On image gradients in digital image correlation

INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2016; **105**:243–260
Published online 14 July 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nme.4971

On image gradients in digital image correlation

J. Neggers[1], B. Blaysat[2,*,†], J. P. M. Hoefnagels[1] and M. G. D. Geers[1]

[1]*Department of Mechanical Engineering, Eindhoven University of Technology, Den Dolech 2, Eindhoven 5612 AZ, The Netherlands*
[2]*Clermont Université, Université Blaise Pascal, Institut Pascal, Clermont-Ferrand BP 10448, 63000, France*

SUMMARY

In digital image correlation (DIC), the unknown displacement field is typically identified by minimizing the linearized form of the brightness conservation equation, while the minimization scheme also involves a linearization, yielding a two-step linearization with four implicit assumptions. These assumptions become apparent by minimizing the non-linear brightness conservation equation in a consistent mathematical setting, yielding a one-step linearization allowing a thorough study of the DIC tangent operator. Through this analysis, eight different image gradient operators are defined, and the impact of these alternative image gradients on the accuracy, efficiency, and initial guess robustness is discussed on the basis of a number of academic examples and representative test cases. The main conclusion is that for most cases, the image gradient most common in literature is recommended, except for cases with: (1) large rotations; (2) initial guess instabilities; and (3) costly iterations due to other reasons (e.g., integrated DIC), where a large deformation corrected mixed gradient is recommended instead. Copyright © 2015 John Wiley & Sons, Ltd.

Received 19 September 2014; Revised 13 May 2015; Accepted 8 June 2015

KEY WORDS: digital image correlation; strain measurement; displacement fields; brightness conservation

1. INTRODUCTION

Full-field displacement measurements are more and more used in experimental mechanics. Affordable digital cameras allow for the acquisition of images during testing, and enable, through smart postprocessing the recovery of displacement and strain maps. Such maps are of great interest in experimental mechanics, because they allow to characterize models, as well as validate finite element (FE) simulations [1–3]. For this purpose, digital image correlation (DIC) is one of the most used image processing techniques. The origins of most DIC methods can be traced back to the Lucas–Kanade algorithm [4, 5] but were first successfully applied to experimental mechanics applications by Sutton *et al.* [6–8]. In DIC, a visible surface pattern imaged with a camera during sample deformation, where an image captured at the reference state, is compared with an image captured at the deformed state. Originally, the displacement field was obtained by maximizing the cross-correlation product of the two images, which can efficiently be carried out in the Fourier space [6–10]. However, to allow non-constant deformations (warps) within the considered zones of interest (ZOI), the cross-correlation product is often interchanged with an iterative optimization routine such as a Gauss–Newton (or Levenberg–Marquardt) algorithm [11–15]. The common approach found in literature is to first linearize the Lucas–Kanade algorithm (also known as the brightness conservation or optical flow equation) and second apply the iterative optimization algorithm, which results in a two-step linearization. The second linearization is in the iterative

*Correspondence to: B. Blaysat, Clermont Université, Université Blaise Pascal, Institut Pascal, BP 10448, 63000 Clermont-Ferrand, France.
†E-mail: benoit.blaysat@univ-bpclermont.fr

algorithm, which requires the linearization of the cost function with respect to the degrees of freedom (DOF). Effectively, such a DIC algorithm is a modified Newton method, because the applied tangent operator is not consistent with the one obtained from applying Newton's method directly to the non-linear brightness conservation. The main difference between the modified and the consistent tangent operators is expressed in the definition of the image gradient, which can be defined in several ways depending on the chosen linearization path.

The goal of this article is to show the origin of the various image gradients by a careful linearization of the brightness conservation and to evaluate the influence of these image gradients on the performance of DIC methods. To achieve this, the linearized system of equations is derived departing from the original non-linear brightness conservation, by directly linearizing with respect to the DOF, resulting in a consistent Gauss–Newton DIC minimization scheme. The obtained general system of equations will be reduced thereafter, to recover the form commonly used in DIC literature, thereby unraveling the underlying assumptions.

In a typical DIC application, three different coordinate spaces can be identified, namely: (1) the coordinates of the material points in the undeformed or reference state; (2) the coordinates of the material points in the deformed state; and (3) the coordinates given by the camera frame (the pixel coordinates). In some cases, some of these coordinate frames are identical, and for the case of small displacements, it is possible to use one coordinate space to approximate all three. In the history of DIC, the small displacement approximation was usually applicable, and a proper inversion of the deformation mapping function was not required. However, not rigorously defining these coordinate spaces has led to some confusion: (1) in papers where it is not clear which coordinates are used for which equation; (2) in papers where the coordinates are erroneously interchanged without proper transformation; and (3) in almost all cases where the impact assuming small displacements is not discussed. The interested reader is referred to the work of Tong where this is discussed thoroughly [16]. If a coordinate transformation affects the unknown displacement field, which is almost always the case in DIC, then it will impact the gradients applied to fields expressed in these coordinates, for example, the image gradient. Many mathematical styles encountered in DIC and image registration literature do not explicitly write the coordinate space and thus hide the critical discussion point, which is the topic of this article. To guide the discussion, the next section recalls the derivation of a DIC algorithm and deliberately uses a verbose mathematical style, which is chosen to emphasize the coordinate frame underlying the different fields. After deriving the DIC algorithm and defining the different image gradients, these image gradients will be applied to two academic examples and a relevant example taken from the literature to test the performance of the various image gradients in terms of accuracy, efficiency, and initial guess robustness.

2. DIGITAL IMAGE CORRELATION ALGORITHM DERIVATION

The derivation of most DIC algorithms departs from the assumed conservation of brightness (or optical flow), which states that the changes in the observed pattern, between images f and g, are due to the movement of material points (i.e., displacement $\underline{U}$), while the brightness of the material points is considered constant

$$f(\underline{x}) \approx g \circ \underline{\Phi}(\underline{x}), \tag{1}$$

$$\underline{\Phi}(\underline{x}) = \underline{x} + \underline{U}(\underline{x}), \tag{2}$$

where f and g are scalar fields of the imaged light intensity recorded by the camera sensor and $\underline{\Phi}(\underline{x})$ is a vector function that maps the reference coordinate $\underline{x}$ to the deformed coordinate. Throughout this article, the coordinate $\underline{x}$ refers to the (Lagrangian) reference coordinates, while the deformed coordinates are always expressed using the mapping function. In DIC, any set of material points in the reference state can be selected, but typically, the points are chosen that coincide with the pixel locations in the first image. It is important to note that the pixel locations are fixed in space and thus represent a Eulerian coordinate space, which only coincides with the reference coordinate space in

DOI: 10.1002/nme

Table I. Nomenclature.

$f(\underline{x})$: reference image	$g(\underline{x})$: current image
$\underline{U}(\underline{x})$: true displacement field	$\underline{u}(\underline{x}, \mathbb{a})$: approximated displacement field
$\underline{\Phi}(\underline{x})$: true mapping	$\underline{\phi}(\underline{x}, \mathbb{a})$: approximated mapping
$\tilde{g} = g \circ \underline{\phi}$: back-deformed image	$r(\underline{x}, \mathbb{a}) = f(\underline{x}) - \tilde{g}(\underline{x}, \mathbb{a})$: residual map
Ω: region of interest	$\underline{\underline{F}} = \underline{\text{grad}}(\underline{\phi})^T$: deformation gradient tensor
$\mathbb{a}$: column of degrees of freedom	$\Psi(\mathbb{a})$: $\mathcal{L}_2(\Omega)$ norm of the residual

the first image. Additionally, throughout this article, the composition operator $\circ$ is used to express the composition of two functions. This style allows a more precise definition of the gradient, which are throughout the article defined as the gradient with respect of the local coordinate $\underline{x}$, that is, $\underline{\text{grad}}(\cdot) = \sum_i (\partial \cdot / \partial x_i) \underline{e}_i$. For example, with this writing style, it is possible to distinguish between

$$\underline{\text{grad}}\,[g \circ \underline{\Phi}] \neq \underline{\text{grad}}(g) \circ \underline{\Phi}, \tag{3}$$

where $\underline{\text{grad}}\,[g \circ \underline{\Phi}]$ is the gradient of the image g after it has been deformed using $\underline{\Phi}$ and $\underline{\text{grad}}(g) \circ \underline{\Phi}$ is the gradient of the image g evaluated at the deformed coordinates. Note that notation are summarized in Table I.

