On the use of adaptive refinement in isogeometric digital image correlation

Citation for published version (APA):

DOI:
10.1002/nme.4952

Document status and date:
Published: 07/12/2015

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 28. Sep. 2023
On the use of adaptive refinement in isogeometric digital image correlation

S. M. Kleinendorst, J. P. M. Hoefnagels*,†, C. V. Verhoosel and A. P. Ruybalid

Department of Mechanical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

SUMMARY

A novel global digital image correlation method was developed using adaptive refinement of isogeometric shape functions. Non-uniform rational B-spline shape functions are used because of their flexibility and versatility, which enable them to capture a wide range of kinematics. The goal of this work was to explore the full potential of isogeometric shape functions for digital image correlation (DIC). This is reached by combining a global DIC method with an adaptive refinement algorithm: adaptive isogeometric GDIC. The shape functions are automatically adjusted to be able to describe the kinematics of the sought displacement field with an optimized number of degrees of freedom. This results in an accurate method without the need of making problem-specific choices regarding the structure of the shape functions, which makes the method less user input dependent than regular global DIC methods, while keeping the number of degrees of freedom limited to realize optimum regularization of the ill-posed DIC problem. The method’s accuracy is demonstrated by a virtual experiment with a predefined, highly localized displacement field. Real experiments with a complex sample geometry demonstrate the effectiveness in practice. Copyright © 2015 John Wiley & Sons, Ltd.

Received 21 December 2014; Revised 27 March 2015; Accepted 6 May 2015

KEY WORDS: digital image correlation; isogeometric analysis; adaptivity; hierarchical splines

1. INTRODUCTION

Digital image correlation (DIC) is a computational technique used for determining displacement fields. DIC was first developed in the 1980s and introduced in the field of solid mechanics by Peters [1], Sutton [2] and Chu [3]; see Pan et al. [4] for an extensive review on the development of the digital image correlation method. DIC has become a standard technique for material scientists to couple experiments and numerical simulations by analyzing displacement fields from mechanical tests, from which strain fields can be determined.

In the DIC algorithm, a system of equations needs to be solved in order to obtain the two-dimensional displacement field. In this system, the unknowns are the vector displacements per pixel. The known variable is the brightness field, which is a scalar value per pixel. Because the number of unknowns exceeds the number of equations, the problem is ill-posed by nature. To overcome this problem, the displacement field is regularized with so-called shape functions. The coefficients associated with the basis functions in this discretization are referred to as degrees of freedom (DOFs).

There are several possibilities for this regularization. A distinction is made between a local and a global approach. In the local method, the region of interest is divided in small subsets, zones of interest, which are correlated in both images in order to find their displacement. The total

*Correspondence to: J. P. M. Hoefnagels, Eindhoven University of Technology, Gemini-Zuid 4.122, 5600MB Eindhoven, The Netherlands.
†E-mail: j.p.m.hoefnagels@tue.nl

Copyright © 2015 John Wiley & Sons, Ltd.
displacement field then is interpolated to obtain full-field information. In the global DIC (GDIC) method, the correlation of the two images is executed over the entire domain at once, hence the term 'global' [5–7]. Typically, continuity of the displacement field is imposed (except when discontinuities are deliberately introduced [8, 9]), making GDIC less sensitive to image noise; see e.g., Hild et al. [10].

In contrast to the local method, where displacements are calculated at every subset center, the global approach provides a displacement at every pixel. Therefore, the information density is generally higher. However, because of the higher amount of regularization, the global method is more reliant on the choice of certain important variables the user makes. In particular, the choice of the shape functions is critical, because these must be able to capture the kinematics of the sought displacement field. This choice can be difficult, as the displacement field is generally unknown in advance. Therefore, GDIC is able to provide more accurate results, but only if the user supplies the algorithm with the correct problem-specific kinematic description.

An interesting type of shape functions is non-uniform rational B-splines (NURBS). NURBS originate from computer-aided design (CAD) modeling, where they are the industry standard. NURBS are able to represent many geometric shapes—in particular conic sections—exactly, whereas standard finite element shape functions approximate them. Hughes et al. proposed to use the geometrical representation of CAD-models directly for finite element analysis: isogeometric analysis [11, 12]. The use of B-splines in a DIC setting was pioneered by Cheng [5]. Recently, Beaubier et al. directly utilized the CAD-representation of their subject for the DIC algorithm [13]. Advantages of NURBS were investigated by Elguedj et al. [14], who concluded that, compared with high-order Lagrange shape functions, NURBS need fewer functions to adequately describe the displacement field, which improves the conditioning of the problem and thereby the noise robustness of the correlation algorithm.

To explore the full potential of NURBS shape functions for DIC, this paper presents a global DIC method that adaptively adjusts the shape functions to capture the kinematics required to compute the displacement field. To this end, GDIC is combined with an adaptive refinement algorithm for the NURBS shape functions: adaptive iso-GDIC. As will be shown, the flexibility of the shape functions and its ability to perform automatic adaptive refinements make the novel approach suitable for various types of DIC problems. Most importantly, it provides an accurate method without the need of supplying the algorithm with specific prior information about the problem. An optimum number of DOFs is pursued, such that enough freedom is allowed to capture the kinematics of the problem, but not too much freedom, so noise robustness is assured. This is an advantage over the existing DIC methods because it combines high-strain field accuracy with user-friendliness and a low level of expertise.

