Digital Image Correlation and Fracture: An Advanced Technique for Estimating Stress Intensity Factors of 2D and 3D Cracks
Stéphane Roux, Julien Réthoré, François Hild

To cite this version:

HAL Id: hal-00381646
https://hal.archives-ouvertes.fr/hal-00381646
Submitted on 6 May 2009

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.
TOPICAL REVIEW

Digital Image Correlation and Fracture:
An Advanced Technique for Estimating Stress Intensity Factors of 2D and 3D Cracks

S Roux,1 J Réthoré2 and F Hild1
1: Laboratoire de Mécanique et Technologie (LMT-Cachan)
ENS de Cachan / CNRS / UPMC / PRES UniverSud Paris
61 Avenue du Président Wilson, F-94235 Cachan Cedex, France.
2: Laboratoire de Mécanique des Contacts et des Structures (LaMCoS)
INSa Lyon / UMR CNRS 5259
20 Avenue des Sciences, F-69621 Villeurbanne Cedex, France.
E-mail: stephane.roux@lmt.ens-cachan.fr,
julien.rethore@lamcos.insa-lyon.fr, francois.hild@lmt.ens-cachan.fr

Abstract. Digital image correlation is a measurement technique that allows one to retrieve displacement fields “separating” two digital images of the same sample at different stages of loading. Because of its remarkable sensitivity, it is not only possible to detect cracks with sub-pixel opening, which would not be visible, but also to provide accurate estimates of stress intensity factors. For this purpose suitable tools have been devised to minimize the sensitivity to noise. Working with digital images allows the experimentalist to deal with a wide range of scales from atomistic to geophysical one with the same tools. Various examples are shown at different scales, as well as some recent extensions to three dimensional cracks based on X-ray Computed micro-tomographic images.

PACS numbers: 46.50.+a Fracture mechanics, fatigue and cracks;
46.80.+j Measurement methods and techniques in continuum mechanics of solids;
07.05.Pj Image processing.

Special issue: “Fracture: from the atomic to the geophysical scale”
1. Introduction

It is a banality to observe that images, and imaging, are playing a more and more important role in everyday life, as well as in science. However, what is less trivial is to utilize the potential of imaging techniques as *measurement devices*. Powerful pattern characterization is obtained from sophisticated segmentation techniques (identifying phases, objects and their size or shape), which are ultimately limited in their spatial resolution to the pixel level. Applications to electronic microscopy or atomic force microscopy images have provided with the years amazing examples of their capabilities [1, 2].

One of the popular measurement techniques using pictures [3] is Digital Image Correlation (DIC). The latter consists in comparing two images of the same scene, typically an object under load, and retrieving the displacement field that allows for the best match. In its spirit and even in some implementations, DIC is very similar to the corresponding tool very frequently used in fluid mechanics, which is known as “Particle Image Velocimetry” or PIV. DIC [4, 5, 6] appeared at about the same time as PIV [7, 8, 9, 10, 11]. However, because of the typical magnitude of strains and displacements that are usually much smaller in solid mechanics than in fluid flow analyses, its development was a bit slower since the demand was more challenging to meet. However, today it can be said that the common performances of such techniques allow one to use them in most instances of experimental mechanics.

When time is added, and a series of images is obtained, the comparison between these images, and their minute differences are analyzed very accurately, and sub-pixel features are revealed. This statement may appear as extremely counter-intuitive, as the intrinsic discretization of the information at the pixel scale may appear as an unbreakable barrier. The key to this paradox is to understand that these sub-pixel distances we are referring to are shared in common by a large number of pixels, and even though one single of them would not be able to provide sub-pixel accuracy, a large collection will succeed.

We are here at the heart of the problem of digital image correlation, namely, the above mentioned spreading of information over many pixels makes the problem ill-posed in the mathematical (Hadamard) sense, and hence calls for regularization. However as this regularization will be relaxed the analysis will be more and more prone to uncertainties. The balance between the level of regularization and the sensitivity to noise motivates the most recent developments and successes of DIC analyses. First, *regularization*, which in non-mathematical terms is a kind of guidance to the type of displacement fields that are looked for, can be engineered to incorporate the best *a priori* knowledge on the mechanical test / material behavior at hand. Second, *noise robustness* is a problem to be tackled as early as possible in the complete process from images to, say, stress intensity factors (SIFs). In the sequel, rather than presenting an exhaustive and historical picture of DIC, and its application to cracks, we will rather focus on the most recent developments and emphasize on the above regularization and
noise robustness aspects.

Let us mention the specific properties of DIC that makes it extremely attractive in the field of experimental mechanics:

- Optical Digital Image Acquisition equipments from CCD to CMOS sensor-based cameras are available with an ever increasing quality and decreasing cost. The resolution (or number of pixels) reaches nowadays 10 Mpixels. The gray level and color coding depth commonly exceed 8 bits (e.g., 12 to 16-bit deep devices can be purchased today).

- Image acquisition is a non contact technique, and even with a large magnification it tolerates a large distance from the sample to the camera. Long distance (far-field) microscopes are located 15 to 35 cm away from the sample surface and yet provide images whose physical pixel size lies in the micrometer range [12, 13]. This non contact character makes it a technique of choice for unfriendly environments (high temperature, humid, or corrosive atmosphere).

- Fast digital cameras also provide images with a very high frequency (from about $10^4$ to more than $10^6$ frames per second, full resolution), and thus it opens the way to detailed analyses of dynamic tests. Exposure time as small as fractions of microseconds are achieved, and yet provide images that may be analyzed with DIC [14, 15].

- As will be seen below, the image surface has to be patterned to be analyzed. This pattern can however be the natural surface of the specimen if it is textured and opaque enough (e.g., concrete material, raw surface of a ceramic) but otherwise, paints are easily sprayed on the surface to provide the required texture without affecting the mechanical behavior, and in a reversible way if required for further analysis.

- Any (other than optical) digital image can be processed in the same way provided distortions, aberrations are not too large and uncontrolled [16, 17, 18].

- Basic DIC dealing with image pairs only analyzes in-plane displacements, however using two (or more) cameras provides the relevant information for the extraction of full three dimensional displacement fields of the surface of a solid [19]. This technique, called “stereo correlation,” will not be developed further in this paper even though it was applied to the analysis of cracked samples [20].

- The result of an analysis, typically a full field measurement of the displacement, contains a considerable amount of data compared with classical instrumentations available on universal testing machines (e.g., extensometers, strain gauges). This calls for specific exploitations in the field of constitutive law identification [21, 22].

In the following, the discussion will focus on the analysis of cracks by means of DIC. The use of full field measurements is of particular interest when dealing with (strong) discontinuities induced by the presence of cracks. Fracture mechanics has benefited from DIC results [23, 24]. For instance, crack tip opening angles [25] or crack tip
opening displacements \([12, 19, 13]\) are measured with a very good accuracy by means of DIC. The technique can be applied to small areas \([18]\), and to large structures such as aeronautical parts \([26]\) or bridges \([27]\) with the \textit{same} algorithms. The fact that displacement uncertainties of the order of \(10^{-2}\) pixel are achieved allows this technique to reveal cracks that cannot be seen if their opening is less than one pixel wide.

All the previous analyses use local estimates of the displacements by considering \textit{independent} interrogation windows. An alternative solution is to use global (or Galerkin) approaches to measure displacement \textit{fields}. It is based on a variational formulation of the conservation of the brightness \([28, 29]\). When particularized to finite element shape functions \([30]\), it allows one to use all the numerical tools to analyze cracks (\textit{e.g.}, X-FEM \([31, 32]\)). This procedure was followed to develop eXtended Digital Image Correlation (or X-DIC \([33, 34]\)) to analyze straight or curved cracks \([36]\).

Besides the qualitative aspect of detecting a crack, the next challenge is to be able to reach a quantitative estimate of mechanically significant properties, such as opening loads in fatigue experiments \([37, 38]\) and/or SIFs \([23]\) by least squares fit using closed-form solutions \([39]\). One related key quantity is the precise determination of the crack tip location, which has a direct consequence on the amplitude of SIFs \([40]\). This problem is, for instance, addressed by using supersingular fields \([38, 41]\). An alternative solution is to resort to \(J\)-integral \([42]\), or its domain version \([43]\) to use more measured data, namely, to reduce its sensitivity to measurement uncertainties \([44]\).

