



Towards enhanced nanoindentation by image recognition

Steffen Brinckmann^{1,a)} , Ruth Schwaiger¹

¹ Microstructure and Properties of Materials (IEK-2), Institute of Energy and Climate Research (IEK), Forschungszentrum Juelich GmbH, 52425 Juelich, Germany

^{a)} Address all correspondence to this author. e-mail: s.brinckmann@fz-juelich.de

Received: 25 November 2020; accepted: 17 March 2021; published online: 30 March 2021

The Oliver–Pharr method is maybe the most established method to determine a material's Young's modulus and hardness. However, this method has a number of requirements that render it more challenging for hard and stiff materials. Contact area and frame stiffness have to be calibrated for every tip, and the surface contact has to be accurately identified. The frame stiffness calibration is particularly prone to inaccuracies since it is easily affected, e.g., by sample mounting. In this study, we introduce a method to identify Young's modulus and hardness from nanoindentation without separate area function and frame stiffness calibrations and without surface contact identification. To this end, we employ automatic image recognition to determine the contact area that might be less than a square micrometer. We introduce the method and compare the results to those of the Oliver–Pharr method. Our approach will be demonstrated and evaluated for nanoindentation of Si, a hard and stiff material, which is challenging for the proposed method.

Introduction

Indentation is widely used to characterize the mechanical material properties of materials. Traditionally, indentation has been used to determine the hardness of a material by pushing a hard sphere or pyramid into the material with a given force or weight. Afterwards, the scientist would measure—by optical microscopy—the contact diameter for the Brinell sphere [1] (for ductile metals) or the diagonal length of the Vickers imprint [2] (for harder metals). Based on that contact area, the hardness is then calculated. In addition to these static indentation methods, dynamic indentation methods exist, e.g., [3].

Starting with the Rockwell indentation technique [4], the hardness was determined based on the depth recorded during the indentation. In the case of Rockwell indentation, a minor and a major force are applied and the difference in depth determines the hardness. Later on, the universal or Martens hardness used multiple loading steps and the unloading curve [5]. In the 1990s, Oliver and Pharr further improved depth-sensing indentation [6]. They combined the Sneddon equation [7] and approximated the contact depth with the help of the unloading stiffness. From the contact depth, Oliver and Pharr evaluated the contact area using an analytical representation. The

Oliver–Pharr method is the current standard [8] when determining Young's modulus and hardness down to an indentation depth in the range of tens of nanometers.

The Oliver–Pharr method requires two calibration steps. Initially the frame stiffness is calibrated, which is followed by the area function. The area function $A_c = m_0 h_c^2 + m_1 h_c + m_3 h_c^{0.5} + m_4 h_c^{0.25} + \dots$ uses a custom series function to map the contact depth h_c to a contact area A_c . It should be noted that the function has little geometric relevance as only the first term h_c^2 and second term h_c relate to conical opening angle and tip rounding (for small contact depths), respectively. As the tip wears during multiple experiments, the tip area function should be re-calibrated using the reference material. A typical reference material is fused silica, which has a low Young's modulus compared to most metals and hard-coating materials. This difference leads to extrapolation of the mechanical properties from a compliant reference material to a stiff sample. Naturally, every extrapolation might result in large differences between intrinsic and apparent properties.

The conventional Oliver–Pharr method uses $h_c = h_{\max} - 0.75 P_{\max} / S_{\max}$ to determine the contact depth h_c from the measured depth h , force P , and unloading stiffness S

at the maximum force with subscript 'max.' Lately, the 0.75 pre-factor is only used for brittle materials that show a pronounced sink-in during indentation while the factor decreases to 0 for ductile metals that show a pronounced pileup during indentation [9]. As such, the user has to know *a priori* the ductility of the material to determine its hardness and Young's modulus. In FEM simulations, a lower yield stress reduces the elastic deformation and increases the plastic deformation, which results in less sink-in and more pileup deformation during indentation. However, sink-in and pileup deformation are not uniquely linked to elasticity and ductility, respectively. For instance, non-ductile nanocrystalline Ni exhibits a pileup. Therefore, we concentrate on the terms *sink-in* and *pileup* in this study.

Moreover, depth sensing has two sources of potential error: surface contact identification and thermal drift. As the nanoindenter tip approaches the surface, its depth is tared at the surface contact. This contact is typically determined from a stiffness or force increase but is especially error prone if surface layers exist. The thermal drift is determined after the indentation and is assumed linear. This thermal drift estimate is used to correct the measured depth. Hence, the contact area and depth determination have a number of sources of uncertainty.

The frame stiffness calibration is of importance as it includes all stiffness contributions other than the contact stiffness.¹ The gantry stiffness, the stiffness of the glue, the stiffness of the nanoindenter tip mounting, and the influence of the sample thickness are summarized in the frame stiffness. Since the contributions from the glue, sample thickness and from the tip mounting vary depending on the user, also the frame stiffness changes from one sample to the next.

Good practice warrants that the calibration uses the same tip and sample mounting, if the contact stiffness—between tip and sample—is larger than 10% of the frame stiffness [10]. According to Sneddon [7], the contact stiffness K_c is given by

$$K_c = \sqrt{\frac{4A_c E_r^2}{\pi}}, \quad (1)$$

where E_r is the reduced modulus. The reduced modulus E_r is given by $1/E_r = (1 - \nu_{\text{tip}}^2)/E_{\text{tip}} + (1 - \nu^2)/E$, where the 'tip' subscript denotes the material properties of the hard and stiff indenter tip and the properties without subscript denote the properties of the sample. E denotes Young's modulus and ν Poisson's ratio.

If we assume a perfect Berkovich tip with an area function of $A_c = 24.5 h_c^2$, we can simplify the area to

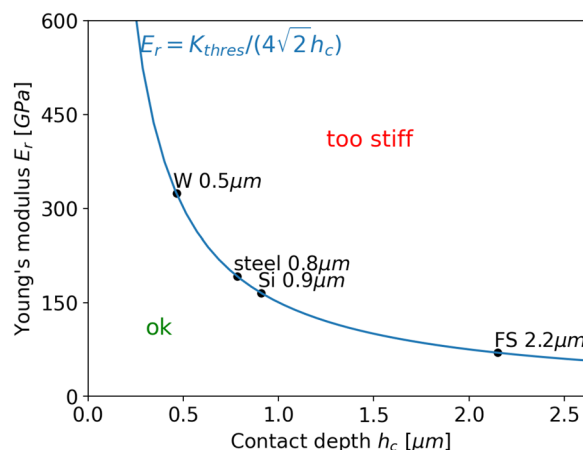


Figure 1: Threshold depth for different reduced moduli. Four characteristic materials have been selected: fused silica with Young's modulus of 72 GPa and Poisson's ratio 0.18, Si with Young's modulus of 179 GPa and Poisson's ratio of 0.27, steel with Young's modulus of 210 GPa and Poisson's ratio of 0.3, tungsten with Young's modulus of 411 GPa and Poisson's ratio of 0.3. The threshold depth is given for each material.