This paper is organized as follows. In Section 2, the isogeometric GDIC (iso-GDIC) methodology is discussed. Here, the two-dimensional global DIC algorithm is introduced, and NURBS shape functions are elaborated, as well as the isogeometric parametrization of the geometry. In Section 3, a virtual experiment is executed to test the performance of NURBS shape functions for a GDIC problem, in comparison with other, more commonly used types of shape functions. The adaptive refinement algorithm and the choices made in its design are detailed in Section 4. In Section 5, a virtual experiment is executed in order to give a proof of concept of the iso-GDIC algorithm with adaptive refinement. Furthermore, the results of this test-case are compared with conventional, non-adaptive DIC algorithms. In Section 6, real experiments with a complex specimen geometry are carried out to explore the method’s performance in practice, after which general conclusions are drawn in Section 7.

2. METHODOLOGY FOR ISOGEOMETRIC GLOBAL DIGITAL IMAGE CORRELATION

In this section, the characteristics of iso-GDIC, are discussed. First, the two-dimensional GDIC algorithm is explained. Next, the isogeometric shape functions, NURBS, are introduced. Besides using NURBS shape functions for the discretization of the displacement field, they are also used for the parametrization of the geometry, which is clarified in the last part of this section.
2.1. Two-dimensional global digital image correlation algorithm

The goal of GDIC is to quantify the displacement field of a deformed specimen by correlating digital images of the specimen. The starting point for a GDIC problem is the reference image and the deformed image. The images are characterized by the intensity of the individual pixels, which is represented by the scalar field $f(x_0)$ for the reference image and $g(x)$ for the image corresponding to the deformed specimen. Material points are assumed to retain the same intensity upon deformation and since every material point needs to be preserved, brightness conservation holds:

$$f(x_0) - g(x_0 + u(x_0)) = \Psi(x_0) \approx 0. \quad (1)$$

Herein, $\Psi(x_0)$ is the residual image, which is zero for all positions for the case that the displacement field $u(x_0)$ is calculated perfectly and no noise is present in the images. In GDIC, the optimal correlation is obtained when minimizing $\Psi(x_0)$ using the least squared approach:

$$u(x_0) = \text{arg min} \int_{\Omega_0} \left[ f(x_0) - g(x_0 + u(x_0)) \right]^2 \, dx_0. \quad (2)$$

Because this is an ill-posed problem, by virtue of the fact that a vector-valued displacement field is sought, while only scalar brightness values are known, and in general, the number of pixels exceeds the number of unique brightness values, the displacement field is regularized as follows:

$$u(x_0) \approx u^*(x_0, \lambda) = \sum_{j=1}^{m} \lambda_j \phi_j(x_0), \quad (3)$$

where $\lambda_j$ are the DOFs and $\phi_j(x_0)$ are their associated shape functions. As explained earlier, many different choices are possible for the shape functions, which is a challenging decision. The shape functions must be able to capture the (often highly inhomogeneous) kinematics of the sought displacement field, which is generally unknown in advance.

Combining equations 2 and 3, the discretized problem is obtained:

$$\begin{cases} 
\Psi(x_0, \lambda) = f(x_0) - g(x_0 + u^*(x_0, \lambda)) \\
\lambda = \text{arg min} \int_{\Omega_0} \Psi(x_0, \lambda)^2 \, dx_0.
\end{cases} \quad (4)$$

This problem is iteratively solved by optimizing the unknowns $\lambda$ by a Gauss–Newton scheme, yielding the following system of equations:

$$M \delta \lambda = b \quad (5)$$

where $\delta \lambda$ is the DOFs update and

$$M_{ij} = \int_{\Omega_0} \left[ (\nabla_0 f \cdot \phi_i) (\nabla_0 f \cdot \phi_j) \right] \, dx_0 \quad (6a)$$

$$b_j = \int_{\Omega_0} \left[ (\nabla_0 f \cdot \phi_j(x_0)) [f(x_0) - g(x(x_0, \lambda))] \right] \, dx_0. \quad (6b)$$

The calculation of the domain integral for $M_{ij}$ and $b_j$ is executed by summation over all pixels, as Gauss quadrature schemes, traditionally used in finite element calculations, cannot be used for DIC because of the highly irregular pattern of the images [7]. Because the image gradient $\nabla_0 f$ and the gray level values $f$ and $g$ are naturally already known at each pixel location, the shape functions $\phi$ are also evaluated there. To this end, a nearest neighbor inverse mapping algorithm is employed, which was found to provide a good combination of sufficient accuracy, robustness and computational speed. This algorithm locates for each pixel the closest local, parameterized coordinate pair and assigns the values of the shape functions at this location to the pixel.
2.2. Non-uniform rational B-splines shape functions

In this work, the potential of NURBS as shape functions for DIC (equation 3) is explored. B-splines are piecewise polynomial shape functions that are defined recursively by the Cox–de Boor relation [15, 16]:

\[
N_{i,0}(\xi) = \begin{cases} 
1 & \text{if } \xi_i \leq \xi < \xi_{i+1} \\
0 & \text{otherwise}
\end{cases} \quad (7a)
\]

\[
N_{i,p}(\xi) = \frac{\xi - \xi_i}{\xi_{i+p} - \xi_i} N_{i,p-1}(\xi) + \frac{\xi_{i+p+1} - \xi}{\xi_{i+p+1} - \xi_{i+1}} N_{i+1,p-1}(\xi), \quad (7b)
\]

where \( \xi \) is the coordinate in the parametric domain. A graphical interpretation is shown in Figure 1a. The basis functions are defined over a knot vector \( \Xi = \{\xi_1, \xi_2, \ldots, \xi_n\} \) and \( Z = \{\xi_1, \xi_2, \ldots, \xi_3, \xi_3\} \). (a) B-splines in the one-dimensional domain. The knots on the parametric domain are given by \( \xi_i \); (b) A B-spline shape function on the two-dimensional domain. Only one shape function is shown in this representation.