Identifying the displacement field that will satisfy Equation (1) is an ill-posed problem, which deteriorates through the always present additional noise field. Therefore, DIC methods approximate the true displacement field with a field represented by a finite number of unknowns as follows:

$$\underline{U}(\underline{x}) \approx \underline{u}(\underline{x}, \mathbb{a}), \tag{4}$$

$$\text{and} \quad \underline{\Phi}(\underline{x}) \approx \underline{\phi}(\underline{x}, \mathbb{a}), \tag{5}$$

where $\mathbb{a}$ is a column of DOF, that is, $\mathbb{a} = [a_1, a_2, \ldots, a_n]^T$. Applying more pixels per DOF allows for attenuation of acquisition noise (e.g., [17]), provided that the discretized displacement field can adequately describe the true displacement field. Accordingly, Equation (1) becomes dependent on the DOF

$$f(\underline{x}) \approx g \circ \underline{\phi}(\underline{x}, \mathbb{a}). \tag{6}$$

2.1. Two-step linearization

First, the two-step linearization, which is common in DIC literature, is revisited in order to carefully detail the steps involved, enabling a detailed comparison with the consistent system of equations derived later. First, Equation (6) is written in a linearized form as follows [4–21]:

$$f(\underline{x}) \approx g(\underline{x}) + \underline{\text{grad}}(g)(\underline{x}) \cdot \underline{u}(\underline{x}, \mathbb{a}). \tag{7}$$

Because of the high non-linearity of the pattern, this linearization is only valid for a limited range (i.e., the gray value spatial fluctuation is strongly non-linear). This is commonly managed by iteratively updating the deformed image, bringing it closer to the reference image. The iteratively updated back-deformed image $\tilde{g}$, for the iteration k, is then defined as a function of the unknown DOF $\mathbb{a}^k$,

$$\tilde{g}^k(\underline{x}, \mathbb{a}^k) = g \circ \underline{\phi}(\underline{x}, \mathbb{a}^k), \tag{8}$$

which is the image g in terms of the reference coordinates $\underline{x}$, with pixel values interpolated at the estimated deformed coordinates $\underline{\phi}$. This back-deformed image is not a real image, it is merely

DOI: 10.1002/nme

a collection of gray values at the current estimation of the deformed coordinates, but for ease of implementation, it is usually stored as an image. Using this back-deformed image, the linearized brightness conservation is written as follows:

$$f(\underline{x}) \approx \tilde{g}^k(\underline{x}) + \underline{\mathrm{grad}}(\tilde{g}^k)(\underline{x}, \mathbb{a}^k) \cdot \delta\underline{u}(\underline{x}, \mathbb{a}^k), \tag{9}$$

$$\approx \tilde{g}^k(\underline{x}) + \underline{\mathrm{grad}}(f)(\underline{x}) \cdot \delta\underline{u}(\underline{x}, \mathbb{a}^k), \tag{10}$$

where the displacement field is now an incremental displacement $\delta\underline{u}(\underline{x}, \mathbb{a})$, consistent with the iterative approach, solving for the DOF $\mathbb{a}$. When close to the solution, $\underline{\mathrm{grad}}(\tilde{g}^k)$ resembles $\underline{\mathrm{grad}}(f)$, which is usually preferred because of its cheaper computational cost [12, 18, 19].

Pioneering DIC implementations used the cross-correlation product to determine the rigid body translations of small ZOI [6–10]. These implementations are rather cost effectively solved in the frequency domain but require that the pixels in the ZOIs move uniformly. To accommodate non-constant deformation fields (i.e., strain) within a ZOI [11–15] or even within the entire region of interest (ROI, denoted Ω) [17–21], it has become common practice to solve for $\mathbb{a}$ by minimizing a cost function iteratively. The benefit is that the displacement field discretization can be chosen tailored to the kinematics of the experiment, at the cost that the satisfaction of Equation (7) now is a coupled problem. For both the ZOI and the ROI implementations, the goal is to find the optimal approximate solution of Equation (7), namely $\underline{u}^{\mathrm{opt}}$. Accordingly, a minimization problem is formulated as follows:

$$\underline{u}^{\mathrm{opt}}(\underline{x}) = \underline{u}(\underline{x}, \mathbb{a}^{\mathrm{opt}}) \text{ with } \mathbb{a}^{\mathrm{opt}} = \underset{\mathbb{a}}{\mathrm{Argmin}}\,\{\Psi(\mathbb{a})\}, \tag{11}$$

where the cost function Ψ is defined as the $\mathcal{L}_2(\Omega)$ norm of the residual of the linearized brightness conservation (Equation (7)),

$$\Psi(\mathbb{a}) = \frac{1}{2}\int_{\Omega}\left[f - \tilde{g}^k - \underline{\mathrm{grad}}(f)\cdot\delta\underline{u}\right]^2 \mathrm{d}\underline{x}. \tag{12}$$

Minimization of Ψ implies

$$\forall i \in [1,n], \quad \Gamma_i(\mathbb{a}^{\mathrm{opt}}) = \frac{\partial\Psi}{\partial a_i}\left(\mathbb{a}^{\mathrm{opt}}\right) = 0. \tag{13}$$

The resulting non-linear system is linearized and solved using a Gauss–Newton algorithm. At each iterative step, the DOF $\mathbb{a}^{k+1}$ are decomposed in an initial guess $\mathbb{a}^k$ and an iterative update $\delta\mathbb{a}$ as follows:

$$\mathbb{a}^{k+1} = \mathbb{a}^k + \delta\mathbb{a}, \tag{14}$$

$$\forall i \in [1,n], \quad \Gamma_i\left(\mathbb{a}^{k+1}\right) = 0 \Rightarrow \Gamma_i\left(\mathbb{a}^k\right) + \frac{\partial\Gamma_i}{\partial\mathbb{a}}\left(\mathbb{a}^k\right)\delta\mathbb{a} = 0. \tag{15}$$

Equation (15) can be written in matrix form as follows:

$$\mathbb{M}\cdot\delta\mathbb{a} = \mathbb{b} \qquad \text{with} \begin{cases} \mathbb{M} = \begin{bmatrix} M_{11} & \dots & M_{1j} & \dots & M_{1n} \\ \vdots & & \vdots & & \vdots \\ M_{i1} & \dots & M_{ij} & \dots & M_{in} \\ \vdots & & \vdots & & \vdots \\ M_{n1} & \dots & M_{nj} & \dots & M_{nn} \end{bmatrix}, \text{and } M_{ij} = \dfrac{\partial\Gamma_i}{\partial a_j}\left(\mathbb{a}^k\right), \\ \mathbb{b} = [b_1, b_2, \dots, b_n]^T, \text{ and } b_i = -\Gamma_i\left(\mathbb{a}^k\right) \end{cases} \tag{16}$$

DOI: 10.1002/nme

thereby arriving at the conventional two-step linearized system of equations:

$$M_{ij} = -\int_{\Omega} \underline{\varphi}_i \cdot \underline{\mathrm{grad}}(f)\ \underline{\mathrm{grad}}(f) \cdot \underline{\varphi}_j \,\mathrm{d}\underline{x}, \tag{17}$$

$$b_i = -\int_{\Omega} \underline{\varphi}_i \cdot \underline{\mathrm{grad}}(f) \left[\underbrace{f - \tilde{g}^k}_{= r\left(\underline{x}, \mathbb{a}^k\right)} \right] \mathrm{d}\underline{x}, \tag{18}$$

where $\underline{\varphi}_i$ is the linearization of $\underline{\phi}$, that is,

$$\forall i \in [1, n],\ \underline{\varphi}_i = \frac{\partial \underline{u}}{\partial a_i^k} = \frac{\partial \underline{\phi}}{\partial a_i^k}. \tag{19}$$

The fields $\underline{\varphi}_i$ are usually referred to as the basis functions. The basis functions are not required to be linearly independent, but if chosen as such, then the approximated displacement field can be written in the familiar form found in literature, for example, [17–21],

$$\underline{u}\left(\underline{x}, \mathbb{a}^k\right) = \sum_{i=1}^{n} a_i^k \underline{\varphi}_i(\underline{x}). \tag{20}$$

This type of algorithm can be considered a forward additive Lucas–Kanade algorithm [4]. Other types of algorithms exist; for instance, the work of Baker and Matthews [5] adds three more, and the work of Tong [16] discusses even more and gives the correct definition of the coordinate systems for all discussed algorithms. However, for this article, the discussion is restricted to the forward additive types of algorithms. These are common in DIC applications because the resulting displacement field is generally given relative to the reference coordinates common in solid mechanics. Additionally, many of the proposed algorithms require an intermediate coordinate system, which requires an extra interpolation step, which further degrades the displacement accuracy. Nevertheless, the concept of performing a single linearization by applying Newton's method directly to the non-linear form of the algorithm applies to all. Consequently, the conclusions made as a result of analyses of the forward additive algorithm may also apply to other algorithms.

2.2. *One-step linearization*

This section discusses the equations starting directly from the non-linear form of the brightness conservation equation (Equation (6)). The result will be compared with the system of equations defined previously (Equations (17) and (18)). Obtaining the solution of the non-linearized brightness conservation starts with formulating a consistent cost function:

$$\Psi\left(\mathbb{a}^k\right) = \frac{1}{2}\int_{\Omega} \left[f\left(\underline{x}\right) - g \circ \underline{\phi}\left(\underline{x}, \mathbb{a}^k\right) \right]^2 \mathrm{d}\underline{x}. \tag{21}$$

Analogous to the two-step linearization, the iterative Gauss–Newton algorithm is applied (Equations (13)–(16)). Reusing the definition of $\underline{\Gamma}$ (Equation (13)), $\mathbb{b}$ becomes, $\forall i \in [1, n]$,

$$b_i = -\int_{\Omega} \frac{\partial \left[g \circ \underline{\phi} \right]}{\partial a_i}\left(\underline{x}, \mathbb{a}^k\right) r\left(\underline{x}, \mathbb{a}^k\right) \mathrm{d}\underline{x}, \tag{22}$$

where, from now on, $\mathbb{a}^k$ is simplified as $\mathbb{a}$. Applying the chain rule on the first term in the integral yields

$$\forall i \in [1, n],\ \frac{\partial \left[g \circ \underline{\phi} \right]}{\partial a_i} = \underbrace{\frac{\partial \underline{\phi}}{\partial a_i}}_{= \underline{\varphi}_i} \cdot \underbrace{\underline{\mathrm{grad}}(g) \circ \underline{\phi}}_{= \underline{G}}. \tag{23}$$

