Robust evaluation of statistical surface topography parameters using focus-variation microscopy

2016 Surf. Topogr.: Metrol. Prop. 4 035003

(http://iopscience.iop.org/2051-672X/4/3/035003)

Surface Topography: Metrology and Properties

OPEN ACCESS

RECEIVED
12 April 2016

REVISED
16 June 2016

ACCEPTED FOR PUBLICATION
18 July 2016

PUBLISHED
19 September 2016

PAPER

Robust evaluation of statistical surface topography parameters using focus-variation microscopy*

E N Grossman[1], M Gould[2] and N P Mujica-Schwann[1,3]

[1] National Institute of Standards and Technology, Boulder CO, USA
[2] Zen Machine and Scientific Instruments, Lyons, CO, USA
[3] Department of Physics, University of Colorado, Boulder CO, USA

E-mail: erich.grossman@nist.gov

Keywords: roughness, autocorrelation length, focus-variation microscopy

Abstract

Spatial bandwidth limitations frequently introduce large biases into the estimated values of rms roughness and autocorrelation length that are extracted from topography data on random rough surfaces. The biases can be particularly severe for focus-variation microscopy data because of the reduced lateral resolution (and therefore dynamic range) inherent in the technique. In this paper, we describe a measurement protocol—something similar to a deconvolution algorithm—that greatly reduces these biases. The measurement protocol is developed for the case of surfaces that are isotropic, and whose topography displays an autocovariance function that is exponential, with a single autocorrelation length. The protocol is first validated against Monte Carlo-generated mock surfaces of this form that have been filtered so as to simulate the lateral resolution and field-of-view limits of a particular commercial focus-variation microscope. It is found that accurate values of roughness and autocorrelation length can be extracted over a four octave range in autocorrelation length by applying the protocol, whereas errors without applying the protocol are a minimum of 30% even at the absolute optimum autocorrelation length. Then, microscopy data on eleven examples of rough, outdoor building materials are analyzed using the protocol. Even though the samples were not in any way selected to conform to the model's assumptions, we find that applying the protocol yields extracted values of roughness and autocorrelation length for each surface that are highly consistent among datasets obtained at different magnifications (i.e. datasets obtained with different spatial bandpass limits).

1. Introduction

In many applications it is essential to estimate the

can be very time-consuming, particularly if samples are anisotropic and/or numerous. Within the last decade however, a gradual conversion has taken place

microscope, and phase-shifting interferometer [7–9], and a recent review [10] describes the current status of calibration and quantitative accuracy for most of the optical 2D techniques. On the other hand, comparatively little has been written, to our knowledge, about focus-variation microscopy (FVM), although some initial studies have appeared [11]. The applicable ISO standard [12] appeared only recently.

This paper does not aspire to be comprehensive or definitive at a metrological level, but is rather a practical description of methods we have found to 'give the right answer in most situations'. It discusses the use of FVM to estimate the rms height variation Sq and the autocorrelation length Sal of a random rough surface. Note that these are statistical parameters, and therefore present a slightly different measurement problem from measuring the height of a step (or some other deterministic feature) accurately. A truly metrological study would also necessarily discuss traceability, and would therefore compare FVM results with those obtained from another technique traceable to international standards. However, that is well beyond the scope of the current work.

Any statistically random surface is described by either of two equivalent functions, the power spectral density (PSD), which is a function of spatial frequency, or the autocovariance function (ACV), which is a function of lag or displacement. By the Wiener-Khintchine relation, the PSD and ACV are Fourier transforms of one another. (The *autocorrelation* function differs from the autocovariance only in the addition of a constant, the square of the mean elevation, that manifests itself in the PSD as a delta function at $f = 0$. In this paper we shall treat all theoretical and experimental elevations as differences from the mean, and only use ACV and PSD's without a delta function at the origin.) Sq and Sal are the two most important topographic parameters of a rough surface for determining its electromagnetic scattering properties. In the most commonly used scattering formalism [13], based on the Kirchoff approximation, only their ratio Sq/Sal (the mean surface slope) enters, but in more sophisticated treatments, this is not the case. Implicit in the use of a single autocorrelation length Sal are the assumptions that (a) the surface's ACV is isotropic, and (b) the ACV has a simple exponential form. It might seem straightforward (if not trivial) to verify these assumptions and to estimate the values of Sq and Sal directly from their definitions, using an FVM dataset, which is a simply a collection of sampled points $z_i(x_i, y_i)$. However, there are a number of practical issues that arise, particularly concerning spatial bandwidth limitations, that create bias in the estimators and/or raise their uncertainties. These are often dramatically manifested in situations where the same field of a particular sample is imaged by the same instrument but at different magnifications, resulting in vastly different estimates of Sq and Sal. This effect arises from the finite lateral resolution of the 3D FVM data (which varies with magnification). This lateral resolution is fundamental to the FVM technique and is always significantly greater (i.e. poorer) than the resolution of the underlying 2D images. It arises because the FVM technique is based on forming an estimate of local contrast at every position within an image. At each pixel's position, such an estimate requires information about the light scattered from neighboring pixels; this introduces correlations that necessarily degrade the lateral resolution (spatial bandwidth). In other words, the vertical resolution provided by the FVM technique comes at some price in lateral resolution (see section 7.2.3 of [2].)

There are other aspects of FVM measurements that can create bias or increase uncertainty in estimates of Sq and Sal. These are discussed in considerable detail from a standards perspective in [14], using as a starting point the 23 metrological characteristics listed in ISO standard 25178-601 [15]. Two of the more notable effects are measurement noise N_M (section 3.1.9 of [12]) and residual flatness z_{FLT} (table 1 of [12]). The importance of these effects can be estimated by various methods, e.g. measuring the same field of a particular sample multiple times, and measuring a 'known-flat' reference sample. Such measurements are a routine part of the FVM measurements we perform on everyday surfaces for our electromagnetic scattering studies. Ceramic gage blocks that have been slightly roughened (to provide visible surface contrast) are used as the flat references. In all cases we have found, as one might expect, that the levels of noise and residual flatness are on the order of the vertical resolu-

uncertainty in estimating Sq and Sal, and therefore this paper focuses exclusively on these effects.

Spatial bandwidth effects in the context of 1D profilometer measurements were treated thoroughly in a classic series of papers by Church and Takacs [17–19]. Reference [17] in particular emphasizes the fact that any surface profile measurement is only sensitive to a limited range of spatial frequencies, extending from $f_{min}=1/d_{max}$ to $f_{max}=1/d_{min}$, where the d's are the corresponding spatial wavelengths. The (lateral) dynamic range of the instrument is $Q = f_{max}/f_{min}$. Comparison between bandwidth effects on profilometer versus optical measurements was highlighted in [20]. The present paper, being oriented toward FVM, differs from these earlier works chiefly in the facts that: (a) the analysis is inherently 2D rather than 1D, and (b) the dynamic range of most FVM measurements is significantly less than that of most 1D stylus measurements, leading to larger correction factors.

This paper is organized as follows. Section 2 and appendix A describe a simple method to estimate the lateral spatial resolution of an FVM, and show that for our particular instrumentation this lateral resolution manifests itself in the data as a Gaussian low-pass filter with surprisingly large $1/e$-width. In section 3, we derive the protocol for estimating the true values of a surface's rms height variation and autocorrelation length, Sq and Sal, from FVM measurements. These symbols for the rms height variation and autocorrelation length have been established as standards for surface metrology [21] work, although many earlier studies, and common practice in the area of electromagnetic scattering, uses σ and L to denote the same quantities. In this paper, Sq and Sal are reserved for the true parameters of the surface, while σ and L are used for estimates based on measured data. First-order estimates are denoted σ_1 and L_1, and are formed by directly applying the statistical definitions to the FVM data. Then corrections that depend on instrumental resolution are applied to arrive at final estimators σ_2 and L_2. These corrections are exact for the case of an instrument with Gaussian point-spread function and a surface with exponential ACV. In section 4, the protocol is validated by generating mock surfaces with specified Sq and Sal values with the correct form of ACV and applying the corrected estimators. In section 5, the same estimators are calculated from FVM data on real-world samples of various outdoor building materials. This shows that, although the corrections derived in section 3 may only be exact for surfaces with exponential ACV, the assumptions are robust enough to give useful approximate estimates for a great many real-world surfaces—those that are isotropic and have a single characteristic lateral scale. Section 6 summarizes our conclusions.

2. Lateral resolution

As will become clear in later sections, it is critical to have accurate quantitative information about the particular microscope's lateral resolution (i.e. its point-spread function or PSF) in order to extract Sq and Sal reliably. In most cases, the accuracy to which this lateral resolution is known determines the final accuracy to which Sq and Sal can be determined. When used in 2D mode, (i.e. as a conventional microscope), the lateral resolution is simply determined by the pixel size of the camera or the quality (diffraction or aberration-limited) of the microscope objective. However, this is completely separate from the lateral resolution when used as an FVM. The lateral resolution of an FVM is almost entirely determined by its composition algorithm rather than the microscope or camera optics. In some cases (though not for the particular instrument used in this study) the lateral resolution is adjustable, and it can be shown [22] that ideally it is identical to the lateral resolution of a coherence scanning interferometer or a confocal microscope.