In all the previous identification analyses, the measurement and identification steps are performed sequentially, namely, displacements are first measured and subsequently post-processed to determine SIFs or even \(J\)-integrals. Since closed-form solutions are available to analyze displacements in the vicinity of the crack tip, they can be used directly in a correlation code. They are referred to as integrated DIC \([45, 40, 38]\) since \textit{a priori} knowledge on the type of mechanical behavior is sometimes possible. As a consequence, SIFs appear as generalized degrees of freedom that are measured directly. In that case, the measurement and identification stages are carried out in one single step. Moreover, this type of approach can be used to combine X-FEM with X-DIC to extract SIFs from displacements that minimize a weighted functional that combines brightness conservation and equilibrium conditions \([41]\). It allows for SIF determinations with very small uncertainties.

Up to now, only 2D displacement fields were discussed. Stereovision \([19]\) is also used to evaluate 3D displacements on external surfaces of a body using camera pairs \([20]\). Although this technique is sometimes referred to as 3D, the fact that the support is two-dimensional is ambiguous, and 3D-DIC will be reserved in the sequel to the determination of 3D displacements in a volume\(\dagger\).

To monitor phenomena \textit{within} opaque materials, X-Ray Computed MicroTomography (XCMT) is a very powerful way of imaging material microstructures in a non-destructive manner \([46]\). In particular, cracks can be observed \([47, 48]\), their closure \([49]\),

\(\dagger\) The latter is sometimes referred to as Volumetric DIC or V-DIC.
Digital Image Correlation consists in analyzing a series of images, from which displacement fields are measured so as to match either a first image considered as a reference one (typically the unloaded stage) and each subsequent image, or (if displacement amplitudes are too large) consecutive pairs of images. In the latter case, the displacement from the reference image is reconstructed as a sum of elementary displacements taking into account non-linearities induced by large displacements. We thus focus here on the analysis of image pairs, one being the “reference” image, while the second is the “deformed” image (Figure 1).

2.1. Images

Usually color is not considered, although future developments may benefit from it. A digital image is represented as a simple scalar valued number (the gray level) at each pixel, and denoted by $f(x)$. The $x$ argument is discretized into elementary pixels, usually in two dimensions, along a regular square grid (the pixel matrix). The gray level value $f$ is also discretized using from 8- to 16-bit encodings. Moreover, for optimal results, the image texture should be a rapidly varying field with sharp contrast and small correlation length. Hence, even though it is tempting to consider theoretically some continuation of the field, ignoring both pixel and gray level discretizations, one should be extremely careful in so doing as $f$ is expected to be very irregular.

In spite of this warning, if one wants to give some meaning to sub-pixel displacements, it is necessary to account for the gray level change if a pixel is moved by an arbitrary value, in particular less than its size. An interpolation scheme has to be specified [56]. Various prescriptions have been used. The most elementary is a bi-
linear interpolation, which consists in dressing each elementary square of pixels with a combination of the four functions \(1, x, y\) and \(xy\) so as to match the gray values at each corner. This interpolation scheme gives satisfactory results even if it gives discontinuous derivatives across edges and at the nodes of the pixel matrix. Higher order interpolations (cubic, quintic spline functions \([57]\)) are also used and they typically lead to improvements usually at the expense of longer computation times. Another choice is given by Fourier decomposition of the original image that naturally provides a continuous field with infinite differentiability, and yet has a small computation cost, when Fast Fourier Transform algorithms are used. The main difficulty of this choice is the aliasing effect of non-periodic boundary conditions on the edge of images or sub-images. In the sequel, for the sake of simplicity, the chosen interpolation scheme is linear along all space directions (bilinear for 2D images and trilinear for 3D volumes).

### 2.2. Optical flow determination

The optical flow of an image sequence is a set of vector fields relating each image to the next one. The basic principle underlying DIC is the fact that the image is only deformed by the in-plane displacement field, without any further alteration of the gray levels. Therefore, if \(g\) denotes the deformed and \(f\) the reference images, these functions are related by

\[
g(x) = f(x + u(x))
\]

This equation is often referred to as the brightness conservation \([58, 59]\). It is easily relaxed if one proposes an alternative form of the image transformation. For instance, an AFM image is not simply a passive texture advection, but is also a topography of...
the surface. Hence it may be modified by out-of-plane displacements.

Other modifications may be observed, such as a change in illumination of the scene, and hence, an affine change in gray levels

\[ g(x) = (1 + a(x)) f(x + u(x)) + b(x) \]  

is often present. The gray level offset \( b \) and rescaling \( a \) may have to be determined in addition to the displacement unknowns. This relaxation of the brightness conservation is often included in order to make the algorithm more robust. (Regularizations on both fields \( a \) and \( b \) will have to be considered and they rely on additional assumptions such as the small values of both fields, and their slowly varying character in space.)

Furthermore, image acquisition is never perfect, and always contains a small amount of noise, so that

\[ g(x) = f(x + u(x)) + \eta(x) \]  

where \( \eta \) is a small amplitude noise, which is often considered as “white”, or spatially uncorrelated. Although these extensions of the brightness conservation will not be considered in the sequel, the noise corruption of the latter will be analyzed.

2.3. Regularization

From Equation (1), it is seen that the problem of determining \( u(x) \) is ill-posed as such, and even more so if the strict conservation is relaxed. Hence, some additional constraints, or regularizations, have to be included. This regularization usually consists in restricting the variability of the displacement field. The most usual form is to assume that \( u(x) \) is piecewise constant or linearly varying with \( x \) over zones, allowing for arbitrary discontinuities across zone boundaries. The extension of these zones, referred to as “Zones Of Interest” or ZOI, thus qualifies the degree of regularization. Other forms may be chosen such as e.g., imposing a global continuity for \( u \) through a decomposition over finite element shape functions, or for application to cracks, one may benefit from the a priori knowledge of analytical displacement fields.

To keep the discussion as general as possible, the regularization consists in a specific choice of a “library” of basis functions, \( \psi_n(x) \), such that \( u \) is restricted to a given functional space

\[ u(x) = \sum_n \omega_n \psi_n(x) \]  

and the DIC algorithm should provide the user with the “best” determination of \( u \) within this space, or equivalently “the best” \( \omega_n \).

Because of image imperfections, or simply because the actual displacement field is not exactly part of the chosen subspace, the brightness conservation cannot be fulfilled exactly, and hence a weaker form is preferred as the minimization of the quadratic difference, \( \Phi^2 \),

\[ \Phi^2(\omega_n) = \int \left[ g(x) - f(x + \sum_n \omega_n \psi_n(x)) \right]^2 \, dx \]  

also called global correlation residual.

2.4. Algorithm

Different strategies are distinguished. When $\psi$ is constant over a zone, and no continuity is searched for, it suffices to focus on this very zone. Moreover, if $\psi$ is constant over the ZOI, minimizing $\Phi^2$ is equivalent to maximizing the cross-correlation between $f$ and $g$ over this zone. This is conveniently computed in Fourier space, and a polynomial fit of the cross-correlation function allows one to locate its maximum with sub-pixel resolution. The advantage of this formulation is that a global search for the maximum is easy to perform and hence this technique is quite tolerant to large displacements [60].

As soon as more complex $\psi$ functions are used the cross-correlation property cannot be used directly as a global search algorithm, but rather through an iterative minimization procedure. A strategy that is quite performing consists in assuming that the sought displacement is small enough to allow for a Taylor expansion up to first order of the functional to minimize [61]. This hypothesis then transforms the problem into a simple quadratic minimization, and hence an elementary linear system is to be solved

$$[M]\{\omega\} = \{d\}$$

with

$$M_{mn} = \int [\nabla f(x) \psi_m(x)][\nabla f(x) \psi_n(x)]dx$$

and

$$d_m = \int [\nabla f(x) \psi_m(x)][f(x) - g(x)]dx$$

where $\{\omega\}$ denotes the column vector gathering all displacement amplitudes $\omega_n$.

