

Basic elements of metrology and introduction to the various techniques

André Chrysochoos, Yves Surrel

▶ To cite this version:

André Chrysochoos, Yves Surrel. Basic elements of metrology and introduction to the various techniques. Michel Grediac (Editor), Francois Hild (Editor). Full-Field Measurements and Identification in Solid Mechanics, Wiley-ISTE, 2012, 978-1-118-57847-6. hal-02057254

HAL Id: hal-02057254 https://hal.science/hal-02057254v1

Submitted on 5 Mar 2019 $\,$

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Chapter 1

Basics of Metrology and Introduction to Techniques

1.1. Introduction

Full-field optical methods for kinematic field measurement have developed tremendously in the last two decades, due to the evolution of image acquisition and processing. Infrared (IR) thermography has also dramatically improved due to the extraordinary development of IR cameras. Because of their contactless nature, the amount of information they provide, their speed and resolution, these methods have enormous potential both for the research lab in the mechanics of materials and structures and for real applications in industry.

As for any measurement, it is essential to *assess* the obtained result. This is the area of metrology. The ultimate goal is to provide the user with as much information as possible about the measurement *quality*. We deal with a specific difficulty for the quality assessment of the optical methods that arises precisely from their full-field nature. The metrology community is far more familiar with point-wise or average scalar measurements (length, temperature, voltage, etc.). As of today, the metrology of full-field optical methods is not yet fully settled. However, the wide dissemination of these techniques will only efficiently occur when users have a clear understanding of how they can characterize the measurement performances of the equipment that vendors put on the market.

First, the goal of this chapter is to present some basic elements and concepts of metrology. It is by no means exhaustive, and only aims at presenting the basics of the

Chapter written by Andrž CHRYSOCHOOS and Yves SURREL.

domain in a simple way, so that the researchers, users, developers and vendors can exchange information based on well-established concepts. Second, we will rapidly present the different optical techniques based on their main characteristics (how information is encoded, interferential or not, etc.).

It should be noted that optical measurement techniques exhibit a non-negligible amount of complexity. Figure 1.1 outlines the typical structure of a measurement chain that leads from a physical field to a numerical measurement field using a camera. It can be seen that there are many steps required to obtain the final result which the user is interested in, and numerous parameters that may impair the result and effects are involved at each step. Most importantly, there are usually numerical postprocessing stages that are often "black boxes" whose metrological characteristics or impact may be difficult or even impossible to obtain from the supplier of the equipment in the case of commercial systems.



Figure 1.1. Outline of the many steps involved in a measurement chain using a camera

1.2. Terminology: international vocabulary of metrology

1.2.1. Absolute or differential measurement

In any scientific domain, terminology is essential. Rather than enumerating the main terms to use (precision, sensitivity, resolution, etc.), let us try to adopt the final

user point of view. What are the questions he generally asks, and in which context? There are, in fact, not so many questions, and each of them leads naturally to the relevant metrological term(s):

1) Is the obtained result "true", "exact" and "close to reality"? How to be confident in the result?

2) Is the equipment "sensitive"? Does it see small things?

These two questions lie behind the separation of metrology into two distinct domains, within which the metrological approach will be different: *absolute* measurement and *differential* measurement.

1.2.1.1. Absolute measurement

Here, we seek the "true" value of the measurand (the physical quantity to measure), e.g. to assess that the functional specifications of a product or system are met. Dimensional metrology is an obvious example. The functional quality of a mechanical part will most often depend on the strict respect of dimensional specifications (e.g. diameters in a cylinder/piston system). The user is interested in the deviation between the obtained measurement result and the true value (the first of the two questions above). This deviation is called the measurement *error*. This error is impossible to know, and here is where metrology comes in. The approach used by metrologists is a *statistical* one. It will consist of evaluating the statistical distribution of the possible errors, and characterizing this distribution by its width, which will represent the average (typical) deviation between the measurement *uncertainty*, which gives the user information about the amount of error that is likely to occur.

It is worth mentioning that some decades ago, the approach was to evaluate a maximum possible deviation between the measurement and the real value. The obtained uncertainty values were irrelevantly overestimated because the uncertainty was obtained by considering that all possible errors were synergetically additive in an unfavorable way. Today, we take into account that statistically independent errors tend to average each other out to a certain extent. In other words, it is unlikely (and this is numerically evaluated) that they act coherently together to impair the result in the same direction. This explains the discrepancy between the uncertainty evaluation formulas that can be found in older treatises and those used today. As an example, if c = a + b, where a and b are independent measurements having uncertainties Δa and Δb , the older approach (maximum upper bound) would yield $\Delta c = \Delta a + \Delta b$, but the more recent approach gives $\Delta c = \sqrt{\Delta a^2 + \Delta b^2}$, these two values being in a non-negligible ratio of $\sqrt{2} = 1.414$ in the case $\Delta a = \Delta b$.

1.2.1.2. Differential measurement

Now, let us consider that the user is more interested in the deviation from a "reference" measurement. In the mechanics of materials field, the reference measurement will most likely be the measurement in some initial state, typically before loading. It is basically interesting to perform measurements in a differential way, and this is done for two reasons:

- As the measurement conditions are often very similar, most *systematic errors* will compensate each other when subtracting measurement results to obtain the deviation; typically, optical distortion (image deformation) caused by geometrical aberrations of the camera lens will be eliminated in a differential measurement.

– Measurement uncertainty is only about the *difference*, which implies that interesting results can be obtained even with a system of poor quality; let us take an example of a displacement measurement system exhibiting a systematic 10% error. Absolute measurements obtained with this system would probably be considered unacceptable. However, when performing differential measurements, this 10% error only affects the difference between the reference and actual measurements. Consider the following numerical illustration: with a first displacement of 100 μ m and a second displacement of 110 μ m; this hypothetical system would provide measured values of 90 and 99 μ m, roughly 10 μ m error each time, probably not acceptable in an absolute positioning application, for example. But the measurement differential displacement is 9 μ m, which is only a 1 μ m error.

Regarding this second example, we should emphasize the fact that it is *not possible* to use relative uncertainty (percentage of error) with differential measurements because this relative uncertainty, ratio between the uncertainty and the measurand, is related to a very small value, the difference of measurements, that is nominally zero. Hence, "percentage of error" is strictly meaningless in the case of differential measurements. Only absolute uncertainty is meaningful in this case.