The instrument used to obtain all the measurement results in this paper is a commercial FVM (Keyence model VHX1000E)[4] equipped with a telecentric × 20–200 zoom lens objective (model VHZ20W) providing nominal magnifications of × 30, × 50, × 100, and × 200. Other magnifications are available but were not used in this study. These nominal magnifications are relative to the 15 inch display and are merely used as labels. The actual magnifications relative to the camera's CCD are listed in table 1. The zoom lens objective

Table 1. Measured FVM lateral resolution.

Nominal mag.	Actual mag.	Optimal res. target	R_0 (μm)	p (μm)	FOV (mm × mm)	A_{eff} (mm^2)	N
× 30	0.70	1A	234.5	7.30	11.7 × 8.74	0.173	590
× 50	1.17	2A	126	4.42	7.05 × 5.28	0.0499	746
× 100	2.35	3A	69.1	2.22	3.54 × 2.66	0.0150	628
× 200	4.69	4A	35.1	1.11	1.78 × 1.34	0.00396	601

specularly reflected, light. This, together with the fact that all samples of interest (to us) have roughness much greater than the wavelength, Sq $\gg \lambda$, guarantees that the detected light is purely incoherent. This greatly simplifies the analysis in [22]. The microscope's other features include image-stitching, high dynamic reserve, and built-in tilt correction. The FOV and sampling distance (pixel pitch) are listed for each magnification in table 1.

For all magnifications, the raw dataset for a single field consists of 1197 × 1597 $z(x, y)$ values (not the original stack of 2D images). In other words, this 'raw' dataset is really the output of the microscope's composition algorithm, which is proprietary and whose details are unknown to the authors. This is the primary reason for directly measuring the lateral resolution, using the resolution targets and analysis procedure described in this section and in appendix A. As described below, we can infer from these measurements that the 'raw' dataset has already been passed through a digital filter with particular properties. We assume that this digital filtering is linear. The filter shape and width are not adjustable, or even specified by the manufacturer, for the particular FVM used in this study. For other instruments, the digital filtering and lateral resolution are at least somewhat more transparent to the user; however, even in these cases it will often be of interest to measure it directly. In any case, the approach taken here is to measure the instrumental PSF directly, using the output of the composition algorithm from a stack of microscope images made on known physical test targets. In other words, we are treating the FVM microscope and its composition algorithm as a single 'black box'. All FVM data were analyzed using Matlab™, although much of the underlying analysis of section 3 was implemented in Mathematica™.

contrast to be maximized, i.e. the z-position at which that particular (x, y) location is in 'best-focus'. Although the details of the composition algorithm are proprietary to the microscope manufacturer, it is clear that calculating any measure of local contrast requires information about the light scattered from a larger (possibly much larger) area than a single pixel. This introduces correlations (possibly strong correlations) between the elevations calculated for neighboring pixels; in other words it degrades the lateral resolution.

In order to describe this quantitatively, a series of resolution targets was specially constructed, as illustrated in figure 1, using conventional CNC machining (not turning), but with very small end mills (down to 200 μm (0.008″) diameter), high spindle speeds (up to 30 krpm), and low feed rates. Each target consisted of a cylindrical post atop a larger cylindrical base, with diameters and heights as listed in the figure. A wide range in diameters and heights is necessary to allow measurement over a range of magnifications. For each magnification, the narrowest posts are laterally unresolved while the widest are laterally well resolved.

The FVM measurements represent the 2D convolution of the post's top-hat topography and the microscope's PSF. When the post is well resolved, the FVM data shows a 'flat-topped' profile for the post and a vertical height that corresponds closely to its nominal height. When it is laterally unresolved, the FVM data displays a profile that corresponds to the PSF of the composition algorithm, and a post height that is reduced from the nominal height. In the limit of a vanishing post diameter, $D_p \to 0$, i.e. when extremely unresolved, the measured FVM profile would ideally be the point-spread function itself (within a multiplicative factor). However, in this limit the post height displayed in actual FVM data becomes so small that it

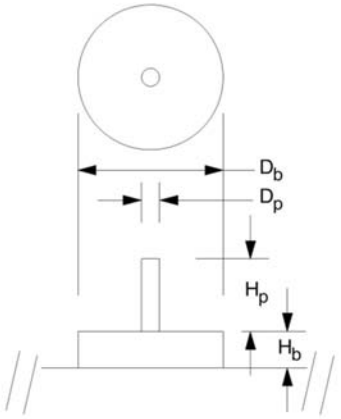
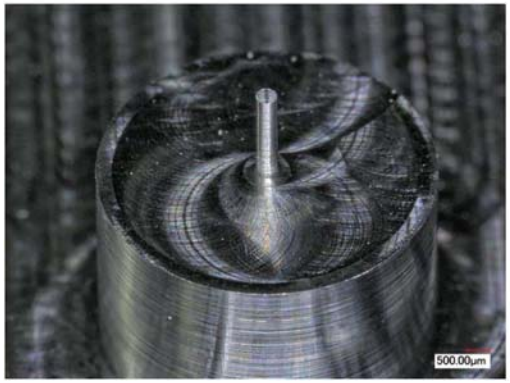


Table 1				
Designation	D_p (mm)	H_p (mm)	D_b (mm)	H_b (mm)
1A	0.50	2.50	8.0	5.0
2A	0.20	0.80	3.20	1.00
3A	0.10	0.50	1.60	0.20
3X	0.10	0.50	1.59	n/a
4A	0.050	0.250	0.80	0.050
4X	0.050	0.250	1.59	n/a
5A	0.020	0.10	0.320	0.008



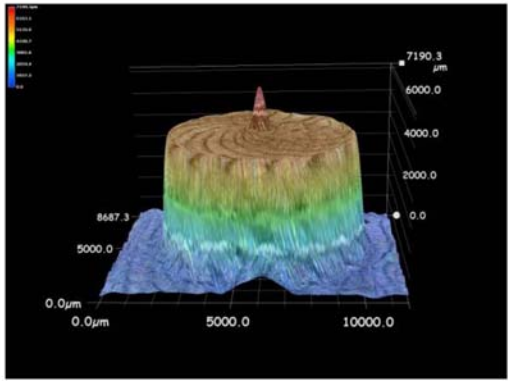


Figure 1. FVM resolution target geometry (upper left) and dimensions (upper right). The base of the 3× and 4× targets was machined, but not deliberately elevated above the background surface. A conventional micrograph at × 30 magnification of target 1 A (lower left) and a 3D rendering of the × 30 FVM data on the same target (lower right).

the Gaussian scales inversely with magnification; i.e. it is a constant fraction of the FOV. This Gaussian fit is a reasonable zeroth-order approximation to the point-spread function. However, for accurate quantitative work it is not satisfactory because the post is not in the fully unresolved limit. If the point-spread function's radius to $1/e$-point is R_0 and the post radius is R_p, then $R_p/R_0 < 1$ may be true, but $R_p/R_0 \ll 1$ is not, so we expect a (downward) correction needs to be applied to the width of the FVM data (whether obtained by a Gaussian fit or some other means), in order to obtain an unbiased estimate of the PSF width. A procedure is described in appendix A to calculate this correction, for R_{eff} to constitute a reasonably precise estimate of R_0 (say within 5%), the target radius needs be quite small, approximately $R_p < 0.4R_0$. As illustrated in figure A1, this in turn implies an apparent post height of only <10% the actual height. The core of the problem is that for any physical FVM target, manufacturing considerations limit the achievable ratio of height to radius. This aspect ratio is 10 for most of the FVM targets in this study, near the limit of what can be accurately produced with conventional machining of such small structures. The FVM measurements $M(r)$ will therefore display a hump of approximately equal height and radius. In this case, estimating R_{eff} requires

