

Mechanical Spectroscopy of Materials using Atomic Force Microscopy (AFM-MS)

M. Petrov¹, D. Canena¹, N. Kulachenkov¹, N. Kumar¹, Pierre Nickmilder², Philippe Leclère²,
Igor Sokolov^{1,3,*}

¹ Departments of Mechanical Engineering, Tufts University, Medford, MA 02155, USA

² Laboratory for Physics of Nanomaterials and Energy, Institute for Materials Science and Engineering, University of Mons, Mons, Belgium

³ Departments of Physics and Biomedical Engineering, Tufts University, Medford, MA 02155, USA

* Corresponding author, igor.sokolov@tufts.edu

Abstract

Here, we present a novel mechano-spectroscopic atomic force microscopy (AFM-MS) technique that overcomes the limitations of current spectroscopic methods by combining the high-resolution imaging capabilities of AFM with machine learning (ML) classification. AFM-MS employs AFM operating in sub-resonance tapping imaging mode, which enables the collection of multiple physical and mechanical property maps of a sample with sub-nanometer lateral resolution in a highly repeatable manner. By comparing these properties to a database of known materials, the technique identifies the location of constituent materials at each image pixel with the assistance of ML algorithms. We demonstrate AFM-MS on various material mixtures, achieving an unprecedented lateral spectroscopic resolution of 1.6 nm. This powerful approach opens new avenues for nanoscale material study, including the material identification and correlation of nanostructure with macroscopic material properties. The ability to map material composition with such high resolution will significantly advance the understanding and design of complex, nanostructured materials.

Keywords: NanoSpectroscopy, Material Identification, Composite Materials, Machine Learning, Atomic force microscopy (AFM), Polymer composites.

Introduction

Identifying material composition at the nanoscale remains a critical challenge in materials science. Accurate recognition of nanostructures is essential for the development of new nanocomposite materials, for uncovering structure-property relationships¹⁻⁵. This understanding is instrumental in the innovation of new composite materials, especially nanocomposites, which are integral to advancements in technology and industry⁶⁻¹⁰. Broad utilization of polymeric materials brings another challenge to understanding microplastic contamination^{11,12}. The spectroscopy capable of identifying the types of plastics down to the nanoscale, i.e., at the very early stages, will be important in this area as well¹³⁻¹⁶.

Traditional spectroscopic techniques such as infrared (IR) spectroscopy^{17,18}, Raman spectroscopy¹⁹, X-ray photoelectron spectroscopy (XPS)²⁰, and micro nuclear magnetic resonance (micro-NMR) spectroscopy²¹ have been the cornerstone for material identification by exploiting unique material properties. However, these methods encounter resolution limitations that preclude material identification at the nanoscale. Optical spectroscopic methods are convenient but are inherently limited by the diffraction limit, while XPS offers exceptional height resolution but lacks lateral resolution.

Recent advancements in spectroscopic methods have begun to bridge the gap to nanoscale resolution. Infrared absorbance spectroscopy, when coupled with atomic force microscopy (AFM-IR), achieves high resolution, albeit on thin films due to its reliance on heating a multi-micron sample area with IR light, while a sharp AFM probe detects the thermal expansion of the sample²²⁻²⁵. AFM-IR is also constrained to ambient conditions and is unsuitable for aqueous environments. Its relatively slow operation limits material identification to specific locations rather than enabling comprehensive spectroscopic imaging. Moreover, resolving interfaces between polymers with differing thermal expansion coefficients—a critical aspect in nanocomposite studies—remains a significant challenge when using AFM-IR because the interface is moving laterally at the moment of heating due to the different heat expansions of the materials of the interface. Tip-enhanced

Raman spectroscopy (TERS) has also emerged as a high-resolution spectroscopic method ²⁶⁻²⁹, yet its application has been largely confined to isolated structures, such as carbon nanotubes or single DNA molecules, due to its complex setup and limited substrate versatility ³⁰. Finally, both AFM-IR and TERS suffer from the risk of changing samples due to excessive laser radiation.

Here, we introduce an innovative AFM mechano-spectroscopic imaging technique (AFM-MS) that uses AFM operating in a sub-resonance tapping ³¹⁻³³, specifically PeakForce QNM version ³³⁻³⁵ with the extension Ringing mode (RM) ^{32,36-39}. RM yields multidimensional data of the mechanical and physical characteristics of a sample surface with unparalleled resolution down to the single nanometer scale. This mode captures a spectrum of mechanical and physical properties, such as the dimensions of material necks formed by the AFM probe during disconnection, the length of molecules pooled from the surface by the AFM probe, the energy dissipated upon the probe detachment, etc. Each RM and PeakForce QNM channel serves as a distinct map, delineating the distribution of specific physical properties across the sample surface. All RM channels can be recorded simultaneously on top of several channels of PeakForce QNM mode. This provides a comprehensive profile of the physical and mechanical properties of the surface recorded at each image pixel. In our study, we demonstrate the use of twelve different surface properties, see the Experimental section for detail.

AFM was first proposed as a spectroscopic technique for distinguishing different materials back in 1994 by measuring the van der Waals component of the force curve ⁴⁰. However, the limited number of distinguishable materials ⁴¹ and the challenges associated with measuring small van der Waals forces hindered the widespread adoption of this approach. Subsequently, piezoelectric force microscopy (PFM) was employed to differentiate between various bacterial species ^{42,43}.

Nevertheless, this method is limited to materials exhibiting piezoelectric properties, which are relatively uncommon, particularly in polymers. In contrast to these earlier techniques, the AFM-MS technique described herein can be applied to a wide range of materials that undergo deformation and exhibit adhesion under an AFM probe without requiring additional stimuli such as electrical voltage or infrared heating of the sample. In this study, we demonstrate the effectiveness of this technique through the identification of binary and ternary material blends. We also present a comparative analysis with confocal Raman microscopy to highlight the strengths and potential limitations of the AFM-MS approach.

