High resolution and large field of view imaging using a stitching procedure coupled with distortion corrections

Ali Rouwane, Damien Texier, Jean-Noël Périé, John-Eric Dufour, Jean Charles Stinville, Jean-Charles Passieux

Highlights

High resolution and large field of view imaging using a stitching procedure coupled with distortion corrections

Ali Rouwane, Damien Texier, Jean-Noël Périé, John-Eric Dufour, Jean-Charles Stinville, Jean-Charles Passieux

- The proposed non-rigid stitching method estimates and corrects part of the optical distortions using gray-levels in overlapping regions.
- The method attempts to correct image stitching as best as possible using minimal information (small overlap, self-calibration, no extra images).
- It improves the quality of the blended image and greatly reduces blurring in the overlaps due to lens distortion.
- The method outperforms classic stitching by reducing displacement/strain artifacts when the obtained mosaics are analyzed with DIC.

High resolution and large field of view imaging using a stitching procedure coupled with distortion corrections

Ali Rouwane[a], Damien Texier[a,*], Jean-Noël Périé[a], John-Eric Dufour[a], Jean-Charles Stinville[b], Jean-Charles Passieux[a]

[a] *Institut Clément Ader (ICA), Université de Toulouse, INSA-ISAE-IMT Mines Albi-UPS-CNRS, Toulouse, France*
[b] *Materials Science and Engineering, University of Illinois at Urbana-Champaign, USA*

Abstract

A numerical approach was implemented to precisely stitch together images from the same projector/camera that form a mosaic by regularly moving either the projector/camera or the scene/sample. Such an imaging approach is used, for example, in automated microscopy. The presence of optical distortions can lead to detrimental blurring artifacts in the overlaps. The present development identifies and corrects non-affine distortion functions using the gray level conservation equation on reduced overlapping regions of adjacent images with sub-pixel accuracy. The present numerical development was first tested on synthetic images with known distortions to confirm that the algorithm is capable of detecting only non-affine distortions. Then, Digital Image Correlation (DIC) was applied to a pair of large laser scanning confocal microscope mosaics (121 images of 1024×1024 pixel^2, *i.e.*, more than 100 MPixel) created using the proposed non-rigid stitching. The method aims to improve the quality of blended images after stitching using the sample pattern in the overlapped regions. This new numerical development significantly minimizes kinematic field artifacts due to lens distortion in overlapped regions.

Keywords: Non-rigid stitching; Distortion compensation; Image blending; Mosaicing; Digital Image Correlation.

1. Introduction

A field of view is an open area an instrument or a person can observe. Large fields of view with a high spatial resolution are of interest due to the combined capability to represent a large scene with high spatial information describing the variability of the scene. Such large fields of view require high-resolution images that can be acquired either from full-field high-resolution sensors, from scanning acquisition systems, or from observation strategies such as the pixel shift technique, sweeping panoramas, mosaicing, or an array of sensors, or even cameras. Each acquisition technique has advantages and drawbacks that make images of the same physical size and numerical resolution different. For instance for optical techniques, commercial full-frame ultra-high resolution sensors can provide up to 100 megapixels (MP) images on

*Corresponding author
Email addresses: rouwane@insa-toulouse.fr (Ali Rouwane), damien.texier@mines-albi.fr (Damien Texier), jean-noel.perie@iut-tlse3.fr (Jean-Noël Périé), jedufour@insa-toulouse.fr (John-Eric Dufour), jcstinv@illinois.edu (Jean-Charles Stinville), passieux@insa-toulouse.fr (Jean-Charles Passieux)

centimeter-size sensors (*e.g.*, 102 MP BSI-CMOS 43.8×32.9 mm^2 medium format sensor with a pitch size of 4.6 µm or HASSELBLAD H6D-400C MS 53.4×40.0 mm^2 sensor with pitch size of 4.6 µm). In recent years, the development of megapixel optical sensors for smartphones has allowed significant technological and manufacturing advances, sometimes at the expense of sensitivity, dynamic range, and physical resolution due to numerical aperture restrictions. Micro-sized smartphone sensors can now be of 200 MP-resolution images with photosites of 0.64 µm, as demonstrated with the Samsung Isocell HP1 sensor. Furthermore, combining such high-resolution sensors with pixel refinement techniques, such as the pixel shift technique, can quadruple the numerical resolution of the images [1]. To circumvent individual sensor performance, an alternative solution for higher-resolution images is the assembly of juxtaposed sensors. For instance, the SLAC laboratory at Stanford University has developed for the Large Synoptic Survey Telescope (LSST) of Vera C. Rubin Observatory project the largest digital photo sensor of 3.2 gigapixels (GP) measuring more than 600 mm on a side and is made up of 189 smaller 16 MP sensors (pitch of 10 µm) [2]. However, this approach can be limited by the diffraction and aberration limits imposed by the numerical aperture of the monocentric lens. Therefore, the development of multi-sensors with a monocentric ball lens and an additional relay lens for each sensor (few GP-resolution images for [3] and up to 50 GP-resolution images for the AWARE-2 camera [4]) or multi-camera array microscopes (MCAM) [5] have demonstrated their ability to generate GP-resolution images. All these techniques have the advantage of acquiring high-resolution images in a single shot, leading to relatively fast acquisition times.

Images can also be constructed using scanning acquisition techniques, such as scanning electron microscopy, laser scanning confocal microscopy, scanning acoustic microscopy, and atomic force microscopy. These techniques are limited by electronic controllers, sometimes encoded in formats not sufficiently resolved to accurately capture the signal at a precise, accurate, and repeatable position, compared to fix arrays of full-field sensors. For instance, scanning electron microscopy (SEM) uses deflection coils controlled by a deflection coil current whose discretion and resolution are given by the controller encoder. Recent developments for scanning electron microscopy allow for micrographs resolution of about 50k×40k pixels in a single acquisition, *i.e.* about 2 GP, with the Atlas 5 functionality proposed by ZEISS on the ZEISS Crossbeam series (ZEISS Crossbeam and Sigma series already offer images of 32k×24k pixels) [6, 7]. However, the larger the images, the longer the acquisition times. Therefore, as for multi-sensor optical means, multi-beam electron microscopy was developed for high-throughput imaging and allows for the acquisition of 91 simultaneous images, leading to an unprecedented field of view size and resolution for fast acquisition times [8, 9].