1.2.2. Main concepts

We will consider in this section the main concepts of metrology. The corresponding terms are *standardized* in a document (*international vocabulary of metrology*, abbreviated as "VIM", [VIM 12]) which can be downloaded freely from the BIPM website.

1.2.2.1. Measurement uncertainty

The VIM definition is as follows:

non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used.

The underlying idea is that the measurement result is a random variable, which is characterized by a probability density function centered at a certain statistical average value. This probability density function is typically bell-shaped. Note # 2 in the VIM states:

The parameter may be, for example, a standard deviation called standard measurement uncertainty (or a specified multiple of it), or the half-width of an interval, having a stated coverage probability.

So the idea is really to *characterize the width* of a statistical distribution. There is obviously a conventional feature to perform this characterization. We may choose the width at half the maximum, the width at 1/*e*, the usual standard deviation, etc.

In industry, the need to assess the quality of a measurement instrument can be slightly different: it is most often necessary to be "sure" that the measurement deviation will not exceed some prescribed value. So, the chosen width will include as much as possible the total spreading width of the function. Of course, no real 100% certainty can be obtained, but we can obtain *confidence intervals* at x%, which are intervals having a probability of x% to include the real measurand value. For example, for a Gaussian distribution of the measurement result probability, an interval of $\pm 3\sigma$ has a probability of 99.5% to englobe the true value. Hence, the concept of *expanded measurement uncertainty*, whose VIM definition is:

product of a combined standard measurement uncertainty and a factor larger than the number one.

The "combined standard measurement uncertainty" is the final uncertainty obtained when all sources of uncertainty have been taken into account and merged. In general, this is what is referred to when we speak of "uncertainty" in short.

Measurement uncertainty is the key notion to use when dealing with absolute measurement.

It should be noted that the two major uncertainty sources are those related to noise, introducing random fluctuations of null average value, and those related to *systematic errors* (calibration errors, poor or insufficient modeling, external permanent biasing influences, etc.). With full-field optical methods, noise can easily be characterized (see section 1.2.2.2).

1.2.2.2. Resolution

The VIM definition is as follows:

smallest change in a quantity being measured that causes a perceptible change in the corresponding indication.

This is the key notion to use when performing a differential measurement. In practice, with full-field optical methods, noise (optical, electronic, etc.) will be the factor limiting the resolution. A practical definition of resolution that we can propose is (non-VIM):

change in the quantity being measured that causes a change in the corresponding indication greater than one standard deviation of the measurement noise.

As it happens, with techniques that use cameras, it is easy to *measure* the measurement noise because a large number of pixels are available to obtain significant statistics. To evaluate the measurement noise, it suffices to make two consecutive measurements of the same field, provided, of course, that it does not vary significantly in the meantime. Let us denote by $m_1(x, y)$ and $m_2(x, y)$ two successive measurements at point (x, y). These values are the sum of the "signal" s (the measurand value) and the noise b, as:

$$\begin{cases} m_1(x,y) = s(x,y) + b_1(x,y) \\ m_2(x,y) = s(x,y) + b_2(x,y) \end{cases}$$
[1.1]

where the signal s is unchanged, but the noise has two different values. The subtraction of the two results yields:

$$\Delta m(x,y) = m_2(x,y) - m_1(x,y) = b_2(x,y) - b_1(x,y)$$
[1.2]

Assuming *statistical independence* of the noise between the two measurements, which is often the case¹, we can evaluate the standard deviation of the noise b (we *always* assume that the noise is stationary, i.e. its statistical properties do not change in time) by using the theorem which states that the variance (square of the standard deviation) of the sum of two statistically independent random variables is equal to the sum of their variances, which gives:

$$\sigma^2(\Delta m) = \sigma^2(b_2 - b_1) = \sigma^2(b_2) + \sigma^2(-b_1) = 2\sigma^2(b)$$
[1.3]

An estimator of the statistical variance of the noise difference $b_2 - b_1$ is obtained by calculating, for all the pixels of the field, the arithmetic variance:

$$\tilde{\sigma}^2(\Delta m) = \frac{1}{N_{\text{pixels}}} \sum_{\text{pixels}} [\Delta m(x, y) - \overline{\Delta m}]^2$$
[1.4]

where $\overline{\Delta m}$ is the arithmetic average of Δm within the field:

$$\overline{\Delta m} = \frac{1}{N_{\text{pixels}}} \sum_{\text{pixels}} \Delta m(x, y)$$
[1.5]

¹ Counter example: atmospheric turbulence has typical correlation times of the order of one or many seconds, possibly larger than the interval between successive measurements if we do not take great care.

We thus obtain for the resolution r:

$$r = \sigma_b \approx \frac{1}{\sqrt{2}} \tilde{\sigma}(\Delta m) \tag{1.6}$$

For the sake of mathematical rigor, we have chosen to denote by an ~ sign the real statistical values, and not their *estimators* obtained from arithmetic averages over a large number of pixels. This number being typically much larger than 10,000 (which corresponds to a sensor definition of 100×100 pixels, a very low definition), we can ignore the difference and consider the estimators to be faithful.

Let us insist on the fact that this estimate of the noise level is very easy to perform, and should be systematically performed during full-field measurements.

Characterizing resolution using the noise level is common in IR thermography. An IR camera is characterized with *noise equivalent temperature difference* (NETD), which is nothing other than the output noise level converted into the input, i.e. converted into the temperature difference that would cause a reading equal to one standard deviation of the noise. Temperature variations will begin to "come out of the noise" when they reach the order of magnitude of the NETD. To change an output noise level into an input level, we use *sensitivity*.

1.2.2.3. Sensitivity

The VIM definition is as follows:

quotient of the change in an indication of a measuring system and the corresponding change in a value of a quantity being measured.