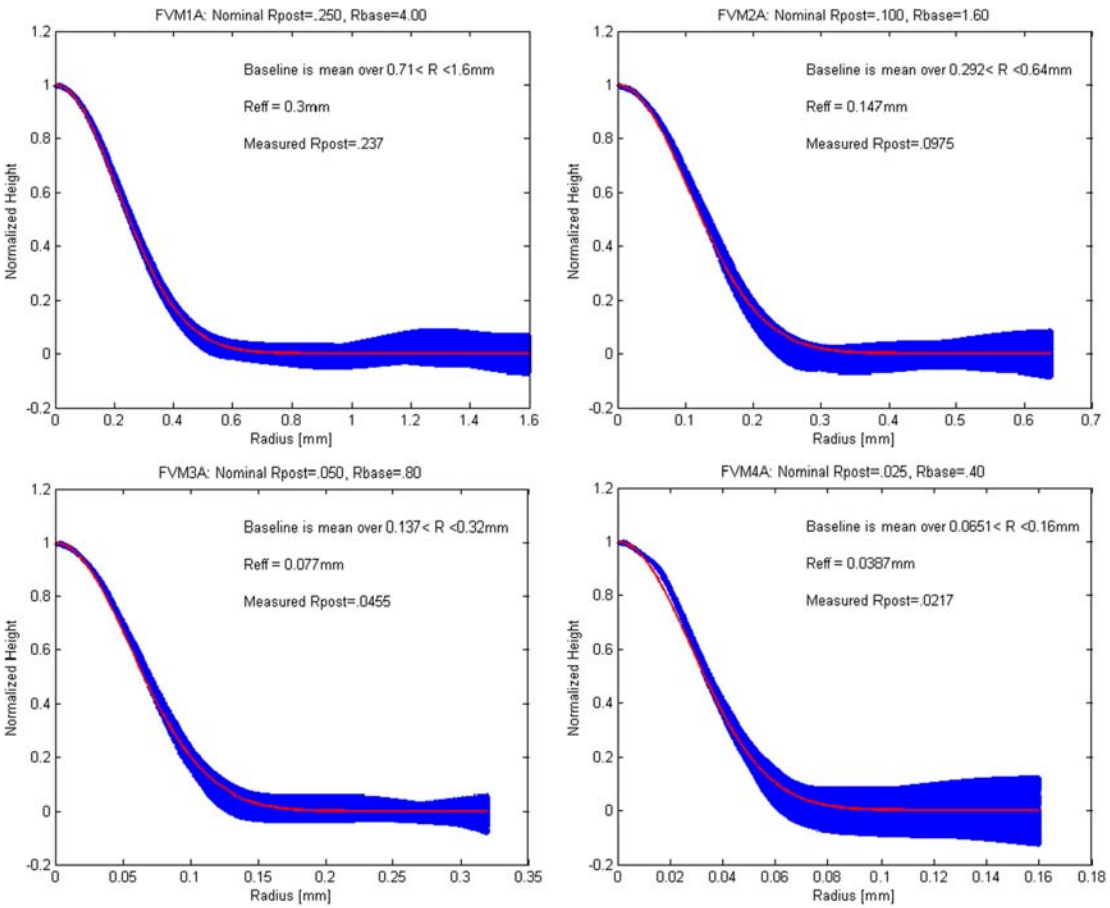


Figure 2. Measured radial profiles of the resolution targets at the four different magnifications (note changes in horizontal scale). All elevation data (azimuths covering 0 to 2π) are indicated as blue points; the red trace is the best-fit Gaussian.

along with the measured[5] pixel size p. Also listed is the inferred FOV of a 1197 × 1597 image with the measured pixel size. (Note that this considerably exceeds the manufacturer's specified FOV for the lens.) The effective area of the Gaussian PSF $A_{\text{eff}} = \pi R_0^2$ is also listed. Finally, $N = \text{FOV}/A_{\text{eff}}$ represents the number of statistically independent elevation measurements in a single FVM image.

The difficulty with using FVM measurements to estimate Sq and Sal is now apparent: the very limited lateral dynamic range. Although the measured PSF's are accurately Gaussian at all magnifications, the

3. Analytic correction of spatial bandwidth limitations

3.1. Exact and approximate analytic corrections

There are four primary spatial bandwidth effects to consider in the extraction of statistical surface parameters. When $R_0 \square L$, the limited lateral resolution (microscope PSF) discussed above will (a) decrease the observed σ^2 by averaging out the vertical surface fluctuations, and (b) increase the observed L by spreading out the fluctuations over a length scale R_0. However, since the PSF is accurately Gaussian with a

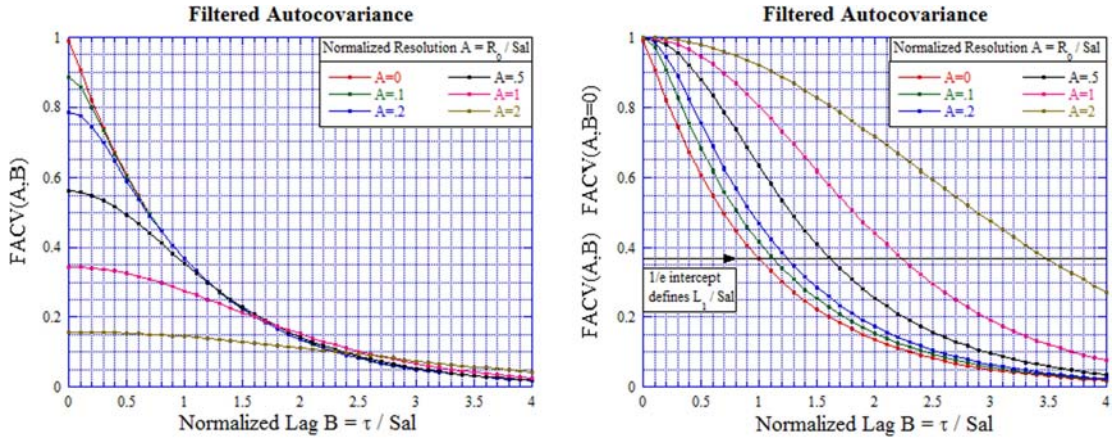


Figure 3. Apparent autocovariance after filtering, FACV$(A,\ B)$ as a function of lag, for various amounts of Gaussian filtering. Both lag and Gaussian filter width are normalized to the *true* autocorrelation length Sal. (Left) Autocovariance normalized to the true rms $(\mathrm{Sq})^2$. (Right) The same curves, but normalized to the *apparent*, i.e. observable, rms, σ_1^2, the $B = 0$ intercept. The $1/e$ points of these curves (roots of equation (5)) constitute the apparent autocorrelation lengths.

commercial instruments, including ours.) Finally, we have found that artifacts are sometimes introduced into FVM data by optical and illumination effects that increase markedly at the edges of the FOV, particularly at low magnification where the FOV is physically very large. All these effects make the low-frequency end of the instrument's spatial bandpass somewhat ill-defined. Therefore, while we can treat the high-frequency effects analytically by signal-processing arguments, the low frequency effects are only treated empirically, through the Monte Carlo simulations.

Let the actual physical surface $z\,(x,\ y)$ be described as a random variable with an isotropic and exponential ACV, with rms height variation Sq and autocorrelation length Sal. These are the true values of the surface's topography parameters. The surface's ACV and PSD are thus given by

$$\mathrm{ACV}(\tau) = \mathrm{Sq}^2 \mathrm{e}^{-\tau/\mathrm{Sal}}, \tag{1a}$$

$$\mathrm{PSD}(f) = \frac{2\pi\,(\mathrm{Sal})^2\,(\mathrm{Sq})^2}{[1 + (2\pi f\,\mathrm{Sal})^2]^{3/2}}. \tag{1b}$$

It is understood that both frequency and lag are defined over a 2D plane, but since we are specializing

$$\begin{aligned}\mathrm{FPSD}(f) &= |\mathrm{e}^{-\pi^2 R_0^2 f^2}|^2 \mathrm{PSD}(f)\\ &= \frac{2\pi\,(\mathrm{Sal})^2\,(\mathrm{Sq})^2}{(1 + (2\pi f\,\mathrm{Sal})^2)^{3/2}} \mathrm{e}^{-2\pi^2 R_0^2 f^2}.\end{aligned} \tag{2}$$

Here, FPSD(f) represents the surface's PSD when affected by the microscope's finite resolution, and the term within the absolute value sign is the properly normalized filter kernel (i.e. the Fourier transform of the microscope's PSF). Most of our calculations have been done in the spatial rather than frequency domain, so the corresponding ACV is

$$\mathrm{FACV}(R_0,\ \tau) = (\mathrm{Sq})^2 \int_0^\infty \frac{\mathrm{e}^{-\frac{A^2x^2}{2}} J_0(Bx)}{(1 + x^2)^{3/2}} x\mathrm{d}x, \tag{3}$$

where we have defined a dimensionless instrumental resolution $A = \dfrac{R_0}{\mathrm{Sal}}$ and a dimensionless lag $B = \dfrac{\tau}{\mathrm{Sal}}$. However, these dimensionless quantities are not directly accessible because they are normalized to the surface's true autocorrelation length, which is unknown. In general, there is no closed form solution for FACV$(A,\ B)$, but at zero lag, i.e. $\tau = B = 0$, the ACV gives the apparent rms, and in this limit (3) does have a closed form solution, namely:

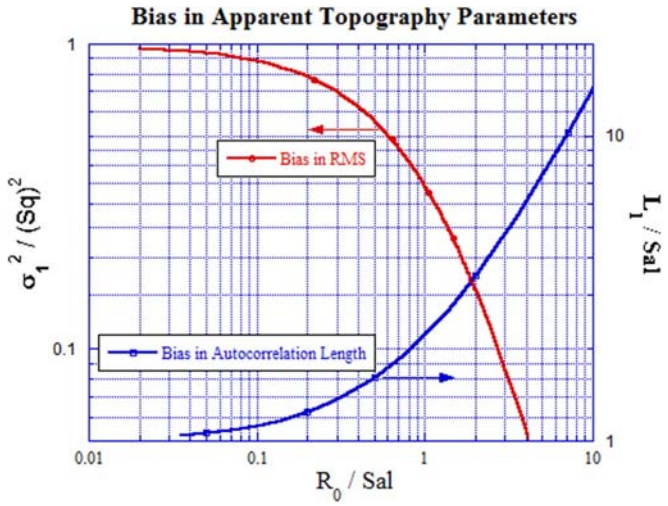


Figure 4. Bias in σ and L. (Left) apparent rms σ_1^2 (normalized to actual rms $(\mathrm{Sq})^2$), as a function of Gaussian filter width. (Right) apparent autocorrelation length normalized to true autocorrelation length, also as a function of Gaussian filter width.

normalized to the *apparent* rms σ_1^2. These plots correspond directly to the schematic illustration of figure 3 of [17]. However, these filtered ACV's are not subject to the requirement $\int_0^\infty \mathrm{FACV}(\tau)2\pi\tau\mathrm{d}\tau = 0$, and therefore do not necessarily go negative at large τ, because the zero-frequency PSD has not been set to zero by requiring a zero mean elevation. Figure 4(left) shows the ratio of apparent rms to true rms, obtained from the $\tau = 0$ intercept of the curves in figure 3. Figure 4(right) shows the ratio of apparent autocorrelation length to true autocorrelation length obtained from the $1/e$ points in the curves in figure 3.

