Pixel-level robust digital image correlation

Corneliu Cofaru,[1]* **Wilfried Philips,**[1] **and Wim Van Paepegem**[2]

[1]*Telin-IPI-IBBT, Ghent University*
St-Pietersnieuwstraat 41, B-9000 Ghent, Belgium
[2]*Department of Materials Science and Engineering, Ghent University,*
Technologiepark-Zwijnaarde 903, B-9052 Zwijnaarde, Belgium
* Cornel.Cofaru@telin.ugent.be

Abstract: Digital Image Correlation (DIC) is a well-established non-contact optical metrology method. It employs digital image analysis to extract the full-field displacements and strains that occur in objects subjected to external stresses. Despite recent DIC progress, many problematic areas which greatly affect accuracy and that can seldomly be avoided, received very little attention. Problems posed by the presence of sharp displacement discontinuities, reflections, object borders or edges can be linked to the analysed object's properties and deformation. Other problematic areas, such as image noise, localized reflections or shadows are related more to the image acquisition process. This paper proposes a new subset-based pixel-level robust DIC method for in-plane displacement measurement which addresses all of these problems in a straightforward and unified approach, significantly improving DIC measurement accuracy compared to classic approaches. The proposed approach minimizes a robust energy functional which adaptively weighs pixel differences in the motion estimation process. The aim is to limit the negative influence of pixels that present erroneous or inconsistent motions by enforcing local motion consistency. The proposed method is compared to the classic Newton-Raphson DIC method in terms of displacement accuracy in three experiments. The first experiment is numerical and presents three combined problems: sharp displacement discontinuities, missing image information and image noise. The second experiment is a real experiment in which a plastic specimen is developing a lateral crack due to the application of uniaxial stress. The region around the crack presents both reflections that saturate the image intensity levels leading to missing image information, as well as sharp motion discontinuities due to the plastic film rupturing. The third experiment compares the proposed and classic DIC approaches with generic computer vision optical flow methods using images from the popular Middlebury optical flow evaluation dataset. Results in all experiments clearly show the proposed method's improved measurement accuracy with respect to the classic approach considering the challenging conditions. Furthermore, in image areas where the classic approach completely fails to recover motion due to severe image de-correlation, the proposed method provides reliable results.

References and links

1. W. H. Peters and W. F. Ranson, "Digital imaging techniques in experimental stress-analysis," Opt. Eng. **21**, 427–431 (1982).
2. M. A. Sutton, W. J. Wolters, W. H. Peters, W. F. Ranson and S. R. McNeill, "Determination of displacements using an improved digital correlation method," Image Vision Comput. **1**, 133–139 (1983).
3. W. H. Peters, W. F. Ranson, M. A. Sutton, T.C. Chu and J. Anderson, "Application of digital correlation methods to rigid body mechanics," Opt. Eng. **22**, 738–742 (1983).
4. B. Pan, A. Asundi, H-M. Xie and J. X. Gao, "Digital image correlation using iterative least squares and pointwise least squares for displacement field and strain field measurements," Opt. Lasers Eng. **47**, 865–874 (2009).
5. J. Zhang, Y. Cai, W. Ye and T.X. Yu, "On the use of the digital image correlation method for heterogeneous deformation measurement of porous solids," Opt. Lasers Eng. **49**, 200–209 (2011).
6. L.B. Meng, G.C. Jin and X.F. Yao, "Application of iteration and finite element smoothing technique for displacement and strain measurement of digital speckle correlation," Opt. Lasers Eng. **45**, 56–73 (2007).
7. C. Tang, L. Wang, S. Yan, J. Wu, L. Cheng and C. Li, "Displacement field analysis based on the combination digital speckle correlation method with radial basis function interpolation," Appl. Opt. **49**, 4545–4553 (2010).
8. Y. Sun, J.H.L. Pang, C.K. Wong and F. Su, "Finite element formulation for a digital image correlation method," Appl. Opt. **44**, 7357–7363 (2005).
9. Y.N. Chen, W.Q. Jin, L. Zhao and F.W. Li, "A subpixel motion estimation algorithm based on digital correlation for illumination variant and noise image sequences" Optik **120**, 835–844 (2009).
10. B. Pan, Z. Wang and Z. Lu, "Genuine full-field deformation measurement of an object with complex shape using reliability-guided digital image correlation," Opt. Express **18**, 1011–1023 (2010).
11. J. Poissant and F. Barthelat, "A novel subset splitting procedure for digital image correlation on discontinuous displacement fields," Exp. Mech. **50**, 353–364 (2010).
12. C. Cofaru, W. Philips and W. Van Paepegem, "Improved Newton-Raphson digital image correlation method for full-field displacement and strain calculation," Appl. Opt. **49**, 6472–6484 (2010).
13. C. Cofaru, W. Philips and W. Van Paepegem, "A three-frame digital image correlation (DIC) method for the measurement of small displacements and strains," Meas. Sci. Technol. **23**, 105406 (14 pp.) (2012).
14. G. Besnard, F. Hild and S. Roux, "'Finite-Element' displacement fields analysis from digital images: Application to Portevin-Le Châtelier bands," Exp. Mech. **46**, 789–803 (2006).
15. J. Réthoré, S. Roux and F. Hild, "From pictures to extended finite elements: extended digital image correlation (X-DIC)," C.R. Mécanique **335**, 131–137 (2007).
16. J. Réthoré, F. Hild and S. Roux, "Extended digital image correlation with crack shape optimization," Int. J. Numer. Meth. Eng. **73**, 248–272 (2008).
17. M. A. Sutton, J-J. Orteu and H. W. Schreier, *Image correlation for shape, motion and deformation measurements* (Springer, 2009).
18. B. Pan, H. Xie and Z. Wang, "Equivalence of digital image correlation criteria for pattern matching," Appl. Opt. **49**, 5501–5509 (2010).
19. B. Peng, Q. Zhang, W. Zhou, X. Hao and L. Ding, "Modified correlation criterion for digital image correlation considering the effect of lighting variations in deformation measurements," Opt. Eng. **51**, 017004 (2012).
20. C. Cofaru, W. Philips and W. Van Paepegem, "A novel speckle pattern - adaptive Digital Image Correlation approach with robust strain calculation," Opt. Lasers Eng. **50**, 187–198 (2012).
21. P.J. Huber and E.M. Ronchetti, *Robust Statistics, 2nd Edition* (John Wiley & Sons, New York (NY), 2009).
22. P.J. Rousseeuw and A.M. Leroy, *Robust Regression and Outlier Detection* (John Wiley & Sons, 1987)
23. F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw and W.A. Stahel, *Robust Statistics: The Approach Based on Influence Functions* (John Wiley & Sons, 1986).
24. H. A. Bruck, S. R. McNeill, M. A. Sutton and W. H. Peters, "Digital image correlation using Newton-Raphson method of partial differential correction," Exp. Mech. **29**, 261–267 (1989).
25. G. Vendroux and W. G. Knauss, "Submicron deformation field measurements: Part 2. Improved digital image correlation," Exp. Mech. **38**, 86–92 (1998).
26. M. J. Black and P. Anandan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," Comput. Vis. Image Underst. **63**, 75–104 (1996).
27. M. Ye, R. M. Haralick and L. G. Shapiro, " Estimating piecewise-smooth optical flow with global matching and graduated optimization," IEEE Trans. Pattern Anal. Mach. Intell. **25**, 1625–1630 (2003).
28. C-H. Teng, S-H. Lai, Y-S. Chen and W-H. Hsu, "Accurate optical flow computation under non-uniform brightness variations," Comput. Vis. Image Underst. **97**, 315–346 (2005).
29. Y-H. Kim, A. M. Martìnez, A. C. Kak, "Robust motion estimation under varying illumination," Image Vis. Comput. **23**, 365–375 (2005).

30. S. Baker, D. Scharstein, J.P Lewis, S. Roth, M. J. Black and R. Szeliski, "A database and evaluation methodology for optical flow," Int. J. Comput. Vis. **92**, 1–31 (2011).
31. X. Ren, "Local grouping for optical flow," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 2008) pp. 1–8.
32. D. Sun, E. B. Sudderth and M. J. Black, "Layered image motion with explicit occlusions, temporal consistency, and depth ordering," Adv. Neural Inf. Process. Syst. **23**, 2226–2234 (2010).
33. D. Sun, E. B. Sudderth and M. J. Black, "Layered segmentation and optical flow estimation over time," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 2012) pp. 1768–1775.
34. Y. Weiss, "Smoothness in layers: Motion segmentation using nonparametric mixture estimation," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 1997) pp. 520–526.
35. M. Werlberger, T. Pock and H. Bischof, "Motion estimation with non-local total variation regularization," *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 2010) pp. 2464–2471.
36. M.J. Black, *Robust incremental optical flow* (PhD. Thesis, Yale University, 1992).
37. C. Cofaru, W. Philips and W. Van Paepegem, "Evaluation of digital image correlation techniques using realistic ground truth speckle images," Meas. Sci. Technol. **21**, 055102 (17 pp.) (2010).

