2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN) | 978-1-6654-9624-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/IPSN54338.2022.00038

MiLTOn: Sensing Product Integrity *without* Opening the Box using Non-Invasive Acoustic Vibrometry

Akshay Gadre
Carnegie Mellon University
agadre@andrew.cmu.edu

Deepak Vasisht
Microsoft and UIUC
deepakv@illinois.edu

Nikunj Raghuvanshi
Microsoft
nikunjr@microsoft.com

Bodhi Priyantha
Microsoft
bodhip@microsoft.com

Manikanta Kotaru
Microsoft
mkotaru@microsoft.com

Swarun Kumar
Carnegie Mellon University
swarun@cmu.edu

Ranveer Chandra
Microsoft
ranveer@microsoft.com

ABSTRACT

This paper asks: "Can we detect whether a fragile product, made of porcelain or glass is damaged as it travels along the supply chain, *without* opening its packaging?" We ask this question in the context of the multi-billion dollar global supply chain industry of fragile products that experience large overheads due to product returns.

This paper presents MiLTOn, a novel acoustic and mm-wave based solution for through-box non-invasive product integrity sensing that is sensitive to even minute sub-mm cracks in the object. MiLTOn is inspired by acoustic vibrometry used for instance to monitor cracks in railroads. Unlike traditional vibrometry, MiLTOn is unique in its ability to sense products non-invasively using an external transducer and microphone, neither of which are in direct physical contact of the object within the box. MiLTOn processes measurements from the microphone to design a robust and environment-independent product signature that can be used to sense presence of product defects. Our extensive evaluation on a large number of fragile products of diverse materials demonstrates 97% accuracy in identifying product damage.

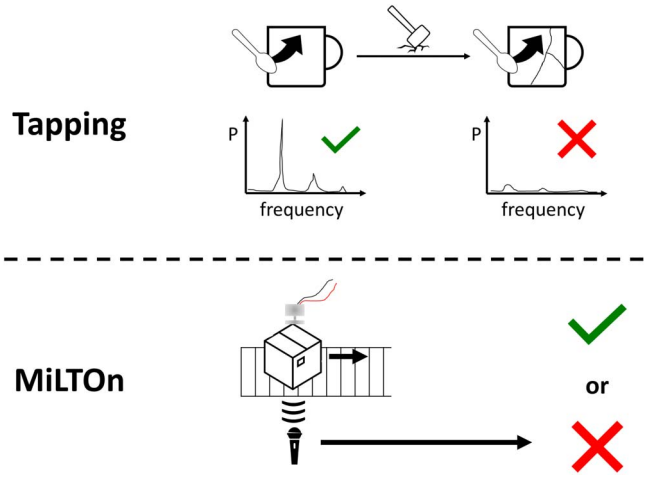


Figure 1: MiLTOn's design principle is to simulate the act of "tapping" a porcelain/glass object and retrieving its acoustic response to test for damage, but non-invasively through its packaging without opening the box.

1 INTRODUCTION

Quality means doing it right when no one is looking
– Henry Ford

Recent growth in the supply chain industry has seen a rapid rise in the number of packages shipped across the world. Just last year, 11.9 billion packages were shipped world wide with a significant increase expected this year. While this increase in supply chain has enabled sellers to cater to a global customer base, it comes with the major problem of returns. A significant number of these packages were returned due to damage to the product such as cracks, breakages, etc. This damage often occurs across the supply chain as the product changes several hands in its journey from seller to consumer. Yet it is currently challenging to pinpoint the entity along the supply chain who must (fairly) bear the cost of the return. This problem is particularly acute for products that are inherently fragile and therefore prone to breakage such as porcelain or glass products (cups, bowls, plates, etc.) – a $40 billion industry globally [4]. This paper is a university-industry collaboration that seeks to address the challenging practical problem of product integrity sensing for this important segment of the global supply chain.

Specifically, this paper asks: "How do we cost-effectively detect if a porcelain or glass product (say a cup) has been damaged (cracked, broken, etc.) inside a package — without opening it?" Note that unlike manufacturing processes, where quality control solutions intend to identify variations compared to an ideal model object, we only seek to identify defects that get introduced in the supply chain to each object individually. Thus, we seek an automated approach to doing so, as manually opening and inspecting a large number of packages at various checkpoints would prove cumbersome, costly, and might in fact add to the risk of damage. Thus, it is imperative for a technological through-box product sensing solution to be cheap, automated and accurate, where system cost justifies the savings from avoided product returns. While there has been rich prior work on through-wall sensing in varied contexts using RF signals such as X-rays[21, 44], mm-wave [24], Wi-Fi [7], ultra-wide band [41],

978-1-6654-9624-7/22/$31.00 ©2022 IEEE
DOI 10.1109/IPSN54338.2022.00038

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.

none of these solutions can cheaply measure internal shears or cracks that are common forms of damage in fragile products.

Our approach to sense product integrity (Fig. 1) is inspired by a common experience: Imagine holding a wine glass and tapping it, – you would hear a distinctive sound. However, this sound changes if the glass was damaged in even subtle ways, such as cracks or broken parts. In fact, even repairing the broken glass does not restore the original distinctive tone. Technically, tapping introduces a mechanical impulse that traverses and reverberates across the object, resulting in a frequency-dependent resonant response unique to its shape, topology, and materials. This approach of studying object properties via mechanical frequency response is called Acoustic Vibrometry. Indeed, acoustic vibrometry has been used to identify rotten fruit [22, 39] and cracked objects [1] to broken railroads [12].

However, these approaches require expensive, carefully-calibrated equipment and the ability to attach isolated physical objects to contact probes. Our use case is very different. We imagine a future where boxes transiting on a conveyor belt in a warehouse are sensed via a cheap, self-calibrating setup made of commodity components that could be deployed at scale. We make the observation that in our application, it suffices to detect only the *presence* of a defect, rather than its exact nature or extent. In other words, we seek to accurately flag boxes that contain potentially damaged products for manual review by an executive. Our focus is therefore to reduce both false positives and negatives of such damage alerts, despite variations in sensor placement and individual transducer properties along the supply chain.

This paper presents MiLTOn[1], an acoustic sensing system that can robustly detect breakage for a variety of objects. Our system consists of two key co-designed components.

Sensing platform: Our acoustic sensing setup uses a contact transducer that conducts an audible chirp to the object by turning the entire box into a sound source. The aggregate sound from the composite box-object system is picked up by a nearby microphone. Unfortunately, this acoustic response of the box-object system is tightly coupled. Thus, our design makes physically-motivated choices in constructing a robust sensing setup to efficiently retrieve the object response that does not require rigorous transducer-dependent calibration every time the system is used (see Sec. 4). Our design is assisted by a mmWave radar and camera based system to rapidly identify where the product is located within the box, to inform the acoustic transducer when it is to be activated.

Damage detection: We leverage the above carefully constructed sensing setup to retrieve the isolated acoustic response of the object, which we call an acoustic product signature, shortly after quality control processes when the product has freshly been manufactured. We deem these responses as "good" responses representing no damage. However, designing a data-driven classifier typically also requires representative "bad" samples. This is fundamentally impossible in a realistic supply chain, as we want the classifier to detect damaged object without knowing what a damaged object's response looks like – a classic chicken and egg conundrum. We solve this problem by first designing a distance metric for the signatures that is robust to ambient noise. With this metric, we show that a straightforward clustering algorithm already yields a sample-efficient anomaly detector. We find it sufficient to train on just a few (5-10) instances of the object during packaging to achieve high accuracy in identifying anomalies. Further, the system may be boot-strapped: when we mis-classify an object as broken, the box gets opened and the new observation can be used to update clustering for that object based on its shipping label. Sample-efficiency and boot-strapping together make the technique quite attractive for an industrial setting: pre-training is not required, and the system learns quickly. Finally, we guide our users based on industry-driven cost-benefit analysis to determine the threshold in our unary classifier to maximize the profitability of MiLTOn.

Limitations: We highlight a few important limitations of MiLTOn: (1) First, our system works mainly for objects which have distinct resonant behavior (porcelain, metallic, wooden, glass). (2) Second, our system will fail if the object does not vibrate due to a highly absorbant enclosure or undetected movement of the object inside the package. (3) We acknowledge that no quality control system or its evaluation is fool-proof, there will always remain some forms of damage that are unforeseen. We elaborate in Sec. 11, how these limitations do not forestall MiLTOn's utility for the supply chain.

Evaluation: We evaluate our system using the Adafruit Large Surface 5W transducer for 50 porcelain and glass products encased in typical closed cardboard boxes that they were shipped in. We simulate the breakage by emulating the various levels of breakages that a supply chain company is concerned with (a) minor breakages (e.g. handles of a cup) (b) repairable breakages (breakages that can be repaired at a low cost) (c) non-repairable breakages (crushed or other damage). Our evaluation demonstrates:

- 96.2% accuracy in detecting damaged porcelain cups based on an extensive case study. Optimizing for cost-savings provides a true positive accuracy of 97.3%.
- 97.92% accuracy in identifying anomalies in objects of different materials, shapes, sizes and across packaging materials.
- An overall reduction in loss due to returns by 56.7% for a typical supply chain company.

Contributions: Our main contributions include:

- A mechanism to make non-invasive acoustic vibrometry practical for identifying product integrity using a commodity contact transducer and a multi-modal setup to generate robust acoustic signature.
- Extensive evaluation with levels of product breakages that occur in different fragile materials.
- A holistic cost-benefit analysis of our system in production considering real supply chain company priorities.