The limiting cases for figure 4 are notable: when the Gaussian filtering is negligible ($R_0 \ll \mathrm{Sal}$), the apparent autocorrelation length approaches the true value $L_1/\mathrm{Sal} \to 1$, and when the Gaussian filtering is dominant and the surface's intrinsic fluctuations are highly filtered, ($R_0 \gg \mathrm{Sal}$), the apparent autocorrelation length is simply $\sqrt{2}\,R_0$, the value obtained by convolving (in 2D) two Gaussians of $1/e$-radius R_0. One can therefore parametrize the strength of the Gaussian filtering in terms of a new variable $a = \dfrac{L_1}{\sqrt{2}\,R_0} - 1$,

This has been done using the numerical solutions of equation (5); the result for L_1/Sal is shown in figure 5. It is clear that in the highly filtered limit, the correction factor has an $a^{-1/2}$ dependence. This is a statement about how fast the solution to equation (5) approaches its limiting value $L_1 \to \sqrt{2}\,R_0$ and is not necessarily obvious from the analytic formulae. A useful fact is that the numerical result can be very well fit with an extremely simple empirical approximation

$$\alpha \equiv \frac{L_1}{\mathrm{Sal}} \approx \left(1 + \left(\frac{5}{a^{3/2}}\right)\right)^{1/3}. \qquad (6)$$

Over the entire interesting range of a, $0.05 \ll a \ll 20$, this approximation has a maximum fractional residual of 3.6%, and an rms residual of 2.4%, which should be adequate for all but the highest precision requirements. The numeric solution of equation (5)—as shown in figure 5 and including its empirical approximation—is the core result of this section. By using it to correct experimental measurements, an accurate estimate of Sal can be obtained (namely $L_2 = L_1/\alpha$). Since R_0 is known *a priori*, figure 4 or equation (4) can then be used to obtain an accurate estimate of Sq, which we denote σ_2, from the experimentally mea-

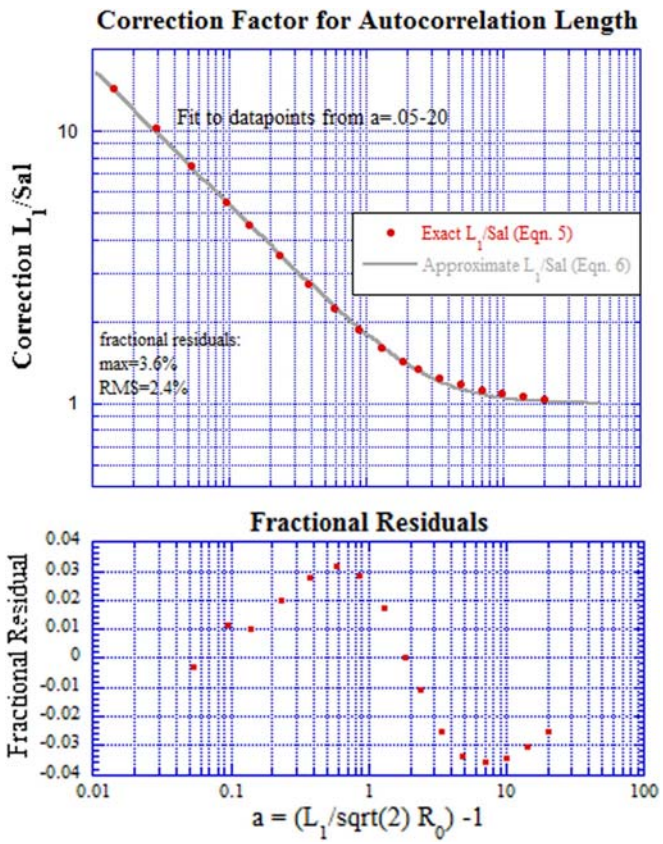


Figure 5. Correction factor for the autocorrelation length, as a function of the observable parameter a. The red dots indicate the numeric solution of equation (5), while the continuous gray curve is the approximation of equation (6).

autocorrelation length (up to $L_1/\mathrm{Sal} \approx 600\%$ based on figure 5) to be reliably corrected.

3.2. Measurement protocol

Based on all the above, an explicit protocol for estimation of Sal and Sq can be summarized as follows:

(1) Determine R_0 as accurately as possible, using the same magnification and microscope settings as will be used in measurement of the subject surface.

(2) Perform the FVM measurements of the subject surface, and extract from them the first-order estimate L_1.

(3) If $L_1 < \sqrt{2}R_0$, then the sampled area of the surface is simply too small to make a meaningful estimate of Sal or Sq. The FVM measurements need to be retaken at lower magnification. (Alternatively, adjacent fields need to be stitched to provide a larger FOV.) We shall see in section 5 that an additional criterion exists relating the surface rms and the microscope's vertical resolution that also needs to be satisfied, but in practice this seems to be violated by only the smoothest samples.

(4) However, if $\sqrt{2}R_0 < L_1 < \mathrm{FOV}/2\pi$, then calcu-

than $\mathrm{FOV}/2\pi$ (this numeric value is justified in the Monte Carlo simulations below),

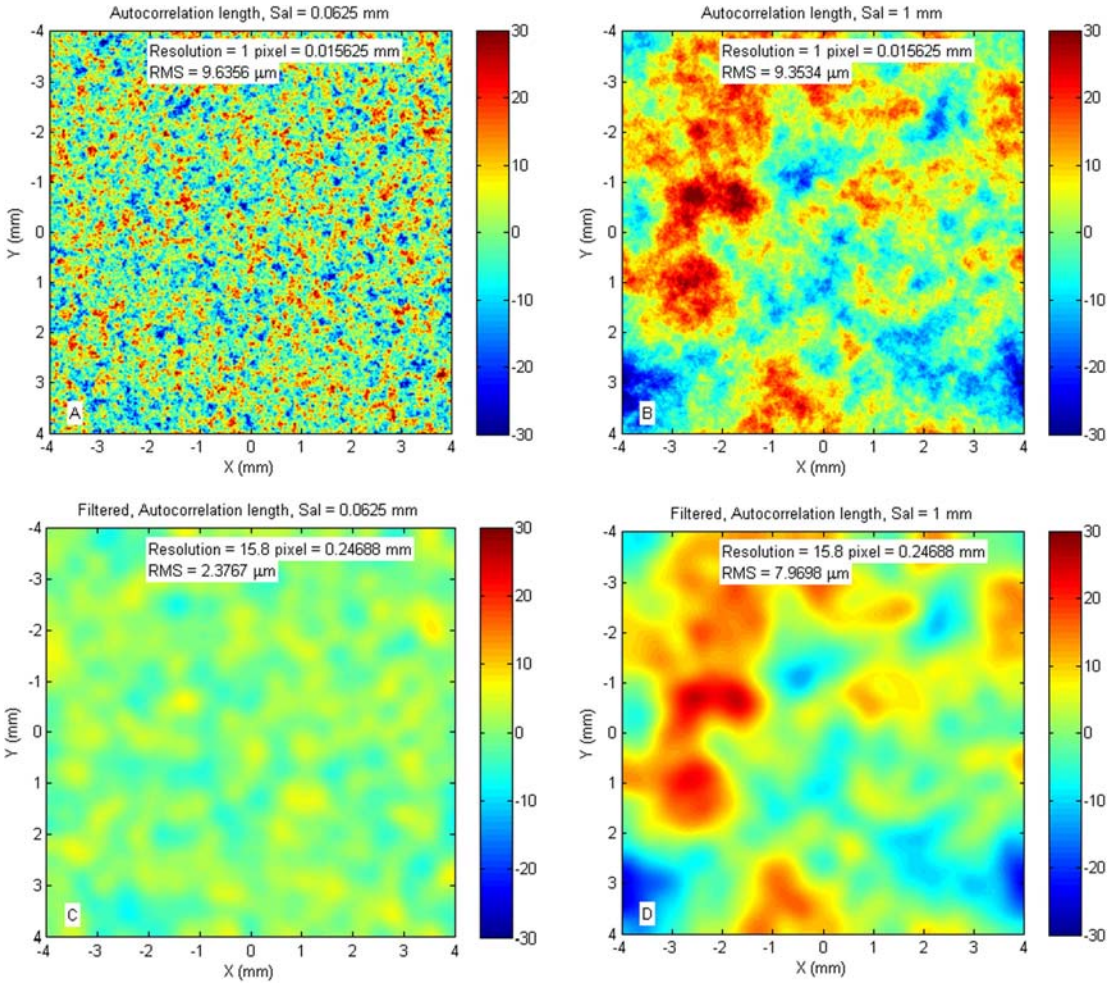


Figure 6. Monte Carlo simulated topographies of exponentially correlated, 8 × 8 mm^2 surfaces, with autocorrelation lengths of Sal = 0.0625 mm (left) and 1.0 mm (right). The vertical (color) scale is held fixed for all four panels, and extends over ±3Sq. The upper panels are unfiltered, with an effective resolution of ~1 pixel, while the lower panels have been filtered using a Gaussian kernel with $1/e$ radius of 0.25 mm ≈ 16 pixels.

develop at least an empirical understanding of the effects of finite FOV, Monte Carlo simulations were performed on exponentially correlated mock topographies. Fields of 512 × 512 pixels were used, which allowed all calculations, even compute-intensive correlations, to be performed in reasonable times (approx. two minutes or less) on a typical desktop

the same mock topographies after the Gaussian filter is applied. Following the protocol outlined in section 3, first-order estimates σ_1 and L_1 were calculated for both the filtered and unfiltered maps. For the filtered maps, the corrections described in section 3 were also applied and σ_2 and L_2 were calculated. This entire procedure was followed for mock topographies with L_0 varying

Table 2. Spatial bandwidth limits for 10% accuracy.