NURBS are rationalized B-splines, which are defined as:

\[
R_i(\xi) = \frac{N_i(\xi)W_i}{w(\xi)}, \quad (8)
\]

where \( W_i \) are scalar control point weights and \( w(\xi) = \sum_{i=1}^{n} N_i(\xi)W_i \) is the weighting function. NURBS are able to represent many shapes exactly, in particular conic sections. However, in this work, all weights \( W_i \) are set to 1, as it was found that B-splines are sufficiently versatile to parametrize the sample geometries. However, the method can easily be extended to applications where non-unity weight factors are required. Furthermore, in this work, so-called open B-splines are used, where the first and last knot are repeated \( p + 1 \) times. However, we do use non-uniform B-splines, such that the knots are not restricted to a uniform distribution.

When taking the tensor product of the shape functions in two directions, with local parametric coordinates \( \xi \) and \( \zeta \), two-dimensional NURBS shape functions are obtained:
Figure 2. A non-uniform rational B-spline (NURBS) curve (a) and NURBS surface (b) $x(\xi)$, constructed from the control net $P$ and the shape functions. The control points are indicated by red dots. (a) NURBS curve; (b) NURBS surface mesh.

\[ \phi_A(\xi) = R_i(\xi)R_j(\xi), \]  

with $A = A(i, j)$ and $\xi = (\xi, \zeta)$. An example of a two-dimensional NURBS shape function is shown in Figure 1b.

2.3. Geometry parametrization

In isogeometric analysis, NURBS are used for both discretization of the sought displacement field and parametrization of the geometry of the sample [11, 12]. The mapping of the parametric coordinate $\xi$ to the coordinate $x$ in the physical domain is carried out by defining control points $p_A$, such that

\[ x(\xi) = \sum_A \phi_A(\xi)p_A = P^T\phi(\xi). \] 

In Figure 2a, an example is shown of a B-spline curve $x(\xi)$, using the shape functions of Figure 1a and the control points $P$, as indicated by the red dots in Figure 2a. In Figure 2b, a B-spline surface mesh is shown, constructed in the same manner, using shape functions in 2D-space, as represented in Figure 1b, and the control net as indicated by the red dots in Figure 2b.

Isogeometric analysis is a powerful tool in representing a large range of displacement fields. The advantage of iso-GDIC is twofold. Firstly, complex sample geometries can be described accurately, and the possibility exists to directly use a CAD-model for the analysis, as Beaubier suggests [13]. This is particularly useful for complex geometries. Secondly, continuity of the shape functions over element boundaries can be controlled by means of knot insertion, which allows for a great flexibility. The multiplicity $m$ of a certain knot determines the continuity of the shape functions at that location, i.e., the continuity equals $C^{p-m}$. This is convenient to represent displacement fields that incorporate lowered continuity, e.g., as resulting from highly localized deformation in a material. If every knot is repeated $p$ times, the element boundaries are $C^0$ continuous, which resembles the situation where traditional finite element shape functions are used, making the approach equivalent to that of Besnard et al. [7]. Furthermore it can easily be shown that if only one element is used, B-splines are equivalent to regular monomials and other globally defined polynomial shape functions up to order $p$ that are expected to work best for smooth, regular displacement fields. Thus, NURBS shape functions provide a generic manner of describing many different forms of displacement fields and form a superset of other types of shape functions, including Lagrange shape functions (traditionally used in the finite element method (FEM)) and globally defined polynomials.
2.4. Blurring step

In (global) DIC approaches, it might occur that the method converges to a local minimum, as opposed to the global minimum (equation 2), leading to an erroneous solution. In order to prevent this, for each increment, the DIC problem is first solved for a blurred version of the images $f$ and $g$, using a Gaussian kernel (size $10 \times 10$ pixels, standard deviation 1.5 pixels). The result is used as an initial guess for the non-blurred images. This method has proven to reduce the bias error of the DIC method [17]. In literature also, other methods to solve this problem are known. Multi-grid algorithms have been used where the DIC problem is, in $n$ steps, first solved for 'coarse-grained' versions of the images [7, 8], i.e., the images are decomposed into super-pixels of $2^n \times 2^n$ pixels by averaging the gray level values of these pixels.

3. ISOGEOMETRIC GLOBAL DIGITAL IMAGE CORRELATION RESULTS AND ANALYSIS

The performance of NURBS shape functions for GDIC is first explored by a virtual experiment, where a comparison with Lagrange (traditional FEM) shape functions [18] and globally defined Chebyshev polynomials [19] is made. Comparison is based on the root mean square (RMS) value of the exact error:

$$
\varepsilon_{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (u_i - u_{i,ref})^2},
$$

with $n$ the total number of pixels. In this experiment, an image $f$ with a random grayscale speckle pattern is subjected to a predefined displacement field $u_{ref}$ to obtain the deformed image $g$. The displacement field, which mimics a localization of the strain field, is expressed by:

$$
u_{ref}(x, y) = a \arctan (x - b),
$$

where $a = u_{max}/\arctan\left(\frac{1}{2}(x_{max} + x_{min})\right)$, such that the displacement field range is $[-u_{max}, u_{max}]$, and $b = \frac{1}{2}(x_{max} - x_{min})$, such that the localization occurs in the center of the domain. The applied displacement field is plotted in Figure 3.

