tends to smoothen the signal. Therefore, to retain the noise-to-signal transition information, it is desirable to deploy filters with time-localized impulse responses. However, localization in both time and frequency is limited according to the Heisenberg uncertainty principle [68]. Here, to achieve the best time-frequency resolution, Gabor filters are employed [69]. The component $\underline{z}_k(n)$ is then obtained as

$$\underline{z}_k(n) = x(n) * \underline{\psi}_k(n), \quad (7.1)$$

where $*$ denotes the convolution operator and the Gabor function is given as

$$\underline{\psi}_k(n) = \frac{1}{\sqrt[4]{\pi}} \sqrt{\frac{\omega_k}{\varsigma}} \exp \left[-\frac{1}{2} \left(\frac{\omega_k n}{\varsigma} \right)^2 \right] \exp(-j\omega_k n), \quad (7.2)$$

with $\omega_k = 2\pi f_k / f_s$, and f_s being the sampling frequency. The parameter ς determines the spread of $\underline{\psi}_k(n)$ in time as well as in frequency. The real part and the magnitude of each extracted complex signal $\underline{z}_k(n)$ are illustrated in Fig. 7.1 for a typical sensor-received signal due to a finger tap on a glass surface.

In this research, the vibration signals are captured by sensors whose frequency responses roll off beyond 10 kHz. In such a low frequency range, only the flexural mode of vibration is present [70], where the wave propagation velocity is proportional to the square root of the vibration frequency. This velocity can be expressed as a function of frequency as [34]

$$c(f) = \sqrt[4]{\frac{E}{12\dot{D}(1-\nu^2)}} \sqrt{2\pi f L_z}, \quad (7.3)$$

7.2 Kurtosis for TOA estimation

where the modulus of elasticity E , the flexural rigidity $\dot{D}$, and the Poisson's ratio ν are material-dependent properties, and L_z is the thickness of the surface. The time, in samples, required for the component of frequency f_k to arrive at the sensor is given by

$$\tau(f_k) = \frac{\zeta}{\sqrt{f_k}}, \quad (7.4)$$

where $\zeta = d/\sqrt[4]{\pi^2 L_z^2 E / [3\dot{D}(1 - \nu^2)]}$, d being the distance from the impact location to the sensor. Considering only signal components within a frequency band centered at f_c , i.e., $f_k \in (f_c - \Delta_f, f_c + \Delta_f)$, $k = 1, \dots, K$, the variation among the TOAs τ_k of $\underline{z}_k(n)$ is limited to $\Delta_\tau \approx \left| \frac{d\tau}{df}(f_c) \right| \Delta_f = \zeta \Delta_f / (2\sqrt{f_c^3})$. By selecting a sufficiently narrow bandwidth Δ_f , the variation Δ_τ becomes negligibly small so that all τ_k for a particular sensor can be considered equal. Such an effort to reduce the effect of dispersion is illustrated in Fig. 7.1, where f_k are selected within 3 kHz to 5 kHz. The TOAs τ_k (estimated using the method to be proposed in the following sections), represented by a dot ($\bullet$) in each subplot, can be seen relatively aligned at a specific time instant. The arrival of the signal can now be characterized by the TOA set $\boldsymbol{\tau} = \{\tau_1, \dots, \tau_K\}$. While a common TOA can be estimated as $\hat{\tau} = \sum_{k=1}^K \tau_k / K$, the spread $\sigma_{\boldsymbol{\tau}} = \left[\sum_{k=1}^K (\tau_k - \hat{\tau})^2 / K \right]^{1/2}$ represents the quality of the estimation, where large $\sigma_{\boldsymbol{\tau}}$ implies high uncertainty of the estimate. In addition, outliers can be removed from $\boldsymbol{\tau}$ so that the signal arrival is more accurately characterized by the TOA profile. In this work, instead of immediately obtaining TOA estimates from $\boldsymbol{\tau}$, the profile will be utilized in combination with those obtained from other sensor outputs to achieve better estimation of TDOAs.

7.2 Kurtosis for TOA estimation

In this section, the arrival time for each f_k -component of $x(n)$ is estimated. While the noise-to-transition in $\underline{z}_k(n)$ may occur gradually, the uncertainty imposed

Figure 7.2: The kurtosis sequences and their respective curvatures. The sequence $\gamma_k(n)$ in (a) is the kurtosis of $\underline{Z}_k(n)$, computed as in (7.9), while the kurtosis sequence in (b) is computed from $\underline{z}_k(n)$. The curvature of $\gamma_k(n)$ in (a) can be utilized to determine the starting point of the rising edge in $\gamma_k(n)$. The starting point (•) corresponds to the peak of $\xi_k(n)$, denoted by o.

by the long transition period is circumvented by estimating its starting point. The starting point is defined as the time instant when the energy of $\underline{z}_k(n)$ deviates from that of the background noise. To capture the appearance of the signal on the background noise, the kurtosis sequence for a sliding length- L frame $\mathbf{z}_k(n) = [\underline{z}_k(n), \dots, \underline{z}_k(n + L - 1)]^T$ is computed. Due to the sensitivity of the kurtosis to outliers, $\text{kurt}\{\mathbf{z}_k(n)\}$ exhibits a large value when the energy of $\underline{z}_k(n)$ exceeds that of the noise. This results in a sharp rising edge in the kurtosis sequence whenever an energy deviation occurs. However, due to the high variation in energy of $\underline{z}_k(n)$ after its arrival, direct computation of $\text{kurt}\{\mathbf{z}_k(n)\}$ results in a non-smooth sequence with multiple peaks, as illustrated in Fig. 7.2(b). This makes it challenging to determine the exact instant when $\underline{z}_k(n)$ arrives. To address this problem, $\underline{z}_k(n)$ is modeled as a realization from a harmonic random process with frequency f_k . Such a random process is defined as

$$\underline{Z}_k(n) = a_k(n)e^{j2\pi f_k n + \varphi_k(n)}, \quad (7.5)$$

7.2 Kurtosis for TOA estimation

where $a_k(n) = |\underline{z}_k(n)|$ and $\varphi_k(n)$ is a random variable uniformly distributed on $[-\pi, \pi]$. Here, instead of $\text{kurt}\{\mathbf{z}_k(n)\}$, the kurtosis for each $\mathbf{Z}_k(n) = [\underline{Z}_k(n), \dots, \underline{Z}_k(n+L-1)]^T$ is computed as [71]

$$\gamma_k(n) = \frac{\kappa_4\{\mathbf{Z}_k(n)\}}{\kappa_2^2\{\mathbf{Z}_k(n)\}}, \quad (7.6)$$

where $\kappa_2\{\mathbf{Z}_k(n)\}$ and $\kappa_4\{\mathbf{Z}_k(n)\}$ are the second and the fourth cumulants of $\mathbf{Z}_k(n)$, respectively.