Considering and possibly correcting image artifacts is the first concern when using imaging for metrology. In scanning electron microscopy and other scanning techniques, artifacts in image reconstruction can arise from beam position accuracy and repeatability, electronic drift (electrical charges of the sample or controllers), but also from astigmatism and spherical aberrations, i.e., optical lens distortions [10]. Non-stationary step-changes in the measured displacements were evidenced using digital image correlation (DIC) techniques, leading to anisotropic artifacts depending on the scanning strategy (continuous

scanning along a direction versus line jumps in the orthogonal direction [10, 11, 12, 13, 14, 15]. Different algorithms to correct artifacts from line jumps are developed, but are not of concern for the present paper [10, 11, 14, 15]. For scanning drift or aberrations, progressive distortions can also be identified and corrected using cross-grating techniques during image acquisition [12, 16]. Recently, Maraghechi et al. [17] proposed a unified integrated DIC (IDIC) framework extended with a series of hierarchical mapping functions to concomitantly correct spatial distortion, drift distortion and scan line shifts. Hardware developments for SEM incorporating phase-locked loop control aimed at minimizing positional errors during e-beam scanning [18].

If time acquisition is not an issue, acquiring successive images at different locations with overlapping regions is generally less expensive and may benefit from more appropriate lenses for aberration and diffraction limits. Sweeping panoramas [19] or image stitching [20, 21] techniques can then be used to reconstruct the multi-region blended image. Large but high-resolution stitched images are reported, for instance, in Refs. [22, 9, 23] for collecting time of hours to days thanks to automated motion and acquisition systems. Mosaics of images generally consist of taking an image sequence from the same viewpoint, leading to homographies (or collineations) relationships between successive images.

Combining multiple images into one panoramic (or stitched) image can be performed with different strategies. They can be classified into two main groups: feature-based [24, 25] and registration-based approaches. The most common approach is based on feature detection. In panoramic image stitching, it is common to choose large overlapping regions to get a sufficiently high number of feature points. These feature points are usually found using robust feature detection techniques such as SIFT [26], Harris [27], SURF [28], and many others (see *e.g.* [29, 30] as an example). Once the features are detected, projective transformations represented by a 3×3 homography matrix are identified using these data points [31]. An outlier filtering using the RANSAC algorithm [32] is also used to get a more accurate homography. We note that the feature matching technique step can also be performed using phase correlation techniques, *i.e.* by maximizing the cross-correlation between multiple feature windows *a priori* fixed.

In the presence of geometric distortions, simple affine/projective transformations that match overlapping regions cannot allow accurate stitching of images, as distortions are usually non-linear. This becomes very problematic, especially when the overlap size is small (which is a practical choice for minimizing the acquisition time and the memory footprint). This induces severe limitations, as it does not dissociate the distortion and the simple affine transformations related to the physical multi-view acquisition (such as the sample/microscope translation). This leads to an inhomogeneous blurring effect that results from averaging different matched gray-level values at the overlapping regions. When performing Digital Image Correlation (DIC) with high-resolution images stitched as such, these distortions can affect the quality of the displacement and strain fields [33, 34].

To limit the effects of distortion, stitching techniques have been developed that look for a deformation of each image in addition to its rotations and translations. They are sometimes referred to as *non-rigid* stitching. The deformations of the images are usually chosen as second [35] or higher order [36, 37]

polynomials whose coefficients are optimized to better align the matching points. The goal of these methods is generally to produce *visually better* mosaics, in other words, so that the imaged objects are neither discontinuous nor blurred. However, such methods are not suitable for digital image correlation, since these deformations, that differ between two adjacent images–and sometimes within a single image–, generate nonphysical strain discontinuities.

Several techniques have been developed to estimate and correct optical, *i.e.*, more physically-sound, distortions in DIC and stereo DIC. The major sources of distortions related to optical lenses are radial, decentering, and prismatic distortions and are taken into account in so-called "camera models" [38]. Most calibration procedures rely on detecting fiducial markers such as circles, ellipses, lines, line intersections, checkerboard, or more complex patterns [39] printed or engraved on calibration objects. In addition, photogrammetric calibration methods enable calibration without very precise and accurate calibration objects [40, 33]. The pattern printed on the sample is considered known and bias-free; it is thus possible to perform calibration with one single image [41]. Using microscopy, *e.g.* optical microscopy, laser scanning confocal microscopy, or electronic microscopy, it is experimentally complex to first acquire or manufacture calibration objects based on the fabrication tolerances meeting the spatial resolution of distortions to be identified. In addition, positioning and moving a calibration object with the same flatness - and electron charges for scanning electron microscopy - in the field of view of the imager prior to the experiment to calibrate distortions is not trivial.

In this study, an algorithm is proposed to make the most of a minimal dataset, *i.e.*, composed solely of mosaic sub-images (no other images, no new experiments with different operating conditions like change in scanning direction or change in overlap, no *a priori* knowledge of the pattern) and with a relatively small overlap (10 %) to maximize the field of view with automated image acquisition for mosaics/grid collection using microscopes. The idea is not only to reconstruct a mosaic image without blurring in the overlapping regions, but also to apply homogeneous physically-based corrections so that the high-resolution images can be used in DIC *a posteriori*. The present algorithm consists of a registration that dissociates the rigid translations from the assumed intrinsic optical distortion of the imaging system. In the spirit of the work of [42], a calibration procedure for high order camera models is proposed given the reduced-size overlap. This type of method is considered as a self-consistent calibration, since it does not require a calibration object. A minimization of a global Sum of Squared Distances (SSD) metric over all the overlapping regions of the mosaic is thus considered. The optimization variables are the position of the images in a global reference coordinate system and the parameters of the considered distortion model. Finally, the corrected sub-images are stitched with standard phase correlation and blending approaches [43, 44]. The present work can be seen as an enhancement of the 2D image stitching as established in the work of Preibisch et al. [45] and Kaynig et al. [46] (available in the image-processing package Fiji). Interestingly, the present correction method is not specific to a microscopy or any image acquisition apparatus, and choices of correction modes may depend on image modality. It is worth reminding that the present model only applies to continuous distortion fields and is not applicable to line jump artifacts that may occur in bilateral scanning microscopy, such as electron microscopy [14, 17].

