Surface Topography: Metrology and Properties

PAPER

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

E N Grossman¹, M Gould² and N P Mujica-Schwann¹,³

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 statistical parameters of a rough surface. The most widespread need is in quality control for manufactured surfaces (this application has driven the development of most commercial instruments). However, the present work is motivated by interest in electromagnetic scattering from common everyday surfaces (roughness levels of 10–500 μm) such as would be encountered in many applications of submillimeter-wave radar. Historically, surface topography has been measured with stylus-based instruments [1]. As a fundamentally one-dimensional (1D) technique, this can be very time-consuming, particularly if samples are anisotropic and/or numerous. Within the last decade however, a gradual conversion has taken place to optical techniques [2, 3] that are fundamentally two-dimensional (2D). International standards are being rewritten [4] as new types of non-contact instrumentation arrive in the commercial marketplace, including the focus-variation microscope ([5, 6] and chapter 7 of [2]). In some cases, only qualitative or descriptive information is needed from 3D imagery, but in others, including analysis of submillimeter-wave scattering, reliable quantitative measurements, including uncertainties, are required. A comprehensive study of measurement noise, resolution and other uncertainties has recently been presented for the coherence scanning interferometer, imaging confocal.
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 $S_q$ and the autocorrelation length $S_l$ 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.) $S_q$ and $S_l$ 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 $S_q / S_l$ (the mean surface slope) enters, but in more sophisticated treatments, this is not the case. Implicit in the use of a single autocorrelation length $S_l$ are the assumptions that (a) the surface’s ACV is isotropic—meaning that the ACV is independent of direction in the 2D lag plane—and (b) that it can be approximated as exponential, so that $ACV(\tau_x, \tau_y) = (S_q)^2 e^{-|\tau|/S_l}$, where $\tau = \sqrt{\tau_x^2 + \tau_y^2}$. This is by far the most commonly assumed form for $ACV(\tau_x, \tau_y)$, although other forms have also been treated by scattering theories, and FVM data can also be used to assess the validity of this approximation for any particular surface.

This form of the ACV is assumed throughout this paper (although some exceptions are mentioned in section 5). Also generally assumed by most scattering theories, and verifiable by FVM, is that the surface’s height distribution is Gaussian. In this paper, we shall also assume that all height distributions are Gaussian. Superficially, it might seem straightforward (if not trivial) to verify these assumptions and to estimate the values of $S_q$ and $S_l$ 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 $S_q$ and $S_l$. 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 $S_q$ and $S_l$. 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 resolution (see equation (8) below, in section 5). On the other hand, the samples of interest to us are all considerably rougher than this, with values of $S_q \gg N_M, 2z_{FLT}$. For such samples, noise and residual flatness are negligible contributors to the uncertainty and bias in estimating $S_q$ and $S_l$, compared to the effects of lateral resolution. For smoother samples, where noise and residual flatness are more significant and one must push to the vertical resolution limits of FVM, other methods have been developed [16] to remove bias, but they are unnecessary for our purposes. For samples whose roughness is much larger than the instrument’s vertical resolution, it is the instrument’s lateral resolution (spatial bandwidth) that dominates the bias and
uncertainty in estimating $S_q$ and $S_{al}$, 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, $S_q$ and $S_{al}$, 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 $\sigma$ and $L$ to denote the same quantities. In this paper, $S_q$ and $S_{al}$ are reserved for the true parameters of the surface, while $\sigma$ and $L$ are used for estimates based on measured data. First-order estimates are denoted $\sigma_1$ 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 $\sigma_2$ 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 $S_q$ and $S_{al}$ values, and with the same field-of-view (FOV) as actual FVM datasets. The same Gaussian filter is applied to these mock surfaces as would be introduced by our FVM instrumentation into real topography data; then the estimators $\sigma_2$ and $L_2$ are computed for the filtered mock surfaces. The range of surfaces for which the first-order estimates are reasonably accurate ($\leq 10\%$) is roughly five octaves wide in $S_{al}$ if there is no instrumental filtering, but vanishes to zero when the biases caused by the FVM’s Gaussian point-spread function are included. By using the corrected estimates $\sigma_2$ and $L_2$, roughly four octaves of accuracy can be recovered. Section 4 also describes a Monte Carlo simulation illustrating the tradeoff between reduced bias and increased uncertainty that applies when using 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—that 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 $S_q$ and $S_{al}$ reliably. In most cases, the accuracy to which this lateral resolution is known determines the final accuracy to which $S_q$ and $S_{al}$ 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) equipped with a telecentric $\times 20–200$ zoom lens objective (model VHZ20W) providing nominal magnifications of $\times 30, \times 50, \times 100$, and $\times 200$. 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 magnification relative to the camera’s CCD are listed in table 1. The zoom lens objective includes ring-light illumination with minimum (maximum) incident angles of $\beta_{\min} = 25^\circ$ ($\beta_{\max} = 29^\circ$), and it has an angular range of detection of $\sigma_D = 18^\circ$, corresponding to a numerical aperture (NA) of 0.31. (These quantities, taken from dimensional measurements of the objective, correspond to the definitions given in section 3.5 of [22].) The numerical aperture corresponds to the NA at the maximum magnification of $\times 200.$ Ring-illumination, common (though not universal) in FVM, is significant because it ensures that the signal is comprised entirely of scattered, rather than