This corresponds to the coefficient s that maps a small variation of the measurand onto the corresponding variation of the measurement reading (the output m)

$$s = \frac{\partial m}{\partial e} \tag{1.7}$$

Sensitivity is not always relevant for optical field measurements. As an example, correlation algorithms directly provide their results as displacement values; their sensitivity is equal to one. On the contrary, the grid method presented later in this AQ1 book provides a phase after the detection step. The relationship between the phase Φ and the displacement u is:

$$\Phi = 2\pi \frac{u}{p} \tag{1.8}$$

where p is the grid pitch because the phase varies by 2π when the displacement is equal to one grid period. Thus, the sensitivity according to the VIM is:

$$s = \frac{2\pi}{p} \tag{1.9}$$

This notion is rarely used in optical field methods. However, in IR thermography, it remains fully meaningful and represents the voltage or current variation at the sensor level relative to the received irradiation.

1.2.2.4. Repeatability

The VIM definition is as follows:

measurement precision under a set of repeatability conditions of measurement

to which we add the definition of measurement precision

closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions

and the definition of repeatability conditions:

condition of measurement, out of a set of conditions that includes the same measurement procedure, same operators, same measuring system, same operating conditions and same location, and replicate measurements on the same or similar objects over a short period of time.

Actually, the underlying idea is fairly simple: repeatability is supposed to characterize the *measurement instrument alone*, having excluded all external influences: same operator, same object under measurement, short delay between successive measurements, etc. In short, repeatability tests will characterize the noise level of the measurement instrument.

When using IR thermography, it is essential to evaluate repeatability only after the radiometric system has reached its thermal equilibrium. The non-stationary feature of the parasitic radiations coming from the optical elements and from the rest of the camera may significantly bias the results. It is noteworthy that the time to thermal equilibrium can be some hours, even in a stabilized environment. This is a noticeable constraint [BIS 03].

The "dual" notion, which supposes, on the contrary, different operators, times, locations and even measurement systems, is *reproducibility*.

1.2.2.5. Reproductibility

The VIM definition is as follows:

measurement precision under reproducibility conditions of measurement

to which we add the definition of the reproducibility conditions:

condition of measurement, out of a set of conditions that includes different locations, operators, measuring systems, and replicate measurements on the same or similar objects.

To emphasize the fact that very much can be changed in this case, Note 1 in the VIM states:

The different measuring systems may use different measurement procedures.

As can be seen, as little as possible should change when performing repeatability tests (in short, noise evaluation), and as much as possible may change when doing reproductibility tests (in short, all possible error sources can be involved).

1.2.2.6. Calibration

The VIM definition is as follows:

operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication.

Thus, calibration enables, from artifacts producing known values of the measurand, the identification of the transfer function of the measurement system. As an example, for an instrument having a linear response, measuring a single standard allows us to identify the slope of the straight line that maps the measurand to the output indication. Note in passing that this slope is nothing other than the sensitivity of the instrument according to the VIM definition presented in section 1.2.2.3.

A measurement standard is a

realization of the definition of a given quantity, with stated quantity value and associated measurement uncertainty, used as a reference.

In IR thermography, the ideal measurement standard is the *black body*, a radiating object with unity emissivity. For cameras, it is practical to use plane black bodies that present isothermal surfaces (e.g. within ± 0.02 K). In this case, temperatures and radiated energies unequivocally correspond. Classic calibration procedures use two black body temperatures that include the temperature range of the scene that is to be measured. The procedure involves a non-uniformity correction (NUC), which sets the offsets of all matrix elements such that the represented thermal scene appears uniform and detects and replaces the bad pixels (bad pixel replacement (BPR)); bad pixels are those which, according to a user-prescribed criterion, provide an output signal that

deviates too far from the average. Recently, this calibration step, which is convenient and fast, has tended to be replaced by a pixel-wise calibration, which is longer and more costly but offers improved precision over a wider measurement range [HON 05]. The interested reader will find in Chapter 16 more information on this subject.

1.2.2.7. Illustration

Let us recall here an elementary result. If we average the results of N nominally identical measurements (in other words, measurement performed in *repeatability conditions*, section 1.2.2.4), we obtain something that can also be considered a random variable, because another set of N measurements would provide a different value for the average. The classic result is that the standard deviation of the fluctuations of the average of N measurements is equal to the standard deviation of the fluctuations of a single measurement divided by \sqrt{N} .



Figure 1.2. Measurement, error and uncertainty

Figure 1.2 illustrates some of the notions presented previously. We have represented on the horizontal axis different values related to one or many measurements of the same quantity, with their uncertainties, namely:

- The true value of the measurand (it is permissible to speak of the "true value", explicitly referred to in the VIM in section 2.11).

- The true value biased by the systematic errors; measurement values will fluctuate around this value.

– The result of a single measurement; many measurements will provide values that will fluctuate according to some probability density function of standard deviation σ_n , where *n* stands for noise.

– A value resulting from the averaging of N measurement results; many such averages will fluctuate with a standard deviation that is \sqrt{N} smaller than for a single measurement.

– The standard deviation σ_m , where *m* stands for "model", which corresponds to the probability density function of systematic errors that could be determined by a proper modeling and investigation of the physics of the instrument.

– The combined uncertainty $\sqrt{\sigma_m^2 + \sigma_b^2}$, which takes into account the uncertainty related to systematic errors, as well as the noise.

- The expanded uncertainty that can be used as a confidence interval.

In practice, it is often the uncertainty component not related to noise (the systematic errors) that is the most difficult to evaluate.

1.3. Spatial aspect

Full-field optical measurement methods have a special aspect for metrology: the fact that they provide a spatial field instead of a single scalar. Concepts adapted to this spatial nature are lacking in the VIM. The most important is *spatial resolution*, which depends on numerous factors, especially all the digital postprocessing that can take place after the raw measurements are obtained, for example, to decrease the amount of noise. Before presenting the possible definitions, let us recall some basic results concerning spatial frequency analysis.

1.3.1. Spatial frequency

The relevant concepts here are commonly used by optical engineers when dealing with *image forming*. Indeed, the process of forming an image from an object using an imaging system can be seen as the transfer of a signal from an input (object irradiance) to an output (image intensity). This transfer will change some features of the signal, according to some *transfer function*. In a way similar to an electronic filtering circuit, which is characterized by a gain curve as a function of (temporal) frequency, an image-forming system can be characterized by the way it will decrease (there is never any amplification) the *spatial frequencies*. In lens specifications found in catalogs, we can sometimes find "gain curves" (called optical transfer function) that correspond exactly to gain curves characterizing electronic amplifiers or filters.