This direct formulation is however generally not reliable because of the poor precision of the Taylor expansion, which is the result of having a fine grained contrasted texture for the reference image. However, if the images $f$ and $g$ are severely low-pass filtered, the small scale irregularity will be dampened and the Taylor expansion may turn out to work. Once a first determination of the displacement was obtained, the deformed image may be corrected so that the remaining displacement has smaller magnitudes. Therefore, less severely filtered images are used to determine the corrections of the displacement, and this is performed iteratively down to the stage where unfiltered images are used. This multiscale procedure is quite robust and easy to implement [60, 30]. It offers the possibility for a wide variety of displacement bases to be considered, and hence most of the sequel will exploit this freedom.

The basis $\psi_n$ was not specified up to now, and many different choices are possible with the present formulation. However, one very convenient form is provided by finite element shape functions. Because of the fact that digital images are typically discretized over a square grid, it is natural to consider Q4 shape functions over square shaped elements. For each element, four shape functions are to be used, namely $(1 \pm x)(1 \pm y)/4$
Figure 2. Displacement field for the SiC sample shown in Figure 1. The vertical (resp. horizontal) component of the displacement field is shown on the left (resp. right). The displacement field has been decomposed over square-shaped bilinear shape functions (Q4 in the language of finite elements) whose size is $\ell = 32$ pixels. The displacements are expressed in pixels (1 pixel = 1.85 $\mu$m).

When the square is mapped onto the $[-1; 1]^2$ reference domain. Each component follows the same decomposition, so that the unknowns $\omega$ carry a double index

$$u^e(x) = \omega_{n\alpha} N_n(x) e_\alpha$$

(9)

where $n$ is an index over the nodes belonging to a given element $e$, $N_n$ the so-called shape functions, and $\alpha = 1, 2$ specifies a space direction (of unit vector $e_\alpha$).

As an example, Figure 2 shows the displacement field obtained from a decomposition over bilinear shape functions with square shaped (Q4) elements of size $\ell = 32$ pixels. In spite of the fact that the entire displacement field is continuous, it is easy to distinguish a linear segment over which the displacement gradient is large, which suggests that a crack may be present. From this figure, one observes that the displacement discontinuity is less than half a pixel wide, and hence it becomes clear that it cannot be seen with bare eyes in Figure 1.

To illustrate the fact that this technique also allows for large displacements to be measured, an example of a crack in rubber (SBR with carbon black) is shown in Figure 3 [62]. The corresponding displacement field is given in Figure 4. The amplitude of the displacement discontinuity is remarkably large.

It is also to be noted that this technique allows for the analysis of pictures taken at very different scales. To illustrate this point, Figure 5 shows a civil engineering frame structure made of reinforced concrete, and Figure 6 the displacement field in a damaged state, as well as the residual field that reveal very clearly the multiple crack pattern. This example shows the benefit of using the residual field in identifying crack geometries. Only the largest cracks can be seen on the original deformed image, but the residual field highlights even minute discontinuities. In such a case, the determination of the stress intensity factors at each crack tip would be difficult, but it would also be inappropriate at the scale shown in the figure. Moreover the full 3D geometry in the bulk of the sample is unknown. The degradation of image registration along cracks can
Figure 3. Reference (left) and deformed (right) pictures of a crack in a rubber specimen (courtesy of S. Mzabi). The pixel size is 9.38 µm.

Figure 4. Vertical (left) and horizontal (right) components of the displacement for the crack in rubber. The displacement field has been decomposed over Q4 elements whose size is ℓ = 32 pixels. The color scales for the displacements are expressed in pixels (1 pixel = 9.38 µm).

be exploited to design an automatic detection procedure for multiple crack patterns as recently proposed by Helm [35].

2.5. Quality evaluation

The minimum achievable value of Φ² gives a simple and global scalar evaluation of the quality of the matching between both images. More interestingly, this objective function is read as the surface integral of a density, and hence rather than a global quality factor, one has naturally access to a field of local residual, ρ(x) = |g(x) − f(x + u(x))|. This is especially interesting as this field can be interpreted when being non uniform (Figure 6), and it provides guidance to enrich the functional displacement basis or to mask a region that cannot be trusted (e.g., if a paint fragment has detached from the surface).

Moreover, this residual field may reveal specific features such as the presence of an invisible crack. In the example of Figure 1, the crack in the “deformed” image cannot be distinguished with bare eyes (its opening is much less than one pixel or 1.85 µm). However, after correction from the displacement field shown in Figure 2, the residual
Figure 5. Large scale view of part of a reinforced-concrete frame (top), zoom on an analyzed region in its reference (bottom left) and deformed (bottom right) state. The pixel size is 370 μm. Note that large cracks can be guessed in the deformed image in spite of the surface marking.

Figure 6. Vertical displacement expressed in pixels (left) and gray level residual map (right) revealing a complex multiple crack pattern (1 pixel = 370 μm).

field is obtained and shown in Figure 7. There, the crack is very clearly seen as being supported by a straight line. An automatic procedure to determine this crack support may be implemented in the form of Radon transforms [34, 54]. One difficulty, which will
be addressed in the sequel, is the determination of the crack tip location with a good accuracy.

![Residual field for the SiC sample shown in Figure 1. It is expressed in difference of gray levels between the reference and corrected deformed images, and should be compared to the full dynamic range of the image (210 gray levels).](image)

**Figure 7.** Residual field for the SiC sample shown in Figure 1. It is expressed in difference of gray levels between the reference and corrected deformed images, and should be compared to the full dynamic range of the image (210 gray levels).

### 2.6. A priori performance evaluation

The success of DIC and its ability to resolve accurately displacements is to be traced back to the consistent collective information carried by a large number of pixels. This is formalized through the regularization that is meant to reduce drastically the number of unknown parameters to capture a given kinematics. The larger the zone over which the displacement gradient can be neglected (the larger the ZOI, or the element), the more accurate the determination of the translation. However, simultaneously, the poorer is the kinematic description, as the dimensionality of the space in which displacements are searched for decreases. Alternatively, if the zones are small, one may accurately capture complicated displacement fields, but the data to be exploited are much less per degree of freedom. Thus, one has to face a difficult compromise. There is no general answer to this issue, and depending on the mechanical test to be analyzed, one will have to adjust the zone or element size so that both spatial resolution and noise level will be judged satisfactory. What was mentioned here for ZOI sizes, equally applies to a finite element decomposition where the element size plays the role of the ZOI. More generally, the ratio of the number of degrees of freedom used to describe the kinematics to the total number of pixels, (e.g., $1/\ell^2$ for a collection of piecewise translations) is the important quantity that will determine the balance between spatial resolution and uncertainty.

It is therefore essential to carry out an *a priori* evaluation of the performance of the technique. One such test is to create a fake (synthetic) image that is translated from the reference one by a fixed displacement, $\Delta$. When this field belongs to the space of searched displacements no systematic errors are expected. However, one can try to blindly determine this displacement for different refinements of the discretization. Systematic errors exist (in particular because of interpolation techniques), but are
always much smaller than the uncertainty (in the sense of standard deviation of the displacement). The latter is the largest for a fake displacement of $\Delta$ that amounts to half a pixel in all space dimensions.

The change of the uncertainty versus element size is shown in Figure 8. The increase in accuracy with $\ell$ is easily seen. This analysis, which is easy to perform as it only requires a reference image, may help determining for the most appropriate element size. It is worth noting that the texture of the SiC sample (Figure 1) is not very easy to analyze, and this explains why rather high uncertainties are observed. Last, it is also worth remembering that a very simple gray level interpolation was used in the present analyses. In practice, a displacement field that does not belong to the space of searched functions will still show a similar behavior for small sizes $\ell$. However, for large $\ell$, the distance between the actual field and searched displacement form will no longer decrease and a systematic increase of the uncertainty will appear [33, 34, 36].

