Localization of taps on solid surfaces for human-computer touch interfaces

Reju, Vaninirappuputhenpurayil Gopalan; Khong, Andy Wai Hoong; Sulaiman, Amir

2012

Reju, V. G., Khong, A. W. H., & Sulaiman, A. (2012). Localization of Taps on Solid Surfaces for Human-Computer Touch Interfaces. IEEE Transactions on Multimedia, 15(6), 1365-1376.

<https://hdl.handle.net/10356/97054>

<https://doi.org/10.1109/TMM.2013.2264656>

© 2012 IEEE. This is the author created version of a work that has been peer reviewed and accepted for publication by IEEE Transactions on Multimedia, IEEE. It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at DOI: [<http://dx.doi.org/10.1109/TMM.2013.2264656>].

Downloaded on 26 Apr 2023 18:35:40 SGT

Localization of Taps on Solid Surfaces for Human-Computer Touch Interfaces

Vaninirappuputhenpurayil Gopalan Reju, *Member, IEEE*, Andy W. H. Khong*, *Member, IEEE*,
and Amir Bin Sulaiman

Abstract

Localization of impacts on solid surfaces is a challenging task due to dispersion where the velocity of wave propagation is frequency dependent. In this work, we develop a source localization algorithm on solids with applications to human-computer interface. We employ surface-mounted piezoelectric shock sensors that, in turn, allow us to convert existing flat surfaces to a low-cost touch interface. The algorithm estimates the time-differences-of-arrival between the signals via onset detection in the time-frequency domain. The proposed algorithm is suitable for vibration signals generated by a metal stylus and a finger. The validity of the algorithm is then verified on an aluminium and a glass plate surface.

Index Terms

Human-computer interface, source localization on solids, tangible interfaces, TDOA estimation.

EDICS Category: 3-INTF

V. G. Reju and A. W. H. Khong are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore-639798. (E-mail: reju@ntu.edu.sg; AndyKhong@ntu.edu.sg Phone: +65-67906462; +65-67906008). A. Sulaiman is with Singapore Airlines, Singapore-486854. (E-mail: amir.su@gmail.com Phone: +65-94555920).

This work is supported by the Singapore National Research Foundation Interactive Digital Media R&D Program, under research grant NRF2008IDM-IDM004-010. US Provisional patent application No. US61/507,717 filed on July 14, 2011.

I. INTRODUCTION

Due to the advent of software applications, there is an increasing demand for an easy-to-use human-computer interface (HCI) [1]. The touch interface is one of the most popular HCIs due to its flexibility and ease of use. However, as the size of the surface increases, existing capacitive or resistive touch screens will become prohibitively expensive due to the costs incurred [2]. For larger surfaces however, optical- and camera-based approaches are more commonly used [1]. These systems often require extensive setup and they may not be easily deployed on commonly available surfaces such as wooden or glass tabletops and glass window panes. This motivated researchers to pursue the development of low-cost touch interfaces so that, by mounting several discrete sensors onto existing surfaces, vibration signals generated by an impact can be utilized for localization [3], [4], [5]. This effectively converts an ordinary surface into a touch interface. Unlike optical-based approaches, touch interfaces based on vibration sensing can be used under different lighting conditions and it allows the touch interface to be scalable in size making it attractive for large-scale applications such as the conversion of display panels to touch screens as well as for health-care [6] or entertainment [7].

As with any touch screen technologies, the main challenge in converting a solid surface to a touch interface is the accurate estimation of the impact location using received signals that are extracted from the sensors. In this work, we consider only surface-mounted sensors. Unlike the use of microphones in air, source localization on solid surfaces is challenging due to dispersion where the velocity of wave propagation is dependent on the frequency of the signal [8], [9]. A signal generated on the surface of a solid due to an impact caused by a stylus or finger contains a number of waves at different frequencies [10]. As these waves propagate at different velocities, received signals will become distorted and different from that generated by the impact. In addition, sensors placed at different locations will receive waveforms that are significantly different due to the differences between source-sensor transfer functions. Coupled with this complex process, multipath reflections from the edges as well as the internal structure of the plate make source localization a challenging task.

For source localization on solid surfaces, one of the commonly used approaches include location template matching (LTM) [3], [5], [11] where signals received from known locations are first stored in a database during calibration. Source localization is then performed by matching the sensor output with signals from the database. Although LTM exploits any dissimilarity between the source-sensor transfer functions to differentiate taps made at different locations, any dissimilarity that occurs *within* the same tap location will render a reduction in localization performance. Such ‘intra-location’ dissimilarity may

arise from any inconsistencies in the style of tapping or any deviation of the test condition from that of the calibration phase. The LTM method also requires significant time and effort for calibration since only pre-recorded signals acquired from known positions can be used to locate a given tap. As a result, LTM based systems are impractical for large-scale applications.

Localization using time-difference-of-arrival (TDOA) methods, on the other hand, does not require extensive calibration and it can be used for source tracking and localization [12]. These methods utilize signals received from multiple sensors where the TDOA between two sensors defines a hyperbolic function with the sensor positions corresponding to the foci of the hyperbola. Using several TDOA estimates, the intersection of these hyperbolae provides an estimate of the source location. It is therefore important to note that an accurate TDOA estimate is required for source localization. Due to waveform distortion caused by dispersion, conventional cross-correlation techniques such as proposed in [13] will not achieve the desired result of source localization on solids. In addition, TDOA-based approaches assume known propagation speed. In heterogeneous materials, however, variation of propagation speeds with frequencies often violates such an assumption.

To estimate TDOA in the presence of dispersion, time-frequency approaches such as the use of wavelet transforms have been proposed [14], [15], [16]. These methods are mainly used for ultrasonic testing of materials where the source excitation signals are of known characteristics. For our case however, the source signals may vary in wave shape and frequency characteristics depending on the style of tapping. An alternative approach involves locating the phase points via cross-correlating the sensor outputs and a single-frequency sinusoidal wave that has been modulated by a Gaussian pulse [17]. Since propagation velocity is frequency dependent, only a single frequency is used for TDOA estimation. Utilizing the phase information of the propagating wave and estimating the TDOA at a particular frequency as opposed to the whole spectrum may also address the problem of dispersion [9]. Dispersive signal technology is another technique used in the commercially available touch screens from the company 3M.

In this paper, we propose an algorithm for estimating the TDOA of impacts made on a flat solid surface. This is achieved by detecting the onset of each received signal. Since a signal generated by an impact on a solid surface is, unlike speech, impulsive, it is feasible to develop an onset detection algorithm that operates in a dispersive environment that is relatively free of background vibration such as for table-tops or window panes. Another advantage of the proposed onset detection for TDOA estimation is that even if reflected signals from the surface borders are added to the sensor output, performance of the algorithm will not be significantly affected by the reflected components. In fact, conventional cross-correlation methods can adversely be affected by these multiple acoustic reflections. The Kullback-Leibler

information discrimination (KLID) algorithm [18] that has been proposed for onset detection assumes that the signal is short-time stationary and the variation in power spectrum between the two neighboring signal frames centered at a time-point is used to quantify the amount of discontinuity at that point. In addition, a wavelet technique for time of arrival estimation that is based on the Gabor wavelet has been proposed for the detection of impacts on solids [14]. It is important to note that unlike onset detection methods that have been proposed for electrocardiography (ECG) and electroencephalography (EEG) signals where the number of sensors and computational complexity are often not of primary concern, high accuracy and low processing latency with limited number of sensors are important factors to consider for our touch interface application.

