AI-DRIVEN SCANNING GHZ ULTRASONIC IMAGING-BASED MEMS METROLOGY

Karan Jha[1], Anuj Baskota[1], Justin Kuo[1], Serhan Ardanuç[1], Scott Zimmerman[1], and Amit Lal[1,2]
[1]Geegah Inc., Ithaca, USA.
[2]School of Electrical and Computer Engineering, Cornell University, Ithaca, USA.

ABSTRACT

Micro-Electro-Mechanical Systems (MEMS) are pivotal in advancing modern technology, yet their mechanical properties measurement remains a challenge. We introduce an AI-driven Ultrasonic Imaging-Based Scanner that offers rapid and accurate mechanical properties assessment of thin-film stacks. This system leverages artificial intelligence to mitigate ultrasonic interference fringes, enhancing image clarity and accuracy. Through sophisticated image processing and AI deep neural networks, our approach not only refines the imaging process but also automates defect detection and quality control with minimal human intervention. This manuscript details the development, implementation, and evaluation of this innovative tool, demonstrating its potential to advance MEMS manufacturing.

KEYWORDS

GHz, ultrasonic imaging, Artificial Intelligence, YOLOv8, Control System,

INTRODUCTION

Micro-Electro-Mechanical Systems (MEMS) devices are integral to numerous applications, from sensors to actuators, owing to their unique high aspect ratio structures and multifunctionality. However, the mechanical properties of these devices, which often dictate their performance and yield, are challenging to measure accurately at the wafer scale. Traditional methods for thin film characterization include wafer curvature [1], Raman spectroscopy to reveal stress effects, photo-acoustic ultrasonic waves while useful, can be too slow, expensive, and require considerable processing for easier utility in research and production. A tool that can image the surface and sub-surface components of MEMS chips, creating images of the structure and the mechanical properties simultaneously would provide a faster path for MEMS devices to be brought to market, due to faster development times and higher yields.

Here, we have introduced an AI driven automated scanning tool that has the potential to provide fast and accurate measurement of mechanical properties of thin-film stacks across a wafer. The tool can visualize and reveal the surface characteristics of MEMS chips by differentiating between structures based on the difference in the acoustic reflection coefficients, which are directly related to the geometric mean of the elasticity and density. This work attempts to addresses a significant challenge of removing the artifact of ultrasonic interference fringes from the images using AI. We use image processing and AI DNN (Deep Neural Net) models to estimate the state of an ultrasonic imaging system and send control inputs based on the estimated state. The imaging system comprises a 3D stage with an optical camera and a GHz ultrasonic imaging array. GHz imaging chip consists of a 128x128 50µm pixel array of 2µm thick Aluminum Nitride (AlN) transducers that allow imaging at a sampling rate of up to 12 frames per second (FPS) (Figure 1a-1d). In our previous work, we successfully demonstrated the use of a GHz ultrasonic imager array to sense fingerprints [2], nematodes [3], soil parameters [4], lens [5], and hydrogel metrology [6].

Here, we also demonstrate a fully automated system that requires minimal to no human intervention and can provide high frame rates in detecting defects and quality control. The imager is non-destructive, it allows a high-resolution scanning of any material with an impedance mismatch with silicon that is deposited on the chip by measuring the ultrasonic pulses transmitted through a thin liquid layer that is squeezed between the sample and the imaging chip. The imaging chip is brought in contact with the sample to be imaged by using positional feedback like other high resolution imaging techniques such as Atomic Force Microscopy [7] and Scanning Probe Microscopy [8].

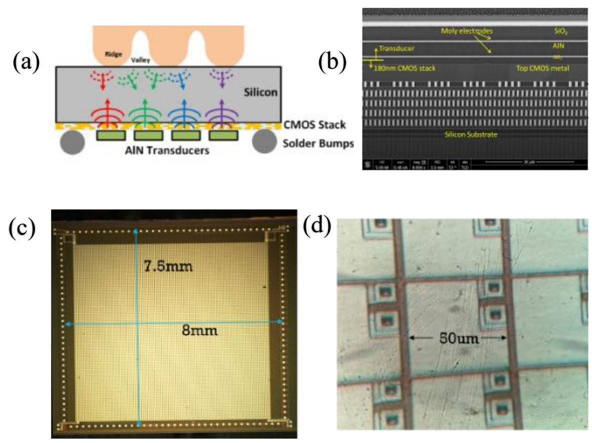

Figure 1: a) The concept of GHz ultrasonic pulses transmitted through silicon and reflected from top. b) AlN on CMOS stack. c) 128x128 50um pixel array. d) Close up of 50 um pixel array

