Abstract—We report the feasibility study of a new optoacoustic sensor for both near-distance ranging and material thickness classification for robotic grasping. It is based on the optoacoustic effect where focused laser pulses are used to generate wideband ultrasound signals in the target. With a much smaller optical focal spot, the optoacoustic sensor achieves a lateral resolution of 93 μm, which is six times higher than ultrasound pulse-echo ranging under the same condition. A new multi-mode broadband PZT (lead zirconate titanate) transducer is built to properly receive the wideband optoacoustic signal. The ability to receive both low- and high-frequency components of the optoacoustic signal enhances the material sensing capability, which makes it promising to determine not only material type but also the sub-surface structures. For demonstration, optoacoustic spectra are collected from hard and soft materials with different thickness. A Bag-of-SFA-Symbols (BOSS) classifier is designed to perform primary material and then thickness classification based on the optoacoustic spectra. The accuracy of material / thickness classification reaches ≥ 99% and ≥ 94%, respectively, which shows the feasibility of differentiating solid materials with different thickness by the optoacoustic sensor.

I. INTRODUCTION AND RELATED WORK

Robust grasping of unknown objects has been a grand challenge for robotics [1] [2]. Sensor-less grasping exists but suffers from efficiency issues [3] [4]. Sensor-based grasping approaches still dominate applications and sensors dictate the quality of information in grasping decision. It is important to detect object relative pose at near distance and recognize material type and even the sub-surface structures (such as the thickness of a thin shell) of the object. The near-distance ranging (< 0.5 cm) helps a robot to properly respond to subtle changes in object pose right before the grasping operation and adjust operations dynamically during grasping. Moreover, the material type and sub-surface structure information can help planner to estimate force distribution, impact characteristics and friction coefficients for better grasping.

Many sensors have been used to assist grasping. Unfortunately, existing sensors for grasping have difficulties in satisfying these requirements. Cameras, LiDARs, or radars suffer from the occlusion caused by closing-in fingers themselves [5] or have a blind zone when the perceived object is too close [6] [7] [8] [9]. Proximity sensors, which include optical, E-field, and ultrasonic types, have been developed for pre-touch ranging recently. However, the optical proximity sensors generally lack the lateral resolution for fine structures [10] [11], and the E-field sensors have difficulties in detecting materials with low dielectric contrast to air [12] [13]. Similarly, pulse-echo ultrasound ranging can measure near distance [14] [15] [16] but has limited lateral resolution due to its widespread signal dispersion pattern. Tactile sensing [17] [18] and force sensing [19] are also popular approaches. However, tactile and force sensing require the robot to touch the object which may change object poses, damage the object, or lead to either slow grasping process or complete failure in grasping.

Previously, we have demonstrated a finger-mounted bi-modal sensor, which combines pulse-echo ultrasound for ranging and optoacoustic generation for material type sensing [20]. It provides robotic fingers with the capability to detect the distance and bulk material type of the target at a close distance before contact occurs. Pulse-echo ultrasound is suitable for short-distance ranging owing to the relatively slow sound speed. However, due to the high contrast of acoustic impedance between air and most bulk materials, it largely loses its effectiveness for material sensing in air, because strong reflection can always occur regardless of the target material. In contrast, optoacoustic generation utilizes short laser pulses to directly excite wideband ultrasound waves (i.e., optoacoustic signals) in the target. The frequency spectrum of the optoacoustic signal is closely related to both the material type and even the sub-surface structures of the target [21] [22]. Therefore, it provides a more effective way for material sensing than pulse-echo ultrasound (Fig. 1).

![Figure 1. A simplified diagram of robotic finger transmitting (a) sound beams (in black) onto target and receiving reflected sound waves (in red), (b) focused laser beams (in green) onto target and receiving induced optoacoustic waves (in red).](image)

However, several issues remain in the bi-modal sensor. First, similar with other pulse-echo ultrasound ranging devices, it has relatively low lateral resolution limited by the acoustic focal spot size. Second, the current air-coupled ultrasound transducers usually have narrow frequency bandwidth. This limits the accuracy of the distance ranging
because the acoustic axial resolution is inversely proportional to the acoustic bandwidth. The narrow bandwidth of the transducer also limits the material sensing capability, because only a small portion of the acoustic spectrum can be received and used for classification. Last, the functioning of the bi-modal sensor requires both ultrasound pulser-receiver and pulsed laser, which makes the entire sensing system complex and costly.