In this work we are exploiting the randomness of the noise signal and impulsive nature of the impact signal. Assuming that the signal-to-noise ratio (SNR) is sufficiently high, the aim of the algorithm is to estimate the point where the randomness of the sensor signals changes, i.e., the onset point. Since the amplitude of a signal can vary significantly and setting an amplitude threshold is challenging, we have used angle information (Hermitian angle) instead of amplitude. Due to dispersion, instead of using the entire frequency components we have selected only a few frequency bins by analyzing the signal in the short-time Fourier transform (STFT) domain. Noting that elements in the STFT domain are complex, angle in a complex vector space is used. The novelty of this work therefore lies in the methodology and its application. For the detection of onset point based on time-frequency analysis, wavelet-based techniques are widely used due to its time-frequency resolution. In this work however, we exploit the STFT due to the following reasons: 1) the discrete Fourier transform (DFT) does not require selection of mother wavelet, 2) base function for the DFT is sinusoidal which is important since we are mitigating the effect of velocity dispersion by selecting only a few adjacent frequency components and 3) the availability of well-established algorithms for the implementation of DFT/FFT such as the sliding DFT [19] when the hop size is one. In addition, we note that the Hermitian angle cannot be used directly on a single-element data such as for our case of onset detection for each channel in the STFT domain. In order to exploit the Hermitian angle, we provide detailed explanation of how to construct the two-element vectors from the received signal of each channel for onset detection. The main contribution of this work is therefore the use of STFT, the selection of a random element (drawn from a uniformly distributed process) and its concatenation with the signal sample to form a vector so that the Hermitian angle between this vector and a reference vector can be utilized for onset detection. In addition, the advantage of the proposed approach is its low complexity for real-time applications. As experiment results will show, our algorithm can achieve good localization performance for an impact generated either by a metal stylus or a finger

Fig. 1. Signals received by four sensors placed at x, y positions (0.1, 0.1) m, (1.1, 0.1) m, (1.1, 0.9) m and (0.1, 0.9) m respectively due to an impact at (0.6, 0.7) m on a glass surface of dimension 1.2 m $\times$ 1.0 m $\times$ 5 mm.

on an aluminium and a glass surface for touch interface applications.

II. PROPOSED TDOA ALGORITHM FOR IMPACT LOCALIZATION

Similar to [9], we employ a set of low-cost surface-mounted Murata PKS1-4A10 shock sensors for localizing an impact made on a solid surface. We note, from a user's perspective, that the impact can be generated via a finger or a stylus. Consider an illustrative example where four sensors are placed at the corners of a glass surface. Figures 1 (a) and (b) show the signals received by the sensors for an impact made by a metal stylus and a finger tap, respectively. It can be seen that the change in amplitude at the onset of the wave arrival is less pronounced for signals generated by the finger. In addition, unlike signals generated by the stylus, no single dominant peak exists for that corresponding to the finger tap. In view of the above, we develop an onset detection algorithm based on the concept of angle in a complex vector space. The algorithm first converts the received signals into the time-frequency (TiF) domain. To minimise the error due to velocity dispersion, only a small range of frequency bins are used for the analysis. The onset of each sensor output is then estimated, using the Hermitian angle (HA), in our proposed TiF-HA onset detection algorithm. The estimated onset and hence the time-of-arrival (TOA) of the signals received by each sensor can be used to compute the TDOA between the signals. Multilateration techniques such as proposed in [20] can then be used for localizing the point of impact.

A. Time-frequency analysis of the received signals

Let $s(n)$ be the source signal generated on a flat surface by an impact where $x_i(n)$ and $x_j(n)$ are the received signals from Sensors i and j placed at different locations on the surface. Defining $h_i(n)$ as the

TABLE I
LIST OF IMPORTANT VARIABLES

Notation	Description
N	Data frame length
M	Number of data frames
K	DFT length
L_i	Length of the leading zeros of $x_i(n)$
$L_{\min}$	Minimum of L_i , $i = 1, 2, \dots$
L_v	Length of leading zeros portion used for noise estimation
k_s	Starting index of the frequency bins used for analysis
k_e	Ending index of the frequency bins used for analysis
Δk	No. of frequency bins used for analysis
$\tau_{i,j}$	TDOA between $x_i(n)$ and $x_j(n)$
L_a	Length of the plate
L_b	Width of the plate
L_c	Thickness of the plate
D	Stiffness of the plate
E	Young's modulus
ν	Poisson's ratio
μ	Absorption coefficient of the plate material
ρ	Density of the plate material
κ and α	Modes of the wave propagation

impulse response between the point of impact and the i th sensor,

$$x_i(n) = h_i(n) * s(n) + v_i(n), \quad (1)$$

where $*$ denotes linear convolution and $v_i(n)$ is the noise picked up by the sensor. As can be seen from Fig. 1, where an impact is generated on a piece of glass plate at (x, y) position (0.6, 0.7) m with sensor positions at (0.1, 0.1) m, (1.1, 0.1) m, (1.1, 0.9) m and (0.1, 0.9) m, respectively, the length of the “leading zeros” varies with the source-sensor distances. In addition, due to dispersion, the waveforms are vastly different and as such, the delay between the two signals cannot be accurately estimated using the conventional cross-correlation method such as described in [13].

In the proposed TiF-HA algorithm, we perform time-frequency analysis by employing the short-time Fourier transform (STFT) of $x_i(n)$. Consider a data frame of length N where

$$\mathbf{x}_i(m) = [x_i(m), \dots, x_i(m + N - 1)]^T, \quad (2)$$

$m = 0, 1, \dots, M$ is the frame index and $(\cdot)^T$ denotes transposition operator. The formulation of (2) implies that the STFT is performed using a one sample shift. Defining $\mathbf{F}_K$ as the $K \times K$ discrete Fourier transform matrix, $K \geq N$, the STFT of $\mathbf{x}_i(m)$ is then given by

$$\begin{aligned}\underline{\mathbf{x}}_i(m) &= \mathbf{F}_K \tilde{\mathbf{x}}_i(m) \\ &= [\underline{x}_i(0, m), \dots, \underline{x}_i(k, m), \dots, \underline{x}_i(K-1, m)]^T,\end{aligned}\quad (3)$$

where $\tilde{\mathbf{x}}_i(m) = [\mathbf{x}_i^T(m) \quad \mathbf{0}_{1 \times (K-N)}]^T$ is a sequence obtained from $\mathbf{x}_i(m)$ after padding it with $K - N$ zeros, $\mathbf{0}_{1 \times (K-N)}$ is the $1 \times (K - N)$ null vector and $k = 0, 1, \dots, K - 1$ is the frequency-bin index. In this paper, frequency-domain signals are denoted by an underscore. It is worth noting that a rectangular windowing function is used for the STFT analysis in order to avoid widening of the signal transition time from the leading-zero period to the signal due to vibration. In STFT analysis the use of a rectangular window will increase spectral leakage. However, since spectral leakage occurs across frequency bins and not time in the time-frequency domain, the use of a rectangular window will not significantly degrade the performance of the proposed onset detection algorithm; additional experiments using Hamming and Hanning windows do not show an improvement in localization performance.

Assuming that $s(n)$ is an impulse and that a tap is made closer to Sensor i compared to Sensor j , we can therefore express the STFT of $x_i(n)$, for the k th frequency bin, as an $M \times 1$ vector

$$\underline{\mathbf{x}}_i(k) = \underline{s}(k) \left[\underbrace{[\underline{h}_i(k, 0), \dots, \underline{h}_i(k, L_i - 1)]}_{\underline{\mathbf{z}}_i(k)}, \underbrace{[\underline{h}_i(k, L_i), \dots, \underline{h}_i(k, M - 1)]}_{\underline{\mathbf{h}}'_i(k)} \right]^T + \underline{\mathbf{v}}_i(k), \quad (4)$$

where $\underline{s}(k)$ is the DFT coefficient of $s(n)$, $\underline{\mathbf{v}}_i(k) = [\underline{v}_i(k, 0), \dots, \underline{v}_i(k, M - 1)]^T$, $\underline{h}_i(k, m)$ is the k th frequency bin of the impulse response from source to the i th sensor while L_i is the length of the “leading zeros”. The subvector $\underline{\mathbf{z}}_i(k)$ corresponds to the STFT coefficients of the “leading zeros” while subvector $\underline{\mathbf{h}}'_i(k)$ corresponds to the non-zero coefficients of the impulse response.