In order to make an objective comparison between the three types of shape functions, they should be compared for the same number of DOFs. Therefore, the RMS value of the error in the displacement field is plotted as a function of the number of DOFs in $x$-direction in Figure 4a. The number of DOF’s in $y$-direction is three (one element with three shape functions) for all situations, because the localization in the displacement field only occurs in $x$-direction. As expected, the error is smaller for a larger number of DOFs. This is because of the complexity of the displacement field: a higher number of shape functions is needed to describe the kinematics, resulting in a lower discretization error. As can be seen, a distinction is made for Chebyshev shape functions between even and odd

![Figure 3. The applied displacement field $u_{ref}$.](image)
Figure 4. Root mean square (RMS) value of the error in the displacement field as a function of the number of degrees of freedom (DOFs) (a) and as a function of noise level (b), plotted for different types of shape functions. The noise level is relative to the amplitude of the gray values of the deformed image $g$. The combinations of shape functions with an equal number of DOFs that are used in the noise analysis are indicated by ellipses in (a). Note that for a clear graphical representation of graph (b) combinations have been chosen for which the RMS-error of the three types of shape functions does not overlap (as is the case for e.g., 15 DOFs). To do so, the combination always consists of even Chebyshev polynomials, element-centered NURBS, and vertex-centered Lagrange shape functions, resulting in the lowest RMS-error for the Lagrange type. However, for the noise sensitivity, the absolute error value is not as important as the trend and the variance in the error, thereby justifying this choice.

For polynomials of even order, an extremum occurs at the center of the domain ($x = 0$), while for polynomials of odd order, an inflection point occurs at this location. Therefore, Chebyshev functions of an odd polynomial order are better able to capture the steep slope of the arctangent displacement field, which also shows an inflection point at $x = 0$. The error for the odd Chebyshev polynomials is therefore smaller than for the even Chebyshev shape functions, especially for a lower number of DOFs, while for a higher number of DOFs, the influence of the even polynomials is canceled out by the odd polynomial contribution to the solution. For Lagrange shape functions and NURBS, a similar distinction is made between element-centered and vertex-centered meshes. In an element-centered mesh, the number of elements is odd, which causes the center of an element to occur at the center of the domain. For a vertex-centered mesh, the number of elements is even, which leads to a vertex, i.e., a node for Lagrange shape functions or a knot for NURBS, located at $x = 0$. A vertex induces a line of lowered continuity ($C^0$ for Lagrange and $C^{p-m}$ for NURBS), which enables the shape functions to better capture the discontinuity in this particular displacement field. Therefore, the RMS-error is lower for the vertex-centered shape functions, as can be seen in Figure 4a.

Also, the influence of noise on the performance of the different methods is investigated. To this end, the virtual experiments are repeated for different levels of artificially imposed noise on the deformed image: a noise level of 100% means that the noise amplitude is of the same order as the gray level values. Three different sets of a constant number of DOFs are used for the three types of shape functions. The RMS value of the error in the displacement field and its standard deviation (error bars) are shown in Figure 4b as a functions of the noise level. It can be seen that for a higher number of DOFs the method is more sensitive to noise: at a certain noise percentage, the method does not converge anymore. This is a well-known property of GDIC [14, 20]. Furthermore, Chebyshev shape functions are observed to be more sensitive to noise than the other two types; as for higher polynomials orders, they do not converge for noise percentages over 20%, while NURBS and Lagrange shape functions go up to 40–50%. The difference in noise-sensitivity between Lagrange and NURBS shape functions is observed to be small. The uncertainty in the RMS-error (error bars in the graph) is slightly smaller for the NURBS, and also, the increase in error is slightly smaller.
Figure 5. Root mean square value of the error in the displacement field plotted against the multiplicity $m$ of the knot in the center of the $x$-domain. An example of the rearrangement of the knot vector in order to reach these multiplicities is shown in the inset on the right side of the graph. In this case, we see a knot vector in $x$-direction with 12 knots, for shape functions of order 2 (the multiplicity of the knots at the edges is $p + 1 = 3$). The knots are redistributed such that the total number of knots, and hence degrees of freedom, remains the same. In the center of the domain, the multiplicity of the knot is increased.

However, from this experiment, we conclude that there is no significant difference in performance between the two types of shape functions that are more locally defined, i.e., NURBS and Lagrange shape functions. The shape of the mesh is more important than the type of the shape functions.

In order to illustrate the influence of the continuity of the shape functions rather than the type, the experiment is repeated, using only NURBS shape functions; however, for a given number of DOFs, the knots along the $x$-axis are redistributed in order to increase the multiplicity of the knot in the center of the domain, where the high gradient in the displacement field occurs, as shown in the inset in Figure 5. The virtual experiment is executed for four sets of constant number of DOFs in $x$-direction: 7, 9, 13, and 17. In each set, the knot vector is rearranged as described earlier. The four sets are again compared by means of the RMS value of the error in the displacement field.

The error is plotted against the multiplicity of the knot in the center of the domain in Figure 5. From this graph, it is again observed that the error decreases if the amount of DOFs increases, because the complex kinematics of this type of displacement field are difficult to capture with polynomial shape functions and require more DOFs. This time, however, it is noted that a significant improvement is achieved if the multiplicity of the center knot increases. This can be explained by the fact that the continuity decreases when knots are repeated. The applied displacement field approaches a step function with $C^{-1}$ continuity in the center of the domain. If the continuity of the shape functions is decreased as well, they are better able to approximate the displacement field.

These virtual experiments illustrate that adding DOFs is particularly useful if they are added at the right location. Moreover, the actual placement of the knots is more important than the amount of DOFs and even more important than the type of shape functions used. An advantageous situation occurs when the number of DOFs remains limited because this enhances noise robustness, and still, complex kinematics can be described. Naturally, the displacement field is unknown in advance, and the user is generally unable to provide the GDIC algorithm with the most suitable mesh. Therefore, in order to optimally benefit from the flexibility of the shape functions used in isogeometric analysis, we propose to combine the iso-GDIC algorithm with adaptive mesh refinement. Note that adaptive refinement could also be combined with a GDIC algorithm using other types of shape functions; however, NURBS are a more general approach because they form a superset of the other types and, furthermore, have the advantage of being able to describe complex geometries exactly.

