Ultrasonic Evaluation of Agra Marble Inlays (Parchin Kari/Pietra Dura) Using Machine Learning for Acoustic Signal Classification: A Heritage NDT Approach

Esma Tuzovic[1,2], Nico F. Declercq[1,2]

[1] *George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, United States*

[2] *IRL 2958 Georgia Tech - CNRS, 2 Rue Marconi, Metz, 57070, France*

This contribution has been accepted based on the submitted abstract. The submitted manuscript was not reviewed and is reproduced here without edits or corrections by the conference board. Neither the EAA nor the AAA take responsibility for its contents.

Abstract

Agra marble inlay work (also known as Parchin Kari, Pacchikari or Pietra Dura) is a traditional craftsmanship wherein carefully shaped hard stones are inserted into an engraved marble matrix, with the Taj Mahal being the best-known example. Despite its historical importance and popular application to modern items, inlayed objects present significant challenges for material classification due to their complexity and heterogeneous nature.

In this paper, we propose an ultrasonic system driven by machine learning algorithms for automated material classification in Agra-type marble inlay work. The acquired immersion ultrasonic data are subjected to pre-processing steps such as alignment, noise-artifact removal, and feature extraction involving temporal, spectral and energy-related parameters, followed by feature ranking and classifier training.

Using a surface-referenced gating approach across both pulse-echo and through-transmission acquisition, the echo-mode classifier distinguished pure marble, bluestone, and metal with approximately 99.5% accuracy under a spatially-blocked evaluation, while the two-mode comparison revealed that echo and transmission are sensitive to different material properties. These findings demonstrate the potential of non-destructive testing with a data-centric workflow to classify complex inlays, and the method extends to other cultural-heritage objects.

Keywords: *marble inlay, machine learning, material classification, ultrasound, nondestructive testing*

Corresponding author: *etuzovic3@gatech.edu*

1. Introduction

The Taj Mahal (India, Agra; constructed between 1631 and 1648 during the reign of Shah Jahan) is well-known around the world not only due to architectural symmetries but also because of the extensive use of marble inlays. The inlay technique, which goes by various names such as Pietra Dura (hard-stone inlay), Parchin Kari, or Pacchikari, involves embedding colored hard stones in white marble to create floral and geometrical patterns, and this craft practice continues even now in the workshops of Agra and provides both architectural panels and touristic souvenirs. Non-destructive investigation of composite stone structures is relevant in material characterization of the inlay itself, conservation planning and authentication/provenance research; ultrasound is a promising method since it can examine below the surface non-invasively, however, its performance in the heterogenous stone-inlay systems is insufficiently reported. This method has already been explored to assess marble quality, but without considering the specific inlay stones case scenarios [1], [2]. The current paper develops a physics-informed signal representation for ultrasonic C-scans that enables reliable machine-learning classification of marble inlay materials.

2. Experimental Setup

As part of the investigation of the Agra marble inlay, two experimental configurations are devised. The first method utilizes the Polar C-Scan System with two distinct transducers, namely an emitter and a receiver (JSR Ultrasonics, DPR500). The two transducers and the vertically placed marble sample are submerged in a water tank, and the emitter sends ultrasonic waves to the sample at a 90° angle in the water medium, while the receiver detects the transmitted signal through the marble. The marble/transducer distances are kept to a few mm to achieve greater precision due to higher undamped amplitudes. Further, the transmitted signal is processed by signal processing software to determine the time of flight and amplitude at every scanned point. The motor-driven arm keeps the same marble/transducer distance during a sweeping operation on the planar surface of the marble while both transducers are mounted on the same arm to ensure equal movement for the emitter and receiver; a planar scan over a 45×30 mm planar surface is performed. The second method employs the Polar C-Scan System with one transducer working as both emitter and receiver (JSR Ultrasonics, DPR500), immersed in the water tank above the horizontally placed marble. Attached perpendicularly to the robot arm, the transducer is set at its focal distance from the marble for maximized precision. The same marble zone is scanned in both c-scan modes, the specific zone was chosen because it contains different materials in a small surface.