2.7. Noise sensitivity

It was mentioned that digital images are typically corrupted by noise, $\eta$, which often is considered as white. A good estimation of it is given by taking different pictures of the reference scene without any loading, and analyzing the raw difference. Hence, it is of interest to investigate the effect of noise on the resulting evaluation of $\omega_n$. Let us denote by $\delta\omega_n$ its variation with respect to a small amount of noise $\eta$. The expectation value of noise, denoted by $\langle \eta \rangle$, is 0. It is however characterized by its variance

$$\langle \eta^2 \rangle = \sigma^2$$

Figure 8. Standard uncertainty as a function of the element size $\ell$ for a finite element decomposition of the displacement field when $\Delta = 0.5$ pixel. A power-law decrease as $A^{1+\alpha}\ell^{-\alpha}$ is typically observed, with $A \approx 3$ pixels and $\alpha \approx 1.5$ for the reference picture shown in Figure 1. An element size $\ell = 32$ pixel (as used in the previous examples) leads to an uncertainty of the order of $8 \times 10^{-2}$ pixel.
Conventionally, one may consider that the reference image is noiseless, if the noise variance on the deformed image is doubled, $2\sigma^2$. By reverting to the linear system to be solved for $\omega$ (at the final step of the algorithm), it is observed that no systematic error (or bias) will result from the noise

$$\langle \delta \omega_n \rangle = 0$$

and the full correlation matrix of the error in $\omega$ reads

$$\langle \delta \omega_m \delta \omega_n \rangle = 2\sigma^2 M^{-1}_{mn}$$

The latter result is quite interesting, as it allows one to estimate the uncertainty level due to the image noise on the displacement directly from the inverse of the matrix $[M]$ to be computed for DIC. Moreover, it can be used to design robust determinations of derived quantities, such as SIFs for cracks (see Section 4.1).

When applied to Q4 elements, the previous result leads to the following estimate of the standard deviation $\sigma_u$ of the nodal displacements induced by a Gaussian white noise [30, 40]

$$\sigma_u = \frac{12\sqrt{2}\sigma p}{7(\langle |\nabla f|^2 \rangle)^{1/2}\ell}$$

where $p$ is the physical pixel size. For the picture of Figure 1, one computes $\langle |\nabla f|^2 \rangle^{1/2} \approx 17$ pixel$^{-1}$, hence $\sigma_u \approx 0.14\sigma p/\ell$, so that when $\ell = 32$ pixels and $\sigma = 1$ gray level, the estimate of $\sigma_u$ is about $5 \times 10^{-3}$ pixel to be compared with $8 \times 10^{-2}$ pixel given by the uncertainty analysis. Thus for such an example, the uncertainty analysis, which is by construction deprived from noise, leads to much higher levels of uncertainty than the noise. Note however that the scaling of these two uncertainties with $\ell$ is different. For a suited kinematic basis and large element sizes, the noise is expected to be dominant. This is a strong incentive to work with large elements, or more precisely with few kinematic degrees of freedom. The difficulty is thus to design a kinematic basis that is tailored to the description of the observed mechanical test.

2.8. Equilibrium gap penalization as a regularization

The above section has given one example of regularization scheme by restricting the searched displacement to a specific subspace. Other procedures can be imagined where the regularization is “softer.” Rather than excluding fields that would live outside a given subspace, one could accept them but give in turn a penalty to a global functional containing the objective function, Equation (5), plus an extra penalization term. This procedure may allow one to incorporate some a priori knowledge on the mechanics of the problem (be it related to the material or the test).

In particular, one may assume that (part of) the body behaves as a linear elastic solid. Let us assume here that we do not wish to exploit any specific situation for which some solutions would be known. One may thus resort to a finite element formulation of the elastic problem, and using the same mesh as that in the DIC analysis. Let us
simply recall that the elastic problem (application of the potential energy theorem) can be cast in a linear form

\[
[K]\{\omega\} = \{f\}
\]

(14)

where \([K]\) is the stiffness matrix, and \(\{f\}\) for internal nodes corresponds to body forces suitably integrated over the elements. Dirichlet (kinematic) boundary conditions may complement this problem prescribing \(\omega\) nodal values on the boundary.

Hence, the norm \(\| [K]\{\omega\} - \{f\} \|^2 \) (summed over all internal nodes) is a measure of the non-satisfaction of the momentum balance equation. It is therefore a good candidate for an extra penalty. This norm is the “equilibrium gap” [63], which opens the way to coupling DIC to identification techniques, but we will leave aside this topic, which is only mildly related to the topic of the present paper.

The combination of the two functionals, \(\Phi^2\), and \(\| [K]\{\omega\} - \{f\} \|^2\) with a respective weight \((\xi, 1 - \xi)\) and a suitable normalization by \(A\) and \(B\) such that \(\Xi(\xi = 0) = \Xi(\xi = 1) = 1\)

\[
\Xi^2 = \xi A\Phi^2 + (1 - \xi)B\| [K]\{\omega\} - \{f\} \|^2
\]

(15)

does fulfill the required goal to keep a reasonably large functional space for the displacements, and incorporate a penalty for deviations from a linear elastic behavior. It is interesting to observe that the problem still remains linear in \(\omega\), even if it is somewhat more computer resource demanding. The system to solve finally reads

\[
\left( [M] + \frac{(1 - \xi)B}{\xi A} [K]'[K] \right) \{\omega\} = \{d\} + \frac{(1 - \xi)B}{\xi A} [K]'\{f\}
\]

(16)

A spectral analysis shows that a specific value of the parameter \(\xi\) corresponds to (the fourth power of) a length scale below which the elastic penalty function will dominate and above which DIC will have a larger weight. Thus, although in a more hidden way, one may consider that this regularization is comparable in spirit to the previous analysis, if the cross-over scale is compared to the element or ZOI size, but now below this length scale the displacement field fulfills the constitutive equations of linear elasticity and the balance conditions. In that sense, the filter implemented here is mechanically-based.

The fundamentals for DIC, which apply to most mechanical problems, have been introduced. It is worth mentioning that even though all the developments were presented to analyze 2D images, they can be generalized to 3D images without any theoretical difficulty. Let us now specialize the discussion to analyses in which cracks are present.

3. Application to cracked bodies

We have seen that having a global formulation of the problem rather than small windows (or ZOIs) treated as independent, and prescribing a continuity of the displacement field was beneficial in terms of uncertainty for displacement measurement. However, one obvious limitation to this additional requirement is when the analyzed displacement field is discontinuous. In particular, cracks (but also localization bands) are not completely
suited to such an approach. A systematic error is unavoidable across the crack faces [36]. It is worth noting that this phenomenon also occurs for local DIC algorithms. It usually leads to spurious results for the ZOIs cut by the crack. In that case, there is no regularization available to correct for these artefacts. Different routes can be followed to use the approaches mentioned in the previous section and yet allow for a discontinuous displacement field.

3.1. eXtended Finite Element enrichment strategy

The first extension of the finite-element DIC for cracks concerns a specific enrichment that consists in introducing a discontinuity across the crack faces in addition to the regular finite-element description. This technique is widely used in computational mechanics, and is known under the name of X-FEM (for eXtended Finite Element Method [31, 32]). Its image correlation counter-part is referred to as X-DIC [33, 34, 36]. In addition to the displacement discontinuity, the enrichment may also contain stress and strain singular contributions to describe the displacement in the vicinity of the crack tip. The main advantage of this technique is that most of the analysis remains unchanged but only a local refinement is included.

The displacement is decomposed over the following basis (Figure 9)

\[
\mathbf{u}(\mathbf{x}) = \sum_{m \in \mathcal{N}, \alpha} \omega_{m\alpha}^{(1)} N_m(\mathbf{x}) e_\alpha \\
+ \sum_{m \in \mathcal{N}_{cut}, \alpha} \omega_{m\alpha}^{(2)} N_m(\mathbf{x}) H_m(\mathbf{x}) e_\alpha \\
+ \sum_{m \in \mathcal{N}_{tip}, \alpha} \omega_{m\alpha}^{(3)} N_m(\mathbf{x}) C_j(\mathbf{x}) e_\alpha
\]  

(17)

where \( H_m(\mathbf{x}) \) relative to node \( m \), of coordinate \( \mathbf{x}_m \) is obtained from the Heaviside function \( \mathcal{H} \) applied to an algebraic measure of the distance to the crack support. The latter is defined by a current point on the crack path \( \mathbf{x}_{\text{crack}} \), and the normal \( \mathbf{n} \) to the crack path

\[
H_m(\mathbf{x}) = \mathcal{H}((\mathbf{x} - \mathbf{x}_{\text{crack}}) \cdot \mathbf{n}) - \mathcal{H}((\mathbf{x}_m - \mathbf{x}_{\text{crack}}) \cdot \mathbf{n})
\]  

(18)

The \( C_j(\mathbf{x}) \) functions account for the crack tip singularity, expressed in terms of the polar coordinates \((r, \theta)\) with respect to the crack tip

\[
C(\mathbf{x}) = \begin{bmatrix}
\sqrt{r} \sin(\theta/2) \\
\sqrt{r} \cos(\theta/2) \\
\sqrt{r} \sin(\theta/2) \sin(\theta) \\
\sqrt{r} \cos(\theta/2) \sin(\theta)
\end{bmatrix}
\]  

(19)

The present formulation, proposed by Zi and Belytschko [64], simplifies the implementation.