DOI: 10.1002/nme

Note that this is the true definition of the image gradient $\underline{G}$, which can be approximated in various ways (as shown next). Finally, $\mathbb{b}$ is expressed as follows:

$$\forall i \in [1,n],\ b_i = -\int_\Omega \underline{\varphi}_i(\underline{x},\mathbb{a}) \cdot \underline{G}(\underline{x},\mathbb{a})\, r(\underline{x},\mathbb{a}) \mathrm{d}\underline{x}. \tag{24}$$

The corresponding tangent operator $\mathbb{M}$ consists of three terms:

$$\mathbb{M} = \mathbb{M}^a + \mathbb{M}^b + \mathbb{M}^c, \tag{25}$$

with $\forall (i,j) \in [1,n]^2$

$$M^a_{ij} = -\int_\Omega \underline{\varphi}_i(\underline{x},\mathbb{a}) \cdot \underline{G}(\underline{x},\mathbb{a}) \frac{\partial r}{\partial a_j}(\underline{x},\mathbb{a})\, \mathrm{d}\underline{x}, \tag{26}$$

$$M^b_{ij} = -\int_\Omega \underline{\varphi}_i(\underline{x},\mathbb{a}) \cdot \frac{\partial \underline{G}}{\partial a_j}(\underline{x},\mathbb{a}) r(\underline{x},\mathbb{a}) \mathrm{d}\underline{x}, \tag{27}$$

$$M^c_{ij} = -\int_\Omega \frac{\partial \underline{\varphi}_i}{\partial a_j}(\underline{x},\mathbb{a}) \cdot \underline{G}(\underline{x},\mathbb{a})\, r(\underline{x},\mathbb{a})\, \mathrm{d}\underline{x}. \tag{28}$$

Expanding $\mathbb{M}^a$ in a manner similar to b_i, one can write, $\forall (i,j) \in [1,n]^2$,

$$M^a_{ij} = \int_\Omega \underline{\varphi}_i(\underline{x},\mathbb{a}) \cdot \underline{G}(\underline{x},\mathbb{a})\, \underline{G}(\underline{x},\mathbb{a}) \cdot \underline{\varphi}_j(\underline{x},\mathbb{a})\, \mathrm{d}\underline{x}. \tag{29}$$

The second term, $\mathbb{M}^b$, involves the derivative of $\underline{G}$ towards the DOF. Applying the chain rule on Equation (23) yields, $\forall j \in [1,n]$,

$$\frac{\partial \underline{G}}{\partial a_j} = \frac{\partial \underline{\phi}}{\partial a_j} \cdot \underline{\underline{\mathrm{grad}(\mathrm{grad}(g))}} \circ \underline{\phi} = \underline{\varphi}_j \cdot \underline{\underline{\mathcal{G}}}, \tag{30}$$

where $\underline{\underline{\mathcal{G}}}(\underline{x},\mathbb{a})$ is the second gradient of image g, evaluated at $\underline{\phi}$. Subsequently, M^b_{ij} becomes

$$M^b_{ij} = -\int_\Omega \underline{\varphi}_i(\underline{x},\mathbb{a}) \cdot \underline{\underline{\mathcal{G}}}(\underline{x},\mathbb{a}) \cdot \underline{\varphi}_j(\underline{x},\mathbb{a}) r(\underline{x},\mathbb{a}) \mathrm{d}\underline{x}. \tag{31}$$

Equations (24), (28), (29), and (31) constitute the consistent solution for the iterative Gauss–Newton minimization of the (non-linear) brightness conservation. Using this formulation should provide the expected fast convergence behavior attributed to Gauss–Newton methods, if close to the solution.

2.3. Recovery of the common digital image correlation formulation

Let us now focus on the assumptions that need to be made to recover the formulations used in DIC literature. Starting with the terms $\mathbb{M}^b$ and $\mathbb{M}^c$, both terms involve the image residual r, which tends to be large at the beginning of the iterative process, possibly driving the minimization into a local minimum. Moreover, $\mathbb{M}^b$ also contains the second gradient of the image, making it highly sensitive to noise. Finally, the adopted basis is often linearly independent, reducing $\mathbb{M}^c$ to zero (i.e., the basis functions φ are independent of the DOF $\mathbb{a}$). Therefore, to improve robustness, one may choose to omit both terms, yielding a relaxed tangent operator, that is,

$$\mathbb{M} \equiv \mathbb{M}^a. \tag{32}$$

DOI: 10.1002/nme

This omission may be justified by the fact that both terms approach zero when approaching the optimal solution in the absence of acquisition noise.

$\mathbb{M}^a$ contains the true image gradient $\underline{G}$, which relates as follows to the gradient of the commonly reported back-deformed image $\underline{\text{grad}}(\tilde{g})$:

$$\underline{\text{grad}}(\tilde{g}) = \underline{\text{grad}}\left[g \circ \underline{\phi}\right], \tag{33}$$

$$= \underline{\underline{\text{grad}}}(\underline{\phi}) \cdot \underline{\text{grad}}(g) \circ \underline{\phi}, \tag{34}$$

thus,

$$\underline{G} = \underline{\text{grad}}(g) \circ \underline{\phi} = \underline{\text{grad}}(\tilde{g}) \cdot \underline{\underline{F}}^{-T}, \tag{35}$$

where $\underline{\underline{F}}$ is the deformation gradient tensor. To simplify this form to the one regularly found in literature, first, small deformations need to be assumed, that is, $\underline{\underline{F}}^T \equiv \underline{\underline{I}}$. Second, $\underline{\text{grad}}(\tilde{g})$ is replaced with $\underline{\text{grad}}(f)$, which has been justified in the literature because $\tilde{g}$ is updated at each iteration and converges towards f as follows [12, 18, 22]:

$$\underline{G}(\underline{x}, \mathbb{a}) \approx \underline{\text{grad}}(f). \tag{36}$$

By substituting the foregoing equation in Equations (24) and (29), the commonly used DIC formulation is retrieved (Equations (17) and (18)). In doing so, a total of four implicit assumptions have been revealed:

$$\begin{aligned} \mathbb{M}^b \equiv \mathbb{0}, \qquad & \underline{\underline{F}}^T \equiv \underline{\underline{I}}, \\ \mathbb{M}^c \equiv \mathbb{0}, \quad & \underline{\text{grad}}(\tilde{g}) \equiv \underline{\text{grad}}(f). \end{aligned} \tag{37}$$

Depending on the DIC application at hand, each assumption may be justified. Applying all four will yield a tangent operator, which is independent of the DOF, and thus can be computed once and for all iterations, reducing a major part of the computational cost. Alternatively, it is possible to partly conserve this computational advantage, without assuming small displacements, by applying the large displacement correction (e.g., $\underline{\underline{F}}$) to the update in the DOF instead of to the image gradient [23]. For the case of integrated DIC, the basis functions typically depend on the DOF and thus require updating for each iteration [24–27]. Therefore, for integrated DIC, the additional cost for applying the consistent tangent operator may be negligible compared with the benefit of converging in less iterations.

The aforementioned discussions lead to six possible tangent operators, each with different levels of approximation (Table II). Additionally, mixed gradients have been proposed in literature as well, for example [18],

$$\underline{G} \approx \frac{1}{2}\left(\underline{\text{grad}}(\tilde{g}) + \underline{\text{grad}}(f)\right), \tag{38}$$

which, because of anti-symmetry in the DOF estimates obtained by $\underline{\text{grad}}(\tilde{g})$ or $\underline{\text{grad}}(f)$, leads to a lower truncation error of the linearization of the pattern and, in the case of acquisition noise, allows for additional noise attenuation by the averaging of both image gradients. A similar mixed gradient is also possible for the proposed large deformation image gradients, for example,

$$\underline{G} \simeq \frac{1}{2}\left(\underline{\text{grad}}(\tilde{g}) + \underline{\text{grad}}(f)\right) \cdot \underline{\underline{F}}^{-T}, \tag{39}$$

resulting in a total of nine tangent operators; see Table II.

DOI: 10.1002/nme

Table II. The discussed variations of the image gradient and their respective assumptions.

	Assumptions			
Image gradient	$\mathbb{M}^b \equiv \mathbb{0}$	$\mathbb{M}^c \equiv \mathbb{0}$	$\underline{\underline{F}}^T \equiv \underline{\underline{I}}$	$\underline{\text{grad}}(\tilde{g}) \equiv \underline{\text{grad}}(f)$
$\mathbb{M} = \mathbb{M}^a + \mathbb{M}^b + \mathbb{M}^c$	No	No	No	No
$\underline{G}_1 = \underline{\text{grad}}(g) \circ \underline{\phi}$	Yes	Yes	No	No
$\underline{G}_2 = \underline{\text{grad}}(\tilde{g}) \cdot \underline{\underline{F}}^{-T}$	Yes	Yes	No	No
$\underline{G}_3 \equiv \underline{\text{grad}}(f) \cdot \underline{\underline{F}}^{-T}$	Yes	Yes	No	Yes
$\underline{G}_4 \equiv \underline{\text{grad}}(\tilde{g})$	Yes	Yes	Yes	No
$\underline{G}_5 \equiv \underline{\text{grad}}(f)$	Yes	Yes	Yes	Yes
$\underline{G}_6 \equiv (\underline{G}_2 + \underline{G}_3)/2$	Yes	Yes	No	Mixed
$\underline{G}_7 \equiv (\underline{G}_4 + \underline{G}_5)/2$	Yes	Yes	Yes	Mixed
$\underline{G}_8 \equiv (\underline{G}_1 + \underline{G}_5)/2$	Yes	Yes	Mixed	Mixed