The difficulty of spatial frequency is related to its vectorial aspect because we have to deal with functions defined on a two-dimensional (2D) domain, e.g. the irradiance

L(x, y) of the object to be imaged. When dealing with a time function f(t), the variable t is defined over a unidimensional space, and the Fourier conjugated variable, the frequency, is a scalar variable. With a function defined on a plane, f(x, y) or $f(\vec{r})$ with vector notation, the Fourier variable is of dimension 2: (f_x, f_y) or \vec{f} .

To understand the notion of spatial frequency, the simplest way is to consider an elementary spatial periodic function, for example $z(x, y) = \cos[2\pi(f_x x + f_y y)]$, where f_x and f_y are constants. This "corrugated iron sheet" function is illustrated in Figure 1.3.



Figure 1.3. Function $\cos[2\pi(f_x x + f_y y)]$

AQ2 We can rewrite its definitive equation as $z(x, y) = \cos(2\pi \vec{f} \cdot \vec{r})$, where \vec{r} is the vector of components (x, y) and \vec{f} is the *spatial frequency vector* of components (f_x, f_y) . It is easy to convince ourselves that:

– The direction of vector \vec{f} is perpendicular to the function z(x, y) isovalue lines.

- The norm (length) of \vec{f} is equal to the reciprocal of the function period $p: ||\vec{f}|| = 1/p.$

The history of instrumental optics has been a race to obtain instruments that transmit as much as possible all the spatial frequencies present in an object, in the same way as audio electronics developed to transmit with high fidelity all the acoustic frequencies present in voice or music signals. In the optical domain, test objects are often targets made up of equidistant straight lines that implement "single frequency" objects² (Figure 1.4). The investigation of these target images, describing whether the image-forming system has correctly transmitted ("resolved") the lines, allows us to determine the instrument *cutoff frequency* beyond which the lines will

² We ignore here the fact that sharp-edged lines have in their frequency spectrum many harmonics with frequencies that are integer multiples of the basic frequency.

not be distinguished one from the other. We can also see in Figure 1.4 trumpetshaped patterns that correspond to increasing spatial frequencies, which is the spatial implementation of what is known as "chirps" in signal processing.



Figure 1.4. Mire ISO 12233

We can suppose that Fourier analysis is perfectly well adapted to deal with all these notions, and that the optical transfer function is nothing else than the filtering function in the Fourier plane (the spatial frequency plane) where the spatial frequencies present in the object are represented.

Figure 1.5 illustrates the process of image forming, starting from a sinusoidal grid object. The instrument gain (or *optical transfer function*) is represented in the central box, at the top.

Light intensity is always non-negative, so a constant average value has to be superimposed onto the sinusoidal variation, yielding the irradiance:

$$I(x) = A\left[1 + \gamma \cos\left(\frac{2\pi x}{p}\right)\right]$$
[1.10]

Visibility (or contrast) is the dimensionless ratio of the modulation and the average. Visibility is unity when "blacks are perfectly black", i.e. when the irradiance minimum reaches zero. The change in the irradiance signal by a system having the transfer function G(f) that attenuates nonzero frequencies causes a loss of visibility. For a spatial frequency f_0 , as represented in Figure 1.5, visibility loss is

$$\frac{\gamma'}{\gamma} = \frac{G(f_0)}{G(0)} \tag{1.11}$$



Figure 1.5. Transmission of sinusoidal grid irradiance through an image-forming system

For certain spatial frequencies, visibility can even vanish completely. Also, if the transfer function G(f) oscillates around zero, visibility will change its sign, causing a phase lag of π of the output signal: minima and maxima will exchange their places; this is called *visibility inversion*. This effect is shown in Figure 1.6, where the visibility changes its sign when crossing the zones where it vanishes.



Figure 1.6. Contrast inversions

Historically, of course, this mathematical background was not available. The spatial resolution of an image-forming system was defined from the minimum distance between the images of two neighboring points that allowed the points to be

distinguished ("resolved"). We recognize the spatial equivalent of the resolution of a measuring instrument as the "smallest distance between two points whose images can be separated". We can state this in other words: "the smallest change in position that causes a perceptible change in the corresponding image", which is the copy *mutatis mutandis* of the VIM definition. This concept relates to what is known in optical engineering as the Rayleigh criterion.

Considering the images of grids having various frequencies instead of considering neighboring points is the dual approach (in the Fourier sense). In this case, we seek the highest frequency that is transmitted by the instrument without perceptible degradation (e.g. without loss greater than 50%). The reciprocal of this spatial frequency is a geometrical wavelength that can be considered as the spatial resolution. This explains why the following definition can be found in [AST 08]:

– "Optical data bandwidth: Spatial frequency range of the optical pattern (e.g. fringes, speckle pattern, etc.) that can be recorded in the images without aliasing or loss of information". We can recognize the notion of performance of an image-forming system, characterized by a cutoff frequency f_c . The bandwidth is the interval $[0, f_c]$.

– "Spatial resolution for optical data: One-half of the period of the highest frequency component contained in the frequency band of the encoded data". This is the quantity $1/2f_c$. The factor 2 is something of a convention and has no precise justification. However, there is always a conventional aspect in these definitions as the considered quantities vary in a continuous way. The above-mentioned Rayleigh criterion also has a conventional aspect. A factor of 2.5 or 1.5, for example, might have been used without the possibility of justifying that choice any more clearly.

For image correlation techniques, the minimum spatial resolution is the size of the subimages used to calculate the correlation, because we obtain only one measurement per subimage position, for the whole subimage area. Of course, the subimage can be (and is, in general) moved pixel by pixel to obtain a displacement field populated with as many pixels as the original image, but this does not change the spatial resolution, as all measurements corresponding to a distance smaller than the subimage size will be correlated and not independent.

To end this section, let us emphasize that we should not use the term "resolution" to designate the number of pixels of the camera sensor, but *definition*. Remember that we speak of high-definition TV (HDTV) to evoke a very large number of pixels in the image. High resolution is obtained with a high magnification of the imaging system. Thus, we can have a very good resolution with a low definition, with a sensor that does not have a very high number of pixels but covers a very small field of view³.

³ This is a little theoretical; in practice, of course, a high definition is usually chosen when a high resolution is available.