Video demo : https://youtu.be/Iawpd2ujZ2E

2 RELATED WORK

2.1 Acoustic Imperfection Sensing

There has been much work done on leveraging the acoustic behaviour of objects to verify their integrity. Prior work [13, 19] have leveraged it to identify whether metallic objects or building materials are cracked or have cavities while others [8, 15] have leveraged them to identify the firmness of fruit or vegetables. Some prior work has also used acoustics in seismology [34], for sensing gear

[1]It is rumoured Schröedinger owned a cat with this name

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.

tooth breakage [40] and tool breakage [23]. Yet, much of this work relies on a carefully calibrated transmitter in direct contact or close proximity with the object within a tightly controlled setting.

There has also been a lot of work done on identifying defects in large scale systems such as railroad [16], conveyor belts [27], industrial equipment [28] using acoustic vibrometry [37]. While all of these solutions identify unique challenges pertaining to each system, they are able to connect their receivers to the object directly or leverage intrinsic vibration in systems such as conveyor belts. In contrast, MiLTOn enables us to inspect the integrity of static products within a box passively without requiring special probes which may in turn affect the packaging or the object.

2.2 Acoustic Imaging Systems

There is rich prior work on using ultrasound or audible acoustic signals for imaging objects [29, 30], localization [26], gesture sensing [17, 25] as well as novel communication systems [11, 43] that enable them. Acoustics is also used in body-sensing applications such as ultra-sound [38, 46, 47] or various health monitoring applications such as breath [10, 35, 48] and heart rate [10, 45] sensing. Such systems benefit from sounds emitted by the body or rely on special gels or liquids to conduct acoustic energy into the body [20]. Unlike this rich prior work, MiLTOn targets the unique problem of non-invasively sensing the integrity of products within a box, without an existing audio source inside.

2.3 Product Testing and Quality Control

A wide-range of solutions are available for product testing and quality control of manufactured goods. Prior work has explored X-ray imaging [21, 44], mm-wave [36, 49], terahertz imaging [50] and varied RF sensing systems [7, 18] for sensing through-obstructions, including in the product sensing context. However, X-ray imaging poses safety concerns and traditional camera imaging [33] for sensing product integrity requires a line-of-sight view of the object. Further, even mm-wave frequencies natively struggle to resolve sub-mm cracks or shears of ceramic or glass [36] due to lack of spatial resolution. Further, internal movement of padding material and varying electromagnetic reflectivity of outside padding (tape and paper stickers reflect mmWave more than cardboard) makes it difficult to isolate the behavior of the object from the packaging and surroundings. MiLTOn addresses this challenge by leveraging the acoustic resonance of products to isolate objects behavior, operates through-box and identifies fine cracks or damage.

3 MILTON - OVERVIEW

Fig. 2 depicts the journey of MiLTOn, when a package arrives at a warehouse or processing hub for either delivery to the customer or on its way to another warehouse. MiLTOn uses a mmWave + camera multimodal sensing system to scan the box barcode and detect it's orientation on the conveyor belt. Since cardboard is also relatively transparent to mmWave frequencies (77-81GHz), we also leverage the mmWave response to get the rough position of the object inside the box. Upon detecting the location of the object inside the box, an acoustic transducer is placed against the box at a point that is closest relative to the product. A nearby microphone then captures the acoustic response and computes a product signature, which is

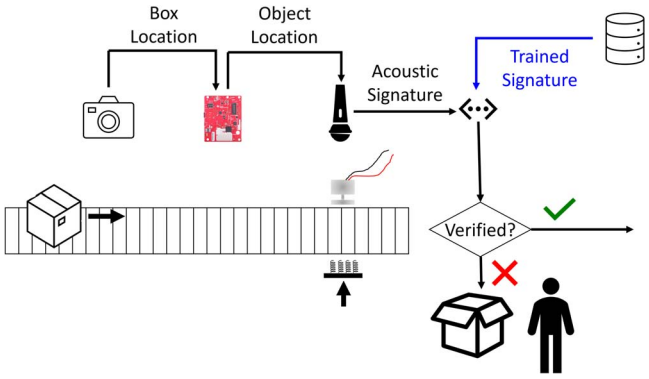


Figure 2: MiLTOn enables real-time integrity check for packages and boxes.

then compared to the prior signature of the same product when it was undamaged. Based on this signature check, the box is either sent to an executive for further inspection or passed on to the next conventional stage in the supply chain. MiLTOn is applied at both the ingress and egress to identify if and where product damage took place in the supply chain cheaply.

In the rest of this paper, we describe the challenges and opportunities in making the above system practical:

(1) Non-invasive through-box acoustic sensing platform: A key challenge in MiLTOn's design is retrieving significant representative acoustic responses through a non-invasive system – i.e., without physical contact with the object in consideration. First, we design a acoustic-transducer based platform that maximizes energy transfer to the object preserving the ability to retrieve its response in the box-object composite response. We further address how our platform can be extended to support multiple transducers simultaneously. We process the received acoustic box-object response at the microphone in the frequency domain to characterize the object of interest. Sec. 4 presents our approach.

(2) Damage Detection: We build an object-specific acoustic signature that is robust and repeatable, regardless of object location within the box or ambient noise. We describe how this signature can be validated against a prior signature (e.g. collected during packaging) based on a signature validation threshold. We further show how our system can learn from failures and adapt this signature if validation reports incorrect results reducing operational costs and improving accuracy. Sec. 5 details our approach.

(3) Cost-Benefit Analysis: An important aspect of any practical quality-control system is whether the benefits of the system outweigh its cost. To this end, we leverage our close industry collaborators to present a case-study for cost-benefit analysis that informs how our system's signature validation threshold that dictates the anticipated false positives vs. false negatives can be tuned. (Sec. 6)

4 ACOUSTIC SENSING PLATFORM

In this section, we describe how to design a platform that induces mechanical wave propagation through the object within a package and receive its acoustic response over-the-air. We also detail how to guide this system using mm-wave and camera imaging to maximize

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.

the transfer of energy from outside of the package to the object within the package to enhance its internal resonance.

4.1 Vibrating the object through-box

To retrieve an acoustic response from the object, we must first ensure that mechanical vibrations from the transducer reach the object. Unlike our anecdotal example of tapping a wine glass, physical contact with the object is not possible given the object is sealed within packaging. We therefore need a mechanism to induce mechanical vibrations in the object through the box.

The need for a contact transducer: From a practical standpoint, an over-the-air setup with a speaker and microphone would be easier to deploy. However, using a speaker to excite the object and box over the air does not work well in practice. The physical reason is that solid objects have an acoustic impedance thousands of times larger than air, and the transmission coefficient is proportional to the air-to-solid impedance ratio. Thus, most of the acoustic energy from the speaker is scattered back by the box. As a result, the direct speaker-to-microphone path can be 1000× stronger than the object response severely reducing the signal-to-noise ratio (SNR) of the received signal, even in low noise conditions.

To overcome this limitation, we draw inspiration from acoustic vibrometry to sense cracks in railroad tracks[12]. A transducer and a time-synchronized acoustic probe are both directly attached to the railroad tracks to vibrate the tracks and extract the resulting acoustic response. We similarly employ a contact transducer attached to the side of the box. However, the received response is complex combination of the responses of the box and the object. Hence, we must treat the box-and-object in combination as a composite system. Our results show that this difficulty does not obviate our goal of breakage detection.

Unlike vibrometry, using a (non-contact) microphone suffices for sensing in our scenario, for the same physical reason as above: the solid-to-air impedance ratio is quite favorable. In effect, the transducer-object-box system is an efficient radiator and a closely placed microphone readily obtains usable signal-to-noise ratio. However, compared to traditional vibrometry outlined above, we obtain a much less "cleaner" signal: there is attenuation from the box, interference from the box response, and additive ambient noise that the microphone picks up. As we will discuss in later sections, it is possible to extract usable breakage information from this composite signal. However, success there depends crucially on the sensor placement and physical setup, which we discuss next.

4.2 Acoustic model

Our sensing setup is carefully designed so that its physics is well-modeled by a feed-forward chain of linear filters as follows:

$$p(t) = s(t) * s_{eq}(t) * h(t) * r(t) * r_{eq}(t), \quad (1)$$

where $p(t)$ is the signal received at the microphone, s is a wideband chirp signal input to the transducer, s_{eq} is the unknown transducer response, h is the response of interest, r the through-the-air response from object (or box) to the microphone and finally, r_{eq} is the microphone's unknown response. Because of the feed-forward structure, one may still estimate the response, h, despite unknown transducer responses by doing a reference measurement without the box: $p_{ref}(t) = s(t) * s_{eq}(t) * r(t) * r_{eq}(t), \quad (2)$

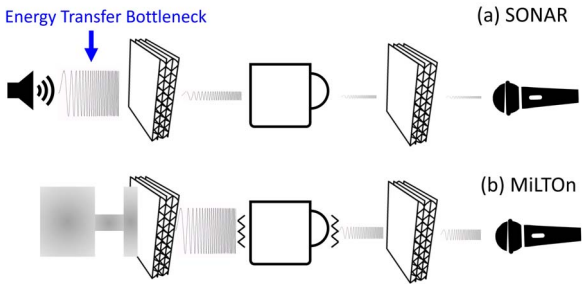


Figure 3: Comparing two systems: (1) A naïve acoustic SONAR and (2) MiLTOn that effectively transforms the box into a speaker.

followed by a deconvolution in frequency domain (where we use capitals to denote Fourier-transformed quantities):

$$H(f) = P(f)/P_{ref}(f) \quad (3)$$

Thus, our sensing system becomes **self-calibrating**: (1) P_{ref} can be periodically updated when no boxes are present on the conveyor belt, and (2) we can use commodity transducers that often deviate from a flat frequency response (i.e. $S_{eq}(f) \neq 1, R_{eq}(f) \neq 1$) since these are deconvolved away. One could employ Weiner deconvolution with prior ambient noise estimates for increased robustness; in our experiments, direct deconvolution sufficed.