Under practical condition, elements in $\underline{\mathbf{z}}_i(k)$ will not be zero and as a consequence, identifying the onset of the signals from these STFTs is challenging. In addition, in the absence of significant background vibration, the magnitude of elements within $\underline{\mathbf{z}}_i(k)$ will be smaller than that of $\underline{\mathbf{h}}'_i(k)$. In a similar manner, the STFT of the received signal $x_j(n)$ can now be expressed as

$$\underline{\mathbf{x}}_j(k) = \underline{s}(k) \left[\underbrace{[\underline{h}_j(k, 0), \dots, \underline{h}_j(k, L_i + \tau_{ij} - 1)]}_{\underline{\mathbf{z}}_j(k)}, \underbrace{[\underline{h}_j(k, L_i + \tau_{ij}), \dots, \underline{h}_j(k, M - 1)]}_{\underline{\mathbf{h}}'_j(k)} \right]^T + \underline{\mathbf{v}}_j(k), \quad (5)$$

where τ_{ij} is the TDOA between the received signals. The aim therefore is to estimate τ_{ij} using STFTs $\underline{\mathbf{x}}_i(k)$ and $\underline{\mathbf{x}}_j(k)$ for which the onset of $x_i(n)$ and hence TOA of the received signals will be estimated.

We note that when computing the STFT, the data frame length must be selected such that $N < L_{\min}$, where $L_{\min}$ is the length of $\underline{z}_i(k)$ corresponding to the signal having the least number of “leading zeros.” This condition is imposed to ensure that the disparity between noise-only segments and the tap is included during analysis. We note that since samples are continuously streamed from the sensor outputs, all samples after the previous tap signal to the onset of the current tap signal are considered as leading zeros and hence there will be sufficient number of leading zeros for the analysis of the signal. In addition, to achieve high frequency resolution, it is desirable to use higher values of K . Doing so however will increase the computational complexity and due to this tradeoff, we have chosen $K = 1024$ and $N \leq K$.

In the proposed algorithm, a small range of frequency bins are used for the analysis and a hop size of one sample is employed. Hence, computation of a K -point DFT using the fast-Fourier transform (FFT) for every new data sample may become computationally expensive. For real-time processing, the sliding DFT [19], [21] or its stabilized versions [19], [21], [22] may be used. In sliding DFT,

$$\underline{x}_i(k, n) = (\underline{x}_i(k, n-1) - x_i(n-K) + x_i(n)) e^{j2\pi k/K} \quad (6)$$

which requires only two real additions and one complex multiplication per DFT coefficient.

B. The use of Hermitian angle for onset detection

According to the mathematical model for flexural vibration, the vertical displacement of a plate at location (a, b) at time instant t due to an impact at location (a_p, b_p) on the plate surface is given by [3]

$$\gamma_p(a, b, t) = \frac{4}{\rho L_a L_b L_c} \sum_{\kappa=1}^{\infty} \sum_{\alpha=1}^{\infty} \frac{\Omega_{\kappa\alpha}}{\xi_{\kappa\alpha}} (e^{-0.5\tilde{\mu}t} \sin(\xi_{\kappa\alpha}t)), \quad (7)$$

where L_a is the length and L_b is the width of the plate, $\xi_{\kappa\alpha} = \sqrt{\omega_{\kappa\alpha}^2 - (\tilde{\mu}/2)^2}$, $\omega_{\kappa\alpha} = (\tilde{\beta}^2 + \beta^2) \sqrt{\frac{D}{\rho L_c}}$, $D = EL_c^3/12(1 - \nu^2)$ is the stiffness of the plate, E the Young's modulus, ν Poisson's ratio and L_c is the thickness of the plate. In addition, $\tilde{\beta} = \pi\kappa/L_a$, $\beta = \pi\alpha/L_b$, $\tilde{\mu} = \mu/(\rho L_c)$, where μ is the absorption coefficient, ρ is the density of the plate material while $\Omega_{\kappa\alpha} = \sin(\tilde{\beta}a) \sin(\beta b) \sin(\tilde{\beta}a_p) \sin(\beta b_p)$ and $\kappa, \alpha \in \mathbb{Z}^+$ are the modes of wave propagation.

Hence, the signal received by a vibration sensor placed at location (a, b) consists of an infinite sum of exponentially decaying sinusoidal waves whose frequency $\xi_{\kappa\alpha}$ is determined by κ , α and material properties. In this work we assume that the sensor transfer functions across all sensors are the same and hence they impose the same effect on all the sensor outputs. Note that the gain factor $\Omega_{\kappa\alpha}/\xi_{\kappa\alpha}$, as shown in Fig. 2, will vary across frequency components and such variation is not smooth across frequency bins. In order to detect the onset, we first map the amplitude of each component using a nonlinear function so

Fig. 2. Variation in gain factor $\frac{\Omega_{\kappa\alpha}}{\xi_{\kappa\alpha}}$ as a function of frequency $\xi_{\kappa\alpha}$ for glass plate where $\kappa, \alpha = 1, \dots, 10$.

that the amplitudes of the signals will be within the range of 0 to 1. To achieve this, we use the notion of Hermitian angle between two complex vectors.

In a complex vector space, the Hermitian angle between a reference vector $\mathbf{y}$ and $\mathbf{u}_i$ is defined as [23]

$$\theta_{\mathbf{y}, \mathbf{u}_i} = \cos^{-1} \left| \frac{\mathbf{y}^H \mathbf{u}_i}{\|\mathbf{y}\| \|\mathbf{u}_i\|} \right|, \quad (8)$$

where $0 \leq \theta_{\mathbf{y}, \mathbf{u}_i} \leq \pi/2$, $\|\mathbf{u}_i\| = \sqrt{\mathbf{u}_i^H \mathbf{u}_i}$ and the superscript $(\cdot)^H$ denotes the conjugate transpose operator. It is important to note that the Hermitian angle cannot be used directly as a measure on a single-element data since the Hermitian angle between two complex scalars is always zero. In order to employ the Hermitian angle as a measure on these single-element data, we append another element $\underline{w}_i(k, m)$ with $\underline{x}_i(k, m)$ to form a vector $\underline{\mathbf{u}}_i = [\underline{x}_i(k, m), \underline{w}_i(k, m)]^T$. In addition, we generate a two-element reference vector $\mathbf{y} = [y_1, y_2]^T$. To determine $\mathbf{y}$ and the value of $\underline{w}_i(k, m)$ for onset detection, we let

$$\underline{x}_i(k, m) = X_i e^{j\phi_i}, \quad (9)$$

$$\underline{w}_i(k, m) = W_i e^{j\psi_i}, \quad (10)$$

$$y_1 = Y_1 e^{j\varphi_1}, \quad (11)$$

$$y_2 = Y_2 e^{j\varphi_2}, \quad (12)$$

where ϕ_i and ψ_i are the phase of $\underline{x}_i(k, m)$ and $\underline{w}_i(k, m)$ respectively while φ_1 and φ_2 are the phase of y_1 and y_2 respectively. Substituting (9)-(12) into (8), we have

$$\theta_{\mathbf{y}, \mathbf{u}_i} = \cos^{-1} \left| \frac{X_i Y_1 e^{j(\varphi_1 - \phi_i)} + W_i Y_2 e^{j(\varphi_2 - \psi_i)}}{\sqrt{X_i^2 + W_i^2} \sqrt{Y_1^2 + Y_2^2}} \right|. \quad (13)$$

Therefore, for an arbitrary value of y_1 and by selecting $y_2 = 0$, when $\underline{w}_i(k, m) \neq 0$, (13) will reduce to

$$\cos \theta_{\mathbf{y}, \mathbf{u}_i} = \left| \frac{X_i}{\sqrt{X_i^2 + W_i^2}} \right|. \quad (14)$$

Fig. 3. Variation in $\cos \theta_{\mathbf{y}, \mathbf{u}_i}$ as a function of X_i and W_i for different ranges of X_i and W_i .