1. Introduction

Digital Image Correlation, shortly known as DIC, is a class of image analysis techniques introduced in the early '80s [1–3] that find spatial correspondences between different digital images. In experimental mechanics, it can be classified as a non-contact optical metrology method that employs digital image analysis to extract full-field deformation measurements of objects that are subjected to external stresses. For in-plane deformation measurements, which are the focus of this work, at least two images have to be considered: one that shows the analyzed object before deformation and one showing the object after the deformation process has ended. By convention, these are called *reference* and *deformed* images respectively. Prior to image capture, the analyzed object is spray-painted so that a seemingly random layer of paint speckles is present on its surface to aid the matching process. Displacements are obtained by matching the image intensity pattern of the reference image to that of the deformed image. In practice, this consists in the optimization of a similarity criterion that penalizes pixel differences between the two images, considering a pre-defined displacement model. This can be done either locally [4–13], using rectangular image regions, also called *subsets*, or globally [14–16] across the entire image. Strain data is subsequently obtained either directly from the displacement field components or through the differentiation of the displacement field.

The accuracy of DIC is impacted to a large extent by experiment related factors such as image noise and lighting variations across the specimen's surface as well as by the characteristics of the deformations that are to be measured, such as their local magnitude fluctuations [17]. Despite recent advances aimed at improving displacement and strain accuracy, DIC lacks an approach that seeks to limit the influence of the various sources of measurement errors in an unified and effective way. If some problematic areas, such as image intensity variations [18,19], image noise [9], displacement discontinuities [11,12] or measurements near the edges of specimens [10,20], individually received attention in literature, others such as the presence of occlusions or localized image artefacts due to strong reflections or shadows, were not discussed. One of the main reasons for the apparent fragmentation in DIC solutions is the fact that current DIC methods optimize quadratic similarity criteria associated with the subset intra-pixel differences. Quadratic similarity criteria are defined in literature as either sums of squared differences (SSD), cross-correlation criteria (CC) or variations which use normalized image intensities [18, 19].These require explicit modelling of the factors that might impact accuracy because they assign an equal importance to all pixels of a reference and deformed subset pair in the motion estimation process. Therefore, the motions of all individual reference subset pixels

have an equal contribution to the final motion estimate of the subset. This is an obvious limitation since it is perfectly reasonable to assume that although most pixels present motions that can be properly modelled, some do not. These pixel motions lower the accuracy of the final motion estimate in direct proportion to both their numbers and magnitudes.

This paper addresses these limitations by introducing a new subset-based DIC approach that uses robust estimation [21, 22] to penalize intra-subset pixel differences between reference and deformed subsets. The pixel differences are weighted in such a way that pixels with motions that do not fit with the overall subset motion contribute less to the final motion estimate, resulting in improved accuracy of the local motion estimates. Such a low-level approach presents two other advantages: the first is that motion outliers at pixel level are treated in a very general manner and no a-priori assumptions or modelling of the image or displacement degradation is needed. The second advantage is that the robust estimation principle can be introduced into virtually all existing DIC methods as a way of enforcing local motion consistency and dealing with displacement outliers. The proposed method uses a Newton-Raphson optimization approach [24, 25] to obtain the displacements relating to material deformation and robust spatial regularization to adaptively increase spatial displacement consistency [12].

The rest of the paper is structured as follows: in Section 2 the mathematical formulation of the prosed method is presented. In Section 3, the proposed method is compared to the classic DIC method - which employs here the sum of square differences - in three experiments. Finally, the conclusions of this work are presented in Section 4.

2. Pixel-level robust DIC

2.1. Robust estimation

The field of robust statistics emerged from the need to address the situations in which the mathematical models developed to characterize certain phenomena through a set of related observations are insufficient to fully handle larger errors or outliers. Similar to many computer vision approaches, DIC focuses on providing the optimal solution for a set of measurements - the digital images - given some exact parametric models as for example, the pixel displacement model inside the subset. There are however situations in which the assumed models cannot correctly describe the observations and in which the integration of the concepts of robust statistics can provide a better final solution to the estimation problem. Because of this, robust estimation is a popular choice in motion estimation algorithms [26–29, 36] to improve accuracy by reducing the effects of motion outliers. The impact of the outliers on the final estimation result can be interpreted through the *influence function* of the estimator [23], which is the first derivative ψ of the function considered for the estimator. Considering the case of the quadratic estimator and its influence function ψ:

$$\rho(x) = x^2, \quad \psi(x) = 2x, \tag{1}$$

graphically illustrated in Fig. 1, it is obvious that the influence of the outliers increases linearly with their size and without bound. Since in DIC the quadratic estimator is used to penalize intra-pixel differences, even a few very large pixel differences - which usually appear if pixels in either of the two subsets are affected by noise or have a different motion - can have a significant impact on the final displacement estimate. It is therefore desirable to use robust estimators that gradually decrease the influence of subset pixel differences as these increase.

In this paper, two robust estimators are investigated. The first is the Welsch estimator, defined along with its influence function as:

$$\rho(x, \sigma) = \frac{\sigma^2}{2} \left[1 - \exp\left(-(x/\sigma)^2\right) \right], \quad \psi(x, \sigma) = x \exp\left(-(x/\sigma)^2\right), \tag{2}$$
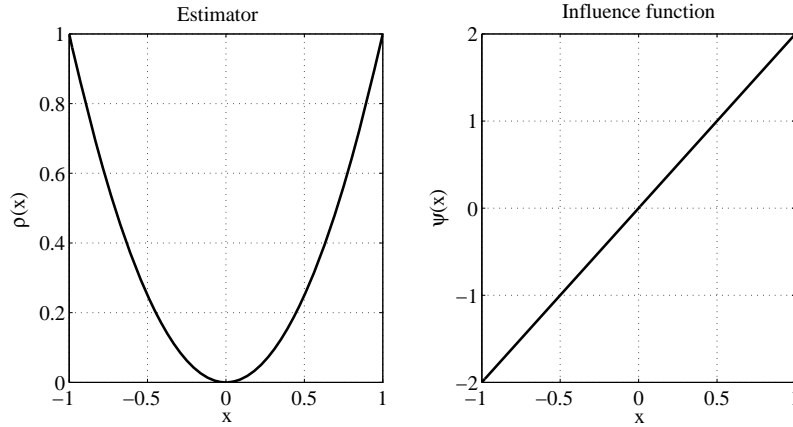
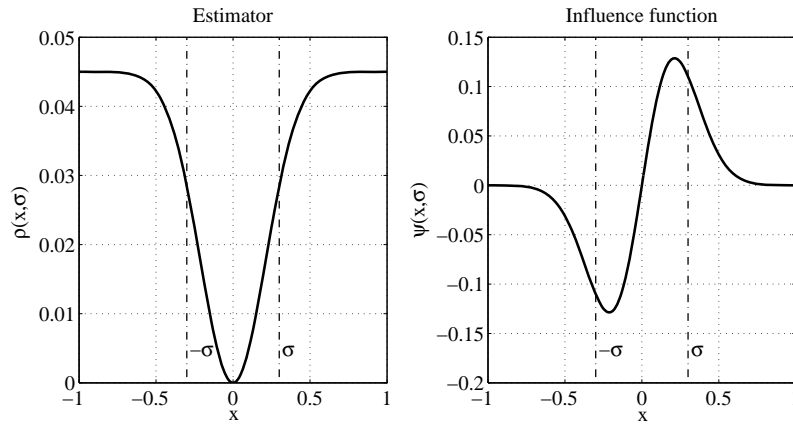Fig. 1: The quadratic estimator (left) and its influence function (right).

Fig. 2: The Welsch estimator (left) and its influence function (right). The shape parameter was considered constant with $\sigma = 0.3$

where ρ is the expression of the estimator, ψ that of its influence function and σ a shape parameter. The second estimator is the Geman-McClure estimator, which is similarly defined as:

$$\rho(x,\sigma) = \frac{x^2}{\sigma + x^2} \ , \quad \psi(x,\sigma) = \frac{2x\sigma}{(\sigma + x^2)^2} \ . \tag{3}$$

The two estimators and their corresponding influence functions are shown in Figs. 2 and 3, respectively. Both are *redescending* estimators (their influence function tends to zero above a certain chosen limit) and twice differentiable. The latter property is desired because it makes their application straightforward with optimization algorithms such as the Newton-Raphson. Furthermore, both estimators are inherently adaptive to the scale of the errors present. This is done through the shape parameter σ that controls the outlier rejection threshold: measurements that generate errors larger than the threshold are treated as outliers.

Let us assume that the threshold τ is the limit beyond which a measurement is treated as outlier, depending on its error with respect to the assumed model. These thresholds can be also interpreted as the inflexion points in the influence function ψ from where it is starting to
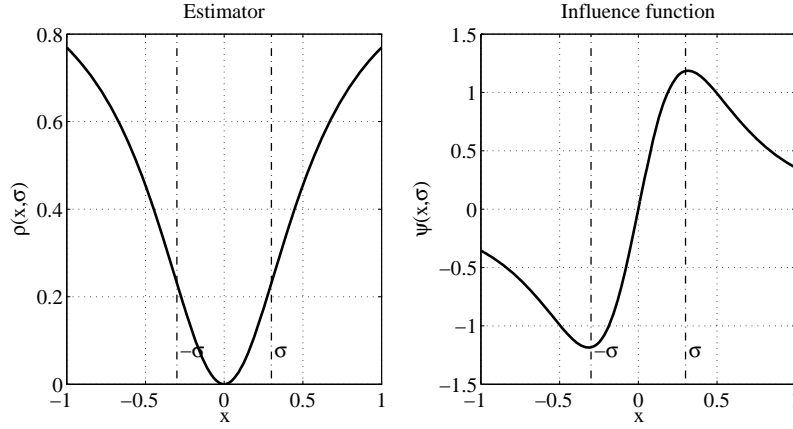
Fig. 3: The Geman-McClure estimator (left) and its influence function (right). The shape parameter was considered constant with $\sigma = 0.3$

decrease. In these points, the second derivative of the estimator is zero. It is straightforward to show that for the Geman-McClure estimator, τ is:

$$\tau = \pm\sqrt{\frac{\sigma}{3}}\,, \tag{4}$$

while for the Welsch estimator,

$$\tau = \pm\frac{1}{\sqrt{2}}\sigma\,. \tag{5}$$

Obviously, having an estimate of the magnitude of the maximum "allowable" errors in the measurements, namely τ, one could set the parameter σ in such a way that any measurements that present errors larger than τ are outliers. For example, setting $\sigma = \sqrt{2}\tau$ in the Lorentzian has the effect of treating all measurements with errors smaller than $-\tau$ or larger than τ as outliers. More importantly, inside the interval $[-\tau,\ \tau]$ the estimators are convex because the second derivatives are larger or equal to zero. For the interval, the estimator ρ is approximately quadratic and its influence function ψ approximately linear. This implies that if σ is chosen large enough as for all the measurements to present errors within the convexity interval, the problem becomes one of convex minimization.