For the validation of the method, the algorithm was first tested on synthetic dataset with known distortions to confirm the validity of the present algorithm and modes that can be identified, *i.e.* non-affine distortions. A methodology is then proposed for selecting the distortion modes to be considered for real images with improved sub-pixel corrections using the proposed approach on a real high-resolution mosaic image. The developed code is available at `https://github.com/arouwane/Distortion-correction` [47].

2. Problem setting

The present numerical development intends to improve stitching operations considering lens distortion functions on reduced-size overlapping regions after an automated acquisition of a mosaic of images. For such a configuration, the imaging system (microscope) is fixed, and the sample is regularly translated in both horizontal and vertical directions to acquire the different sub-images.

Let f_1 and f_2 be two adjacent distorted images of the considered sample (see Fig. 1). Registration of f_1 and f_2 is usually performed by satisfying the gray-level conservation equation $f_1 = f_2 \circ \boldsymbol{\phi}$ using similarity criteria [44] or using features matching [31]. In the gray-level conservation equation, the unknown parameter is the transformation field $\boldsymbol{\phi}$ which, in the presence of distortions, is a complex composition of the translation and the intrinsic distortions related to the imaging system. If no distortion is considered, then $\boldsymbol{\phi}(\mathbf{x}) = \mathbf{x} + \Delta\mathbf{t}$ where $\Delta\mathbf{t}$ is the translation between both images. In this case, the sum of squared differences dissimilarity can be considered as follows:

$$\boldsymbol{\phi}^\star = \underset{\boldsymbol{\phi}}{\arg\min} \quad \sum_{\mathbf{x}_i \in \Omega} \left(f_1(\mathbf{x}_i) - f_2 \circ \boldsymbol{\phi}(\mathbf{x}_i)\right)^2 . \tag{1}$$

In the presence of distortions, $\boldsymbol{\phi}(\mathbf{x}) = \mathbf{x} + \boldsymbol{\Delta}\mathbf{u}$ where $\boldsymbol{\Delta}\mathbf{u}$ is a complex composition of both the translation $\boldsymbol{\Delta}\mathbf{t}$ and the intrinsic distortions related to the imaging system (given by the distortion degrees of freedom vector denoted $\mathbf{d}$). To dissociate the translations from the supposed common distortion, one can reformulate the previous dissimilarity as follows:

$$\boldsymbol{\Delta}\mathbf{t}^\star, \mathbf{d}^\star = \underset{\boldsymbol{\Delta}\mathbf{t}, \mathbf{d}}{\arg\min} \quad \sum_{\mathbf{x}_i \in \Omega} \left(f_1 \circ \mathbf{P}\left(\mathbf{x}_i, \mathbf{d}\right) - f_2 \circ \mathbf{P}\left(\mathbf{x}_i - \boldsymbol{\Delta}\mathbf{t}, \mathbf{d}\right)\right)^2 , \tag{2}$$

where $\mathbf{P}$ is the distortion correction transformation of the considered acquisition system. It is a two-dimensional vector field, defined as the inverse of the distortion transformation.

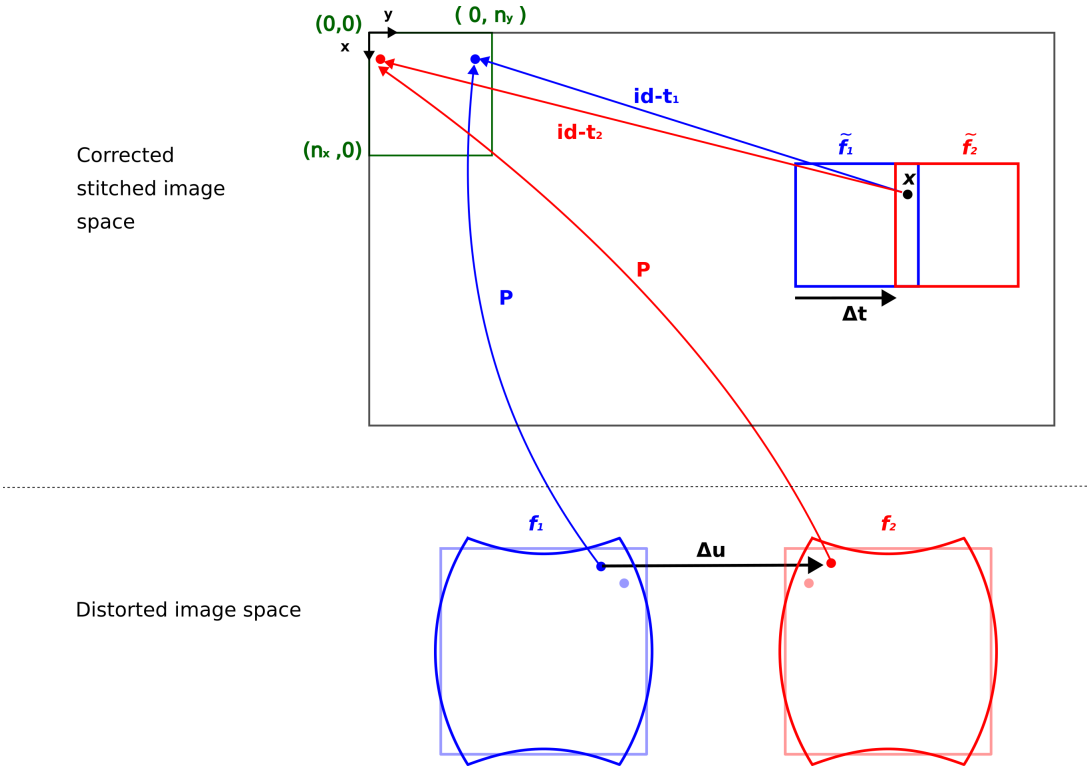

Figure 1: Schematic illustration of the image stitching problem with the presence of distortions.