We illustrate the variation of $\cos \theta_{\mathbf{y}, \mathbf{u}_i}$ for different ranges of X_i and W_i in Fig. 3. In addition, the histogram of $\cos \theta_{\mathbf{y}, \mathbf{u}_i}$ for these cases are plotted in Fig. 4. It can be seen that when X_i and W_i vary within the same range, such as shown in Fig. 3 (a) and (d) where they vary within the range of -1 to 1 and -10 to 10 respectively, the distribution of the values of $\cos \theta_{\mathbf{y}, \mathbf{u}_i}$ is more uniform within 0 to 1 compared to when X_i and W_i vary significantly with respect to each other (Fig. 3 (b) and (c)). We also note that Fig. 3 (a) is identical to Fig. 3 (d) since, from a geometric viewpoint, X_i , W_i and $\sqrt{X_i^2 + W_i^2}$ forms a right-angled triangle and hence the angle between X_i and $\sqrt{X_i^2 + W_i^2}$ will remain the same even if X_i and W_i are scaled by the same factor. Hence if we select $\underline{w}_i(k, m)$ such that the amplitudes of $\underline{w}_i(k, m)$ are in the range of $\underline{x}_i(k, m)$ during the leading zeros period of $\underline{\chi}_i(k)$, the variation in Hermitian angle calculated during the leading zeros period will be high. On the contrary, the variation in Hermitian angle will be low during the remaining period. If we repeat the same procedure within a selected range of frequency bins, the standard deviation of Hermitian angles across the frequency bins will be high during the leading zeros period and low for the remaining period. We note that $\underline{w}_i(k, m)$ can be a constant for all the values of k and m . However, if the noise is coloured, the standard deviation in Hermitian angle during the leading zeros period will be low and in order to avoid this we have used random values as explained later in the next paragraph.

Hence the procedure for detecting the onset of $x_i(n)$ begins by removing the DC offset of $x_i(n)$ when $E[\mathbf{v}_i(n)]$ is subtracted from $x_i(n)$, where $\mathbf{v}_i(n)$, of length $L_v < L_{\min}$, is a segment of the data taken from the “leading zeros” portion of $x_i(n)$. We then employ the Hermitian angle that is computed between $\mathbf{y} = [1 + j, 0]^T$ and vector $\underline{\mathbf{u}}_i = [\underline{x}_i(k, m) \quad \underline{w}_i(k, m)]^T$ where the sample $\underline{w}_i(k, m)$ is drawn from a

Fig. 4. Histogram of $\cos \theta_{y,u_i}$ shown in Fig. 3.

Fig. 5. Hermitian angles $\theta_{y,u_i}(k, m)$, in degrees, calculated for two different signals (a) for signal generated due to stylus impact and (b) for the finger tap.

uniformly distributed process in the TiF domain. To determine the value of $\underline{w}_i(k, m)$ and to maximize the standard deviation of the Hermitian angles during the “leading zeros” period, a set of uniformly distributed samples is generated in the TiF domain itself. This is achieved by constructing a $K \times M$ complex matrix $\underline{\mathbf{W}}_i$ such that elements of this matrix satisfy

$$\sum_{k=k_s}^{k_e} \sum_{m=0}^{L_v-1} |\underline{w}_i(k, m)|^2 = \alpha \sum_{k=k_s}^{k_e} \sum_{m=0}^{L_v-1} |\underline{x}_i(k, m)|^2 \quad (15)$$

where α is the proportionality constant which determines the power of $\underline{w}_i(k, m)$, k_s and k_e defines the starting and ending indices of the frequency bins while $\underline{w}_i(k, m)$ is the (k, m) th element of $\underline{\mathbf{W}}_i$. As explained earlier, elements of $\underline{\mathbf{W}}_i$ that are generated using a uniform distribution within the interval -0.5 to 0.5 before applying the constraint in (15), will allow the algorithm to achieve a high variance of the Hermitian angle during the “leading zeros” period and vice versa.

We explain how $\underline{\mathbf{W}}_i$ can assist the algorithm to achieve high variance in terms of the Hermitian angle during the “leading zeros” period. From (1), the concatenation of $\underline{x}_i(k, m)$ and $\underline{w}_i(k, m)$ can now be

expressed as

$$\begin{aligned}
\mathbf{u}_i &= [\underline{x}_i(k, m) \quad \underline{w}_i(k, m)]^T \\
&= \begin{bmatrix} \underline{s}(k) & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \underline{h}_i(k, m) \\ \underline{v}_i(k, m) \\ \underline{w}_i(k, m) \end{bmatrix} \\
&= \begin{bmatrix} \underline{s}(k) \\ 0 \end{bmatrix} \underline{h}_i(k, m) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \underline{v}_i(k, m) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \underline{w}_i(k, m).
\end{aligned} \tag{16}$$

Assuming $s(n)$ is an impulse, at all points in the TiF plane $\underline{s}(k)$ will be a constant while $\underline{h}_i(k, m)$, $\underline{v}_i(k, m)$ and $\underline{w}_i(k, m)$ are both time and frequency dependent. During the “leading zeros” period, $\underline{h}_i(k, m) = 0 \forall k, m$ and therefore

$$\mathbf{u}_i = [\underline{v}_i(k, m) \quad \underline{w}_i(k, m)]^T. \tag{17}$$

Since elements $\underline{w}_i(k, m)$ are uniformly distributed and are independent of $\underline{v}_i(k, m)$, the Hermitian angle between $\mathbf{y}$ and the resultant vector

$$\begin{bmatrix} \underline{x}_i(k, m) \\ \underline{w}_i(k, m) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \underline{v}_i(k, m) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \underline{w}_i(k, m) \tag{18}$$

will vary significantly between the maximum range from 0 to $\pi/2$ rad. We note that for this to occur, it is desirable for $\underline{w}_i(k, m)$ to have approximately the same power as $\underline{v}_i(k, m)$. This is to avoid either term from being dominant since if that is the case, $\theta_{\mathbf{y}, \mathbf{u}_i}$ will predominantly be influenced by the dominant term which, in turn, will result in a low variation of the Hermitian angle that is undesirable. We therefore propose to use an empirically determined value of $\alpha = 0.0625$.

When $\underline{h}_i(k, m) \neq 0$, i.e., from the onset point onwards, all three terms on the right of (16) will be non-zero. Hence $\theta_{\mathbf{y}, \mathbf{u}_i}$ will vary across different points in the TiF plane. However, this variation is expected to be smaller than that during the “leading zeros” period since there is now lesser influence of the second and third terms in (16). It has also been shown in [24] that the Hermitian angle between two complex vectors will remain the same even if the vectors are multiplied by any complex scalar. The above implies that, for this period after the onset,

$$\theta_{\mathbf{y}, \mathbf{u}_i} \approx \theta_{\mathbf{y}, [\underline{s}(k) \ 0]^T \underline{h}_i(k, m)} = \theta_{\mathbf{y}, [\underline{s}(k) \ 0]^T}. \tag{19}$$

Since $\underline{s}(k)$ is non-varying, it is therefore expected that the variation of $\theta_{\mathbf{y}, \mathbf{u}_i}$ is low during the period after the onset.