$A_c \approx 8\pi h_c^2 = 25.1 h_c^2$ accepting an error of 1% in Young's modulus. With the area simplification, we obtain a reduced modulus $E_r = K_c / (4\sqrt{2} h_c)$. If we use a typical frame stiffness $K_{\text{frame}} = 8.5 \cdot 10^6 \pm 1.56 \cdot 10^6 \text{ N/m}$ [10], we notice that the contact stiffness between tip and sample should be lower than $K_c = 8.5 \cdot 10^5 \text{ N/m}$. The maximum contact depth h_c that fulfills the stiffness requirement, is plotted as a function of the reduced modulus in Fig. 1. The figure shows that the contact depth is limited for stiff materials while compliant materials allow a larger contact depth.

In passing it should be noted that a number of nanoindenter vendors do not allow the user to vary the frame stiffness depending on the current glue compliance or do not allow the calibration of the frame stiffness by the user.

In summary, over the last decades, the indentation measurements have initially used optical methods to determine the contact area and hardness. Then, instrumented indentation and the Oliver–Pharr method make use of the depth measurements as well as stiffness and area calibrations to determine the hardness and Young's modulus. Challenges to the accurate representation of the frame stiffness exist for hard and stiff materials, which are used in hard coatings.

Currently, image recognition is employed widely in common consumer products. Faces [11] and finger-prints [12] are recognized in electronics to give access rights to the consumer. Street signs are recognized in cars to inform the driver of speed limits and potential hazards [13].

Here, we introduce Enhanced Nanoindentation by Image Recognition (ENIR) and use image recognition to determine the contact area and from it the hardness and Young's modulus.

¹ The frame stiffness has naturally non-linear terms, which are ignored in this study because their quantification is difficult due to limited contact at low contact forces.

In this approach, the frame stiffness is evaluated using cyclic unloading for each sample, i.e., from each set of measurements. The contact area after the final unloading is determined by image recognition based on micrographs. It should be noted that a number of macro-indenter vendors use image recognition to determine the contact area and hardness (without Young's modulus) [14, 15]. However, their algorithms are proprietary, and publicly accessible documentation does not exist.

To establish the ENIR method and evaluate its accuracy, we use a hard and stiff material with homogenous material properties: single-crystal Si wafer with a $\langle 100 \rangle$ normal orientation. The properties of Si wafers have been reported in a number of studies. Hess [16] determined that Poisson's ratio is 0.27 using an Nd:YAG laser with a wave length of 355 nm on an undoped and polished sample. The hardness of a $\langle 111 \rangle$ normal surface was determined as 11.7 GPa using atomic force microscopy (AFM) after indentation with 0.1 mN [17]. The hardness of the $\langle 100 \rangle$ surface of an undoped single crystal was evaluated as 13 GPa by indenting with a normal force of 0.2 mN [18]. Nanoindentation and the analysis of the load-displacement curve for a $\langle 001 \rangle$ surface have obtained a hardness of 7–9 GPa [19].

Nanoindentation reports the isotropic Young's modulus, while the sample might be anisotropic. Nanoindentation with 0.2 mN on a $\langle 100 \rangle$ surface reported a Young's modulus of 179 GPa and with 15 mN a Young modulus of 202 GPa [18]. The authors found that doping significantly influences the modulus. Microindentation on the $\angle 111$ resulted in a range of Young's moduli of 163–188 GPa [20]. Hess reported a Young modulus of 160 GPa for the $\langle 111 \rangle$ Si-surface using an Nd:YAG laser [16].

The goal of this study is to introduce the ENIR method, which does not require frame stiffness and area function calibration on a reference material to determine the hardness and Young's modulus from image recognition. It will be highlighted that the surface contact identification is not necessary and that the undesirable thermal drift is less influential in the ENIR method. Its applicability and uncertainty compared to the Oliver–Pharr method will be evaluated.

Theory for enhanced nanoindentation using image recognition

The approach consists of two steps. In the first step, the frame stiffness for the given sample and tip mounting is evaluated from cyclic unloading segments of a multitude of indentations. Thereafter, the area is evaluated by microscopy, and the Young modulus and hardness are obtained.

Following [10, 21], the compliance (inverse of stiffness) of the frame and contact are added to obtain the measured compliance. The contact compliance of the material with its reduced modulus E_r depends on the contact area, see Eq. 1. If

the definition of hardness is employed $H = P/A_c$, the measured compliance C_{measured} is given by

$$C_{\text{measured}} = C_{\text{frame}} + C_c = C_{\text{frame}} + \frac{\sqrt{\pi H}}{2E_r} \frac{1}{\sqrt{P}}, \quad (2)$$

$$= C_{\text{frame}} + a \frac{1}{\sqrt{P}}$$

where a is a constant that can also be obtained as $1/a^2 = \frac{K_c^2}{P}$. Eq. 2 is typically used for the frame stiffness calibration [10], because the intercept is easily determined by linear regression. Alternatively, the K_c^2/P ratio can be evaluated for every indentation depth and the frame stiffness can be adjusted to result in a constant K_c^2/P . The K_c^2/P approach requires iterating the frame stiffness, evaluating the contact stiffness and calculating the ratio as a function of depth. As such, that K_c^2/P approach is rather cumbersome.

In the ENIR method, the C_{frame} is evaluated from multiple unload cycles of all indentations using the same sample and tip mounting, i.e., all experiments on the same sample. It should be noted that this evaluation—as any frame stiffness calibration—requires that the material properties H and E_r are constant as a function of depth. Hence, the ENIR method is not applicable for materials with gradients. For metals, the penetration depth has to surpass the domain of the indentation size effect [22]. For thin films, quasi-constant properties are required and the indentation depth is limited by 10% of the film thickness.

Figure 2a shows a typical nanoindentation with cyclic unloading (5 segments) and a maximum load of 5 mN in Si. The compliance in Fig. 2b shows that the frame compliance is much smaller than the contact compliances. Measurements with a force of less than 1 mN are excluded from the fit because contact is not well established for low indentation depth.