In the present paper, a brief development of the gray-level conservation equation was considered when performing image registration between two adjacent distorted images. The relationship between the registered displacement field $\mathbf{\Delta u}$ and the distortion correction $\mathbf{P}$ was defined. Let $\tilde{f}_1$ and $\tilde{f}_2$ be the undistorted images corresponding to f_1 and f_2 (see again Fig. 1). Supposing that $\mathbf{\Delta t}$ is the exact translation of $\tilde{f}_2$ relative to $\tilde{f}_1$, then $\tilde{f}_2 = \tilde{f}_1 \circ (\mathbf{id} - \mathbf{\Delta t})$ where $\mathbf{id}$ is the identity operator. In the presence of distortions, $\tilde{f}_1$ and $\tilde{f}_2$ are composed by the inverse projector $\mathbf{P}^{-1}$ (provided that the distortion field is invertible). Performing image registration in the overlapping region subsequently consists of solving the following equation:

$$f_1 \circ \mathbf{P}^{-1} = f_1 \circ (\mathbf{id} - \mathbf{\Delta t}) \circ \mathbf{P}^{-1} \circ (\mathbf{id} + \mathbf{\Delta u}). \tag{3}$$

This results in:

$$\mathbf{P}^{-1} - \mathbf{P}^{-1} \circ (\mathbf{id} + \mathbf{\Delta u}) + \mathbf{\Delta t} = 0 \tag{4}$$

Eq. (4) establishes the relationship between the unknown distortion field defined by the transformation $\mathbf{P}$ and the registered displacement field between two overlapping images $\mathbf{\Delta u}$. When the distortion field $\mathbf{P}$ is such that it can perfectly superimpose to itself, the registered displacement field $\mathbf{\Delta u}$ is reduced to a pure rigid body translation. If $\mathbf{P}$ is such that $\mathbf{\Delta u}$ is a pure rigid body translation for all points in the overlapping region, then the stitching artifacts disappear and the stitching is perfect. Consequently, it is not possible to characterize such a distortion using only the gray levels in the overlap. The condition for this is explained below.

First, let us notice that equation (4) can be rewritten as:

$$\mathbf{P} - \mathbf{P} \circ (\mathbf{id} + \mathbf{\Delta t}) - \mathbf{\Delta u} = 0. \tag{5}$$

Then, using a first order Taylor expansion of (5), it can be shown that:

$$\mathbf{J_P \Delta t} = \mathbf{\Delta u}. \tag{6}$$

Since $\mathbf{\Delta}t$ is constant, if the Jacobian $\mathbf{J_P}$ is constant, so is $\mathbf{\Delta u}$. The distortion mode $\mathbf{P}$ will be measurable if its Jacobian is not constant in the overlap. Conversely, this means that any (component of the) distortion field that is affine in the overlap cannot be calibrated using only overlapping gray-levels. In practice, modes that almost fulfil this condition (i.e., are almost affine in the overlap) will also be difficult to measure. But that's not a big problem for stitching, since these affine (or quasi-affine) modes don't produce stitching artifacts. In the following, we just have to be careful not to introduce affine or quasi-affine modes in the parameterization of the distortion field to ensure that the calibration problem is well posed.

In the rest of the paper, the word *affine* will designate affine and quasi-affine functions in the overlap, in other words, functions that won't be measurable in this context. The word *calibration* is used but is a bit of a misnomer, since the proposed method will consist of estimating and correcting only the part of the distortion field that results in a non-rigid $\mathbf{\Delta u}$, as it is the source of the stitching artifacts, as explained in Section 3. Eq. (4) is a complex non-linear equation that can be numerically solved, but a first registration step is, however, needed to determine the $\mathbf{\Delta u}$ field. This step is implicitly considered in Eq. (2). The registration cost function used to identify the distortion field will be based on Eq. (2) and detailed in Section 4.

3. Non-parametric distortion modes

Before solving the distortion identification problem, the distortion model must be first set. The general form of the distortion correction can be defined as follows:

$$\mathbf{P}(\mathbf{x}, \mathbf{d}) = \mathbf{x} + \mathbf{N}(\mathbf{x})\mathbf{d}, \tag{7}$$

where $\mathbf{x} = (x, y)$ is the 2D position in the deformed (or distorted configuration denoted $\tilde{\Omega}$, see *e.g.* Fig. 1), $\mathbf{N}(\mathbf{x})$ is the 2D distortion shape function matrix and $\mathbf{d}$ are the distortion correction modes. The distortion shape functions can derive from a physical or a parametric model or even a constrained mesh-based (or B-spline) model. One of the most common models is the Brown-Conrady model, which is a camera distortion model generally used to account for optical aberrations related to spherical lenses. It is a geometric distortion model representing radial and tangential variations of the field of view [38], as depicted in Fig. 2. A simple first-order Brown-Conrady correction field can be defined as follows [41]:

$$\begin{aligned} \mathbf{P}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \mathbf{d}_0\tilde{x}r^2 + \mathbf{d}_1(r^2 + 2\tilde{x}^2) + 2\mathbf{d}_2\tilde{x}\tilde{y} \\ \mathbf{d}_0\tilde{y}r^2 + 2\mathbf{d}_1\tilde{x}\tilde{y} + \mathbf{d}_2(r^2 + 2\tilde{y}^2) \end{pmatrix} \\ &= \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 2\mathbf{d}_2\tilde{x}\tilde{y} + 3\mathbf{d}_1\tilde{x}^2 + \mathbf{d}_1\tilde{y}^2 + \mathbf{d}_0\tilde{x}\tilde{y}^2 + \mathbf{d}_0\tilde{x}^3 \\ 2\mathbf{d}_1\tilde{x}\tilde{y} + \mathbf{d}_2\tilde{x}^2 + 3\mathbf{d}_2\tilde{y}^2 + \mathbf{d}_0\tilde{y}\tilde{x}^2 + \mathbf{d}_0\tilde{y}^3 \end{pmatrix}, \end{aligned} \tag{8}$$

where $\tilde{x} = (x - x_c)/l$ and $\tilde{y} = (y - y_c)/l$ and $r = \sqrt{\tilde{x}^2 + \tilde{y}^2}$. l is a characteristic scaling length (in pixels) used for normalization. l can be chosen as the maximum length of the mosaic sub-images. (x, y) represents the coordinates in the image coordinate system. (x_c, y_c) is the distortion center (of zero distortion), which is assumed to be fixed at the image center.