	Unfiltered		Filtered		Filtered and corrected	
	σ_1	L_1	σ_1	L_1	σ_2	L_2
Minimum Sal(pixels)	1.56	2.49	n/a	62.9	<4	<4
Maximum Sal(pixels)	69.5	60.1	n/a	78.0	73.1	38.2
Fractional bandpass (octaves)	5.5	4.6	0	0.3	□4.2	□3.3

Sq is given by $f_{max}/f_{min} = \mathrm{FOV}/8p$, which is 6 octaves for our simulations. Thus our criterion is slightly more conservative than that of [10], but this is easily explained by differences in assumptions regarding the instrumental transfer function, 2D versus 1D treatment, etc. However, the basic point remains the same: the range of surfaces (i.e. range of autocorrelation lengths) for which the instrument and a first-order algorithm can provide accurate estimates is limited: by finite lateral resolution at short lengths, and by finite FOV at large lengths.

An important point is that by adjusting the microscope magnification, any sample (whose Sal is some fixed length in mm) can be moved into the 'good' ~5 octave range of bandwidth. Changing the microscope magnification effectively shifts the 'good' bandpass of the microscope along the horizontal axis; the FOV and resolution R_0 change by equal factors. Microscope magnification is typically adjustable in discrete increments, in our case differing by approximately a factor of $\sqrt{2}$, (i.e. half an octave), so a 'good' bandwidth of 5 octaves suffices for accurate estimates, even if the magnification is not optimal. Moreover, the width of the plateau ensures that several adjacent magnification values will all yield consistent estimates of Sq and Sal.

As stated above, each mock topography (such as figures 6(a) and (b)) was spatially filtered using a Gaussian kernel. The $1/e$ width for the filter was scaled to provide equivalent dynamic range to that of the microscope. Thus, while the microscope FOV is

Monte Carlo-generated topographies like those in figure 6, when filtered with Gaussian kernels as described above. For the fine-featured surfaces (Sal □ R_0), major errors are introduced into the first-order estimates by the filtering. The first-order estimate σ_1 is biased low while the first-order estimate L_1 is biased high. Moreover, the errors introduced by the filtering for small Sal (i.e. at high frequencies) overlap with the errors introduced by the finite FOV for larg Sal (i.e. at low frequencies). Therefore, in contrast to the unfiltered case (red points), there is no plateau, no fraction of the available spatial bandwidth for which first-order estimates σ_1 and L_1 are reasonably accurate. The σ_1 is guaranteed to be biased low at all magnifications, while the L_1 could be biased high or low. Thus, when measuring an unknown sample, even with an optimal choice of magnification, errors greater than 30% in .. are expected. If a magnification is used that is even slightly non-optimal, much larger errors are possible. This is reflected in the central columns of table 2, indicating there is no range of <10% accuracy for rms roughness, and where the range of accuracy for autocorrelation length is so narrow as to be useless. As mentioned earlier, the dynamic range of these mock topographies (i.e. the ratio between their FOV and resolution) has been chosen to match (approximately) the dynamic range of our commercial focus-variation microscope. Therefore an alternate way of describing the message of figure 8 is: *A single (i.e. unstitched) FVM field does not have sufficient dynamic range to allow accurate extraction of* Sq *and* Sal*, unless the correction factors described earlier are applied.*

The correction factors described above (i.e. σ_2/σ_1 and L_2/L_1) correct this problem and produce unbiased estimates over their range of applicability. This is shown by the blue points in figure 8, and in the final two columns of table 2. For autocorrelation lengths from 4 to 32 pixels, and possibly 64, (i.e. over 3–4 octaves of bandwidth) the corrected values σ_2 and L_2 accurately reproduce the actual values.

4.2. Uncertainties

The above protocol corrects the biases introduced into σ_1 and L_1 by the limited lateral resolution, by exploiting

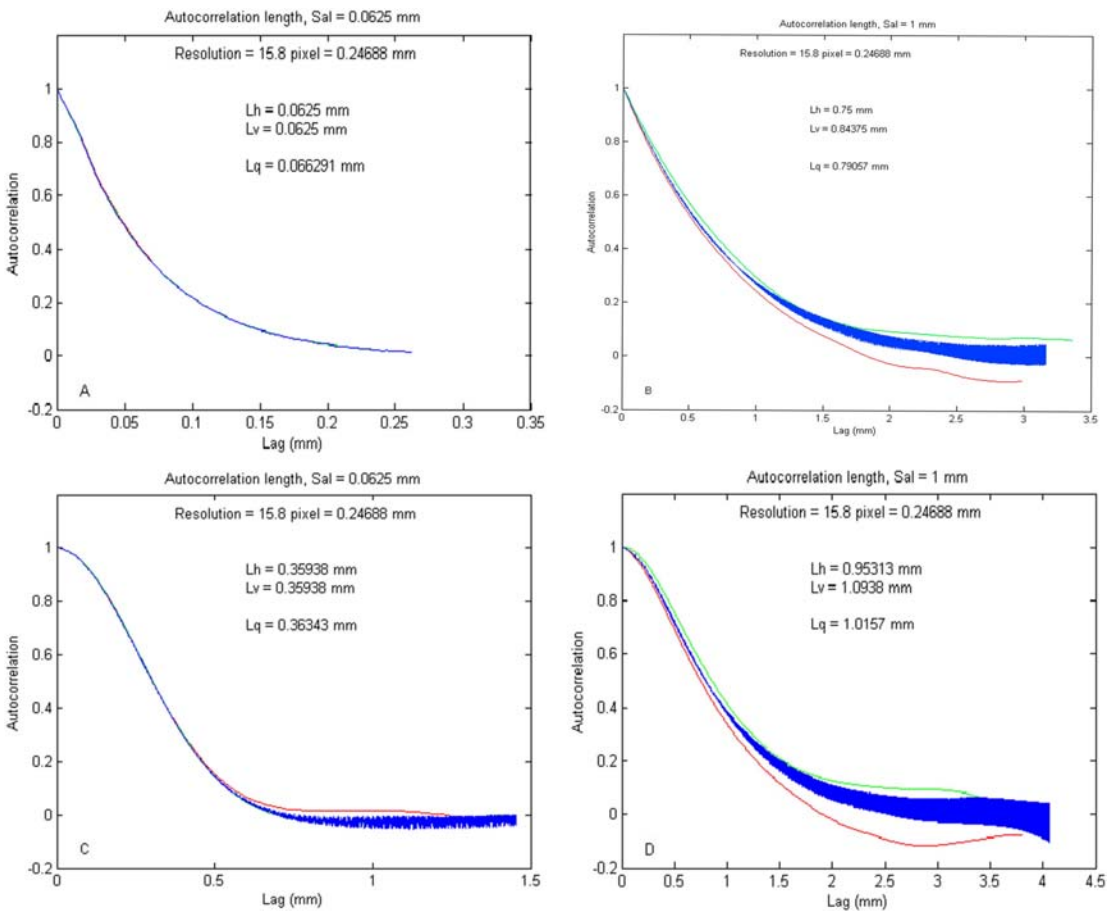
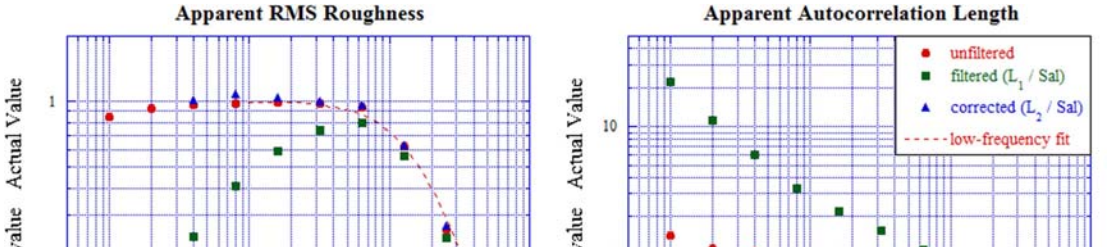


Figure 7. First-order autocovariance functions $\mathrm{ACV}^{(1)}$, normalized to σ_1^2 at the origin, corresponding to the four Monte Carlo topographies in figure 9: (a) unfiltered, with Sal = 0.0625 mm (b) unfiltered Sal = 1.00 mm, (c) filtered, with Sal = 0.0625 mm, and (d) filtered, with Sal = 1.00 mm . L_h and L_v, from the red and green traces respectively, are autocorrelation lengths in the horizontal and vertical directions respectively, while L_q, from the blue points is the autocorrelation length from the average of all points in the lag plane (overwhelmingly not in the cardinal directions).