2.2. Robust subset similarity criterion and optimization

Let $f(x,y)$ be a reference subset in the image that contains the analyzed material specimen before deformation and let $g(x',y')$ be its corresponding deformed subset, in the image showing the specimen after deformation. Both subsets are of size $N \times M$ pixels with $x' = x + u(x,y)$, $y' = y + v(x,y)$, where $u(x,y)$ and $v(x,y)$ are the horizontal and vertical displacements defined as:

$$u(x,y) = p_1 + p_3(x - x_0) + p_5(y - y_0)\,, \tag{6a}$$

$$v(x,y) = p_2 + p_4(x - x_0) + p_6(y - y_0)\,. \tag{6b}$$

In the equations above, x_0 and y_0 are the horizontal and vertical coordinates of the subset centre and $\mathbf{p} = (p_1,\ldots,p_6)^\mathsf{T}$ the first order displacement component vector which represents the

sought motion estimation solution. To obtain the solution, the proposed robust subset similarity criterion, expressed as:

$$C(\mathbf{p}) = E_D(\mathbf{p}) + \mu_S E_S(\mathbf{p}) , \tag{7}$$

is optimized. Here, μ_S is a parameter that adjusts the strength of the regularization, E_D is the image data term that penalizes subset differences at pixel level and E_S is a regularization or smoothness term. The data term uses the Welsch estimator to penalize pixel differences and is of the form:

$$E_D(\mathbf{p}) = \sum_{x=1}^{M} \sum_{y=1}^{N} \frac{\sigma_D^2}{2} \left[1 - \exp\left(-\left(f(x,y) - g(x,y;\mathbf{p}) \right)^2 / \sigma_D^2 \right) \right] , \tag{8}$$

where σ_D is the shape parameter of the estimator. The smoothness term employs the Geman-McClure estimator, to penalize local displacement component differences. It is defined as:

$$E_S(\mathbf{p}) = \sum_{i=1}^{6} \sum_{k=1}^{8} \frac{\left(p_i - p_{ik} \right)^2}{\sigma_i + \left(p_i - p_{ik} \right)^2} \tag{9}$$

and adaptively adjusts how much each displacement component p_i is influenced by its spatial neighbours p_{ik}, found in its 8-connected neighbourhood, through the outlier rejection parameter σ_i. The neighbouring displacement components p_{ik} are assumed to have been calculated in the previous iteration. The calculus of both data term and smoothness term robust rejection parameters is detailed in Section 2.3.

The minimization of Eq. (7) through the Newton-Raphson method has the solution at the t-th iteration:

$$\mathbf{p}^{(t)} = \mathbf{p}^{(t-1)} - \mathbf{H}_C^{-1}\left(\mathbf{p}^{(t-1)}\right) \mathbf{J}_C\left(\mathbf{p}^{(t-1)}\right) , \tag{10}$$

where $\mathbf{J}_C$ and $\mathbf{H}_C$ are the Jacobian and Hessian matrices associated to the subset similarity criterion with

$$\mathbf{J}_C(\mathbf{p^{(t-1)}}) = \left(\frac{\partial}{\partial p_1} C\left(\mathbf{p}^{(t-1)}\right) \quad \cdots \quad \frac{\partial}{\partial p_6} C\left(\mathbf{p}^{(t-1)}\right) \right) \tag{11}$$

and

$$\mathbf{H}_C(\mathbf{p}) = \begin{pmatrix} \frac{\partial^2}{\partial p_1^2} C\left(\mathbf{p}^{(t-1)}\right) & \cdots & \frac{\partial^2}{\partial p_1 \partial p_6} C\left(\mathbf{p}^{(t-1)}\right) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial p_6 \partial p_1} C\left(\mathbf{p}^{(t-1)}\right) & \cdots & \frac{\partial^2}{\partial p_6^2} C\left(\mathbf{p}^{(t-1)}\right) \end{pmatrix} . \tag{12}$$

The general form of the Jacobian elements at the t-th iteration is:

$$\frac{\partial}{\partial p_i} C\left(\mathbf{p}^{(t-1)}\right) = -\sum_{x=1}^{M} \sum_{y=1}^{N} \left[\frac{\partial}{\partial p_i} g\left(x,y;\mathbf{p}^{(t-1)}\right) \left(f(x,y) - g\left(x,y;\mathbf{p}^{(t-1)}\right) \right) \right.$$
$$\left. \exp\left(-\left(f(x,y) - g\left(x,y;\mathbf{p}^{(t-1)}\right) \right)^2 / \left(\sigma_D^{(t-1)}\right)^2 \right) \right] \tag{13}$$
$$+ \mu_S \sum_{k=1}^{8} \frac{2\sigma_i^{(t-1)} \left(p_i^{(t-1)} - p_{ik}^{(t-1)} \right)}{\left(\sigma_i^{(t-1)} + \left(p_i^{(t-1)} - p_{ik}^{(t-1)} \right)^2 \right)^2} ,$$

with $i = 1, \ldots, 6$. The Hessian elements are of the form:

$$\frac{\partial^2}{\partial p_i \partial p_j} C(\mathbf{p}^{(t-1)}) = \sum_{x=1}^{M} \sum_{y=1}^{N} \left[\frac{\partial}{\partial p_i} g(x,y;\mathbf{p}^{(t-1)}) \frac{\partial}{\partial p_j} g(x,y;\mathbf{p}^{(t-1)}) \right.$$
$$\left. \exp\left(-\left(f(x,y) - g(x,y;\mathbf{p}^{(t-1)})\right)^2 / (\sigma_D^{(t-1)})^2\right) \right], \tag{14a}$$

$$\frac{\partial^2}{\partial p_i^2} C(\mathbf{p}^{(t-1)}) = \sum_{x=1}^{M} \sum_{y=1}^{N} \left[\left(\frac{\partial}{\partial p_i} g(x,y;\mathbf{p}^{(t-1)}) \right)^2 \right.$$
$$\left. \exp\left(-\left(f(x,y) - g(x,y;\mathbf{p}^{(t-1)})\right)^2 / (\sigma_D^{(t-1)})^2\right) \right] \tag{14b}$$
$$+ \mu_S \sum_{k=1}^{8} \frac{(2\sigma_i^{(t-1)})^2 - 6\sigma_i^{(t-1)}(p_i^{(t-1)} - p_{ik}^{(t-1)})^2}{\left(\sigma_i^{(t-1)} + (p_i^{(t-1)} - p_{ik}^{(t-1)})^2\right)^3},$$

with $i, j = 1, \ldots, 6$ and $i \neq j$. Note that the terms $\partial^2 C(\mathbf{p}) / \partial p_i^2$ are only present on the first diagonal of the Hessian matrix since the second order partial derivatives of E_S with respect to two different displacement components are zero.

The method first divides the reference image into overlapping rectangular subsets and finds the best match for each reference subset in the deformed image through cross-correlation coefficient maximization. The resulting integer pixel displacements represent the initial solution in the Newton-Raphson optimization. The latter updates at each iteration the entire motion vector field excluding the locations where convergence has already been reached. As convergence criterion, a difference smaller than 10^{-5} between consecutive iterations for all displacement components of a reference subset has been chosen. Once convergence is reached, its six components are prevented further updating. However, they are used in subsequent iterations in the calculus of their neighbour's regularization terms. If in three successive iterations the total number of motion vectors that reached convergence does not change, the algorithm stops. To obtain the intensity values of the iteratively deformed subset $g(x,y,\mathbf{p})$ at non-integer coordinates, bicubic spline interpolation is used.

2.3. Outlier rejection

One of the major advantages of the proposed method is that it can address a large variety of sources of errors in displacement accuracy generically. Therefore, the effects of these sources are not specifically modelled for. Consequently, the optimal values for the outlier rejection thresholds τ are unknown. These have to be locally adapted - through the shape parameters σ - to the errors in such a way that the similarity criterion remains approximately convex at each iteration. The approach taken is that of iteratively refining σ to construct convex approximations of the similarity criterion at each iteration using the local pixel and displacement differences. Practically, the minimization process can be regarded as a two-step procedure where first, a coarse approximation of the displacement solution is obtained by forcing the similarity criterion to be convex, which is equivalent to applying the quadratic estimator. Subsequently, once the solution is close to the minimum σ is gradually adapted, allowing for the outliers to be removed while still allowing the algorithm to converge.