We made two key observations that ensure that our system may be well-approximated as linear by Eq. 1. First, the transducer cannot have the object's weight transferred to it, so its safest to attach it horizontally. We observed that loading the transducer even slightly with the object's weight will result in strong feedback breaking the feed-forward model. In such a situation, one would need to do a fully coupled characterization of the composite transducer-object system. This is clearly infeasible. Second, the transducer has a flat plate that attaches and drives the object - this should ideally be small in comparison to the size of the box. A large plate acts as a strong acoustic radiator, and direct over-the-air paths to the microphone introduce a significant additive term in Eqs. 1 and 2. This once more introduces significant complexity in robustly solving for the object response $H(f)$. These fundamental considerations also motivate our transducer placement, described next.

4.3 Positioning the Transducer

For the moment, assume that we have the precise location of the box as well as the location of the object within it at high accuracy (we describe our approach to achieve this in Sec. 4.4). We would then be required to identify the exact location along the known box location for maximizing the energy transfer. Fortunately, as long as the approximate orientation of the object and placement within the package is known while being packed, the optimal location to attach the transducer can be manually inferred through an exhaustive search beforehand. Given that during mass manufacturing, all objects are traditionally packed in the exact same manner, this location could be pre-modeled and pre-defined. In our evaluation case-study focused on coffee mugs, we identify that the optimal location to attach the transducer is always horizontal and as near to the top of the cup as possible.

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.

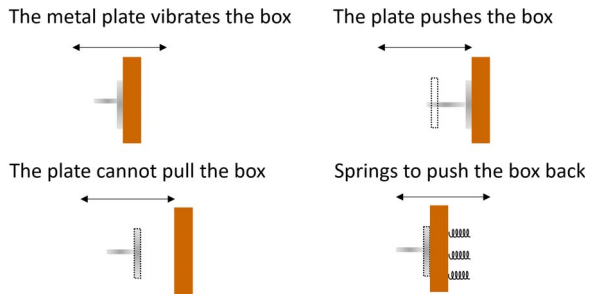


Figure 4: MiLTOn uses springs to hold the box in place while the transducer is attached to the box's side.

Ensuring contact: Unfortunately, attaching a transducer horizontally to the box leads to another challenge. Given that the transducer is simply a vibrating plate, it will push the box outwards at first but will not be able to pull the box when returning back inwards. In other words, rather than vibrating the box horizontally, a transducer would merely push it away. Fig. 4 demonstrates this problem. Further, the obvious remedy to use a rigid grabber to hold the object in place against the transducer will hinder box's vibrations and introduce non-linearities in Eq. 1. Thus, we need a solution that keeps the object in place but remains elastic enough to enable vibration of the box to the maximum possible amplitude.

We hence leverage multiple helical springs to push the object back against the transducer (see Fig. 4). In a warehouse setting, an automated arm can be used to attach springs that push the box against the transducer to achieve the same behavior. One may however be concerned about how this would impact the resonance of the box (as helical springs with attached mass are known to have resonance frequencies). However, since the amount of energy transferred from the box to the spring setup is minuscule (and can be distributed across multiple springs), this does not affect the acoustic vibration of the box. In fact, a careful analysis of springs of this nature [31] have shown even in cases where these resonance frequencies exist, they are outside the acoustic frequencies that we operate in (>100 Hz and <20 kHz).

Extending to multiple transducers for multiple objects: When there is a single object inside a box, it is indeed possible to identify the ideal location of the transducer using the below mentioned camera-mmWave hybrid setup. However, in many cases, there may be multiple objects inside the box out of which a single or a few of the objects may be damaged. Our initial results in Sec.10.4 shows that when the damaged cup is far from the transducer, the accuracy of MiLTOn does reduce. While it does improve as the number of damaged objects increase, it is necessary for a solution when the damaged cup location may not be known independently. In such a situation, we take multiple measurements across locations along the edge of the package as the package travels across the conveyor belt. We then assume each of these measurements as independent signatures and flag a damaged product if any of the signature changes beyond a certain threshold.

Our evaluation verified the ability of our platform to detect the resonant peaks of the object against the reference of tapping the object and analyzing its sound. (See Fig. 8)

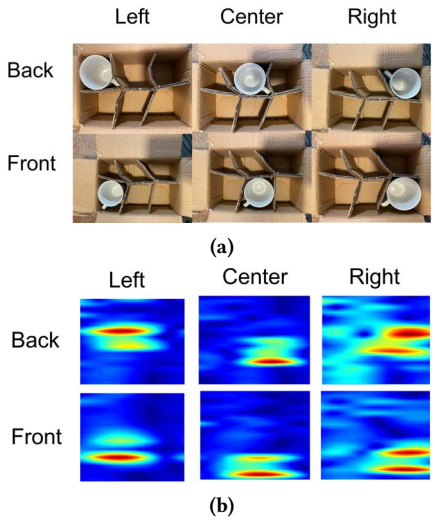


Figure 5: mmWave radar beamforming enables MiLTOn to identify the approximate location of the product – here a porcelain coffee mug. (Note: box is closed during evaluation)

4.4 mmWave+Camera for Box Tracking

Two important issues remain in positioning the transducer correctly relative to the object-in-box system. First, how do we know precisely where on the conveyor belt the box is located? Second, how do we know if, and by how much, the product has moved within the box itself relative to its location at the time of manufacture? Answering both of these questions is critical to place the transducer in its optimal location for energy transfer.

Camera-based Box Positioning: MiLTOn addresses the first challenge by using a commodity camera to sense the box's location at cm-accuracy. We use a state-of-the-art SIFT-based object recognition and image segmentation algorithm [6] that identifies the spatial coordinates and bounds of the box along the conveyor belt. The camera sensing system is placed ahead of the mm-wave radar platform along the conveyor belt so that the latter can be deployed next to scan the contents of the box.

mmWave-based Object Positioning: We next activate our mmWave radar platform to collect I/Q samples relative as the box passes by to compute an RF through-box image. Note that cardboard is known to be transparent to mmWave signals [32]. For mmWave processing, we use the standard Bartlett algorithm [36] across the 4 RX antenna on the mmWave radar that results hotspots that match the spatial bounds of the product within the box. We use the phase and amplitude of each distance bin response in the mmWave radar antenna to create a heatmap of responses across locations and angle-of-arrivals (see Fig. 5 for a few representative examples and Sec. 8.2 for detailed results). We then compute the optimal location to place the transducer by matching the mm-wave RF image observed for the box with a template RF image.

5 DAMAGE DETECTION

In this section, we show how to design a robust signature from the received acoustic signals, that characterizes the state of the

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.

object. We capture an *initial product signature* during the packaging process (Sec. 5.1). We then validate the measured signatures across the supply chain to the initial product signature (Sec. 5.2). Finally, we only open the box for inspection when the validation process flags a potentially broken product. MiLTOn is also designed to adapt to incorrectly flagged packages (Sec. 5.3) by reinforcing the product signature to avoid the same error from recurring.

5.1 Computing the Signature

Our process for computing the signature relies on one or more receptions of signals from the acoustic transducer. We design our system to allow for multiple receptions to compute the initial signature shortly after initial packaging, while subsequent signatures could be computed with a single reception.

What should the transducer transmit? A näive solution would be to use the mechanical transducer to send an acoustic impulse to directly measure the impulse response. Unfortunately, it is impossible to transmit an instantaneous impulse from a band limited transducer. Further, the transducer is an acoustically unequalized object, i.e. it has different gains across frequencies. Thus, we need to design a waveform to emulate the behavior of our mechanical impulse within the capabilities of the transducer.

MiLTOn uses a wide-band chirp which sweeps all frequencies between 100 Hz and 20 kHz over a time duration of 5 s. Our rationale for using chirps is two-fold: (1) First, it naturally spans the range of frequencies desired to allow for rich signatures for package-sized objects; (2) It circumvents the band-limited nature of the transducer since the instantaneous bandwidth of a chirp is extremely low.

Designing the Signature: At a high level, our signature is a processed acoustic channel that captures the mean and variance of the acoustic response of the composite box-object system in the frequency-domain. Rather than storing all frequencies exhaustively, we only select those whose measurements are robust and repeatable across measurements and are likely to represent the frequency response of the product's material. We compute the response of the box-plus-object *system* rather than isolating the object itself for two reasons: (1) First, it is rather challenging to isolate the object's sole response, given that the acoustic wave propagates in complex ways across both the box and the object; (2) Second, regardless, any change in the product will feature in the response of the box-plus-product system – which is our objective in any case.

Mathematically, we compute the acoustic signature as follows: (1) First, we use Eq. 3 to compute the acoustic response in the frequency domain, $H_i(f)$, where $i = 1, \ldots, n$ for each of n receptions collected from n chirps; (2) Second, we compute the mean and standard deviation $\mu(f)$ and $\sigma(f)$ of each measurement; (3) Finally, we drop any measurements in the $(\mu(f), \sigma(f))$ tuple over 20 kHz or below 100 Hz to filter out frequencies that are either too large or too small to have the acoustic response featured. In practice, we may filter this bandwidth further, based on any prior information of the frequency response for a product.

Making the Signature Noise-resilient: To make our signature robust to noise, we perform a few additional pre-processing steps. We are particularly interested in two sources of error: (1) Ambient noise or narrowband interferers that perturb specific measurements of the frequency response; (2) Frequencies that are inherently poor in informativeness owing to high variance across signature measurements, often due to transducer imperfections.