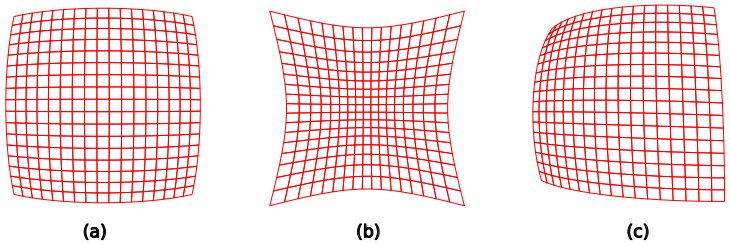

Figure 2: Examples of lens distortions obtained from the Brown-Conrady model of Eq. (8). (a) Positive radial distortion. (b) Negative radial distortion. (c) Positive radial and tangential distortions.

The radial and tangential transformations modeled by this parametric model are limited and cannot allow identifying general transformations that derive from complex optical paths, as is the case for microscopes with corrected lenses. In the context of this work, a simple polynomial description is rather considered. Therefore, the displacement field is expressed in the basis given by Eq. (9) for a cubic polynomial model (without the constant shift mode).

$$\{\tilde{x}, \quad \tilde{y}, \quad \tilde{x}\tilde{y}, \quad \tilde{x}^2, \quad \tilde{y}^2, \quad \tilde{x}^2\tilde{y}, \quad \tilde{x}\tilde{y}^2, \quad \tilde{x}^3, \quad \tilde{y}^3\}. \tag{9}$$

For the complete cubic polynomial model of Eq. (9), there are 9 distortion modes for each direction. In order to illustrate the different distortion modes and to evidence "the identifiability" of the polynomial modes of Eq. (9), we consider visualizing a distorted grid, its horizontal and vertical shifts.

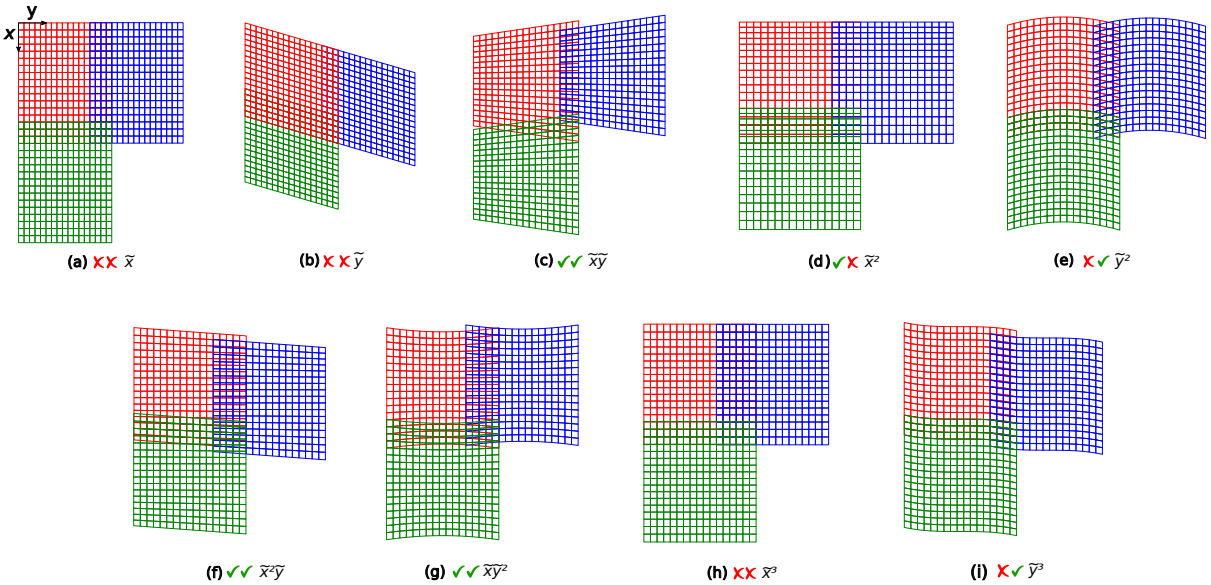

Figure 3: Vertical deformation of a structured grid (representing the image domain) with the different distortion modes of Eq. (9) for both the horizontal and vertical overlapping cases. (a)-(i) represent respectively the modes of Eq. (9) with positive coefficients.

Visually, affine and quasi-affine modes are the modes for which a translation can perfectly align the grid lines on the overlapping regions (see *e.g.* Figs. 3(a)-(b)-(h) for which the lines remain parallel for all translation values). In other words, there exists a translation that induces a problem with multiple solutions for the distortion parameters. As mentioned above, these affine modes cannot be measured as they induce a rigid-body distortion displacement vector. Affine and non-affine modes are represented as red cross and green check symbols in Fig. 3 for both directions, respectively. For clarity purposes, large distortion magnitudes were considered to illustrate each distortion mode in Fig. 3. Theoretically, one can choose the following modes:

$$\mathbf{N}(\mathbf{x}) = \begin{pmatrix} \mathbf{N}^x(x,y) & \mathbf{0} \\ \mathbf{0} & \mathbf{N}^y(x,y) \end{pmatrix} \text{ with } \begin{cases} \mathbf{N}^x(x,y) = \begin{pmatrix} \tilde{x}\tilde{y}, & \tilde{x}^2, & \tilde{y}^2, & \tilde{x}^2\tilde{y}, & \tilde{x}\tilde{y}^2, & \tilde{y}^3 \end{pmatrix} \\ \mathbf{N}^y(x,y) = \begin{pmatrix} \tilde{x}\tilde{y}, & \tilde{x}^2, & \tilde{y}^2, & \tilde{x}^2\tilde{y}, & \tilde{x}\tilde{y}^2, & \tilde{x}^3 \end{pmatrix} \end{cases} . \tag{10}$$

Remark 1. *Due to small distortion magnitudes and small overlaps, all the stretching modes in the dominant direction are not considered. (x, x^2, x^3 for the x direction and y, y^2, y^3 for the y direction). In this study, only significantly non-affine distortions are assumed to be relevant.*