To illustrate the results obtained with such an approach, Figure 10 shows the displacement field obtained for the case study of Figure 1. The discontinuous enrichment is activated without affecting the estimation of displacements away from the crack.
Figure 9. Schematic plot of the regular Q4 mesh, the nodes (shown as gray disks) supporting a discontinuous enrichment, and the nodes (shown as black squares) supporting a singular (crack-like) field.

Figure 10. Vertical (left) and horizontal (right) components of the displacement field for the SiC sample of Figure 1 as resolved with the eXtended Q4-Digital Image Correlation. Q4 elements are used with a size $\ell = 32$ pixels. The color scales for the displacements are expressed in pixels (1 pixel = $1.85 \mu m$).

Because genuine discontinuities are taken into account, one may resort to large element sizes, and hence benefit from a more accurate determination of the displacement (see Figure 8). This is the main benefit of this formulation, in the sense that the proposed library of displacement fields is more suited to the actual displacement.

However, it is to be mentioned that the present setting, directly transposed from the initial formulation proposed for numerical computations with cracks [31, 32], suffers from a rather poor conditioning. Many degrees of freedom are added (see Equation (17)), and hence the accuracy of the enriched degrees of freedom is poor. This is felt by considering the displacement discontinuity across the crack faces as shown in Figure 11.
Figure 11. Normal displacement discontinuity across the crack face as determined with the eXtended Q4-Digital Image Correlation (solid line: X-DIC) and using the elastic regularization (dashed line: XI-DIC). 1 pixel = 1.85 µm.

3.2. Equilibrium gap penalization

To mend this poor conditioning, different strategies may be followed. Either one may artificially reduce the number of enriched degrees of freedom, but the cost is that the numerical scheme may become more cumbersome to construct or less matched to the regular discretization. A second option is to keep the previous enrichment, but penalize the rapid variation of the displacement discontinuity. This solution corresponds to adding to the objective functional a second one that is a quadratic norm of derivatives of the displacement jump $[u]$ along the crack faces $\Gamma$, for instance

$$P = \int_{\Gamma} \left| \frac{\partial [u]}{\partial s} \right|^2 ds$$

(20)

where $s$ is a curvilinear abscissa along the crack path, is a simple choice to implement. However the solution expected for an elastic material is $[u] \propto s^{1/2}$, which will also be penalized by the above form.

A way to circumvent this difficulty is to choose a form that will not affect the expected elastic solution. This is exactly the purpose of the above-introduced equilibrium gap regularization. Doing so, one will not only regularize the enriched degrees of freedom but also the regular ones simultaneously. The fact that tractions vanish on the crack face is automatically included in the penalization. Figures 11 and 12 show the obtained results for the reference case of the SiC specimen. The displacement field was smoothed quite significantly, and little noise remains. It is also to be noted that the regularization does not affect the external boundary, as no information is available there to supplement image correlation. This example illustrates what was meant by a mechanically-based filter. In other words, it is an integrated approach in the sense that some $a$ priori information on the mechanical nature of the problem was added. It is therefore an eXtended and Integrated approach to DIC (or XI-DIC [41]). In this reference, the following SIF value $K_I = 3.02$ MPa.m$^{0.5}$ was estimated.
Figure 12. Vertical (left) and horizontal (right) component of the displacement field for the SiC sample of Figure 1 as resolved with the eXtended Q4-Digital Image Correlation and using the equilibrium gap penalization. Q4 elements are used with a size $\ell = 16$ pixels, and $\xi = 0.25$ (see Equation (15)), 1 pixel = 1.85 $\mu$m.

3.3. Direct use of closed-form solutions

Another option is to use directly in the DIC formulation, the exact expression of the displacement field, solution to the elastic problem of a cracked medium. In fact, this approach is much richer than it may appear at first glance.

The displacement field in an infinite solid containing a semi-infinite crack along the $x$-axis ($x < 0$) is written as a combination of a double infinite series. It is convenient to use polar coordinates $(r, \theta)$ for the current point, or its affix in the complex plane $z = re^{i\theta}$, where the origin is located at the crack tip, and to express the displacement field $\mathbf{u}$ as a complex function, $\mathbf{u} = u_x + iu_y$. The elementary solutions read

$$\phi_n^I(r, \theta) = ir^{n/2} \left( \kappa e^{in\theta/2} - \frac{n}{2} e^{i(4-n)\theta/2} + \frac{n}{2} + (-1)^n e^{-in\theta/2} \right)$$

$$\phi_n^{II}(r, \theta) = ir^{n/2} \left( \kappa e^{in\theta/2} + \frac{n}{2} e^{i(4-n)\theta/2} - \frac{n}{2} - (-1)^n e^{-in\theta/2} \right)$$

where $\kappa$ is Kolossov’s constant, namely, $\kappa = (3-\nu)/(1+\nu)$ for plane stress or $\kappa = (3-4\nu)$ for plane strain conditions, and $\nu$ Poisson’s ratio. The labels, mode I and II, are chosen on the basis of the following symmetries

$$\phi_n^I(z) = \phi_n^I(z)$$
$$\phi_n^{II}(z) = -\phi_n^{II}(z)$$

All these fields are homogeneous functions of the distance to the crack tip $r$. With this convention, all elastic displacement fields with finite strain energy densities are generated with $0 \leq n < \infty$. All odd indices label fields with a discontinuity across the crack faces, whereas even indices correspond to continuous fields. To mention just a few examples, $n = 0$ corresponds to translations either parallel (hence termed I) or perpendicular (labeled II) to the crack path, $n = 1$ are the classical mode I and mode II displacement fields whose amplitudes are proportional to the corresponding SIFs. The case $n = 2$ gives access to either the T-stress component for mode I, or
rigid body rotation about the crack tip for mode II (no discontinuity is involved here since 2 is even). Larger indices correspond to “subsingular” or higher order fields that may capture the remote heterogeneity of the loading, but do not affect the mechanical loading at the crack tip. This family of fields is thus the appropriate basis function to describe the displacement field for a traction free crack in an elastic solid.

Furthermore, negative indices are generally discarded as they would correspond to a diverging energy density at the crack tip. However, the linear elastic behavior is often an approximation in the immediate vicinity of the crack tip and hence, it is appropriate to exclude a small region around the crack tip where non-linearities (e.g., plasticity, damage, debonding) are expected to be predominant. If such a small disk centered on the crack tip is removed, its boundary supports an (unknown) distribution of forces that has to activate the above fields, including the negative index ones.

Since the displacement fields that correspond to a crack in an elastic solid are known, it is natural to use them as basis functions. One has to consider also rigid body motions, but as mentioned above, they are included in the decomposition, namely, $\phi_{0}^{j}$ correspond to a translation, whereas $\phi_{2}^{j}$ is a rotation about the crack tip. The advantage of this procedure is that only very few degrees of freedom are used to describe the kinematics, typically $\phi_{j}^{n}$ are used with $j = I$ or $II$ and $0 \leq n \leq 4$, or about 10 different fields, whereas a large number (of the order of $10^{6}$) of pixels are at hand. Hence small uncertainty levels are to be expected. Moreover, the amplitudes of $\phi_{j}^{1}$ directly give access to the sought SIFs in mode $j$.

Figure 13 shows the resulting displacement field for the SiC sample. Let us underline that this procedure can be seen as being along the same line as the regularized approach based on a finite-element penalization (XI-DIC [41]). However, the solution to the elastic problem is not performed numerically, but rather analytical solutions are used. In that sense, it is also an integrated approach (or I-DIC [45, 40]). Moreover, two of the generalized degrees of freedom are directly the sought quantities (here SIFs). Conversely, for XI-DIC, the measured displacement field needs to be post-processed to estimate SIFs.