It can easily be seen from (7.5) that $E\{\underline{Z}_k(n)\} = 0$, which yields $E\{\mathbf{Z}_k(n)\} = \sum_{i=n}^{n+L-1} E\{\underline{Z}_k(i)\} = 0$. The second and fourth cumulants are then respectively reduced to

$$\begin{aligned} \kappa_2\{\mathbf{Z}_k(n)\} &= E\{\mathbf{Z}_k(n)\mathbf{Z}_k^*(n)\} \\ &= \frac{1}{L} \sum_{i=n}^{n+L-1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \underline{z}_k(n) \underline{z}_k^*(n) d\varphi_k(n) \\ &= \frac{1}{L} \sum_{i=n}^{n+L-1} |\underline{z}_k(n)|^2, \end{aligned} \quad (7.7)$$

and

$$\begin{aligned} \kappa_4\{\mathbf{Z}_k(n)\} &= E\{\mathbf{Z}_k(n)\mathbf{Z}_k^*(n)\mathbf{Z}_k(n)\mathbf{Z}_k^*(n)\} \\ &= \frac{1}{L} \sum_{i=n}^{n+L-1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \underline{z}_k(n) \underline{z}_k^*(n) \underline{z}_k(n) \underline{z}_k^*(n) d\varphi_k(n) \\ &= \frac{1}{L} \sum_{i=n}^{n+L-1} |\underline{z}_k(n)|^4, \end{aligned} \quad (7.8)$$

where $\underline{Z}_k^*(n)$ and $\underline{z}_k^*(n)$ are the complex conjugates of $\underline{Z}_k(n)$ and $\underline{z}_k(n)$, respectively. As a result, the kurtosis sequence is obtained as

$$\gamma_k(n) = \frac{\frac{1}{L} \sum_{i=n}^{n+L-1} |\underline{z}_k(i)|^4}{\left(\frac{1}{L} \sum_{i=n}^{n+L-1} |\underline{z}_k(i)|^2\right)^2}. \quad (7.9)$$

It is worth noting from (7.9) that $\gamma_k(n)$ is an energy-based measure. Therefore, similar to existing energy-based TOA techniques such as TiF-HA [44], TOA-HAD,

Figure 7.3: Frequency components extracted from a typical vibration signal induced by a finger tap on glass surface where white Gaussian noise is added to the signal at SNR of 10 dB. The black solid lines represent the real parts of the frequency components while the dashed lines represent their respective magnitudes. The thick gray lines represent the kurtosis sequences $\gamma_k(n)$ obtained from (7.9). The black dots represent the TOAs estimated using $\gamma_k(n)$.

and STFT-Logistic, utilization of $\gamma_k(n)$ also requires high signal-to-noise ratio (SNR). The effect of SNR on TOA estimation is shown in Fig. 7.3, where $x(n)$ is contaminated with white Gaussian noise. While the overall SNR is 10 dB, when considering each component individually, the SNR decreases with frequency. This is due to the lower power of high-frequency components in comparison with low-frequency ones. As can be seen in Figs. 7.1 and 7.3, an empirically determined SNR of at least 10 dB is necessary for $\gamma_k(n)$ to be sufficiently well conditioned so that an accurate TOA estimate can be obtained. As a result, TOA estimation will only be performed on $\underline{z}_k(n)$ that satisfies the SNR requirement.

To estimate the SNR for each $\underline{z}_k(n)$, the silent period is first estimated from the signal $x(n)$. Applying kurtosis on the sliding frame $\mathbf{x}(n) = [x(n), \dots, x(n +$

7.3 Arrival time estimation using curvature

$L - 1)]^T$, a kurtosis sequence is obtained as

$$\begin{aligned}\gamma(n) &= \frac{\kappa_4 \{\mathbf{x}(n)\}}{\kappa_2^2 \{\mathbf{x}(n)\}} \\ &= \frac{\frac{1}{L} \sum_{i=n}^{n+L-1} [x(i) - \bar{x}(n)]^4}{\left(\frac{1}{L} \sum_{i=n}^{n+L-1} [x(n+i) - \bar{x}(n)]^2 \right)^2},\end{aligned}\quad (7.10)$$

where $\bar{x}(n) = \sum_{i=n}^{n+L-1} x(i)/L$. An initial estimate of the onset time, which separates the silent period and the signal period, can then be obtained as

$$\tau_{\text{on}} = \underset{n}{\operatorname{argmax}} \gamma(n). \quad (7.11)$$

The silent period is then estimated as $[1, \tau_{\text{on}} - \Delta]$, where Δ is a buffer length chosen empirically to guarantee that the signal content is absent during the silent period. Similarly, the signal period is determined as $[\tau_{\text{on}} + \Delta', N]$ for some empirical value Δ' . The SNR of each $z_k(n)$ is subsequently estimated as

$$\text{SNR}_k = \frac{\sum_{i=\tau_{\text{on}}+\Delta'}^N z_k(n) z_k^*(n) / (N - \tau_{\text{on}} - \Delta' + 1)}{\sum_{i=1}^{\tau_{\text{on}}-\Delta} z_k(n) z_k^*(n) / (\tau_{\text{on}} - \Delta)}. \quad (7.12)$$