---

4 Use of this product does not constitute an endorsement by NIST. Other manufacturers’ products could perform equally well or better in this application.
FVM data were analyzed using Matlab.

In other words, we are treating the FVM microscope and its images made on known physical test targets. In even in these cases it will often be of interest to measure at least somewhat more transparent to the user; however, the sample. The core of the FVM implementation is from below the best focus for the lowest point of the central plane.

The elements, the digital filter shape and width are not adjustable 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.

The basic principle of FVM is straightforward: a sequence of $N_z$ digital images (each image comprised of $N_x \times N_y$ pixels) is acquired as the sample’s position, relative to the microscope objective, is stepped through the microscope’s depth of focus. The sequence extends from below the best focus for the lowest point of the sample, to above the best focus for the highest point of the sample. The core of the FVM implementation is then a ‘composition’ algorithm that uses the $N_z$ images as input and produces a single $N_x \times N_y$ array of elevations as output. Each elevation is the (interpolated) z-position at which the algorithm finds a measure of local 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 $\mu$m (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 blends into the vertical noise. Thus, the best FVM data for determining the point-spread function at each magnification turns out to be in the ‘marginally resolved’ regime: it is data from the smallest post that still shows a height well above the (vertical) noise. For each magnification, only the FVM data obtained from that one optimal target, indicated in figure 2 and table 1, was used to extract the PSF.

For each of these targets, the measured $z(r)$, normalized so that the baseline (i.e. mean height within a ring centered on the base) lies at $z = 0$ and the maximum measured height lies at $z = 1$, is well fit by a Gaussian, as shown in figure 2. Moreover, the width of

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

Table 1. Measured FVM lateral resolution.
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 \( R_\text{p} \), then \( R_\text{p}/R_0 < 1 \) may be true, but \( R_\text{p}/R_0 = 1 \) 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, thereby allowing the extraction of \( R_0 \) with good accuracy (<3% error) from profile data on posts as large as \( R_\text{p}/R_0 \approx 2 \). This procedure uses an effective radius \( R_\text{eff} \) —defined entirely from measurements \( M(r) \) including the measured value at the origin \( M(r = 0) \)— given by \( (\pi R_\text{eff}^2)M(r = 0) = 2\pi \int_0^\infty M(r)r \, dr \). A universal plot of (theoretical) \( R_\text{eff} \) can be directly calculated from (A1) for a succession of normalized post sizes as shown in figure A2.

Figure A2 also provides some insight into why this slightly elaborate process, (estimating \( R_0 \) from marginally resolved, not fully unresolved, targets) is needed in the first place. The figure shows that, in order for \( R_\text{eff} \) to constitute a reasonably precise estimate of \( R_0 \) (say within 5%), the target radius needs be quite small, approximately \( R_\text{p} < 0.4R_0 \). 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 \( R_\text{eff} \) requires accurate location of the baseline, and a combination of limited vertical resolution, limited accuracy of tilt correction, and limited physical flatness of the machined base makes this problematic. Either a wider post diameter is required, or a smaller \( R_0 \) (obtained by using a higher magnification). We have found the best results are obtained for normalized post diameters in the range of \( 0.75 < R_\text{p}/R_0 < 1.5 \). In this regime, the required correction for finite post diameter is quite substantial, namely 10%–60%.