Applying this procedure for different crack geometries (e.g. varying the crack tip position, or its orientation) and minimizing the global residual allows for the optimization of these geometrical parameters [40]. It is also possible to evaluate Poisson’s ratio using the same methodology. Interestingly enough, the fields have an affine structure in $\kappa$ (see Equations (21-22)) and hence splitting them into two contributions, one can directly estimate $\kappa$ (or Poisson’s ratio).

4. Evaluation of Stress Intensity Factors

Except for the integrated approach (in which measurement and identification steps are performed simultaneously), SIFs are not a direct output of the measurement stage. The displacement fields need to be post-processed to evaluate SIFs as accurately as possible since any measured displacement has an uncertainty level.
4.1. Interaction integral

There exist many ways to evaluate SIFs, in particular developed for their use in numerical simulations. One of the most popular is the interaction integral technique [65, 66], and was also shown to be applicable to measured displacement fields [43]. It consists in constructing the following domain (Ω) integral

\[ I = -\int_{\Omega} \left[ \sigma^a_{m,k}\delta_{kj} - (\sigma^a_{ij}u^{a}_{i,k}\sigma^a_{ij}u^{a}_{i,k}) \right] q_{k,j} \, dx \]  

(24)

where \( \mathbf{u} \) and \( \mathbf{\sigma} \) are the displacement and stress fields related through the elastic constitutive law, \( \mathbf{u}^a \) and \( \mathbf{\sigma}^a \) those of an auxiliary elastic problem with the same elastic properties, \( \delta \) Kronecker’s delta, and \( \mathbf{q} \) a vector field called the Virtual Crack Extension (VCE) field. The latter is equal to \( \mathbf{0} \) outside the domain \( \Omega \), it is parallel to the crack path, and is equal to the unit vector pointing along the propagation direction at the crack tip. No other constraints (apart from continuity and differentiability) are prescribed to the VCE field, which remains for a large part arbitrary. If the arbitrary problem is chosen to be a pure mode \( j = I \) (resp. \( II \)) crack solution, i.e., \( \mathbf{u}^a = 1/(2\sqrt{2\pi\mu})\phi^j_1 \) (where \( \mu \) is Lamé’s coefficient), the above interaction integral is equal to \( 2K_j(1 - \nu^2)/E \) (for plane strain).

It is important to note the linearity properties of the above interaction integral, whose discretized form reads

\[ K_j = A^j_{ik}\omega_i q_k \]  

(25)

where \( [A^j] \) is associated with the considered mode \( j = I \) or \( II \), \( \{\omega\} \) collects all measured degrees of freedom, and \( \{\mathbf{q}\} \) all nodal values of the VCE field. In Section 2.7, the sensitivity of the measured displacement field was analyzed with respect to image noise. A white noise \( \eta \) on the images leads to an uncertainty on the measured displacement, \( \delta \mathbf{u} \), of zero mean, and covariance that is directly related to matrix \( [M^{-1}] \) (see Equation (12)).
Consequently, the evaluation of SIFs is also affected by a “noise,” of zero mean, and variance
\[
\langle \delta K_j \delta K_j \rangle = \sigma^2 A^j_{ik} M^{-1}_{il} A^j_{lm} q_k q_m \\
\equiv \sigma^2 B^j_{km} q_k q_m
\]
where the latter writing (defining \([B^j])\) underlines the quadratic dependence of the \(\delta K\) covariance matrix on the VCE. Since the VCE field \(q\) is largely arbitrary, one may take advantage of the above form to minimize the variance of \(\delta K_j\) [44]. It consists in solving the linear system
\[
[B^j]\{q\} = 0
\]
under the above mentioned boundary conditions. It is to be stressed that the computation of the VCE is to be done for each image and integration domain \(\Omega\) (since matrix \([B^j]\) depends on matrix \([M^{-1}]\)). It provides the best determination of the SIF tailored to the image texture in such a way that the sensitivity to image noise is minimized.

Figure 14 shows the VCE field computed for the SiC image. In contrast to the common use that consists in choosing a priori a specific form (Gaussian, cone, truncated pyramid [67, 68]) and using a vector field parallel to the crack face, the above formulation yields a non trivial orientation and magnitude without any a priori choice [44]. The resulting uncertainty is much less than what would be obtained from such an a priori VCE field, as shown in Figure 15. Although quite close, the optimal VCE is different for mode I and II. Note that small domain sizes display deviations from the theoretical power-laws that are nothing but statistical fluctuations. They do not compromise the validity of the analysis.

Let us note that the interaction integral method may be rephrased in the following way, focusing on its linearity with respect to the measured displacement field
\[
K_j = e^j_i \omega_i
\]
where \(\{e^j\} = [A^j]\{q\}\) is an extractor (vector) field whose scalar product with the displacement field provides the SIF. The extractor field shows the spatial weight distribution assigned to each degree of freedom of the kinematic measurement. Figure 16 shows the corresponding extractor field for the same SiC case. Amazingly, most of the weight is carried by the crack face, so that even if the intrinsic formulation appears to be a domain integral, in fact, the actual measurement of the SIF is quite close to what would be done by studying only the crack opening discontinuity. Hence it appears to be much more fragile than what could have been anticipated. The mode I SIF value is estimated to be 2.7 ± 0.4 MPa.m\(^{1/2}\) with the conical VCE.

4.2. Other characteristic crack features

An important property of the decomposition (21-22) is that the fields \(\phi_n\) are sorted according to their rate of increase with the distance to the crack tip. Therefore large negative order fields vanish very quickly with \(r\), and can be omitted from the analysis. Similarly, large positive indices do not affect the crack tip, and are only useful for accounting for very inhomogeneous boundary conditions. Thus, only a small number of those fields are kept in the sequel. The displacement field is therefore searched as

\[
u(z) = \sum_{n,j} \omega^i_j \phi^n_j(z)
\] (29)

Another property of interest is the fact that the derivatives of these fields with respect to the crack tip position along the \(x\) axis, (equal to \(-\partial \phi^n_j/\partial x\)), belong to the same series

\[
-\frac{\partial \phi^n_j}{\partial x} = \frac{n}{2} \phi^{n-2}_j
\] (30)
for $j = I$ or $II$. From the latter equation, the following expressions are also obtained

$$\frac{\partial^m \phi_j^I}{\partial x^m} = -\frac{(2m-2)!}{2^{2m-1}(m-1)!} \phi_j^{1-2m}$$

(31)

Although the crack path can generally be distinguished using the previously described DIC algorithms, the precise location of the crack tip is much more difficult to identify (sometimes, as shown in the present paper, the crack is invisible, and this problem is extremely difficult). Let us assume that the crack was mis-positioned by a small distance $d$ along its axis. The new decomposition of the displacement field reads

$$u(z + d) = \sum_{n,j} \tilde{\omega}_j^n \phi_j^n(z + d)$$

(32)

When linearized about the exact crack position, Equation (32) yields the following inter-relationships

$$\omega_j^n = \tilde{\omega}_j^n - \frac{n + 2}{2} \tilde{\omega}_j^{n+2} d$$

(33)

In particular, a crack in an infinite medium such that $\omega_j^n = \delta_{n1} K_0$, will display a correction term $\tilde{\omega}_j^{-1} = dK_0/2$. It is therefore natural to interpret the $-1$-amplitude $\tilde{\omega}_j^{-1}$ as resulting from a mispositioning of the crack tip. The positioning of the crack tip is thus computed from

$$d = \frac{2\tilde{\omega}_j^{-1}}{\tilde{\omega}_j^1}$$

(34)

(so that $\omega_j^{-1} = 0$). This mis-positioning also induces a SIF variation $\Delta \omega_j^1 = \tilde{\omega}_j^1 - \omega_j^1$

$$\Delta \omega_j^{-1} = \frac{3\tilde{\omega}_j^{-1} \tilde{\omega}_j^3}{(\tilde{\omega}_j^1)^2}$$

(35)
Digital Image Correlation and Fracture

that depends on three parameters \( \tilde{\omega}^{-1}_j, \tilde{\omega}_j^1, \text{and } \tilde{\omega}_j^2 \). Equation (35) shows that at least three quantities need to be identified when the crack tip position is unknown.