4. ADAPTIVE REFINEMENT

We have seen that on one hand, NURBS shape functions provide a flexible tool in describing both simple and complex kinematics of displacement fields, and on the other hand, optimized placement
of DOFs in a DIC problem is computationally favorable. Therefore, it is expected that iso-GDIC combined with an adaptive refinement algorithm yields a powerful method for solving a large variety of DIC problems. In this section, a two-dimensional adaptive refinement algorithm of NURBS shape functions is explained, as well as some important choices on criteria used in the algorithm.

4.1. Hierarchical mesh refinement

A hierarchical approach for refining the mesh is employed; see [21] for details. The key idea to hierarchical refinement is to replace some of the shape functions in the initial basis by shape functions of the refined basis. The bases are constructed hierarchically, which implies that multiple levels of basis functions exist, representing subsequent levels of refinement of the underlying geometry. This is illustrated in one dimension in Figure 6. In Figure 6(a), the initial basis is shown and in Figure 6(b), the refined basis.

Now we could mark elements for refinement; however, in hierarchical mesh refinement, this strategy is not always suitable [22]. It can occur that selecting an element for refinement does not result in any refinement. For example, if the second element in Figure 6a is selected, there are no shape functions in the initial basis nor the refined basis that completely lie within this domain. Hence, no shape functions are replaced and no refinement occurs. Therefore, it is more suitable to mark shape functions for refinement rather than elements.

Suppose, for example, that shape functions 1 and 7 are selected for refinement, as indicated by the dashed lines in Figure 6a. They are replaced by the shape functions of the refined basis that completely lie within the same domain, as indicated by the solid lines in Figure 6b. The resulting set of shape functions is shown in Figure 6c. It should be noted that in the one-dimensional case, the same effect could be obtained by knot insertion. In the two-dimensional case, the refinement remains local; this in contrast to tensor product B-splines.

4.2. Refinement criteria

In order to determine in which area a finer mesh is required, i.e., which shape functions must be refined, an error estimator is needed. In DIC, the accuracy of the method is based on the residual

![Figure 6](image-url)

Figure 6. The concept of hierarchical refinement. In the top figure, an initial basis is shown, and in the middle figure, a uniformly refined basis. The dashed basis functions in the initial basis are replaced by the solid shape functions of the refined basis. In the lower figure, the result of this refinement step is shown. Element boundaries (knot locations) are indicated by vertical black dashed lines.
ψ(x₀) = f(x₀) - g(x₀ + u(x₀)). Therefore, we propose to let the selection of shape functions be based on the following refinement indicator:

\[
C_j = \frac{1}{(f_{max} - f_{min})} \frac{\delta_{f,global}}{\delta_{f,j}} \frac{\int_{\Omega_j} |\psi(x_0)| \phi_j(x_0) dx_0}{\int_{\Omega_j} \phi_j(x_0) dx_0}.
\]

In this expression, \(j\) is the index of a shape function \(\phi_j\) and \(\Omega_j\) indicates its area of support. The refinement indicator is based on the residual \(\psi(x_0)\) and is scaled with the shape function itself, in order to prevent shape functions with a larger support always being favored for refinement and to scale the influence of the residual with the value of the shape function. Furthermore, the indicator is scaled with the mean intensity gradient \(\delta_f = \frac{1}{mn} \sum_{i=1}^{m} \sum_{k=1}^{n} \sqrt{\sum_{f} |\nabla f(x_{i,k})|}\) of the entire image (\(\delta_{f,global}\)) divided by the mean intensity gradient of the area underneath the shape function (\(\delta_{f,j}\)). This diminishes the influence of contrast differences in different areas in the pattern. The indicator is also normalized with the range of gray values (\(f_{max} - f_{min}\)). This makes the indicator independent on the absolute gray values, which allows for a general criterion based on, for example, the noise level.

It is important that refinement does not take place if the residual due to correlation mismatch is lower than the noise amplitude. The noise level heavily depends on the measurement system used to capture the images, i.e., the camera, lighting conditions, microscope lenses, and so on. Therefore, the value of the threshold should be determined for every experimental data set separately. The noise level can be evaluated by correlating two or more images with no deformation and determining the average residual with respect to the gray value amplitude. Moreover, a safety factor should be included in the threshold. This safety factor cannot be too low, in order to make sure the residual is above noise level, but must not be too high, because aberrant features must not be neglected. Based on extensive experimentation, a safety factor of 4 is recommended. In this work, a noise level of about 1.5% was found for the images taken with the measurement system used, and thus, a normalized residual below 6% is assumed to be acceptable, i.e., shape functions with \(C_j > 0.06\) are marked for refinement.

In order to optimize the selection procedure, a second criterion is taken into account. Only shape functions with a refinement indicator \(C_j\) larger than \(\bar{C} + 1.5\sigma\) are refined, with \(\bar{C}\) the mean of the indicators of all shape functions and \(\sigma\) the standard deviation. This criterion makes the algorithm suitable for problems where no refinement is required. Moreover, this is useful in an experimental setting, where typically, a sequence of images is correlated. The first couple of images may embody merely homogeneous deformation where no refinement is desirable. Refinement only takes place if both criteria are met.