3. ILLUSTRATION ON ACADEMIC EXAMPLES

Before discussing the test case results, the implementation of the various image gradients is discussed, starting with the most straightforward implementation ($\underline{G}_5$) gradually extended to the most involved implementation $\mathbb{M}$. The gradient operator is implemented by a forward–backward numerical differentiation scheme for the inner pixels; the edge pixels are differentiated by forward or backward differentiation depending on the side of the image; however, these are outside the ROI and are not expected to influence the results. For the sub-pixel interpolation scheme, a two-dimensional (2D) cubic spline interpolation is used, wherever the images (or derivatives) are to be evaluated at non-integer positions. With these $\underline{G}_4$, $\underline{G}_5$ and $\underline{G}_7$ can be implemented. To correct these for large deformations, the deformation gradient tensor is required, which is obtained by applying the numerical differentiator on the current displacement field as follows:

$$\underline{\underline{F}}^T = \underline{\underline{\text{grad}}}(\underline{u}) + \underline{\underline{I}} \tag{40}$$

where $\underline{\underline{I}}$ is the 2D unity tensor. Because the deformation gradient tensor is only 2D in this case, it is efficiently inverted algebraically; however, this inversion is required for each pixel, making it still relatively costly. With $\underline{\underline{F}}^{-T}$ known $\underline{G}_2$, $\underline{G}_3$ and $\underline{G}_6$ can be implemented. To implement $\underline{G}_1$, the image gradient of g is required at the estimated deformed coordinates. The gradient of g is constant for all iterations and thus can be computed once and for all. However, to obtain $\underline{G}_1$, this image gradient requires interpolation for each iteration. The same is true for the second gradient required to implement $\mathbb{M}$; see Equation (31). Interpolating images (or their derivatives) will always introduce some error. Consequently, the implementation of $\underline{G}_1$ is expected to be slightly less accurate than, for instance, $\underline{G}_6$ but more straightforward to implement because it does not require the computation of $\underline{\underline{F}}^{-T}$.

3.1. One-dimensional test case

For this test case, a one-dimensional (1D) pattern is generated and artificially deformed with a quadratic displacement field to obtain images f and g (Figure 1a). To evaluate the performance of all image gradients, a number of DIC options have to be chosen. Note that the conclusions resulting from Section 2 are general for DIC and thus apply for both subset-based (or local) DIC and global DIC implementations, yet, for the examples, only global implementations are investigated. For this test case, a cubic polynomial basis is chosen for the displacement parameterization, such that the basis can capture the exact solution of the applied displacement field. Figure 1b shows the convergence performance for the nine tangent operators, which are all initiated using a zero initial guess for all parameters. For each curve, two regimes can be identified: an initial regime far from the solution where convergence is slow followed by a regime with rapid convergence close to the solution. Considering the second (rapid) regime, different convergence rates can be identified. The

DOI: 10.1002/nme

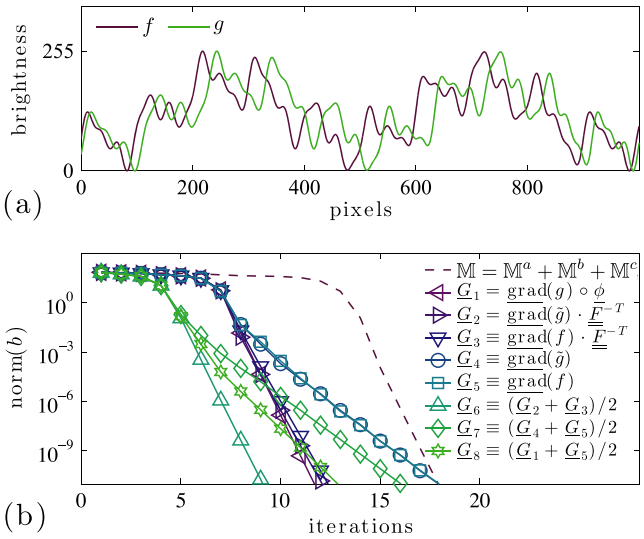

Figure 1. (a) A one-dimensional example, where a non-linear (quadratic) displacement field is applied to the pattern f to generate g, and (b) the convergence performance for nine tangent operators.

fully consistent operator ($\mathbb{M} = \mathbb{M}^a + \mathbb{M}^b + \mathbb{M}^c$) and all other formulations using the large deformation formulations (LDF, i.e., $\underline{G}_1$, $\underline{G}_2$, $\underline{G}_3$, and $\underline{G}_6$) show an improved convergence rate with respect to the small deformation formulations (SDFs, i.e., $\underline{G}_4$, $\underline{G}_5$, and $\underline{G}_7$). In the initial (slow) regime, the formulations appear in three groups, transitioning from slow to rapid convergence at iterations 12, 7, and 4, respectively. In agreement with the foregoing discussion, the fully consistent operator converges poorly when far from the solution, even for the relatively smooth pattern applied in this example case. The single image gradient formulations (i.e., $\underline{G}_1$ to $\underline{G}_5$) form the center group, transitioning at the seventh iteration. The third group are the mixed gradients (i.e., $\underline{G}_6$ to $\underline{G}_8$), which converge faster in the slowly converging regime. This slow regime, where the mixed gradients outperform the single gradients, depends on the quality of the initial guess; the regime will not be observed for a routine started with a good enough initial guess. Evidently, these results are specific for this test case.

3.2. *Rotation test case*

For this test case, 25 reference patterns are generated by randomly placing Gauss-curve shaped speckles (Equation (41)) with a radius of two pixels ($\sigma = 2$) at an average distance of 20 pixels to the nearest speckle (Figure 2). The resulting correlation length for all patterns is $\zeta = 2.9 \pm 0.2$ px. Because the speckle pattern is defined by a sum of closed-form Gauss functions, it can be rotated analytically before evaluating it at the pixel positions to create the reference and deformed images, which are (512×512) px^2. Each of the 25 reference patterns is rotated with an angle θ, which ranges from zero to π.

$$f(x, y) = \sum_i \exp\left(\frac{(x - x_{ci})^2}{-2\sigma^2} + \frac{(y - y_{ci})^2}{-2\sigma^2}\right). \tag{41}$$

To describe the displacement field, a mesh is applied with a single four-noded bilinear element, as shown in Figure 2a. Consecutively, all discussed image gradients are applied to solve for the displacement field between the reference image (e.g., Figure 2a) and a rotated image (e.g., Figure 2d), where in all cases the DIC algorithm is initiated with nodal displacements computed using the rotation angle θ corresponding to the test case, that is, the perfect initial guess. Figure 2e shows the average number of iterations required to converge for each of the discussed image gradients as a function of the rotation angle. The results are averaged over the 25 different patterns and for both

DOI: 10.1002/nme

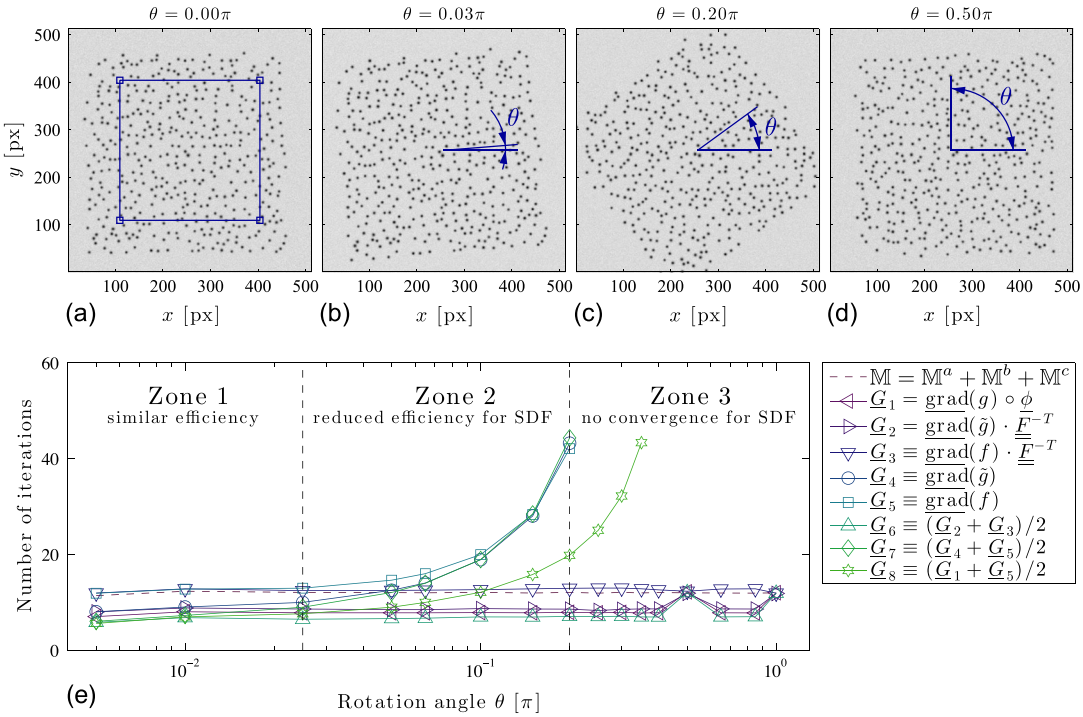

Figure 2. (a) One of the 25 generated reference images (f). (b–d) The corresponding 'deformed' images (g) at various stages of rotation, with the rotation angle indicated by θ. (e) The number of iterations required to converge for each of the defined image gradients as a function of rotation angle. SDF, small deformation formulation.

positive and negative rotation angles to ensure statistical independence of the results. In zone 1, the DIC routine converges for all patterns for all image gradients, with a small difference in the total number of iterations required between the image gradients. In zone 2, the SDF (i.e., $\underline{G}_4$, $\underline{G}_5$, and $\underline{G}_7$) image gradients suffer a reduction in performance, while the LDF (i.e., $\mathbb{M}$, $\underline{G}_1$, $\underline{G}_2$, $\underline{G}_3$, and $\underline{G}_6$) image gradients perform equally well as in zone 1. In zone 3, the performance for the SDF image gradients (for any of the 25 patterns) degrades further up to the point where convergence lost. Although case dependent, these results illustrate clearly that it might be important to take large displacements into account. For cases where the rotation angle is more than 45°, it is unlikely that a small displacement formulation image gradient, like $\underline{\text{grad}}(f)$, will drive the iterative scheme to the correct solution.