To address these issues, this paper reports a new optoacoustic sensor to achieve both distance ranging and material/structure sensing (Fig. 2). A pulsed laser beam is focused onto the target surface after being reflected by the parabolic mirror. Upon absorption, a wide-band (from DC up to several MHz) optoacoustic signal can be generated on the target surface. Because there is no single ultrasound transducer that can have such a wide-band response, two transducers are employed to receive the optoacoustic signals. The lower-frequency components of the optoacoustic signal are received by a microphone, while the higher-frequency components (after being reflected by the parabolic mirror) are received by a specially-designed hollow stacked transducer with wide bandwidth. The center hole in the transducer allows the excitation laser pulses to pass through and reach target to induce the optoacoustic signal. This co-axial design makes the optical focal spot of the laser and acoustic focal spot of the transducer overlap with each other to improve the detection sensitivity of the optoacoustic signal.

Compared with the previous design, the new sensor design has several advantages. First, the (much) smaller laser focal spot and also (much) wider bandwidth of the optoacoustic signals and also the receiving transducers provide both higher lateral and axial ranging resolution. Second, the ability to receive both low- and high-frequency components of the optoacoustic signal enhances the material sensing capability. This makes it promising to determine both material type and even the sub-surface structures (such as thin shells and voids, etc.). Third, without using the ultrasound pulser-receiver, the entire sensor system will become simpler and lower cost. For demonstration, a prototype optoacoustic sensor is designed, fabricated and tested. Our experimental results show that the optoacoustic sensor can provide better ranging resolution and material/structure sensing capability than the previously-reported bi-modal sensor [20].

II. SENSOR CONSTRUCTION AND TESTING

Fig. 3 (a) illustrates the fabricated prototype optoacoustic sensor. It consists of a 3D-printed housing, a 90-degree parabolic mirror, a microphone with a bandwidth of 0 ~ 80 kHz, and a home-made hollow stacked PZT (lead zirconate titanate) transducer to provide a reception bandwidth of 80 kHz ~ 1 MHz. PZT is the most commonly used piezoelectric material for making (air-coupled) ultrasound transducers due to its excellent electromechanical coupling efficiency. However, a single PZT substrate mainly has two modes of vibration: thickness-mode and flexural-mode, leading to two narrow resonance peaks in its reception bandwidth. By bonding multiple pieces of PZT substrates with different thickness together, the frequency response of the stacked layers will be composed of a mixture of all flexural-mode and thickness-mode vibrations tuned by various acoustic coupling and damping mechanisms and also non-linear effects. As a result, this kind of multi-mode and multi-frequency operation can provide much wider bandwidth than that of each single substrate.

To achieve a reception bandwidth of 80 kHz ~ 1 MHz, three metal-coated PZT substrates are prepared and stacked together (Fig. 3 (b)). The thicknesses of PZT substrates are 2.1 mm, 3.2 mm and 4.3 mm to provide thickness-mode resonance around 1.1, 0.75 and 0.5 MHz, and flexural-mode resonance around 0.1, 0.25 and 0.35 MHz, respectively. Conductive epoxy is applied between the three PZT substrates as the bonding material and onto the back surface of the stacked transducer as the backing material to damp the self-resonance. A thick layer of epoxy is also applied onto the back surface of conductive epoxy as a second backing layer to further increase the damping. A 50-ohm micro co-axial cable is soldered onto the PZT substrate for electrical connection (Figs. 3 (b) and 3 (c)). The piezoelectric polarity of the three PZT substrates is arranged in such way that they are electrically connected in parallel to the co-axial cable, such that all the response from the three PZT substrates is received simultaneously.

![Figure 2](image1.png)

Figure 2. A diagram of the optoacoustic sensor for distance ranging & thickness differentiation mounted onto a robotic finger.