The values of \( R_0 \) obtained by the iterative procedure described in appendix A are shown in table 1. For each magnification, the extracted value of \( R_0 \) is listed,
along with the measured pixel size $p$. Also listed is the inferred FOV of a $1197 \times 1597$ 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 $\sigma_q$ and $\sigma_a$ is now apparent: the very limited lateral dynamic range. Although the measured PSF’s are accurately Gaussian at all magnifications, the values of $1/e$ radius $R_0$ amount to nearly 32 pixels, independent of magnification. At least for this particular instrument, the price in lateral resolution paid for the FVM elevation information is very high! The effective number of statistically independent measurements is only $N \sim 600$, corresponding to $22 \times 28$ pixels in each dimension, even though the initial images are nearly 2 Mpixel in size. By comparison, the discussion in [17] assumed for numerical estimates a 1D dataset with 1024 independent measurements.

5 The microscope’s internal (lateral) measurement scale was calibrated at each magnification against a NIST-traceable gage block of appropriate size. The microscope’s internal scale never varied by more than 3% from the gage block measurement.

### 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 $R_0 \gg L$, the limited lateral resolution (microscope PSF) discussed above will (a) decrease the observed $\sigma^2$ 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 known width, general signal-processing arguments can be used to model and correct for these two spatial bandwidth effects. This is done below in a closely analogous fashion to the discussions in [17–19]. The other two important effects relate to the low frequency portion of the surface’s power spectrum. When $\text{FOV} \ll L$, the limited FOV will (c) decrease the observed $\sigma^2$ and (d) decrease the observed $L$, simply because the measurement domain only includes a fraction of the area of a typical fluctuation. Analysis of the low frequency fluctuations is further complicated because a linear tilt and in many cases a quadratic curvature are removed from FVM data prior to analysis. (Tilt correction capability is built into many

![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 $2\pi$) are indicated as blue points; the red trace is the best-fit Gaussian.](image)
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 \( \text{Sq} \) and autocorrelation length \( \text{Sal} \). These are the true values of the surface’s topography parameters. The surface’s ACV and PSD are thus given by

\[
ACV(\tau) = \text{Sq}^2 e^{-\tau/\text{Sal}},
\]

\[
PSD(f) = \frac{2\pi (\text{Sal})^2 (\text{Sq})^2}{1 + (2\pi f \text{Sal})^2}.
\]

It is understood that both frequency and lag are defined over a 2D plane, but since we are specializing to the case of isotropic surfaces it suffices to calculate in terms of \( f = \sqrt{f_x^2 + f_y^2} \) and \( \tau = \sqrt{\tau_x^2 + \tau_y^2} \). In this case, Fourier integrals over the 2D plane reduce to 1D integrals with a Bessel function in the integrand, so that for example:

\[
PSD(f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\text{Sq})^2 e^{2\pi i (f_x \tau_x + f_y \tau_y)} e^{-\tau/\text{Sal}} d\tau_x d\tau_y
\]

\[
= \int_{0}^{\infty} (\text{Sq})^2 J_0(2\pi f \tau) e^{-\tau/\text{Sal}} \tau d\tau.
\]

The filtering introduced by the instrument’s Gaussian PSF is most easily treated in the spatial frequency plane, where it is represented by a simple multiplication:

\[
\text{FPSD}(f) = |e^{-2\pi R_{\text{f}} f|^2} PSD(f)
= \frac{2\pi (\text{Sal})^2 (\text{Sq})^2}{1 + (2\pi f \text{Sal})^2} e^{-2\pi R_{\text{f}} f|^2}.
\]

Here, \( \text{FPSD}(f) \) 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

\[
\text{FACV}(R_0, \tau) = (\text{Sq})^2 \int_{-\infty}^{\infty} e^{-Bx^2/2} J_0(Bx) dx,
\]

where we have defined a dimensionless instrumental resolution \( A = \frac{R_0}{\text{Sal}} \) and a dimensionless lag \( B = \frac{\tau}{\text{Sal}} \). 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 \( \text{FACV}(A, B) \), 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:

\[
\text{FACV}(A, B = 0) = \sigma^2
= (\text{Sq})^2 \left[ 1 - \frac{1}{e} \sqrt{\frac{\pi}{2}} A e^{-A^2} \text{Erfc} \left( \frac{A}{\sqrt{2}} \right) \right].
\]

where \( \text{Erfc}(x) \) is the complementary error function. Moreover, equation (3) can be solved numerically to give the apparent autocorrelation length, namely the value of \( B \) for which

\[
\frac{\text{FACV}(A, B)}{\text{FACV}(A, B = 0)} = e^{-1}.
\]

This gives the correction factor \( \alpha \equiv \frac{L_1}{\text{Sal}} \).

