Ultrasonic based proximity detection for handsets

Pablo Peso Parada*, Rahim Saeidi*

* Cirrus Logic Inc, London, UK. Email: pablo.pesoparada@cirrus.com

Abstract—A novel approach for proximity detection on mobile handsets which does not require any additional transducers is presented. The method is based on transmitting a chirp and processing the received signal by applying Least Mean Square (LMS), where the desired signal is the transmitted chirp. The envelope of three signals (estimated filter taps, estimated output and error signal) are characterized with a set of 12 features which are used to classify a given frame into one of two classes: proximity active or proximity inactive. The classifier employed is based on Support Vector Machine (SVM) with linear kernel. The results show that over 13 minutes of recorded data, the accuracy achieved is 95.28% using 10-fold cross-validation. Furthermore, the feature importance analysis performed on the database indicates that the most relevant feature is based on the estimated filter taps.

I. INTRODUCTION

Handset designs have rapidly changed over the last years. The main trend on these designs is to increase the screen area without increasing the phone size reaching more than 91% of screen-to-body ratio. The main challenges to increase this ratio are the different sensors placed on the front of the phone such as front camera, earpiece, home button or infra-red sensor. This paper focuses on how the infra-red sensor can be removed. This infra-red (IR) sensor is used to detect proximity during calls and lock the screen of the phone when it is next to the ear thus avoiding the problem of user triggering undesired commands, e.g. hang-ups, when it is in contact with the face or ear. Therefore, proximity detection is active when there is an object, such as the head, close to the screen of the phone and inactive otherwise.

Proximity detection based on ultrasound can be used to remove the IR sensor since it employs the transducers already built-in, i.e. earpiece and microphone, and therefore enables the new generation of bezel-free handsets.

When using ultrasound signals on handsets, one of the main challenges is the efficiency of the earpiece at this high frequency range. The maximum frequency of ultrasound signals is constrained by the frequency response of transmitter (earpiece) and receiver (microphone). The earpiece and microphone are designed for audio/speech applications, thus providing an acceptable frequency response in the audible frequency range. However, in the ultrasound range, i.e. frequencies greater than 20 kHz [1], the performance is not guaranteed, and moreover it should be taken into consideration that the higher the frequency of the signal the higher the attenuation due to the propagation through the air [2]. The complexity of using ultrasound signals on smart phone devices is further increased by the variability of the frequency response from earpiece to earpiece in the frequency range above 20 kHz, the non-linearities of the earpiece in the ultrasound frequency range creating distortions in the audible frequency band and the headroom limitation to transmit the ultrasound pulse when the earpiece is driving acoustic signals such as speech in a call.

One common application of ultrasound signals in low power devices is the estimation of the Time Of Flight (TOF) which is defined as the time elapsed between the transmission and reception of the pulse. This measurement is directly related to proximity detection since this detection can be triggered when a minimum TOF is detected. Different methods are proposed in the literature to estimate the TOF. These can be divided into two groups depending on the emitted signal: continuous (harmonic, periodic) or discontinuous (pulse) signal [3]. The latter approach is more often used in the literature due to its lower power consumption. This group can as well be subdivided into two main approaches: measurement based on threshold detection and based on cross-correlation. Another review of a wide range of techniques to compute TOF is also provided in [4]. The authors classify these techniques in two types: time-domain methods and Fourier-domain phase-based methods. On one hand, in the former group, two sub-classes are considered: threshold detection and cross-correlation. The cross-correlation technique is usually referred as matched filter when the pulse is narrow band, e.g. tone, and referred as pulse compression when the pulse is wide band, e.g. chirps, because the cross-correlation produces a pulse whose width is inversely proportional to the bandwidth of the chirp. On the other hand, Fourier-domain phase-based methods can be implemented using a single-frequency, multi-frequency or chirp signals. The main idea is to derive the TOF from the phase of the received signal. The accuracy and TOF range depends on the number of frequencies and the bandwidth. In addition, hybrid methods based on both approaches are also reviewed.