Results and discussion

Nano mechano-spectroscopy (AFM-MS) approach

We demonstrate the method using a complex blend of three polymers. **Figure 1** shows a schematic of the suggested spectroscopy. Images (maps) of different physical/mechanical properties of known polymers (Fig.1a) are used to train an ML algorithm, which is then used to identify polymers in the composite sample. Since the data size of these images can be quite large (every single pixel of the image is a data point), there are no restrictions related to the small data size problem, which is typical for the ML processing of AFM images⁴⁴. Both regular and deep-learning ML methods can be used. Further, we will use supervised classification methods. Note that unsupervised methods, like principal component analysis (PCA), can be used for some particular cases. However, it is not working well in general, including the example of three polymers shown here (SI Figure S1).

Two approaches are proposed for working with the data, which differ in the databases used for training the ML algorithm (Fig.1b). In Method I, the training is done using AFM images of pure polymers, while in Method II, the training is done using information from known locations of known polymers within the composite samples of the study. Both methods are used in all material spectroscopies. Method I is more generic and allows building a broad database of nanomechanical properties of polymers. Such a database will allow the identification of specific polymers within a sample that contains unknown polymers. Although Method II seems less generic, it has multiple advantages, which will be discussed later. The machine learning part is similar for both approaches. The trained ML algorithm is applied then to each pixel of an unknown sample (Method I) or unknown parts of the sample (Method II). Depending on the type of classifier, we can identify either the polymers or the probability of belonging to the particular polymer at each image pixel. The neural networks and decision trees used here presented probabilities for each of the three polymers, which were encoded by RGB colors in Figure 1b.

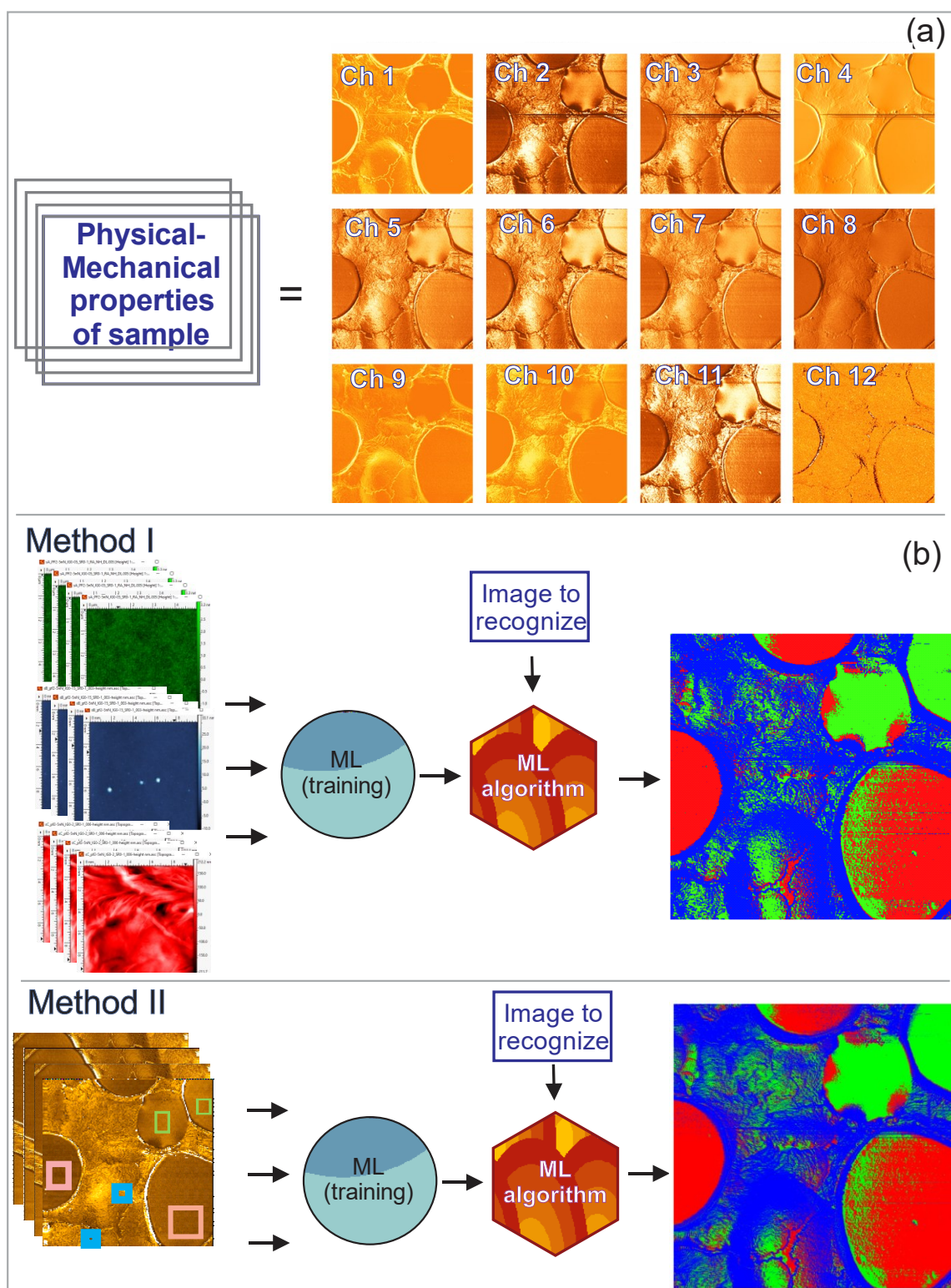


Figure 1. A schematic of the nano mechano-spectroscopy. (a) Physical and mechanical properties of the samples are imaged simultaneously using 12 different imaging channels (see SI Figure S2 for detail). An example of a $5 \times 5 \mu\text{m}^2$ blend of three different polymers, polystyrene (PS), polyethylene