The area evaluation by image recognition is explained in the next section. Here, the determination of Young's modulus and hardness are shown. If the measured compliance is the sum of the frame and the contact compliance, and if the Sneddon equation [7] is employed, we obtain the reduced modulus E_r

$$E_r = \sqrt{\frac{\pi}{4A_c}} \frac{1}{C_{\text{measured}} - C_{\text{frame}}} \quad (3)$$

From this equation, Young's modulus is evaluated by $(1 - \nu^2)/E = 1/E - (1 - \nu_{\text{tip}}^2)/E_{\text{tip}}$, where the 'tip' subscript denotes the properties of the tip material. ν is Poisson's ratio of the sample material. The hardness is determined as $H = P_{\text{max}}/A$. Hence, we need to measure the compliance C_{measured} , force P_{max} , and contact area A_c at the final unloading to determine Young's modulus and hardness.

The entire approach does not evaluate the contact depth or contact area function. Therefore, the contact of the tip with

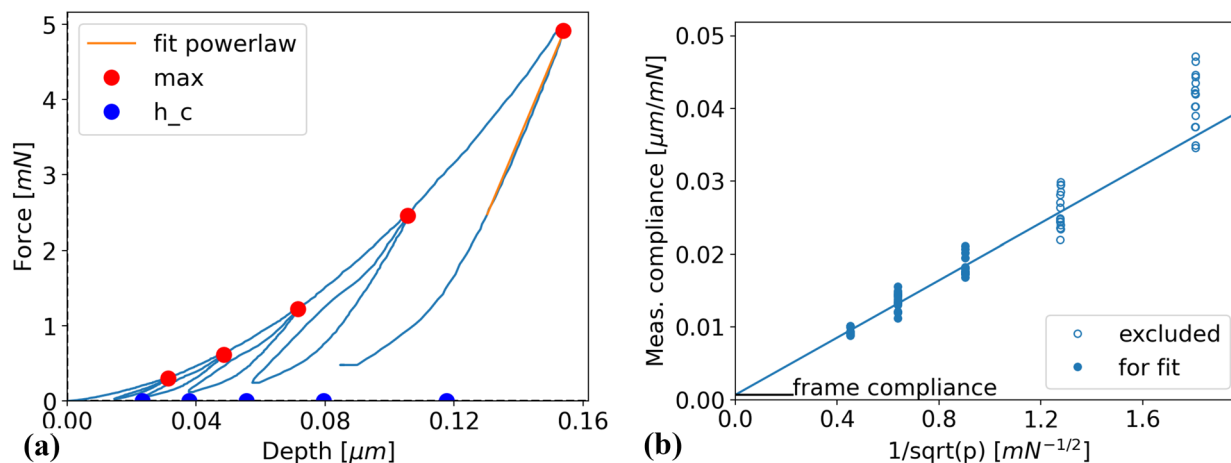


Figure 2: (a) One typical nanoindentation measurement with 5 unloading cycles on Si. The individual maximum forces and contact depth are denoted by red and blue dots, respectively. The unloading slope is determined from a power law, as shown in orange at the final unloading. (b) Measured compliance as a function of the load for 16 indentations. The linear fit used the filled datapoints as the open datapoints are not used for fitting and only shown for reference. The frame compliance is positive but close to zero.

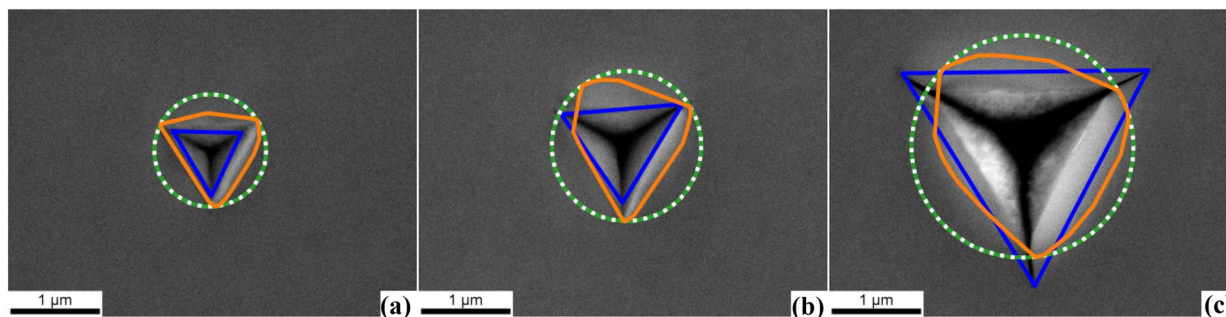


Figure 3: Original images and superposed contours (see Fig. 7). (a) after 5 mN indentation; (b) after 10 mN indentation; (c) after 20 mN indentation.

the surface does not have to be identified from the load–displacement curve, and the drift influences the results less in the ENIR method.² Furthermore, the method does not depend on the sink-in/ductility of the material, i.e., is the contact depth equal to the maximum depth as for ductile materials or less than the maximum depth as for brittle materials [9]. In addition, the frame stiffness is not considered a fixed calibration constant, but it depends on the tip and sample mounting. As such, the frame stiffness is evaluated for each sample, i.e., set of measurements, separately.

Results

The nanoindenter was used to indent the Si-wafer with a matrix of 4×4 indents with maximal normal forces of 5, 10, and 20 mN, each. The loading segment included 5 loading–unloading cycles

to determine the stiffness as a function of the applied normal force. As shown in Fig. 2a, the force–displacement curves have some rough unloading curves and a significant hysteresis during the unloading–reloading cycle. Jang et al. [23] report that this elbow behavior is connected to a volume-reducing phase transformation, which occurs for Berkovich indenter tips especially at low loads. Those authors also note that a pop-out mechanism occurs at forces above 20 mN. Indeed, the elbow hysteresis dominates the measurements of the force domain that is employed in this study.

Initially, we focus on the area identification from SE images. To this end, we imaged 10, 8, and 4 imprints of nanoindentations with normal forces of 5, 10, and 20 mN, respectively. The number of scans was chosen to allow for statistical analysis; while it was obvious early on that the measurements with a 20 mN normal force were less significant. These scans were analyzed as described in the previous section, and images were rejected if they failed the quality requirements. Ideally, those imprints would be scanned again in a fully automatic configuration. After

² The drift influences the apparent unloading stiffness in the Oliver–Pharr and ENIR methods.

rejection, 8, 6, and 4 images remained for 5, 10, and 20 mN, respectively. Fig. 3 shows representative images for each normal force group and the identified contours.