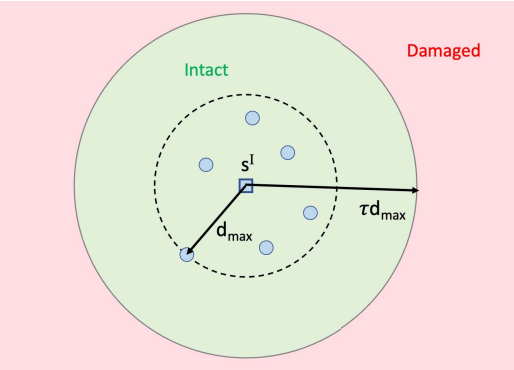


Figure 6: MiLTOn computes the threshold for the distance of the signature observed on the field with the initial product signature to assess damage, using all collected individual measurements during initial packaging.

Mathematically, we model these by performing the following two steps: (1) First, we measure an additional metric $c(f)$ – called the confidence-metric, that is a normalized signal-to-interference-plus-noise ratio (SINR) across measurements at a given frequency. (2) Second, we perform outlier rejection to remove frequencies where $\frac{\sigma(f)}{\mu(f)}$ exceeds a threshold to drop measurements in extremely noisy frequencies. At the end of this process, the values $(\mu(f), \sigma(f), c(f))$ measured across a subset of frequencies $f \in F$ represent the acoustic signature of the product. Note that even our cheap transducer and microphone setup only had sub-1dB variation across 90.8% of the frequencies within 100 Hz to 20 kHz.

5.2 Validating the signature

Having computed the acoustic signature, we now seek to design a mechanism to compare two signatures of the same product – one the initial signature $s^I = (\mu^I(f), \sigma^I(f), c^I(f))$ computed during packaging and another $s = (\mu(f), c(f))$ collected at a warehouse in the supply chain. We note that the latter signature lacks the standard deviation $\sigma(f)$, since we typically take only one measurement at a time in each warehouse facility per product. We now need a mechanism to accurately compare how dissimilar two signatures are relative to each other, i.e. a distance metric.

Defining a Distance Metric: We compare the two signatures by computing the L-2 norm of the difference in their μ values. We weight each difference in μ values by two factors: (1) $1/\sigma^I(f)$ to account for noise and interference in the initial product signature per frequency; (2) $c(f)$ and $c^I(f)$, the respective confidence values in the initial and newly observed product signatures. Mathematically, the distance is:

$$d(s^I, s) = \left\| c(f)c^I(f)\frac{\mu(f) - \mu^I(f)}{\sigma^I(f)} \right\|_2$$

Defining detection threshold: Next, based on the distance calculated above, we need to make a decision on whether this distance is sufficiently large to declare the product damaged or otherwise. In other words, we need to define an optimal threshold on the distance that accurately identifies if the product is damaged.

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.

To compute this threshold, we rely on the individual channel measurements H_i(f) collected shortly after the manufacture. Recall that each of these acoustic channel measurements were used collectively to form the initial product signature $s^I = (\mu(f), \sigma(f), c(f))$. However, we can also process these channels one at a time to form individual signatures $s_i = (\mu_i(f), c_i(f))$. When we observe the space of these individual signatures s_i in relation to the initial product signature s_I, (Fig. 6), it is clear that the individual signatures s_i would be spatially separated relative to s_I within some circle. This circle has a radius defined by the maximum distance across each individual signature s_i relative to the initial product signature s^I, that is: $d_{max} = \max_i d(s^I, s_i)$. In other words, the maximum distance circle shown in Fig. 6 provides a sense of how much signatures can deviate across many measurements for an undamaged product.

Assessing Damage: Let us assume that we encounter a signature s for a product at a warehouse. We declare that a product is damaged if $d(s^I, s) > \tau d_{max}$ and not damaged otherwise, where τ is a threshold parameter. The value τ is a parameter that offers error tolerance in our initial data from the first stop of the supply chain. The choice of the correct value of the threshold parameter τ directly impacts our false positives vs. false negatives and overall cost savings. The value of τ is scaled inversely based on the number of data points n collected at the factory i.e. the more the confidence in their signature, the stricter the threshold. We discuss how a supply chain company would choose this parameter in Sec. 6.

5.3 Learning from Mistakes

There are two types of mistakes MiLTOn can make in practical settings. First, consider false positives where MiLTOn fails to detect a damaged object inside a package. In this case, there is no value in learning from the mistake as all the associated costs have been incurred. However, in the case of false negatives, when we open a box and find an undamaged product, we can use this information to reinforce our signature.

Sample Augmentation: One would think that MiLTOn can only actively learn from undamaged boxes that were opened. However, by opening a box, we are not only getting the information that the box is undamaged now, but also for all previous stages in the supply chain (including in previous warehouses, hubs, etc.). This is valuable information because it immediately indicates to MiLTOn that the signature of this product as well as *all* previously recorded signatures of this product correspond to that of an undamaged product. This improves the number of samples available for MiLTOn reducing the effective validation cost of our system.

Bootstrapping: MiLTOn uses the above observation to include all these signatures into the set of initial individual signatures as shown in Fig. 6. It can then repeat the process described in Sec. 5.2 to update d_{max}, if needed, based on the distance of these newly added signatures as well as reduce τ to account for the number of newly added signatures. The net effect would be an expansion of the circle in Fig. 6 to accommodate these previously outlying measurements. We could even start with a new product with no measurements, and our system would trigger the box to be opened and update its estimates throughout the supply chain.

It is indeed true that as a package moves across warehouses without being checked, it will amortize a lot of unverified data (could be false positives) polluting the detection threshold. However, MiLTOn makes a conscious decision by allowing it. MiLTOn identifies that the real cost to the supply chain company is opening the box repeatedly (shown in Sec. 6) and thus purposefully makes it amicable to avoid opening the box while remaining robust to detect product damage to a sufficient degree.

6 COST-BENEFIT ANALYSIS

In this section, we address a question common to any industrial system: "Does the system result in cost-savings for a supply chain company to deploy?". We further assess how the threshold τ should be decided by a retailer to maximize profit for its supply chain.

6.1 A Simplified Economic Model

We set some broad design principles and a simplified economic model that influence our system's threshold. First, let's analyze the losses a retailer incurs for every returned item today. The first cost is the obvious cost of replacing and shipping the cup. However, there is a much more pertinent cost that is often overlooked – the reputation loss due to damaged product delivery. A survey by a major shipment company estimated that 50% of customers are less likely to buy from a retailer that shipped them damaged products[3]. Even though the reputation loss is intangible, it is so extensive that retailers use several steps to minimize it today – for example, cash-equivalent coupons as apology. Industry experts estimate that the real cost of replacing a damaged item is, on average, 17 times the cost of shipping it[2]. We call the sum of the above losses as the total cost of replacing a damaged item – the *damage cost*. A retailer effectively faces this cost if a customer receives a damaged object.

In contrast, our design adds an additional cost of enabling the system - the *validation cost* of labor for opening flagged packages to verify the damage. If the item is indeed damaged, the retailer replaces the cup and incurs additional replacement cost of replacing it. Note that *replacement cost* is lower than cost above because the retailer no longer needs to compensate the customer and early detection reduces the shipping cost as well.

However, when MiLTOn incorrectly tags a cup as damaged, the retailer incurs an additional *validation cost* over the *damage cost*, it faces for losing customer confidence. As an example, consider a scenario where MiLTOn correctly tags a cup as damaged, the net cost savings for a retailer or logistics provider in that case is: *damage cost - replacement cost - validation cost* (always positive).

The final cost involved is the fixed cost of installing and maintaining our system. However, this is subsumed by validation cost for the most part since the latter is recurrent and often involves labor versus our system that is automated and relatively low cost. We further explore how this model corresponds to actual numbers in our case-study on cups in Sec. 9.5.

Cost of Latency: Another important factor that governs the profitability of a system like MiLTOn is the impact on the regular operation of the supply chain. In our case, it is the latency of the mechanical arm and time to send the acoustic signal on the order of seconds. At first glance, this seems high for any supply chain that operates on lower latency bounds. However, there are several other processes such as putting the package from storage on the conveyor belt, manually scanning packages as well as moving

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.

packages across warehouses that take significantly more time to perform. We find that we can reduce the latency burden of MiLTOn in a similar way that supply chain companies do for these processes – parallelism. Similar to how a supply chain company would hire multiple employees to scan packages, we would operate multiple replicas of MiLTOn. This would reduce the latency and operational costs of MiLTOn while increasing the capital expenses to deploy MiLTOn. We evaluate the impact of these expenses in Sec. 11.3.

6.2 Tuning the Threshold

It is in the best interest of the retailer to choose our threshold multiplier, τ, to optimize for the net cost savings. Around such an optimum, as we lower the value of τ, the number of false negatives increase and customers get more damaged items. On the other hand, if the value of τ is too high, the number of false positives increases, we incur more validation cost unnecessarily, but less damaged items get shipped to customers. Therefore, for each threshold τ, we empirically evaluate the anticipated false positives and false negatives. We then tune our threshold so that net-on-net, the cost reductions are maximized with our system compared to the prevalent system. We call this choice of threshold the *profit optimum*. We note that the profit-optimum may not always be the same as the canonical *accuracy optimum* where the false positive and false negative rates are weighed equally.