4. APPLICATION TO THE BENCHMARK PROBLEM OF BORNERT *ET AL.*

To evaluate the performance of the various image gradients in more detail, a more representative test case is used in the next section. Bornert *et al.* [28] evaluated the performance of various DIC algorithms on a series of images subject to a non-linear (sinusoidal) displacement field. They made the images used for their study available [29] allowing other researchers to re-evaluate the results under different conditions and connect their findings to the results of the original paper. The well-defined reference displacement field and noise fields allow for direct evaluation of the accuracy of the obtained results. Additionally, one of the key differences between the proposed image gradients is the existence of a large deformation correction in some of the gradients; therefore, the test case presented by Bornert *et al.* is interesting for this analysis because the underlying displacement field is non-linear. The data and images discussed in [28] are chosen because they are representative for the types of experiments where DIC is often applied to. The published images by Bornert *et al.* are all artificially generated images using well-established algorithms to generate the speckle pattern

DOI: 10.1002/nme

as realistically as possible. The advantage of generated images versus real images captured from experiments is that the displacement field that is to be identified is exactly known. Additionally, it is possible to control other experimental uncertainties such as acquisition noise to evaluate the performance of the algorithm under various levels of noise.

As discussed in Section 2, in DIC, a non-linear problem is linearized and iteratively solved. One contribution to the non-linearity of the problem is the pattern, which is deliberately irregular and chaotic. The pattern correlation length can be seen as an indicator of the strength of this non-linearity. Bornert *et al.* supplied image series for three different patterns (i.e., fine, medium, and coarse). The patterns were chosen such that their correlation lengths ζ are approximately double at each level $4\zeta_{\text{fine}} \approx 2\zeta_{\text{medium}} \approx \zeta_{\text{coarse}}$. The correlation contours are shown in Figure 3, which are computed by finding the distance at which the autocorrelation function equals $1/2$ for each radial direction. The contours appear reasonably circular, indicating that the speckle pattern is isotropic and can be expressed by a single correlation length.

Bornert *et al.* applied a sinusoidal displacement field in the x-direction and no displacement in the y-direction. Consequently, the displacement field is non-linear and without any rotation. The displacement field is expressed as follows:

$$\underline{u}(\underline{x}) = \alpha p \sin\left(\frac{2\pi}{p}\underline{x}\cdot\underline{e}_x\right)\underline{e}_x, \tag{42}$$

where $\underline{x}$ is in pixels, p is the period of the sine wave in pixels, and $2\pi\alpha$ is the amplitude of the xx-component of the strain. Reference [29] contains images for a range of periods $p = [10, 20, 30, 60, 130, 260, 510]$, but only the last three cases are given for all speckle sizes and all noise levels. Consequently, in this article, only the last three are considered in the evaluation. Additionally, for those cases, the displacement is increased incrementally, with $\alpha = [0.001, 0.005, 0.01, 0.02]$ corresponding to maximum strain levels of $\varepsilon_{\max} = [0.63\%, 3.1\%, 6.3\%, 12.6\%]$. Because the discussion of the image gradient is pivoting around small or large deformations, the last increment with the largest strain is used in most of the results later on. The images are of size 512×512 px^2, and the ROI is centered in the image with a size of 472×472 px^2, that is, with an unused boundary of 20 px at each side. Within this ROI, two types of shape functions are applied: harmonic and four-noded FEM quadrilaterals (Q4). For the Q4 shape functions, regular bilinear shape functions are applied on a regular mesh of 9×9 nodes, resulting in 81 DOF per direction of displacement, that

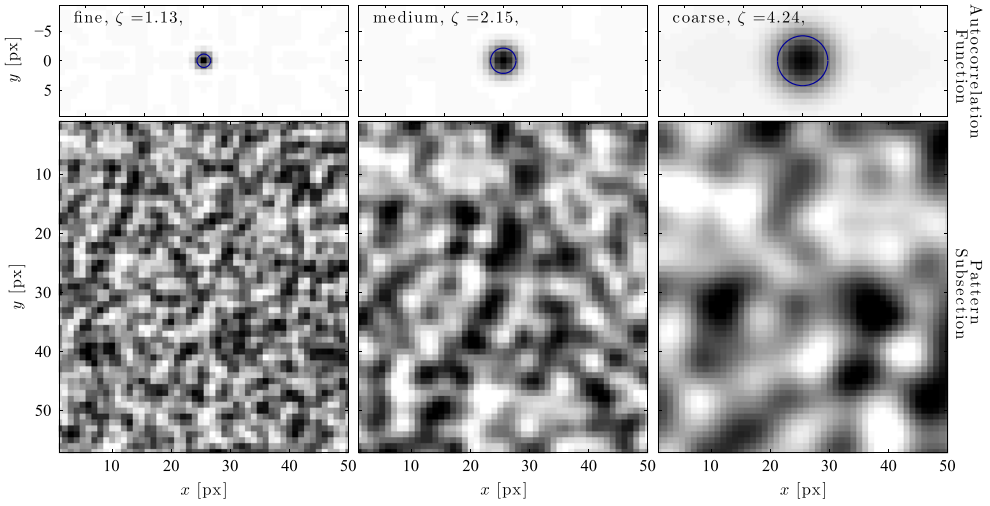

Figure 3. The three speckle patterns (i.e., fine, medium, and coarse) shown enlarged in the lower three figures, with their respective autocorrelation functions (ACF) shown previously at the same scale. The correlation contour is plotted inside the ACF, defined as the radial distance at which the ACF equals $1/2$. The average of the radial distance is the correlation length ζ given in the respective figures.

DOI: 10.1002/nme

is, $2 * 81$ in total. This set of shape functions is referred to as φ_{81q}. The applied mesh is shown in Figure 4, showing that the boundary elements are larger by a factor of $\sqrt{2}$ to balance the uncertainty of the boundary nodes with respect to the center nodes [17]. The second type of basis functions is generated by the dyadic product of two 1D harmonic shapes defined as follows:

$$\varphi_i(\underline{x}) = P_\alpha(x) \otimes P_\beta(y), \qquad \forall \alpha, \beta \in [1, 2, \ldots, n], \quad (43)$$

$$P_k(x) = \frac{1 + (-1)^k}{2} \sin(k\pi x/p) + \frac{1 + (-1)^{k-1}}{2} \cos((k-1)\pi x/p), \quad \forall k \in [1, 2, \ldots, n] \quad (44)$$

where n is the number 1D harmonic functions including the first mode $P_1 = 1$. The sets of harmonic shape functions are created in three variations. In the first set $n = 3$, which, through the dyadic combinations, results in 9 DOF (per displacement direction, i.e., 18 in total), ergo, this set is called φ_{9h}. The other two variations are called φ_{81h} and φ_{169h} corresponding to the respective number of DOF, that is, $n = 9$ and $n = 13$.

4.1. Displacement accuracy

Figure 5 shows the displacement accuracy for all image gradients for the case where $\alpha = 0.02$ and $p = 510$ for the four shape function sets. The displacement accuracy is defined as follows:

$$\mathcal{E}_u = \frac{1}{N} \sum_{i=1}^{N} \|\underline{u}(\underline{x}_i) - \underline{U}(\underline{x}_i)\| \quad (45)$$

where $\underline{u}(\underline{x}_i)$ is displacement vector at the ith pixel obtained by the DIC algorithm after convergence. $\underline{U}(\underline{x}_i)$ is the reference counterpart of $\underline{u}(\underline{x}_i)$ obtained by evaluating Equation (42), and N is the total number of pixels in the ROI. The usual $\mathbb{R}^2$ Euclidean norm $\|\cdot\|$ is used here to compute the vector length. The displacement accuracy can be seen as the average distance (in pixels) of the obtained