7.3 Arrival time estimation using curvature

While $\gamma_k(n)$ exhibits a sharp rising edge on noise-to-signal transition, there is no knowledge of the correct threshold value that $\gamma_k(n)$ would exceed at the transition point. Thresholding is therefore not an appropriate approach for determining the starting point of the rising edge. Here, this point is determined by utilizing the curvature of $\gamma_k(n)$. The curvature of a function measures how fast a function changes its direction, and a sharp rising edge corresponds to such a change being sudden. The curvature of $\gamma_k(n)$ at point n is given by

$$\xi_k(n) = \frac{\Delta^2 \gamma_k(n)}{[1 + (\Delta \gamma_k(n))^2]^{3/2}} \quad (7.13)$$

where $\Delta\gamma_k(n) = \gamma_k(n+1) - \gamma_k(n)$ and $\Delta^2\gamma_k(n) = \Delta\gamma_k(n+1) - \Delta\gamma_k(n) = \gamma_k(n+2) - 2\gamma_k(n+1) + \gamma_k(n)$ are, respectively, the first and second discrete difference of $\gamma_k(n)$. The arrival time of $z_k(n)$ is estimated as

$$\tau_k = \underset{n}{\operatorname{argmax}} \xi_k(n). \quad (7.14)$$

It is worth noting that for a noisy function, its curvature fluctuates significantly since $\xi_k(n)$ is sensitive to the sign of $\Delta^2\gamma_k(n)$, and therefore is not a good measure for edge detection. However, the envelope $|z_k(n)|$ is a relatively low-band signal with its maximum frequency negligibly small compared to f_k . The curvature sequency $\xi_k(n)$ can therefore be utilized as shown in Fig. 7.2(a). This is also the advantage of $\gamma_k(n)$ over other non-smooth measures such as the kurtosis directly computed from $z_k(n)$, denoted by $\gamma'_k(n)$, whose curvature $\xi'_k(n)$ fluctuates significantly as illustrated in Fig. 7.2(b).

7.4 TDOA estimation

Consider the signals $x_1(n), \dots, x_R(n)$ captured by R number of surface-mounted sensors. For each $x_i(n)$, $i = 1, \dots, R$, denote by τ_i^k the TOA estimated for its f_k -component $z_i^k(n)$ using (7.14). The TDOA between $z_i^k(n)$ and $z_j^k(n)$ for each sensor pair (i, j) is given by

$$\tau_{i,j}^k = \tau_i^k - \tau_j^k. \quad (7.15)$$

Since the frequencies f_k are chosen within a narrow band where velocity dispersion is negligible, $\tau_{i,j}^k$ are all equal to a common value $\tau_{i,j}$. In practice, $\tau_{i,j}^k$ will be randomly distributed in a neighborhood of $\tau_{i,j}$. The TDOA between $\{z_i^k\}$ and $\{z_j^k\}$ can be estimated as

$$\hat{\tau}_{i,j} = \frac{1}{K} \sum_{k=1}^K \tau_{i,j}^k. \quad (7.16)$$

Note that the quality of the estimate $\hat{\tau}_{i,j}$ in (7.16) depends on the consistency of $\tau_{i,j} = \{\tau_{i,j}^1, \dots, \tau_{i,j}^K\}$. To achieve better accuracy in source localization, only

7.5 Computational complexity

$\tau_{i,j}$ that exhibit sufficiently low spreads are utilized. This results in a binary weighting matrix $[w_{i,j}]$ given by

$$w_{i,j} = \begin{cases} 1, & \text{if } \sigma_{\tau_{i,j}} < \tau, \\ 0, & \text{otherwise,} \end{cases} \quad (7.17)$$

where $\sigma_{\tau_{i,j}} = \left[\sum_{k=1}^K (\tau_{i,j}^k - \hat{\tau}_{i,j})^2 / K \right]^{1/2}$, and τ is an empirical threshold, with $w_{i,j} = 0$ effectively implies $\hat{\tau}_{i,j}$ having no effect on the impact location estimate.

Denoting by f_c the center frequency of the band where f_k are selected from, since dispersion is negligible within the band, all f_k -components propagate at the same speed $c(f_c)$. Estimation of f_c can be achieved during calibration as described in [44]. The source location (u_s, v_s) can then be estimated by minimizing the error function

$$(\hat{u}_s, \hat{v}_s) = \underset{(u,v)}{\operatorname{argmin}} \sum_{i,j} \left(\frac{d_i(u,v) - d_j(u,v)}{c(f_c)} - \frac{\hat{\tau}_{i,j}}{f_s} \right)^2 w_{i,j}, \quad (7.18)$$

where $d_i(u,v) = \sqrt{(u - u_i)^2 + (v - v_i)^2}$ is the distance from an arbitrary location (u,v) to the i th sensor positioned at (u_i, v_i) . Such minimization can be achieved using existing optimization algorithms such as Levenberg-Marquardt [72] or Nelder-Mead [65].

7.5 Computational complexity

To compute the complexity of TOA-HPP, recall that K components are first extracted from the signal, and TOA estimation is subsequently performed on each $\underline{z}_k(n)$. The complexity for filtering when done in the frequency domain is $\mathcal{O}(N \log N)$, where N is the signal length. Component extraction is, therefore, of $\mathcal{O}(KN \log N)$ complexity. On the other hand, for each $\underline{z}_k(n)$, the corresponding sequences $\gamma_k(n)$ and $\xi_k(n)$ can be computed in $\mathcal{O}(N)$. This results in $\mathcal{O}(KN)$ complexity for the second stage and $\mathcal{O}(KN \log N)$ overall for the proposed TOA-HPP algorithm.

Figure 7.4: Performance comparison of the proposed algorithm with that of TiF-HA [44] and STFT-Logistic using (a) RMSE and (b) standard deviation of the localization errors (STDE), both measured in meter.

7.6 Experiment Results

The performance of the TOA-HPP algorithm is evaluated using different sets of real data collected on a glass plate of dimension $1.2 \text{ m} \times 1.0 \text{ m} \times 5.0 \text{ mm}$. On the plate surface, impacts are generated by either a finger or a metal stylus at the locations $(0.3, 0.3)$, $(0.3, 0.5)$, $(0.3, 0.7)$, $(0.5, 0.3)$, $(0.6, 0.5)$, $(0.6, 0.7)$, $(0.9, 0.3)$, $(0.9, 0.5)$, and $(0.9, 0.7)$ (all the dimensions are in meters). At each location, a set of five impacts are generated, and hence a total of forty-five test cases are performed for each experiment setup. The induced vibrations are captured by eight surface-mounted Murata PKS1-4A10 piezoelectric shock sensors mounted at the corners and the midpoints of the edges, 0.1 m away from the plate boundaries. The sensor outputs are subsequently digitized at a sampling frequency of $f_s = 96 \text{ kHz}$. In order to quantify the overall performance of each algorithm on a set of data, the root-mean-square error (RMSE) and the standard deviation (STDE) of the localization errors for all impacts in the set are employed.