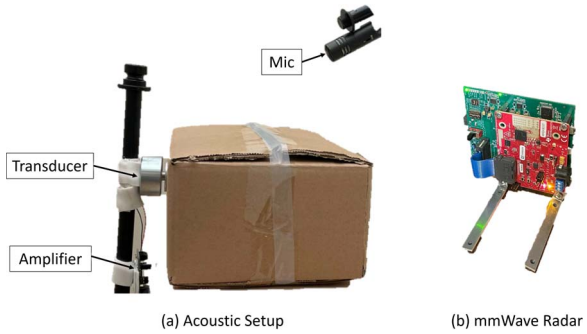
In Sec. 9.5, we make these economic trade-offs concrete in the context of a case study – a manufacturer of porcelain cups, through our simplified model developed in coordination with our industry collaborator. We study how our choice of threshold impacts and can ensure net-profitability from the deployment of our system.

7 IMPLEMENTATION AND EVALUATION

We implemented MiLTOn using Adafruit Large Surface Transducer with Wires - 4 Ohm 5 Watt being fed in acoustic signals via Adafruit Stereo 20W Class D Audio Amplifier - MAX9744 with 6dB amplification (Fig. 7(a)). At the receiver, we use a AKG P170 Small-diaphragm Condenser Microphone. We connect both of these via RME Babyface Pro FS to the computer . We use the TI AWR 2243 mmWave Radar (Fig. 7(b)) for identifying the location of the cup inside the box. Our code is built in MATLAB to perform real-time integrity check on the products inside the box.

For each evaluation, our acoustic system transmits a 5 second long chirp spanning bandwidth from 100 Hz to 24 KHz. Our receiver receives the signal at 48KHz sampling rate. We then use these signals to perform integrity check of the product and output a boolean result : **Positive** (No need to open the box) or **Negative** (Open the box to check integrity)

To evaluate our system, we consider various parameters such as materials, training data, and type of object. For materials, we use 4 cups/mugs of different materials (plastic, glass, wood and porcelain). For type of object, we use various objects of glass and ceramic that are typically shipped such as children toys and wine/drink glasses. Finally, for an in-depth analysis of the robustness of our system, we perform a detailed evaluation across 30 types of breakages on "10 Strawberry Street" Catering Mug Set. All evaluation is performed inside a closed box with single-wall E-flute corrugated cardboard walls. The same undamaged cup is broken when switching from



(a) Acoustic Setup (b) mmWave Radar

Figure 7: Experimental Testbed

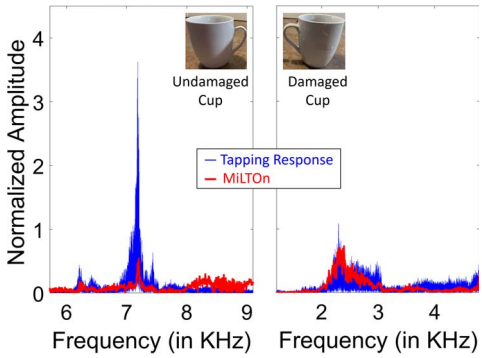


Figure 8: MiLTOn's ability to replicate the response as heard upon tapping the cup

positive to negative data collection. Note that we evaluate our system on objects which are small enough to resonate from a single transducer while operating in a cost range where more expensive and complicated systems such as X-ray would be infeasible. Detecting damage for objects that do not resonate readily under acoustic signals (e.g. objects made of cloth, Styrofoam, etc.) remains out of the scope for this paper.

Across our evaluation, we use four terms:

(1) **True Positive**: Not Damaged and Box not opened
(2) **False Positive**: Damaged and Box not opened
(3) **True Negative**: Damaged and Box opened
(4) **False Negative**: Not Damaged and Box opened

8 MICROBENCHMARKS

In this section, we evaluate the basic primitives behind MiLTOn.

8.1 Correlation with Tapping

One of the key intuitions that led to MiLTOn was the fact that we can replicate the effect of tapping the cup by using a transducer. To evaluate this, we use a spoon and collect a sound clip of tapping the cup. We also collect an acoustic measurement for verification. The cup is then broken into two clean parts using a hammer and another acoustic signature is collected. Finally, we attempt to repair the cup using a scotch tape and evaluate the acoustic behavior.

Result: Our evaluation demonstrates that the cup response even after repair looks significantly different from the original response. Further, as shown in Fig. 8, this difference in the responses can also be seen in the acoustic signatures with corresponding peaks at the

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.

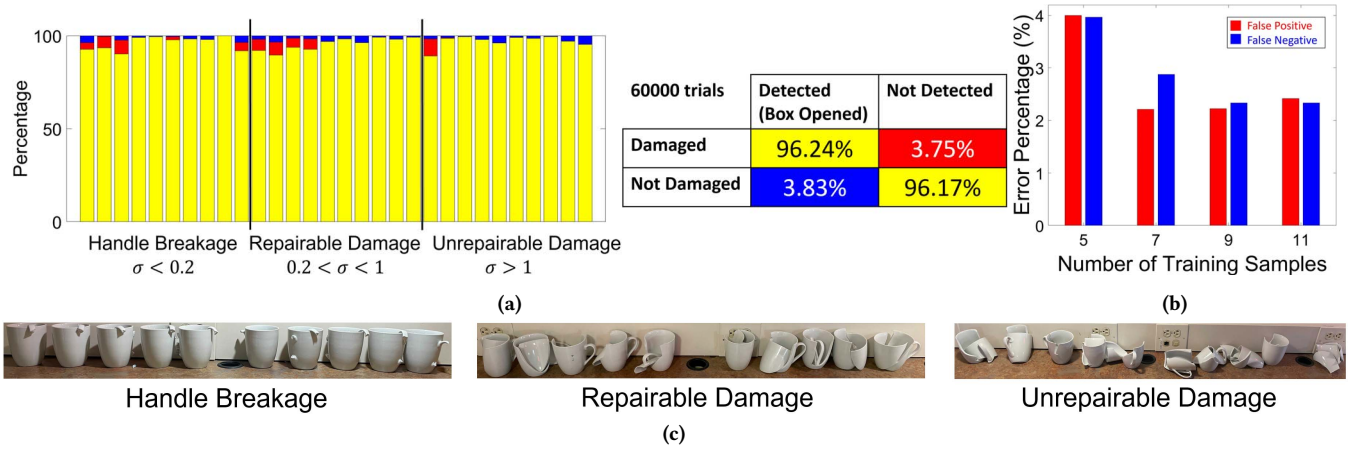


Figure 9: Porcelain Cups Case Study: (a) Accuracy across the evaluated breakages ; (b) Effect of Training Samples on MiLTOn performance; (c) Cracked cups after evaluation.

resonance modes. This shows the ability of MiLTOn to effectively measure acoustic behavior of objects and, in turn, their integrity.

8.2 mmWave Location Classification

We evaluate the ability of mmWave Radar to identify the location of the object reflection (as the location of the box is known). To evaluate this, we take the example of packaging typically seen in tableware industry, a box with 6 compartments tightly holding six articles (See Sec.4.4). We then place the cup in each of the compartments and record their radar responses. Using the Bartlett-based beamforming algorithm across the 4 antennae of the mmWave radar, we attempt to identify the location of the cup.

Result: Fig. 5 shows that the radar can distinguish signatures across compartments in a reasonably sized box. We achieve an accuracy of 4.75cm in identifying the location of the cup inside the box. However, since we know the bounding box across dimensions of where the box is located, we can get perfect accuracy by mapping the location to the closest compartment. Note that this may not work for a large number of extremely small objects (such as marbles) or boxes that reflect mmWave signals strongly (such as metal).

9 PORCELAIN CUPS : A CASE STUDY

In this section, we present our main results based on the evaluation across 60,000 trials for a single commodity object – a porcelain cup.

9.1 Motivation and evaluation

We chose porcelain cup as our object of study since it provides diversity across multiple axes. First, the cup can be broken in vastly different kind of ways ranging from a chip to breaking into powder. Second, a large number of cups are transported across the world every day. It is estimated that over a million cups are shipped around the world every month. Finally, they are really cheap, making financial parity of our system even more difficult to achieve.

Thus, to evaluate MiLTOn, we used 30 cups from "10 Strawberry Street". We then identified typical breakages seen in the actual supply chain for the same manufacturer across bad reviews[5]. We identified three typical types of damages seen in the review pictures and comments: (1) Handle Breakage (2) Repairable Damage (3) Unrepairable Damage. Since there are no well accepted universal standards for understanding the degree of damage in the supply chain, we borrow a metric typically used in measuring cracks on pavements to quantify these damages to the porcelain cup. The surface cracking metric (SCM)[9] is measured as $100\frac{\sum_i^N l_i w_i}{A}$ where l_i is the length of the crack, w_i is the width of the crack and A is the surface area of the pavement. Since the width of the cracks are extremely difficult to measure, we adapt a simplified version of the crack metric $\sigma = \frac{\sum_i^N l_i^2}{A}$. For our cups, the external surface area (A) is 110 sq. in. By measuring the size of the cracks, the ranges for the three categories are: (1) Handle Breakage : $\sigma < 0.2$; (2) Repairable Damage: $0.2 < \sigma < 1$; (3) Unrepairable Damage: $\sigma > 1$.

We collect 15 data points for each of the 30 cups using our system (rattling the box after every collected data) as undamaged data points. Then we use a hammer to break the cups into aforementioned 3 categories of damages. After performing this breakage, we collect 10 different acoustic samples for each of the broken cups. We, then, train our acoustic signature on 5 randomly chosen good data points that we collected earlier and evaluate the remaining 10 good and 10 bad data points for the system. We repeat this whole process 100 times to ensure robustness across training data points.