1.3.2. Spatial filtering

A very common operation for reducing the spatial noise that is present in the image is spatial filtering, which consists of replacing every pixel in the image by an average, weighted or not, performed with respect to its neighbors. This operation is very well described using the Fourier transform, and it is easy to calculate the consequences on spatial resolution.

This kind of filtering is basically a convolution operation. For the sake of simplicity, let us use a one-dimensional description with continuous variables, as sampling introduces some complexity without changing the basic concepts. The simplest filtering is the moving average, where each point of a signal g(x) is replaced by the signal average over a distance L around this point:

$$\overline{g}(x) = \int_{x-\frac{L}{2}}^{x+\frac{L}{2}} g(u) \mathrm{d}u$$
[1.12]

This expression can be described as a convolution product with a rectangle function $\Pi(x)$ defined as:

$$\Pi(x) = \begin{cases} 0 & \text{if } x < -1/2 & \text{or } x > 1/2 \\ 1 & \text{else} \end{cases}$$
[1.13]

Equation [1.12] can, indeed, be rewritten as:

$$\overline{g}(x) = \int_{-\infty}^{+\infty} g(u) \Pi\left(\frac{x-u}{L}\right) du = (g * \Pi_L)(x)$$
[1.14]

where * denotes a convolution product and $\Pi_L(x) = \Pi(x/L)$. The effect of such a filtering can be seen in the Fourier space (or frequency space), by taking the Fourier transform of the previous equation:

$$\hat{\overline{g}}(f_x) = \hat{g}(f_x) \times \hat{\Pi}_L(f_x)$$
[1.15]

which gives:

$$\hat{\overline{g}}(f_x) = \hat{g}(f_x) \times L\operatorname{sinc}(Lf_x)$$
[1.16]

where sinc is the "cardinal sine" function: $\operatorname{sinc}(u) = \operatorname{sin}(\pi u)/\pi u$. The initial frequency spectrum (the Fourier transform of signal g) is attenuated by the transfer function, in this case $\operatorname{sinc}(Lf_x)$ which reaches its first zero for $Lf_x = 1$, or $f_x = 1/L$ (Figure 1.7).

Gaussian filtering is also used very frequently:

$$\overline{g}(x) = \int_{-\infty}^{+\infty} g(u) G_{\sigma}(x-u) \mathrm{d}u$$
[1.17]

where

$$G_{\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$
[1.18]

is a normalized Gaussian function having a standard deviation σ . Its Fourier transform is:

$$\hat{G}_{\sigma}(f_x) = \exp(-2\pi^2 \sigma^2 f_x^2) = \sigma' \sqrt{2\pi} G_{\sigma'}(f_x)$$
[1.19]

where $\sigma' = 1/2\pi\sigma$. It is also a Gaussian function, which has the property of decreasing very rapidly when its variable increases. This provides interesting properties for noise filtering because spatial noise is mostly present at high frequencies (typically, noise changes from one pixel to another, which corresponds to the highest frequency in a sampled image).



Figure 1.7. Cardinal sine function $sinc(Lf_x)$, which is the transfer function of *a moving average over a length* L

More generally, any (linear) filtering can be described as a convolution with a certain function that is usually called the filtering *kernel*.

Numerically, linear filtering is efficiently implemented in the Fourier plane because efficient discrete Fourier transform algorithms (the so-called fast fourier transform – FFT) are available. In the Fourier plane, the convolution is replaced by a simple multiplication with the transfer function, the Fourier transform of the filtering kernel. However, these FFT algorithms can only be used if the image topology is rectangular, in other words when no invalid pixels or holes exist in the image. Holes are zones with no measurements and cannot be simply replaced by zero. If holes are present, some kind of interpolation has to be used to fill them so that no discontinuity arises along their edges, because these possible discontinuities will introduce adverse effects in the signal spectrum. This is not a trivial operation, and it may finally be simpler to implement the convolution.

Successive filterings correspond to successive convolution products. As the convolution product is associative, this can be considered as a single filtering using a

kernel that is the convolution product of all the different kernels used. As an example, if three filterings having kernels $K_1(x)$, $K_2(x)$ and $K_3(x)$ are used:

$$\overline{g}(x) = \{ [(g * K_1) * K_2] * K_3 \}(x) = [g * (K_1 * K_2 * K_3)](x)$$
[1.20]

we obtain something equivalent to a single filtering with kernel $(K_1 * K_2 * K_3)(x)$. It can be shown [ROD 78] that with a high number of successive filterings with the *same* filter, we tend toward Gaussian filtering; this is a consequence of what is known as the central-limit theorem.

As for spatial resolution, each filtering widens an object point, since the result of the filtering of a point located at x_0 and described by a Dirac function $\delta(x - x_0)$ is simply the filtering kernel translated at the same point, because:

$$\delta(x - x_0) * K(x) = K(x - x_0)$$
[1.21]

It can also be shown [ROD 78] that the "width" Δ_{f*g} of the convolution product of two functions f and g is the quadratic sum of the widths of each function:

$$\Delta_{f*g} = \sqrt{\Delta_f^2 + \Delta_g^2} \tag{1.22}$$

This equation allows us to calculate how the spatial resolution evolves as a function of successive spatial filterings.

To end this section, let us mention that this approach for the study of spatial resolution by the frequency transfer function was the basis of the work performed by the GDR 2519 research group, reported in [BOR 09], with respect to image correlation.

1.4. Classification of optical measurement techniques

In this section, we propose a systematic classification of different optical techniques for the measurement of kinematic fields. This classification is based on a limited number of basic key concepts.

In this context, infrared thermography has a slightly special position. It enables the measurement of a surface temperature, and not a kinematic value such as displacement, strain, slope and curvature.

In fact, the signal output by infrared sensors depends on the irradiance received from emitting bodies, and its transformation into a temperature is not a trivial task. The signal-to-energy relationship first involves Stefan–Boltzmann's and Planck's laws. Then, ignoring any parasitic radiation, translating received energies into the temperature of the target is only possible if the target has a known emissivity, which is a necessary condition. Other aspects of course need to be taken into account to ensure the accuracy of the temperature measurement. They are presented in Chapter 16 of this book.