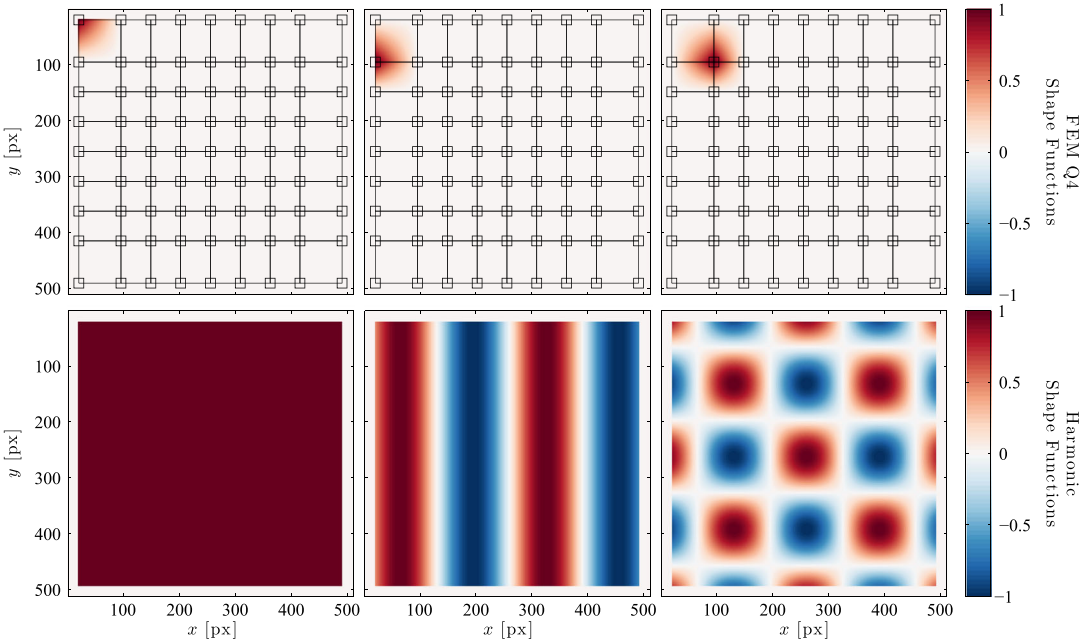

Figure 4. Three example shape functions for each family of shape functions, that is, (bottom) two-dimensional harmonic shapes and (top) four-noded FEM quadrilaterals (Q4).

DOI: 10.1002/nme

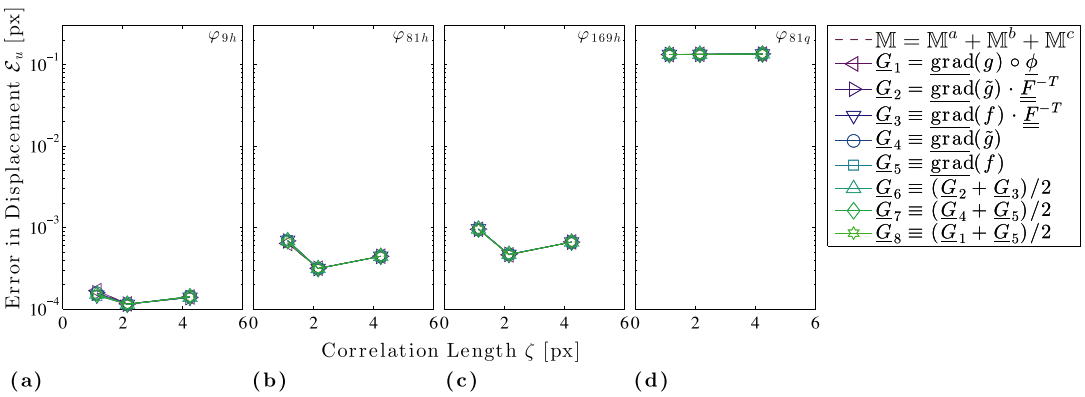

Figure 5. The displacement error as a function of the speckle size (correlation length) for all four tested shape function sets (harmonic (a–c) and FEM Q4 (d)) for $\alpha = 0.02$ and $p = 510$. Note that all considered image gradients perform equivalently for these cases.

solution to the perfect solution. The important conclusion of Figure 5 is that the same accuracy is observed for all discussed image gradients, for the evaluated cases. This conclusion is applicable to most of the current applications of DIC, provided that convergence is found. Consequently, current results in the literature, obtained by using SDFs, are not challenged. Therefore, the discussion on which implementation to choose reduces to computational efficiency and initial guess robustness only, at least in the absence of severe rotations.

Figure 5 also shows that the most accurate result is obtained when using φ_{9h}. This is expected because this basis contains the exact solution (within a single degree of freedom) and only a small number of additional DOF. Adding extra (kinematically unnecessary) DOF only makes the method more sensitive to noise and interpolation errors, where for this case, only the latter are present. The sensitivity to interpolation errors is also expressed by the optimum accuracy for the medium correlation length speckle size ($\zeta \approx 2.15$). Shorter correlation lengths result in larger image gradients making DIC methods more sensitive and thus more accurate. At the same time, shorter correlation length patterns present more difficulties for the sub-pixel interpolation methods reducing the accuracy and leading to an optimum in accuracy for these particular cases. The relatively bad performance of the Q4 shape functions is due to the coarse mesh, which cannot properly accommodate the sinusoidal displacement field, that is, it is an under-kinematic description. This is not representative for FE-based DIC methods in general.

4.2. Acquisition noise

In Reference [29] also, images are provided with three levels of acquisition noise, that is, $\eta = [0.8, 2.5, 4.8]\,\sigma_{\text{GV}}$, where σ_{GV} is the standard deviation of the noise field in gray value intensity. These noise levels are relatively low, with the largest approximately 1.9% of the dynamic range of the pattern (i.e., 255). This is nevertheless sufficient to evaluate the influence of acquisition noise on the evaluated image gradients. Figure 6 shows the results for a single basis (φ_{81h}) for $\alpha = 0.02$ and $p = 260$. Again, no significant difference is observed between the discussed image gradients. This means that, at least for the tested case, the displacement accuracy of all the image gradients is equal if convergence is obtained. This confirms that a particular image gradient implementation can be selected on efficiency and robustness arguments only (in the absence of large rotations).

4.3. Convergence rate

The academic example discussed in Section 3 showed that the proposed image gradients can be grouped in two ways: (1) based on their performance far from the solution where the mixed formulations performed the best ($\underline{G}_6$ and $\underline{G}_7$); and (2) based on their performance close to the solution, where the LDFs performed the best ($\underline{G}_1$, $\underline{G}_2$, $\underline{G}_3$, and $\underline{G}_6$). For each combination of the discussed

DOI: 10.1002/nme

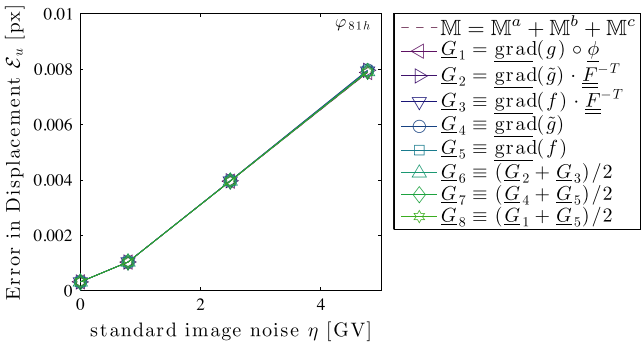

Figure 6. The displacement accuracy as a function of image noise for the intermediate harmonic shape function set for $\alpha = 0.02$ and $p = 260$. Note that all evaluated image gradients perform equally.

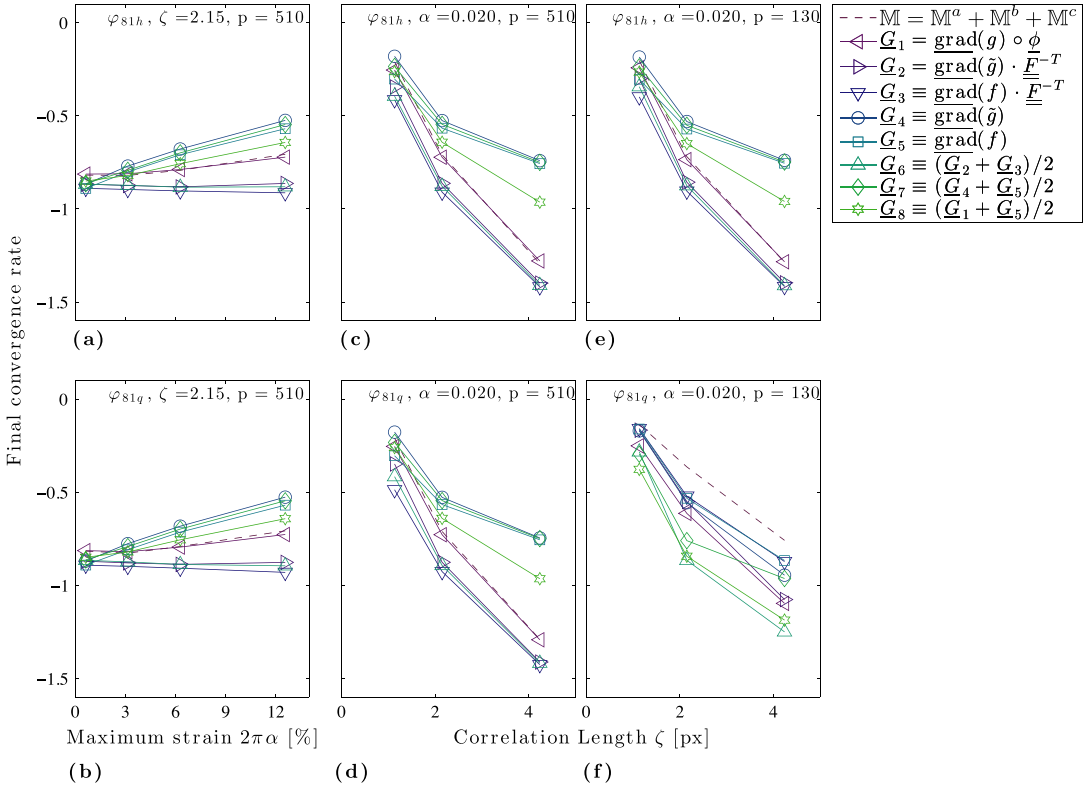

Figure 7. (a,b) The final convergence rate, obtained by determining the logarithmic decay in the two-norm of the right-hand member $\mathbb{b}$, as a function of the maximum strain $2\pi\alpha$ for both families of basis functions. The displacement period p is indicated at the top of the respective figures. (c–f) The final convergence rate as a function of the correlation length for the harmonic (φ_{81h}) and the FEM Q4 (φ_{81q}) basis.

cases (variation on φ, α, p, and ζ), curves similar to Figure 1b with a bilinear representation (on a log scale) are superposed to capture the initial slope, the transition iteration, and the final slope. The obtained final convergence rate for each image gradient implementation for several test cases is presented in Figure 7. The presented values for the final convergence rate are the logarithmic decay in the two-norm of the right-hand member $\mathbb{b}$ per iteration, with -1 indicating a linear convergence rate.