9.2 Accuracy

As shown in Fig. 9a, our evaluation shows that we achieve a 96.21% accuracy in identifying broken and unbroken cups. Note that for the purposes of this evaluation, the threshold is optimized for accuracy (unlike as mentioned in Sec. 6 to maximize profit). Only 3.75% of the damaged cups are not detected by our system which based on various observations cited in Sec. 6 corresponds to 0.22% of total cups reach the customer damaged (compared to 6% prior).

9.3 Noise Resilience

We evaluate the resilience of MiLTOn's performance in presence of ambient noise in Fig. 10. As seen in the figure, we can clearly see that even when the SNR of the signal is around 9 dB, accuracy of MiLTOn remains above 90%. Note that most of our experiments were conducted at a SNR of around 15.7 dB (quiet conference room).

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.

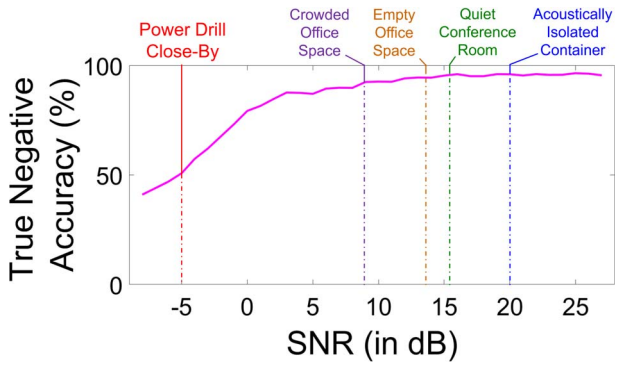


Figure 10: Noise Resilience of MiLTOn

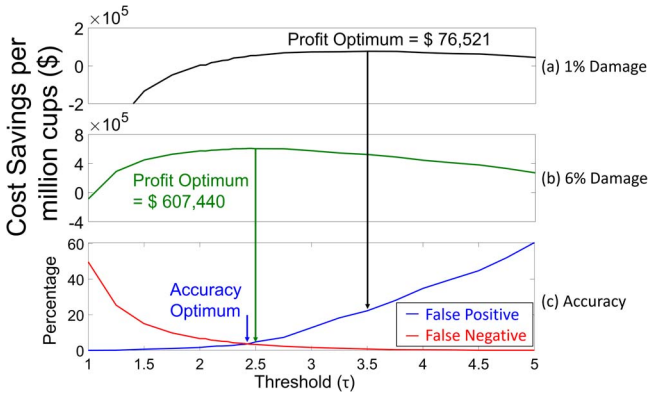


Figure 11: Cost-Benefit Analysis of MiLTOn

Further, we reiterate that MiLTOn is designed to be deployed in a closed box (acoustically isolated container) to reduce the effects of ambient noise similar to a X-ray (albeit cheaper and much easier to deploy). In fact, a simple makeshift isolation chamber (a cardboard equivalent of a covered chamber for X-rays) can provide reasonable acoustic shielding of the area of the system from loud ambient noise. Further, more precise and directional microphones can overcome the noise limitations of our proof-of-concept setup.

9.4 Effect of Training Data

To evaluate the effect of MiLTOn across training data, as we increase the training samples, we proportionally reduce the number of bad points chosen to ensure a 50-50 distribution between good and the bad points. As shown in Fig. 9b, both false positives and negatives decrease as we increase the number of training samples. However, beyond a certain number of training samples the return on investment by training on more and more data are limited. This occurs around 9 training samples in our dataset. Thus, in a practical setting, it would be perceptive of a retailer to identify this threshold and not unnecessarily train on all available samples.

9.5 Cost Savings

It is critical to understand how MiLTOn's performance translates to cost savings for an enterprise to achieve profitability. We perform this analysis based on the publicly known costs and metrics described in Sec. 6. We use standard metrics to evaluate the costs involved in shipping ceramic cups. In this case study, each cup costs $3 and costs $0.82 to ship per cup. In addition, the labor costs are assumed to be $15/hr. Therefore, the effective cost of a broken cup reaching the user is $16.94. The effective cost of a broken cup that is detected early is $5.48, which includes the shipping cost, the replacement cost, and the labor cost of verifying damage.

Therefore, any cup that is detected early saves the retailer $11.46. For false positives, the retailer pays the entire damage cost ($16.94). For false negatives, the retailer must pay an additional labor cost of $4.16 for labor. Finally, according to public reports, the percentage of damaged shipments varies from 6% to 15%. Higher damage rates make our system more profitable. For this section, however, we assume the lower end of this damage range: 6% as well as a 1% worst case for analysis.

We use these metrics to estimate the effective cost savings provided by MiLTOn and plot the results in Fig. 11 as a function of the threshold τ. We use this analysis to show that the profit optimum value of τ is different from the equal error value. Based on our analysis, we expect to save a company up to $607,440 if 6% of its packages are being returned due to damages, or $76,521 if only 1% are, for every million cups shipped.

The cost of deploying MiLTOn is $300 for mmWave radar and $50 for acoustic setup as we can reuse existing cameras on the supply chain. An estimate of cloud storage and compute cost for a million signatures varies between $30-$90 a year. Further, suppose we deploy the system in 100 warehouses of a large supply chain company, the total cost comes to roughly $40,000. This means that even if 6% of a supply chain packages are being returned due to damage, we can still see an effective profit of $567,440 every year. We further discuss how this is a lower bound for the potential profit in Sec.11.3.

10 ROBUSTNESS ANALYSIS

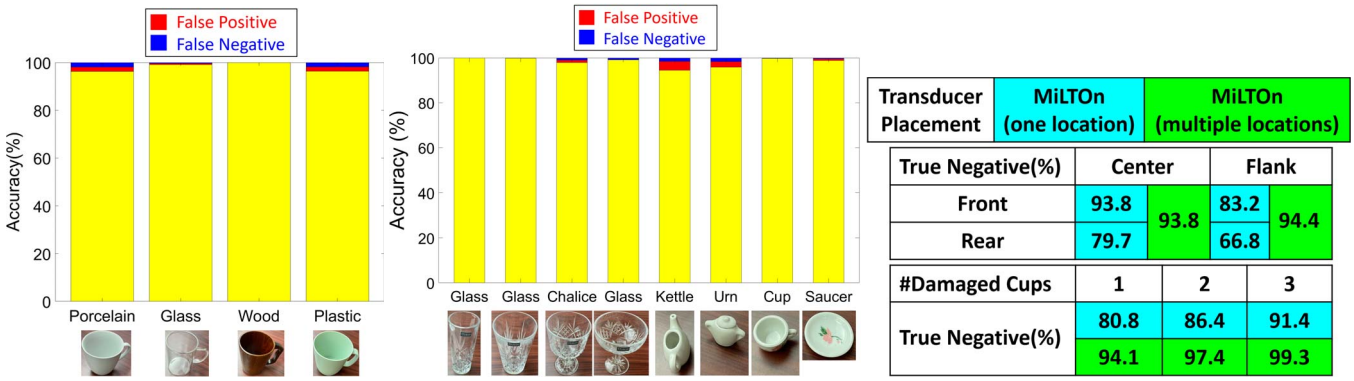
In this section, we present robustness analysis of MiLTOn and evaluate accuracy across materials, object type, size, padding and presence of multiple objects in a package.

10.1 Accuracy across Materials

We study the effect of the material of the object on MiLTOn's performance. To evaluate this, we used 4 cups/mugs of similar sizes made from 4 different materials – porcelain, glass, wood and plastic. We then collected 15 acoustic measurements from each of the mugs before breaking them with a hammer. We collect 10 acoustic measurements after breakage while giving the box a strong shake to emulate transportation. We use 5 out of the 15 good measurements to train our signature and evaluate across the remaining 20 measurements. We repeat this selection of 5 training signatures 100 times to prevent training data bias.

Result: Our results shown in Fig.12a show a high accuracy in detecting the damage across the materials. This is due to the fact that our acoustic signature is a combination of the box-object acoustic system and a broken object also affects the behavior of the box in turn improving the accuracy of the system. Across the 4 materials, we achieve an aggregate accuracy of 97.91%. The most surprising result here was the fact that the wooden cup had only it's handle broken, but showed the widest girth between the training and testing behavior reaffirming our understanding of the system.

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.



Transducer Placement	MiLTOn (one location)		MiLTOn (multiple locations)	
True Negative(%)	Center		Flank	
Front	93.8	93.8	83.2	94.4
Rear	79.7		66.8	

#Damaged Cups	1	2	3
True Negative(%)	80.8	86.4	91.4
	94.1	97.4	99.3

(a) (b) (c)

Figure 12: Robustness Analysis: (a) Accuracy of MiLTOn across materials of the object; (b) Accuracy of MiLTOn across different types of objects; (c) MiLTOn Multi-cup evaluation: 6 cups in a box; (top) shows the benefit of using multiple transducers over a single transducer at center front across location of damage; (bottom) shows the accuracy vs. number of cups damaged

Padding	Shaved Paper	Crumpled Paper	Packing Peanuts
Accuracy (%)	98.1	97.5	97.9

Figure 13: Effect of packaging on MiLTOn

10.2 Accuracy across Products

Another important dimension to evaluate is how the acoustic signature changes across shape and size of the products. Thus, we picked two classes of objects that deviate from our case study. A set of large and heavy glassware, and a set of articles typical children's porcelain toy sets. Evaluation steps are same as Sec. 10.1.