1.4.1. White light measurement methods

In this family, the measurand is encoded in the spatial variation of light intensity, this variation being obtained in a non-interferometric way; in other words, we are excluding everything related to interferometric fringes. Purely optical aspects may be restricted to the image-forming process: measuring displacements by using image correlation has nothing really "optical" in nature, but is only based on a geometrical phenomenon that is present in images.

1.4.1.1. Encoding techniques

The measurand encoding, shown in Figure 1.1, is mainly of two kinds:

Encoding through a random signal: the signal is characterized by its local random variation, acting as a signature; the receiving system will have to identify this signature to complete the measurement; in this category lie all the methods using image correlation, including the so-called (laser) speckle correlation; these methods are presented in Chapter 6.

Encoding through a periodic signal, more precisely through the modulation of the phase (or, equivalently, of the frequency) of a spatial sinusoidal signal, called a "carrier"; as an example, we can cite the displacements of the lines of a grid deposited on a substrate, or the displacements of the lines of a grid reflected by a mirror-like surface where the local slopes are not uniform; these methods are presented in Chapter 3.

It should be noted that *each technique exists under both forms* (random encoding or phase encoding), even if the terminology does not always help to recognize this fact. There are of course major differences in performance resulting from these different encoding approaches:

- The practical realization of random encoding, which is essentially noise, is *difficult to fully characterize*, using complex notions such as statistical moments (there are plenty of such "moments") or average power spectral density (PSD); the correlation method for in-plane displacement measurements often requires paint to be sprayed on the body under examination to produce a "speckle", a random contrasted pattern; it is obvious that this manual process cannot easily be made repeatable, and the "quality" of the obtained pattern is difficult to characterize.

- On the other hand, encoding through the modulation of a spatial periodic carrier can be fully characterized by its frequency, its local phase, its local average level, its local contrast and its harmonic content.

- Random encoding *is not quantitative by itself*; the information lies within a local contrast morphology or signature that the detection system has to identify.

– On the contrary, encoding by the phase modulation of a carrier is *quantitative* because the signal (e.g. the displacement of a grid line) corresponds to a number which is the amount of phase modulation. Phase detection techniques are very well established and efficient (they mostly rely on Fourier analysis) and allow easy characterization.

- A drawback of phase encoding is its *periodicity*: the same code value periodically repeats as the signal increases, which leads to *ambiguities*. Removing these ambiguities corresponds to what is called *phase unwrapping* that suppresses the 2π jumps appearing in the detection process that only outputs the phase modulo 2π . Phase unwrapping is, in fact, simply *numbering the fringes* (or lines, depending on the method).

- In general, correlation methods are much easier and cheaper to implement; either the part under test is sufficiently textured to allow the use of digital image correlation, or it suffices to spray paint on it. This is much simpler than gluing or engraving or in some way depositing a grid onto the surface, not to mention the fact that grids can only be placed on flat or cylindrical surfaces.

1.4.1.2. Examples

Almost all white light techniques consist of measuring a position or a displacement in an image. A simple geometrical analysis of the measurement system allows us to understand the measurement principle.

1) Measurement of in-plane displacements: here, the object undergoing displacements is simply observed using a camera. The displacements of marked points of the object correspond to displacements of image points. The primary output of the technique is the displacement field of the image points. The available techniques are the following.

i) Grid technique: a grid has to be deposited on the surface; the local displacements will modulate the phase of the grid acting as a periodic carrier. Its main drawback is the required surface preparation.

ii) $Moiré^4$ is only an additional layer on the grid technique, since the nature of the encoding is the same. It consists of decreasing the spatial frequency of the carrier using a beat phenomenon, exactly as stroboscopy decreases the temporal frequency of a vibration.

iii) Deflectometry: in this technique, derived from the previous one, the object reflects a fixed regular grid pattern. Distortion of the grid image is caused by non-uniformity of the surface slopes. Thus, this technique enables slope fields to be measured. It is necessary that the surface be at least partially reflective, in which case

⁴ The first letter should not be capitalized within a sentence, as it is a common noun and not the name of an inventor.

the technique is quite easy to implement since no surface preparation is required. If not, a revealing fluid can be wiped onto the surface to make it reflective.

iv) Digital image correlation: this is the speckle implementation of the same concept, and the surface points are marked using the surrounding local random contrast pattern.

2) Shape measurement: Shape measurement techniques derive from displacement measurement techniques. The fundamental geometrical arrangement (Figure 1.8) is the same for stereocorrelation, where two cameras observe the object from two different points of view, and for fringe projection (or structured light projection). If the position in space of the optical centers C_1 and C_2 are known, knowing the position of points M'_1 and M'_2 allows the position of point M to be determined. The difference between the two techniques lies in the trajectory and nature of the information:

i) With stereocorrelation, the information starts from point M and propagates toward points M'_1 and M'_2 , located on two camera sensors; the analysis software has to recognize the local signature around M in the images to pair M'_1 and M'_2 .

ii) With fringe projection, the information (which is quantitative as it is a fringe phase, proportional to the abscissa or ordinate of M'_1) starts from M'_1 , which is most often a video projector pixel, and propagates toward M and afterward M'_2 . The analysis software has only to detect the phase and unwrap it⁵ to know exactly the emitting point because its coordinates are known from the phase.



Figure 1.8. Shape measurement using stereocorrelation or fringe projection

1.4.2. Interference methods

The temporal phase $\psi(t) = \omega t + \phi$ of an electromagnetic wave in the visible domain, such as the one emitted by an helium-neon ($\lambda = 633$ nm) or YAG ($\lambda = 532$ nm) laser, vibrates at a frequency around one hundred terahertz (10^{14} Hz).

⁵ In this technique, unwrapping should also preserve an absolute phase origin.

There does not exist any sensor capable of recording temporal variations at such high frequencies, and they are only sensitive to the average light power value. This is why the interference phenomenon is useful to obtain phase information. The nature of interference is a beat phenomenon between two periodic signals of same frequency, as in stroboscopy or moiré, that will remove the term ωt present in ψ . The formula that gives the light intensity in two-wave interferences is:

$$I = I_0 (1 + \gamma \cos \Phi) \tag{1.23}$$

where γ is the visibility (or contrast) and the phase Φ of the interference fringes is:

$$\Phi = \psi_2 - \psi_1 = \phi_2 - \phi_1 = 2\pi \frac{\delta}{\lambda}$$
[1.24]

where δ is the optical path difference between the two interfering beams and λ is the wavelength in the propagation medium, most often air. In fact, an interferometric measurement is always performed between an initial state that acts as the reference (due to imperfections, we never have an initial state without any interference fringes) and a final state. It is therefore always an *interference fringe phase variation*

$$\Delta \Phi = (\phi_{2,\text{final}} - \phi_{1,\text{final}}) - (\phi_{2,\text{initial}} - \phi_{1,\text{initial}})$$
[1.25]

which is used (double difference).