Figure 1. Scanned marble zone

3. Reflection (Pulse-Echo) C-scan Processing

3.1 A-scan interpretation

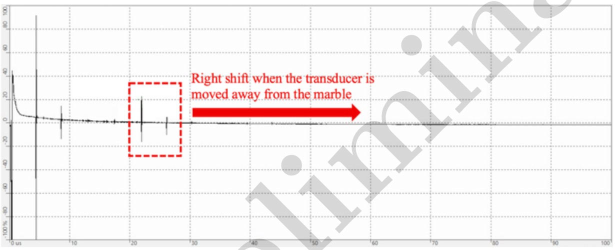

Figure 2. Random point's A scan

When placing the emitting transducer above the marble sample – in echo mode – at a random point, the A-scan of Figure 2 is obtained. This A-scan is roughly identical regardless of the exact coordinates of the robot arm. Changes in the A-scan curve reflecting the surface material (pure marble or precious stone inlay or metal) will only be visible when zooming into the exact fraction of the impacted portion of the signal.

Corresponding author: etuzovic3@gatech.edu

The first task would be to understand the A-scan obtained in Figure 2, and what each of the peaks/reflections represent. When keeping the transducer on the same vertical axis, and doing a pure vertical translation of the transducer, the following behavior is observed:

- The peaks in the red zone have an increase in TOF when the transducer is moved vertically upwards (away from the marble piece) and subsequently move to the right side of the curve.
- All other peaks remain at the same position.

This shows that the two peaks in the red zone represent the first reflection from the marble surface and the second reflection from the backside reflection. The smaller peaks between these two main peaks represent the A-scan inside the marble, which will be our area of interest to analyze the spectra of the different materials covering the surface of the marble piece.

The first few peaks represent the reflections due to the geometry of the transducer itself, and the subsequent smaller peaks after the red zone peaks represent subsequent reflections of the marble piece or transducer [3].

Given our objective to characterize the different surface of the marble, we will focus on understanding what is happening between the two main peaks in the red zone.

3.2 Time-of-flight (TOF) map

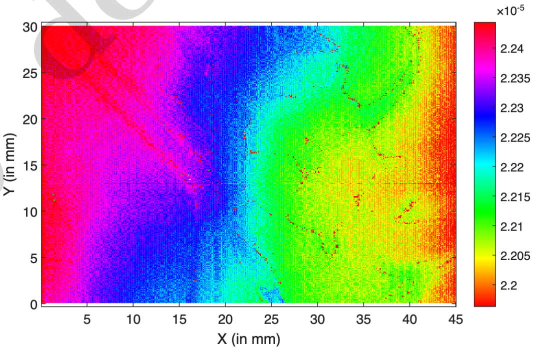

Figure 3. TOF map for C-scan reflection mode

Figure 3 shows that the time of flight is not uniform across areas where we would have expected alignment and a unified color (same zone, same material, same distance). This variation is due to 2 main components:

- Regardless of the precision, there will always be a slight tilt that will be noticeable in the range of microseconds.
- The marble piece doesn't have a perfectly planar surface; some areas are slightly thicker than others.

The color variation is not uniform in a single front/direction, this can be seen by comparing the vertical red color gradient on the right to the green C shaped gradient in the middle of the scan. If tilt was the only parameter for the gradient non uniformity, the color variation would follow the same linear gradient fade pattern through the scan. Using the TOF as a characterizing feature for the different materials will therefore necessitate some alignment to make sure that the signals are comparable.

3.3 Amplitude map

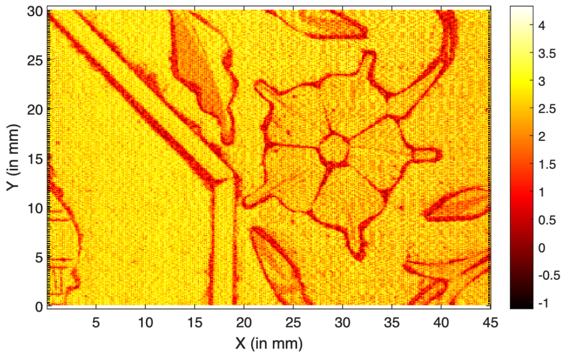

Figure 4. Amplitude map for C-scan reflection mode

The amplitude map is much more robust to slight stand-off (tilt and thickness) variations. The damping resulting from a few microseconds TOF difference doesn't have any notifiable impact on the amplitude profile. Subsequently, the amplitude variations act as a much more significant material differentiator when viewing the surface of the marble, as per the scan Figure 4.