oxide (PEO), and polyvinyl pyrrolidone (PVP) is presented. (b) Two methods of material recognition are shown. Method I: The training of the machine learning algorithm is done using the AFM images of the corresponding pure polymers prepared in the same conditions. Method II: The training of the machine learning algorithm is done using the areas of known polymers within the sample of study (shown on the image with rectangles of different colors). Note that the AFM data can be used directly for the training of ML algorithms without any preprocessing. An example of the application of both methods to the blend of three different polymers is shown. Three different colors were used to identify particular polymers: PS is red, PVP is green, and PEO is blue. Two machine learning algorithms, (shallow) Neural Networks and Random Forest, were used here and showed identical results.

Comparing the results of both methods, one can see very similar results. Nevertheless, there are some differences in the complex blend of PEO and PVP polymers (blue and green). This is attributed to the different times the mix and pure polymer (needed for Method I) samples were stored after preparation. It demonstrates the limitation of Method I, the need for strict control of the sample preparation, including environmental factors when creating the database of materials for training. Both PEO and PVP are sufficiently hydrophilic and, therefore, for example, sensitive to environmental humidity. Method II, on the other hand, is more robust, as all parts of the sample were prepared simultaneously under the same conditions, and images were collected in the same environment during scanning.

Table 1 describes a more detailed comparison of these two methods. For example, the imaging environmental conditions and sample preparation must be precisely the same as those used to collect the external database of polymers. This is because the physical and mechanical characteristics of polymers can change substantially when those parameters are not kept the same. For example, the mechanical properties of polymers can strongly depend on temperature. Hydrophilic polymers are sensitive to humidity, etc. The AFM imaging can result in different values of the recorded sample properties. For example, the map of adhesion between the AFM probe and sample depends upon the adhesion energy but also on the viscosity of the polymer and, therefore, on the time of the probe-sample contact, i.e., the speed of tapping⁴⁵. Finally, the roughness of the sample surface created by different sample preparation methods can result in

substantially different values of the recorded sample properties. For example, high roughness will result in lower stiffness of the material at small indentation depths ⁴⁶.

Table 1. Comparison of two approaches for identification of components in a sample. Method I uses an existent database of nanomechanical properties of pure components (polymers). Method II uses the data collected on the part of known locations of particular pure components (polymers) within the AFM image of the study.

Requirement	Method I	Method II
Knowledge of individual (polymeric) components of the sample	Not required (optional)	Required
The external database of components (polymers)	Required	Not required
AFM imaging under a controlled environment	Frequently required	Not required
AFM imaging is restricted to the same imaging conditions as used in the external material (polymer) database	Required	Not required
Sample preparation is limited to the one used for the external material database	Required	Not required

The advantages and limitations of both methods are not unique to the described mechano-spectroscopy. For example, databases of materials exist in Raman spectroscopy ⁴⁷. When dealing with a blend of different polymer phases, in particular, at a small scale (Raman microscopy), the existing database can usually be used reliably only for a qualitative evaluation of the material types. A possible discrepancy with the pre-existing Raman database of polymers is also caused by environmental factors (temperature, humidity) and sample preparation (different material processing, residual solvents, etc.). Therefore, Method II is rather frequently used in Raman microscopy. Another example of Method II is described in ref. ²⁰, in which bacterial cells were studied. Biological cells are susceptible to the environment. Therefore, it seems to be impractical to implement Method I when studying such an environment-dependent object.

Figure 2 demonstrates the details of the work of the trained ML algorithms for material recognition. Two samples comprising the same three polymers (PEO, PVP, PS) were prepared: a freshly prepared blend of the polymers and the same sample aged for two weeks after preparation.

When applied to each pixel of the image, the algorithm calculates the probability that the pixel belongs to each particular type of material/polymer. The left part of the image demonstrates the maps of the distribution of these three probabilities separately for each of the three polymers. The rightmost images of Figure 2 show an RGB combination of all three probabilities, in which red, green, and blue channels correspond to each of the three polymers, respectively. One can see that the freshly prepared sample demonstrates a clear separation between the three polymers. The same sample after two weeks shows a clear inter-diffusion of polymers. This is expected because of the hydrophilic nature of PVP and PEO polymers, which were exposed to environmental humidity during storage. It demonstrates the ability of the presented technique to study the surface diffusion of different materials.