To further illustrate the above, Figs. 5 (a) and (b) show the Hermitian angles across time and frequency for $x_1(n)$ illustrated in Fig. 1 (a) and (b), respectively. As can be seen from Fig. 5, the variation of Hermitian angles during the “leading zeros” period is higher compared to those after the onset point. More importantly, there exist many spurious peaks for the signal generated by a finger after the onset. This makes onset detection a challenging task especially for a finger tap.

We next quantify the variation of Hermitian angles by defining the standard deviation

$$\begin{aligned}\sigma_i(m) &= \sqrt{\frac{1}{\Delta k + 1} \sum_{k=k_s}^{k_e} (\theta_{\mathbf{y}, \mathbf{u}_i}(k, m) - \mu_i(m))^2}, \\ \mu_i(m) &= \frac{1}{\Delta k + 1} \sum_{k=k_s}^{k_e} \theta_{\mathbf{y}, \mathbf{u}_i}(k, m)\end{aligned}\tag{20}$$

computed across frequency bins from k_s to k_e , where $\Delta k = k_e - k_s$. Defining

$$\boldsymbol{\sigma}_i(m) = [\sigma_i(0), \dots, \sigma_i(L_i + \tau_{ij} - 1), \sigma_i(L_i + \tau_{ij}), \dots, \sigma_i(M - 1)]^T,\tag{21}$$

it is therefore expected that $\sigma_i(m_1) > \sigma_i(m_2)$ for $m_1 \in \{0, \dots, L_i + \tau_{ij} - 1\}$ and $m_2 \in \{L_i + \tau_{ij}, \dots, M - 1\}$. Hence, the detected transition point in $\boldsymbol{\sigma}_i(m)$ computed over a complete period of the signal corresponds to the onset of the signal. The differences between these onset times will result in the TDOAs of the signals. We note that in (20) the standard deviation is calculated across different frequency bins under the assumption that the variation in signal velocity across a small range of adjacent frequency bins will be negligibly small. This assumption is verified in Fig. 6 where the contour plot of the absolute value of the DFT coefficients of one of the sensor outputs, expressed in dB, is shown in the time-frequency plane. This plot is generated using $K = 1024$ frequency bins at a sampling frequency of $f_s = 96$ kHz. For clarity, only frequencies of up to 10 kHz are shown. It is evident that the high-frequency components arrive at the sensor earlier than the low-frequency components, i.e., the velocity of the signal increase with frequency. However, for a small bandwidth this velocity difference is negligible.

As an illustrative example, Figs. 7 (a) and (c) show $x_i(n)$ and its corresponding $\boldsymbol{\sigma}_i(m)$ respectively. An offset of 128 samples observed between the transition point in $\boldsymbol{\sigma}_i(m)$ and the onset of $x_i(n)$ is due to the effect of STFT analysis. This offset is equivalent across all channels and can be ignored during the analysis since the same STFT analysis window length is used. The main advantage of using $\boldsymbol{\sigma}_i(m)$ instead of directly using $x_i(n)$ for onset detection is now apparent; unlike in $x_i(n)$, whose amplitude can vary significantly, the variation of $\theta_{\mathbf{y}, \mathbf{u}_i}(k, m)$ and hence that of $\boldsymbol{\sigma}_i(m)$ will be limited to between 0 and $\pi/2$ rad. This makes the selection of the threshold relatively simpler and robust.

Fig. 6. Contour plot of the absolute value of the DFT coefficients of one of the sensor outputs, expressed in dB, in the time-frequency plane. (a) The whole signal length and (b) the expanded view of the frame index from 250 to 350.

Fig. 7. Different stages of the TiF-HA algorithm (a) sensor output, $x_i(n)$, (b) histogram of $\sigma_i(m)$, (c) standard deviation of the Hermitian angles, $\sigma_i(m)$, (d) cumulative histogram, (e) detected transition points and (f) histogram of the cumulative histogram.

C. Determination of threshold

The problem of detecting the transition point in $\sigma_i(m)$ is now translated to one that involves finding an appropriate threshold, as shown in Fig. 7 (c). The aim of this threshold is to separate samples in $\sigma_i(m)$ into two groups; one corresponding to the “leading zeros” region while the other corresponding to the active region of the signal. We employ the histogram of $\sigma_i(m)$ to estimate the threshold. Defining

B as the number of histogram bins, the histogram therefore satisfies the relationship

$$M = \sum_{b=1}^B F_i(b) \quad (22)$$

where $F_i(b)$ denotes, for the i th channel, the number of occurrence that corresponds to each histogram bin index b . The histogram corresponding to $\sigma_i(m)$ illustrated by Fig. 7 (c) is plotted in Fig. 7 (b). As can be seen, the larger values of $\sigma_i(m)$ for $m = 0, \dots, L_i + \tau_{ij} - 1$ compared to the rest of the period, $m = L_i + \tau_{ij}, \dots, M - 1$, will result in a histogram bin index $b_{\min}$ that contains the least number of occurrence in $F_i(b)$. This corresponds to the threshold value that is to be applied to $\sigma_i(m)$. Hence the objective here is to estimate a threshold value $b_{\min}$ where

$$b_{\min} = \arg \min_b F_i(b), \quad b_1 < b < b_2 \quad (23)$$

such that the transition point of $\sigma_i(m)$ determined by $b_{\min}$ corresponds to the signal onset and that b_1 and b_2 are the bin indices nearest to the cluster centroids obtained by clustering elements of $\sigma_i(m)$ into two clusters using k-means clustering. The analysis of (23) is limited between b_1 and b_2 in order to avoid the minima at both ends of $F_i(b)$.

To find $b_{\min}$, we note that $F_i(b)$ may contain local minima and evaluating (23) directly may not achieve our desired accuracy. We therefore obtain the cumulative histogram of $F_i(b)$ such that the cumulated sum $U_i(b)$ satisfies

$$U_i(b) = \sum_{p=b_1}^b F_i(p), \quad b_1 < b. \quad (24)$$

As shown in Fig. 7 (d), the abscissa corresponding to the point of inflection of $U_i(b)$ will be equivalent to $b_{\min}$. For the estimation of this inflection point, we compute the number of occurrence $\mathcal{N}_i$ for which $\sigma_i(m)$ occurs for a given range of $U_i(b)$ as illustrated in Fig. 7 (f). Since the point of inflection occurs when $\sigma_i(m)$ occurs the most number of times for a given range of $U_i(b)$, the point of inflection can be determined by finding $U_i(b)$ that gives the maximum $\mathcal{N}_i$. Therefore, the bin center corresponding to the maximum number of $\mathcal{N}_i$ shown in the abscissa of Fig. 7 (f) is equivalent to the ordinate value of the inflection point shown in Fig. 7 (d). Using this value, the threshold and hence $b_{\min}$ can be determined.

As an illustrative example, for the case shown in Fig. 7 (f), the maximum $\mathcal{N}_i$ corresponds to $U_i(b) = 163$ and from Fig. 7 (d), we note that the threshold of $b_{\min} = 7.33$ is determined. In the event if more than one maxima exist with equal number of $\mathcal{N}_i$, computation of $\mathcal{N}_i$ will be repeated using half the number of bins until a single maxima exist.

D. Detection of signal onset

Having determined the threshold, we next apply it on $\sigma_i(m)$ to determine the transition point in $\sigma_i(m)$. This can be achieved by determining the frame indices where the elements in $\sigma_i(m)$ cross the threshold. However, in some cases such as shown in Fig 7 (e) for a finger tap, more than one transition points may be detected due to spurious peaks present in $\sigma_i(m)$. To eliminate these spurious peaks, at the points where the transitions are detected, the sample of $\sigma_i(m)$ at those points are replaced by the mean of its two adjacent samples, provided if these adjacent samples are higher or lower than the sample at the point where the transition is detected, i.e.,

$$\sigma_i(m) = \begin{cases} 0.5(\sigma_i(m-1) + \sigma_i(m+1)) & \text{if } \sigma_i(m-1) > \sigma_i(m) \\ & \sigma_i(m+1) > \sigma_i(m) \\ 0.5(\sigma_i(m-1) + \sigma_i(m+1)) & \text{if } \sigma_i(m-1) < \sigma_i(m) \\ & \sigma_i(m+1) < \sigma_i(m) \\ \sigma_i(m) & \text{otherwise} \end{cases} \quad (25)$$

This process is to keep the samples at the actual transition point of $\sigma_i(m)$ unaltered since the adjacent samples at this point will be above and below the threshold level.