This doesn't however suggest that it can also act as the best differentiator to distinguish between the different materials.

4. Transmission C-Scan Processing

For the transmission scan, a set of transducers with different emitting frequencies were tested, and the frequency resulting in the best scan output was maintained, this was defined as the 5 Mhz emitting transducer. The image doesn't have the exact same layout as the initial marble piece image, this is because in transmission mode, the C-scan mirrors the image horizontally and vertically. We chose to keep the image without mirror processing. This should not however impact the results that were obtained.

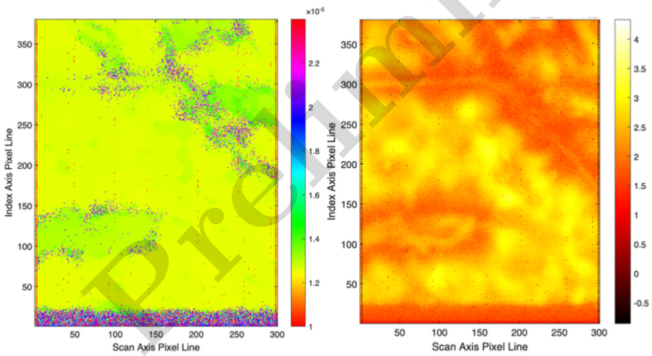

Figure 5. Amplitude and TOF maps for C-scan reflection mode.

The amplitude and TOF scans in the transmission use case highlight different characteristics of the marble piece. Whereas echo mode/reflection highlights mainly characteristics at the surface of the marble piece (reflected echoes, cracks, interfaces, where reflections occur), transmission focuses more on the consistency of the marble (porosity, attenuations, and how much energy passes through). The difference is most visible by comparing the amplitude texture of marble in echo vs transmission modes.

5. Acoustic Material Classification from Ultrasonic C-scan Data

5.1 Overview

The observations above indicate that simple TOF or amplitude maps alone are insufficient for robust material identification. We therefore reformulate the problem as supervised classification of individual A-scans.

In this research, we have designed a classifier for detection of the material contained inside a sub-layer (~2 mm) under the surface of an approximately 8 mm thick slab of marble, which could be pure marble, bluestone, or metal, from ultrasonic A-scans generated using C-scanner. Two acquisition modes have been studied separately, namely, echo and transmission modes. The key physical problem in this regard was that the thickness of the slab is not constant; therefore, the arrival time of the diagnostic wave would vary from one pixel to another.

The baseline approach compared the A-scans to a unique template for each material type based on peak alignment and normalized cross correlation. This worked at a chance level. The diagnosis indicated that the problem was a representation problem and not a modeling one; aligning the A-scans based on their single maximum peak did not work well with variation in thickness, and combining misaligned A-scans into a single template lost its coherence because of incoherent addition.

5.2 Signal Conditioning and Feature-based reformulation

Prior to feature extraction, each echo-mode A-scan is passed through a signal-conditioning pipeline. Each A-scan is surface-aligned, filtered, gated around the region of interest, transformed into temporal and spectral descriptors, standardized, and finally classified using ridge regression and k-nearest neighbors. The following paragraphs describe these steps in detail. The raw signal is first leveled to correct the tilt introduced by the geometry of the transducer. Impulse artifacts are then suppressed by hard clipping: any sample whose magnitude exceeds the threshold, $|x(t)| > 90$, is set to zero, removing saturated acquisition spikes and electrical artifacts without distorting the underlying waveform. The de-spiked signal is next filtered in the frequency domain. Its spectrum $Y(f) = \mathcal{F}\{x(t)\}$ is multiplied by a bandpass mask $M(f)$, defined as

$$M(f) = 1, \quad f_{low} \leq |f| \leq f_{high}$$

$$M(f) = 0, \quad otherwise \qquad (1)$$

with the cut-offs set relative to the Nyquist frequency as

$$f_{low} = 0.05 \cdot \frac{F_s}{2}$$

$$f_{high} = 0.80 \cdot \frac{F_s}{2} \qquad (2)$$

The filtered waveform is recovered through the inverse transform,

$$x_f(t) = \mathcal{F}^{-1}\{Y(f) \cdot M(f)\} \qquad (3)$$

which discards out-of-band noise. Finally, the conditioned signal is cropped to retain only the area of interest, producing the input used to train the classifier.