![Figure 3](image2.png)

Figure 3. (a) Photograph of the fabricated prototype optoacoustic sensor. (b) Schematic of the hollow stacked PZT ultrasound transducer. (c) Photograph of the fabricated hollow stacked PZT transducer.

The hollow stacked PZT transducer is the key component of the optoacoustic sensor. Because the higher-frequency components of the optoacoustic signal play a critical role in both ranging and material/structure sensing. Therefore, the performance of the stacked PZT transducer will largely determine that of the entire sensor. An optoacoustic testing is conducted to characterize the overall reception bandwidth of the hollow stacked PZT transducer (Fig. 4 (a)). A Q-switched 532-nm Nd:YAG pulsed laser is used as the light source. The
laser pulse duration is 15 ~ 25 ns and the pulse energy is around 50 μJ. A piece of thick black tape is used as a standard target to generate short-pulse optoacoustic signals with a wide and flat bandwidth. After reception by the transducer, the frequency components of the optoacoustic signal are modulated and therefore can be used to characterize the acoustic reception bandwidth of the transducer. A function generator provides a trigger signal to synchronize the firing of the Q-switched laser and the data acquisition of the oscilloscope. The transmitted laser pulse is reflected and focused onto the black tape by the parabolic mirror. After the laser pulse illuminates the black tape, the induced optoacoustic signal propagates along the reversal path, and is received by the stacked transducer and amplified by the preamplifier. The amplified signals are captured and recorded by the oscilloscope. Fig. 4 (b) shows a representative frequency spectrum of the optoacoustic signal from the black tape, indicating a continuous reception bandwidth of 80 kHz ~ 1.1 MHz. The individual peaks correspond to the resonance frequencies of both thickness-mode and flexural-mode vibration of the three PZT substrates. The frequency components at 80 ~ 400 kHz are stronger than those at 0.5 ~ 1.1 MHz. This can be explained by the fact that under the same excitation condition, the flexural-mode vibration of a thin plate has larger amplitude than the thickness mode.

The setup of optoacoustic distance ranging is similar to that shown in Fig. 4 (a), except that a thin copper wire with a diameter φ = 0.1 mm is used as target to explore the accuracy of optoacoustic ranging for tiny targets. The copper wire is mounted onto a two-axis stage and supported by an adjustable Z-axis stage (Fig. 5 (a)). The laser pulses are reflected and focused by the parabolic mirror with a travel distance of \( L_1 + L_2 + d \), the same as the travel distance of the optoacoustic signal from the source point on the copper wire to the stacked PZT transducer. In this setup, \( L_1 \) and \( L_2 \) are 14 mm and 6.35 mm, respectively. The distance \( d \) between the parabolic mirror and the copper wire is decreased from 9 mm to 4.5 mm with a decrement of 0.5 mm till the target contacts the 3D-printed fixture. At each distance, one-way time delay \( t \) is determined from the optoacoustic signal after 128 times averaging to improve the signal-to-noise ratio (SNR). Fig. 5 (b) shows a representative optoacoustic signal received by the stacked PZT transducer, indicating the measured time delay between the initial rise of the “Trigger” and “Optoacoustic Signal” for distance \( d \) calculation.

$$ L = ct, \tag{1} $$

where \( c \) is the sound velocity in air, which is \( \approx 340 \text{ m/s} \).

III. RANGING EXPERIMENT AND RESULTS

A. Optoacoustic Distance Ranging

When a short laser pulse is incident on the target surface, optoacoustic signal is induced with neglectable delay from laser trigger. So the time delay \( t \) between the trigger and the received optoacoustic signal is equal to the one-way travel time between transducer and target. Therefore, the travel distance of the optoacoustic signal can be calculated as

\[ d = c \cdot t - L_1 - L_2 \]

![Figure 4: A diagram of the general testing setup for the optoacoustic sensor. (b) Representative frequency spectrum of the optoacoustic signal generated by the black tape and received by the stacked PZT transducer.](image)