However, if the false transition is due to two or more adjacent spikes in $\sigma_i(m)$, the above method will not remove all the spikes. Therefore, in addition to the above, $\sigma_i(m)$ is smoothed by a three-point moving average filter. After the spurious spike elimination stage, the threshold detection process, described in Section II-C, is repeated until a single transition point is detected or a predefined number of cycles is reached. If the algorithm fails to determine the threshold within, for example, ten cycles, the corresponding sensor output will be discarded. We note that when the signal-to-noise ratio (SNR) of the sensor output is high, such as for impacts generated by a metal stylus, spurious spikes will less likely to occur in $\sigma_i(m)$ as illustrated in Fig. 8 (b).

E. Location estimation via onset detection

The detected transition point will be considered as the onset of the sensor output and only those signal onsets which have been estimated successfully will be used for TDOA estimation. The estimated TDOA $\hat{\tau}_{i,j}$, in terms of number of samples, between two sensor outputs will be computed as the difference between indices of the onsets. Having estimated the $\hat{\tau}_{i,j}$, one could utilize a pre-estimated velocity of the signal or a closed-form solution where both the source location and velocity of the signal are estimated simultaneously [4], [25], [26], [27]. In the first method, source signals are first generated by the user at pre-defined locations on the surface. Assuming known sensor positions, we define d_i as the distance

Fig. 8. Signal picked up by a sensor due to an impact generated by a metal stylus. (a) Sensor output and (b) standard deviation of the Hermitian angles.

between the source and the i th sensor and $c(f)$ as the velocity of the signal, $c(f)$ can then be estimated by

$$\hat{c}(f) = \arg \min_{c(f)} \sum_{i,j} \left(\frac{d_i - d_j}{c(f)} - \hat{\tau}_{i,j} \right)^2. \quad (26)$$

where $\hat{\tau}_{i,j}$ is the TDOA estimated across frequency bins k_s to k_e using the TIF-HA algorithm described above. Such minimization can be performed using, for example, the Levenberg Marquadt algorithm [28]. Once $\hat{c}(f)$ is estimated, the unknown source position (x_s, y_s) , can be estimated by minimizing the error function

$$(\hat{x}_s, \hat{y}_s) = \arg \min_{x_s, y_s} \sum_{i,j} \left(\frac{d_i - d_j}{\hat{c}(f)} - \hat{\tau}_{i,j} \right)^2 \quad (27)$$

$$d_i = \sqrt{(x_s - x_i)^2 + (y_s - y_i)^2}$$

where (x_i, y_i) is the position of the i th sensor. The above process of velocity estimation is often referred to as calibration and in this work, we employ such process prior to localizing a tap made by a finger or stylus.

III. SIMULATION AND EXPERIMENTAL RESULTS

Experiments are conducted on aluminium and glass surfaces of dimensions $0.6 \text{ m} \times 0.6 \text{ m} \times 2.5 \text{ mm}$ and $0.6 \text{ m} \times 0.6 \text{ m} \times 5 \text{ mm}$, respectively. They are securely mounted on a table with the help of four clamps fastened to the sides of the plates. To acquire the vibration signals due to an impact, eight Murata PKS1-4A10 piezoelectric shock sensors are mounted on the plate surface 0.1 m away from the edges as shown in Fig. 9. These sensors, being most sensitive in the neighborhood of 2 to 4 kHz, are mounted on the surface using double-sided adhesive tape which can easily be detached to allow portability and scalability for HCI

Fig. 9. Point of impacts and sensor locations on aluminium and glass surfaces. The impact locations for algorithm evaluation, pre-estimation of velocity and the sensor positions are denoted by $\bullet$, $\times$ and $\blacksquare$ respectively.

TABLE II
EXPERIMENTAL CONDITIONS

Parameters	TiF-HA	KLID [18]
Sampling frequency	96 kHz	96 kHz
FFT size, K	1024	256
Data frame length, N	256	256
L_v	256	-
α	0.0625	-
Data length	1024	2048

applications. The values of different parameters used for the experiments are given in Table II. Data is acquired using an 8-channel 24-bit simultaneous sampling data acquisition system at a sampling frequency of 96 kHz to ensure good time-delay resolution. It is also important to note that data from all sensors have to be synchronized prior to processing. For real-time implementation, we have used the ADZS-21469 EZLITE Sharc development and its daughter board. Prior to real-time implementation, we verified that the eight channels are time synchronized by transmitting a white Gaussian noise simultaneously into all channels and cross-correlating the received signals. These inter-channel cross-correlations exhibit peaks at zero-lag verifying that the signals are time-synchronized.

In order to estimate the location of an impact, we used pre-estimated velocities derived from sixteen locations denoted by $\times$ in Fig. 9. Each location is tapped a total of five times giving eighty sets of eight-channel received data. Knowing the tap locations and sensor positions, velocities of the signal across frequencies are then estimated as described in Section II-E. Since the velocity of the signal varies across frequencies, for TiF-HA, we estimated the velocity for 3 to 10 kHz. The estimated velocities are shown in Fig. 10 where $\bullet$ and $\times$ denote the velocities for aluminium and glass plate, respectively. It can be seen that, apart from a few outliers, velocity of the wave propagation varies linearly across frequencies

Fig. 10. Variation of estimated velocities with frequencies. The estimated velocities are denoted by $\bullet$ and $\times$ whereas solid lines represent the velocity after smoothing.

and hence, we approximate these velocities using a straight line as shown in Fig. 10. This straight line is obtained using weighted linear least squares and a first degree polynomial [29]. Since speed estimation was performed by tapping the plate manually, large number of points are used in this one-time process.

To evaluate the localization performance, we used forty-five sets of eight-channel sensor outputs corresponding to five taps at each of the nine tap locations denoted by $\bullet$ as shown in Fig. 9. It is important to note that, unlike LTM-based approaches where only a discrete set of tap locations can be localized, localization of taps using TiF-HA is *not* limited to those denoted by $\bullet$ shown in Fig. 9. In addition, for each experimental result presented, we evaluated the performance of the algorithm using a metal stylus as well as the index finger. These tests are important as they illustrate the feasibility and ease of use of the user interface.

Since we are evaluating $\sigma_i(m)$ over a range of frequency-bins as shown in (20), we first investigate the influence of Δk on the localization accuracy over a frequency range of $f_c = 3$ to 10 kHz, where f_c is the frequency corresponding to the center frequency bin, $(k_s + k_e)/2$, used for analysis. As TiF-HA uses the standard deviation of Hermitian angles across frequency bins, small values of Δk will not achieve a reliable onset point detection. On the contrary, having a large value of Δk will increase the computational cost in addition to the higher localization errors due to dispersion. We therefore evaluate how the localization accuracy varies across center frequencies $3 \leq f_c \leq 10$ kHz for different values of Δk . In this work we evaluate the performance of TiF-HA for $\Delta k = 32$ and 64 where one bin corresponds to 93.75 Hz.