4. Identification algorithm

Let $(f_i)_{i\in[\![1,m]\!]}$ be a set of distorted images representing a high-resolution mosaic of images of a planar sample. A regular grid of images was acquired in an automated process. To identify the distortion correction field, a global functional is considered by summing the squared gray-level differences over all the overlapping regions:

$$\underset{\mathbf{p}}{\arg\min} \quad \sum_{i=1}^{n} \sum_{\mathbf{x}\in\Omega_i} \left(f_i^0 \circ \mathbf{P}\left(\mathbf{x} - \mathbf{t}_{f_i^0}, \mathbf{d}\right) - f_i^1 \circ \mathbf{P}\left(\mathbf{x} - \mathbf{t}_{f_i^1}, \mathbf{d}\right)\right)^2 . \tag{11}$$

f_i^0 is the left (or top) image of the overlap Ω_i if Ω_i is vertical (or horizontal). f_i^1 is the right (or bottom) image of the overlap Ω_i if Ω_i is vertical (or horizontal). $\mathbf{t}_{f_i^1} > 0$ and $\mathbf{t}_{f_i^0} > 0$ are the coordinates of the images f_i^0 and f_i^1. The total vector of unknown parameters (translations and distortion modes) denoted $\mathbf{p}$ is defined as follows:

$$\mathbf{p} = \left(\mathbf{d}_1, ..., \mathbf{d}_n, {\mathbf{t}_\mathbf{x}}^1, {\mathbf{t}_\mathbf{x}}^1, ..., {\mathbf{t}_\mathbf{x}}^m, {\mathbf{t}_\mathbf{y}}^m\right)^T \in \mathbb{R}^{n+2m} \tag{12}$$

where $\mathbf{d}_1, ..., \mathbf{d}_n$ are the distortion correction parameters related to the optical system and ${\mathbf{t}_\mathbf{x}}^i, {\mathbf{t}_\mathbf{y}}^i$ are the translations (or position) of the image i. n represents the number of distortion parameters, and m is the number of images. Problem (11) is solved using a Gauss-Newton algorithm [48, 49]. The operators of this descent scheme are obtained from the assembly of elementary operators defined on each overlapping region. The local Hessian matrix $\mathbf{H}^i$ and right-hand side of the Gauss-Newton system $\mathbf{b}^i$ relative to an overlap Ω_i at iteration k are defined as follows:

$$\mathbf{H}^{k,i} = \sum_{\mathbf{x}\in\Omega_i} \mathbf{J}^{k,i}(\mathbf{x})\mathbf{J}^{k,i}(\mathbf{x})^T \quad \text{and} \quad \mathbf{b}^{k,i} = -\sum_{\mathbf{x}\in\Omega_i} \mathbf{J}^{k,i}(\mathbf{x})^T \left(f_i^0(\mathbf{x}) - f_i^1(\mathbf{x})\right) \tag{13}$$

where $\mathbf{J}^{k,i}$ is the local Jacobian matrix, computed for each region and defined as follows:

$$\mathbf{J}^{k,i}(\mathbf{x}) = \left(\boldsymbol{\nabla} f_i^0 \circ \mathbf{P}\left(\mathbf{x} - \mathbf{t}_{f_i^0}, \mathbf{d}\right)\right)^T \boldsymbol{\nabla}\mathbf{P}^0\left(\mathbf{x} - \mathbf{t}_{f_i^0}^k, \mathbf{d}^k\right) - \left(\boldsymbol{\nabla} f_i^1 \circ \mathbf{P}\left(\mathbf{x} - \mathbf{t}_{f_i^1}, \mathbf{d}\right)\right)^T \boldsymbol{\nabla}\mathbf{P}^1\left(\mathbf{x} - \mathbf{t}_{f_i^1}^k, \mathbf{d}^k\right). \tag{14}$$

The gradient matrices of the unknown transformation (relative to the distortion parameters and the two translations of the images) are defined as follows:

$$\boldsymbol{\nabla}\mathbf{P}^0 = \begin{pmatrix} | & | \\ \left(\frac{\partial \mathbf{P}_x}{\partial d_i}\right)_i & \left(\frac{\partial \mathbf{P}_y}{\partial d_i}\right)_i \\ | & | \\ -\frac{\partial \mathbf{P}_x}{\partial x} & -\frac{\partial \mathbf{P}_y}{\partial x} \\ -\frac{\partial \mathbf{P}_x}{\partial y} & -\frac{\partial \mathbf{P}_y}{\partial y} \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\nabla}\mathbf{P}^1 = \begin{pmatrix} | & | \\ \left(\frac{\partial \mathbf{P}_x}{\partial d_i}\right)_i & \left(\frac{\partial \mathbf{P}_y}{\partial d_i}\right)_i \\ | & | \\ 0 & 0 \\ 0 & 0 \\ -\frac{\partial \mathbf{P}_x}{\partial x} & -\frac{\partial \mathbf{P}_y}{\partial x} \\ -\frac{\partial \mathbf{P}_x}{\partial y} & -\frac{\partial \mathbf{P}_y}{\partial y} \end{pmatrix}. \tag{15}$$

The global operators of the Gauss-Newton linear system are parallelly assembled by adding the contribution of the operators of each local overlapping region. Concerning the optimization algorithm, a standard multi-level Gaussian blurring strategy is considered to gradually filter high-frequency displacement fields [50, 51]. When distortion is appropriately corrected, the global dissimilarity indicator is minimal and should only quantify the gray-level noise in the original images. In other words, the algorithm is considered efficient if the gray level disparity (defined by Eq. (16)) on all the overlapping regions remains in the same order of magnitude as twice the image noise.

$$\sigma\left(f_i^0 \circ \mathbf{P}\left(\mathbf{x} - \mathbf{t}_{f_i^0}, \mathbf{d}\right) - f_i^1 \circ \mathbf{P}\left(\mathbf{x} - \mathbf{t}_{f_i^1}, \mathbf{d}\right)\right) \tag{16}$$

σ corresponds to the standard deviation of a distribution. In practice, this disparity should not exceed 4 gray-level values. We note that a brightness offset correction is performed when computing the gray-level disparity on each overlapping region.