A 3-D stage is used to control the imaging chip. Edge detection and image classification estimate the state of the stage and the sample (Figure 2a, 2b). The immediate image captured by the imaging chip is examined using YOLOv8 [9] to determine the state of the 3D stage. Based on the estimated state of the stage, control inputs are sent using serial connection. YOLOv8 is the state-of-the-art neural networks-based architecture typically used for object detection and classification. However, by updating the output layer, which is the final layer, we can modify it to classify images based on the features present in the images. YOLOv8, through a deep convolution layer network can understand deep-lying features of the image to provide a powerful tool for image classification [10].

The YOLOv8 based image classifier determines when the chip structures are visible in the images. Once clear structures are visible, the imaging is stopped. The obtained image is noisy due to electronic and environmental sources. More importantly, interference fringes are present as expected from the wave nature of the ultrasonic propagation through layered media with thicknesses comparable to wavelengths. We attempt to employ a U-Net architecture-based machine learning model to de-noise and de-fringe the obtained ultrasonic images in order to extract the underlying structures and defects on or within the sample (Figure 2c).). The sample is a surface micromachined chip fabricated in PolyMUMPS technology [11], which has released polysilicon structures from a few micrometers to millimeters (Figure 2d).

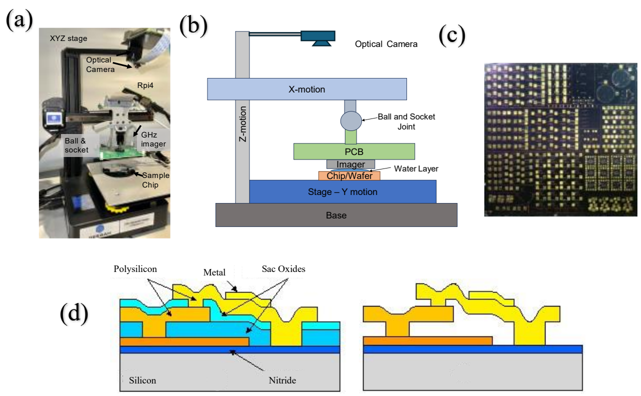

Figure 2: a) An XYZ scanner with optical aligner and imager that can obtain images over a die or a wafer by scanning. b) A schematic detailing the components of the XYZ scanner. c) Optical image of a 1cm x 1cm sample MUMPS chip that was imaged using the scanner. d) Cross section view of the PolyMUMPS surface micromachining technology [11].

METHODS AND EXPERIMENTS

We use a 3-dimensional stepper-motor-based stage to bring the imager in contact with the sample placed on the base of the stage (Figure 2a,2b). The imaging chip is attached to the stage using a ball-and-socket joint to allow the board to align parallel to the surface of the sample.

Stage Control System

We designed a closed loop system to control and automate the motion of the stage without requiring human input. An optical camera captures the sample, and the stage can be used to augment the ultrasonic imaging. The captured optical image is processed to estimate the X-Y coordinate of the sample. Once the imagine chip is aligned with the sample, the stage starts descending, moving the imaging chip closer to the sample at $500\,\mu m/s$. Rolling averages and standard deviation of the 20 previous frames are continuously recorded. A sufficient deviation between the rolling averages and the current frame triggers liquid contact. The stage's descendance is slowed down to $50\,\mu m/s$ once the fluid contact with the imager chip is established (Figure 3a).

As the stage descends further down, the imager aligns itself parallel to the sample's surface through the ball-and-socket joint (2a,2b). After the fluid contact, we use a fine-tuned YOLOv8 model to detect the frames where we can see interference fringes, which appear due to the interference between the imaging surface and the sample. The structures start appearing in the captured ultrasonic images at larger fringe sizes (Figure 3a).