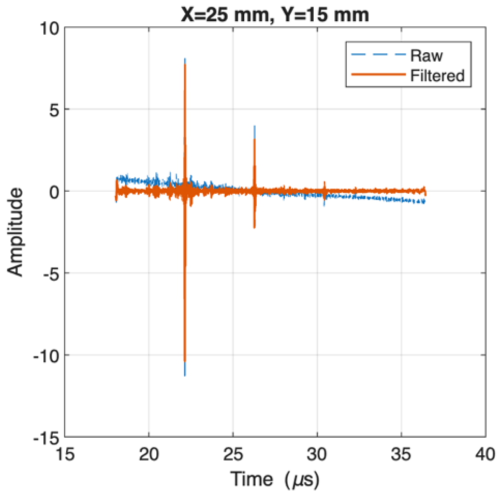

Figure 6. Raw and filtered A-scan signals for a random point on the sample

The pipeline was refitted to work based on delay-invariant characteristics (shape of magnitude spectrum, spectral centroid/bandwidth/rolloff/flatness, band energy ratios, cepstral coefficients, and envelope characteristics) using a regularized linear (ridge) classifier in conjunction with k-nearest-neighbor reference, both coded from scratch in basic MATLAB without any external libraries. Importantly, testing was done using a spatially blocked train-test partition, where training and testing were done using distinct spatial blocks of each region, buffered by a strip, since the use of random partition is prone to information leakage caused by the neighboring near duplicate pixels. Figure 7 shows which area range was selected for each material (used for both training and testing).

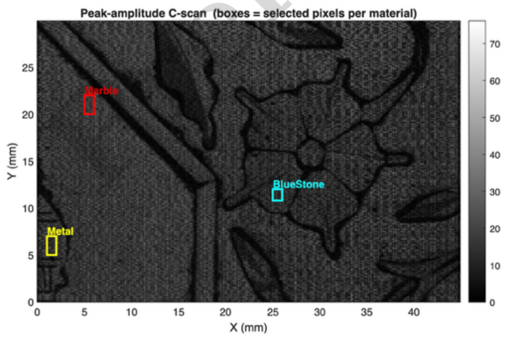

Figure 7. Selected zones for training and testing the classifier

5.3 Surface-referenced gating and modes comparison

Different A-scans are shown in Figure 8. Following the physical nature of the sample, the key methodological move involved relating each measurement to some detectable characteristic of the signal rather than to absolute time. A relevant preliminary move here involved the handling of the reference signals – in the original approach all materials were collapsed to an average A-scan and compared to these three fixed references, which turned out to be a brittle system since one averaged template does not account for the natural pixel-to-pixel variations within a material, as well as being contaminated by any misalignment in the averaging process. It was instead decided that each A-scan should be considered on its own, so each material has a collection of examples, not just one representative one. With regard to echo mode, the front surface of the sample was detected as the first echo for each A-scan and used solely for the time reference, and then a narrow gate was placed at a fixed distance after it (so it moves along the thickness variation), thus capturing the second echo from the embedded layer (~0.35 μs after the first).

The candidate regions have been verified quantitatively via per-sample Fisher discriminant ratio prior to adopting them, thus verifying that separability was indeed localized precisely where the domain reasoning suggested.