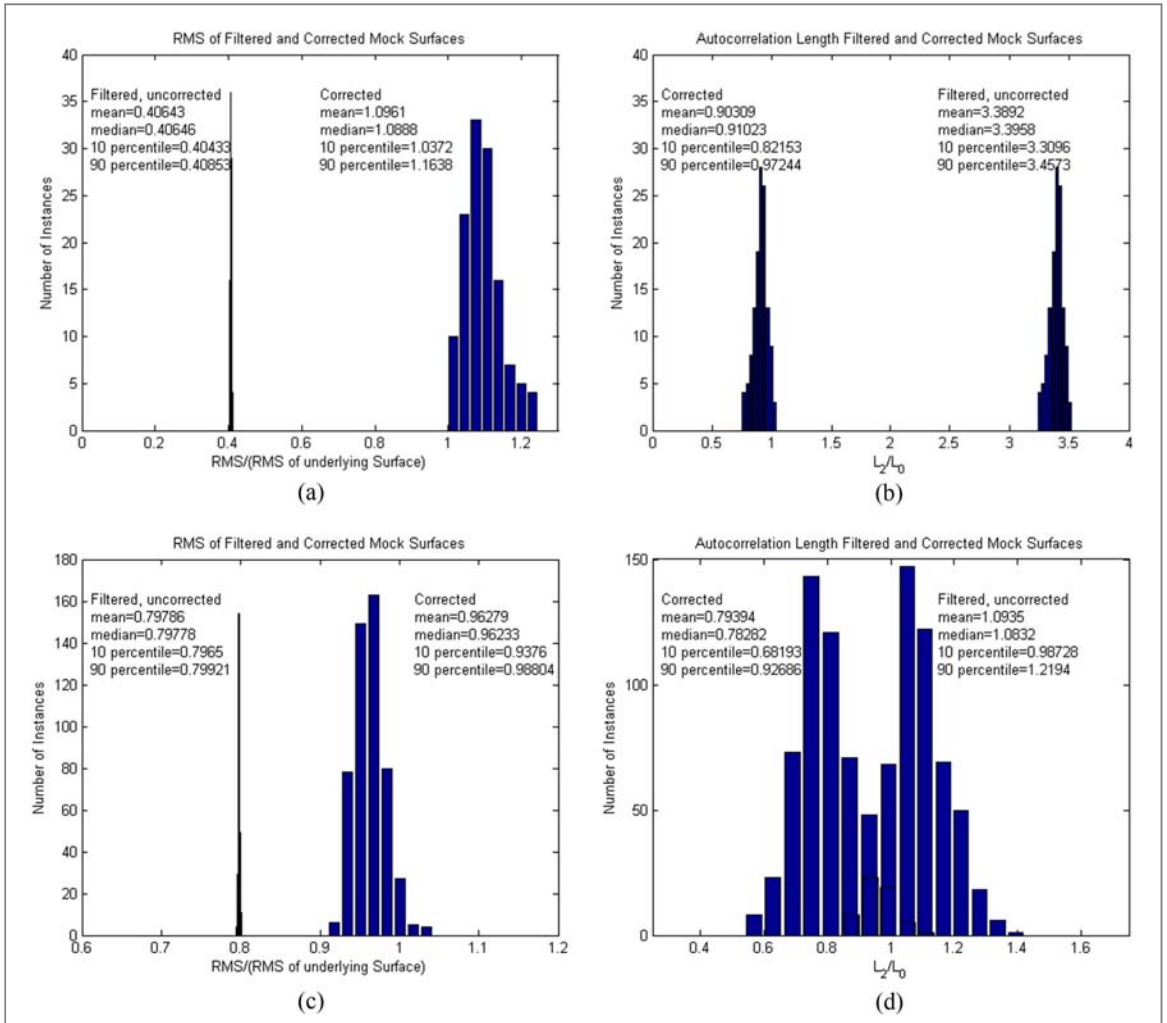


Figure 9. Histograms of extracted parameters from the Monte Carlo simulations, before and after applying the correction algorithm. (a) and (b) display rms and autocorrelation length respectively, for surfaces with Sal $=$ 0.125 mm, leading to large corrections; (c) and (d) display the same for surfaces with Sal $=$ 1.0 mm, for which the corrections are relatively small.

fractional uncertainty in the extracted value of L_2. For example, if a is experimentally determined as $a = 0.1 \pm 0.03$, then the correction factor is $\alpha = 5.35 + 1.15/-0.6$ and the fractional uncertainty in L_2 is $(+21\%/-11\%)$.

To illustrate this more explicitly, and to obtain quantitative estimates of uncertainties in the extraction algorithm for realistic parameters, a separate

roughness for many applications. The same holds true for the estimated values of autocorrelation length. The final conclusion is that over this 3-octave range in roughness scale (i.e. 8:1 in Sal), the corrected estimates σ_2 fall within $\pm 3\%$ of their mean values 80% of the time, and the corrected estimates L_2 fall within $\pm 9\%$ of their mean values 80% of the time. The mean values are still slightly biased from the true values

$$F(\mathrm{Sal}) = \left[1 + \left(\frac{\pi \mathrm{Sal}}{\mathrm{FOV}}\right)^2\right]^{-1} \quad \text{for } \sigma,$$

$$F(\mathrm{Sal}) = \left[1 + \left(\frac{2\pi \mathrm{Sal}}{\mathrm{FOV}}\right)^2\right]^{-1/2} \quad \text{for } L. \qquad (7)$$

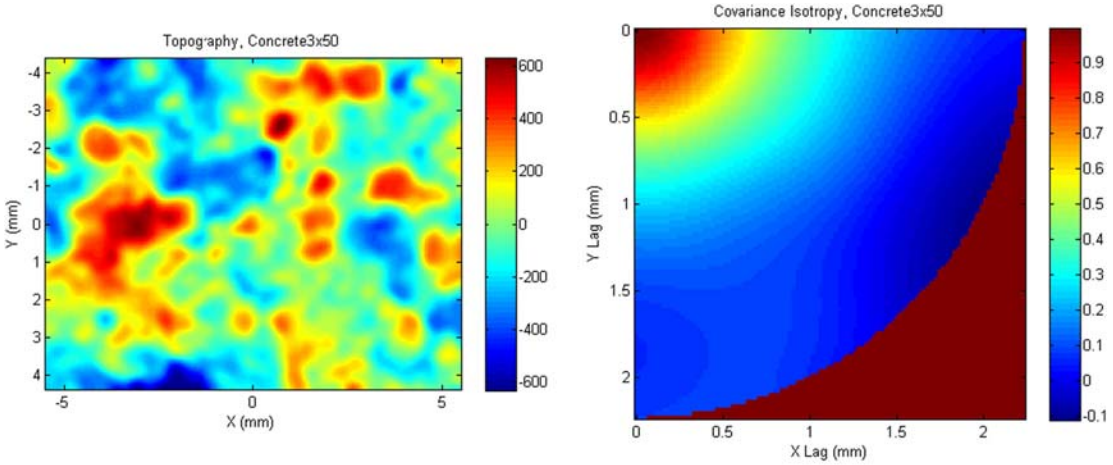
For comparison, Church [17] states 'when $L_1 \square$ FOV$/8\pi$ then L_1 may be taken as approximately equal to Sal and σ_1 is essentially equal to Sq.' The estimates in (7) are somewhat less stringent than this. They imply that biases in (first-order estimates of) σ and L remain below 5% as long as $L_1 \square FOV/6.1\pi$ and that they remain below 10% for $L_1 \square FOV/4.1\pi$. However, they describe the measurement biases for surfaces whose *only* low-frequency instrumental artifact is finite FOV. Real FVM measurements also include a (necessarily imperfect) correction for tilt and sometimes a correction for large-scale curvature. Moreover, FVM imagery is often ring-illuminated. Therefore, non-uniformity in illumination intensity and direction begins to appear at length scales comparable to the objective lens diameter (entrance aperture). This will be approximately the FOV at the lens's lowest magnification. Finally, although FVM objectives are very highly corrected, any residual off-axis aberrations in the objective will be greatest at the edges of the FOV. For these reasons, the measurement biases described by (7) are really best-case figures, and some judgment needs to be exercised when extracted values of autocorrelation length approach FOV$/4\pi$. We have found that increasing the FOV by 'stitching' together multiple microscope fields is a reliable check that measurement biases are not excessive.

Some FVM fields (including the unstitched fields from our instrument) are not square. In these cases the biases introduced by low frequency effects will be slightly different in horizontal and vertical directions, potentially making an isotropic surface appear slightly anisotropic, with different L and σ in the horizontal and vertical directions. However, the fits given in equation (7) imply that this effect is very small. The spurious anisotropy only becomes significant when the bias introduced by the FOV already is large. If the $L_1 <$ FOV$/2\pi$ requirement of the protocol (section 3, step (3)) is observed, then this effect can be neglected.

Eleven samples of common outdoor building materials were obtained, and each was imaged by FVM at four different magnifications: × 30, × 50, × 100, and × 200. The overall FOV was held approximately constant for all magnifications, at 11.1 × 8.2 mm, by stitching together smaller subfields for the higher magnifications. (Thus, there were 2 × 2 subfields at × 50, 4 × 4 at × 100, and 8 × 8 at × 200.) In stitched datasets, the edges of the various subfields were not perfectly aligned, so the first step in analysis was to find the largest inscribed rectangle within the dataset. In addition, this rectangular dataset was cropped, eliminating all data within one lateral resolution length (i.e. 32 pixels, independent of magnification) of the boundary. Next, if either dimension of the resulting dataset was >1024 pixels in size, the dataset was averaged and downsampled by a factor of × 5 in each dimension. This was simply to accelerate the later calculations. It had negligible effect on any resulting statistics because the original data were so heavily oversampled to begin with (by approximately a factor of × 32 in each dimension). Then the mean and best-fit linear tilt were removed from the dataset. The result was the basic topography map of the rough surface. Then the first order values σ_1 and L_1 were calculated. (See appendix B for explicit formulae).