In the data term, the outlier rejection threshold is set so that at each iteration, the pixels in the reference and deformed subsets which produce differences above the median absolute pixel difference between the two subsets, are be treated as outliers. Using the relationship between

the outlier rejection threshold and the robust shape parameter for the Welsch estimator from Eq. 5, we obtain that:

$$\sigma_D^{(t-1)} = \sqrt{2}\,\text{median}\left(\left|f(x,y) - g\left(x,y;\mathbf{p}^{(t-1)}\right)\right|\right). \tag{15}$$

To avoid automatically rejecting too many pixels by setting the threshold too low - especially when close to convergence or if pixels are not affected by large errors - a minimum rejection threshold value has to be set. Its value is twice the median of all absolute pixel differences across all reference and deformed subset pairs from the previous iteration. Although $\sigma_D^{(t-1)}$ is calculated and used at iteration t, the notation associates it to the previous iteration because in calculating it, the previous displacement component solution $\mathbf{p}^{(t-1)}$ is used.

In the regularization term, there is a total of six σ_i parameters, each corresponding to one of the six displacement components in the linear model. Each parameter is calculated as a multiple of the standard deviation $\tilde{\sigma}_i$ of the smoothness residuals:

$$r_{ik} = p_i - p_{ik}, \quad \text{for } i = 1,\ldots,6 \,;\, k = 1,\ldots,8 \tag{16}$$

with

$$\tilde{\sigma}_i = \sqrt{\frac{1}{8-1}\sum_{k=1}^{8}(r_{ik} - \bar{r}_i)^2}, \quad \bar{r}_i = \frac{1}{8}\sum_{k=1}^{8} r_{ik} \tag{17}$$

and updated at each iteration in the minimization process.

For each displacement component, the extent to which neighbours contribute to its value is given by the influence function of the robust estimator, which decreases rapidly towards zero when $|r_{ik}| > |\tau_i| = \sqrt{\sigma_i/3}$. Small σ_i values, equivalent to a strong correlation between neighbouring estimates in the area around the displacement component currently estimated, lead to strong regularization using the closest neighbouring values. Alternatively, larger residual values reflect strong changes in the displacement component's neighbourhood, and the influence that the neighbours exert on the component being estimated is limited. For the first two experiments presented in this paper, $\sigma_i = 15\tilde{\sigma}_i$ while for the last one, concerning the Middlebury [30] dataset, $\sigma_i = 7\tilde{\sigma}_i$. The constant multiplication factor in σ_i is chosen empirically. This comes as a consequence of the fact that in robust estimation, the outlier rejection thresholds have to be specified prior to the estimation. In other words, an estimate of the magnitude of the errors associated with outliers has to be known in advance - in this case, at the beginning of each iteration. The values of the thresholds are however not directly specified because displacement component differences can vary significantly in size during the course of the minimization. Thresholds are instead expressed as multiples of the standard deviation $\tilde{\sigma}_i$ of the smoothness residuals, whose values are linked to the magnitude of the local displacement component differences. There is no constraint on the values that the multiplication factors can take. However, if these are too large, the effects of regularization are minimal because the weights assigned to neighbouring components are be very small. In these situations, if the regularization strength parameter would be increased accordingly, the effect of the regularization term would be to smooth the displacement components regardless of the local similarities between them. If the multiplication factors are chosen very small, neighbouring displacement components would have to be very similar to the central one to influence it. Similar previous regularization implementations [12, 20] preserved well large displacement discontinuities while at the same time smoothing minor local displacement component variations, effectively filtering out displacement calculus noises.

3.   Experimental details and results

The proposed robust method is compared in terms of displacement accuracy to the classic Newton-Raphson DIC approach with sum of square differences (SSD) similarity criterion in three experiments. Two experiments are DIC specific and employ speckle images while the third one is more generic and relies on four grayscale image sequences from the Middlebury optical flow evaluation dataset.

The first experiment aims to numerically compare the two methods in terms of displacement accuracy. The deformed image in the experiment is obtained from a 512×512 pixel reference image by artificially deforming it using radial basis function interpolation [37] and then degrading it by introducing areas of missing image data and image noise. The reference image is part of a larger 1024×1024 pixel image showing a plastic film specimen onto which a speckle pattern has been previously applied. The second experiment is a real deformation experiment in which the displacements present on the surface of a deforming plastic film are measured with the two DIC methods. The plastic specimen is bonded on a rectangular piece of transparent glass of similar width which has previously been cut in two, horizontally. On the opposite side of the glass, another plastic specimen is also bonded. Both reference and deformed images are sized 1024×1024 pixels. Although no numerical accuracy figures are available, the nature of the results suffices in providing a clear picture on the displacement quality differences present between the methods.

The reference speckle image from the numerical experiment and both images from the real experiment are obtained in a similar manner: to create the speckle patterns, the surfaces of two thin plastic films are sprayed with a matt black paint, followed by a very fine dust of white matt paint. The paint adheres to the film surface and provides a non-flaking thin layer which does not change the properties of the film. The resulting speckle pattern contained densely packed, small speckles, of maximum 5 pixels in diameter. Illumination was provided by two slide projectors with 100 W halogen bulbs and integrated infrared filters, reducing the unwanted thermal spectrum. To capture the images, a Hitachi P110 CCD camera with a resolution of 1024×1024 pixels and 8-bit grayscale intensity quantization was used. Through the optical arrangement, one pixel displacement in the image plane corresponds to a 23.8 μm displacement in the object plane. Noise levels are low throughout the images. Image noise is added artificially only to the images corresponding to the first experiment. For the second experiment, the specimen was stretched by means of a specially designed low vibration hydraulic loading device.

The purpose of the third experiment is to assess how DIC methods, which are designed to work exclusively with highly textured images, perform when applied on more generic images, with potentially less image texture content. The experiment will also compare the accuracy of DIC methods to that of standard optical flow methods. To this end, four grayscale image sequences of two frames each from the Middlebury dataset are employed: 'Dimetrodon', 'Hydrangea', 'Venus' and 'Grove2'. These are sized 584×388, 584×388, 420×380 and 640×480 pixels respectively. The 'Dimetrodon' and 'Hydrangea' sequences consist of real world images with non-rigid motions and complex occlusions. The maximum horizontal and vertical displacements do not exceed approximately three pixels in the 'Dimetrodon' and and ten pixels in the 'Hydrangea' sequence. The latter also contains very localized motion discontinuities and variations in the areas corresponding to the plant itself. The 'Venus' sequence contains stereo images. The vertical displacements are absent while the horizontal ones present sharp discontinuities and have amplitudes up to approximately ten pixels. The 'Grove2' sequence contains computer generated images. Displacements do not exceed four pixels in amplitude and present complex spatial variations with sharp and local discontinuities in the foliage areas. The background presents lower amplitude, non-rigid displacements with smooth spatial variations. The ground truth optical flow for all four image pairs is available.
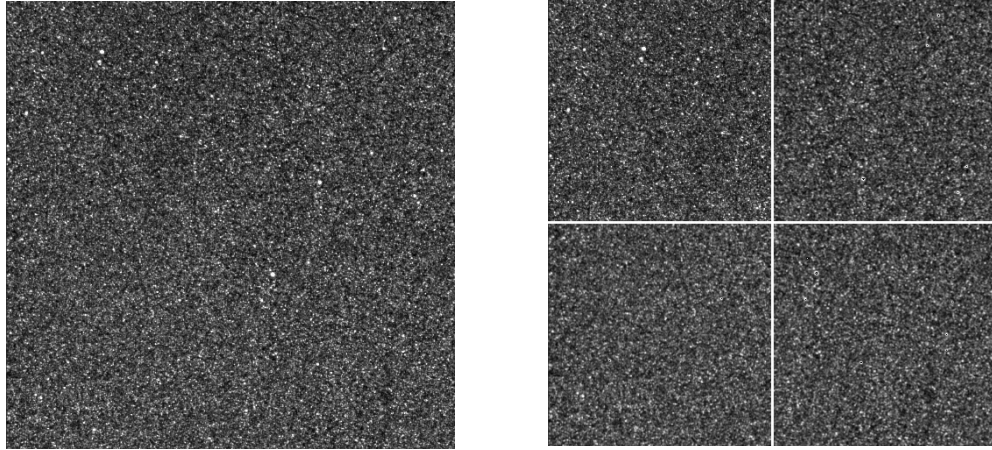
Fig. 4: The reference (left) and deformed image (right) used in the first experiment.

In evaluating the proposed DIC approach, only displacement data is used. Although the evaluation can also focus on strain accuracy, the latter is largely dependent on displacement accuracy. Therefore, any quality improvements found in the calculated displacements are to be reflected in the strains as well. Moreover, a displacement-based evaluation is preferable because strain calculation generally involves post-processing the displacement data which can make differences in accuracy as well as localized displacement errors more difficult to identify and assess.