Finally, it might occur that the level of refinement causes elements to become smaller than a few pixels. This has a negative influence on the conditioning of the problem and increases the noise-sensitivity (Section 3). In order to prevent this, the pattern of image \(f\) is evaluated to determine the correlation length \(\ell_c\). The correlation length is an in-plane length scale of the pattern and is defined as the length where the autocorrelation peak of image \(f\) equals a certain threshold, here taken as \(\frac{1}{e}\), where \(e\) is Euler’s number. Elements should contain at least several pattern features; therefore, a threshold is placed at \(10\ell_c \times 10\ell_c\) pixels. If an element is smaller than this size, all shape functions associated to this element are marked to have reached the highest level of refinement, and no further refinement is allowed.

5. ANALYSIS OF ADAPTIVE ISOGEOMETRIC GLOBAL DIGITAL IMAGE CORRELATION

A virtual experiment is executed to give a proof of concept of the adaptive iso-GDIC method. Moreover, the virtual experiment allows for comparison of the new method with that of an iso-GDIC algorithm that does not utilize adaptive refinement, and a local DIC implementation. In this test-case, an image of a speckle pattern on a tensile bar from a real experiment is artificially deformed in order to obtain image \(g\).
Figure 7. Evolution of the applied displacement as a function of the \( x \)-coordinate. The displacement is applied in four increments, starting from a linear displacement, towards the final arc-tangent field.

Figure 8. Evolution of the image and the mesh after applying the displacement. The reference image \( f \) is shown along with intermediate image \( g_3 \) and the final deformed image \( g_4 \). The evolved meshes are shown on top of the images. (a) \( f \), # DOFs = 35; (b) \( g_3 \), # DOFs = 62; (c) \( g_4 \), # DOFs = 130.

The displacement field \( u \) in \( x \)-direction again represents a tensile experiment with strong localization of the displacement, approaching a step function with \( C^{-1} \) continuity, and is given by equation 12. The final displacement is applied in four increments, which are illustrated in Figure 7. The corresponding reference image \( f \), intermediate image \( g_3 \), and final image \( g_4 \) are shown in Figure 8.

It is recognized that this type of displacement generally is challenging for global DIC methods (except when shape functions are used that are able capture discontinuities [8, 9]), while local DIC methods are better equipped for handling discontinuities because continuity of the solution is not imposed. The displacement is applied in four increments, ranging from a linear displacement field to the final arc-tangent displacement. All four reference displacement fields in \( x \)-direction, \( u_{\text{ref}} \), are plotted on the reference image \( f \) in Figure 9a. The displacement \( v \) in \( y \)-direction is zero everywhere in the image domain and disregarded in the analysis.

In order to let the virtual experiment optimally resemble a real experiment, a specific procedure for generating the synthetic images is followed; see Neggers [23]. Here, the virtual sample space, on which the artificial displacement field is defined, is larger than the field of view, i.e., the final image size. Starting from an initial image, i.e., virtual sample space, of \( 1300 \times 1018 \) pixels, the size of the final image is chosen \( 650 \times 509 \) pixels.

The iso-GDIC algorithm with adaptive refinement, as described in the previous section, is applied to this problem. NURBS shape functions of the second-order are used. In Figure 8, the evolution of the mesh is depicted. It can be seen that refinement indeed takes place in the central area, where the large gradient in the applied displacement field occurs. It is noted that the region of refinement is wide initially. This is because of the fact that shape functions are refined rather than elements: the
The entire support of the shape function, which consist of multiple elements, is refined. The calculated displacement fields for all four increments are shown in Figure 9b.

In order to investigate the influence of adaptive refinement, the virtual experiment is repeated without adaptive refinement. To be able to perform a comparison as fair as possible, the amount of DOFs should be the same for both cases. For the non-adaptive iso-GDIC method, a uniform mesh of 10 by 9 elements is chosen and second-order NURBS shape functions are used, which yields approximately the same amount of DOFs as the adaptive iso-GDIC experiment after refinement (130 DOFs for the refined problem vs. 132 DOFs for the non-adaptive DIC method). In Figure 9c, the resulting displacement field is shown.

The displacement field is also calculated using a local DIC method (using MathID software [24]), choosing optimized settings: subset size 21 pixels, step size 7 pixels, SSD correlation coefficient, bicubic spline interpolation, and affine shape functions. The correlated displacement field is depicted in Figure 9d. It can be observed that for the last two increments, the displacement is not correlated in the center region of the domain. This is because in local DIC, only the displacement of the subset centers is calculated and only small deformations of the subsets can be captured. The variation in displacement is large in the center area of the domain, and therefore, local DIC is not able to correlate the subsets located in this area. Note that the loss of subsets in the central area has already been minimized by using an optimum subset size and step size. Another problem that occurs in local DIC is loss of information at the top and bottom edges of the domain because subsets of which the center falls outside the domain cannot be correlated.
Figure 10. Error fields for the four increments and different digital image correlation (DIC) methods. The error is defined as the difference between the reference displacement field and the calculated displacement field $\varepsilon = u - u_{ref}$ and expressed in terms of pixels. The four rows correspond to the four increments in applied displacement, ranging from a linear, homogeneous, displacement field (top) to the final arc-tangent field (bottom). The first column (a) corresponds to the adaptive isogeometric global DIC (iso-GDIC) method, (b) to the non-adaptive iso-GDIC implementation and the last column (c) to the local DIC method. It should be stressed that finding an optimal setting for local DIC to perform best is laborious and requires experience. In the adaptive iso-GDIC algorithm, important choices are already made in the development of the method, which ensures accurate results, without much expertise of the user needed.