Here the performance of TOA-HPP is compared with that of TiF-HA [44] and STFT-Logistic, which have been shown to outperform other methods such as the KLID-based [43] or scalogram-based [41] by significant margins. It can be seen

7.7 Concluding Remark

from Fig. 7.4 that the proposed algorithm and STFT-Logistic outperform the TiF-HA in both cases. This is because TiF-HA assumes abrupt noise-to-signal transitions which is unlikely the case for vibration signals captured on large plates. In addition, STFT-Logistic only modestly outperforms TiF-HA for taps generated by fingers, which may include vibrations due to flesh, nail and bone. For such taps, the transition becomes complex and is not correctly described by the logistic function. This is reflected in the modest margin of STFT-Logistic against TiF-HA. The algorithm proposed here, however, captures the starting instant of the transition period, and therefore is not affected by the complex transition. As a result, the proposed TOA-HPP outperforms both STFT-Logistic and TiF-HA by decent margins.

7.7 Concluding Remark

In the proposed TOA-HPP, frequency components of each sensor-received signal are first modeled as harmonic random processes. The theoretical kurtosis is subsequently derived for each process, which exhibits a sharp rising edge when the signal energy deviates from that of the background noise. Due to the smoothness of the derived kurtosis sequence, its curvature can be utilized to determine the starting point of the rising edge as the arrival time of the signal. While the noise-to-signal transition may be gradual, the proposed algorithm captures the starting point of the transition period and therefore avoids complications in transition modeling. Experiment results show that the proposed TOA-HPP outperforms existing methods, especially when the transition is gradual and complicated as in the case for impacts generated by fingers.

Chapter 8

Conclusions and Future Research Direction

In this thesis, algorithms for impact localization on rigid surfaces have been discussed. By providing methods for accurately estimating the location of user interaction, daily objects can be converted to touch interfaces. The two main approaches discussed in this thesis include LTM and TDOA-based localization, and each approach faces different challenges. For LTM algorithms, one of the challenges lies in the search for location-dependent features which can be extracted as the signature of a signal for use in location matching. On the other hand, for TDOA-based localization, accurate TDOA estimation is essential. This is, however, a challenging task since the vibration signal is distorted as it reaches the sensor, due to dispersion and multipath.

8.1 Conclusions

In the first part of the thesis, methods belonging to the LTM family are discussed. The CC-LTM and Z-LTM methods, which utilize the time-reversal theory and the classical plate theory respectively are reviewed in Chapter 2. While the CC-LTM method deploys the signal itself as its signature in location matching, Z-LTM extracts a vector of variances as a location-dependent feature of the signal. With the scope of applications being for touch interfaces made of solid plates, utilization of the classical plate theory for LTM has shown to be advantageous, with

8.1 Conclusions

Z-LTM outperforming CC-LTM. However, extraction of feature also implies an increase in sensitivity to background noise, since any distortion to the extracted feature may result in poor location matching. The CC-LTM method therefore outperforms Z-LTM in the presence of noise. In Chapter 3, the BLC-LTM algorithm is proposed, which combines the advantages of both approaches. From analysis of the classical plate theory, it has been shown that band-limited components (BLCs) of a vibration signal can be employed as its location-dependent feature. A cross-correlation based measure has also been proposed that makes use of the presence of dispersion in the vibration signal for distinguishing signals generated from different locations. By utilizing multiple BLCs, the proposed measure is robust to band-limited noise. As a result, BLC-LTM outperforms existing algorithms by significant margins in such situations.

The second part of this thesis is devoted to TDOA-based localization. Chapter 4 reviews various techniques for TDOA estimation, direct and indirect. In the approach for direct TDOA estimation, the TDOA is computed by correlating the sensor outputs followed by detecting the maximum peak of the correlation function. This approach, however, is sensitive to waveform distortion due to multipath. The indirect approach for TDOA estimation focuses only on the signal portion that contains its arrival. While circumventing the multipath problem, methods in this category need to estimate the TOA of each sensor output, which is a challenging task, especially when the noise-to-signal transition is gradual. The TOA-HAD method for TOA estimation proposed in Chapter 5 attempts to monitor the transition using a statistical framework where the time-domain signal is converted to HADs which vary in time. With the noise-to-signal transition reflected via the increase of signal contribution as well as decrease of noise contribution in HADs, a common threshold can be derived for simultaneous determination of TOAs across all the sensors. This helps in reducing variance in the resultant TDOA estimates, hence improving localization performance. The

TOA-HAD, however, employs a linear function to model the increase of signal contribution during the transition period, which is not entirely valid. This limits its accuracy in TOA estimation for signals received at sensors far away from the impact source.

In Chapter 6, modeling of the noise-to-signal transition is discussed. Here, the use of a logistic function is proposed to model the energy variation throughout the transition period. By fitting a logistic function to the sensor data, its translation parameter, which determines the location of the function along the time axis, is estimated. Shown to be robust to noise within the transition period, the estimated translation parameters are subsequently used for TDOA estimation in place of TOA estimates. This results in decent localization performance on large surfaces. Transition modeling, however, sees limitation when the signal content within the transition period is highly complicated. This is the case for impacts generated by fingers where multiple vibrations are generated and propagated to the sensors. To circumvent such complications, the TOA-HPP algorithm proposed in Chapter 7 focuses on the early portion of the transition period. By modeling frequency components of the signal as harmonic random processes the TOA of each component is estimated as the beginning point of the corresponding transition period. Utilizing the property of kurtosis being sensitive to outliers, the starting point of the transition period is determined as when the sensor output's energy deviates from that of the background noise, corresponding to the maximum peak in the kurtosis sequence. The points determined for all the components constitute a TOA profile of the signal, which, in turn, is combined with those profiles obtained for other sensors in TDOA estimation. As shown in the experiment results, the performance of TOA-HPP outperforms that of STFT-Logistic by a decent amount for the case of finger taps on large surfaces.