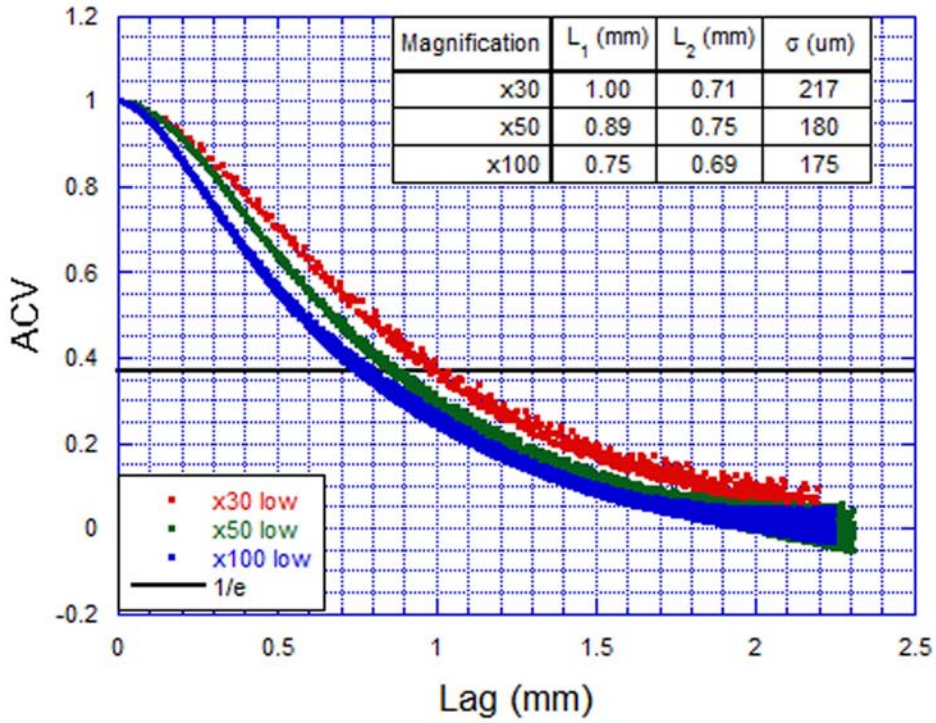
The calculated normalized autocorrelation was actually a 2D array of ACV as a function of lag (τ_x, τ_y). This array was first plotted directly, as a test of the surface's intrinsic isotropy. Then, using the assumption that it was sufficiently isotropic, all points in the array were plotted as a 1D function of $\tau = (\tau_x^2 + \tau_y^2)^{1/2}$. The point where this function fell to $1/e$ of its 0-lag value was identified as L_1. Finally, the algorithm

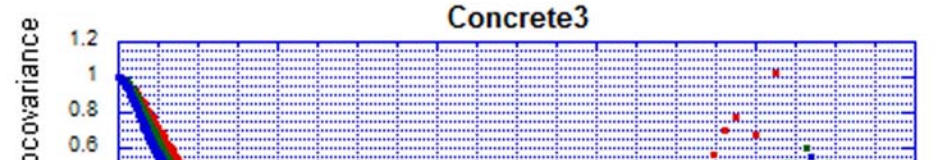
be exponential (for example it might have two different scale lengths, or even fractal statistics), or its height distribution may not be Gaussian. Indeed, one could argue that the greatest value of a careful statistical analysis of FVM imagery lies in identifying those surfaces for which 'standard' statistical measures Sq and Sal are *not* meaningful (and why they are not). Non-etheless, a significant effort was made to apply the above algorithm to real surface topographies, and thus to extract a single σ_2 and L_2 for each.



Concrete3

Magnification	L_1 (mm)	L_2 (mm)	σ (um)
x30	1.00	0.71	217
x50	0.89	0.75	180
x100	0.75	0.69	175





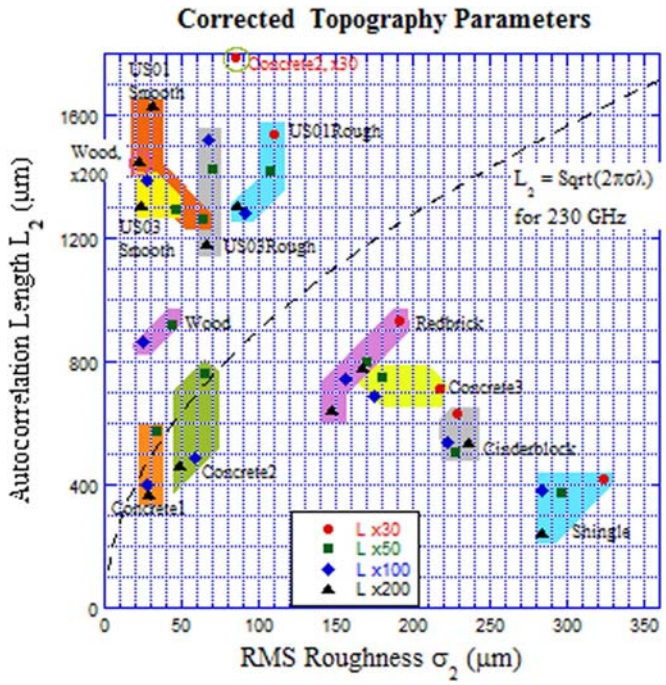


Figure 11. Estimates of rms roughness and autocorrelation length obtained from FVM measurements, using the protocol described in section 3, on 11 physical samples of outdoor building materials.

Table 3. Extracted rms roughness and autocorrelation length for outdoor building materials.

	Rms roughness σ_2 (μm)				Autocorrelation length L_2 (mm)			
Sample	× 30	× 50	× 100	× 200	× 30	× 50	× 100	× 200
Concrete1	n/a	34.0	27.0	28.3	n/a	0.578	0.401	0.370
Concrete2	84.2	64.8	58.1	49.0	1.79	0.763	0.492	0.462
Concrete3	217	180	175	167	0.713	0.754	0.691	0.784
Cinderblock	229	227	222	236	0.630	0.508	0.540	0.542
Red brick	191	169	156	148	0.935	0.800	0.747	0.643
Shingle	323	296	284	284	0.419	0.378	0.386	0.247
US01 Rough[a]	110	107	90.8	86.6	1.54	1.42	1.28	1.31
US01 Smooth	n/a	63.4	26.8	31.6	n/a	1.27	1.39	1.63
US03 Rough	n/a	70.2	67.3	65.9	n/a	1.43	1.52	1.19
US03 Smooth	n/a	46.3	26.8	23.1	n/a	1.29	1.39	1.31
Wood	n/a	44.0	24.1	22.6	n/a	0.922	0.866	1.45

[a] US01 etc denote varieties of commercial building stone.

autocorrelation length estimates are obtained for (at × 100), reflecting the intuitive fact that wood's

determination of vertical resolution in an FVM is beyond the scope of this paper, (and it is not specified bt the manufacturer for this particular commercial instrument), we have found that a quadratic scaling with magnification

$$\delta z \approx \frac{80\ \text{mm}}{(\text{magnification})^2}, \tag{8}$$

provides a reasonable estimate in most cases. At this level, vertical artifacts (for example the residual flatness of [14]) may be introduced into the FVM data by the instrument and these have a lateral scale that is completely unrelated to the sample. In other words, samples that are too smooth (Sq too low) cannot be measured with too low a magnification, regardless of their autocorrelation length. This is essentially a third criterion that must be met by an FVM dataset, in addition to the FOV and resolution criteria stated in section 3.2, in order to obtain valid σ and L measurements.

The dashed line in figure 11 indicates the condition $\frac{\sigma\lambda}{L^2} = \frac{1}{2\pi}$ for $\lambda = 1.3$ mm (corresponding to a frequency of 230 GHz). It indicates a family of surfaces whose mean surface curvature is comparable to a mm-wave radar wavelength. The ubiquitous Kirchoff approximation for rough surface scattering is a low-curvature approximation [13, 23], corresponding to the region in σ–L space lying above the curve in figure 11. It describes the angular distribution of scattered light under directional illumination, i.e. the balance between specular (mirror-like) reflection and perfectly diffuse, Lambertian scattering. The fact that many of the outdoor building materials fall outside that region indicates that more sophisticated treatments of scattering than the Kirchoff approximation are required in order to model the qualitative appearance of everyday outdoor scenes, when viewed at millimeter wavelengths.

6. Conclusions

FVM provides much less lateral dynamic range (fractional spatial bandwidth) than one might expect based on typical camera formats. For our particular instrument of a surface with an exponential, isotropic ACV and Gaussian height distribution and an instrument with a Gaussian PSF. As shown by the Monte Carlo simulation in section 4, nearly all of the spatial bandwidth that is lost in the FVM's poor lateral resolution can be recovered by these correction factors, which perform a similar function to a deconvolution algorithm. Although the correction factors are mathematically correct only for isotropic, exponentially correlated surfaces, we have found that the measurement protocol yields σ and L values that are consistent across magnification for nine out of eleven real-world samples of outdoor building materials that we have examined. The exceptions correspond to well understood violations of the measurement protocol's assumptions. Use of this measurement protocol (including the correction factors) therefore enables unstitched FVM data to be used for the reliable and robust evaluation of σ and L of unknown surfaces, on instruments where it would otherwise be subject to very large and magnification-dependent biases.