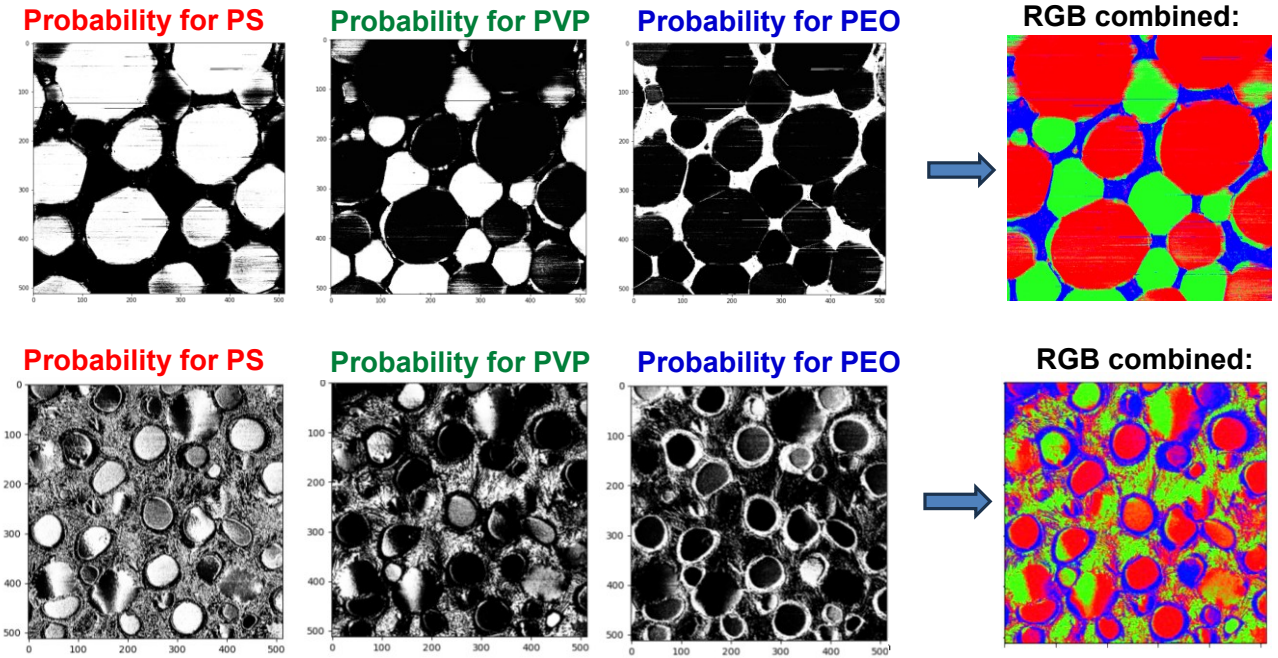


Figure 2. Details of the work of machine learning classification algorithms. The trained ML algorithm calculates the probability of each pixel representing a particular material. RGB combined images: The probabilities are encoded by RGB colors and combined to present the distribution of particular polymers over the sample. Two samples of $20 \times 20 \mu\text{m}^2$ are shown. The top row: A freshly prepared blend of three polymers (PS, PVP, PEO). The bottom row: The same sample two weeks after the preparation; polymer diffusion makes the sample more challenging for recognition.

It is interesting to note that AFM-MS is rather tolerable to horizontal streak artifacts. These artifacts are clearly seen in channels 2-4 in Fig.1a and the top probability images of Fig.2 as horizontal lines. These artifacts are typically caused either by vibrational noise or a small piece of dirt picked up by the AFM probe (which is quickly self-cleaned). It is possible to eliminate these artifacts or at least minimize them by multiple imaging of the same area. However, it has not been done on purpose to demonstrate this feature of AFM-MS, tolerance to these artifacts. After combining multiple channels with a machine learning algorithm, the streak artifacts either substantially vanish or completely disappear in the classification image, as shown in Figs. 1 and 2 (RGB images).

The lateral spectroscopic resolution of the presented technique can theoretically be defined by the area of contact of one single AFM touch. The area, in turn, is defined by the radius of the AFM probe, load force (indentation depth), the adhesion energy between the AFM probe and sample, and the speed of indenting. Certainly, the two materials should be sufficiently different from the physical/mechanical points of view. The radius of curvature of the AFM probe used in this work is of the order 2-3 nm, whereas the indentation depth of the polymers used in this work is not more than 2-4 nm. Thus, the lateral resolution of the described technique is expected to reach the level of single nanometers.

Although the spectroscopic resolution of AFM-MS is theoretically defined as a single-pixel resolution, it can be quite misleading due to the variation of the measured properties across each material. As demonstrated in SI Figure S3, there is an illusion that a human eye can distinguish between PEO, PS, and PVP polymers even in single channels. However, the histogram of the distribution of physical properties (SI Figure S4) reveals significant overlap among the properties of the three polymers. Consequently, relying on a single channel for spectroscopic separation would result in incorrect identification in many pixels. To achieve nearly unambiguous identification of the specific polymer, a combination of multiple channels is necessary.

It is instructional to demonstrate the spectroscopic resolution using mechanically distinctive materials. **Figure 3** demonstrates the spectral resolution of ~ 1.6 nm by distinguishing silica material of a nanoparticle embedded in epoxy resin. This is in agreement with the above theoretical

estimations. Presumably, the value of this resolution is limited by the radius of curvature of the AFM probe.