Filtered ACV’s, normalized to \((\text{Sq})^2\), the intrinsic rms, are shown at left in figure 3 for various amount of Gaussian filtering. At right, they are shown when
normalized to the apparent rms $\sigma_1^2$. These plots correspond directly to the schematic illustration of figure 3 of [17]. However, these filtered ACV’s are not subject to the requirement $\int_{-\infty}^{\infty} FACV(\tau)2\pi d\tau = 0$, and therefore do not necessarily go negative at large $\tau$, 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 ($R_0 \ll \text{Sal}$), the apparent autocorrelation length approaches the true value $L_1/\text{Sal} \rightarrow 1$, and when the Gaussian filtering is dominant and the surface’s intrinsic fluctuations are highly filtered, ($R_0 \gg \text{Sal}$), 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 $a = L_1/\sqrt{2}R_0 - 1$, where $a \ll 1$ is the heavily filtered limit and $a \gg 1$ is the well-resolved limit. The critical point is that $a$ is defined in terms of the observed quantities only: $L_1$ and the known resolution $R_0$.

The curves in figure 4 are universal for a Gaussian PSF (of any width) and exponential autocorrelation function, and give the desired correction factors $\sigma_1^2$ and $L_1/\text{Sal}$. However, they do not directly provide an algorithm for obtaining $\text{Sq}$ and $\text{Sal}$ from measurement because they are presented as a function of $R_0/\text{Sal}$, and while $R_0$ is known a priori, $\text{Sal}$ is the quantity sought, not the quantity obtained directly from measurement. (That is $L_1$.) The dependences illustrated in figure 4 need to be inverted, and displayed as a function of $R_0/L_1$.

This has been done using the numerical solutions of equation (5); the result for $L_1/\text{Sal}$ 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 $L_1 \rightarrow \sqrt{2}R_0$ 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

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

Over the entire interesting range of $a$, $0.05 \leq a \leq 20$, 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 $\text{Sal}$ 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 $\text{Sq}$, which we denote $\sigma_2$, from the experimentally measured $\sigma_1$.

Of course, there is a limit on how far this procedure can be pushed to correct for the biases created by limited resolution. That limit is determined by how accurately $R_0$ is known, since uncertainty in $R_0$ translates directly into uncertainty in $a$, thru $da = (1 - a)\frac{dR_0}{R_0}$. In the unresolved range, $a < 1$, this means (referring to figure 5) that uncertainty in the extracted value of $\text{Sal}$ diverges leftward of a certain point, when $a$ is less than the fractional uncertainty in $R_0$. (This is the reason for the emphasis given in section 2 and appendix A on experimental determination of $R_0$.) Nonetheless, a reasonable accuracy in $R_0$, say 5%, permits very substantial errors in
autocorrelation length (up to $L_i/\text{Sal} \approx 600\%$ based on figure 5) to be reliably corrected.

### 3.2. Measurement protocol

Based on all the above, an explicit protocol for estimation of $\text{Sal}$ and $\text{Sq}$ 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 estimates of rms and autocorrelation length, namely $\sigma_1$ and $L_1$. (The explicit formulæ to do this, for discretely sampled data fields are listed in appendix B.)

3. If $L_1$ is not ‘significantly’ greater than $\sqrt{2}R_0$, then the surface fluctuations are too unresolved to proceed. To be more quantitative: assuming $R_0$ is known to 10% accuracy, then calculating an $L_1$ value that exceeds $\sqrt{2}R_0$ by only 10% (or less) implies that only an upper limit on $\text{Sal}$ can be determined, namely for this example $\text{Sal} < R_0/4$ (from figure 5). There is simply not enough resolution to determine $\text{Sal}$, and FVM measurements need to be retaken at higher magnification. On the other hand, if $L_1$ is not significantly less than $\text{FOV}/2\pi$ (this numeric value is justified in the Monte Carlo simulations below), then the sampled area of the surface is simply too small to make a meaningful estimate of $\text{Sal}$ or $\text{Sq}$. 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 $\sqrt{2}R_0 < L_1 < \text{FOV}/2\pi$, then calculate $a = \frac{L_1}{\sqrt{2}R_0} - 1$, and from that calculate the correction factor $\alpha$ from equation (6). Then $L_2 = L_1/\alpha$ forms an unbiased estimate of $\text{Sal}$. Use this to form an unbiased estimate of $A = \text{Sal}/\text{Sal}$. Finally, using this estimate of $A$ and the extracted $\sigma_1$ from step (2), use equation (4) to calculate $\sigma_2$, which is an unbiased estimate of $\text{Sq}$.