In [5] several methods proposed previously in the literature are compared to a method based on an adaptive Finite Impulse Response (FIR) filtering using Least Mean Squares (LMS) error criterion. The coefficients of the filter are interpolated and employed to compute the delay by searching for local maxima. The results indicate that LMS approach performs better than cross-correlation when several echoes are received. In [6] multiple adaptive filtering methods based on LMS are analysed with regards to accuracy and computational cost.

In [1] Laguerre basis functions are employed to build a generalized correlation function. This approach does not need a priori knowledge about the pulse since it applies the Laguerre basis functions to the signal which decay exponentially to zero

(similar to the received pulses). It estimates the TOF as well as the characteristics of the ultrasound echo envelope employing a Kalman filter where the measurement of true Kalman state is derived from the Laguerre functions. These functions are also employed in [7] in order to classify echoes with different dumping factors. The K-nearest neighbour classifier is employed to classify 6 different dumping factors. In this case, the features are derived from the Laguerre coefficients after applying Linear Discriminant Analysis (LDA) and the Laguerre coefficients are obtained by expanding the echo envelope into an orthonormal set of Laguerre functions.

The TOF in [8] is computed employing a wavelet network. This network is a function constructed with a finite number of wavelets with different characteristics, i.e. weight, translation and dilation. These characteristics are trained with a set of examples in order to output the correlation function and thus obtain the TOF on unseen data.

In [9] a system based on Code Division Multiple Access (CDMA) is proposed to improve TOF estimation. The authors in [10] apply a pulse compression technique and the maximum of the cross-correlation is used to estimate the TOF. The novelty in this paper is on reducing the costs of this cross-correlation. This consists of a recursive cross-correlation operation of single-bit signals and an average filtering. The signals are converted to single-bit by a delta-sigma modulator.

Several methods have been proposed in the literature to perform proximity detection on a handset using existing transducer and ultrasounds. In [11] two LMS adaptive filters are used to model the acoustic impulse response from the earpiece to the microphone. One of these filters is constantly estimating the response, whereas the other filter is only active when proximity is inactive. The step size of the first filter is relatively higher than the second filter to account for rapid changes, i.e. handset movements. The proximity detection flag is triggered when the difference between the responses of each filter is higher than a predetermined threshold. The method proposed in [12] employs two microphones distant from each other such that one of them is covered by the head's user when proximity is active. The signals of both microphones are compared by means of the energy present in higher frequencies normalized by the total energy. The motivation for this comparison is that when the head is close to the handset it acts as a low pass filter. Therefore, the proximity flag is activated when difference between normalized energy of both microphone is above a given threshold. A method purely based on correlation between the transmitted and received signal is proposed in [13]. The proximity detector is triggered when the correlation lag is lower than an upper bound and its amplitude is between a predefined range of values.

Motivated by the performance achieved with the LMS approach shown in [5] and its low complexity, we propose in this work a proximity detection algorithm based on LMS. However, instead of exploiting the local maxima of the estimated filter coefficients, we propose to model the shape of this filter in addition to the error and estimated output signals for both classes, i.e. when the proximity detection is active and

when it is inactive. Then, a set of features from these three signals is extracted for a given frame and classified based on two values: the distance to the hyperplane that separates both classes, using the Support Vector Machine (SVM) approach, and the classification output in the previous frame.

The remainder of this paper is organized as follows. In Section II the proposed algorithm is presented, followed by the evaluation performance in Section III and the conclusion of this work in IV.

II. ALGORITHM

The method proposed in this section is based on modelling the reflections of a transmitted ultrasonic pulse. This modelling is performed using the LMS filter. Once these reflections are characterized, the proximity detection decision is derived from the characteristics of estimated LMS signals.

Figure 1 shows the complete block diagram of the proposed approach. Firstly, the ultrasound signal and potentially audio is sent to the earpiece ("Chirp + Audio"). A high pass filter is employed to remove low frequencies of the transmitted and received audio ("HPF") which may distort the estimated response due to the presence of noise in the transmission channel at these frequencies, e.g. background noise or speech. The received audio, captured in non-overlap blocks of 4096 samples, is then averaged over 2 blocks ("Blocks avg.") to reduce the noise and enhance the desired signal. This signal, together with the high pass filtered transmitted signal, is used in the LMS approach to compute the estimated received signal $\hat{\boldsymbol{y}}_j$, the error signal $\hat{\boldsymbol{e}}_j$ and the estimated impulse $\hat{\boldsymbol{h}}_j$ at j frame which are averaged over the last 4 estimates to reduce noisy estimations. The features described in Section II-B are computed and averaged for a period of 20 pulses in order to have a smoother behaviour of these features. Finally, the classification is based on the pre-trained SVM introduced in Section 2.3. This classification provides the binary decision regarding the proximity detection. Table I shows the parameters used in this work.