Result: Our result in Fig.12b shows the accuracy across the objects. Clearly, the heavier and larger glassware demonstrates significantly better accuracy than the smaller ceramic objects. We surmise that this is due to the fact that bigger and heavier objects have a larger contribution to the acoustic response of the box-object composite system than the lighter and smaller ones. Another key observation in this evaluation was that the signature remains robust to orientation changes as the glass products were removed and placed back in a different orientation between measurements. Averaged across the groups, we see a 99.23% accuracy (0.56% false positives) for the glassware compared to the 97.23% accuracy (3.65% false positives) for the porcelain toy set.

10.3 Resilience to Packaging

We evaluate the effect of lightweight packaging on MiLTOn's operation in real world. We evaluate our system for three packaging materials – shaved paper, crumpled paper, and packing peanuts. Evaluation steps are same as Sec. 10.1.

Result: As depicted in Fig. 13, across the three packaging materials, MiLTOn's performance remains above 97% accuracy. We observe that padding creates small air pockets in the box that add behavior at relatively high frequencies (8 KHz and above, cup-sized objects resonate around 2.4 and 7.2 KHz), however as these behaviors are highly dynamic, across signatures their contribution gets diluted. We believe that the accuracy remains unaffected as enough energy reaches the cup in a typical setting for it to resonate and contribute it's unique signature. If the cup does not vibrate and resonate (in an adversarial setting), we surmise the accuracy will drop drastically.

10.4 Multiple Objects

Finally, before presenting our detailed study on porcelain cups, we wanted to analyze how the accuracy of MiLTOn varies when there are multiple objects inside the same box. We use 6 cups in a box as the base line and first create a single breakage for each cup location to evaluate MiLTOn. We also study the impact of transducer location on the accuracy of the system. Finally, we also attempt to detect multiple breakages within a box simultaneously.

Result: Our results in Fig.12c (top) shows that the accuracy of MiLTOn is heavily affected by the location of transducer. When the transducer is placed optimally (green), the accuracy in finding damage to objects remains above 90%. However, when it is not placed at the right location, the system becomes more inaccurate. This behavior is expected as the amount of energy transferred to cups on the rear and flank is significantly attenuated (rear more so than flank). Thus, at an average the accuracy of detecting damage across locations when the transducer is located statically at center-front location (blue) is 80.8%. However, as multiple cups get damaged (Fig.12c (bottom)), this accuracy increases due to large changes in the cumulative received acoustic signals.

To improve accuracy further when sensing multiple cups, rather than placing the transducer at only one location, we consider an alternative denoted as MiLTOn (multiple locations) in Fig. 12c that either uses multiple transducers or takes multiple measurements at multiple locations from a single transducer. By doing so, we demonstrate that the accuracy of the system remains above 90% (green) when multiple cups are damaged, ensuring effectiveness.

11 DISCUSSION

11.1 Limitations

While our work solves many practical challenges in making acoustic vibrometry a reality over the air, there are several limitations

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.

of our work that need to be addressed to make the system more ubiquitous and robust to various factors. First, our evaluation is limited to the scenarios where object under consideration has a resonant behavior. This means other commodity objects such as clothes, groceries will require more sophisticated solutions to detect spoilage or tears. However, the class of resonant objects does span a wide variety of products from porcelain and metallic products to wooden and glass products. We believe an acoustic synthetic-aperture radar approach may allow for disambiguating individual objects, which we leave for future work. Second, our system struggles if the material of the enclosure itself is highly absorptive as the amount of energy reaching the object is too little to make it resonate. Another scenario which may affect the ability of MiLTOn to correctly isolate the behavior of the object is when the object moves inside a sparsely occupied box. While mmWave should be able to detect this movement and flag it, it will remain difficult to ascertain if the object is broken or not. However, while the response SNR of the object-box system is greater than 9 dB, MiLTOn can achieve greater than 90% accuracy. Finally, we acknowledge that no quality control system can be truly exhaustive in its analysis – there will always remain some forms of damage that are unforeseen.

11.2 Who deploys MiLTOn?

We ask the question of who has the correct incentives to deploy MiLTOn with many diverse stakeholders and companies often involved along the supply chain. While a detailed economic analysis is beyond the scope of this paper, we believe MiLTOn can be a win-win system if it augments the existing supply chain insurance framework that various stakeholders already pay into (usually to mitigate business interruption). Imagine insurance firms that offer incentives for companies to deploy MiLTOn by reducing premiums and rewarding companies that handle packages with less damage caused. We believe that much like how big data from phone accelerometers [14, 42] has revolutionized the auto insurance sector, a similar opportunity exists in supply chain insurance with MiLTOn.

Another key factor to take into consideration is the cost effectiveness of MiLTOn. Indeed, for an expensive or large object (e.g., a designer vase), it might be more cost-effective to open the box at every location. Further, the cost of the system would make sense to deploy where the quantity of the objects is sufficiently large to amortize the initial cost of a few hundred dollars by sheer scale. Much of the compute and storage can be offloaded to a cloud service to further reduce the costs of deployment for the end user.

11.3 Improving MiLTOn latency and cost

A key bottleneck in real world adoption of many sensing solutions for supply chain lies in the latency and cost of the system. MiLTOn's current implementation that relies on commodity hardware faces similar bottlenecks. However, MiLTOn can improve significantly with custom hardware purchased at scale. For e.g. MiLTOn uses a 5 second chirp to sense the behavior of the object which gives a resolution of 0.2 Hz in the received signal power. With custom hardware, this can be reduced by at least a factor of 20 (frequency resolution of 4 Hz) with little effect on the accuracy of the system and in turn reducing the latency to 250 ms. Further, we use off-the-shelf prototypes for mmWave Radar and acoustic transducers which cost roughly 10× the cost of a customized solution. This can reduce the cost of deploying the system as well as make multiple parallel chains cost-effect for reducing the impact on latency of the spring-loaded arm.

11.4 Future application scenarios

While we believe MiLTOn is an exciting first step in identifying package anomalies, it's underlying approach can scale across various industry verticals. One such vertical is automated robotic warehouses, where all ingress, assessment and dispatch is performed using smart algorithms. Our approach can provide an additional sensing modality to track object integrity as a package is being carried around by a robot. Failures can be investigated to further improve reliability guarantees of the robots. Another direction of exploration that logically follows MiLTOn is leveraging the vast literature in acoustic vibrometry and sensing to improve the capabilities of MiLTOn beyond integrity checks to new objectives such as material sensing and food quality identification. Each of the above can build upon MiLTOn's learnings.

Further, this highlights the strengths of using multiple modalities to detect damage or mutilated products. MiLTOn's three modalities provide three complementary information sources measuring different properties of the object. The camera detects visible damage to the box externally. mmWave signals can detect macro changes in electro-magnetic properties of the object (detecting a brick instead of a iPhone) and can see through the box to locate the objects. Acoustic signals detect the mechanical properties of the object (typically mechanical stresses cause breakages) to perform integrity checks. Thus, future applications can leverage more modalities to complement MiLTOn and improve the capabilities of the system.

12 CONCLUSION

This paper presents MiLTOn, an acoustic solution for non-invasive product testing that verifies the integrity of fragile porcelain and glass products as they move through the supply chain. Assisted by mm-wave and camera systems, MiLTOn can monitor even tiny sub-millimeter cracks in the product using principles of acoustic vibrometry. Different from prior research in vibrometry, MiLTOn can achieve this without physical contact with the product within the box and operates non-invasively with commodity sensors. To do so, MiLTOn designs a novel mechanism that treats the box itself as an acoustic speaker and processes its acoustic response to assess damage of the product within it. Our extensive evaluation demonstrates high accuracy in determining product damage for fragile goods. While this paper focuses on fragile goods (primarily glass and porcelain tableware), we believe that there is potential for extensive future work in extending our approach to other diverse domains including metallic goods and agricultural products, all of which are known to offer distinctive acoustic responses.

ACKNOWLEDGEMENTS

We thank our shepherd and the anonymous reviewers for their valuable feedback and insights. We thank our colleagues at Microsoft Research for Industry and WiTech Lab for their constructive inputs and help. This research was done at Microsoft. Akshay was supported in part by NSF (awards 2106921, 2030154, 2007786, 1942902).

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] 2000. Acoustic resonance analysis in manufacturing. https://www.ndt.net/article/v05n01/hertlin/hertlin.htm.

[2] 2019. How much a damaged pack can really cost your business. https://www.amcor.com/insights/blogs/how-much-a-damaged-pack-can-really-cost-your-business.

[3] 2019. Rethinking Packaging: A DHL perspective on the future of packaging in the logistics industry. https://www.dhl.com/content/dam/dhl/global/core/documents/pdf/glo-core-rethinking-packaging-trend-report.pdf.

[4] 2019. Tableware Market Size. https://www.grandviewresearch.com/industry-analysis/tableware-market.

[5] 2021. Customer Review for 10 Strawberry Street. https://www.amazon.com/10-Strawberry-Street-CATERING-12-MUG-W-Catering/.

[6] 2021. Introduction to SIFT. https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_feature2d/py_sift_intro/py_sift_intro.html.

[7] Fadel Adib and Dina Katabi. 2013. See through walls with WiFi!. In *ACM SIGCOMM 2013*. 75–86.

[8] Delan Zoe H Arenga and Jennifer C Dela Cruz. 2017. Ripeness classification of cocoa through acoustic sensing and machine learning. In *IEEE International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*. 1–6.

[9] Danilo Balzarini, James Erskine, and Michael Nieminen. 2021. Use of the Pavement Surface Cracking Metric to Quantify Distresses from Digital Images. *Transportation Research Record* 2675, 9 (2021), 1534–1544.

[10] Nam Bui, Nhat Pham, Jessica Jacqueline Barnitz, Zhanan Zou, Phuc Nguyen, Hoang Truong, Taeho Kim, Nicholas Farrow, Anh Nguyen, Jianliang Xiao, Robin Deterding, Thang Dinh, and Tam Vu. 2019. eBP: A wearable system for frequent and comfortable blood pressure monitoring from user's ear. In *ACM MobiCom*. 1–17.