### 4. Monte Carlo simulations

#### 4.1. Correction of resolution-induced biases

In order to validate the above protocol for correcting the biases introduced by finite resolution, and to
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 computer. Random, exponentially correlated surfaces with a given $S_{a_l}$ and $S_q$, and with a Gaussian elevation distribution, were generated by calculating the corresponding $FSD(f)$, namely equation (1b), over a 512 × 512 pixel field in frequency space, assigning random phases to each point in frequency space uniformly distributed over $[-\pi, \pi]$, and finally performing an inverse Fourier transform. This mock topography was then filtered using a Gaussian kernel to simulate the FVM instrumental resolution. Figures 6(a) and (b) display examples that have been scaled laterally to $8 \times 8 \text{ mm}^2$ FOV, with fixed rms $S_q = 10 \mu m$ for the Gaussian elevation distribution, and autocorrelation lengths of $S_{a_l} = 0.062$ mm and $S_{a_l} = 1.0$ mm respectively. Figures 6(c) and (d) show the same mock topographies after the Gaussian filter is applied. Following the protocol outlined in section 3, first-order estimates $s_1$ 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 $s_2$ and $L_2$ were calculated. This entire procedure was followed for mock topographies with $L_0$ varying from $1/64 \text{ mm}$ (1 pixel) to $8 \text{ mm}$ (512 pixels) in factors of two.

When there is no Gaussian filtering, $s_1$ and $L_1$ themselves provide reasonably accurate estimates of $S_q$ and $S_{a_l}$. This remains true over a substantial range in $S_{a_l}$. This is manifested in figure 8 as a plateau in the unfiltered (red) points at $s_1/S_q = 1$ and $L_1/S_{a_l} = 1$ that is some 4.5 octaves wide. To be somewhat more quantitative, if we ask over what range of surfaces (i.e. maximum and minimum values of $S_{a_l}$) the first-order estimates $s_1$ and $L_1$ lie within 10% of their true values, we find the values given in the first two columns of table 2. For comparison, the corresponding guideline given in [10] is that the ratio of maximum to minimum spatial frequency for accurate determination of $S_{a_l}$ and $S_q$.

![Figure 6. Monte Carlo simulated topographies of exponentially correlated, 8 × 8 mm² surfaces, with autocorrelation lengths of $S_{a_l} = 0.0625$ mm (left) and 1.0 mm (right). The vertical (color) scale is held fixed for all four panels, and extends over ±3$S_q$. 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.](image-url)
Sq is given by \( f_{\text{max}}/f_{\text{min}} = \text{FOV}/8p \), 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’ bandpass 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 S\( q \) 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 1197 \( \times \) 1597 pixels, with resolution \( R_0 \approx 32 \) pixel (see table 1), the filtered mock topographies were 512 \( \times \) 512 pixels, with resolution \( R_0 = 16 \) pixel. The effect of this filtering on the topographies is shown in figures 6(c) and (d), while its effect on the ACV’s is shown in figure 7. In terms of qualitative appearance, the filtered mock topographies greatly resemble actual FVM images of rough surfaces (see section 5 below), as they should.

When the first-order estimates \( \sigma_l \) and \( L_1 \) are calculated from the filtered topography using the same formulae (equations (3a)-(3d) in appendix B), the undesirable effects of finite lateral resolution become obvious. Figure 8 displays (as the green points) these first-order estimates relative to their true values, for Monte Carlo-generated topographies like those in figure 6, when filtered with Gaussian kernels as described above. For the fine-featured surfaces (\( \text{Sal} \leq R_0 \)), major errors are introduced into the first-order estimates by the filtering. The first-order estimate \( \sigma_l \) 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 \( \sigma_l \) and \( L_1 \) are reasonably accurate. The \( \sigma_l \) 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 auto-correlation 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. \( \sigma_2/\sigma_1 \) and \( \sigma_2/L_4 \)) 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 \( \sigma_2 \) and \( L_2 \) accurately reproduce the actual values.