Parameter	Value
Sampling frequency	96 kHz
Pulse length	4096 samples
Pulse amplitude	-14 dBFs
Chirp initial and final freq.	25 kHz - 35 kHz
HPF cut frequency	20 kHz
Block avg. length	2
Estimate avg. length	4
Feature avg. length	20

TABLE I
MAIN PARAMETER OF THE ALGORITHM.

In this section, the LMS algorithm is firstly introduced. Then, a description of the features employed to characterize the impulse response is presented followed by the classifier used to detect proximity.

A. Least mean square adaptive filtering

The acoustic transfer function $\boldsymbol{h}(n)$ provides information about how the transmitted signal from the earpiece $x(n)$ is modified in order to receive the signal $y(n)$ at the microphone
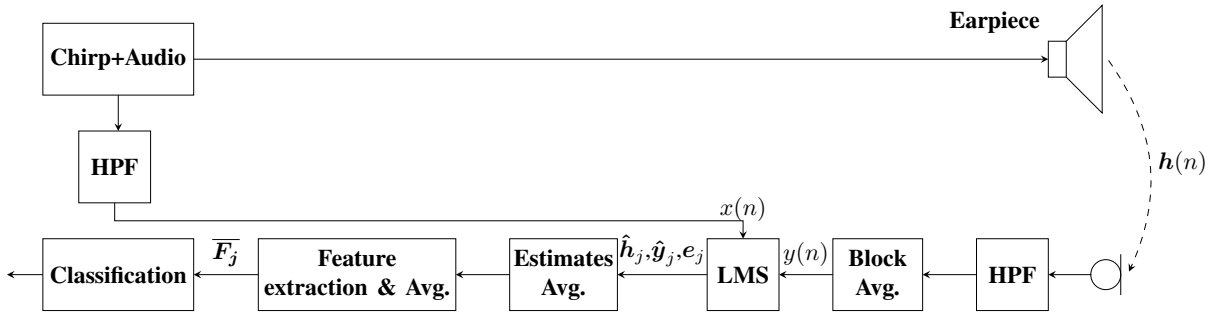
Fig. 1. Block diagram of the proposed method. HPF stands for high pass filter and LMS for least mean squares method.

position assuming no external noise sources are present, and therefore it gives an indication of the number and magnitude of the reflections occurred while transmitting the pulse $x(n)$. These reflections can potentially be used to estimate proximity.

In order to compute an estimate $\hat{\boldsymbol{h}}(n)$ of $\boldsymbol{h}(n)$ the adaptive filtering LMS [14] is employed in this work. The estimated received signal is given by

$$\hat{y}(n) = \sum_{m=0}^{M-1} \hat{h}_m(n)x(n-m) = \hat{\boldsymbol{h}}^T(n)\boldsymbol{x}(n) \quad (1)$$

where $\hat{\boldsymbol{h}}(n) = [\hat{h}_0(n), \hat{h}_1(n), \cdots, \hat{h}_{M-1}(n)]^T$ is an M tap FIR filter estimation of the acoustic feedback $\boldsymbol{h}(n)$ and $\boldsymbol{x}(n) = [x(n), x(n-1), \cdots, x(n-M+1)]^T$ comprises the last M transmitted samples.