Figure 1.9. Different interaction modes between incident light and a surface: specular (mirror-like) reflection, diffusion, diffraction

1.4.2.1. Light-surface interactions

Depending on the way the incident light, having smooth (plane or spherical) or rough (random) wave fronts, interacts with the surface under examination, different interferometry techniques can be used. The different possible interactions are outlined in Figure 1.9.

- In the first case, the surface is a mirror. All the energy is reflected in a single direction, following Snell's law, and the wave front remains smooth.

- In the second case, the surface is rough (diffusive) and the energy spreads in all directions. The wave front shape becomes very complicated, which is the reason why they are referred to as 'random'; this will create spatial noise in an image obtained by an imaging system placed downstream. This spatial noise is called speckle.

- In the last case, a diffraction grating⁶ is deposited (in general, molded) onto the surface so that the first diffracted order is normal to the surface. In this case, a noticeable amount of energy is re-emitted along the normal direction, and the wave front remains smooth.

1.4.3. Sensitivity vector

When a surface point that is illuminated by a monochromatic wave undergoes a small displacement characterized by the vector \vec{u} , the phase variation of the beam coming from the light source and re-emitted toward the observer T is calculated from the variation in the optical path. It is easily shown that this phase variation $\Delta \phi$ is

$$\Delta \phi = \vec{g}.\vec{u} \tag{1.26}$$

where the difference between the illumination and observation wave vectors $\vec{g} = \vec{k}_i - \vec{k}_o$ is called the *sensitivity vector*. Its direction indicates *the displacement component the measurement will be sensitive to*. In the case of a mirror-like reflection, the sensitivity vector is normal to the surface; thus only out-of-plane displacements can be measured (Michelson interferometry).



Figure 1.10. Sensitivity vector

1.4.4. Synthetic sensitivity vectors

After a measurement, an interference fringe phase variation field is obtained, which is locally proportional to the sensitivity vector \vec{g} . If different mesurements of the same kinematic field can be made with different illumination conditions, different phase variations for the same displacement are obtained:

$$\begin{cases} \Delta \phi_{0} = \vec{g}_{0}.\vec{u} \\ \Delta \phi_{1} = \vec{g}_{1}.\vec{u} \\ ... \\ \Delta \phi_{N-1} = \vec{g}_{N-1}.\vec{u} \end{cases}$$
[1.27]

⁶ A grating is a very dense pattern of straight parallel lines. A grating has the property of splitting an incident ray into many different rays (called diffraction orders) having different directions.

With a simple linear combination of these equations, that is of the obtained phase images, the displacement vector appears factorized with the corresponding linear combination of sensitivity vectors, which is called a *synthetic sensitivity vector*:

$$\sum_{i=0}^{N-1} \alpha_i \Delta \phi_i = \left(\sum_{i=0}^{N-1} \alpha_i \vec{g}_i\right) . \vec{u} = \vec{G} . \vec{u}$$
[1.28]

Such a linear combination may also appear in an analog way with a setup that uses *simultaneously* different illumination directions, as in interferometric moiré (see section 1.4.5.2).

1.4.5. The different types of interferometric measurements

These simply correspond to the various possible interactions of the incident beam with the surface presented in Figure 1.9 and to the different ways of choosing the interfering beams. In the following, we will use the term "object beam" to designate the incident beam; it is the beam that illuminates the object.

1.4.5.1. Interference between the object beam and a fixed reference beam

In this case, we have for the double difference formula [1.25]:

AO3

$$\Delta \Phi = \Delta \phi_{\rm obj} - \Delta \phi_{\rm ref} = \Delta \phi_{\rm obj} = \vec{g}.\vec{u}$$
[1.29]

so it can be seen that the measured quantity is the *displacement component along the* sensitivity vector \vec{g} . For example, the measurement of the approximate out-of-plane displacement using interferometry in diffuse light (speckle interferometry) is shown in Figure 1.11, where the measured displacement component u_g is aligned along the bissector of the illumination and observation directions. It can be seen that there is some degree of approximation, as this bissector is not exactly normal to the surface. Existing commercial systems may or may not take this effect into account.



Figure 1.11. Interferometry in diffuse light for the measurement of out-of-plane displacements ("out-of-plane speckle" setup)

1.4.5.2. Interference between two object beams coming from the same point

The two beams come from the same object point, but with different incident illumination directions, so the sensitivity vectors are different. Then the double difference formula [1.25] is written:

$$\Delta \Phi = \Delta \phi_2 - \Delta \phi_1 = (\vec{g_2} - \vec{g_1}).\vec{u} = \vec{G}.\vec{u}$$
[1.30]

Thus, we obtain, in an "analog" way, a synthetic sensitivity vector. As a typical example, we have what is often referred to as interferometric moiré, but is more relevantly called *grating interferometry*, outlined in Figure 1.12.



Figure 1.12. Measurement of in-plane displacements: grating interferometry (also called interferometric moiré); corresponding sensitivity vector

Regarding grating interferometry, the main drawback is the difficulty in depositing the grating on the surface. A typical maximum value of the grating size that can be placed is 40–50 mm. The main advantage is the absence of speckle noise in the image, leading to excellent spatial resolution for the measurements.

The same approach can be used with diffuse light, as indicated in Figure 1.13. The major difference (not speaking of the absence of the complicated process of depositing a grating onto the surface) is that the incident beams are not collimated in general, in order to investigate large surfaces.