### Table 2. Spatial bandwidth limits for 10% accuracy.

|                | Unfiltered | Filtered | Filtered and corrected |
|----------------|------------|----------|------------------------|
|                | \( \sigma_l \) | \( L_1 \) | \( \sigma_l \) | \( L_1 \) | \( \sigma_2 \) | \( L_2 \) |
| 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 | \( \geq 4.2 \) | \( \geq 3.3 \) |
Figure 7. First-order autocovariance functions $ACV^{(1)}$, normalized to $\sigma^2_1$ 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 8. (a) (left). Deviation between extracted values of rms roughness from a $512 \times 512$ pixel mock topography and the actual value $Sq$. Red points indicate first-order estimates (i.e. $\sigma/\sqrt{Sq}$) from the unfiltered topography, that has a resolution $\approx 1$ pixel. They show a ~5 octave wide plateau of accuracy, where $\sigma/\sqrt{Sq} = 1$. Green points indicate first-order estimates $\sigma/\sqrt{Sq}$ from the mock topographies that have been filtered to simulate the instrument’s actual lateral resolution. The plateau has disappeared, and all estimates are biased low by a large factor. Blue points indicate the corrected estimates, i.e. $\sigma/\sqrt{Sq}$, after the correction protocol of section 3 is applied. Most of the plateau of accuracy is recovered. (b) (right) The corresponding deviation between extracted values of autocorrelation length and the actual value $Sal$. Again, the first-order estimates on filtered mock topographies show no plateau of accuracy whatsoever, although for small $Sal$ the estimates are biased high and for large $Sal$ they are biased low. Again, the correction protocol of section 3 recovers most of the plateau of accuracy.
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 Monte Carlo simulation was run, in which 128 different mock surfaces were generated, each with the same $S_q$ and $S_{al}$. Each surface was then filtered with the same Gaussian kernel described above, and the filtered topography used as input to the algorithm. Histograms of the extracted rms and autocorrelation length, before and after applying the correction factor, are shown in figures 9 for 2 cases: one in which the correction is relatively small, and one in which it is quite large. Comparing $\sigma_1$ and $\sigma_2$, it is clear that correction of the bias (whether it be a 20% or a 250% bias) comes at the expense of increased dispersion. However, the dispersion in the corrected values is still small enough to constitute a useful measurement of surface 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 $S_{al}$), the corrected estimates $\sigma_2$ 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 because of low frequency effects (and because of the approximations embedded in the algorithm).

4.3. Low-frequency (finite FOV) effects

Figure 8 shows that for the Monte Carlo-generated surfaces, the biases introduced into estimates of $\sigma$ and $L$ by the finite FOV, i.e. for surfaces whose $L_0$ approaches the FOV, are the same whether the estimates are made on filtered or unfiltered surfaces, or whether or not the estimates have been corrected by the algorithm. Fairly good (but purely empirical) estimates of these biases are indicated by the dashed curves in figures 8(a) and (b), calculated from:
For comparison, Church [17] states ‘when \( L_1 \leq \text{FOV}/8\pi \) then \( L_1 \) may be taken as approximately equal to \( \text{Sal} \) and \( \sigma_1 \) is essentially equal to \( \text{Sq} \).’ The estimates in (7) are somewhat less stringent than this. They imply that biases in (first-order estimates of) \( \sigma \) and \( L \) remain below 5% as long as \( L_1 \leq \text{FOV}/6.1\pi \) and that they remain below 10% for \( L_1 \leq \text{FOV}/4.1\pi \). 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 \( \text{FOV}/4\pi \). 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 \( \sigma \) 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 \( L_1 < \text{FOV}/2\pi \) requirement of the protocol (section 3, step (3)) is observed, then this effect can be neglected.

5. Application: FVM image analysis of physical rough surfaces

The algorithm described above is mathematically sound; its validation against numerically generated random topography provides an intuitive and graphical demonstration of that. However, to show its practical usefulness requires validation against the topography of real physical surfaces measured by FVM microscopy. Of course, for a real surface there is no guarantee that its topography satisfies the assumptions made in the above analysis. For example, a rough surface may be anisotropic (different autocorrelation lengths in the \( x \) and \( y \) directions), or its ACV may not 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 \( \text{Sq} \) and \( \text{Sal} \) are not meaningful (and why they are not). Non-ethereal, a significant effort was made to apply the above algorithm to real surface topographies, and thus to extract a single \( \sigma_1 \) and \( L_1 \) for each.