3.1. Numerical DIC experiment

The reference and deformed images from the first experiment are shown in Fig. 4. The numerical deformed image is obtained by first dividing the reference image into four rectangular regions sized 256×256 pixels and applying a rigid body displacement to three of the four regions. The upper right region is displaced horizontally 2.5 pixels, the lower left region is displaced vertically 2.5 pixels while the lower right region is displaced both horizontally and vertically with 2.5 pixels. This implies that vertical displacements are present only in the lower half of the deformed image (a downwards motion) while horizontal displacements are present only in the right half of the image (rightwards motion). After the application of the rigid body displacements, in the areas that separate the four image regions of different motion, all pixels are assigned the value 255. The width of the "occluded" regions is 3 pixels and the length is 512 pixels. Finally, normally distributed noise of zero mean and standard deviation 3 is added to both reference and deformed images. Figure 5 shows the displacement distribution schematic and image difference. It can be observed that in the regions with displacement discontinuities, both missing image data and image noise are present. This situation, although artificially induced in this experiment, can occur in real experiments as when for example visible cracks suddenly develop or if the speckle pattern is significantly affected by deformations. Furthermore, localized reflections can also lead to regions with missing image information by producing image areas of high intensity, as it will be seen in the second experiment.

To compare the proposed and traditional DIC approaches, the displacements in the centres of subsets 15×15, 21×21, 27×27 and 33×33 pixels in size are calculated. The subsets overlap so that subset centres are spaced 5 pixels both horizontally and vertically. The proposed method is investigated both with displacement regularization ($\mu_S = 1000$), and without ($\mu_S = 0$). For
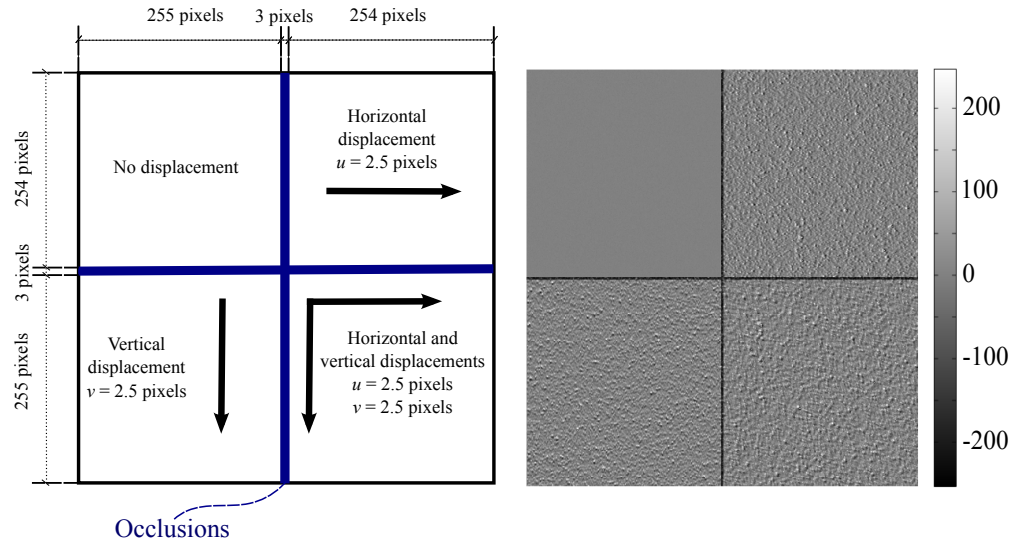
Fig. 5: The numerical displacement schematic (left) and image difference (right) corresponding to the first experiment.
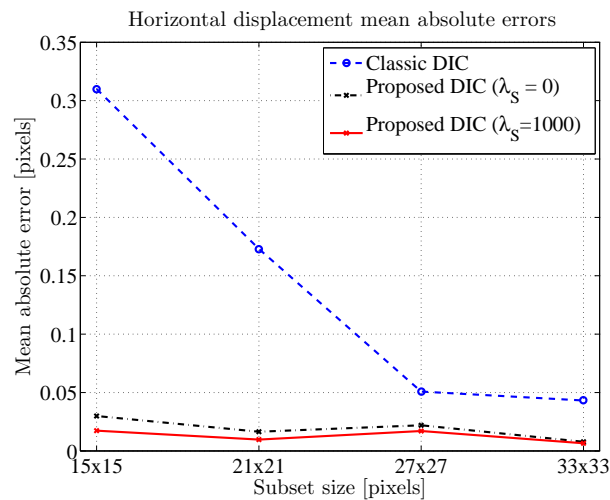

Fig. 6: The mean absolute horizontal displacement errors for the evaluated DIC methods in the first experiment.

both DIC methods used, the subset centre displacements where the convergence condition was not met were not taken into consideration. From a computational expense perspective, the robust method without regularization is approximately two times slower than the classic one. The use of regularization further increases the total running time with approximately 40%.

In Figs. 6 and 7, the horizontal and vertical mean absolute displacement errors for the classic DIC approach and proposed one with and without regularization are shown. The two figures clearly show the accuracy advantages that the robust DIC approach presents with respect to the classic least squares one. For the 15×15 pixel subsets, the classic DIC horizontal displacement error of 0.31 pixels is approximately ten times larger than that of the robust DIC approach of
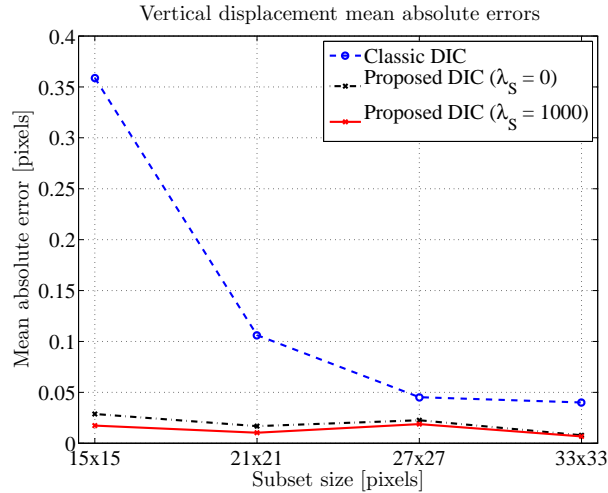
Fig. 7: The mean absolute vertical displacement errors for the evaluated DIC methods in the first experiment.

0.0298 pixels. Regularization further reduces the error in this case to 0.017 pixels. As subset sizes increase the accuracy gap narrows. For the 33×33 pixel subsets, the classic DIC horizontal error of 0.043 pixels is more than 4 times larger than the error of the robust approach which is 0.00784 pixels. The mean absolute vertical displacement errors present similar differences. The classic DIC approach errors are larger than the ones of the robust approach with a factor that varies between 12, for 15×15 pixel subsets, and 6 for 33×33 pixels subsets. Regularization further improves displacement quality, in particular for lower sized subsets, where small local displacement discontinuities due to image noise are reduced. The vertical displacement errors of the robust method are reduced by regularization between 40% for 15×15 pixel subsets and 15% for the 33×33 pixel subsets. This is because larger subset sizes inherently reduce the influence of image noise on the displacement estimates, local motion consistency is higher and therefore the effects of displacement regularization are limited. Nevertheless, as previously noted [12], regularization represents an important element for improving accuracy whenever measuring displacements with high frequency spatial variations, which require a reduction in subset size.

The proposed DIC approach also reduces the number of subsets for which motion cannot be calculated because of convergence issues. Figure 8 shows the variation with subset size of the number of non-converging displacement solutions, for the two evaluated DIC methods. The numbers of non-converging displacements for the proposed method with and without regularization decrease from 13, for the smallest subset size to 0 for the largest subset size. This evidently indicates that it can largely handle the problematic occluded areas which separate the image areas in which large displacement discontinuities are present. In contrast, the least-squares criterion based classic DIC approach cannot recover the displacements for up to 316 subsets when using subsets sized 21×21 pixels. As the subset size increases, this number decreases, reaching 71 for the 33×33 pixel subsets. All subsets whose motion cannot be recovered by the classic DIC approach are overlapping the white image pixels artificially introduced. It is important to observe that whenever large magnitude pixel errors are present in either the reference or the deformed subsets, the least-squares criterion fails to produce accurate results. In comparison, the robust subset similarity criterion adaptively assigns lower weights
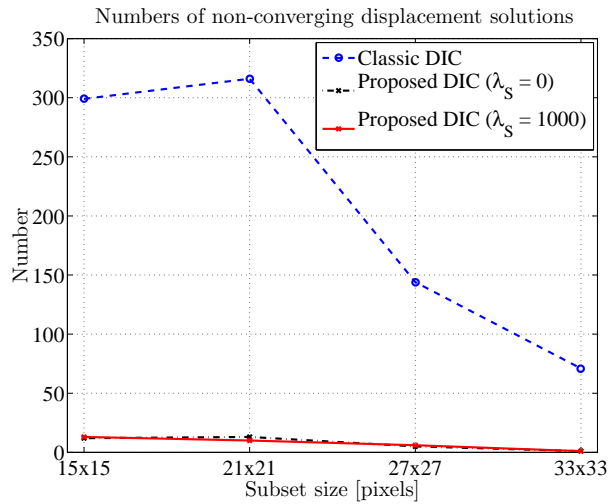
Fig. 8: The variation with subset size of the number of non-converging displacement solutions in the first experiment.