5. Proof of concept: Virtual 2D multiview with distortions

A virtual (synthetic) distorted multi-view of a speckle image is first considered, as illustrated in Fig. 4. This approach aimed at confirming modes, *i.e.* affine and non-affine modes, that can be identified with the present algorithm since ground-truth distortions and image translations are known. The whole high-resolution image, exempt from distortion, is subdivided into different individual images, denoted here as "sub-images". The sub-images were purposely subdivided to have an overlap of 10 %. Each sub-image is then distorted using known distortion modes to test the validity and robustness of the present calculation scheme. This approach aimed at confirming that it is possible to perfectly stitch distorted images even from a dataset that can not be used to detect affine modes. Three distortion models (see *e.g.* Fig. 5) were tested to validate the robustness of the distortion identification. To identify the distortion, different sets of modes were considered from Eq. (9). Table 1 synthesizes performance indicators for the different selections of distortion modes.

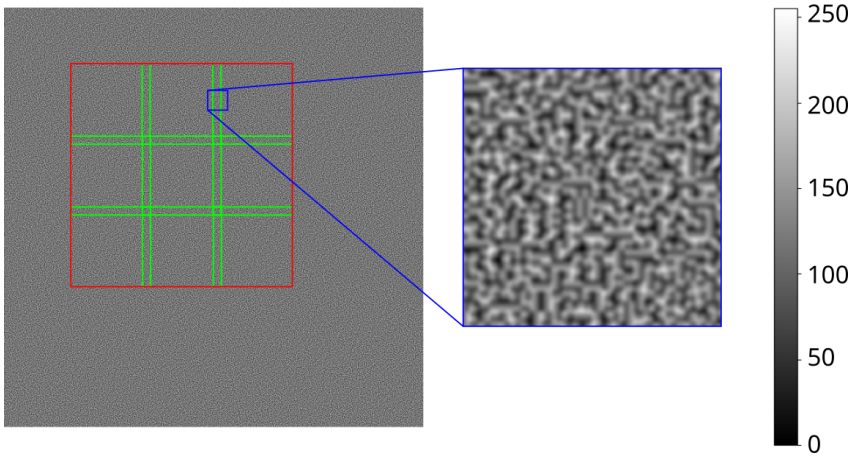

Figure 4: 3×3 synthetic multi-view image acquisition of a speckle texture.

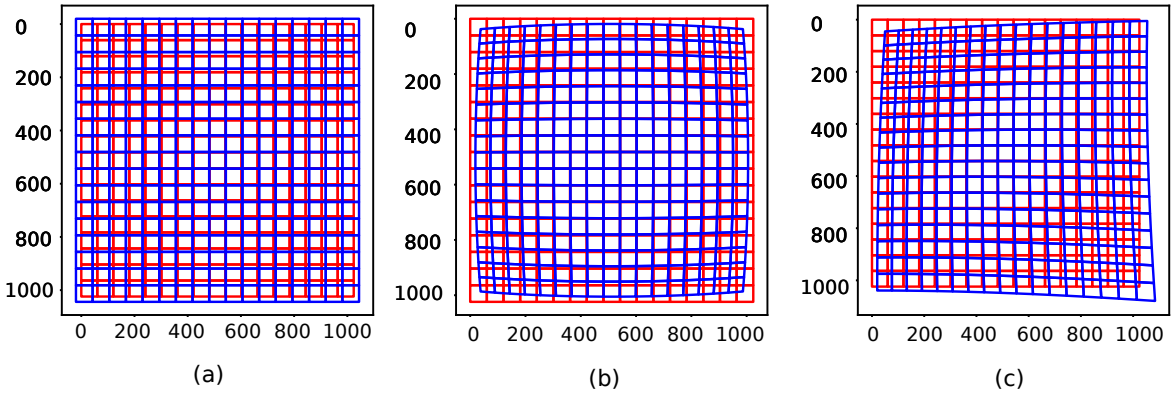

Figure 5: Deformation of a structured grid with three different distortion configurations. (a) Positive linear stretch. (b) Barrel distortion. (c) Pincushion + Tangential distortion. In red: reference configuration. In blue: deformed configuration.

Starting from a zero value vector $\mathbf{d}$ of the distortion parameters, the Gauss-Newton algorithm has an expected theoretical convergence. Let's now consider the first case corresponding to Fig. 5(a). It corresponds to a stretching distortion of the different sub-images by only considering x and y modes in x and y directions, respectively. The optimization algorithm defined in section 4 identifies a set of a zero distortion field and a set of translations that does not correspond to the reference translations (see the first row in Table 1). This proves that distortions due to affine modes (stretching effects in this case) can be compensated by overestimating the translation shift. Gray-level disparity is perfectly minimized, and images are perfectly stitched even if an (affine) part of the distortion field is not corrected. This confirms the choice of removing the corresponding two stretching modes. Finally, we consider two other multi-view distorted acquisitions using the Brown-Conrady model of Eq. (8). One is a barrel distortion (Fig. 5(b)), and the other is a pincushion-tangential distortion (Fig. 5(c)). When using Eq. (9), all the non-zero modes and the reference translations are accurately identified, and the gray-level disparity is perfectly minimized. However, if one uses the incomplete polynomial model of Eq. 10 then the Brown-Conrady distortion cannot be accurately identified (see again Table 1). To conclude, the analysis of

known distortions in synthetic images aimed at validating the present algorithm as a tool for stitching distorted images even if the full distortion model can not be estimated.

Model used for distorting the images	Identification modes	Gray-level disparity	Absolute Translation error (pixels)	Absolute Distortion parameters error
$(\tilde{x})\,, \quad (\mathbf{d}_0 = 8)$ $(\tilde{y})\,, \quad (\mathbf{d}_1 = 8)$	Eq. (10)	0.09	8	-
$\left(\tilde{x}\tilde{y}^2, \quad \tilde{x}^3\right)$ $\left(\tilde{y}\tilde{x}^2, \quad \tilde{y}^3\right)$	Eq. (10)	1.9	7	0.5
$\left(\tilde{x}\tilde{y}^2, \quad \tilde{x}^3\right)$ $\left(\tilde{y}\tilde{x}^2, \quad \tilde{y}^3\right)$	all except x and y	0.09	10^{-4}	10^{-4}
$\left(\tilde{x}\tilde{y}, \tilde{x}\tilde{y}^2, \tilde{y}\tilde{x}^2, \tilde{x}^3\right)$ $\left(\tilde{x}\tilde{y}, \tilde{x}\tilde{y}^2, \tilde{y}\tilde{x}^2, \tilde{y}^3\right)$	Eq. (10)	0.25	0.5	10^{-2}
$\left(\tilde{x}\tilde{y}, \tilde{x}\tilde{y}^2, \tilde{y}\tilde{x}^2, \tilde{x}^3\right)$ $\left(\tilde{x}\tilde{y}, \tilde{x}\tilde{y}^2, \tilde{y}\tilde{x}^2, \tilde{y}^3\right)$	all except x and y	0.09	10^{-4}	10^{-4}

Table 1: Identification of the distortions used to generate the synthetic images using different polynomial modes.