Localization errors quantified using root-mean-square error (RMSE) are plotted as shown in Figs. 11 and 12 for aluminium and glass surface, respectively. For each of these plots, subplots (a) and (b) illustrate RMSE for impacts generated by a metal stylus and finger, respectively. Results shown in these figures are averaged across the forty-five sets of data. In order to evaluate the robustness of the algorithm, the

Fig. 11. Variation of RMSE and standard deviation of localization errors with center frequency on aluminium surface. (a) RMSE when the impacts are due to stylus, (b) RMSE when the impacts are due to finger, (c) standard deviation in estimation error when the impacts are due to stylus and (d) standard deviation in estimation error when the impacts are due to finger.

Fig. 12. Variation of RMSE and standard deviation of localization errors with center frequency on glass surface. (a) RMSE when the impacts are due to stylus, (b) RMSE when the impacts are due to finger, (c) standard deviation in estimation error when the impacts are due to stylus and (d) standard deviation in estimation error when the impacts are due to finger.

standard deviations of the estimation error are also shown in Figs. 11 and 12, below the corresponding RMSE plots. In the above figures, higher values of the standard deviation corresponding to the higher values of RMSE are due to the presence of a few outliers in the estimation error. From these results, we note that location estimates are reliable over a center frequency range of 3 to 7 kHz and the error reduces with increasing Δk . Additional experiments showed that beyond $\Delta k = 64$, any improvement in performance is not justifiable for the additional computational cost and dispersive effect of wave propagation.

Comparing subplots (a) and (b) for Figs. 11 and 12, we note that the proposed TiF-HA algorithm is more robust in localizing impacts due to a stylus compared to a finger. For the case of finger taps,

Fig. 13. Variation of impact localization error with signal velocity for $\Delta k = 64$ at $f_c = 4968$ Hz.

Fig. 14. Variation of RMSE with SNR on a simulated glass plate.

the localization error is significant for higher frequencies particularly when the value of Δk is small. However, for suitable values of Δk and f_c TiF-HA is able to localize a finger tap with reasonable accuracy of less than 2 cm. Although such accuracy may not be feasible for small hand-held devices, the proposed algorithm offers an attractive solution for large surfaces where high resolution is not required. Since the algorithm uses estimated velocity for localization of the impact, the effect of the estimated velocity on localization error is shown in Fig. 13. As expected, any errors arising from velocity estimation will affect the accuracy of localization. We also note that the velocity corresponding to the minimum error for aluminium and glass plate are 1724 m/s and 1638 m/s respectively which are equivalent to the estimated velocities in Fig. 10 at $f_c = 4968$ Hz.

In order to evaluate the robustness of the algorithm against noise, we evaluate its performance using simulated sensor outputs on a 5 mm-thick glass plate for various levels of added white Gaussian noise. In order to simulate the sensor output we used the mathematical model for flexural vibration described in (7) [30] and assumed the sensor transfer function in Laplace domain as $0.00035/(s^2 + 0.34\omega s + \omega^2)$ where $\omega = 12943$ rad/s [31]. The different parameters for the mathematical model for flexural vibration are $E = 65 \times 10^9$ Pa, $\nu = 0.17$, $\rho = 2190$ kg/m³, $\mu = 3000$, $\kappa = 1, \dots, 100$ and $\alpha = 1, \dots, 100$. The size of the plate, tap locations and sensor positions are the same as that shown in Fig. 9. Fig. 14 (a)

Fig. 15. Performance comparison of the proposed algorithm with that of KLID [18] and wavelet based method [14]. (a)RMSE in localization and (b) standard deviation of localization error.

Fig. 16. Localization of the impacts due to finger taps on glass surface using (a) TiF-HA and (b) KLID algorithms. The sensor coordinates, actual tap locations and estimated locations are denoted by ■, ● and × respectively.

and (b) show the simulated sensor outputs before and after the addition of noise at 15 dB SNR. The estimation error is evaluated over a frequency range of 3 to 10 kHz for $\Delta k = 32$ and 64. The RMSE obtained over the frequency range at different noise levels are shown in Fig. 14 (c). We note that the RMSE reduces with increasing SNR as expected. In addition, comparing with experimental results shown in Fig.12, the localization error is modestly higher than those obtained using real recorded data. This is because we have only used $\kappa, \alpha = 1, \dots, 100$ in the simulated data whereas in the real case, κ and α may not be limited to these values. Nevertheless, experimental results validates the performance of the proposed algorithm.

We next compare, using recorded data, the localization performance of the proposed algorithm with KLID algorithm proposed for onset detection [18] and a wavelet based method which had been proposed for impact localization on plate surfaces [14]. In this experiment we have used signal components in the frequency range 3 to 7 kHz. These time of arrivals are then used in conjunction with the Levenberg-

Marquardt optimization algorithm [28] for the estimation of source location. The localization performance of TiF-HA, KLID [18] and wavelet method [14] on aluminium and glass surfaces are shown in Fig. 15. In this experiment, unlike TiF-HA, a single estimated velocity is used for KLID. This velocity is computed using the mean of the velocities estimated across eighty taps during calibration for source localization whereas the wavelet-based method does not require any velocity information. Figure 15 (a) compares the RMSE whereas in Fig. 15 (b), the standard deviation of the estimation errors are compared. From these results we note that, for an impact generated by the stylus on an aluminium surface, the RMSE for TiF-HA is only modestly higher than KLID by 3.5 mm. The TiF-HA algorithm, on the other hand, outperforms KLID on a glass when a stylus is used. This improvement can be observed by a lower RMSE of 14.4 mm for TiF-HA (compared to 18.6 mm for KLID) and a lower standard deviation of 7.8 mm for TiF-HA (compared to 14.8 mm for KLID). In all the cases both TiF-HA and KLID outperforms the wavelet based method.

The performance improvement of TiF-HA over KLID is also evident from results presented for impacts generated by a finger. As can be seen from Fig. 15, TiF-HA can achieve a 10.3 mm and 9.2 mm reduction in terms of RMSE compared to KLID for aluminium and glass surfaces, respectively. The standard deviations of TiF-HA for these surfaces are significantly lower compared to KLID and the wavelet-based algorithm. It is also important to note that the performance of TiF-HA for each material is consistent across taps generated by a stylus and a finger. These results illustrate the robustness of TiF-HA due to the determination of threshold presented in Sections II-C and II-D.

To further illustrate its potential application for human-computer interface, we illustrate the estimated tap locations for TiF-HA and KLID on a glass surface in Fig. 16 (a) and (b), respectively. We chose to present results for a glass surface since this material is most commonly used for a touch screen. Results for the TiF-HA algorithm are obtained using $f_c = 4968$ Hz and $\Delta k = 64$. These results indicate that the localization errors of the proposed TiF-HA algorithm are lower which, in turn, reinforces its suitability for HCI.

Although the above experiments have been conducted using eight sensors, further experiments conducted showed that the performance of the proposed algorithm generally degrades with reducing number of sensors as expected. More specifically, when the algorithm is incorporated with the joint source localization technique [4], the performance degradation is more significant than the calibration method. In addition, the localization performance is expected to degrade with reducing sampling frequency.

IV. CONCLUSION

We have proposed an algorithm for impact localization on flat solid surfaces using surface-mounted sensors via time-frequency analysis. The algorithm is suitable for the localization of impacts generated either by a metal stylus or a finger. The main advantages of the algorithm are its simplicity, robustness and accuracy in localizing these impacts. The proposed algorithm have been implemented and tested on a SHARC AD21469 DSP processor and the algorithm is found to be suitable for real-time human computer interface applications. Mathematical analysis of the algorithm for optimum performance has not been performed in this work and will be addressed in near future.