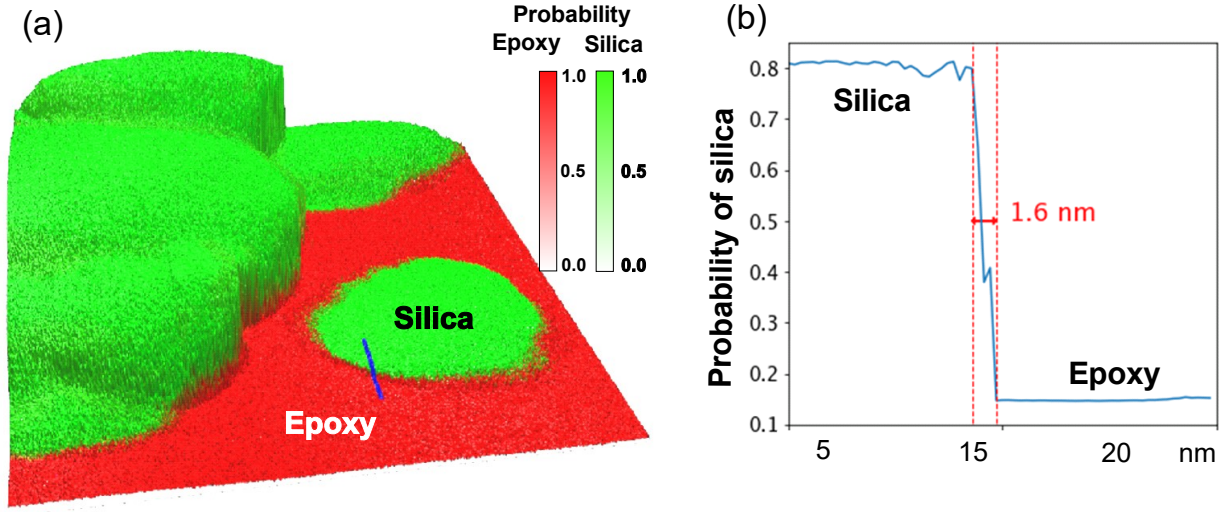


Figure 3. Experimental demonstration of the lateral spectral resolution of the described nano-mechano-spectroscopy. (a) 3D height AFM image of 200x200 nm² is shown. The probabilities of identification of material as epoxy (red) and silica (green) are superimposed on the height image using different colors. (b) A cross-section of the probability of identifying material as silica is shown. The specific location of this cross-section is shown in panel (a) by a solid blue line. The width of the probability change of 1.6 nm, which can be called the lateral spectral resolution²³, is shown.

It is useful to compare the described AFM-MS with another technique, confocal Raman spectroscopy, which is broadly used to identify the type of materials. Figure 4a shows an example of a 4 μm polystyrene bead partially embedded in epoxy resin. The lateral resolution of Raman microscopy is limited by the diffraction limit of ~ 200 nm. However, the actual lateral resolution can be confusing because the Raman signal comes from a noticeable thickness of the sample (Fig.4 b). For example, AFM identifies only 2 μm of the polystyrene bead exposed above the epoxy (Fig.4 d,f)). However, Raman microscopy detects the signal from polystyrene of the full size of 4 μm (Fig.4 e,f). This shows the substantial difference between the Raman technique and the described mechano-spectroscopy. The AFM-MS provides substantially higher both lateral and vertical resolutions over the sample surface. At the first-order approximation, one can say that the vertical

resolution of AFM-MS can be approximated by the depth of the sample deformation by the action of the AFM probe, which is 2-4 nm in the considered examples. It comes at the cost of not being able to detect materials hidden beneath the surface. The vertical resolution of Raman microscopy is hard to estimate because it is defined by many parameters, including the type of material, numerical aperture of the objective, and magnification. Realistically, it could be estimated as $\sim 0.5 \mu\text{m}$.

The synergy between Raman confocal spectroscopy and the described AFM-MS should be noted. Large parts of the sample can be identified with Raman spectroscopy, as shown in Figure 4. It can directly be used to train the classification algorithms of Method II. AFM-MS applied to such a sample will then reveal the location of particular materials on the sample surface with nanoscale accuracy.

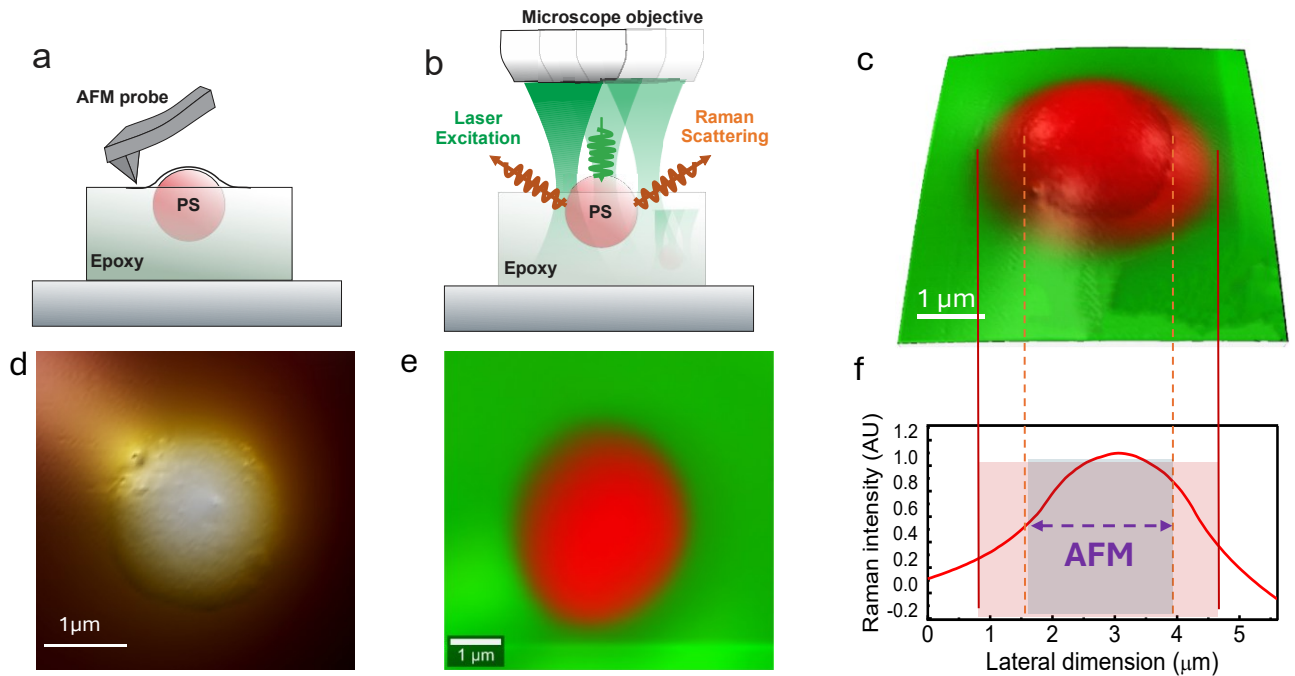


Figure 4. Comparison with confocal Raman microscopy/spectroscopy and mechano-spectroscopy: an example of a polystyrene sphere partially embedded in epoxy. (a) A schematic of the sample. The AFM probe can easily image the outstanding part of the polystyrene bead. Substantial differences in roughness and adhesion (not shown) allow easy identification of the bead above the epoxy. (b) A diagram showing the work of Raman microscopy. The finite depth of focus of the laser excitation light may result in a Raman signal coming from both polystyrene and epoxy. (c) A 3D AFM image of the