While LTM algorithms such as BLC-LTM exhibit certain robustness to noise, due to the intensive calibration, the LTM approach is more suitable for small

8.2 Future Research Directions

surface. For localization on large surfaces, the use of TDOA-based algorithms proves to be a more appropriate choice. Among the proposed algorithms in this category, TOA-HAD exhibits the most consistent performance across different combinations of plate materials and tapping styles, while STFT-Logistic is designed for large surfaces and taps that exhibit simple noise-to-signal transitions. The TOA-HPP algorithm, as opposed to the other two, is specifically designed for complex transitions such as the case of a finger tap on a large glass plate. It is worth noting that, since TOA estimation requires high SNR in general, the performance will degrade if the source-sensor distance becomes too large. A way to maintain the quality of localization is to increase the number of sensors while leaving out sensor outputs with low SNRs.

One of the limitations of the algorithms proposed in this thesis that is worth mentioning is that they do not support tracking of sliding movements. This is because the sensors used in this work respond predominantly to vibrations along the direction normal to the plate surface, which, in turn, are generated due to the movement of the finger/stylus in the same direction when an impact is exerted on the surface. On the contrary, when the finger/stylus slides on the surface, vibrations in the direction normal to the surface are insufficiently significant to be captured by the sensor. Alternatives such as directional sensors, which are capable of capturing vibrations parallel to the surface, can be used. Application of the proposed algorithms on the output signals of such sensors, however, requires more investigation.

8.2 Future Research Directions

There are two main directions for this research to go forward from this point.

One is to extend the application to media of larger dimensions while maintaining high level of localization accuracy. The challenges include the reduction in SNR and increase in dispersion due to the longer range of wave propagation. A

possible approach is first to develop pre-processing techniques that can compensate for the dispersion present in received signals, subsequently followed by signal enhancement, before estimating the TOA/TDOA information for source localization. This requires choosing an appropriate time-frequency representation (TFR) of the signal that are covariant with its group delay shift [73] in order to reveal the characteristics of its dispersion curve. Once the dispersive characteristic is available, time or frequency warping [74] can be employed to remove dispersion for subsequent localization using the proposed methods or existing ones for non-dispersive environments. It is worth noting that in a finite dispersive media such as those in touch interface applications, the phenomenon of multipath is also present. This phenomenon is partially accounted for distortions of the signal, adding difficulty to interpreting its TFR. Methods to circumvent this problem, therefore, demand further investigations.

The other direction is to design and develop algorithm suitable for localizing simultaneous multiple impacts. The overlapping of the elastic waves from more than one sources together with the interference due to their reflections at the edges raise highly challenging problems in separating and localizing the sources. There exist several algorithms [75] for multiple impact localization in plates. These algorithms rely on the separation in arrival times of multiple waves at each sensor to recognize each source. Such separation process, however, requires material dependent parameters, which limits the portability of the algorithms. More investigation should be made to develop algorithms that can handle more narrow time separation between impacts as well as depend less on the material properties.

Appendix A

Hermitian angle distribution estimation

Different choices of $f_{\underline{w}}(w)$ and their effects on the resultant estimates $\widehat{F}_{\Theta}(\theta; m)$ and $\widehat{f}_{\Theta}(\theta; m)$ are to be compared in this section. If $f_{\underline{w}}(w)$ is chosen to be uniform over a square region centered at the origin of the complex plane, we have $f_{\underline{w}}(w) = 1/(4\vartheta_0^2) \forall w$ s.t. $\Re\{w\} \leq \vartheta_0$ and $\Im\{w\} \leq \vartheta_0$, as illustrated in Fig. A.1(a). Here, $\Re\{w\}$ and $\Im\{w\}$ are, respectively, the real and imaginary parts of the complex variable w . The magnitude distribution is then given by

$$f_{\|\underline{w}\|}(u) = \begin{cases} \pi u/(2\vartheta_0^2), & \text{if } 0 \leq u \leq \vartheta_0 \\ (\pi/2 - 2\phi_u) u/\vartheta_0^2, & \text{if } \vartheta_0 < u \leq \sqrt{2}\vartheta_0 \\ 0, & \text{otherwise} \end{cases}, \quad (\text{A.1})$$

where $\phi_u = \arccos(\vartheta_0/u)$. An approximation of the HAD is obtained by substituting (A.1) into (5.3) and (5.4), which results in

$$\begin{aligned} \widehat{F}_{\Theta}(\theta; m) = & \frac{\pi \tan^2 \theta}{4\vartheta_0^2 \Delta k} \sum_{k \in D_{\theta} \cup D'_{\theta}} \|\underline{x}(k, m)\|^2 \\ & - \frac{1}{\Delta k} \sum_{k \in D'_{\theta}} \left[\arccos \left(\frac{\vartheta_0 \cot \theta}{\|\underline{x}(k, m)\|} \right) \frac{\|\underline{x}(k, m)\|^2}{\vartheta_0^2 \cot^2 \theta} \right. \\ & \left. - \frac{\sqrt{\|\underline{x}(k, m)\|^2 \tan^2 \theta - \vartheta_0^2}}{\vartheta_0} \right] + \frac{|D'_{\theta}|}{\Delta k} \end{aligned} \quad (\text{A.2})$$

Figure A.1: Distribution of complex uniform reference noise over the circle of radius ϑ centered at the origin on the complex plane.