In Figure 9, it can be observed that in the first two increments, all methods yield approximately the same solution. In order to objectively compare the three methods, the exact error in the displacement field is determined: $\varepsilon = u - u_{ref}$. In Figure 10, the error fields are shown. In the first two increments, the displacement field is quite homogeneous, and it is observed that no large difference in the error field exists between the three methods. In the third increment, the local method starts to fail: the concentration of displacement is too large and subsets in the center region are not able to correlate. This becomes even more apparent in the final increment, where the steep arc-tangent displacement field is applied. Furthermore, in this increment, a difference between the adaptive method and the non-adaptive method is observed. The refined shape functions are better able to capture the high gradient in displacement in the center of the domain than the non-adaptive isogeometric shape functions, resulting in a smaller region where the error is high. This indicates that the novel, adaptive method offers a more accurate solution than the non-adaptive method for this problem.
while the number of DOFs, and hence the computational effort, is similar for both methods. Moreover, the adaptive iso-GDIC method performs well for the entire range of displacement fields. However, the real strength of the novel method is that this solution was obtained without supplying problem-specific information to the program beforehand: the area where a finer mesh is required is determined autonomously.

6. APPLICATION IN EXPERIMENTAL SETTING

Real experiments are performed in order to investigate the performance of the adaptive iso-GDIC method to practical problems. Uniaxial tensile tests are performed in a microscopic setup.

6.1. Test setup and sample

The setup for the experiment consists of a tensile stage, a microscope, a camera and a computer. The microscope used is a Carl Zeiss Discovery.V20 stereo microscope with a Zeiss PlanApo S 1.5x FWD 30mm lens. The Zeiss AxioCam CCD camera is connected to a computer. Images are recorded using AxioVision 4.7 software. The tensile stage (Kammrath & Weiss tensile/compression module with a 500N load cell) is driven by a controller to make sure a tensile speed of 5 μm/s is maintained.

Two different test specimen are used. The first sample is a polycarbonate double notched 1 mm thick tensile specimen. The second sample has the same material and geometry; however, a hole is present. With this material and these geometries, strong localizations of strain and displacement are expected, which makes them a good test-case for demonstrating the adaptive refinement process. A random speckle pattern is applied using spraypaint to make the samples suitable for the DIC analysis.

6.2. Results

6.2.1. Double notched tensile sample. The reference image $f$ and the final image $g$, corresponding to the deformed specimen, are shown in Figure 11. More images are recorded during the experiment, and the correlation procedure is carried out in several increments using these intermediate images. The adaptive iso-GDIC algorithm is executed, using second-order B-spline shape functions. The resulting mesh refinement can also be seen in Figure 11.

The displacement $u$ in $x$-direction and the displacement $v$ in $y$-direction are shown in Figure 12a on the initial image $f$. It can be seen that, indeed, localization takes place in the center of the domain, as was expected from the geometry of the sample. In the bottom image of Figure 12a, it can be seen that the $y$-displacement is largest in the central area, at the bottom and top of the domain, indicating that necking occurs here. This is also the region where refinement takes place.
Figure 12. Calculated displacement fields $u$ in $x$-direction (top row) and $v$ in $y$-direction (bottom row). The displacement is expressed in terms of pixels. The calculated fields are shown for the adaptive method (a), the non-adaptive isogeometric global digital image correlation (DIC) method (b) and the local DIC algorithm (c). Note that the $y$-coordinate and its associated $v$-displacement point in the downward direction.

Figure 13. Residual image $\Psi(x_0) = f(x_0) - g(x_0 + u(x_0))$ for the adaptive iso-GDIC method at the onset of localization before refinement (a) and after refinement (b). The residual is expressed in terms of gray values.

The algorithm is also executed with a conventional GDIC implementation, using a mesh of 14 by 9 evenly spaced elements and second-order B-spline shape functions, resulting in the same number of DOFs as the final refined mesh as shown in Figure 11b. The resulting displacement field components $u$ and $v$ are depicted in Figure 12b. Because of the homogenous nature of the displacement field, the adaptive algorithm in this case only provides marginally better results than the non-adaptive method, provided the number of DOFs is equilibrated. When compared with the results using the initial mesh of 5 by 3 elements, a substantial improvement in accuracy is observed. This can be seen in Figure 13, where the residual image $\Psi(x_0)$ is compared for the adaptive method before and after refinement, using the meshes of Figures 11a and 11b, respectively. The residual is indeed highest in the central area of the sample where localization takes place. After refinement, this area is smaller than before refinement, and also, the residual level is lower, indicating an improvement of the accuracy. In the next section, we further study the accuracy improvement of adaptive iso-GDIC compared with the initial mesh.
Correlation is repeated using a local DIC method (using MatchID software, subset size 35 pixels, step size 10 pixels, Sum of Squared Differences (SSD) correlation coefficient, bicubic spline interpolation, and irregular shape functions). Note, however, that settings, e.g., subset size and interpolation scheme, have again been optimized to yield the lowest loss of subsets as possible. The results are shown in Figure 12c. It can be seen that in the area where localization takes place, the local DIC method is not able to correlate all subsets. This is because the variation in displacement within these subsets is too large and the displacement of the subset center cannot be found.

This indicates that adaptive refinement is a useful tool in solving this type of problems. But above all, the solution was found automatically: the algorithm was not supplied with specific foreknowledge about the problem, as usually is needed for a global DIC method to perform well in this type of problem.

### 6.2.2. Double notched tensile sample with hole.

The tensile sample before and after deformation can be seen in Figure 14. These images, along with a number of intermediate images taken during the experiment, are used in the correlation process as reference image \( f \) and deformed image \( g \). The meshing procedure is similar to that described in Section 2.3; however, some additional steps are taken in order to make the mesh conforming the hole. Four elements and their corresponding shape functions, which are originally located in or near the hole, are removed. Next, the control points of the elements surrounding the hole are translated to the edge of the hole, such that the mesh forms a fit around the hole. In order to describe the hole accurately, at least 20 elements around the hole are required, resulting in a quite fine initial mesh; see Figure 14a. This is a simple meshing procedure, resulting in a parametrization that is \( C^1 \) continuous everywhere. Other, more complex, parametrization techniques, including coupled NURBS patches, may also be used. As was investigated in Section 3, the accuracy in a localized displacement field, which is to be expected from this sample geometry, can be improved by using lower continuity lines at the localization positions, which could be achieved by a multipatch approach with \( C^0 \) coupling. However, in this study, we want the algorithm to obtain the optimal solution autonomously, without assuming localized behavior in advance.