These results are extremely useful. Even if the actual crack tip is known (or seen), the presence of confined non-linearities should prevent from assimilating the true crack tip with its extrapolation from the far field displacement in the elastic domain. Therefore, the above procedure is the most reliable way of defining the effective crack tip position, consistent with the traditional interpretation of the SIF (even in the presence of confined crack-tip non-linearities). In particular, the fields \( \phi^{-1}_j \) may be seen as the limit of the superposition of two crack fields with SIF respectively equal to \( K \) and \( -K \) and located at a distance \( d \) apart, as \( d \) tends to 0, and \( K \) diverges so that the product \( Kd \) remains constant. Thus this field is referred to as a “dipolar” crack field. This type of procedure will be used in Section 5 to determine SIF fields along the crack front, the latter being determined by canceling out the \( \omega^{-1}_j \) contributions associated with \( \phi^{-1}_j \) field.

The same interpretation also holds for lower orders. For instance, the fields \( \phi^{-3}_j \) are quadrupoles. Williams’ series terms of order \([1, -1, -3, \ldots]\) are nothing but a multipolar expansion of the displacement field. This leads to a physical interpretation of the amplitudes in the case of confined non-linearities. They can always be decomposed (in the far-field elastic domain) as a superposition of elementary cracks of SIF \( K_j(s) \) whose tip is positioned at \( s \). The amplitudes \( \omega^{-2n}_j \) are thus the moments of order \( n \) of the distribution \( K_j(s) \)

\[
\omega^{-2n}_j = \frac{(-1)^{n+1}(2n)!}{(2n-1)2^{2n}(n!)^2} \int s^n K_j(s) \, ds
\]  

(36)

It was shown that the actual crack tip position is defined such that \( \omega^{-1}_j = 0 \), or here at position \( s_0 \) such that

\[
\int K_j(s)(s-s_0) \, ds = 0
\]  

(37)

The next term in this expansion, when centered at \( s_0 \), is an intrinsic measurement of the region over which \( K \) has to be distributed in order to match the far-field elastic displacement. It is thus interpreted as (proportional to) the size squared \( r_{FPZ}^2 \) of the fracture process zone (FPZ), namely if \( r_{FPZ} \) is defined as

\[
r_{FPZ}^2 = \frac{\int K_j(s)(s-s_0)^2 \, ds}{\int K_j(s) \, ds}
\]  

(38)

then, the following estimate (assuming \( \omega^{-1}_j = 0 \)) is obtained

\[
r_{FPZ}^2 = -\frac{\omega^{-3}_j}{\omega^{-1}_j}
\]  

(39)

Higher order moments give further indications on the details of the process zone. However, they correspond to displacement fields that vanish as \( r^{-5/2} \) or faster, and hence they may be difficult to capture from image analyses.
4.3. Optimal extracting fields

We have just seen that the measurement of SIFs is nothing but the evaluation of one term in a series of parameters that are useful to characterize mechanically the crack, i.e., not only its loading (namely, the SIFs themselves), but also the location of its tip, and even the extent of the FPZ. It is thus tempting to develop a general methodology to estimate any term $\omega^n_j$. For this purpose it is possible to parallel the derivation of the interaction integral formalism. However, this pathway is not followed in the sequel, but rather a more general procedure.

From the interaction integral formalism, the fact that it is written in the form of an extractor field (see Equation (28)) is retained. Thus for any quantity of interest, $a$, one wishes to derive the extractor $e$ such that

$$ a = \int_{\Omega} e \cdot u \, dx $$

(40)

In practice $e$ is discretized in the same way as the displacement so that, $e = \{e_i\}$, $u = \{u_i\}$, $a = \{e\}^t\{u\}$. The quantity of interest $a$ will be any of the amplitudes $\omega^n_j$. The extractor should provide the “best” estimate for $a$. By best, it is meant that $e$ should:

- provide a faithful measurement of $a$,
- not be affected by variations of other parameters than the one being searched for,
- be robust to noise (i.e., measurement uncertainties).

The first two requirements are rewritten as a set of orthogonality (and normalization) conditions. If $e$ is relative to the parameter $\omega^n_j$, then for all $k$ and $m$

$$ \{e\}^t\{\Phi^m_k\} = \delta_{jk}\delta_{mn} $$

(41)

where $\{\Phi^m_k\}$ represents the discretized field $\phi^m_k$. The treatment of noise robustness is similar to the previous analysis and leads to the minimization of the sensitivity

$$ S = \{e\}^t[M^{-1}]\{e\} $$

(42)

namely, a linear problem under constraints that is solved by using Lagrange multipliers.

This formulation is very convenient as no constraint exists on the domain over which $\{e\}$ is defined. Thus one might exclude arbitrary regions where the DIC analysis is not possible (e.g., a region being masked by some obstacle, a spurious light reflection). This is to be contrasted with the interaction integral that should contain the crack tip and faces. Moreover, as in the previous example, the extractor field is tailored to the image texture (i.e., matrix $[M^{-1}]$ appears again), it exploits at best its strength and limits its deficiencies.

Figure 17 shows the values of the mode I SIF as estimated with the different techniques presented so far, as a function of the domain size. We note a good convergence of the different values around 2.9-3.0 MPa$\sqrt{m}$, with a much more stable determination for the optimal and least squares extraction than for the interaction integral. These values are in very good agreement with that obtained from the analysis
of COD measurements (2.9 MPa√m [13]). It is worth noting that the least squares extraction technique, although seemingly straightforward and less sophisticated than the interaction integral formalism, is actually very close to the optimal extraction technique. The latter is shown here to have much poorer performances presumably because of the very concentrated weight given to the crack face displacements.

5. Three dimensional analyses

For 3D pictures (e.g., obtained by XCMT or magnetic resonance imaging), the same type of local [69, 70, 71] or global procedures may be implemented. For instance, 8-noded elements with tri-linear shape functions are considered (it is then referred to as C8-DIC [52] or XC8-DIC [54] when an enriched kinematics is implemented).

This case is illustrated with the example of a fatigue crack growing in a cast iron specimen [53, 54, 55]. In the present case, cast-iron samples were selected to provide adequate (random) markers. The graphite nodules provide a random texture in the images (Figure 18). A testing machine inside the tomograph [72] allows for taking 3D pictures of the specimen during a fatigue tensile test. The crack is visible on the deformed picture of the specimen, see Figure 18. The 6 mm long specimen has an almost square cross-section, 1.6 mm wide, 1.55 mm thick. The analyzed region is slightly smaller than the actual sample size since specimen edges were kept aside from the DIC analysis.

Albeit the crack can be seen in the deformed image at maximum load, its actual geometry is difficult to obtain from classical segmentation techniques of image processing. The reason is that the graphite nodules have a low attenuation level, and hence can hardly be distinguished from the crack. However, even without having an
Figure 18. Reference (load 5 N) and deformed (load 100 N) 3D images of a cracked sample. The load is applied along the longitudinal direction (normal to the crack). The voxel size is 3.5 µm.

*a priori* definition of the crack surface, one can perform a global DIC analysis using C8 finite-element shape functions (Figure 19). By itself, this treatment provides a reasonable determination of the displacement where the crack discontinuity can be seen. However, the displacement in elements traversed by the crack is smeared because of the continuous kinematics imposed by the shape functions. The 3D residual map is computed, and it does show large values only across the crack surface. Thus rather than segmenting the deformed image itself, one segments the residual map where most of the graphite nodules have been erased. For this purpose, a specific procedure was developed to identify a single surface with no nodules that supports the crack, but also extends further [55]. What remains to be done is to define the crack front. A first guess is obtained from a visual inspection of the residual or deformed image.

Let us note that the crack geometry is non-symmetric. The reason for this is that the specimen has been cut out from a larger one, three times as wide, where a fatigue pre-crack was grown. The sample shown here was cut out from the edge of this first sample, and hence the deviation of the crack front from a straight line is due to plasticity developing more close to the free side of the specimen. The effect of the non-uniform plastic process zone was one of the motivation for the study. In addition, the random microstructure of the graphite nodules also introduces some heterogeneity. We refer the interested reader to Ref. [53] for a more detailed discussion of the study, and we rather focus here on the DIC analysis. Last, the dissymmetry of the crack, and the severe constraints of the fatigue testing machine designed to fit in a tomograph, resulted in a loading that was not pure uniaxial tension. A small (and unanticipated) bending moment could be detected from the analysis of the kinematics away from the crack. This detail actually illustrates one of the major advantage of DIC, namely, a mechanical test in reality may differ from the *a priori* imagined loading state. Using DIC is not only a
way to check the actual loading state, but it also provides actual kinematic boundary conditions to use for comparison purposes with a numerical modeling. It therefore offers a greater tolerance to non-ideal testing conditions.