As shown in Fig. 3a, the imprint is visible and the pileup is minimal at an indentation force of 5 mN. Under these conditions, the convex hull (orange) and the enclosing circle well describe the contact area. However, the threshold for the imprint corners (blue line, see also dark area in Fig. 7) is too restrictive and the identified contact area is too small. As such, the *triangle of corners* algorithm is of less use in this condition. For 10 mN, the area identification by the *triangle of corners* (blue line) is of mixed quality: while the upper corners are well identified, the lower corner is not well identified. The convex hull that includes the *contact and pileup* area is too large, as significant pileups are observed. The circle and the corresponding equilateral triangle best describe the contact area in this force group.

For imprints with a normal force of 20 mN, we observe cracks at the corners of the imprint. These cracks are identified as dark area and included in the *triangle of corners* area. Jang et al. [24] observed cracks during indentation of Si and used these cracks to calculate its fracture toughness in their early study. In their later study, the same group of authors quantified the threshold for cracks under a Berkovich indenter as 10–20 mN [25]. That threshold has the same value in this study.

Since the cracks are identified as dark area and contribute to the *triangle of corners* area, this area is too large and not useful for large normal forces. The convex hull and the *contact and pileup* area are also too large, because a significant pileup has formed, which cannot be neglected. On the other hand, the *triangle in circle* algorithm proves to be the best approximation of the contact area.

Concluding, the *triangle in circle* algorithm is the superior algorithm for all normal forces. In the following, we quantify the area using the different algorithms and normalize them with the normal force and show these values as hardness.

The inspection of the imprints, see Fig. 3, has revealed that the *triangle of corners* algorithm overestimates the contact areas for large forces and underestimates them for small forces. This behavior results in an apparent hardness that has a strong normal force dependence, see blue line in Fig. 4. The *contact and pileup* algorithm naturally leads to a larger apparent contact area because it includes the pileup. At small indentation forces, the influence of the pileup is less significant which leads to an appropriate area and consequently apparent hardness. However, the pileup renders this algorithm less suitable for higher normal forces. The *circle in triangle* algorithm results in an intermediate contact area and apparent hardness. It should be noted that the reported hardness by this algorithm is slightly increasing with increasing normal forces. This behavior is not expected since Si has a constant hardness in this normal force regime and a size effect is not expected.

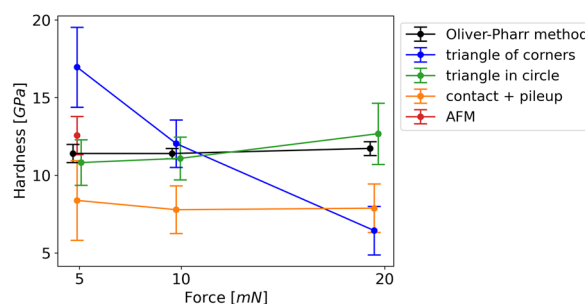


Figure 4: Normal force normalized by the identified projected contact area. A triangle connects the corners of the dark areas: triangle of corners; the convex hull that surrounds the imprint and the pileup: contact + pileup; the triangle that is inscribed in the circle, which surrounds the convex hull: triangle in circle. For reference also see Fig. 7. The results of the Oliver–Pharr method are shown in black while the hardness as determined by AFM measurements is shown in red. The scatter shows the standard deviation for all measurement types. The data are laterally off-set to allow for more clarity.

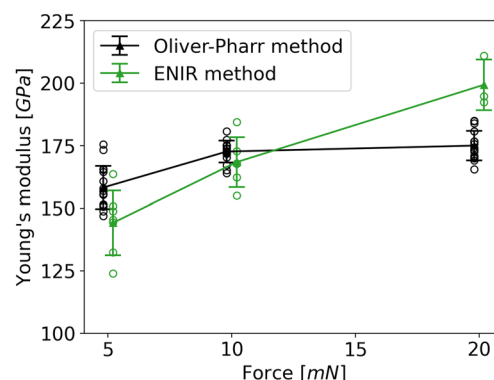


Figure 5: Young's modulus as a function of the indentation force according to the ENIR method and the Oliver–Pharr method. The data are laterally off-set to allow for more clarity.

The measurements according to the Oliver–Pharr method exhibit a constant hardness. The AFM-based and *triangle in circle* algorithm give an average hardness which is similar to that of the Oliver–Pharr method.

The AFM measurements have the smallest scatter, while the scatter is largest for the *triangle of corners* algorithm since it is difficult to identify the corners for small imprints. The scatter for the *triangle in circle* algorithm is lowest for small indentation forces while the *contact and pileup* algorithm has the lowest scatter for large forces. Compared to all image recognition-based methods, the Oliver–Pharr method has the lowest hardness scatter.

From the previous results, we conclude that the *triangle in circle* algorithm is the best for the SEM images and will be employed for the subsequent Young's modulus evaluation. Using Eq. 3, we calculate first the reduced modulus and then Young's modulus, as shown in Fig. 5. Using the present ENIR method,

we obtain Young's modulus that is within the scatter of Young's modulus values determined by the Oliver–Pharr method.

The present approach results in a minor size effect (i.e., smaller indentation forces result in a smaller apparent Young's modulus). This apparent size effect is smaller when using the Oliver–Pharr method. The scatter using the ENIR method is large for high normal forces for which cracks developed during nanoindentation. Moreover, the scatter is higher for the ENIR method than in the Oliver–Pharr method even at low indentation forces which do not result in cracks. The quantification of the scatter will be a central part of the next section.

Discussion

The hardness and Young's modulus according to the ENIR method are similar to the values determined using the Oliver–Pharr method. The similarity of both methods is also given for small contact areas $< 1 \mu\text{m}^2$. At 20 mN the apparent contact area is too small compared to the physical contact area. As a result, the hardness is reported as too high and Young's modulus is reported as too low.

Image analysis can identify the development of cracks, where nanoindentation is not sufficiently sensitive. Indentation into Si [25] has found no phase transformation because only a small volume was deformed although the same authors [23] also observed phase transformations in other studies. These phase transformations during unloading result in a more complicated unloading curve in Si than for other materials.

The Oliver–Pharr and this ENIR method rely heavily on the unloading stiffness which is in most materials well described by the power law $f = B(h - h_f)^m$ with the exponent m larger than 1 but less than 2. However, if an elbow behavior occurs as in the case of Si, then the exponent can significantly increase and the non-linear fitting becomes more cumbersome and leads to a larger scatter.