to these pixel pairs when calculating the displacements, thus being able to significantly reduce convergence problems. In Figs. 9 and 10, the horizontal and vertical displacements recovered with the classic and proposed DIC approach (without regularization) for subset sizes of 15×15 and 33×33 pixels are shown. For the smaller subsets, the classic DIC approach presents large errors both in the areas where displacement discontinuities are present as well as in those where displacements are constant and only occluded pixels are present. Although it uses the same image information and no explicit image degradation model, the proposed method largely reduces displacement errors just by adjustments of the influence of large error yielding pixels. These can be further reduced through regularization while at the same time preserving the horizontal and vertical displacement discontinuities. Increasing subset size reduces the errors in the image areas of constant motion however smooths the displacements in the areas where discontinuities are present. This effect can be clearly seen in Fig. 10a where, for both horizontal and vertical displacement fields, there is a clear region of transition from the 0 to the 2.5 pixel displacements. This is due to the fact that the quadratic criterion assigns equal importance to all pixels within a subset. As the subset starts overlapping image areas of different displacement, the latter gradually start to influence the final displacement estimate. The size of the displacement transition area in the displacement fields is therefore directly proportional to the size of the subset overlap as well as subset size. The larger the subsets and the overlap between them, the larger the area of transition between displacements. The proposed robust method reduces this displacement transition area to one displacement at most. This generally occurs when the subset falls approximately in the middle of the two distinct displacement image regions. In these cases, the number of pixels of different displacements is approximately the same and the robust estimator acts largely as the quadratic one. This situation can be avoided by lowering the error rejection threshold of the estimator, forcing the displacement solution to converge to the displacement of the lowest error yielding pixels. In this evaluation, the error rejection threshold calculation was not specifically designed to minimize displacement transitions in the presence of sharp discontinuities, this area of research necessitating further future investigations. Nevertheless, the robust subset similarity criterion presents much better displacement discontinuity preservation properties than the quadratic one.
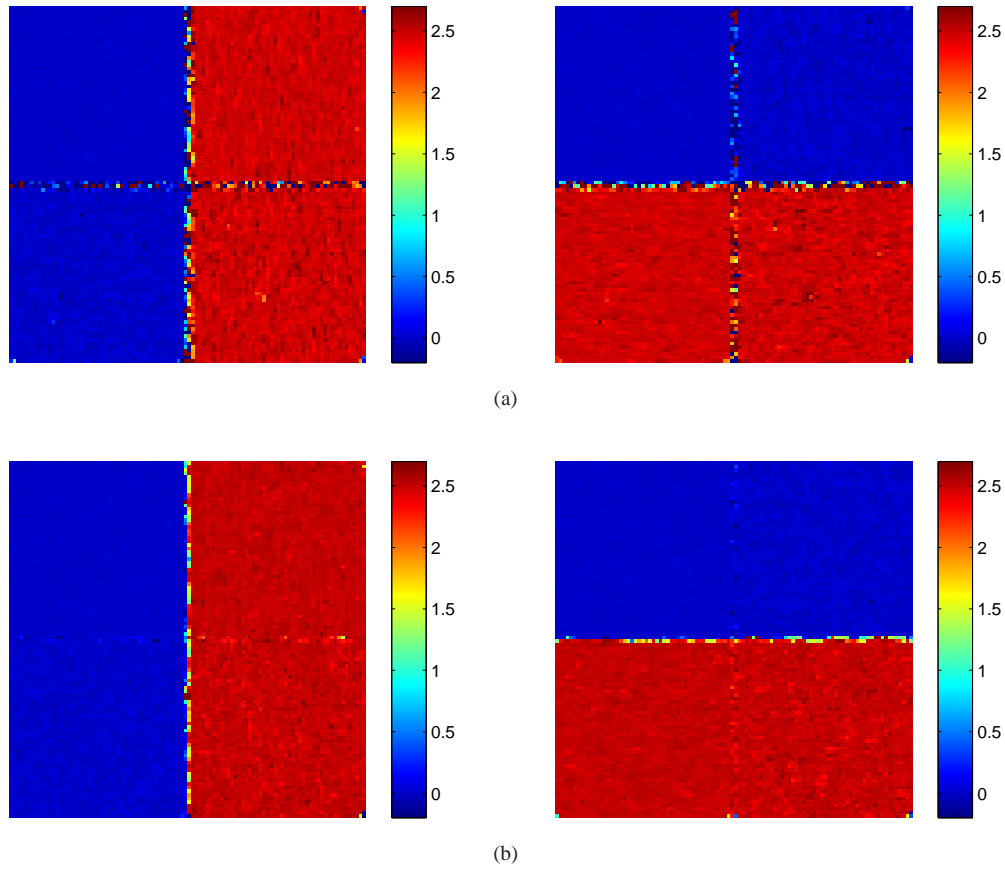
Fig. 9: The horizontal (left) and vertical (right) displacements calculated with (a) the classic and (b) proposed robust DIC approaches ($\mu_S = 0$) using a subset size of 15×15 pixels, in the first experiment.

3.2. *Real DIC experiment*

The reference and deformed images used in the second, real experiment are shown in Fig. 11. In the figure are also shown in detail parts of the reference and deformed images 300×200 pixels in size which contain the lateral crack developed in the specimen. By comparing both images, it becomes clear that in the vicinity of the crack, localized reflections are present. These appeared as a result of small pieces of the torn plastic specimen reflecting the incoming light. The back side of the second plastic film (bonded on the other side of the glass) can be seen exposed where the front side film is ruptured as well as all around the outer edges of the front film. Of particular interest is the displacement accuracy in the vicinity of the crack, where both localized displacement discontinuities as well as reflections are present, situation simulated and investigated in the first experiment. Displacements are calculated using 15×15 pixel subsets with a 5 pixel horizontal and vertical step between neighbouring subset centres. The relatively small subset size is well suited for the speckle pattern and provides a good compromise between accuracy and spatial resolution of the displacement data. For the proposed method, the regularization strength parameter was set at $\mu_S = 3000$.

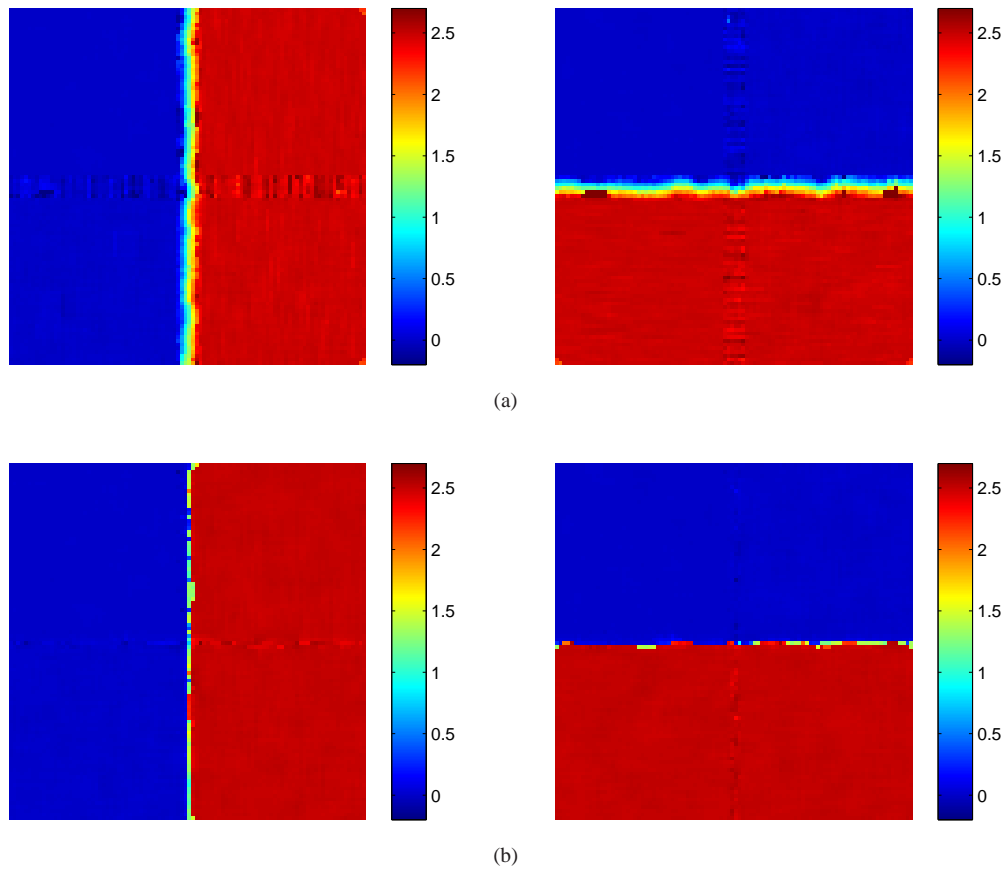For the classic DIC method, 583 displacement solutions did not converge while for the pro-

(a)

(b)

Fig. 10: The horizontal (left) and vertical (right) displacements calculated with (a) the classic and (b) proposed robust DIC approaches ($\mu_S = 0$) using a subset size of 33×33 pixels, in the first experiment.

posed method, the number was of 270. It should be noted that in both cases, many of the non-converging displacement solutions correspond to image regions not belonging to the plastic specimen's surface but rather to the background. The horizontal and vertical displacements obtained with the two DIC methods are shown in Fig. 12.