DOI: 10.1002/nme

Figure 7a and 7b shows that the convergence rate of the SDF degrades for increasing maximum strain amplitude $2\pi\alpha$, while the LDF convergence rates remain approximately independent of maximum strain. This is in line with the assumptions required to go from the consistent LDF gradient to the SDFs. In these figures, it can be observed that both $\mathbb{M}$ and $\underline{G}_1$ show the same trend as the other LDF algorithm but with reduced performance, likely due to the additional interpolation steps required to compute these implementations. Mixing a large and small deformation formulation as is carried out in $\underline{G}_8$ is also possible, and as expected, a performance of such a hybrid gradient is in between the SDF and LDF results. Both types of basis functions, for example, harmonic and FE, show the same behavior confirming that this conclusion is independent of the chosen basis.

Figure 7c and 7d shows the convergence rate as a function of correlation length for a displacement field with a long wavelength (i.e., $p = 510$). For short correlation lengths (fine pattern), all image gradients perform relatively similar, with the LDFs (i.e., $\mathbb{M}$, $\underline{G}_1$, $\underline{G}_2$, $\underline{G}_3$, and $\underline{G}_6$) performing slightly better than the SDFs (i.e., $\underline{G}_4$, $\underline{G}_5$, and $\underline{G}_7$). The difference in performance between both groups becomes more pronounced for images with larger speckle sizes. These observations lead to the conclusion that the advantage of using an LDF algorithm degrades for patterns with small speckle sizes. For this long wavelength displacement field ($p = 510$), there is no observable difference between the harmonic and FE basis functions. However, for the shorter wavelength displacement field ($p = 130$), Figure 7e and 7f shows a difference in trend between the FE basis and the harmonic basis. Figure 7f shows that the convergence rate for almost all gradient formulations is decreased for the FE basis, but the impact on the single gradients $\underline{G}_1$ to $\underline{G}_5$ is more severe than for the mixed gradients. It is unclear what is the precise cause of this, but it is suspected that this behavior is related to the fact that for this short wavelength, the under-kinematic displacement description (coarse mesh) is responsible for the shift in behavior.

These observations highlight that, for the tested cases, the LDF gradients, like $\underline{G}_3$ and $\underline{G}_6$, converge with less iterations, especially for cases where the underlying displacement field contains large deformations. The benefits of the LDF gradients become more pronounced for image patterns, which have larger speckle sizes. This is probably because these image gradients are less numerically unstable and therefore emphasize the advantage of an LDF gradient more. It should be noted that the discussed convergence rate is expressed in iterations and not in computation time. The LDF gradients need more operations to compute and thus are computationally more expensive per iteration than, for instance, $\underline{G}_5$. The total computational efficiency is a combination of the number of iterations and the cost per iteration, which can vary significantly between DIC implementations (e.g., subset and global).

The computational cost for this test case, using the FE shape functions with 9, 81, and 169 nodes, is presented in Figure 8. The total cost is split into a preprocessing cost and the cost per iteration. The preprocessing cost contains the computation of all objects, which do not change during the iterative processes, for example, gradient images (without deformation) and correlation matrix. Most of these objects change during the iterative process for the LDFs. These objects are therefore accounted for

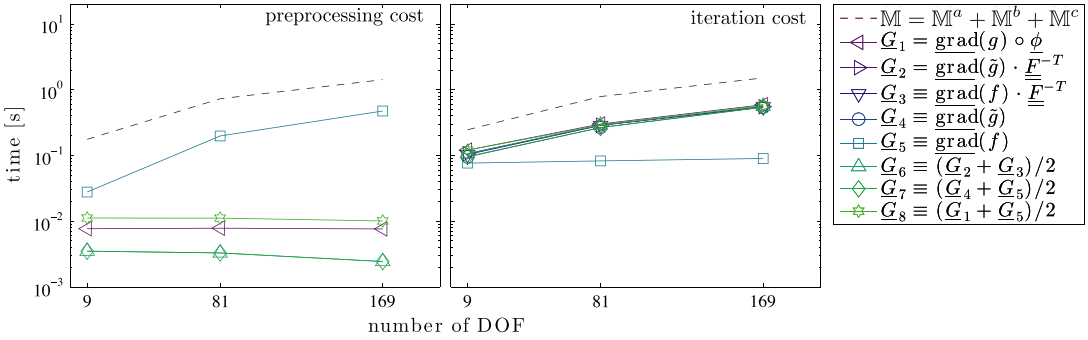

Figure 8. The computational cost for each algorithm split into the computation time required to preprocess the iterative procedure and the time required per iteration.

DOI: 10.1002/nme

in the iteration cost. Figure 8 shows that, for larger meshes, one iteration of the $\underline{G}_5$ algorithm is nearly 10 times faster than all the LDF algorithms. Consequently, for cases where the $\underline{G}_5$ algorithm is converging properly, it will almost always be the preferred option. Interestingly, the variation in cost for the LDF algorithms is small ($\mathbb{M}$ excluded). Therefore, for those cases that benefit from the LDF algorithms, $\underline{G}_6$ is the recommended option because it typically requires the least amount of iterations. It is emphasized that the particular implementation was written in MATLAB (MathWorks, Natick, MA, USA) not focusing on computational efficiency. Consequently, the computational time may change significantly when changing to an optimized implementation, especially when using a low-level, compiled programming language. Nevertheless, it is expected that the main conclusions drawn here remain valid.

4.4. Initial guess robustness

To evaluate the initial guess robustness of the discussed image gradients, the DIC routines are initialized with an initial guess for the DOF, which is at a specific distance to the solution. To have a fair comparison for all test cases, the initial guess distance is implemented as a perturbation field added to the ideal initial guess, which best fits the known reference displacement (i.e., Equation (42)). This perturbation field is defined as a rigid body translation of distance γ at 45°, 135°, 225° and 315° with respect to the x-axis. The results for the four diagonal directions are averaged and presented in Figure 9.

The first observation made from this figure is that, for all image gradients, the maximum allowable initial guess distance $\gamma_{\max}$ increases with increasing speckle size. Additionally, it can be observed that the LDF gradients consistently require less iterations to converge, especially for the larger correlation length images (i.e., Figure 9b and 9c). For the larger initial guess distances, no convergence is found, and the iterative procedures are aborted. The second observation is that the mixed formulations (e.g., $\underline{G}_6$) consistently converged for larger initial guess distances than the single formulations. Furthermore, the choice of the adopted image gradient can be changed throughout the iterative process, and a scheme where initially an expensive formulation like $\underline{G}_6$ is used before switching to the more efficient formulation of $\underline{G}_5$ may present itself as an intelligent compromise.

It is a common practice to use pyramidal, or multi-scale, approaches in DIC to cope with inexact initial guesses (e.g., [17, 18]). In such an approach, the images are blurred using a super pixel technique where a coarse-scale image is formed from the average of groups of pixels. This effectively increases the correlation length, providing a larger convergence radius to the algorithm. The obtained solution is then used to initialize the algorithm at a finer scale, until ultimately the original, non-blurred image is used. Typically, each step of a multi-scale approach returns a solution that is within a few pixels of the solution of the next step. Consequently, an efficient recipe

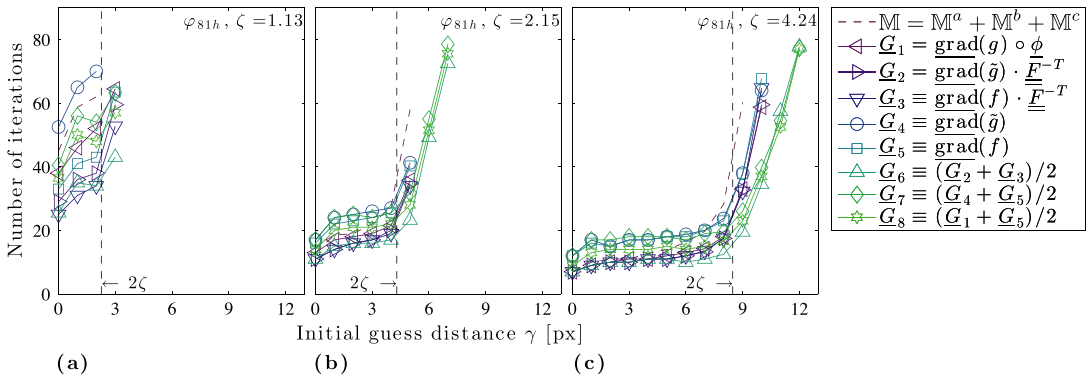

Figure 9. The number of iterations required to converge, as a function of initial guess distance γ, for all image gradients for all three speckle size images (correlation length ζ) for $\alpha = 0.02$. The results are averaged for the three evaluated periods p and for all four directions of applying the initial guess distance.