The scatter in the measured contact stiffness also leads to a high scatter in the calculated frame stiffness, which has a large influence on the reported Young's moduli but is too often ignored. The scatter in the frame stiffness arises from the difficulty of extrapolation to infinite normal forces. In such an extrapolation, results from higher forces reduce the scatter in the intercept evaluation compared to low forces. Therefore, the frame stiffness evaluation should use as high forces as possible whether in the Oliver–Pharr method on fused silica or in the ENIR method on the present sample. However, a large range between low and high forces is also helpful for reducing the scatter in the determination of the frame stiffness.

Figure 6 compares the measured compliance for a Si and fused silica sample. At low forces, the compliance is elevated compared to the series of spring model [21] because the contact

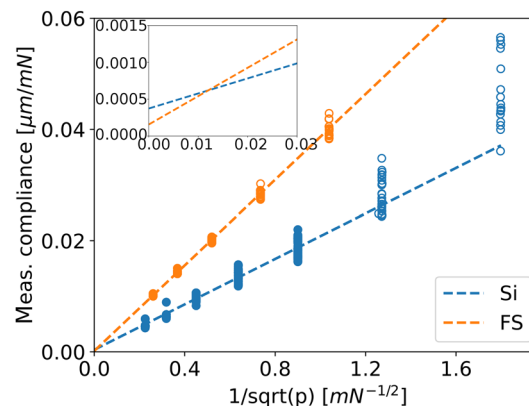


Figure 6: Frame compliance determination for a fused silica (FS) and Si sample as a function of the normal force. The filled symbols represent measurements at higher forces which were used for linear fitting. The open symbols represent other measurements at low forces, which were not used for linear fitting.

is not well established, and those datapoints are excluded from the linear fit. The slopes of both fits differ because the ratio of Young's modulus and hardness differ between both materials. However, one would expect that both curves intersect the y -axis at the same value. As the inset shows, these linear fits do not have the same y -intersection. Moreover, since the y -intersection of the compliance is close to zero, the stiffness is large and small differences in compliance result in severe differences in stiffness.

The scatter in the frame compliance and stiffness is severe as the frame stiffness can range from 10^6 to 10^7 N/m. However, the present and Oliver–Pharr method use a single frame stiffness and ignore this uncertainty, which is hidden from the reported data (Figs. 4 and 5). Increasing the number of indentations into the material in question has a limited effect on the standard deviations, i.e., measuring the same object with the same uncertainty, results in the same standard deviation.

The Oliver–Pharr method also assumes that the tip area is described by a unique analytical function. However, the determination of the series factors results in an uncertainty, which is difficult to quantify since the different factors have different weights. As an estimate, we assume that the area function has an uncertainty of 5%, which is in line with the typical accuracy of area function fitting [9, 10].

To evaluate the uncertainty of Young's modulus, we propagate the standard deviation through the equations and determine the individual uncertainties $\Delta y = dy/dx|_x \Delta x$, in which the derivative of the function is evaluated at the average configuration and is multiplied by the individual uncertainty Δx . For the Oliver–Pharr method, 5 equations are evaluated:

- (1) Stiffness correction: $1/K_c = 1/K_{\text{measure}} - 1/K_{\text{frame}}$
- (2) Depth correction: $h_{\text{max}} = h_{\text{max}} - F_{\text{max}}/K_{\text{frame}}$
- (3) Contact depth: $h_c = h_{\text{max}} - 0.75F_{\text{max}}/K_c$

TABLE 1: Uncertainty analysis for the Oliver–Pharr method using values from indentation with a 10 mN normal force on Si.

	<i>x</i>	Value	Unit	$dy/dx x$	Unit	Δx	Unit	Δ	Unit
(1)	K_{measure}	156.57	mN/ μm	1.13	–	3.72	mN/ μm	4.196	mN/ μm
	K_{frame}	2680.27	mN/ μm	3.85	10^{-3}	1.39	N/ μm	5.472	mN/ μm
	K_c	166.28	mN/ μm					6.895	mN/ μm
(2)	F_{max}	9.75	mN	0.373	nm/mN	38.0	μN	< 1	nm
	h_{max}	0.22	μm	1.00	–	2.30	nm	2	nm
	K_{frame}	2680.27	mN/ μm	1.36	nm ² /mN	1.39	N/ μm	2	nm
	h_{max}	0.21	μm					3	nm
(3)	F_{max}	9.75	mN	4.51	nm/mN	380	nN	0.171	nm
	h_{max}	0.21	μm	1.00	–	2.98	nm	2.978	nm
	K_c	166.28	mN/ μm	265	nm ² /mN	6.90	mN/ μm	1.824	nm
	h_c	0.17	μm					3	nm
(4)	h_c	0.17	μm	10.7	μm	3.50	nm	0.038	μm^2
	Factors	32.08							
	A_c	0.90	μm^2					0.039	μm^2
(5)	K_c	166.28	mN/ μm	0.935	1/ μm	6.90	mN/ μm	6.446	GPa
	A_c	0.90	μm^2	86.5	mN/ μm^4	0.0394	μm^2	3.410	GPa
	E_r	155.46	GPa					7.293	GPa

TABLE 2: Uncertainty analysis for the Enhanced Nanoindentation using Image Recognition (ENIR) method using values from indentation with a 10 mN normal force on Si.

	<i>x</i>	Value	Unit	$dE_r/dx x$	Unit	Δx	Unit	Δ	Unit
	A_c	0.90	μm^2	86.5	mN/ μm^4	0.07	μm^2	6.055	GPa
	C_{measure}	6.39	nm/mN	25.9	N ² / μm^3	0.152	nm/N	3.923	GPa
	C_{frame}	0.373	nm/mN	25.9	N ² / μm^3	7.46	pm/N	0.193	GPa
	E_r	155.46	GPa	4.64			%	7.217	GPa

- (4) Contact area: $A_c = 24.5 h_c^2$ and 5% additional uncertainty

- (5) Sneddon equation: $E_r = K_c \sqrt{\pi/4A_c}$

These evaluations result in the following uncertainty table for the Oliver–Pharr method. We use the italic font for input variables that are derived from the 10 mN indentations into the Si sample, showcased in this study. Since the uncertainties of the individual rows are uncorrelated, the summation in the last column uses the Euclidean norm to obtain the last row of each section.