polystyrene sphere embedded in epoxy (shown in panel d) superimposed with the map of the Raman signal (shown in panel e). (f) Shows the size of the outstanding part of the polystyrene bead as well as the location of the detectable Raman signal associated with polystyrene.

It should be noted that the idea of combining multiple physical properties to resolve a very large number of materials comes from “artificial nose”. It was demonstrated that by combining a limited number of sensors (“receptors”) specific to particular molecules, one could obtain a combinatorial sensor that can distinguish a large number of different stimuli/smells ⁴⁸⁻⁵⁰. Therefore, we expect the nano mechano-spectroscopy technique to be able to distinguish a very large number of polymers, though we demonstrated the approach using only three polymers. The limits of the utility of the described AFM-MS are yet to be determined.

Conclusions

This study introduces an innovative nano mechano-spectroscopy technique (AFM-MS) that utilizes physical and mechanical properties of materials to identify the materials. It allows the locating of specific material components within a composite sample up to a single nanometer spatial resolution. In this respect, the new spectroscopy surpasses the capabilities of existing techniques used for material identification. The unprecedented resolution of this spectroscopy is expected to have wide-ranging applications in material science and engineering, encompassing polymers, biomaterials, nanocomposites, and other complex materials. Furthermore, the AFM-MS technique is a relatively fast imaging technique, which allows obtaining the spectroscopic images of a sample surface with the speed of a regular sub-resonance AFM tapping. Compared to other techniques, AFM-MS is rather versatile; it can be used both in air and in a liquid medium with a minimum nondestructive energy exchange with the sample study. The unique information obtained through this technique can be synergistically combined with other characterization methods, such as confocal Raman microscopy, to provide a more comprehensive analysis of materials.

Experimental section

Materials

Polystyrene (PS), polyethylene oxide (PEO), and polyvinyl pyrrolidone (PVP) were purchased (all by SigmaAldrich, Hoeilaart, Belgium) and used as is. The polymer blends were prepared as follows. All polymer solutions were prepared in Chloroform (10mg/mL). The stock solution was then blended with 1:1:1 ratio. The samples for the AFM-MS study were then prepared by drop casting ($500 \mu\text{L}/\text{cm}^2$) on a freshly cleaned Si wafer and then allowed to dry for 24 hours with a solvent annealing process (a small beaker of Chloroform placed in the vicinity of the drying samples in a closed chamber). The samples with embedded particles were prepared as follows ⁵¹. Washed and dried $4 \mu\text{m}$ polystyrene beads (BaseLine Chromtech, Tianjin, China) were spread over semi-rigidified epoxy (Loctite EA E-30CL, Henkel Corporation, CT, USA). Silica nanoparticles (Pureon Inc., NC, USA) were partially embedded in epoxy (Loctite EA E-30CL, Henkel Corporation, CT, USA) as follows. The silica nanoparticles were freeze-dried to obtain a powder. A layer of epoxy was applied to a glass slide, which was allowed to semi-solidify. Once the epoxy reached a semi-solid state, the silica nanoparticle powder was evenly spread over it. The sample was then left to solidify thoroughly for 48 hours.

Atomic force microscopy

Atomic force microscopes Icon and Bioscope Catalyst (Bruker Nano, Inc., Santa Barbara, CA) working in PeakForce QNM mode with Ringing mode extension (NanoScience Solutions, Inc., Woburn, MA) were used. ScanAsyst AFM probes for working in air (by Bruker Nano, Inc., Santa Barbara, CA) were used. The radius of the AFM probe was found by using TipCheck sample (BudgetSensors, Sofia, Bulgaria). It was in the range between 3 to 4 nm. The spring constant of the cantilever was measured using the thermal tuning method, and it was in the range between 0.3 – 0.4 N/m. The scanning was done in air in humidity between 30 to 40%. The scanning speed was 0.1-0.3 Hz. The PeakForce was chosen to be 1-2 nN. The speed and the amplitude of PeakForce tapping were set to 1 kHz and 300 nm, respectively. The following channels were recorded and used in the described method: PeakForce QNM: Adhesion, Deformation, Dissipation, Peak Force; Ringing mode: Restored adhesion, RM Adhesion, RM Peak Force, Viscoelastic adhesion, Neck height, Disconnection length, Energy loss during disconnection, Dynamic phase shift.

Machine learning algorithms

First, let us describe the structure of the data obtained with AFM on a sample. Each AFM scan M consists of K channels $M^{(k)}$:

$$M = \{ M^{(1)}, \dots, M^{(K)} \}, \text{ where } k=1..K. \quad (1)$$

Each k^{th} channel is represented by an image which consists of $N \times N$ pixels located at Cartesian coordinates (x_i, y_j) . So,

$$M^{(k)} = \bigcup_{i=1}^N \bigcup_{j=1}^N \{ z_{i,j}^{(k)} \}, \quad (2)$$

where $z_{i,j}^{(k)}$ is the value of the k^{th} channel at pixel (x_i, y_j) , $i, j = 1..N$. Since all channels are recorded at the same pixels, the coordinates of the pixels do not carry the channel index.