1.4.5.3. Interference between object beams coming from two contiguous points (differential interferometry, shearography)

A splitting system placed in the imaging arm of the setup enables us to obtain on the sensor two images of the object, laterally shifted by a small quantity $\delta \vec{l}$ (Figure 1.14). The double difference formula [1.25] becomes in this case:

$$\Delta \Phi = \Delta \phi_2 - \Delta \phi_1 = \vec{g}.(\vec{u} + \delta \vec{u}) - \vec{g}.\vec{u} = \vec{g}.\delta \vec{u}$$
[1.31]



Figure 1.13. Measurement of in-plane displacements ("in-plane speckle setup")



Figure 1.14. Differential grating interferometry, or "shearography" if no grating is present

What is measured is then a *differential displacement*, which is very interesting for mechanical engineering, since the differential displacement basically relates to strain. With different illumination directions, we can acquire different phase fields that can be combined in order to obtain synthetic sensitivity vectors enabling the measurement of in-plane and out-of-plane displacements. Once again, grating interferometry or interferometry in diffuse light can be used (in other words, differential interferometry can be used with or without a grating), with the already mentioned drawbacks and advantages. The term "shearography" [HUN 07] is in general reserved for realizations in diffuse light.

Regarding terminology, it should be noted that the term "shear" which appears in many papers dealing with differential interferometry (and from which "shearography" derives) designates the lateral displacement introduced by the splitting component in the imaging system. It has nothing to do with "shear" as the non-diagonal strain component as understood by mechanical engineers. When aligning the lateral shift direction with the x and y axes, and using different illumination directions, all the derivatives of all component displacements can be obtained. Thus, we can obtain all the in-plane strain components and all the slope⁷ (out-of-plane displacement gradient) components.

To conclude, let us emphasize a very interesting feature of differential interferometry: it is basically insensitive to vibrations, as vibrations introduce rigid body displacements but not displacement gradients, making the technique very usable in industrial environments. This is why shearography is now becoming accepted as a non-destructive technique in the aeronautical industry.

1.4.6. Holography, digital holography

Holography is not, strictly speaking, a measurement technique; it is a *recording* technique for an electromagnetic field, both in amplitude and in phase, so that it can be reconstructed afterward. The principle is to make the field interfere with a so-called "reference" beam, and to record the resulting interference fringes. The recording used to be photographic, with special high-resolution photographic plates, but is now mostly digital.

Holography offers the possibility to create interferences between waves that do not exist simultaneously, for example waves diffused by an object before and after deformation. Thus, holography can be seen as a complement to all the interferometric techniques presented before.

It should be noted that there is no real difference between digital holography and the interferometric setup using a smooth reference beam like the one presented in Figure 1.11, as it is in both cases the recording of interferences between an object and a reference beam.

The only difference is that in cases called "speckle interferometry", the recording takes place in a location where the object has already been imaged, so the object reconstruction required in digital holography is not necessary. When we remove the camera lens shown in Figure 1.11 and reconstruct the object from the recorded interference fringes, we perform digital holography.

⁷ We are speaking here of slopes introduced by mechanical loading.

1.4.7. Conclusion

We have presented in this section an approach to classify a large number of optical techniques in terms of their fundamental features (white light or interferometric techniques, random encoding or phase modulation, nature of the measurand). Table 1.1 proposes a very short synthesis, briefly indicating their performances. Of course, such a table is only a very coarse presentation of reality. However, it is hoped that it may be useful for a first approach to the choice of a full-field optical measurement technique.

White light			
Measurand	Random encoding	Phase modulation	Remarks
In-plane disp.	Image correlation	Grid (with or without moiré)	Coupling with out- of-plane disp.
Shape	Stereocorrelation	Fringe projection (structured light)	Transverse calibration of camera required
3D disp.	Stereocorrelation with deposited speckle	Impossible	id.
Slopes	Not used	Deflectometry	Coupling with shape
Interferometry			
Measurand	Reflected light	Diffuse light	Diffracted light
In-plane disp.	Impossible	In-plane speckle	Interferometric moiré
Out-of-plane disp.	Michelson interferometry	Out-of-plane speckle	Not used, but a setup is conceivable
Differential setup (slopes, strain)	e.g. Nomarski microscopy	Shearography	Grating differential interferometry
Performances			
	White light, random	White light, phase mod.	Interferometry
Simplicity	++	+	_
Cost	-	-	+
Performances	_	-+	++

Table 1.1. Synthesis

1.5. Bibliography

- [AST 08] Standard Guide for Evaluating Non-Contacting Optical Strain Measurement Systems, DOI: 10.1520/E2208-02, ASTM International, 2008.
- [BIS 03] BISSIEUX C., PRON H., HENRY J.-F., "Pour de véritables caméras matricielles de recherche", *Contrôles Essais Mesures*, pp. 39–41, January 2003.

AO4

- [BOR 09] BORNERT M., BRÉMAND F., DOUMALIN P., DUPRÉ J.-C., FAZZINI M., GRÉDIAC M., HILD F., MISTOU S., MOLIMARD J., ORTEU J.-J., ROBERT L., SURREL Y., VACHER P., WATTRISSE B., "Assessment of digital image correlation measurement errors: methodology and results", *Experimental Mechanics*, vol. 49, no. 3, pp. 353–370, June 2009.
- [HON 05] HONORAT V., MOREAU S., MURACCIOLE J.-M., WATTRISSE B., CHRYSOCHOOS A., "Calorimetric analysis of polymer behaviour using an IRFPA camera", *International Journal on Quantitative Infrared Thermography*, vol. 2, no. 2, pp. 153–172, 2005.
- [HUN 07] HUNG M.Y.Y., CHEN Y.S., NG S.P., SHEPARD S.M., HOU Y., LHOTA J.R., "Review and comparison of shearography and pulsed thermography for adhesive bond evaluation", *Optical Engineering*, vol. 46, 2007.

AQ5

- [ROD 78] RODDIER F., Distributions et transformation de Fourier, McGraw-Hill, Paris, 1978.
- [VIM 12] International Vocabulary of Metrology–Basic and General Concepts and Associated Terms, available at http://www.bipm.org/en/publications/guides/vim.html, JCGM, 2012.

Author Queries

- [AQ1] The term [unity] has been changed to [one]. Please check.
- [AQ2] The sentence [We can rewrite....] OK as edited?
- [AQ3] The term [bissector] is not found in dictionary. Do you mean [bisector] instead?
- [AQ4] Please provide the volume number for the reference [BIS 03].
- [AQ5] Please provide the page range for reference [HUN 07].