The robust DIC method presents a much smoother displacement variation across the whole displacement field as a result of regularization. This is more visible in the case of the horizontal displacements which are smaller, and therefore, the local displacement variations are larger relative to the displacement amplitude. Horizontal displacements are smaller than one pixel throughout the surface of the specimen while vertical ones vary approximately between -4 and 2 pixels. Vertical motion presents an upward orientation (negative displacement values), consistent with the direction of application of the load. Underneath the lateral crack, the motion is downwards, as a result of unloading in the film. Significant differences in the estimated displacements can be observed in the immediate vicinity of the specimen crack. Figure 13 shows a detail 36×31 displacement points in size containing the vertical displacements around this area. The motion vectors corresponding to the same points are shown in Fig. 14. From the two images it is clear that at the tip of the crack , the robust method presents many more reli-

Front-side   Back-side   Lower half of the
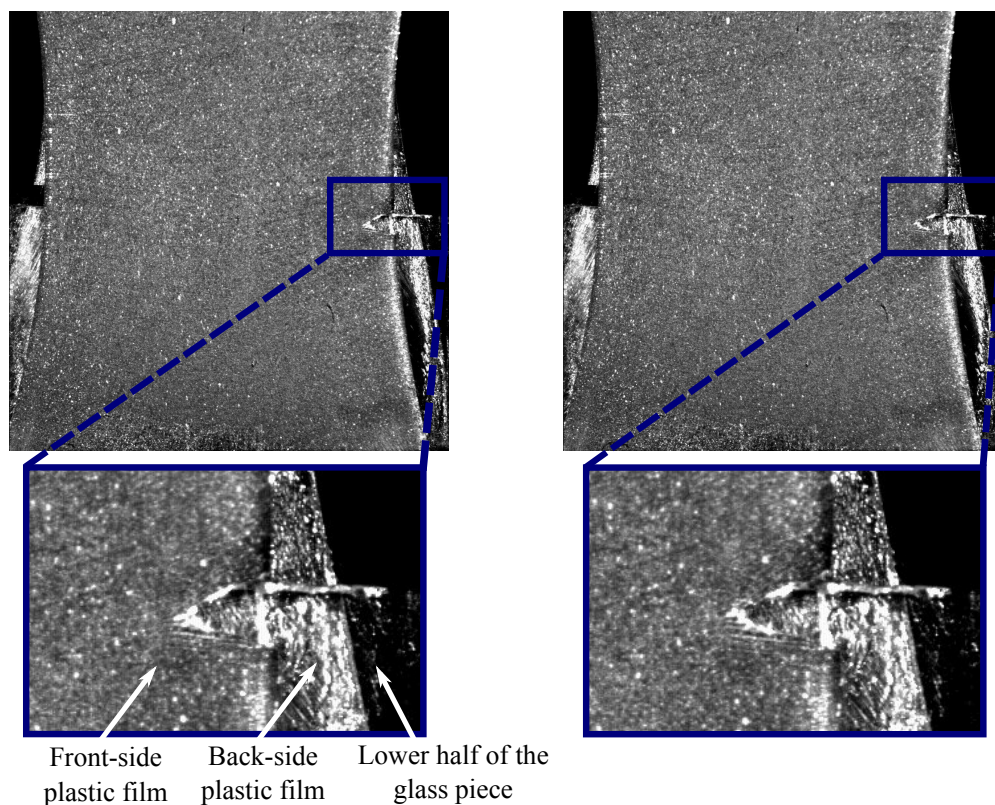plastic film plastic film   glass piece

Fig. 11: The reference (left) and deformed image (right) used in the second experiment.

able displacement estimates than the classic method. It is also important to note that, although the robust method presents displacement solutions that have not converged, indicated by the red markers in Fig. 14, their values are more consistent with neighbouring displacements that correctly converged. Spatial consistency between converging and non-converging displacement values can be noted also in Fig. 13. The displacements calculated with the classic method however are completely unreliable and have very large, erroneous values. These values appear in Fig. 13 in dark shades of red or blue which indicate that the values at the respective points fall outside the displacement amplitude scale to which the colours are mapped. A solution to improve the robust method accuracy at the tip of the crack would be to lower error rejection threshold for subsets with many large intra-pixel differences. This would allow for larger numbers of pixels in the reference and deformed subsets to be treated as outliers as well. At this point this approach still requires further investigation.

3.3. Middlebury dataset experiment

When applied to the Middlebury dataset images, the proposed and classic DIC methods employed 21×21 pixel subsets with a 5 pixel step in both horizontal and vertical directions between subsets. The regularization parameter for the robust DIC method was fixed at $\mu_S = 500$. No other special optimization or modification of the two methods was made for the experiment. For both methods, image-sized displacement fields are obtained by using the six displacement components corresponding to each subset and the average endpoint error (AEE) between them and the ground truth fields is calculated. The endpoint errors corresponding to the two DIC
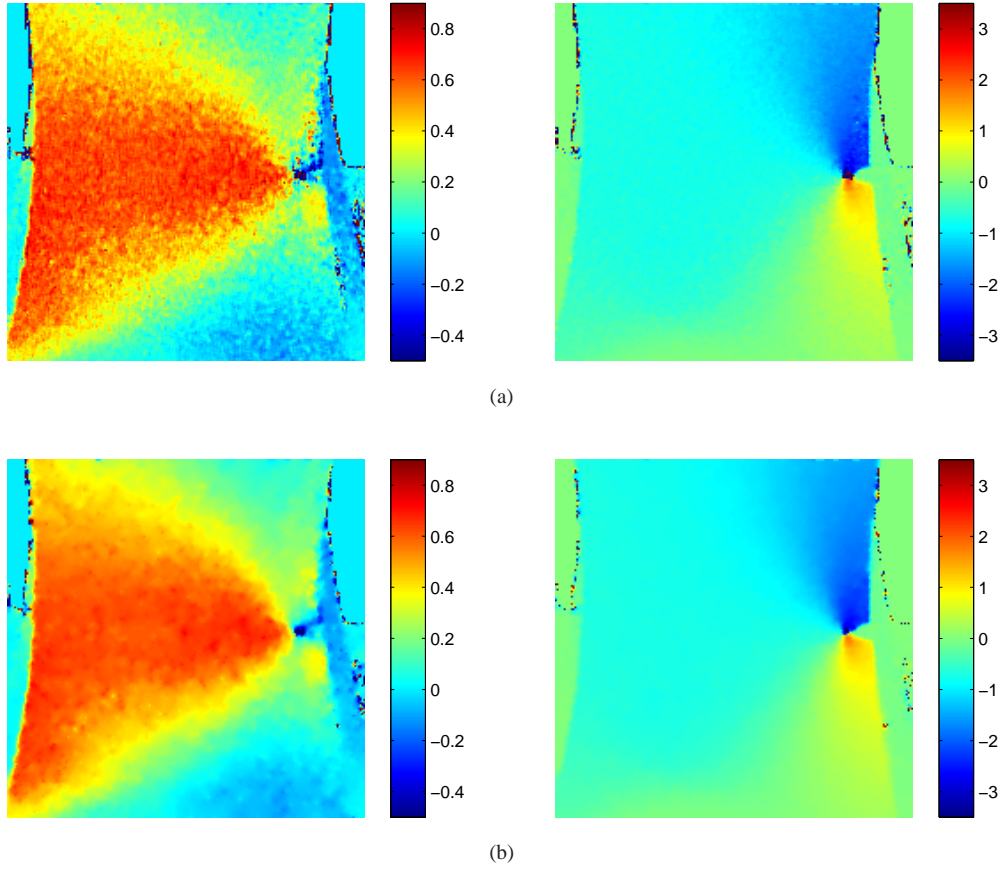
(a)

(b)

Fig. 12: The horizontal (left) and vertical (right) displacements calculated with (a) the classic and (b) proposed robust DIC approaches ($\mu_S = 3000$) using a subset size of 15×15 pixels, in the second experiment.
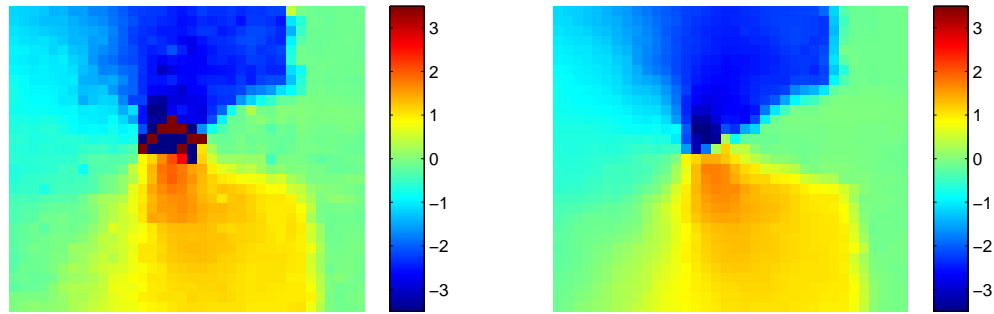

Fig. 13: Vertical displacements (36×31 displacement points detail) from the second experiment calculated with the classic (left) and proposed robust (right) DIC approaches.