Acknowledgments

The authors are grateful to the Department of Homeland Security Science and Technology Directorate and to the Systems Technology Office of DARPA for support of this work. Ms Nina Basta, and Drs Aric Sanders and Richard Chamberlin provided helpful comments regarding operation of our focus-variation microscope and analysis of its data.

Appendix A. Width of point-spread function

To be more precise, let the FVM point-spread function be a Gaussian

$$S(r) = \frac{1}{\pi R_0^2}\exp\left(-r^2/R_0^2\right)$$

that integrates to 1 and has $1/e$ radius R_0. Let the resolution target be represented by a 'top hat' function

$$G(r) = 1, \quad r \square R_p,$$

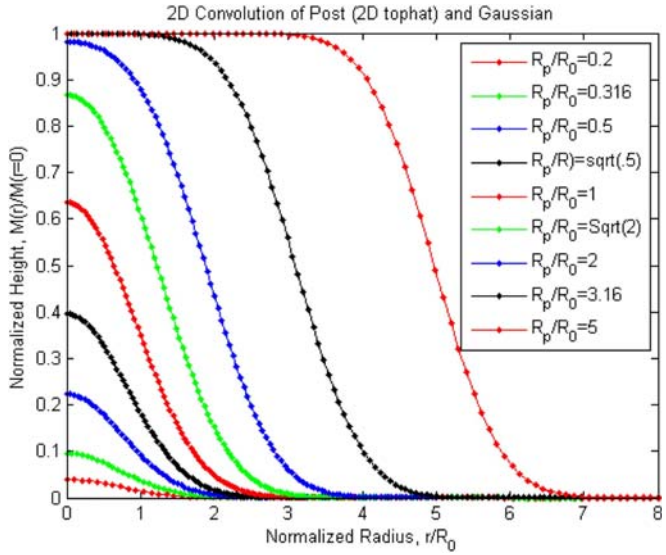


Figure A1. Expected form of the FVM profile from a resolution target, namely convolution of a Gaussian point-spread function and a top hat topography.

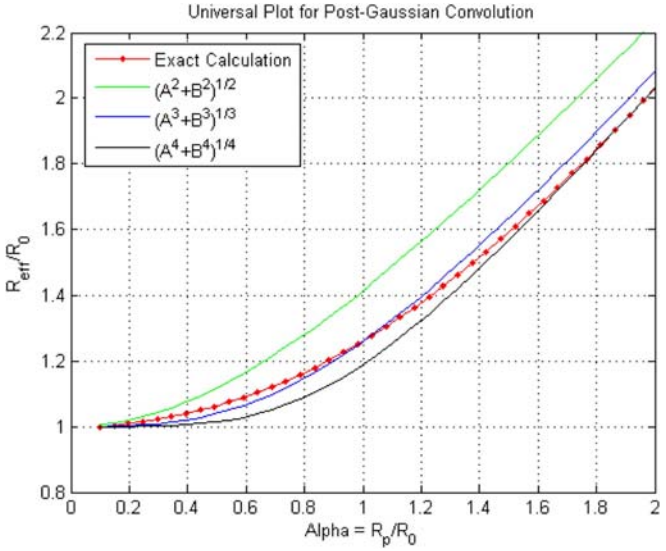


Figure A2. Universal plot of effective radius for FVM profile of a resolution target, along with simple analytic approximations.

σ_1 and L_1 can be made simply from the mean, rms, and autocorrelation of the raw FVM data:

$$\bar{z} = \frac{1}{N_x N_y}\sum_{ij} z_{ij},$$

$$\sigma_1^2 = \left(\frac{1}{N_x N_y}\sum_{i,j}(z_{ij} - \bar{z})^2\right),$$

$$P_{mn} = \frac{\sum_{i=1}^{N_y}\sum_{j=1}^{N_x}(z_{ij} - \bar{z})(z_{i-m,j-n} - \bar{z})}{(N_y - |m|)(N_x - |n|)} \quad \text{where}$$

$$-(N_y - 1) \square m \square (N_y - 1),$$

$$-(N_x - 1) \square n \square (N_x - 1).$$

The expression in the denominator of P_{mn} is simply the number of 'overlaps' between z_{ij} and a displaced (by (m, n) pixels) version of itself. Note that P_{mn} uses 'zero-based indexing' and has dimension $(2N_y - 1) \times (2N_x - 1)$. The expression in the numerator of P_{mn} is simply the Matlab definition of autocorrelation (i.e. xcorr(A, A)). The first-order estimate of the ACV is simply the symmetrized version of P_{ij}:

$$\mathrm{ACV}_{ij}^{(1)} = \frac{1}{4}(P_{m,n} + P_{-m,n} + P_{m,-n} + P_{-m,-n}) \quad \text{for}$$

$$i = m = 0 \ldots N_y, j = 0 \ldots N_x.$$

By definition, the value of $\mathrm{ACV}^{(1)}$ at the origin $\mathrm{ACV}_{00}^{(1)} = \sigma_1^2$, as it must. The first-order estimate L_1 is the value of $\tau_{ij} = \sqrt{x_j^2 + y_i^2}$ where the value of $\mathrm{ACV}^{(1)}$ falls to $1/e$ of that value: $\mathrm{ACV}^{(1)}(\tau_{ij} = L_1) = \dfrac{\sigma_1^2}{e}$.

References

[1] Leach R K 2010 *Fundamental Principles of Engineering Nanometrology* (Oxford: Elsevier)

[2] Leach R K E 2011 *Optical Measurement of Surface Topography* (Berlin: Springer)

[3] Jiang X, Scott P J, Whitehouse D J and Blunt L 2007 Paradigm shifts in surface metrology: I. Historical philosophy *Proc. R. Soc.* A **463** 2049–70

[4] ISO 25178-6 2010 *Geometrical Product Specifications (GPS)—Surface Texture: Areal Part 6: Classification of Methods for Measuring Surface Texture* (Geneva, Switzerland: ISO)

[5] Danzl R, Helmli F and Scherer S 2009 Focus variation—a new technology for high resolution optical 3D surface metrology *Application of Contemporary Non-Destructive Testing in Engineering (Ljubljana, Slovenia)*

[6] Danzl R, Helmli F and Scherer S 2011 Focus variation—a robust technology for high resolution optical 3D surface metrology *J. Mech. Eng.* **57** 245–56

[7] Giusca C L and Leach R K 2013 Calibration of the scales of areal surface topography instruments: III. Resolution *Meas. Sci. Technol.* **24** 105010

[8] Giusca C L, Leach R K, Helary F, Gutauskas T and Nimishakavi L 2012 Calibration of the scales of areal surface topography-measuring instruments: I. Measurement noise and residual flatness *Meas. Sci. Technol.* **23** 035008

[9] Giusca C L, Leach R K and Helery F 2012 Calibration of the scales of areal surface measuring instruments: II. Amplification, linearity and squareness *Meas. Sci. Technol.* **23** 065005

[10] Leach R K, Giusca C L, Haitjema H, Evans C and Jiang X 2015 Calibration and verification of areal surface texture measuring instruments *CIRP Ann. Manuf. Technol.* **64** 797–813

[11] Giusca C L, Claverly J D, Sun W, Leach R K, Helmli F and Chavigner M P J 2014 Practical estimation of measurement noise and flatness deviation on focus variation microscopes *CIRP Ann.— Manuf. Technol.* **63** 545–8

[12] ISO 25178-606 2015 *Geometrical Product Specification (GPS)—Surface texture: Areal— Part 606: Nominal Characteristics of Non-contact (Focus Variation) Instruments* (ISO)

[13] Beckmann P A and Spizzichino A 1987 *The Scattering of Electromagnetic Waves from Rough Surfaces* 2nd edn (Norwood, MA: Artech House)

[14] Haitjema H 2015 Uncertainty in measurement of surface topography *Surf. Topography: Metrol. Prop.* **3** 35004

[15] ISO 25178-601 2010 *Geometrical Product Specifications (GPS) — Surface Texture: Areal, Nominal Characteristics of Contact (Stylus) Instruments* (ISO)

[16] Haitjema H and Morel M A A 2005 Noise bias removal in profile measurements *Measurement* **38** 21–9

[17] Church E L 1986 Comments on the correlation length *Proc. SPIE* **680** 102–11

[18] Church E L and Takacs P Z 1988 Instrumental effects in surface finish measurement *Proc. SPIE* **1009** 46–55

[19] Church E L and Takacs P Z 1991 The optimal estimation of finish parameters *Proc. SPIE* **1530** 71–85

[20] Leach R and Haitjema H 2010 Bandwidth characteristics and comparisons of surface texture measuring instruments *Meas. Sci. Technol.* **21** 032001

[21] ISO 25178-2 2012 *Geometrical Product Specifications (GPS)—Surface Texture: Areal Terms, Definitions and Surface Texture Parameters* (ISO)

[22] Nikolaev N, Petzing J and Coupland J 2016 Focus variation microscope: linear theory and surface tilt sensitivity *Appl. Opt.* **55** 3555–65

[23] Voronovich A G 1999 *Wave Scattering from Rough Surfaces* vol 17 2nd edn (Berlin: Springer)