![Figure 5: (a) A diagram of the optoacoustic ranging test setup. (b) Captured optoacoustic signal, showing the measured delay between “Trigger” and “Optoacoustic Signal”.](image)

Fig. 6 (a) shows the time-delay-calculated distance vs. the real distance \( d \) from 9 mm to 4.5 mm. Fig. 6 (b) shows the deviations of the time-delay-calculated distance from the real one. When the target is close to the laser focal point (\( d = 6 \) mm), the optoacoustic signal arrives at almost the same time, resulting in the smallest deviation and highest accuracy in the time-delay-calculated distance. Therefore, the delay at distance \( d \) of 6 mm is set as the reference for evaluating the deviation at other distances. The deviation is smaller than 0.12 mm when target is within or close to the laser focal zone (when \( d \) ranging from 7.5 mm to 4.5 mm). When the target moves far away from the focal point and even outside the focal zone (when \( d \) is up to 8 mm or even larger), the asynchronous arrival of optoacoustic signal leads to larger
deviation and lower accuracy of distance estimation. For example, the deviation increases when target is far outside the focal zone (when \( d \) ranging from 7.5 mm to 9 mm), reaching its maximum of 0.45 mm at the distance of 9 mm.

![Figure 6](image)

**Figure 6.** (a) Comparison between time-delay-calculated (in black) and real (in red) distance \( (d) \). (b) Deviation of calculated distance from real distance.

### B. Optoacoustic Lateral Resolution

Compared with pulse-echo ultrasound ranging, one major benefit of optoacoustic ranging is its higher lateral resolution for smaller objects, defined by the focal spot size of laser instead of ultrasound. Under the same focusing condition, the focal spot size of laser or ultrasound is proportional to the wavelength. The visible optical wavelength of 400 – 700 nm is much smaller than the acoustic wavelength of 100s of µm in air (at MHz), thereby resulting in a much smaller focal spot and therefore much higher lateral resolution.

The setup to experimentally quantify the lateral resolution is the same as that in Fig. 5 (a), where the \( \phi \approx 0.1 \) mm thin copper wire is scanned laterally by the two-axis stage at a fixed height. The profile of optoacoustic signal amplitude along the scan path is captured and recorded. The FWHM (full-width at half-maximum) value of the Gaussian-fitted profile is used to determine the optoacoustic focal diameter. After repeating the linear scan at different distance \( (d) \) (from 5 mm to 9 mm), the optoacoustic lateral resolution is determined by the minimal optoacoustic focal diameter. The Cu wire is scanned from 0 mm to 0.30 mm with 0.02-mm step at distances \( (d) \) from 5.0 mm to 9.0 mm. The acoustic focal spot reaches its minimum at a distance \( (d) \) of 6.0 mm. Then the scanning is repeated from 0 mm to 0.15 mm with 0.01-mm step at \( d = 6.0 \) mm. The Gaussian-fitted profile of optoacoustic signal amplitude across the scanning range and the FWHM are shown in Fig. 7, indicating the focal spot diameter and lateral resolution is around 93 µm, which is much smaller than the previous pulse-echo ultrasound focal spot of 570 µm [20]. The measured FWHMs are smaller than 0.14 mm at distances \( (d) \) from 5.0 mm to 8.0 mm, and those at \( d > 8.0 \) mm are larger than 0.15 mm, showing the depth of focus is around 3.0 mm.

![Figure 7](image)

**Figure 7.** Optoacoustic lateral resolution determined from Gaussian-fitted profile of received signal amplitude across the scanning range at \( d = 6.0 \) mm.

### IV. Thickness Classification Experiment and Results

#### A. Optoacoustic Spectra Acquisition

To demonstrate the sub-surface structure sensing, optoacoustic spectra are collected with the same experimental setup shown in Fig. 4 (a) from aluminum, plastic and paper sheets (painted with black ink) with different thickness (Table I). These three materials are chosen as an example of materials with high, medium, and low stiffness, respectively. To compensate the difference in the thickness, the height of Z-axis stage is adjusted until the laser focal spot lands on the top surface of each sheet. The two optoacoustic signals received by the microphone and stacked PZT transducer are averaged by 128 times to improve the signal to noise ratio. A representative combined optoacoustic waveform received from an aluminum block is shown in Fig. 8. It consists of a time series of pulses, which is due to multiple reflections between the target surface and the transducer/microphone surface. The first pulse representing the original optoacoustic signal is used for the thickness differentiation. The representative optoacoustic spectra from aluminum, plastic, and paper sheets are shown in Fig. 9.