[11] Charles J Carver, Charles J Carver, Zhao Tian, Hongyong Zhang, Kofi M Odame, Alberto Quattrini Li, and Xia Zhou. 2020. Amphilight: Direct air-water communication with laser light. In *USENIX NSDI*. 373–388.

[12] Robin Clark. 2004. Rail flaw detection: overview and needs for future developments. *Ndt & E International* 37, 2 (2004), 111–118.

[13] Debaditya Dutta, Hoon Sohn, Kent A Harries, and Piervincenzo Rizzo. 2009. A nonlinear acoustic technique for crack detection in metallic structures. *Structural Health Monitoring* 8, 3 (2009), 251–262.

[14] Jakob Eriksson, Lewis Girod, Bret Hull, Ryan Newton, Samuel Madden, and Hari Balakrishnan. 2008. The pothole patrol: using a mobile sensor network for road surface monitoring. In *ACM MobiSys 2008*. 29–39.

[15] N Galili, I Shmulevich, and N Benichou. 1998. Acoustic testing of avocado for fruit ripeness evaluation. *Transactions of the ASAE* 41, 2 (1998), 399.

[16] Z Gong, EO Nyborg, and G Oommen. 1992. Acoustic emission monitoring of steel railroad bridges. *Materials evaluation* 50, 7 (1992).

[17] Yasha Iravantchi, Yang Zhang, Evi Bernitsas, Mayank Goel, and Chris Harrison. 2019. Interferi: Gesture sensing using on-body acoustic interferometry. In *ACM CHI*. 1–13.

[18] Chitra R Karanam and Yasamin Mostofi. 2017. 3D through-wall imaging with unmanned aerial vehicles using WiFi. In *ACM/IEEE IPSN 2017*. 131–142.

[19] Jin-Yeon Kim and Stanislav I Rokhlin. 2002. Surface acoustic wave measurements of small fatigue cracks initiated from a surface cavity. *International journal of solids and structures* 39, 6 (2002), 1487–1504.

[20] Brian Klucinec. 1996. The effectiveness of the Aquaflex gel pad in the transmission of acoustic energy. *Journal of athletic training* 31, 4 (1996), 313.

[21] Nachiket Kotwaliwale, Karan Singh, Abhimannyu Kalne, Shyam Narayan Jha, Neeraj Seth, and Abhijit Kar. 2014. X-ray imaging methods for internal quality evaluation of agricultural produce. *Journal of food science and technology* 51, 1 (2014), 1–15.

[22] Sangdae Lee and Byoung-Kwan Cho. 2013. Evaluation of the firmness measurement of fruit by using a non-contact ultrasonic technique. In *IEEE Conference on Industrial Electronics and Applications (ICIEA) 2013*. 1331–1336.

[23] Xiaoli Li, Shen Dong, and Zhejun Yuan. 1999. Discrete wavelet transform for tool breakage monitoring. *International Journal of Machine Tools and Manufacture* 39, 12 (1999), 1935–1944.

[24] Zhengxiong Li, Fenglong Ma, Aditya Singh Rathore, Zhuolin Yang, Baicheng Chen, Lu Su, and Wenyao Xu. 2020. Wavespy: Remote and through-wall screen attack via mmwave sensing. In *IEEE S&P 2020*. 217–232.

[25] Jian Liu, Yingying Chen, Marco Gruteser, and Yan Wang. 2017. Vibsense: Sensing touches on ubiquitous surfaces through vibration. In *IEEE SECON 2017*. 1–9.

[26] Manni Liu, Linsong Cheng, Kun Qian, Jiliang Wang, Jin Wang, and Yunhao Liu. 2020. Indoor acoustic localization: a survey. *Human-centric Computing and Information Sciences* 10, 1 (2020), 1–24.

[27] Xiangwei Liu, Deli Pei, Gabriel Lodewijks, Zhangyan Zhao, and Jie Mei. 2020. Acoustic signal based fault detection on belt conveyor idlers using machine learning. *Advanced Powder Technology* 31, 7 (2020), 2689–2698.

[28] Xiaoqin Liu, Xing Wu, and Chang Liu. 2011. A comparison of acoustic emission and vibration on bearing fault detection. In *IEEE International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE) 2011*. 922–926.

[29] Wenguang Mao, Mei Wang, and Lili Qiu. 2018. Aim: Acoustic imaging on a mobile. In *ACM MobiSys 2018*. 468–481.

[30] Wenguang Mao, Mei Wang, and Lili Qiu. 2019. Mobile Imaging Using Acoustic Signals. *GetMobile: Mobile Computing and Communications* 22, 4 (2019), 35–38.

[31] Jairo Aparecido Martins, Daniel Leite, and Estaner Claro Romão. 2020. Studying Resonant Frequencies of a Helical Spring with and Without Axial Loads. *Journal of Failure Analysis and Prevention* 20, 4 (2020), 1301–1307.

[32] Koji Mizuno, Haruki Matono, Yoshihiko Wagatsuma, Hideo Warashina, Hiroyasu Sato, Seiko Miyanaga, and Yukio Yamanaka. 2005. New applications of millimeter-wave incoherent imaging. In *IEEE MTT-S International Microwave Symposium Digest, 2005*. IEEE, 4–pp.

[33] Arun Mohan and Sumathi Poobal. 2018. Crack detection using image processing: A critical review and analysis. *Alexandria Engineering Journal* 57, 2 (2018), 787–798.

[34] Tom Parker, Sergey Shatalin, and Mahmoud Farhadiroushan. 2014. Distributed Acoustic Sensing–a new tool for seismic applications. *first break* 32, 2 (2014).

[35] Nhat Pham, Tuan Dinh, Zohreh Raghebi, Taeho Kim, Nam Bui, Phuc Nguyen, Hoang Truong, Farnoush Banaei-Kashani, Ann Halbower, Thang Dinh, et al. 2020. WAKE: a behind-the-ear wearable system for microsleep detection. In *ACM MobiSys 2020*. 404–418.

[36] Akarsh Prabhakara, Vaibhav Singh, Swarun Kumar, and Anthony Rowe. 2020. Osprey: A mmWave approach to tire wear sensing. In *ACM MobiSys*. 28–41.

[37] MT Reiten and R Glenn Wright. 2008. Laser doppler vibrometry use in detecting faulty printed circuit boards. In *2008 IEEE AUTOTESTCON*. IEEE, 33–36.

[38] Nirupam Roy and Romit Roy Choudhury. 2016. Listening through a vibration motor. In *ACM MobiSys 2016*. 57–69.

[39] Sarah Schotte, Nele De Belie, and Josse De Baerdemaeker. 1999. Acoustic impulse-response technique for evaluation and modelling of firmness of tomato fruit. *Postharvest biology and technology* 17, 2 (1999), 105–115.

[40] Avinashchandra Singh, Donald R Houser, and Sandeep Vijayakar. 1998. Detecting gear tooth breakage using acoustic emission: a feasibility and sensor placement study. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 80388. American Society of Mechanical Engineers.

[41] Sukhvinder Singh, Qilian Liang, Dechang Chen, and Li Sheng. 2011. Sense through wall human detection using UWB radar. *EURASIP Journal on Wireless Communications and Networking* 2011, 1 (2011), 1–11.

[42] Margareta Teodorescu and Jan Budík. [n.d.]. The Age of Big Data Analytics: How Big Data is Reshaping the Insurance Industry. ([n. d.]).

[43] Francesco Tonolini and Fadel Adib. 2018. Networking across boundaries: enabling wireless communication through the water-air interface. In *ACM SIGCOMM 2018*. 117–131.

[44] Gerard Van Dalen, Peter Nootenboom, Lucas J Van Vliet, Lennard Voortman, and Erik Esveld. 2007. 3-D imaging, analysis and modelling of porous cereal products using x-ray microtomography. *Image Analysis & Stereology* 26, 3 (2007), 169–177.

[45] Anran Wang, Dan Nguyen, Arun R Sridhar, and Shyamnath Gollakota. 2021. Using smart speakers to contactlessly monitor heart rhythms. *Communications Biology* 4, 1 (2021), 1–12.

[46] Anran Wang, Jacob E Sunshine, and Shyamnath Gollakota. 2019. Contactless infant monitoring using white noise. In *ACM MobiCom 2019*. 1–16.

[47] Peter NT Wells. 2006. Ultrasound imaging. *Physics in Medicine & Biology* 51, 13 (2006), R83.

[48] Xiangyu Xu, Jiadi Yu, Yingying Chen, Yanmin Zhu, Linghe Kong, and Minglu Li. 2019. Breathlistener: Fine-grained breathing monitoring in driving environments utilizing acoustic signals. In *ACM MobiSys 2019*. 54–66.

[49] Flora Zidane, Jerome Lanteri, Julien Marot, Laurant Brochier, Nadine Joachimowicz, Helene Roussel, and Claire Migliaccio. 2020. Nondestructive Control of Fruit Quality via Millimeter Waves and Classification Techniques: Investigations in the Automated Health Monitoring of Fruits. *IEEE Antennas and Propagation Magazine* 62, 5 (2020), 43–54.

[50] David Zimdars and Jeffrey S White. 2004. Terahertz reflection imaging for package and personnel inspection. In *Terahertz for Military and Security Applications II*, Vol. 5411. International Society for Optics and Photonics, 78–83.

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on September 23,2022 at 19:31:54 UTC from IEEE Xplore. Restrictions apply.