Eleven samples of common outdoor building materials were obtained, and each was imaged by FVM at four different magnifications: \( \times 30, \times 50, \times 100, \) and \( \times 200 \). The overall FOV was held approximately constant for all magnifications, at \( 11.1 \times 8.2 \text{ mm} \), by stitching together smaller subfields for the higher magnifications. (Thus, there were \( 2 \times 2 \) subfields at \( \times 50, 4 \times 4 \) at \( \times 100, \) and \( 8 \times 8 \) at \( \times 200 \).) 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 \( \times 5 \) 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 \( \times 32 \) 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 \( \sigma_1 \) 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 \( (r_x, r_y) \). 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 \( r = (r_x^2 + r_y^2)^{1/2} \). The point where this function fell to \( 1/e \) of its 0-lag value was identified as \( L_1 \). Finally, the algorithm described above and the separately obtained values of microscope resolution (section 2 and table 2) were used to obtain corrected estimates \( \sigma_2 \) and \( L_2 \).

An example of this procedure is illustrated in figure 10 for one particular sample of weathered concrete. A good measure of the algorithm’s effectiveness is consistency in the value of autocorrelation length extracted from data obtained at different magnifications. As shown for this sample by the table in figure 10, the variation between the \( \times 30, \times 50, \) and \( \times 100 \) magnifications in terms of \( \Delta = \frac{\max(L)}{\min(L)} - \frac{\min(L)}{\min(L)} \) is 28% for \( L_1 \), i.e. before correction, and 9% for \( L_2 \), i.e. after the correction algorithm. In other words, the algorithm improves the consistency in estimated
autocorrelation length by about a factor of three. This inconsistency reflects the dominant uncertainty in estimated autocorrelation length, namely the limited lateral resolution of the FVM data (which is the only significant thing differing between magnifications). Similar improvements in uncertainty in

\[ \frac{\text{FOV}}{2\pi} \]

Figure 10. Analysis of the FVM topography measurements on a particular sample of weathered concrete. The transition from densely to more sparsely sampled points in the ACV corresponds to \( \frac{\text{FOV}}{2\pi} \), where edge effects make autocorrelation estimates unreliable.
autocorrelation length estimates are obtained for other samples.

Finally, the data for all eleven outdoor building material samples are shown in figure 11 and table 3. For nine of the samples, all of the datasets that met the criteria of the measurement protocol (section 3.2) yielded values for \( \sigma_2 \) and \( L_2 \) that were highly consistent across magnifications. (Several of the \( \times 30 \) datasets did not meet the FOV criterion of the protocol and are therefore not included on the plot.) The two exceptions are the samples ‘wood’ and ‘concrete2’. The former is enormously anisotropic. Its covariance isotropy plot displays half-widths that are 2.7 mm and 0.21 mm in the horizontal and vertical directions respectively (at \( \times 100 \)), reflecting the intuitive fact that wood’s grain introduces anisotropy into the finished surface. While the modest amounts of anisotropy (up to 2:1 in \( L_{h}^{(1)}:L_{v}^{(1)} \)) found in some of the other samples does not seem to affect the protocol’s effectiveness, this extreme level of anisotropy clearly requires a more sophisticated treatment. For the second exception, ‘concrete2’, the protocol yields values of \( \sigma_2 \) and \( L_2 \) that are obviously anomalous at the lowest magnification. This dataset does (barely) meet the FOV and resolution criteria of the protocol. The anomalous values appear to be due to a different effect, namely the sample’s rms roughness being comparable to the vertical resolution of the FVM. Although the precise and robust

![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.](image)

| Sample          | Rms roughness \( \sigma_2 \) (\( \mu m \)) | Autocorrelation length \( L_2 \) (mm) |
|-----------------|-----------------------------------------|-----------------------------------|
|                 | \( \times 30 \) | \( \times 50 \) | \( \times 100 \) | \( \times 200 \) | \( \times 30 \) | \( \times 50 \) | \( \times 100 \) | \( \times 200 \) |
| 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\(^a\) | 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.
determination of vertical resolution in an FVM is beyond the scope of this paper, (and it is not specified by the manufacturer for this particular commercial instrument), we have found that a quadratic scaling with magnification
\[ \delta z \approx \frac{80 \text{ mm}}{\text{(magnification)}^2} \]

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 \( \sigma \) and \( L \) measurements.