#### TABLE I. Different Thickness of Aluminum, Plastic and Paper Sheets for Optoacoustic Differentiation

<table>
<thead>
<tr>
<th>Material</th>
<th>Thickness (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminu m</td>
<td>Block 6.35</td>
</tr>
<tr>
<td>Plastic</td>
<td>8.5</td>
</tr>
<tr>
<td>Paper</td>
<td>0.56</td>
</tr>
</tbody>
</table>

![Figure 8](image)

**Figure 8.** A representative combined optoacoustic waveform received from an aluminum block.
Figure 9. Representative optoacoustic spectra from (a) aluminum, (b) plastic, and (c) paper sheets.

B. Thickness Differentiation

The thickness differentiation is performed with a Bag-of-SFA-Symbols (BOSS) classifier [23], with high accuracy and low time complexity [24]. BOSS feature is used on the top of a 1-nearest-neighbor (1-NN) classifier. Feature obtaining has three major steps: firstly, a fixed size sliding window converts input sequence into a group of slices; secondly, Symbolic Fourier Approximation (SFA) transforms the slices into a set of symbols; finally, a histogram of the symbols is generated to represent the sequence. For classification, the 1-NN classifier compares the histogram of input sequence with existing ones, and then exports the label with the most similar histogram. Parameter search used by the ensemble version of BOSS classifier gets an optimal window size and SFA symbol length, making it parameter tuning free. The BOSS classifier is robust to noise and free from sequence alignment. Therefore, it’s an ideal algorithm to classify the optoacoustic signals and validate the feasibility of thickness differentiation by optoacoustics.

To show the difference between optoacoustic signals of the materials with different thickness, The BOSS classifier is trained to primarily identify the materials of aluminum, plastic & paper, and then differentiate the thickness of the same material. The ratio of testing data to training data is 1/3. The experimental data are transformed into BOSS histogram, serving as feature set for classification. After 50 random trials, BOSS classifier gives confusion matrices (Fig. 10) indicating an accuracy of ≥99% for material differentiation and an overall accuracy for ≥94% of thickness classification. This preliminary result demonstrates the feasibility of differentiating solid materials with different thickness by the optoacoustic sensor.

Figure 10. BOSS classifier averaged confusion matrix of (a) materials of aluminum, plastic and paper, (b) aluminum in different thickness, (c) plastic in different thickness, and (d) paper in different thickness.

V. CONCLUSION AND FUTURE WORK

In this paper, we demonstrated a new optoacoustic ranging & material thickness sensor for robotic grasping. It can perform distance ranging with a lateral resolution of 93 µm and a maximal deviation less than 0.12 mm within a 3-mm pre-touch work distance. It is capable of differentiating the type and thicknesses of hard, medium and soft materials before contact occurs. Therefore, this new sensor opened new possibilities of robotic fingers for robust and nimble grasping. In the future, we plan to test more materials and sub-surface structures to optimize the ranging and sensing performance. Several issues will still need to be addressed to make it more suitable for real applications. First, like any other optical approaches, the optoacoustic sensor would fail on targets that are transparent or light-colored with low optical absorption. This issue can be addressed by adding an optoacoustic transmitter into the sensor. Upon the illumination of the excitation laser pulse, the optoacoustic transmitter can send
ultrasound pulse to target for pulse-echo ultrasound ranging and sensing. Second, the work distance of optoacoustic ranging is limited by the small optical focal length and depth, which can be extended by using tunable focal lenses. We will also integrate the sensor on robot fingers to develop perception algorithms to assist grasping.

ACKNOWLEDGMENT

The authors would like to thank for the inputs and feedback from Xiaoyu Duan and Yue Ou.

REFERENCES