and

$$\begin{aligned} \hat{f}_{\Theta}(\theta; m) = & \frac{2 \tan \theta \sec^2 \theta}{\vartheta_0^2 \Delta k} \left[\sum_{k \in D_{\theta} \cup D'_{\theta}} \frac{\pi}{4} \|\underline{x}(k, m)\|^2 \right. \\ & \left. - \sum_{k \in D'_{\theta}} \arccos \left(\frac{\vartheta_0 \cot \theta}{\|\underline{x}(k, m)\|} \right) \|\underline{x}(k, m)\|^2 \right], \end{aligned} \quad (\text{A.3})$$

where $D_{\theta} = \{k \mid k \in [k_b, k_e]; 0 \leq \|\underline{x}(k, m)\| \leq \vartheta_0 \cot \theta\}$, $D'_{\theta} = \{k \mid k \in [k_b, k_e]; \vartheta_0 \cot \theta < \|\underline{x}(k, m)\| \leq \sqrt{2}\vartheta_0 \cot \theta\}$, $D''_{\theta} = \{k \mid k \in [k_b, k_e]; \|\underline{x}(k, m)\| > \sqrt{2}\vartheta_0 \cot \theta\}$, and $|D''_{\theta}|$ is the cardinality of D''_{θ} . The expressions of $\hat{F}_{\Theta}(\theta; m)$ and $\hat{f}_{\Theta}(\theta; m)$ are computationally expensive due to the complicated form of $f_{\|\underline{W}\|}(u)$ for $\vartheta_0 < u \leq \sqrt{2}\vartheta_0$.

Consider limiting $f_{\underline{W}}(w)$ over a circular support. Let $\underline{W}$ be uniformly distributed over a circle of radius ϑ_0 centered at the origin of the complex plane, i.e., $f_{\underline{W}}(w) = 1/(\pi\vartheta_0^2) \forall w$ s.t. $\|w\| \leq \vartheta_0$, as illustrated in Fig. A.1(b). The magnitude distribution is then given by

$$f_{\|\underline{W}\|}(u) = \begin{cases} 2u/\vartheta_0^2, & \text{if } 0 \leq u \leq \vartheta_0 \\ 0, & \text{otherwise} \end{cases}. \quad (\text{A.4})$$

A. Hermitian angle distribution estimation

The resultant HAD estimates are then obtained as

$$\hat{F}_{\Theta}(\theta; m) = \frac{\tan^2 \theta}{\vartheta_0^2 \Delta k} \sum_{k \in D_{\theta}} \|\underline{x}(k, m)\|^2 + 1 - \frac{|D_{\theta}|}{\Delta k} \quad (\text{A.5})$$

and

$$\hat{f}_{\Theta}(\theta; m) = \frac{2 \tan \theta \sec^2 \theta}{\vartheta_0^2 \Delta k} \sum_{k \in D_{\theta}} \|\underline{x}(k, m)\|^2. \quad (\text{A.6})$$

Comparing (A.5) and (A.6) with (A.2) and (A.3), respectively, the expression $\hat{F}_{\Theta}(\theta; m)$ and $\hat{f}_{\Theta}(\theta; m)$ can be simplified by limiting $f_{\underline{w}}(w)$ over a circular support.

Author's Publication

- [C1] Nguyen Quang Hanh, V. G. Reju, and A. W. H. Khong, "Noise robust source localization on solid surfaces," in *Proc. 10th Int. Conf. Information, Communications and Signal Processing (ICICS)*, 2015, pp. 1–5 (Best Paper Award).
- [C2] Nguyen Quang Hanh, V. G. Reju, and A. W. H. Khong, "Source localization on solids utilizing logistic modeling of energy transition in vibration signals," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2016.
- [C3] Nguyen Quang Hanh, V. G. Reju, and A. W. H. Khong, "On TOA estimation of vibration signals for localizing impacts on solid surfaces," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2017.
- [J1] Nguyen Quang Hanh, V. G. Reju, and A. W. H. Khong, "On the use of rigid surfaces for human-computer interface applications," submitted to *IEEE Trans. Multimedia*.

References

- [1] V. P. Minotto, C. R. Jung, and B. Lee, “Multimodal multi-channel on-line speaker diarization using sensor fusion through SVM,” *IEEE Trans. Multimedia*, vol. 17, no. 10, pp. 1694–1705, 2015.
- [2] G. Zen, L. Porzi, E. Sangineto, E. Ricci, and N. Sebe, “Learning personalized models for facial expression analysis and gesture recognition,” *IEEE Trans. Multimed.*, vol. 18, no. 4, pp. 775–788, 2016.
- [3] Z. P. Bian, J. Hou, L. P. Chau, and N. Magnenat-Thalmann, “Facial position and expression-based human-computer interface for persons with tetraplegia,” *IEEE J. Biomed. Health. Inf.*, vol. 20, no. 3, pp. 915–924, 2016.
- [4] C. Wang, Z. Liu, and S. C. Chan, “Superpixel-based hand gesture recognition with kinect depth camera,” *IEEE Trans. Multimed.*, vol. 17, no. 1, pp. 29–39, 2015.
- [5] G. Wu and W. Kang, “Robust fingertip detection in a complex environment,” *IEEE Trans. Multimed.*, vol. 18, no. 6, pp. 978–987, 2016.
- [6] G. Walker, “Touch sensing,” in *Interactive displays: Natural human-interface technologies*. John Wiley & Sons, Ltd, 2014, pp. 27–105.
- [7] F. Miedl and T. Tille, “3-D surface-integrated touch-sensor system for automotive HMI applications,” *IEEE/ASME Trans. Mechatron.*, vol. 21, no. 2, pp. 787–794, 2016.
- [8] J. Paradiso and N. Checka, “Passive acoustic sensing for tracking knocks atop large interactive displays,” in *Proc. IEEE Sensors*, 2002, pp. 521–527.
- [9] B. Daku, J. Salt, and L. Sha, “An algorithm for locating microseismic events,” in *Proc. IEEE Canad. Conf. Elect. Comput. Eng.*, vol. 4, May 2004, pp. 2311–2314.