We discuss the uncertainty in each equation/section separately. The stiffness correction leads to a contact stiffness $\Delta K_c = 6.895 \text{ mN}/\mu\text{m} \sim 4.15\%$ and both parts, K_{measure} and K_{frame} , contribute roughly equally. The depth correction leads to an uncertainty of 3 nm which is 1.4%. This value arises from the uncertainty of the frame stiffness and the depth. The force uncertainty plays an insignificant role in this section. The contact depth equation does not significantly increase the uncertainty as it remains at 3 nm, within the rounding of nanometer. The contact area has an uncertainty of $0.039 \mu\text{m}^2$ which is 4.39%. Finally, the uncertainty of the reduced modulus is 7.293 GPa which is 4.69% at a reduced modulus of 155.46 GPa.

This reduced modulus corresponds to Young’s modulus of 167 GPa for Poisson’s ratio of 0.27.

The here reported standard deviation is larger than the standard deviation as evaluated from Young’s moduli of each individual nanoindentation, see Fig. 5. That standard deviation 4.323 GPa is deflated compared to the full uncertainty since the former ignores the uncertainty of the frame correction, depth correction, and contact area and assumes those values as a constant.

Now, we evaluate the uncertainty of the ENIR method which relies on image analysis. Here, the same values are used as in Table 1 and the uncertainty of the contact area is chosen such as to result in the same reduced modulus uncertainty as for the Oliver–Pharr method. As a result, the uncertainty of the area recognition should be $0.07 \mu\text{m}^2 \sim 7.8\%$. This area uncertainty is much larger than the uncertainty of the contact area in the Oliver–Pharr method because multiple equations and parameters contribute to the uncertainty in that method. On the other hand, only one equation accumulates uncertainty in the ENIR method, as shown in Table 2.

The required uncertainty of the area measurement is not met for the ENIR method. We obtained an uncertainty on the order of 30% for the area evaluation from SEM images. For AFM measurements, the uncertainty is 10%. As such, the uncertainty

of the ENIR method and algorithms is greater than the uncertainty of the Oliver–Pharr method.

However, the uncertainty evaluation for the Oliver–Pharr method (Table 1) is a best case scenario. For ductile metals, a factor of 0 for Eq. 3 in Table 1 is suggested [9]. If that factor is uncertain, then the uncertainty of the reduced modulus increases to 25%. In that case, the standard deviation using the ENIR method is superior to the Oliver–Pharr method.

The ENIR method for the identification of Young’s modulus and hardness requires image acquisition. Confocal and light microscopy are not sufficient if the contact area is on the order of $1\ \mu\text{m}^2$ as the lateral resolution of those techniques is too limited. Concluding, AFM and SEM are the remaining techniques for these hard materials. AFM is slow but integrated in our nanoin-denter. In addition, AFM data have the 3D information and can be used reliably for the evaluation of the physical contact. As such, the uncertainty is 10%.

SEM imaging would be especially helpful if one uses in-situ nanoindentation with, e.g., a Hysitron PI 89 SEM PicoIndenter or a KLA Nano Flip. That equipment allows the user to indent the surface and then turn the sample to the SEM column for imaging. That approach would benefit from an integration of the nanoindenter and the SEM software to automatically capture indents after indentation.

The ENIR method has one significant advantage over the Oliver–Pharr method: imaging—whether SEM or AFM—supplies more information that helps the scientist to determine the material properties. In this study we have seen that minor cracks develop at higher loads. These cracks did not result in severe compliance and that compliance change is not observed in nanoindentation. However, images could be used to invalidate Young’s modulus and hardness measurements if cracks are observed.

Moreover, if the imprint is concave, the material deforms by sink-in as observed in the present study. If the imprint results in a large pileup, the imprint becomes convex. That information can supplement a nanoindentation study that conventionally only allows to study the energy under the force–displacement curve during nanoindentation.

Conclusions

In this study, we introduced the ENIR method that allows the evaluation of the mechanical material properties by using the frame stiffness correction and Sneddon equation. The contact area was determined from secondary electron images and AFM measurements. We used the ENIR method to determine Young’s modulus and hardness of Si, which are inside the scatter of those reported by the Oliver–Pharr method.

Comparing to the Oliver–Pharr method, the ENIR method has a number of advantages but also disadvantages:

- The ENIR method does not require the area function calibration or the separate frame stiffness calibration. Additionally, the ENIR method does not require to identify the surface contact from the load–displacement curve and is insensitive to the thermal drift. In the ENIR method, the frame stiffness is automatically evaluated for each set of measurements, i.e., each sample.
- Generally, the Oliver–Pharr method results in less uncertainty. To achieve comparable uncertainty using ENIR, the uncertainty of the area measurement must not exceed 7.8%. However, the algorithms for SEM images and AFM measurements resulted in uncertainties of 30% and 10%, respectively.
- Improved image quality is beneficial for reducing the uncertainty. Additionally, improved algorithms are required, which measure the contact area with less scatter.
- As the present approach uses a secondary information stream, i.e., images, more information can guide the scientist. In this study, cracks were visible in the images while the nanoindenter measurements did not show a reduction of the stiffness due to crack formation, i.e., nanoindentation was not sufficiently sensitive to fracture.
- Images can guide the scientist about the presence of pileups (convex shape) or the presence of sink-in deformation (concave shape).

Therefore, the Enhanced Nanoindentation using Image Recognition (ENIR) method has a great potential for hard and stiff materials when used in conjunction with the conventional Oliver–Pharr method. In order to realize the full potential of the ENIR method, the area recognition algorithms should be improved and the method should be applied for a number of materials.

Experimental procedure and data analytics

In this study, we used a MTS Nanoindenter XP with a Berkovich tip. The tip was calibrated prior to testing using a fused silica reference. The nanoindenter is equipped with a DME DualScope DS 95-200 Atom-Force-Microscope (AFM) with a C-26 controller. For imprint identification, we also use a Zeiss Merlin Scanning Electron Microscope (SEM) with a Gemini column and secondary electron (SE) and in-lens detectors. The SE detector was used for imaging as this detector shows a better topography contrast while it has a lower resolution and requires a longer dwell time than the in-lens detector.