The dashed line in figure 11 indicates the condition \( \frac{\sigma \lambda}{L^2} = \frac{1}{2\pi} \) for \( \lambda = 1.3 \text{ mm} \) (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 Kirchhoff approximation for rough surface scattering is a low-curvature approximation [13, 23], corresponding to the region in \( \sigma-L \) 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 Kirchhoff 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 for example, we found that the 1197 \( \times \) 1597 pixel camera format really only provides \( \sim 24 \times 28 \) statistically independent elevation measurements (table 1). Though other FVM implementations will certainly have different numerical characteristics, a large loss of dynamic range is inherent in the FVM technique. Any composition algorithm will degrade the lateral resolution. The reduced dynamic range introduces large biases into estimates of rms roughness and autocorrelation length for random rough surfaces. However, we have derived correction factors for \( \sigma \) and \( L \) that greatly reduce these biases. The correction factors, incorporated into the overall measurement protocol in section 3.2, are exact for the assumptions 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 \( \sigma \) 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 \( \sigma \) 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
\[ G(r) = 1, \quad r \leq R_p, \]
\[ = 0, \quad r > R_p. \]

Then the measured FVM profile is given by the (2D) convolution of a Gaussian and top hat:
\[ M(r) = S(r) \ast G(r) = \frac{2}{R_0^2} \int_0^{R_p} e^{-r'^2/R_0^2} I_0 \left( \frac{2rr'}{R_0^2} \right) r' \, dr', \]

where \( I_0 \) is the modified Bessel function of order 0. This convolution is illustrated graphically in figure A1, showing the expected transition from Gaussian to ‘flat-topped’ profiles as \( R_p/R_0 \) varies from less than unity (post is unresolved) to greater than unity (post is...
resolved). Note that for \( R_p/R_0 < 1 \), the measured height at the origin is less than (possibly much less than) the actual post height. Unless \( R_p/R_0 \ll 1 \), the apparent \( 1/e \) radius \( R_{\text{eff}} \) is larger than the actual \( R_0 \); a downward correction is required in order to obtain an unbiased estimate of \( R_0 \) from measurements.

The exact formula (A1) is too complex to be used to calculate the correction directly from measured data. However, as shown in figure A2, a relatively accurate approximation is given by the very simple formula \( R_{\text{eff}}^3 = R_0^3 + R_p^3 \). This approximation has a maximum fractional error of 3% within the range \( R_p/R_0 < 2 \). If desired, this estimate for \( R_0 = (R_{\text{eff}}^3 - R_p^3)^{1/3} \) can be iteratively refined by using it to calculate a normalized post diameter \( R_p/R_0 \), reading the exactly calculated value of \( R_{\text{eff}}/R_0 \) off of the universal plot in figure A2, and combining that with the measured value of \( R_{\text{eff}} \).

Appendix B. Notation and explicit formulae for first-order estimates

To give explicit formulae, let the FVM data be given by a matrix \( z_{ij} (x_i, y_j) \), where the rows comprise elevation data at a series of horizontal sample positions \( x_j (j = 1 \ldots N_x) \) and the columns comprise elevation data at a series of vertical sample positions \( y_i (i = 1 \ldots N_y) \). We seek to estimate the values of \( \sigma_0 \) and \( L_0 \) even though the \( z_{ij} \) have been corrupted by instrumental bandwidth effects. First-order estimates
\[ \bar{z} = \frac{1}{N_x N_y} \sum_{i,j} z_{ij}, \]
\[ \sigma_1^2 = \frac{1}{N_x N_y} \sum_{i,j} (z_{ij} - \bar{z})^2, \]
\[ P_{mn} = \frac{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} (z_{ij} - \bar{z})(z_{i-m,j-n} - \bar{z})}{(N_y - |m|)(N_x - |n|)} \quad \text{where} \]
\[ -(N_x - 1) \leq m \leq (N_x - 1), \]
\[ -(N_y - 1) \leq n \leq (N_y - 1). \]

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. \( \text{xcorr}(A, A) \)). The first-order estimate of the ACV is simply the symmetrized version of \( P_{ij} \):

\[ ACV_{ij}^{(1)} = \frac{1}{4} (P_{m,n} + P_{m,-n} + P_{-m,n} + P_{-m,-n}) \quad \text{for} \]
\[ i = m = 0 \ldots N_x, j = 0 \ldots N_y. \]

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

### 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 I 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, Haitijama 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 Chavignier 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* (Norwood, MA: Artech House)
[14] Haitjema H 2015 Uncertainty in measurement of surface topography Surf. Topogr. Metrol. Prop. 4 035008
[15] ISO 25178-601 2012 *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 6–53
[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)