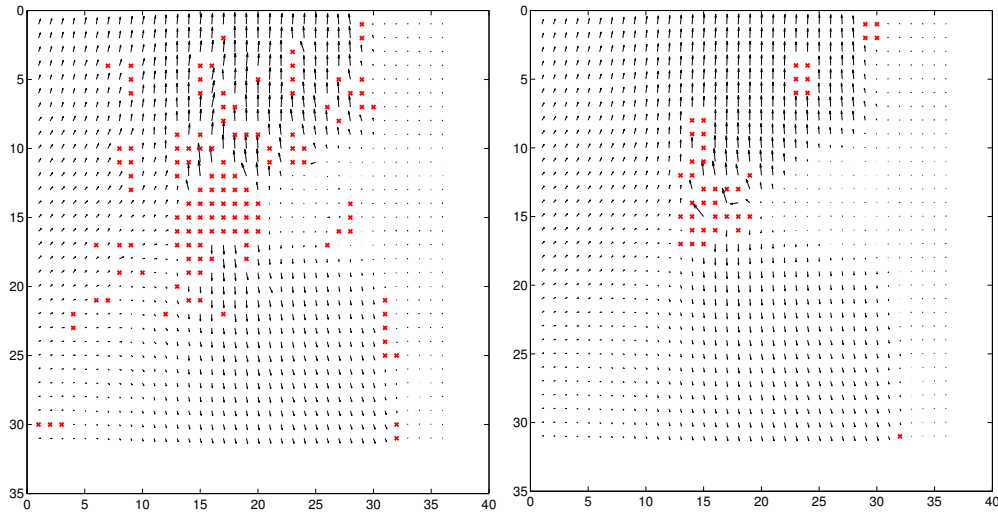Fig. 14: Motion vectors (36×31 displacement points detail) from the second experiment calculated with the classic (left) and proposed robust (right) DIC approaches. The locations marked in red represent displacement solutions that did not converge.

methods along with those of other optical flow methods are presented in Table 1. The resulting pixelwise displacement fields for the two methods are illustrated in Fig. 15. It should be noted that displacements which were not reliably calculated (e.g. the convergence conditions of the Newton-Raphson method were not met) were not used in the calculus of the AEE for either of the methods. These displacements can be easily noted in the displacement results as they present a sharply distinct colour coding compared to the neighbouring displacements.

From the results presented in Fig. 15 it is clear that the classic DIC method approach cannot reliably recover displacements in significantly more image regions than the proposed one. Moreover, except for the 'Hydrangea' image sequence, where the proposed DIC method is slightly outperformed by the classic one, the numerical figures from Table 1 support this observation. Both DIC methods yield poor results at the edges of the images and in the areas where there is lack of texture doubled by motion discontinuities, which is the case in the 'Venus' sequence. These are situations that do not present interest in DIC experiments where object texture is always present. In the image areas where sufficient texture is present, such as those corresponding to the plant in 'Hydrangea' and the background in 'Grove2', both methods yield reliable displacements. A particular effect of using subsets is the inability to recover finer displacement spatial variations or the exact boundaries of motion discontinuities. This is due to the fact that a subset step larger that one pixel has been used and in consequence the obtained displacement fields are a piece-wise estimation of the real, underlying displacements. These effects are present mostly in the 'Hydrangea', 'Venus' and 'Grove2' sequences. These are also the sequences where the proposed DIC method's performance is most impacted compared to the best performing optical flow methods. The largest displacement errors are found at the edges of the plant for 'Hydrangea', the edges of all objects in 'Venus' and the edges of the foliage and branches in the 'Grove2' sequence. In the case of the 'Dimetrodon' sequence, displacement discontinuities, which correspond to the edges of the toy dinosaur, are smaller in size and hence, the impact of the errors at the generated by motion boundaries is smaller. This explains why the proposed method performed similarly to the better performing state-of-the art optical flow methods.

(a) 'Dimetrodon'

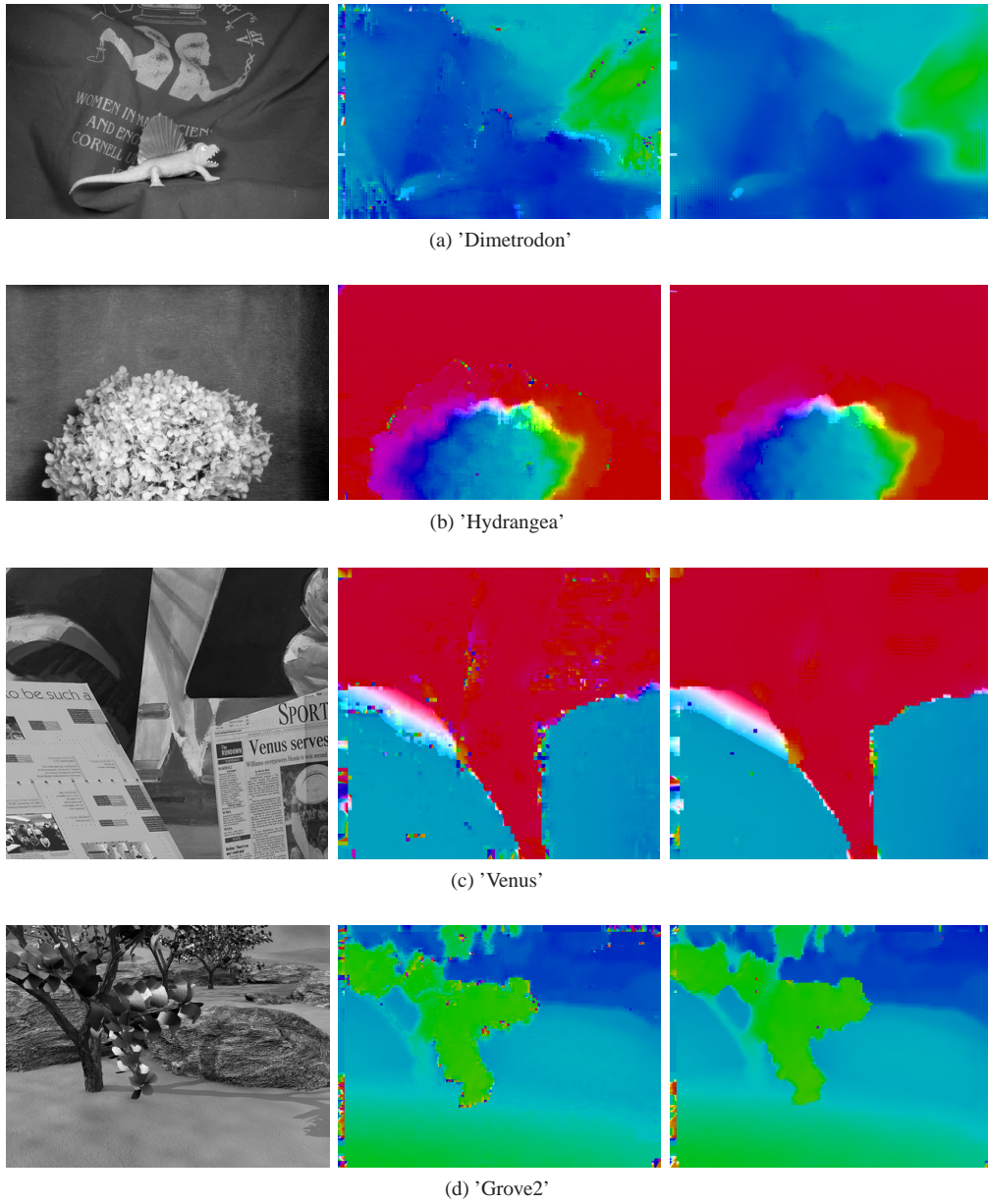(b) 'Hydrangea'

(c) 'Venus'

(d) 'Grove2'

Fig. 15: First image of the each sequence (left column) and pixelwise optic flow fields recovered with the classic (middle column) and proposed robust (right column) DIC approaches in the third experiment. Both DIC methods employed a 21×21 subset size with a step size of 5 pixels.

Table 1: Average endpoint errors (AEE) for the classic and proposed robust DIC methods for the Middlebury training dataset. Both DIC methods employed a 21×21 pixel subset size and a 5 pixel step size. The errors corresponding to other state-of-the-art optical flow methods are presented for reference. The results marked with * are taken from [31] and [32] respectively.

Method	Average endpoint error (AEE)			
	'Dimetrodon'	'Hydrangea'	'Venus'	'Grove2'
Proposed DIC	0.17	0.23	0.36	0.27
Classic DIC	0.24	0.22	0.58	0.61
Black and Anandan* [26]	0.39	0.32	0.54	n./a.
Sun (Layers++, 2 frames) [33]	0.15	0.16	0.21	0.09
Weiss* [34]	0.18	0.25	0.51	0.22
Werlberger (TV-L^1) [35]	0.27	0.22	0.39	0.20
Werlberger (NL-TV-TNCC) [35]	0.16	0.15	0.29	0.14

4. Conclusions

In this paper, a new, fully robust DIC method that addresses the limitations of the, extensively used, quadratic subset similarity criteria, has been proposed. The method is designed to adaptively weigh image pixel differences when estimating displacements so that only pixels that best fit the overall motion inside a given subset contribute to the displacement solution. It also increases motion consistency through spatial displacement regularization, which preferentially penalizes small displacement discontinuities with respect to larger ones. This has the effect of smoothing small displacement discontinuities, which are generally a result of image and calculus noises while maintaining larger ones. In the method, two robust estimators are employed: the Welsch estimator in the pixel-level treatment outliers and the Geman-McClure estimator for regularization. The resulting robust subset similarity criterion is optimized using the popular Newton-Raphson optimization algorithm. The proposed DIC method has been compared to the classic one - that employs the least squares subset similarity criterion - in three experiments: the first experiment is numerical and simulates rigid body translations and displacement discontinuities in the presence of small occlusions and image noise. The second experiment uses speckle images captured in a real experiment consisting in the deformation of a plastic film specimen under uniaxial load. In the third experiment, the proposed and classic DIC approaches are compared to other optical flow methods using four image sequences from the Middlebury evaluation dataset. Results in all experiments indicate that the proposed method significantly outperforms the classic one in terms of accuracy. As a direct result of the use of robust estimation, not only the accuracy of displacements in the presence of image noise, displacement discontinuities and image occlusions is improved, but the number of non-converging displacement solutions is also significantly reduced.