The correlation procedure is executed using adaptive refinement of second-order NURBS shape functions. The adaptively refined mesh is shown in Figure 14b. It can be seen that refinement takes place in the area where the largest deformation, i.e., localization, is expected; the least wide material regions.

The final calculated displacement fields \( u \) in \( x \)-direction and \( v \) in \( y \)-direction can be seen in Figure 15a. From the deformed image, it is observed that necking occurs in the small regions around the hole. Indeed in the bottom image in Figure 15a, displacement field \( v \), this localization is visible: the displacement in \( y \)-direction is largest in this area. Furthermore, a localization in \( x \)-displacement...
Figure 15. Calculated displacement fields $u$ in $x$-direction (top row) and $v$ in $y$-direction (bottom row). The displacement is expressed in terms of pixels. The calculated fields are shown for the adaptive method (a), the non-adaptive isogeometric global digital image correlation (DIC) method (b) and the local DIC algorithm (c).

Figure 16. Residual image $\Psi(x_0) = f(x_0) - g(x_0 + u(x_0))$ for the adaptive isogeometric global digital image correlation method before refinement (a) and after refinement (b). The residual is expressed in terms of gray values. For illustration, an increment is chosen at the onset of localization, where localization still only occurs in the least wide region of the sample, i.e., above the hole, and not yet in the bottom region.

in the same area is recognized in the top image in Figure 15a. Displacement in both directions combined indicates shearing.

The displacement fields $u$ and $v$ are also calculated using a non-adaptive iso-GDIC algorithm. Second-order NURBS shape functions are used along with the initial mesh from the adaptive refinement method, as shown in Figure 14a. The resulting displacement is shown in Figure 15b. It is observed that the displacement field shows the same global characteristics as that calculated with the DIC algorithm with adaptive refinement. However, the refined shape functions are better able to capture the localization in the necking area; comparing Figure 15a and b, the sudden change in displacement is better visible in case adaptive refinement is used. This is also observed when comparing the residual images $\Psi(x_0)$ for the method before refinement and after refinement, using the meshes in Figure 14a and b, respectively; see Figure 16. It can be seen that the residual again is highest in the area where the sample shows localization of the displacement field, i.e., the area above the hole. After refinement, the residual is significantly lower in this area, indicating that the accuracy of the calculation has increased.
Correlation of the images taken during the experiment is again repeated using an optimized local DIC algorithm (MatchID, subset size 53 pixels, step size 10 pixels, SSD correlation coefficient, bicubic spline interpolation, and irregular shape functions). The resulting displacement fields are shown in Figure 15c. It is observed that in the area where localization takes place, correlation fails. The variation in displacement is rather large in this area, and subsets located here fail to correlate.

This again shows the advantages of the adaptive iso-GDIC method: accurate results are obtained, without the need of supplying problem-specific information.

7. CONCLUSIONS

A novel method was developed where iso-GDIC was combined with an adaptive refinement algorithm: adaptive iso-GDIC. The potential advantage of this method is that it is flexible and can be used for a wide variety of DIC problems. But perhaps even more important is that solutions are found autonomously, which means that the user does not have to make problem-specific choices, e.g., the shape functions in global DIC or subset size in local DIC. The NURBS shape functions are able to capture the kinematics of many displacement fields, depending on the structure of the knot vector. In combination with adaptive refinement, the shape functions are autonomously adjusted to be able to describe the kinematics of the sought displacement field with an optimized number of DOFs.

A hierarchical mesh refinement algorithm is used in order to prevent inefficient refinement, providing DOFs only where needed, which is beneficial for the conditioning of the ill-posed DIC-problem. Iterating towards the optimum number of DOFs is incorporated in the procedure of marking shape functions for refinement. Shape functions are merely refined if the residual in its region of support is significantly higher than the average.

A virtual experiment shows that for a test-case with a strong localization of the displacement, where refinement is desired, the new method yields good results compared with more conventional DIC implementations, in particular a non-adaptive iso-GDIC method and local DIC. Most importantly, this result is obtained autonomously. The user does not need to supply the program with specific prior information about the problem, as is required for the more conventional DIC method to obtain accurate results.

Moreover, real experiments with a complex geometry are executed to demonstrate that the novel method also performs well in practice. A notched tensile sample and a specimen with a hole were chosen, which show a localized displacement field under deformation. Mesh refinement also appears in the regions were localization is expected.

In all, it was shown that the iso-GDIC adaptive refinement algorithm has some major advantages. The most important advantage is that the number of DOFs is optimized autonomously, which ensures accurate results, without much problem-specific input and thus expertise of the user needed. Furthermore, this type of shape function is very flexible, which makes the method rather robust and applicable to a wide range of DIC problems. Last, it was shown in specific cases that the method provides better results than more conventional DIC methods, without adaptive refinement. This indicates that adaptive iso-GDIC is a promising technique for accurately solving a wide range of DIC problems.

ACKNOWLEDGEMENTS

The research of S.M. Kleinendorst and J.P.M. Hoefnagels was funded by the Netherlands Organization for Scientific Research (NWO) under the VIDI scheme.

The research of C.V. Verhoosel was funded by the Netherlands Organization for Scientific Research (NWO) under the VENI scheme.

REFERENCES