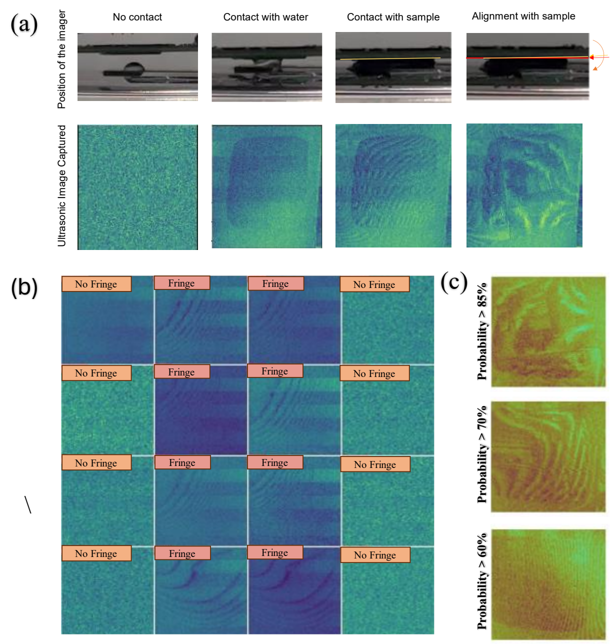

Figure 3: a) Imager approaching the MEMS die with wafer layer. b) Fringe identification using YOLOv8. c) Stopping point for probability threshold 60%, 70%, and 85%.

Training YOLOv8 for image classification

A YOLOv8 image classification model was fine-tuned to classify images based on whether structures are visible in those images. 230 images were annotated using ROBOFLOW based on the visibility of fringes. We used transfer learning by initializing the weights from the -pretrained *yolov8n-cls* model which is the YOLOv8 model with a classification head and further training it for 5 epochs on the training data. The model could achieve 100% accuracy in detecting the images with fringes as confirmed over 3 trials with the number of frames ranging between 55 and 105 (Figure 3b). We used a probability threshold to control the size of fringes (Figure 3c).

Training a UNet Denoiser

Training data was generated synthetically using Python code generating images that can be fed into the model. Real data was generated by randomly assorting shapes on a 64x64 grid. Noisy data was developed by adding gaussian noise and synthetic fringe patterns to the real images. The

brightness of the background and foreground was randomly chosen to emulate experiment conditions. The fringes were generated by randomly selecting one of the four functions (Figure 4a-4d):
$F(x,y) = \sin(\nu xy + \phi)$, $F(x,y) = \sin(\nu x + \phi)$, $F(x,y) = \sin(\nu y + \phi)$, and $F(x,y) = \sin(\nu x + \phi) + \sin(\nu y + \phi)$

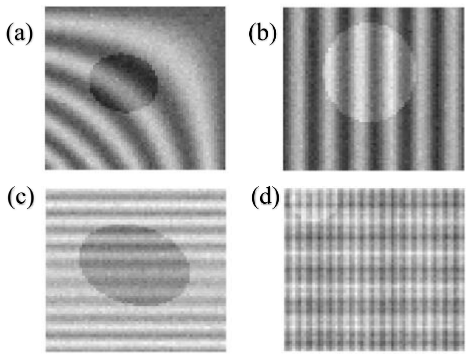

Figure 4: Fringes generated using one of the four functions a) $F(x,y) = sin(\nu xy + \phi)$, *b)* $F(x,y) = sin(\nu x + \phi)$, *c)* $F(x,y) = sin(\nu y + \phi)$, *and d)* $F(x,y) = sin(\nu x + \phi) + sin(\nu y + \phi)$

where $F(x,y)$ represents the fringe intensity at the coordinates (x,y). The fringe parameters ν (fringe frequency) and ϕ (phase shift) were also randomly chosen. The randomness in generating fringes makes the model robust to different kind of fringe patterns it could encounter in ultrasonic imaging. We also added images with multiple fringe patterns superimposed to train the model on images with large fringes and minimal visibility of structures. We add some gaussian noise on top to emulate the process noises that are encountered during the Ultrasonic imaging (Figure 5b middle).

A gaussian blur is applied to each captured frame prior to processing by the deep neural network (DNN) to smooth images by reducing detail and noise. This smoothing was achieved by convolving the image with a Gaussian function, effectively reducing high-frequency components such as noise and minor variations in pixel intensity that are not pertinent to the structural analysis of MEMS devices with dimensions considered in this study ranging from 1 µm to 10 µm . The primary objective of implementing Gaussian blur was to minimize the impact of these undesirable elements on the AI's performance, particularly in enhancing the model's ability to focus on significant structural features rather than noise. By doing so, the system's overall ability to accurately detect and classify defects based on acoustic reflection coefficients was markedly improved.