-
- [10] I. Zorych and Z.-H. Michalopoulou, “Particle filtering for dispersion curve tracking in ocean acoustics,” *J. Acoust. Soc. Amer.*, vol. 124, no. 2, pp. EL45–EL50, 2008.
 - [11] Y. T. Chan, T. K. C. Lo, and H. C. So, “Passive range-difference estimation in a dispersive medium,” *IEEE Trans. Signal Process.*, vol. 58, pp. 1433–1439, 2010.
 - [12] F. Foroozan and A. Asif, “Time reversal based active array source localization,” *IEEE Trans. Signal Process.*, vol. 59, pp. 2655–2668, Jun. 2011.
 - [13] W. Rolshofen, D. T. Pham, M. Yang, Z. B. Wang, Z. Ji, and M. Al-Kutubi, “New approaches in computer-human interaction with tangible acoustic interfaces,” in *Proc. Int. Virtual Conf. IPROMs*, 2005.
 - [14] D. T. Pham, M. Al-Kutubi, Z. Ji, M. Yang, Z. B. Wang, and S. Catheline, “Tangible acoustic interface approaches,” in *Proc. Int. Virtual Conf. IPROMs*, 2005.
 - [15] D. T. Pham, Z. Ji, M. Yang, Z. Wang, and M. Al-Kutubi, “A novel human-computer interface based on passive acoustic localisation,” in *Proc. 12th Int. Conf. Human-Computer Interaction (HCI)*, 2007, pp. 901–909.
 - [16] R. K. Ing, N. Quieffin, S. Catheline, and M. Fink, “Tangible interactive interface using acoustic time reversal process,” *J. Acoust. Soc. Amer.*, vol. 117, no. 4, pp. 2560–2560, 2005.
 - [17] K. Poletkin, X. Yap, and A. W. H. Khong, “A touch interface exploiting the use of vibration theories and infinite impulse response filter modeling based localization algorithm,” in *Proc. IEEE Int. Conf. Multimedia and Expo (ICME)*, 2010, pp. 286–291.
 - [18] X. Yap, A. W. H. Khong, and W. S. Gan, “Localization of acoustic source on solids: A linear predictive coding based algorithm for location template matching,” in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, 2010, pp. 2490–2493.
 - [19] K. R. Arun, X. Yap, and A. W. H. Khong, “A touch interface exploiting time-frequency classification using Zak transform for source localization on solids,” *IEEE Trans. Multimed.*, vol. 13, pp. 487–497, 2011.

REFERENCES

- [20] C. Knapp and G. C. Carter, “The generalized correlation method for estimation of time delay,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 24, no. 4, pp. 320–327, 1976.
- [21] S. M. Ziola and M. R. Gorman, “Source location in thin plates using cross-correlation,” *J. Acoust. Soc. Amer.*, vol. 90, pp. 2551–2556, 1991.
- [22] T. Kosel, I. Grabec, and F. Kosel, “Intelligent location of simultaneously active acoustic emission sources: Part I,” *Aircr. Eng. Aerosp. Technol.*, vol. 75, pp. 11–17, 2003.
- [23] Q. Wang and F. Chu, “Experimental determination of the rubbing location by means of acoustic emission and wavelet transform,” *J. Sound Vibration*, vol. 248, no. 1, pp. 91–103, 2001.
- [24] M. Yang, D. Pham, M. Al-Kutubi, Z. Ji, and Z. Wang, “A new computer interface based on in-solid acoustic source localization,” in *Proc. IEEE Int. Conf. Industrial Informatics*, 2009, pp. 131–135.
- [25] M. Fink, “Time reversal of ultrasonic fields - Part I: Basic principles,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 39, pp. 555–566, Sep. 1992.
- [26] R. K. Ing, N. Quieffin, S. Catheline, and M. Fink, “In solid localization of finger impacts using acoustic time-reversal process,” *Appl. Phys. Lett.*, vol. 87, pp. 1–3, 2005.
- [27] G. Ribay, D. Clorennec, S. Catheline, M. Fink, R. Kirkling, and N. Quieffin, “Tactile time reversal interactivity: Experiment and modelization,” in *Proc. IEEE Int. Ultrasonics Symp.*, vol. 4, 2005, pp. 2104–2107.
- [28] G. Ribay, S. Catheline, D. Clorennec, R. K. Ing, N. Quieffin, and M. Fink, “Acoustic impact localization in plates: Properties and stability to temperature variation,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 54, pp. 378–85, Feb. 2007.
- [29] M. R. Bai and Y. K. Tsai, “Impact localization combined with haptic feedback for touch panel applications based on the time-reversal approach,” *J. Acoust. Soc. Amer.*, vol. 129, pp. 1297–305, Mar. 2011.
- [30] K. F. Graff, *Wave Motion in Elastic Solids*. Dover Publications, 1991.
- [31] E. Ventsel and T. Krauthammer, *Thin Plates and Shells: Theory, Analysis, and Applications*. CRC Press, 2001.

-
- [32] R. Szilard, *Theories and Applications of Plate Analysis: Classical, Numerical and Engineering Methods*. Wiley, 2004.
- [33] R. D. Mindlin, “The Influence of rotatory inertia and shear on the flexural motions of isotropic elastic plates,” *J. Appl. Mech. Trans. ASME*, vol. 18, pp. 31–38, 1951.
- [34] I. A. Viktorov, *Rayleigh and Lamb Waves: Physical Theory and Applications*. Plenum Press, 1967.
- [35] Z. Su and L. Ye, “Fundamentals and Analysis of Lamb Waves,” in *Identification of Damage Using Lamb Waves*. Springer London, 2009, vol. 48, pp. 15–58.
- [36] L. Cohen, “Time-frequency distributions - A review,” *Proc. IEEE*, vol. 77, no. 7, pp. 941–981, 1989.
- [37] A. Janssen, “The Zak transform: A signal transform for sampled time-continuous signals,” *Philips J. Res.*, vol. 43, no. 1, pp. 23–69, 1988.
- [38] H. Bolcskei and F. Hlawatsch, “Discrete Zak transforms, polyphase transforms, and applications,” *IEEE Trans. Signal Process.*, vol. 45, pp. 851–866, 1997.
- [39] J. R. O’Hair and B. W. Suter, “The Zak transform and decimated time-frequency distributions,” *IEEE Trans. Signal Process.*, vol. 44, pp. 1099–1110, 1996.
- [40] B. Berglund, P. Hassmen, and R. F. S. Job, “Sources and effects of low-frequency noise,” *J. Acoust. Soc. Amer.*, vol. 99, no. 5, pp. 2985–3002, 1996.
- [41] L. Gaul and S. Hurlebaus, “Identification of the impact location on a plate using wavelets,” *Mech. Syst. Signal Process.*, vol. 12, no. 6, pp. 783–795, 1998.
- [42] F. Ciampa and M. Meo, “Acoustic emission source localization and velocity determination of the fundamental mode A_0 using wavelet analysis and a Newton-based optimization technique,” *Smart Mater. Struct.*, vol. 19, no. 4, p. 045027, 2010.
- [43] K. R. Arun, E. Ong, and A. W. H. Khong, “Source localization on solids using Kullback-Leibler discrimination information,” in *Proc. 8th Int. Conf. Information, Communications and Signal Processing (ICICS)*, 2011, pp. 1–5.