REFERENCES

- [1] S. J. V. Nichols, "New interfaces at the touch of a fingertip," *Computer*, vol. 40, pp. 12 – 15, Aug. 2007.
- [2] P. I. S. Lei and A. K. Y. Wong, "The multiple-touch user interface revolution," *IT Professional*, vol. 11, pp. 42 – 49, Jan. 2009.
- [3] 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. Multimedia*, vol. 13, pp. 487 – 497, June 2011.
- [4] J. Zheng, K. W. K. Lui, and H. C. So, "Accurate three-step algorithm for joint source position and propagation speed estimation," *Signal Processing*, vol. 87, pp. 3096 – 3100, Dec. 2007.
- [5] D. T. Pham, Z. Wang, Z. Ji, M. Yang, M. Al-Kutubi, and S. Catheline, "Acoustic pattern registration for a new type of human-computer interface," in *Proc. Virtual Conf., IPROMs*, 2005.
- [6] M. Akay, S. Member, I. Marsic, A. Medl, and G. Bu, "A system for medical consultation and education using multimodal human/machine communication," *IEEE Trans. Information Technology in Biomedicine*, vol. 2, pp. 282 – 291, Dec. 1998.
- [7] A. Crevoisier and P. Polotti, "Tangible acoustic interfaces and their applications for the design of new musical instruments," in *Proc. Int'l Conf. on New Interfaces for Musical Expression*, pp. 97 – 100, 2005.
- [8] W. Sachse and Y. Pao, "On the determination of phase and group velocities of dispersive waves in solids," *Journal of Applied Physics*, vol. 49, pp. 4320 – 4327, Aug. 1978.
- [9] A. Sulaiman, K. Poletkin, and A. W. H. Khong, "Source localization in the presence of dispersion for next generation touch interface," in *Proc. Int'l Conf. on Cyberworlds*, pp. 82 – 86, 2010.
- [10] 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'l Conf. on Multimedia and Expo*, pp. 286 – 291, 2010.
- [11] R. K. Ing, N. Quieffin, S. Catheline, and M. Fink, "In solid localization of finger impacts using acoustic time-reversal process," *Applied Physics Letters*, vol. 87, pp. 1 – 3, Nov. 2005.
- [12] D. T. Pham, Z. Ji, M. Yang, Z. Wang, and M. Al-Kutubi, "A novel human-computer interface based on passive acoustic localisation," *Lecture Notes in Computer Science*, vol. 4551, pp. 901 – 909, 2007.
- [13] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 24, pp. 320 – 327, Aug. 1976.
- [14] L. Gaul and S. Hurlebaus, "Identification of the impact location on a plate using wavelets," *Mechanical Systems and Signal Processing*, vol. 12, no. 6, pp. 783 – 795, 1998.

- [15] H. Jeong and Y. Jang, "Wavelet analysis of plate wave propagation in composite laminates," *Composite Structures*, vol. 49, pp. 443 – 450, Aug. 2000.
- [16] H. Jeong and Y. Jang, "Fracture source location in thin plates using the wavelet transform of dispersive waves," *IEEE Trans. Ultrasonics, Ferroelectrics and Frequency Control*, vol. 47, pp. 612 – 619, Aug. 2002.
- [17] S. M. Ziola and M. R. Gorman, "Source location in thin plates using cross-correlation," *J. Acoust. Soc. Am.*, vol. 90, pp. 2551 – 2556, Nov. 1991.
- [18] M. Last and R. Shumway, "Detecting abrupt changes in a piecewise locally stationary time series," *Journal of Multivariate Analysis*, vol. 99, pp. 191–214, Feb. 2008.
- [19] E. Jacobsen and R. Lyons, "The sliding DFT," *IEEE Signal Processing Magazine*, vol. 20, pp. 74 – 80, Mar. 2003.
- [20] H. B. Lee, "A novel procedure for assessing the accuracy of hyperbolic multilateration systems," *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-11, pp. 2 – 15, Jan. 1975.
- [21] E. Jacobsen and R. Lyons, "An update to the sliding DFT," *IEEE Signal Processing Magazine*, vol. 21, pp. 110 – 111, Jan. 2004.
- [22] K. Duda, "Accurate, guaranteed stable, sliding discrete Fourier transform," *IEEE Signal Processing Magazine*, vol. 27, pp. 124 – 127, Nov. 2010.
- [23] K. Scharnhorst, "Angles in complex vector spaces," *Acta Applicandae Mathematicae*, vol. 69, p. 95–103, Nov. 2001.
- [24] V. G. Reju, S. N. Koh, and I. Y. Soon, "Underdetermined convolutive blind source separation via time-frequency masking," *IEEE Trans. Audio, Speech & Language Processing*, vol. 18, pp. 101–116, Jan. 2010.
- [25] H. Chen, G. Wang, Z. Wang, H. C. So, and H. V. Poor, "Non-line-of-sight node localization based on semi-definite programming in wireless sensor networks," *IEEE Trans. Wireless Communications*, vol. 11, no. 1, pp. 108 – 116, 2012.
- [26] G. Wang and H. Chen, "An importance sampling method for TDOA-based source localization," *IEEE Trans. Wireless Communications*, vol. 10, no. 5, pp. 1560 – 1568, 2011.
- [27] Q. Shi, C. He, H. Chen, and L. Jiang, "Distributed wireless sensor network localization via sequential greedy optimization algorithm," *IEEE Trans. Signal Processing*, vol. 58, no. 6, pp. 3328 – 3340, 2010.
- [28] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *The Quarterly of Applied Mathematics*, vol. 2, pp. 164 – 168, 1944.
- [29] W. S. Cleveland and S. J. Devlin, "Locally-weighted regression: an approach to regression analysis by local fitting," *Journal of the American Statistical Association*, vol. 83, no. 403, pp. 596–610, 1988.
- [30] E. Ventsel and T. Krauthammer, *Thin Plates and Shells: Theory, Analysis, and Applications*. New York: Marcel Dekker, 2001.
- [31] J. Yu and C. Lan, "System modeling and robust design of microaccelerometer using piezoelectric thin film," in *Proc. IEEE Int'l. Conf. Multisensor Fusion and Integration for Intelligent Systems*, 1999.

separation.

Vaninirappuputhenpurayil Gopalan Reju received the B.Tech. degree in electrical and electronics engineering from University of Kerala, India, in 1992, M.Tech. degree in electronics engineering from Cochin University of Science and Technology, India, in 1994 and Ph.D. from Nanyang Technological University, Singapore in 2010. He was with Cochin University of Science and Technology from 1995 to 2002 as a Lecturer. He is currently a post-doctoral research fellow at Infocomm Centre of Excellence, Nanyang Technological University. His research interests include array signal processing and blind source

Andy W. H. Khong is currently an Assistant Professor in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. Prior to that, he obtained his Ph.D. ('02-'05) from the Department of Electrical and Electronic Engineering, Imperial College London, after which he also served as a research associate ('05-'08) in the same department. He obtained his B.Eng. ('98-'02) from the Nanyang Technological University, Singapore. His postdoctoral research involved the development of signal processing algorithms for vehicle destination inference as well as the design and implementation of acoustic array and seismic fusion algorithms for perimeter security systems. He has also published works on adaptive filtering, tangible human-computer interfaces, acoustic blind channel identification and equalization for speech dereverberation as well as acoustic source localization.

Andy W. H. Khong is currently an Assistant Professor in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. Prior to that, he obtained his Ph.D. ('02-'05) from the Department of Electrical and Electronic Engineering, Imperial College London, after which he also served as a research associate ('05-'08) in the same department. He obtained his B.Eng. ('98-'02) from the Nanyang Technological University, Singapore. His postdoctoral research involved the development of signal processing algorithms for vehicle destination inference as well as the design and implementation of

Amir Bin Sulaiman received B.AppSc. degree in Information Technology from Royal Melbourne Institute of Technology, Australia, in 2005 and MSc degree in Signal Processing from Nanyang Technological University, Singapore, in 2010. He was with Nanyang Technological University from 2009 to 2011 as a Research Associate. He is currently working as Senior Software Engineer with Singapore Airlines. His research interests include acoustics, and array signal processing.