The estimated transfer function $\hat{\boldsymbol{h}}(n)$ is computed using LMS by minimizing the mean squared error at sample n as follows

$$\arg\min_{\boldsymbol{h}(n)} E\{e^2(n)\} = \arg\min_{\boldsymbol{h}(n)} E\{[y(n)-\hat{y}(n)][y(n)-\hat{y}(n)]\}$$
$$(2)$$

In order to minimize the cost function in (2), an estimate of the exact gradient of the error is used, the so called stochastic gradient [14], obtaining the following update equation for $\hat{\boldsymbol{h}}(n)$

$$\hat{\boldsymbol{h}}(n+1) = \hat{\boldsymbol{h}}(n) + 2\mu\boldsymbol{x}(n)e(n) \quad (3)$$

Finding the learning rate μ that guarantees the convergence using this updating procedure becomes tricky since the weight updates are proportional to the magnitude of the input $x(n)$ and therefore sensitive to the input scaling. Instead, the normalized LMS (NLMS) introduces a normalization factor in the weights adaptation to avoid this problem

$$\hat{\boldsymbol{h}}(n+1) = \hat{\boldsymbol{h}}(n) + \mu\frac{\boldsymbol{x}(n)e(n)}{\boldsymbol{x}(n)^T\boldsymbol{x}(n) + \epsilon} \quad (4)$$

Figure 2 presents $\hat{\boldsymbol{h}}(n)$ with $M = 256$ for two cases: when the proximity is active and when it is inactive. As introduced previously, the magnitude of the first taps of the estimated response $\hat{\boldsymbol{h}}(n)$ is higher when the target is present compared to the situation where the target is not present which indicates that $\hat{\boldsymbol{h}}(n)$ contains relevant information to detect the proximity state. In next section, a set of features extracted from

this estimated transfer function in addition to the estimated received signal $\hat{y}(n)$ and the LMS error $e(n)$ is presented in order to determine the proximity state.
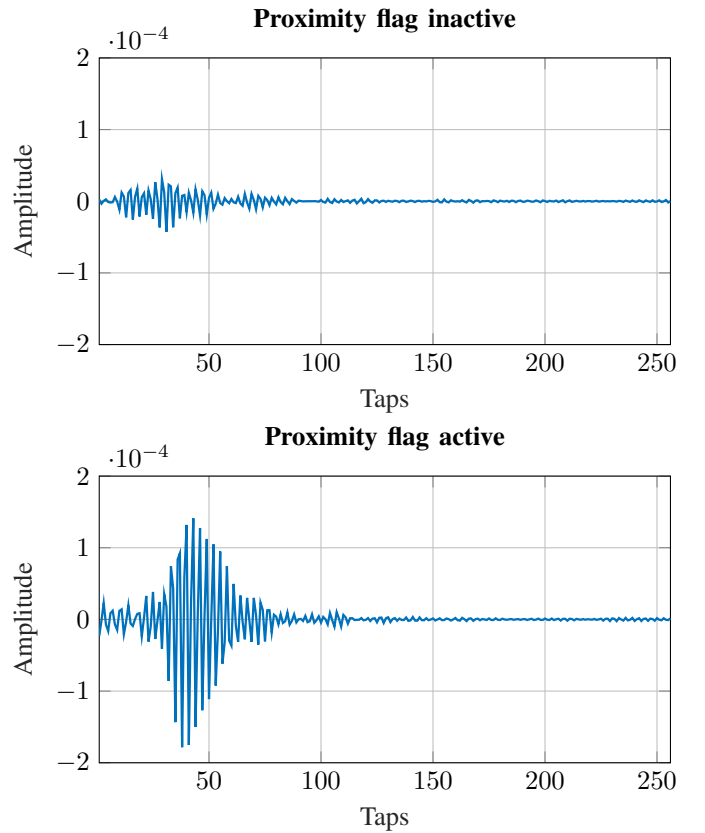

Fig. 2. Example of the $\hat{h}$ estimated for proximity detection inactive (top) and active (bottom).

B. Feature extraction

The feature vector refers to a vector $\boldsymbol{F}_j$ of different values $\{F_{1,j}, F_{2,j}, \cdots, F_{L,j}\}$ computed from the LMS output signals, namely $\hat{\boldsymbol{h}}_j$, $\hat{\boldsymbol{y}}_j$ and $\boldsymbol{e}_j$. This vector is calculated for the jth received pulse, thus the length of $\hat{\boldsymbol{y}}_j$ and $\boldsymbol{e}_j$ is 4096 samples whereas $\hat{\boldsymbol{h}}_j$ comprises 256 taps. The features proposed are based on the first 4th statistical moments, i.e. mean, variance, skewness and kurtosis. The mean describes the central tendency of a distribution, whereas the variance and skewness characterize the width and asymmetry of the

distribution around this mean and the kurtosis characterizes the peakedness of the distribution [15]. These statistical features are extracted from the envelopes of the signals considered as probability density functions following