The as-collected plain Tiff images contain a speckle noise, which is often removed by a median filter [26] with an appropriate radius and recursion. In this study, we employed a non-linear denoising filter [27] which resulted in superior image quality compared to the median filter, as shown in Fig. 7b. It should be noted that an appropriate adaptation of the denoising

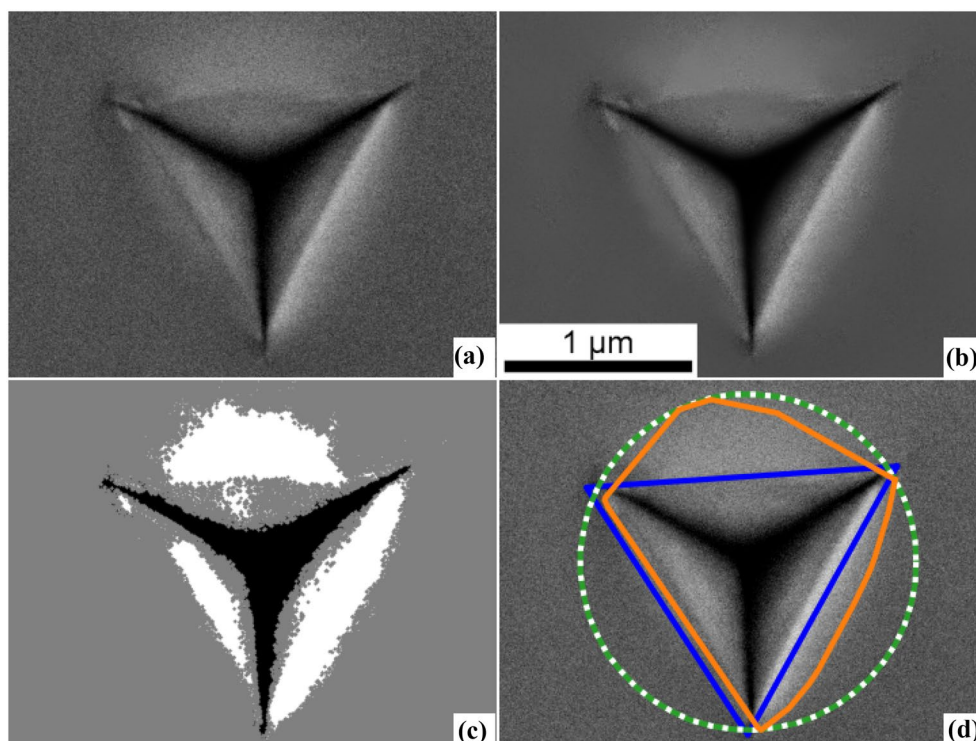


Figure 7: Procedure of contact area identification from SEM images. (a) central part of the full-sized original image which is twice as large as the domain, which is shown here; (b) after non-linear filtering; (c) after applying a threshold for bright and dark areas; (d) identified triangle from dark area (blue), convex hull of bright area (orange) and surrounding circle (green-white) superposed on the original image.

parameters (filter strength of 10, template window size of 21 pixel and a search window size of 63 pixel) severely simplifies the subsequent image processing steps. After filtering, the minimum and maximum of the background gray values were determined from a 60-pixel wide frame. These minimum and maximum thresholds were used to identify the dark (corner of imprints) and bright areas (imprint sides and pileup), respectively (see also Fig. 7c). A morphological closing operation with a 2-pixel kernel was executed on the dark and bright binary images to remove small outlying speckles.

From the dark area, the contour of the main shape was determined and the Ramer–Douglas–Peucker algorithm [28, 29] was used to identify its triangular shape, i.e., the connection of the corners. Finally, the triangle area was determined in units of pixel and converted into an area unit using the pixel-size from the meta data of the original image. This *triangle of corners* area is the first approximation of the contact area.

The bright area has an irregular shape as it includes a part of the imprint and the pileup. The convex hull [30] of the bright area includes the entire imprint and the pileup. As such, the convex hull area—converted into an area unit—is the second approximation of the contact area: *contact and pileup*. Using the convex hull, we determined the circle that encloses this hull [31]. Once the radius r of the enclosing circle is known, the area of an equilateral triangle is determined $A = 3\sqrt{3}/4r^2$.

This area, converted into an area unit, is the third contact area approximation: *triangle in circle*. Since the equilateral triangle is analytically determined, no triangle with specific corner points is determined and can be plotted.

It should be noted that these algorithms failed in a few cases. Criteria on the image quality and the results of the algorithm were defined to automatically filter out failures. If the maximum and minimum background gray values are above 235 and below 90, then the initial raw image is considered too bright or too dark, respectively. If the ratio of “*contact and pileup*” area and “*triangle of corners*” area are above 3 or below 1/3, then the algorithms failed. In these cases, the results of the image analysis were discarded.

The identification using the AFM measurements requires manual interaction as AFM measurements often include horizontal scars. To this end, the raw AFM measurements are cropped to exclude any artifact and to ensure that the 30-pixel wide frame does not include particles that hinder further processing. In the second step, the topography is improved to obtain a horizontal background: the beginning and end of each line scan are used to fit a line, which is subtracted from the line scan. Afterwards, the non-linear denoising algorithm [27] was employed. Since the noise in AFM measurements is periodic, a low pass filter or an FFT filter would also be successful. Here, we focus on one denoising algorithm for all raw data.

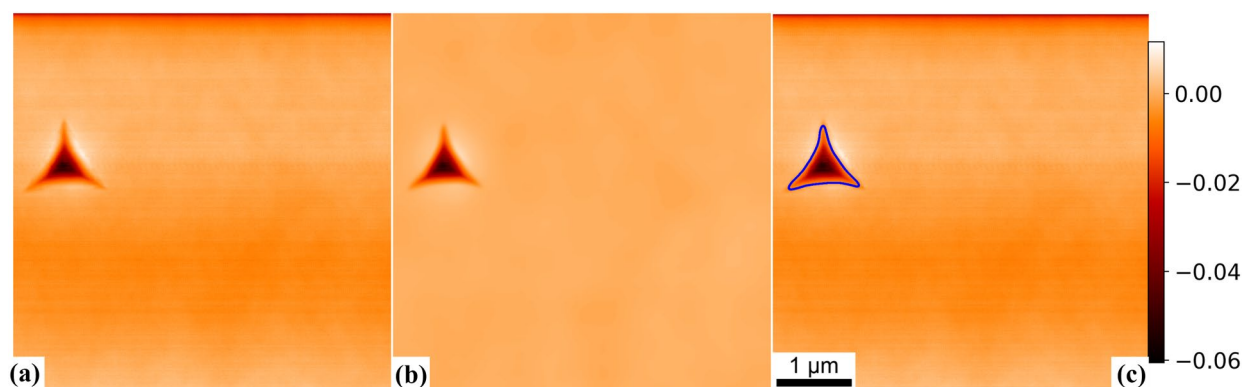


Figure 8: Procedure of contact area identification from AFM measurements. (a) original measurement, which is flipped by the horizontal plane; (b) after leveling and non-linear filtering; (c) identified contour superposed on the original image. Please note that the area of the convex hull is evaluated from the contour identified.