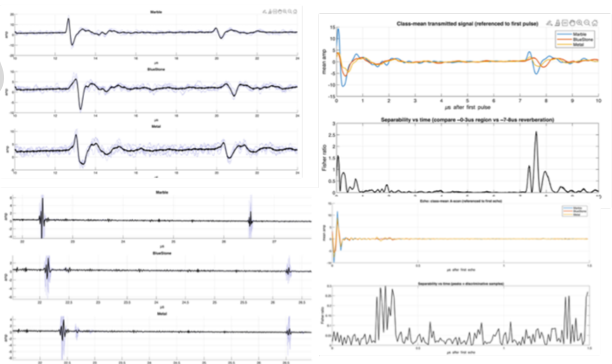

Figure 8. A-scans for different materials - in both modes (echo and transmission) - and corresponding Fischer ratio

For each gated signal, the corresponding feature vector was then reduced to fixed dimensionality, and classification was carried out in the feature space via two complementary classifiers: a regularized linear (ridge) classifier [4], and a k-nearest neighbor (k-NN) classifier [5], assigning to each query signal the majority class among its k nearest neighbors within the standardized feature space (by Euclidean distance). When the aligned gated waveform itself was used as the feature vector, the echo classifier was able to classify with an accuracy of ~99.5%, with all the remaining errors occurring between the most physically difficult materials to distinguish (bluestone and marble, having

very similar impedance). Similarly, the transmission data was analyzed using surface references with a Fisher analysis revealing that the discriminant region was not the first transmitted pulse itself, but a later reverb appearing ~7-8 µs later – the pulse having been reflected three times through the slab. However, a per-pair Fisher analysis revealed a hidden problem with this technique: while marble could be separated with high confidence (~90%), bluestone and metal were inseparable (pairwise separability less than 0.1 for the whole signal).
This represents a physical, not an algorithmic, limit, since transmission is vulnerable to bulk attenuation, while the difference between blue stone and metal lies in the acoustic impedance, which transmission is completely blind to. The echo mode, which interprets the reflected amplitude, will distinguish these two materials; transmission can't do this. Since the two modes rely on different characteristics of the materials, and since the target discrimination depends on impedance, the echo mode is correct here.

6. Conclusion

A classifier for classifying the material contained within the near surface layer of the approximate 8 mm marble slab – pure marble, bluestone, and metal – using ultrasonic C-scan A-scans by pulse-echo and through-transmission scanning approaches has been created. The final echo classifier achieved accuracy of 99.5% based on the stringent spatially blocked test, able to differentiate all three materials, while the transmission approach could only distinguish the marble from other two materials. This is because the success was less related to the classifier used and more to the matching representation of the signal with the underlying physics of the measurement.
Transmission is responsive mainly to bulk attenuation, so that it separates the marble from the embedded materials, but it does not distinguish between bluestone and metal, whose unique characteristic – acoustic impedance – is not revealed by the transmitted wave. Echo detects the amplitude of the reflection from the embedded interface, determined directly by the impedance mismatch, and thus has access precisely to the information that transmission lacks. Consequently, the two modes are complementary detectors of different physical characteristics, and the echo is the right choice in cases like this one, where impedance determines the distinction to be made. Overall, the results show that physically-informed signal gating and a representation matched to a simple classifier are enough to achieve near-perfect material discrimination — and, equally importantly, to recognize when an acquisition mode is physically incapable of the required distinction

7. Acknowledgments

The authors acknowledge support from the French Lorraine Université d'Excellence (LUE, project I-META ANR-15-IDEX-04-LUE).

References

[1] T. Weiss, P. N. J. Rasolofosaon, and S. Siegesmund, "Ultrasonic wave velocities as a diagnostic tool for the quality assessment of marble," Geological Society, London, Special Publications, vol. 205, pp. 149–164, 2002.

[2] M. El Boudani, N. Wilkie-Chancellier, L. Martinez, R. Hébert, O. Rolland, S. Forst, V. Vergès-Belmin, and S. Serfaty, "Marble characterization by ultrasonic methods," Procedia Earth and Planetary Science, vol. 15, pp. 249–256, 2015.

[3] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, Fundamentals of Acoustics, 4th ed. New York, NY, USA: John Wiley & Sons, 2000.

[4] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," Technometrics, vol. 12, no. 1, pp. 55–67, 1970.

[5] P. Cunningham and S. J. Delany, "k-Nearest neighbour classifiers – A tutorial," ACM Computing Surveys, vol. 54, no. 6, art. 128, pp. 1–25, 2021.