So each AFM scan now can be written as

$$M = \bigcup_{i=1}^N \bigcup_{j=1}^N \{ \{ z_{i,j}^{(1)} \}, \dots, \{ z_{i,j}^{(K)} \} \}. \quad (3)$$

Different AFM scans will be presented using an index s , where $s=1..S$, S is the number of different scans. Usually different scans are collected on different areas of the sample. In this case, the database recorded on the sample reads as

$$M = \bigcup_{s=1}^S \bigcup_{i=1}^N \bigcup_{j=1}^N \{ \{ z_{i,j}^{(1;s)} \}, \dots, \{ z_{i,j}^{(K;s)} \} \}. \quad (4)$$

In some cases it is useful to consider only a small portion p of the total image restricted by a specific choice of pixels $\{i, j; p\}$. For example, it can be a zoomed area specific to a particular material. This portion of the database M_p reads as follows:

$$M_p = \bigcup_{s=1}^S \bigcup_{i=1}^N \bigcup_{j=1}^N \{ \{ z_{i,j;p}^{(1;s)} \}, \dots, \{ z_{i,j;p}^{(K;s)} \} \}. \quad (5)$$

In the machine learning terminology, vector $\{ \{ z_{i,j;p}^{(1;s)} \}, \dots, \{ z_{i,j;p}^{(K;s)} \} \}$ for specific i, j, s, p is a feature of the instance dataset. To train a classifier in supervised machine learning, each feature should be assigned a particular class $C^{(l)}$, where $l=1..L$; L is the number of classes in consideration. So the database D used for building the classifier has a class assignment for each feature on the database

$$D = \bigcup_{s=1}^S \bigcup_{i=1}^N \bigcup_{j=1}^N \{ \{z_{i,j}^{(1;s)}\}, \dots, \{z_{i,j}^{(K;s)}\}; C_{i,j}^{(l)} \} \quad (6a)$$

or

$$D_p = \bigcup_{s=1}^S \bigcup_{i=1}^N \bigcup_{j=1}^N \{ \{z_{i,j;p}^{(1;s)}\}, \dots, \{z_{i,j;p}^{(K;s)}\}; C_{i,j}^{(l)} \}, \quad (6b)$$

dependent on the case of equations 4 or 5.

Method 1:

In this method, the classification algorithm is created on samples containing datasets of pure polymers. In the case of three polymers ($L=3$) analyzed in this work, the total dataset is $\{D^{PS}, D^{PVP}, D^{PEO}\}$, where $D^{polymer}$ is the database for each particular class (polymer) described by equation 6a. The number of images per sample, $S=1$, was sufficient due to the homogeneity of the samples.

The whole dataset was split into training and validation sets. The validation set was typically around 10-20% of the entire dataset. A classifier CL was trained on the training set and then tested on the validation set to ensure that the classifier was not overtrained and could distinguish each pure polymer well. The trained this way classifier CL was then applied to every pixel (i,j) of the polymer blend AFM image-dataset M^{mix} (given by equation 4). The obtained vector of probabilities was the result of application of the classifier:

$$\{ \text{Proba}^{PS}, \text{Proba}^{PVP}, \text{Proba}^{PEO} \} = CL[\bigcup_{s=1}^S \bigcup_{i=1}^N \bigcup_{j=1}^N \{ \{z_{i,j}^{(1;s)}\}, \dots, \{z_{i,j}^{(K;s)}\} \}]. \quad (7)$$

In the considered example, $S=1$, $N=512$ (or 256), $K=12$.

Method 2:

Here, to develop a classifier to be applied to the unknown mixed sample, we used a zoomed area in which the known polymers were located. In contrast with Method I, the used database is now described by equation 6b, in which $p = \{PS, PVP, PEO\}$. (An example of a particular location of those zoomed areas is shown in Figure 1.) Similarly, trained classifier CL is then applied for the rest of the image(s) to obtain probabilities of each pixel to belong to a particular class:

$$\{ \text{Proba}^{PS}, \text{Proba}^{PVP}, \text{Proba}^{PEO} \} = CL[\bigcup_{s=1}^S \bigcup_{i=1}^N \bigcup_{j=1}^N \{ \{z_{i,j;p}^{(1;s)}\}, \dots, \{z_{i,j;p}^{(K;s)}\} \}], \quad (8)$$

where $\bar{p}$ is the indicator of the pixels complementary to all used zoomed areas, i.e, all pixels of the image with the exception of the zoomed areas used to build a classifier.

In the considered example, $S=1$, $N=512$ (or 256), $K=12$.

In both of these methods, the pixels of the analyzed sample are assigned a vector of classification probabilities exemplified by equations 7 and 8. The dimension of this vector equals to the number of components in the sample of study. In the example considered here, we deal with three components/polymers. Three dimensions (or less) can be conveniently represented by the RGB color model, and the result of classification can be immediately plotted using the so-called classification map, which is used in the figures of this work.

The classifiers used in this work were either a shallow Neural Network or Random Forest. A Python code of the used classifier is shown in the Supplementary materials. Random forest as implemented within Scikit-learn python package and the Neural Networks were built using PyTorch package. A variety of shallow Neural Network architectures were considered for classification. The final architecture consisted of two fully connected hidden layers with 7 and 5 neurons, respectively. The input number of neurons depended on the number of AFM channels used for classification.

Data availability

Raw data are available from the corresponding author upon request; an example of data and the processing ML code (in Python) are uploaded as Supplementary data.