- Mean

$$\mu = \sum_{i=0}^{N-1} Z_i \cdot P_{Z_i} \qquad (5)$$

- Variance

$$\sigma^2 = \sum_{i=0}^{N-1} (Z_i - \mu)^2 \cdot P_{Z_i} \qquad (6)$$

- Skewness

$$s = \sum_{i=0}^{N-1} \frac{(Z_i - \mu)^3}{\sigma^3} \cdot P_{Z_i} \qquad (7)$$

- Kurtosis

$$k = \sum_{i=0}^{N-1} \frac{(Z_i - \mu)^4}{\sigma^4} \cdot P_{Z_i} \qquad (8)$$

where Z_i is the ith sample position of a total of N samples and P_{Z_i} is the probability density function (pdf) at Z_i. These pdfs are defined as the envelopes of the signals normalized such that $\sum_{i=0}^{N-1} P_{Z_i} = 1$, where the envelope is computed following the Teager-Kaiser method applied every sample on $\hat{h}$, and every 50 samples on $\hat{y}$ and e. This envelope can be calculated as $\chi(n)^2 - \chi(n-1) \cdot \chi(n+1)$ for the signal $\chi(n)$. This down sampling reduces the computational cost of estimating the moments of the envelope.

The length of the feature vector $\boldsymbol{F}_j$ is 12, i.e. $L = 12$, and the elements of $\boldsymbol{F}_j$ are computed as follows

- $\{F_{1,j}, F_{2,j}, F_{3,j}, F_{4,j}\}$: First 4th moments of the envelope distribution of the LMS estimated response $\hat{\boldsymbol{h}}_j$.
- $\{F_{5,j}, F_{6,j}, F_{7,j}, F_{8,j}\}$: First 4th moments of the envelope distribution of the LMS desired signal estimate $\hat{\boldsymbol{y}}_j$.
- $\{F_{9,j}, F_{10,j}, F_{11,j}, F_{12,j}\}$: First 4th moments of the envelope distribution of the LMS error $\boldsymbol{e}_j$.

A time moving average filter of order $Q-1$ is applied to this feature vector in order to increase the robustness against outliers. This filtering is performed by

$$\overline{\boldsymbol{F}_j} = \frac{1}{Q} \sum_{i=0}^{Q-1} \boldsymbol{F}_{j-i} \qquad (9)$$

C. Classification

Once the feature vector $\overline{\boldsymbol{F}}_j$ is computed, the proximity decision is based on a linear regression derived from SVM approach [16]. This method seeks for a hyperplane to separate classes, i.e. active vs. inactive, such that the margin m between classes is maximized as shown in the example of Figure 3. This hyperplane is defined as $\overline{\boldsymbol{F}}_j \cdot \boldsymbol{W} + b = 0$, where $\overline{\boldsymbol{F}}_j$ is the feature vector defined in Section II-B. The hyperplanes defined by the support vectors of each class, this is $\overline{\boldsymbol{F}}_j \cdot \boldsymbol{W} + b = 1$ and $\overline{\boldsymbol{F}}_j \cdot \boldsymbol{W} + b = -1$ for proximity active and inactive respectively, can potentially be used as the threshold to switch between one proximity state to another proximity state. This is similar to
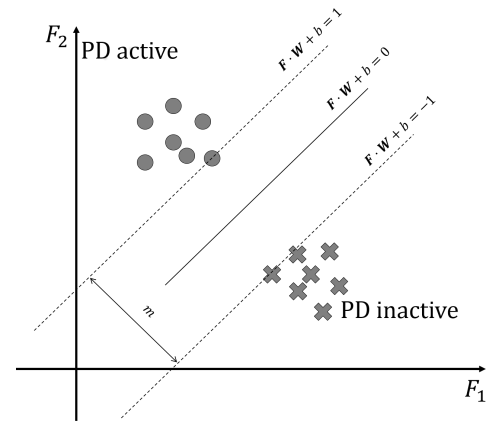

Fig. 3. SVM illustration.

hysteresis thresholding and it is referred in this work as dual-threshold.