DOI: 10.1002/nme

may be to use $\underline{G}_6$ for the first step, when the initialization is far from the solution, and then switch to the fast $\underline{G}_5$ algorithm for the consecutive steps, provided that there are no large rotations in the considered case.

5. CONCLUSIONS

By following the direct linearization of the brightness conservation, a fully consistent tangent operator is obtained for the iterative minimization procedure. Within this tangent operator, a large deformation corrected image gradient appears. This large deformation gradient showed better convergence performance when close to the solution. The consistent minimization scheme can be reduced to the one typically discussed in DIC literature if four assumptions are made: (1) a small residual ($\mathbb{M}^b \equiv \mathbb{0}$); (2) linearly independent basis functions ($\mathbb{M}^c \equiv \mathbb{0}$); (3) small deformations ($\underline{\underline{F}} \equiv \underline{\underline{I}}$); and (4) a good initial guess ($\underline{\text{grad}}(\tilde{g}) \equiv \underline{\text{grad}}(f)$). The results on the evaluated test cases show that the solutions obtained from each of the eight researched image gradients are equivalent. In other words, the accuracy does not depend on the chosen image gradient, if convergence is found.

The large deformation correction to the image gradient is not well known to the DIC community. Probable reasons for this are the two-step linearization and the delicate notation of the coordinate system. However, the large deformation correction is not new, because it is common in the solid mechanics field, often referred to as a pullback, and it is correctly observed by Baker and Matthews [5] in their compositional algorithm. Yet, the results in this paper on the influence of the large deformation correction on the performance of DIC on representative cases are novel, and a clear recommendation for future DIC implementations is at hand.

Because the accuracy is unaffected by the choice of the image gradient, if convergence is obtained, the remaining arguments for selecting an image gradient for a particular DIC case are based on efficiency and initial guess robustness and the presence of severe rigid body rotations. The discussion regarding computational efficiency is highly case and code specific, and general conclusions in this topic are often poorly supported. Nevertheless, the mixed large deformation formulation (LDF) (i.e., $\underline{G}_6 = \frac{1}{2}(\underline{\text{grad}}(f) + \underline{\text{grad}}(\tilde{g})) \cdot \underline{\underline{F}}^{-T}$) converged in less iterations for most, if not all, evaluated cases. However, an iteration using this gradient is more expensive than an iteration that uses the classical small deformation formulation (SDF_ (i.e., $\underline{G}_5 = \underline{\text{grad}}(f)$). The main reason why the latter is so efficient is that this gradient is constant for all iterations, causing other parts in the DIC algorithm (e.g., the correlation matrix $\mathbb{M}$) to be constant as well. This significantly reduces the cost for all consecutive iterations. Consequently, even though this implementation typically uses many iterations, for the tested cases, this implementation was almost always the most efficient (if converged). Because all other discussed gradients require updating of the image gradient (and $\mathbb{M}$) during the iterative process, the associated loss of efficiency is more significant than the variations in computational efficiency between the other gradients. To conclude, the image gradient $\underline{G}_5$, common in DIC literature, is the preferred implementation for most DIC cases, except for the following three cases where the mixed large deformation gradient $\underline{G}_6$ is the preferred option:

(1) Large rotations: for cases where rotations exceed 45°, the small displacement formulations are unlikely to converge, and a large displacement formulation is required.
(2) Initial guess robustness: mixed gradient formulations, using both the reference and the deformed image, improve the stability of the implementation when far from the solution.
(3) Costly iterations: if the DIC iterations are already costly because of other reasons, for instance, for integrated DIC routines (where computing the basis functions for each iteration tends to outweigh the normal DIC costs), reducing the number of iterations is obviously advantageous.

ACKNOWLEDGEMENTS

This work has been supported by the Dutch Technology Foundation (STW) and the Dutch Organization for Scientific Research (NWO).

DOI: 10.1002/nme

REFERENCES

1. Bonnet M, Constantinescu A. Inverse problems in elasticity. *Inverse Problems* 2005; **21**(2):R1–R50.
2. Grédiac M, Pierron F. Applying the virtual fields method to the identification of elasto-plastic constitutive parameters. *International Journal of Plasticity* 2006; **22**(4):602–627.
3. Avril S, Bonnet M, Bretelle AS, Grédiac M, Hild F, Ienny P, Latourte F, Lemosse D, Pagano S, Pagnacco E, Pierron F. Overview of identification methods of mechanical parameters based on full-field measurements. *Experimental Mechanics* 2008; **46**:381–402.
4. Lucas BD, Kanade T. An iterative image registration technique with an application to stereo vision. *Proceedings DARPA Image Understanding Workshop*, Vancouver, 1981; 121–130.
5. Baker S, Matthews I. Lucas-Kanade 20 years on: a unifying framework. *International Journal of Computer Vision* 2004; **56**(3):221–255.
6. Chu T, Ranson W, Sutton M, Peters W. Applications of digital-image-correlation techniques to experimental mechanics. *Experimental Mechanics* 1985; **42**:232–244.
7. Sutton M, Mingqi C, Peters W, Chao Y, McNeill S. Application of an optimized digital correlation method to planar deformation analysis. *Image and Vision Computing* 1986; **4**:143–150.
8. Sutton MA, Turner JL, Bruck HA, Chae TA. Full-field representation of discretely sampled surface deformation for displacement and strain analysis. *Experimental Mechanics* 1991; **31**(2):168–177.
9. Wattrisse B, Chrysochoos A, Muracciole JM, Némoz-Gaillard M. Analysis of strain localization during tensile tests by digital image correlation. *Experimental Mechanics* 2001; **41**:29–39.
10. Wang Y, Lava P, Coppieters S, Houtte P, Debruyne D. Application of a multi-camera stereo DIC set-up to assess strain fields in an Erichsen test: methodology and validation. *Strain* 2013; **49**:190–198.
11. Bruck H, McNeill S, Sutton M, Peters W. Digital image correlation using Newton-Raphson method of partial differential correction. *Experimental Mechanics* 1989; **29**:261–267.
12. Vendroux G, Knauss WG. Submicron deformation field measurements: part 2. Improved digital image correlation. *Experimental Mechanics* 1998; **38**(2):86–92.
13. Cheng P, Sutton M, Schreier H, McNeill S. Full-field speckle pattern image correlation with B-spline deformation function. *Experimental Mechanics* 2002; **42**:344–352.
14. Abanto-Bueno J, Lambros J. Experimental determination of cohesive failure properties of a photodegradable copolymer. *Experimental Mechanics* 2005; **45**:144–152.
15. Wang YQ, Sutton MA, Bruck HA, Schreier HW. Quantitative error assessment in pattern matching: effects of intensity pattern noise, interpolation, strain and image contrast on motion measurement. *Strain* 2009; **45**:160–178.
16. Tong W. Formulation of Lucas-Kanade digital image correlation algorithms for non-contact deformation measurements: a review. *Strain* 2013; **49**(4):313–334.
17. Hild F, Roux S. Comparison of local and global approaches to digital image correlation. *Experimental Mechanics* 2012; **52**:1503–1519.
18. Besnard G, Hild F, Roux S. Finite-element displacement fields analysis from digital images: application to Portevin-LeChâtelier bands. *Experimental Mechanics* 2006; **46**:789–803.
19. Hild F, Roux S. Digital image correlation: from displacement measurement to identification of elastic properties – a review. *Strain* 2006; **42**:69–80.
20. Réthoré J, Hild F, Roux S. Extended digital image correlation with crack shape optimization. *International Journal for Numerical Methods in Engineering* 2008; **73**:248–272.
21. Hild F, Roux S, Gras R, Guerro N, Marante M., Flórez-López J. Displacement measurement technique for beam kinematics. *Optics and Lasers in Engineering* 2009; **47**:495–503.
22. Réthoré J, Roux S, Hild F. An extended and integrated digital image correlation technique applied to the analysis of fractured samples. *European Journal of Computational Mechanics* 2009; **18**:285–306.
23. Guery A, Latourte F, Hild F, Roux S. Characterization of SEM speckle pattern marking and imaging distortion by digital image correlation. *Measurement Science and Technology* 2014; **25**(1):15401.
24. Mathieu F, Hild F, Roux S. Image-based identification procedure of a crack propagation law. *Engineering Fracture Mechanics* 2012; **103**:48–59.
25. Réthoré J, Muhibulla J, Elguedj T, Coret M, Chaudet P, Combescure A. Robust identification of elasto-plastic constitutive law parameters from digital images using 3d kinematics. *International Journal of Solids and Structures* 2013; **50**:73–85.
26. Mathieu F, Leclerc H, Hild F, Roux S. Estimation of elastoplastic parameters via weighted FEMU and integrated-DIC. *Experimental Mechanics* 2014; **55**(1):105–119. DOI: 10.1007/s11340-014-9888-9.
27. Neggers J, Hoefnagels JPM, Geers MGD, Hild F, Roux S. Time-resolved integrated digital image correlation. *International Journal for Numerical Methods in Engineering*; **103**(3):157–182. under review.
28. Bornert M, Brémand F, Doumalin P, Dupré JC, Fazzini M, Grédiac M, Wattrisse B. Assessment of digital image correlation measurement errors: methodology and results. *Experimental Mechanics* 2008; **49**(3):353–370.
29. Bornert M, Brémand F, Doumalin P, Dupré JC, Fazzini M, Grédiac M, Wattrisse B. Addendum to: assessment of digital image correlation measurement errors: methodology and results. *Experimental Mechanics* 2013; **49**(3):11340.