Acknowledgments

I.S. acknowledges support from NSF CMMI 1937373, 2224708 and NIH R01 CA262147 grants. The confocal Raman microscope was purchased with the help of the NSF CBET 1428919 grant. The research in Mons was supported by the F.R.S. – FNRS Grands Equipments (40007941) – 2022.

Author contributions

Mikhail Petrov: Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization. **Daniel Canena:** Investigation, Formal analysis, Visualization. **Nikita Kulachenkov:** Investigation, Formal analysis, Visualization. **Nishant Kumar:** Investigation, Formal analysis, Visualization. **Pierre Nickmilder:** Investigation. **Philippe Leclère:** Conceptualization, Methodology, Resources, Funding acquisition, Supervision, Writing – review & editing. **Igor Sokolov:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing, Resources, Funding acquisition, Project administration, Supervision.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/...>

Online content

Methods (including Python code), source data, supplementary information.

References

1. Scherf, U., and List, E. J. W., *Advanced Materials* (2002) **14** (7), 477
2. Jancar, J., *et al.*, *Polymer* (2010) **51** (15), 3321
3. Le, T., *et al.*, *Chem Rev* (2012) **112** (5), 2889
4. Cubuk, E. D., *et al.*, *Science* (2017) **358** (6366), 1033
5. Yao, Y. G., *et al.*, *Science* (2022) **376** (6589), 151
6. Zeng, Y. W., *et al.*, *Nature* (2022) **602** (7895), 91
7. Klonos, P. A., *et al.*, *Polymer* (2024) **296**
8. Toader, G., *et al.*, *Polymers-Basel* (2024) **16** (4)
9. Wongvasana, B., *et al.*, *Polymers-Basel* (2022) **14** (18)
10. Zhang, M. X., *et al.*, *Angew Chem Int Edit* (2019) **58** (48), 17412
11. Nava, V., *et al.*, *Nature* (2023) **619** (7969), 317
12. Wang, Y. Q., and Wang, Y. X., *Heliyon* (2024) **10** (12)
13. Cao, X. S., *et al.*, *Acs Nano* (2024)

14. Fang, C., *et al.*, *Nano Express* (2024) **5** (2)
15. Muthusubramanian, B., *et al.*, *Environ Sci Pollut R* (2023) **30** (49), 107533
16. Ozden, S., *et al.*, *Mater Today* (2022) **59**, 46
17. Shi, L. X., *et al.*, *Nature Methods* (2020) **17** (8), 844
18. Nasse, M. J., *et al.*, *Nature Methods* (2011) **8** (5), 413
19. Cheng, J. X., and Xie, X. S., *Science* (2015) **350** (6264)
20. Krishna, D. N. G., and Philip, J., *Appl Surf Sci Adv* (2022) **12**
21. Allert, R. D., *et al.*, *Chem Commun* (2022)
22. Dazzi, A., and Prater, C. B., *Chem Rev* (2017) **117** (7), 5146
23. Wang, L., *et al.*, *Sci Adv* (2017) **3** (6)
24. Piergies, N., *et al.*, *Nano Res* (2020) **13** (4), 1020
25. Wang, Z. Q., *et al.*, *Macromolecules* (2019) **52** (24), 9639
26. Yeo, B. S., *et al.*, *Small* (2009) **5** (8), 952
27. Pandey, Y., *et al.*, *Nanoscale* (2024) **16** (22), 10578
28. Cai, Z. F., *et al.*, *Angew Chem Int Edit* (2024)
29. Filez, M., *et al.*, *Advanced Materials* (2024) **36** (5)
30. Bailo, E., and Deckert, V., *Chemical Society Reviews* (2008) **37** (5), 921
31. Smolyakov, G., *et al.*, *Carbohydr Polym* (2016) **151**, 373
32. Dokukin, M. E., and Sokolov, I., *Scientific reports* (2017) **7**, 11828
33. Dokukin, M. E., and Sokolov, I., *Langmuir* (2012) **28** (46), 16060
34. Efremov, Y. M., *et al.*, *Soft Matter* (2019) **15** (27), 5455
35. Li, L., *et al.*, *Nanoscale* (2019) **11** (11), 4707
36. Prasad, S., *et al.*, *Advanced NanoBiomed Research* (2021) **1** (8), 2000116
37. Tilinova, O. M., *et al.*, *Cells* (2024) **13** (4)
38. Petrov, M., and Sokolov, I., *Biomedicines* (2023) **11** (1)
39. Prasad, S., *et al.*, *Advanced Nanobiomed Research* (2021) **1** (8)
40. Sokolov, I. Y., *Surface Science* (1994) **311** (3), 287
41. Sokolov, I., *J VAC SCI TECHNOL A* (1996) **14** (5), 2901
42. Thompson, G. L., *et al.*, *Nanotechnology* (2012) **23** (24)
43. Nikiforov, M. P., *et al.*, *Nanotechnology* (2009) **20** (40)
44. Sokolov, I., *Physical Chemistry Chemical Physics* (2024) **26** (15), 11263

45. Sweers, K. K. M., *et al.*, *Nanoscale* (2012) **4** (6), 2072
46. Liu, Y. K., *et al.*, *Nanoscale* (2020) **12** (23), 12432
47. Vandenabeele, P., *et al.*, *Chem Rev* (2007) **107** (3), 675
48. Malnic, B., *et al.*, *Cell* (1999) **96** (5), 713
49. Bushdid, C., *et al.*, *Science* (2014) **343** (6177), 1370
50. C. Bushdid, *et al.*, *Science* (2024) **383** (6685), eado6457
51. Volkov, D. O., *et al.*, *Applied Surface Science* (2011) **257** (20), 8518