Each U-Net model begins with an input layer that accepts images of size 64x64 pixels with a single channel, reflecting our grayscale ultrasonic images. The architecture comprises a symmetrical structure with an encoder and decoder path, each consisting of multiple convolutional layers followed by batch normalization and max-pooling or up-sampling layers, respectively. The deepest layer separating the encoder and the decoder paths is called the bottleneck which itself consists of convolutional layers to understand deep lying features. The encoder path includes four stages of convolutional layers with increasing filter sizes (64, 128, 256, 512), each followed by batch normalization and max pooling. This design progressively reduces the spatial dimensions while increasing the depth, enabling the network to encode the input images into a compact, feature-rich representation. At the deepest layer, the model employs a large set of 1024 filters, which serves as the bottleneck. This layer captures the most abstract representations of the input data. The decoder path mirrors the encoder, progressively applying up-sampling followed by concatenation with the corresponding encoder outputs (skip connections), convolutional layers (512, 256, 128, 64 filters), and batch normalization. These steps help the network to reconstruct the detailed spatial information lost during encoding. The final layer of the model is a convolutional layer with a single filter applied with a sigmoid activation function, producing the denoised output image.

For training, we initialized and compiled each U-Net model with the Adam optimizer [12] and binary cross entropy loss function. The learning rate is initially set at 0.001 and reduced by a factor of 0.1 after every 100 epochs. Each model is trained for 500 epochs with a batch size of 12, using shuffled batches to ensure varied exposure to training data across epochs. The training process includes verbose logging for real-time performance tracking. After obtaining the predictions from each model the final pixel prediction is obtained by averaging across all the models.

RESULTS AND DISCUSSION

The system described in the previous sections was used to image gold pads on MUMPS dies as well as defects on wafers (Figure 5a, 5c). We apply a probability threshold for controlling the size of the fringes where our stage stops descending. The results of different probability thresholds are shown in Figure 3c.

The results of image-denoising of ultrasonic images are shown (Figure 5b-5d). The model takes about 50ms to make each prediction while obtaining a training time Signal-to-Noise ratio (SNR) of 30dB and a test time SNR of 15dB. By denoising, we were able to improve the SNR from -4dB to 15dB. We further used this method to extract structures that were not introduced to the model during initial training (Figure 5e).

CONCLUSIONS

This study has presented a novel AI-driven GHz Ultrasonic Imaging-Based MEMS Metrology Scanner, which marks a significant advancement in the field of MEMS device manufacturing and quality control. By integrating AI deep neural networks with advanced ultrasonic imaging technology, we have developed a system capable of delivering fast, accurate, and non-invasive analysis of mechanical properties at the wafer scale. This system not only enhances the clarity and precision of the imaging task but also significantly reduces the need for human intervention by automating key aspects of the defect detection and quality control processes.

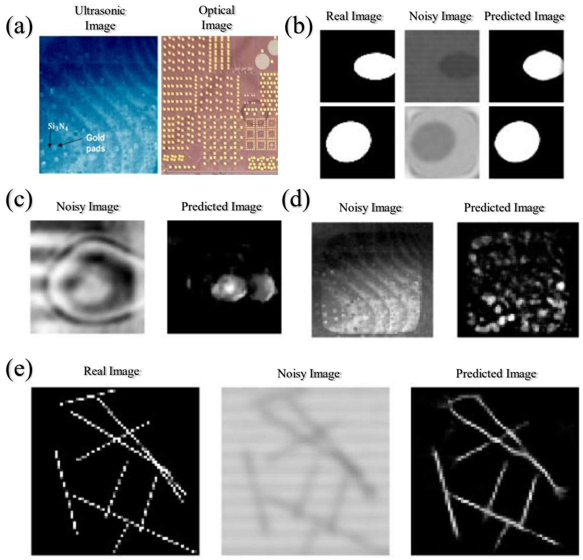

Figure 5: Ultrasonic Image captured of a segment of the sample MUMPS chip. b) Data generated for training with real (left), images with fringes and noise (middle), and predicted images (right). c) Performance of model extracting defects, and d) model extracting the structures on MUMPS chips in real Ultrasonic imaging experiments. e) Performance of the model on denoising newer patterns.