After denoising, the measurements were shifted to result in a zero-average height. Then the contact area was determined by iterations of the threshold depth: starting from a 2nm depth threshold below zero, the binary image of this height threshold was determined and a morphological closing operation with a kernel size of 2 was applied. The contours of the binary image were determined and those contours excluded that had fewer than 6 pixel. If the resulting number of contours was larger than one, the iteration was continued by multiplying the depth threshold with 1.5.

The final contour can have a concave shape, see Fig. 8c, especially in a material that exhibits a sink-in as Si does. However, the nanoindenter tip is convex and the contact area of this convex shape is similarly convex. Therefore, the convex hull of the identified contour was evaluated and the area of this convex hull is used as contact area. Concluding, only one area measure is reported for the AFM measurements.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Data availability

Data will be made available on reasonable request.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format,

as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. J.A. Brinell, Way of determining the hardness of bodies and some applications of the same. *Teknisk Tidskrift* **5**, 69 (1900)
2. R.L. Smith, G.E. Sandly, An accurate method of determining the hardness of metals, with particular reference to those of a high degree of hardness. *Proc. Inst. Mech. Eng.* **102**(1), 623–641 (1922)
3. Standard Test Method for Leeb Hardness Testing of Steel Products. Standard, American Society for Testing and Materials (2012)
4. S.P. Rockwell, The testing of metals for hardness. *Trans. Am. Soc. Steel Treat.* **2**(11), 1013–1033 (1922)
5. W. Weiler, Zur Definition einer neuer Härteskala bei Ermittlung des Härtewertes unter Prüfkraft. *Materialprüfung* **28**(7–8), 217–220 (1986)
6. W.C. Oliver, G.M. Pharr, An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments. *J. Mater. Res.* **7**(06), 1564–1583 (1992)
7. I.N. Sneddon, The relation between load and penetration in the axisymmetric Boussinesq problem for a punch of arbitrary profile. *Int. J. Eng. Sci.* **3**(1), 47–57 (1965)

8. Metallic materials—Instrumented indentation test for hardness and materials parameters—Part 1: Test method. Standard, International Organization for Standardization, Geneva, CH (2015)
9. KLA, *Nano Indenter G200X User's Guide* (2020)
10. Agilent Technologies, *Nano Indenter G200 User's Guide* (2013)
11. R. Gross, I. Matthews, S. Baker, Eigen light-fields and face recognition across pose, in *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*. (IEEE, 2002), pp 3–9
12. S. Juh, Pointing device having fingerprint image recognition function, fingerprint image recognition and pointing method, and method for providing portable terminal service using thereof, (November 10 2005). US Patent App. 10/520,651
13. G. Piccioli, E. De Micheli, P. Parodi, M. Campani, Robust method for road sign detection and recognition. *Image Vis. Comput.* **14**(3), 209–223 (1996)
14. QATM, *Qness 60 evo: Härteprüfung und Mikroskopie nahtlos verbunden und hochautomatisiert* (2020)
15. INNOVATEST, *Nemesis 5100: Automatisch, Vollautomatisch* (2020)
16. P. Hess, Laser diagnostics of mechanical and elastic properties of silicon and carbon films. *Appl. Surf. Sci.* **106**, 429–437 (1996)
17. B. Bhushan, Nanotribology and nanomechanics of mems devices, in *Proceedings of Ninth International Workshop on Micro Electromechanical Systems* (1996), pp. 91–98
18. B. Bhushan, X. Li, Micromechanical and tribological characterization of doped single-crystal silicon and polysilicon films for microelectromechanical systems devices. *J. Mater. Res.* **12**(1), 54–63 (1997)
19. H. Sjöström, I. Ivanov, M. Johansson, L. Hultman, J.-E. Sundgren, S.V. Hainsworth, T.F. Page, L.R. Wallenberg, Reactive magnetron sputter deposition of cnx films on si(001) substrates: film growth, microstructure and mechanical properties. *Thin Solid Films* **246**(1), 103–109 (1994). (**Special Issue: Frontiers in Thin Film Research**Special Issue: FrontiersSpecial Issue: Frontiers in Thin Film Research)
20. M.T. Kim, Influence of substrates on the elastic reaction of films for the microindentation tests. *Thin Solid Films* **283**(1), 12–16 (1996)
21. D.S. Stone, K.B. Yoder, W.D. Sproul, Hardness and elastic modulus of tin based on continuous indentation technique and new correlation. *J. Vac. Sci. Technol. A: Vac. Surf. Films* **9**(4), 2543–2547 (1991)
22. W.D. Nix, H. Gao, Indentation size effects in crystalline materials: a law for strain gradient plasticity. *J. Mech. Phys. Solids* **46**(3), 411–425 (1998)
23. J.I. Jang, M.J. Lance, S. Wen, T.Y. Tsui, G.M. Pharr, Indentation-induced phase transformations in silicon: influences of load, rate and indenter angle on the transformation behavior. *Acta Materialia* **53**(6), 1759–1770 (2005)
24. J. Jang, M.J. Lance, S. Wen, G.M. Pharr, Evidence for nanoindentation-induced phase transformations in germanium. *Appl. Phys. Lett.* **86**(13), 131907 (2005)
25. J. Jang, G.M. Pharr, Influence of indenter angle on cracking in Si and Ge during nanoindentation. *Acta Materialia* **56**(16), 4458–4469 (2008)
26. T. Huang, G.J.T.G.Y. Yang, G. Tang, A fast two-dimensional median filtering algorithm. *IEEE Trans. Acoust. Speech Signal Process.* **27**(1), 13–18 (1979)
27. A. Buades, B. Coll, J.-M. Morel, Non-local means denoising. *Image Process. Line* **1**, 208–212 (2011)
28. U. Ramer, An iterative procedure for the polygonal approximation of plane curves. *Comput. Graph. Image Process.* **1**(3), 244–256 (1972)
29. D.H. Douglas, T.K. Peucker, Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: Int. J. Geogr. Inf. Geovis.* **10**(2), 112–122 (1973)
30. D.G. Kirkpatrick, R. Seidel, The ultimate planar convex hull algorithm? *SIAM J. Comput.* **15**(1), 287–299 (1986)
31. James Joseph Sylvester, A question in the geometry of situation. *Q. J. Pure Appl. Math.* **1**(1), 79–80 (1857)