REFERENCES

- [44] V. G. Reju, A. W. H. Khong, and A. B. Sulaiman, "Localization of taps on solid surfaces for human-computer touch interfaces," *IEEE Trans. Multimed.*, vol. 15, pp. 1365–1376, Oct. 2013.
- [45] G. C. Carter, A. H. Nuttall, and P. G. Cable, "The smoothed coherence transform," *Proc. IEEE*, vol. 61, no. 10, pp. 1497–1498, Oct 1973.
- [46] E. J. Hannan and P. J. Thomson, "Estimating group delay," *Biometrika*, vol. 60, no. 2, pp. 241–253, 1973.
- [47] C. Zhang, D. Florencio, and Z. Zhang, "Why does PHAT work well in low noise, reverberative environments?" in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2008.
- [48] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley-Cambridge Press, 1996.
- [49] K. Kishimoto and H. Inoue, "Time frequency analysis of dispersive waves by means of wavelet transform," *J. Appl. Mech.*, vol. 62, no. 4, pp. 841–846, 1995.
- [50] L. Cohen, "The uncertainty principle in signal analysis," in *Proc. IEEE Int. Symp. Time-Frequency and Time-Scale Analysis*, Oct 1994, pp. 182–185.
- [51] D. Gabor, "Theory of communication," *J. Inst. Elect. Eng. - Part III: Radio and Commun. Eng.*, vol. 93, pp. 429–441, Nov. 1946.
- [52] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed. Academic Press, 2008.
- [53] M. Haase and J. Widjajakusuma, "Damage identification based on ridges and maxima lines of the wavelet transform," *Int. J. Eng. Sci.*, vol. 41, pp. 1423 – 1443, 2003.
- [54] P. Bloomfield, *Fourier Analysis of Time Series: An Introduction*, 2nd ed. John Wiley & Sons, Inc., 2005.
- [55] J. D. Cryer and K.-S. Chan, *Time Series Analysis with Applications in R*. Springer New York, 2008.
- [56] M. Last and R. Shumway, "Detecting abrupt changes in a piecewise locally stationary time series," *J. Multivar. Anal.*, vol. 99, pp. 191–214, Feb. 2008.

-
- [57] R. Peck, C. Olsen, and J. Devore, *Introduction to Statistics and Data Analysis*, 5th ed. Boston, MA, USA: Cengage Learning, 2016.
- [58] R. G. Lomax and D. L. Hahs-Vaughn, *An Introduction to Statistical Concepts*, 3rd ed. New York: Routledge, 2012.
- [59] Q. H. Nguyen, V. G. Reju, and A. W. H. Khong, “Source localization on solids utilizing logistic modeling of energy transition in vibration signals,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2016.
- [60] Z. Su and L. Ye, *Identification of Damage Using Lamb Waves: From Fundamentals to Applications*. Springer Verlag, 2009.
- [61] M. Kaphle, A. C. C. Tan, D. P. Thambiratnam, and T. H. T. Chan, “Identification of acoustic emission wave modes for accurate source location in plate-like structures,” *Struct. Control Health Monitoring*, vol. 19, no. 2, pp. 187–198, 2012.
- [62] J. C. Bezdek, R. A. Chandrasekhar, and Y. Attikouzel, “A geometric approach to edge detection,” vol. 6, pp. 52–75, Feb 1998.
- [63] G. Dinse, “An EM algorithm for fitting a four-parameter logistic model to binary dose-response data,” *J. Agric. Biol. Environ. Stat.*, vol. 16, no. 2, pp. 221–232, 2011.
- [64] F. Cavallini, “Fitting a logistic curve to data,” *College Math. J.*, vol. 24, no. 3, pp. 247–253, 1993.
- [65] J. A. Nelder and R. Mead, “A simplex method for function minimization,” *Comput. J.*, vol. 7, no. 4, pp. 308–313, 1965.
- [66] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, “Convergence properties of the Nelder–Mead simplex method in low dimensions,” *SIAM J. Optim.*, vol. 9, pp. 112–147, May 1998.
- [67] M. H. Wright, “Nelder, Mead, and the other simplex method,” *Documenta Mathematica*, pp. 271–276, 2012.
- [68] K. Gröchenig, *Foundations of Time-Frequency Analysis*. Birkhäuser Boston, 2001.

REFERENCES

- [69] H. Jeong and Y. Jang, “Fracture source location in thin plates using the wavelet transform of dispersive waves,” *IEEE Trans. Ultrason. Ferroelectr., Freq. Control*, vol. 47, no. 3, pp. 612–619, 2000.
- [70] S. A. H. Sabzevari and M. Moavenian, “Locating the acoustic source in thin glass plate using low sampling rate data,” *Ultrasonics*, vol. 70, pp. 1 – 11, 2016.
- [71] J. M. Mendel, “Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications,” *Proc. IEEE*, vol. 79, no. 3, pp. 278–305, Mar 1991.
- [72] D. W. Marquardt, “An algorithm for least-squares estimation of nonlinear parameters,” *J. Soc. Ind. Appl. Math.*, vol. 11, no. 2, pp. 431–441, 1963.
- [73] A. Papandreou-Suppappola, R. Murray, and G. Boudreaux-Bartels, “Group delay shift covariant quadratic time-frequency representations,” *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2549–2564, 2001.
- [74] R. Baraniuk, “Warped perspectives in time-frequency analysis,” in *Proc. IEEE Int. Symp. Time-Frequency and Time-Scale Analysis*, Oct 1994, pp. 528–531.
- [75] J. H. Park, Y. S. Moon, and S. K. Lee, “Source localization of multiple impacts in a plate,” *Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci.*, vol. 226, no. 2, pp. 423–438, 2012.