Our experiments and results demonstrate the scanner's ability to discern differences in the acoustic reflection coefficients, which directly correspond to variations in material properties such as elasticity and density. In our study, we were capable to distinguish air (~420 RayI), water (1.5 MRayI) , Si_3N_4 (~32 MRayI), and Gold pads (~62 MRayI). This capability is critical in environments where precision is paramount and can lead to considerable improvements in the yield and reliability of MEMS devices. Furthermore, the application of our AI-driven approach to remove artifacts such as ultrasonic interference fringes has proven effective, resulting in clearer and more accurate images.

The implementation of the YOLOv8 model and the U-Net architecture for image classification and denoising, respectively, has set a new standard in the processing of ultrasonic imaging data. These models have not only improved the operational efficiency of our imaging process but have also showcased the potential of AI to transform traditional manufacturing practices by enabling more precise control mechanisms and predictive capabilities. The current state of our technology is capable of extracting defects and less complicated structures (Figure 5c,5e). The model is favorably adaptable to identifying unseen patterns (Figure 5e). However, its performance in extracting relatively more complex structures, as in the case of MUMPs chips remains subpar (Figure 5d). Future work begs refinement of the denoising AI to be able to understand complicated structures in ultrasonic images.

Our methods provide improvement over traditional methods like Capacitive Scanning Acoustic Microscopy (CSAM) by allowing imaging at higher frequencies which provides higher resolution. We also automate the process of imaging eliminating the need of human intervention while simultaneously allowing for an increased throughput for quality control processes and the capacity to deal with more complex geometries by integration of Artificial Intelligence.

In conclusion, the AI-driven GHz Ultrasonic Imaging-Based MEMS Metrology Scanner embodies a transformative approach to MEMS metrology. It offers a glimpse into the future of manufacturing technologies where AI and machine learning seamlessly integrate to enhance the capabilities and performance of traditional systems. As we continue to refine these technologies, we anticipate broad implications for the fields of microfabrication and industrial quality control, promising a new era of manufacturing precision and efficiency.

ACKNOWLEDGFMENTS

We would like to thank our colleagues at Geegah Inc. and the School of Electrical and Computer Engineering at Cornell University for their invaluable contributions to this project. We also extend our gratitude to our funding agencies and partners who supported this research. Their collaboration and support have been instrumental in achieving the results detailed in this manuscript.

REFERENCES

1. Flinn, Paul A. "Principles and applications of wafer curvature techniques for stress measurements in thin films." *MRS Online Proceedings Library (OPL)* 130 (1988): 41.
2. J. Hoople, J. Kuo, M. Abdel-moneum, and A. Lal, "Chipscale GHz ultrasonic channels for fingerprint scanning," in 2015 IEEE IUS, Oct. 2015.
3. J. Kuo, et al., "Gigahertz Ultrasonic Imaging of Nematodes in Liquids, Soil, and Air," *2021 IEEE IUS*, Sep. 2021
4. A. Baskota, et al.,"Gigahertz Ultrasonic Multi-imaging of Soil Temperature, Morphology, Moisture, and Nematodes," 2022 IEEE MEMS, Jan 2022
5. J. Hwang, et al., "Sonic Fourier Transform Imaging Using GHz Ultrasonic Transducer Array," in 2024 IEEE MEMS, Jan 2024
6. Y.A.Manav, et al., "Towards GHz Ultrasound Enabled Noninvasive Hydrogel Metrology for Mechanobiology," in 2024 MEMS, Jan 2024
7. Dufrêne, Yves F. "Atomic force microscopy, a powerful tool in microbiology." *Journal of bacteriology* 184.19 (2002): 5205-5213.
8. Bian, Ke, et al. "Scanning probe microscopy." *Nature Reviews Methods Primers* 1.1 (2021): 36.
9. Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
10. Pham, Thi-Loan, and Van-Hung Le. "Ovarian Tumors Detection and Classification from Ultrasound Images Based on YOLOv8." *Journal of Advances in Information Technology* 15.2 (2024).
11. Hristov, M. & Veleva, Plamena & Denishev, Krassimir & Grozdanov, Vladimir & Gaydazhiev, Dobromir. "DESIGN AND INVESTIGATION OF AN ACCELEROMETER."
12. Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).

CONTACT

Karan Jha, *karan@geegah.com