III. EVALUATION

A. Experimental Setup

The entire database used in this work comprises 13 minutes of data recorded with a Samsung Galaxy S6 handset. In 52% of the frames the proximity detection is active, whereas the remaining of the frames it is inactive. The data was captured along different days following the next procedure: first, the user places the handset in the desired proximity detection state, i.e. active or inactive, and then records the data and labels it as the desired proximity detection state. Next recordings are carried out the same manner and at the end of the process all the recordings are concatenated. Therefore this database does not capture the transitions between proximity detection active and inactive and vice versa which can be extremely subjective and could lead to ambiguous decisions.

The evaluation metric used to measure the performance of the proposed algorithm is the accuracy, i.e. $1 - \Upsilon$ where Υ is the error rate, computed from a k-fold cross-validation applied to the entire database with $k = 10$. This procedure randomly splits the entire database into k partitions with the same number of observations. Then $k - 1$ partitions are taken for training and the remaining partition is used to compute the evaluation metric. This process is repeated k times using each of the available partitions for evaluation and the reported evaluation metric Υ is the average of all k metrics computed for each partition. Furthermore the error Υ is broken down into false positive and false negative to find the most common error type. In addition, a feature importance analysis is carried out to find those features that are more relevant for this classification task. The feature importance analysis approach employed in this paper is based on the weighted support machine vector of the model by the coefficients [17].

B. Results

1) Classification performance: A cross-validation test is performed on this database using 10-fold evaluation achieving on average an accuracy of 93.69% with a 2.46% of false

Rank	Feature index	Classification accuracy	Description
1	3	78.25	Skewness of $\hat{\boldsymbol{h}}_j$
2	1	82.82	Mean of $\hat{\boldsymbol{h}}_j$
3	11	89.71	Skewness of $\boldsymbol{e}_j$
4	10	92.14	Variance of $\boldsymbol{e}_j$
5	5	92.93	Mean of $\hat{\boldsymbol{y}}_j$
6	2	92.88	Variance of $\hat{\boldsymbol{h}}_j$
7	9	93.17	Mean of $\boldsymbol{e}_j$
8	6	93.12	Variance of $\hat{\boldsymbol{y}}_j$
9	8	93.33	Kurtosis of $\hat{\boldsymbol{y}}_j$
10	7	93.46	Skewness of $\hat{\boldsymbol{y}}_j$
11	4	93.68	Kurtosis of $\hat{\boldsymbol{h}}_j$
12	12	93.69	Kurtosis of $\boldsymbol{e}_j$

TABLE II

SUMMARY OF THE IMPORTANCE OF THE FEATURES IN THE SET $\boldsymbol{F}_j$. IT IS RANKED FROM THE MOST IMPORTANT TO THE LEAST IMPORTANT.

positive and 3.85% of false negative. This suggests that the detection of the inactive state is more challenging.

In order to reduce errors, a dual-threshold is used instead of using only a single-threshold, i.e. hyperplane derived by SVM. The dual-threshold approach employs the hyperplanes defined by the support vectors of each class to change the proximity state, therefore it keeps the previous state in case $-1 < \overline{\boldsymbol{F}}_j \cdot \boldsymbol{W} + b < 1$ and it changes the state to active or inactive in case $\overline{\boldsymbol{F}}_j \cdot \boldsymbol{W} + b \geq 1$ or $\overline{\boldsymbol{F}}_j \cdot \boldsymbol{W} + b \leq -1$ respectively. This dual-threshold approach increases the accuracy from 93.69% to 95.28% with a false positive rate of 1.28% and false negative rate of 3.44%.

A matched filter approach based on [13] was initially tested, however it did not provide meaningful results due to the reduced chirp amplitude employed to avoid audible earpiece distortions.