![Figure 19](image1.png)

**Figure 19.** Displacement component (expressed in voxels) along the loading direction obtained from Digital Image Correlation using C8 elements whose size is \( \ell = 32 \) voxels (left). Corresponding residual field (right). Its level should be compared to the dynamic range of the pictures (i.e., 256 gray levels). The voxel size is 3.5 \( \mu \text{m} \).

With this first approximation of the surface of discontinuity, an enriched X3D-DIC procedure is designed and run to obtain a more reliable description of the displacement map as shown in Figure 20.

![Figure 20](image2.png)

**Figure 20.** Geometry of the crack surface (left). Displacement component (expressed in voxels) along the loading direction obtained from Digital Image Correlation using an enriched kinematics (right). The volume shown here lies strictly inside the sample to avoid edge effects. The voxel size is 3.5 \( \mu \text{m} \).

If needed, the resulting residuals are again processed to get a better determination of the crack surface. The weak point of the procedure is that the crack front has to be defined by visual inspection. However, this problem was tackled in two-dimensions in
the preceding section using a specific extractor (i.e., the $\phi_I^{-1}$ field). The transposition to three dimensions is straightforward. Each cut of the sample by a plane whose normal is parallel to the mean crack front (except if the plane gets close to the specimen edge) is expected to behave as a 2D plane strain problem for in-plane displacement and as an anti-plane problem for mode III. Therefore the application of an extraction strategy allows one to estimate both the SIF and the best determination of the crack front location (by canceling out the $\phi_I^{-1}$ contributions, see Section 4.2). An example of such a procedure leads to the results shown in Figure 21. In the present case, the levels of mode I SIFs are significantly larger than those of modes II and III. The fact that mode II and III SIFs do not vanish is related to the loading conditions, which were not under pure mode I. Furthermore, it is also shown that a visual estimation of the crack tip location is slightly under-estimated when compared with the extraction technique.

To confirm the validity of the present procedure, a full three dimensional linear elastic finite element simulation was carried out on the actual geometry of the specimen and crack. Kinematic boundary conditions extracted from the DIC analysis were also prescribed on the top and bottom sides. Stress intensity factors for all three modes were computed all along the front. In Figure 21, these estimates are reported as curves while the measured estimates are shown as symbols for two load levels. An excellent general agreement is obtained between those estimates, for all modes, and a slight discrepancy for mode I under the highest load presumably because of a rather large plastic process zone developing over the remaining ligament (based on the elastic simulation, half of the ligament area exceeds the yield stress).

A number of developments are yet to be performed for refining these tools, and in particular a totally automated procedure would be welcome to extract SIFs. However, it is important to note that now, a full three-dimensional characterization of cracks was carried out, and a series of tools is operational and was validated on a test case.

Among the limitations that have been identified so far, from the experimental point of view, 3D-DIC techniques are restricted to the class of materials whose microstructure is contrasted enough so that it can be imaged with a tomograph. Contrary to two dimensions where a speckle pattern can be created on the surface of the examined specimen, this option is not possible in three dimensions, and passive tracers are essential for DIC. This is by far the most limiting feature. Other limitations, apart from those of the tomographic technique itself (such as imaging artefacts in the reconstruction process, or beam hardening effects), comes from the quantitative use of imaging. Most of the use of tomography is relative to single 3D images and stability of the set-up is not a real concern. However, for DIC applications, it is essential that the two images to be registered are directly comparable. For instance due to thermal expansion, the distance between the beam source and the specimen may vary by tiny amounts. Yet, the divergent beam geometry used for laboratory tomographs is such that this small distance change may result in a significant image dilation (strains easily of the order of a few $10^{-3}$ are observed). This apparent strain should not be considered as a real one. Even though consequences of this artefact are limited in some cases (e.g., when
Figure 21. Profiles of mode I, II and III SIFs along the crack front as obtained from 2D cuts through the sample using a least-squares extraction technique (left) for two load levels (top left: medium load, top right: highest load). Determination of the crack front position from the supersingular amplitude extractor (thick white curve) as compared to the (smoothed) hand determination (thin white line) (bottom). The two fronts are drawn over the residual field supported by the crack surface. The crack surface lies above the crack front.

large strain levels are observed [73]), a special attention should be paid to the set up architecture to minimize this effect.

On the theoretical side, for the post-processing of the displacement field, a specific decomposition of the SIFs along the crack front on a given basis should be developed, following standard practice in the field of computational mechanics. The issue of a better handling of the vicinity of the free edge is a theoretical problem that has been long standing, still awaiting for a satisfactory treatment. The recourse to a plane stress elasticity problem is not quite legitimate. However it becomes critical here as resorting to a “sane” mechanical solution is a key to the quality of the SIF determination. However
this problem is much more general than the specific technical application addressed here, and in the absence of a better solution, at least the proposed description has the merit of sticking to the standard use of SIFs for both experimental and numerical analyses.

6. Summary and outlook

Digital Image Correlation is a powerful and cost effective technique to have access to full field kinematic measurements. Moreover, it can be tailored to bridge an essential gap between experiment and numerical modeling. If this statement holds true for most applications of experimental solid mechanics, it is even more appropriate for discussing fracture mechanics. Through the examples mentioned in this review, it is apparent that the considerable body of knowledge and experience accumulated over the years in computational fracture mechanics can readily be applied to the analysis of experiments. Relying on analytical or numerical elastic modeling is extremely beneficial in lowering uncertainties and the detrimental effect of noise inherent to any measurement technique. Moreover, the experimental/computational interaction is not one-way. Having to cope with a natural noise level much greater than typical uncertainties met in computation can help elaborate new noise-robust strategies that may be valuable in computational mechanics as well.

Strikingly, even “invisible” (with bare eyes) cracks can be “seen,” they can be located, the displacement quantified, the crack tip positioned and SIFs measured with a good accuracy. For instance, as shown herein, this is true for 2D surface measurements when crack opening displacements are less than the pixel size, and for 3D opaque materials when imaged by, say, XCMT. Although most approaches yielded essentially surface measurements, a very recent application was carried out to analyze three dimensional cracks based on tomographic imaging where the complete profile of SIFs for all modes all along the crack front was determined even with lab scale tomographs. This opens the way for reliable identifications and/or validations of 3D crack propagation models.

In the examples discussed herein, only SIFs were estimated. However, other quantities may be identified. For instance, cohesive laws are identified by using measured displacement fields [74, 75], or enriched kinematics since they can be incorporated in the measurement basis. The analysis and modeling of fracture process zones is also made possible, especially when low or very low scale imaging techniques are considered. This is of particular interest in the analysis of composite materials, concrete, metals and steels, polymers and elastomers. More generally, the coupling with damage models should be made possible with the development of identification techniques that account for the gradual degradation of material properties [76, 63, 77].

DIC may also be used to drive an experiment [78]. Because of heavy image processing and relatively lengthy image storage, the overall working frequency is about 1 Hz. This low cycle time requires a two-loop-cascade control scheme (an inner quick (wired) displacement-controlled loop is driven by a slow outer DIC-controlled loop).
Consequently, only quasi-static tests were performed up to now. It is envisioned that a real-time system is feasible when the pictures are not stored and computations optimized. The same type of procedure may be applied to other types of displacement fields (e.g., in the presence of cracks [79]). Recent software developments allow such DIC computations for 1 Mpixel-images to be performed mostly on Graphical Processing Units of PCs in 0.05 s [80] instead of 50 s when implemented in Matlab [81]. This opens the way for much faster and more complex control strategies.

Acknowledgments

The support of CETIM Foundation for the project PROPAVANFIS (“Advanced methods for the experimental and numerical analyses of crack propagations under complex loadings”) is gratefully acknowledged. The authors also acknowledge useful discussions with the members of the PROPAVANFIS group, namely, Drs. N. Limodin, W. Ludwig, J. Rannou, J.-P. Tinnes, and Profs. M.-C. Baëetto-Dubourg, J.-Y. Buffière, A. Combescure, and A. Gravouil.
References


[65] D. M. Parks, A stiffness derivative finite element technique for determination of crack tip stress