2) Feature importance: The feature importance ranking is shown in Table II in a decreasing order. This method uses the single-threshold to perform the classification. It shows in the second column the corresponding feature index, a value from 1 to 12. The third column includes the classification accuracy using this feature in addition to all previous more relevant features and in last column describes briefly each feature. This table indicates that the most important feature for this classification task is based on $\hat{\boldsymbol{h}}_j$. Figure 2 shows an example of how $\hat{\boldsymbol{h}}_j$ changes when the proximity flag is active and when it is inactive, and it suggests that the mean and skewness increases when proximity is inactive compared to the active state. Table II also indicates that the statistic less relevant for this classification task is the kurtosis.

IV. CONCLUSIONS

We have presented a novel approach to perform proximity detection on handsets. This approach is based on transmitting ultrasound chirps and processing the received reflections. The power of these transmitted signals is limited to avoiding any audible artefact in the earpiece. The accuracy achieved with this method is 93.69% on a 13 minutes long database. This accuracy is increased to 95.28% when dual-threshold is used. The feature importance analysis suggests that the most

important feature is the skewness of the estimated impulse response. Only using this feature for classification purposes provides 78.25% accuracy. The proposed approach is robust to modification of gain factors, such as earpiece or microphone gain, since it models the shape of the received signals rather than energies and does not create any audible artefact.

REFERENCES

[1] Leopoldo Angrisani, Aldo Baccigalupi, and Rosario Schiano Lo Moriello, *Ultrasonic-Based Distance Measurement Through Discrete Extended Kalman Filter*, INTECH Open Access Publisher, 2009.

[2] ISO 9613-1:1993, "Acoustics – Calculation of the absorption of sound by the atmosphere," Standard, International Organization for Standardization, Geneva, CH, 1993.

[3] Daniele Marioli, Claudio Narduzzi, Carlo Offelli, Dario Petri, Emilio Sardini, and Andrea Taroni, "Digital time-of-flight measurement for ultrasonic sensors," *IEEE Transactions on Instrumentation and Measurement*, vol. 41, no. 1, pp. 93–97, 1992.

[4] J. C. Jackson, R. Summan, G. I. Dobie, S. M. Whiteley, S. G. Pierce, and G. Hayward, "Time-of-flight measurement techniques for airborne ultrasonic ranging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 60, no. 2, pp. 343–355, February 2013.

[5] S. A. Pullano, A. S. Fiorillo, N. Vanello, and L. Landini, "Obstacle detection system based on low quality factor ultrasonic transducers for medical devices," in *IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, May 2016, pp. 1–4.

[6] H. C. So and P. C. Ching, "Comparative study of five lms-based adaptive time delay estimators," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 148, no. 1, pp. 9–15, Feb 2001.

[7] A. M. Sabatini, "A digital-signal-processing technique for ultrasonic signal modeling and classification," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 1, pp. 15–21, Feb 2001.

[8] D. Grimaldi, "Time-of-flight measurement in ultrasound transducer by means of wavelet networks," in *Proceedings of the Second IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, Sept 2003, pp. 365–370.

[9] M. Hazas and A. Hopper, "Broadband ultrasonic location systems for improved indoor positioning," *IEEE Transactions on Mobile Computing*, vol. 5, no. 5, pp. 536–547, May 2006.

[10] Shinnosuke Hirata, Minoru Kuribayashi Kurosawa, and Takashi Katagiri, "Accuracy and resolution of ultrasonic distance measurement with high-time-resolution cross-correlation function obtained by single-bit signal processing," *Acoustical Science and Technology*, vol. 30, no. 6, pp. 429–438, 2009.

[11] Ville Myllyla, "Acoustical proximity detection for mobile terminals and other devices," Apr. 1 2003, US Patent 6,542,436.

[12] Geert Langereis, "Apparatus and method for detecting usage profiles of mobile devices," July 1 2010, US Patent 12/829,117.

[13] Geert Langereis, Twan Van Lippen, Peter Dirksen, and Frank Pasveer, "Proximity sensor, in particular microphone for reception of sound signals in the human audible sound range, with ultrasonic proximity estimation," Mar. 19 2013, US Patent 8,401,513.

[14] Simon O Haykin, *Adaptive filter theory*, Pearson, 5th edition, 2013.

[15] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery, *Numerical Recipes*, Cambridge University Pr., 2007.

[16] Chih-Chung Chang and Chih-Jen Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011.

[17] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine Learning*, vol. 46, pp. 389–422, 2002.