6. Application: Improvement of image registration accuracy for high resolution stitched images

6.1. DIC using confocal microscopy

In the present section, the algorithm will be tested on a real mosaic of micrographs with relatively regular translations thanks to microscope automation. Since ground-truth distortions are not known, a methodology is proposed to better select non-affine distortion modes using DIC results to minimize image deformation on the overlapped regions. For the present application, real images were acquired using a LEXT OLS 5100 laser scanning confocal microscope (LSCM) from OLYMPUS. The LSCM operates with a 405 nm wavelength laser and acquires images with a resolution of 1024 × 1024 pixels in a stitching/grid collection mode. A MPLAPON100XLEXT lens provided individual images with a field of view of 128 × 128 µm. A mosaic of 11 × 11 images was acquired in an automated manner thanks to a motorized XY stage. The adjacent images have an overlap of 10 %. The whole region of interest is depicted in Fig. 6. For information, each intensity LSCM micrograph is a reconstruction of a stack of 2D scans acquired at different height positions thanks to a motorized Z stage on the optical path (height pitch of 120 nm for a height range of quarters to hundreds of micrometers. The reconstructed intensity LSCM micrographs correspond to the brightest signal for each position in the 2D image over the Z integration. LSCM has the particularity to provide 3D images, but the out-of-plane information is not considered here [52]. The candidate material was an Alloy 718 (polycrystalline nickel-based superalloy).

A microtensile specimen was machined from the Alloy 718 plate and then polished down to a colloidal silica finish and an automated polishing jig to ensure an appropriate flatness and face parallelism of the microtensile specimen [53]. The microtensile specimen was then slightly pre-oxidized to form a speckle pattern compatible with the resolution of the LSCM. Mosaics of images were performed before and after tensile testing in the central region of the microtensile specimen to test the stitching algorithm with distortion corrections. The tensile test interrupted at 0.2 % plastic strain was carried out using an Instron 5800R electromechanical tensile testing machine with a load-cell of 5 kN using a constant crosshead displacement rate.

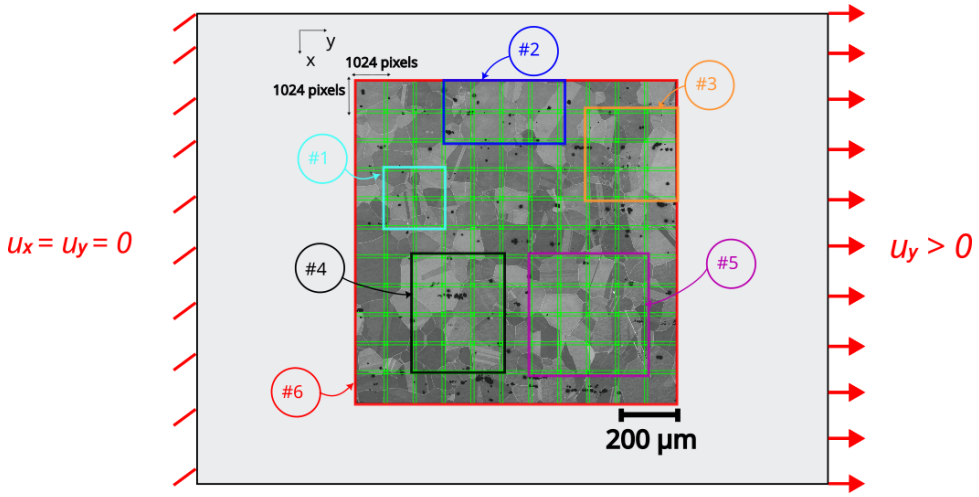

Figure 6: Illustration of the mosaic of reconstructed intensity LSCM micrographs (11 × 11 images) with a resolution of 1024 × 1024 pixels each for applying the stitching with distortion correction problem.

In this section, we consider applying the suggested identification algorithm detailed in Section 4 to approximate the unknown distortion transformation of the LSCM. To perform image stitching with distortion corrections, two strategies are considered. The first solution consists of selecting a small region in the mosaic of images, *i.e.*, a subset of images, and evaluating the distortion field using overlapping regions. A minimum of 2 × 2 subsets of images is necessary to identify the distortion parameters. For instance, different sub-set of images are illustrated in Fig. 6 and labeled from region #1 to region #5. The distortion correction field is then applied to all the images of the mosaic. For this purpose, one has to apply the correction field to the images. Sub-pixel gray-level interpolation was proceeded by bi-cubic spline interpolation. The second option would be to apply the optimization algorithm on the total map to identify at the same time the sub-image translations and the distortion field (region #6 in Fig. 6). This latter calculation is computationally longer than the calculation of a subset of images but adjusts the distortion parameters on all the overlapping regions. It is considered the reference calibration of the distortion model since the minimization of the gray-level conservation is minimal with this solution.

6.2. A first validation approach for the practical mode choice

As aforementioned, mode choice can be difficult, especially due to ignorance of the distortion function. Non-necessary modes can generate artifacts at the sub-pixel levels that digital image correlation (DIC)

techniques can track and quantify. However, this DIC technique requires comparison between a reference image and another to evaluate the displacement field and then image deformations. Therefore, the present optimization of mode choice needs the acquisition of two consecutive mosaics of images of the same region, denoted as *calibration region* and depicted in Fig. 7. Different sets of distortion modes were thus tested to highlight over-compensating distortion corrections leading to additional sub-pixel artifacts. As established in Eq. (10), odd linear stretching modes must be removed.

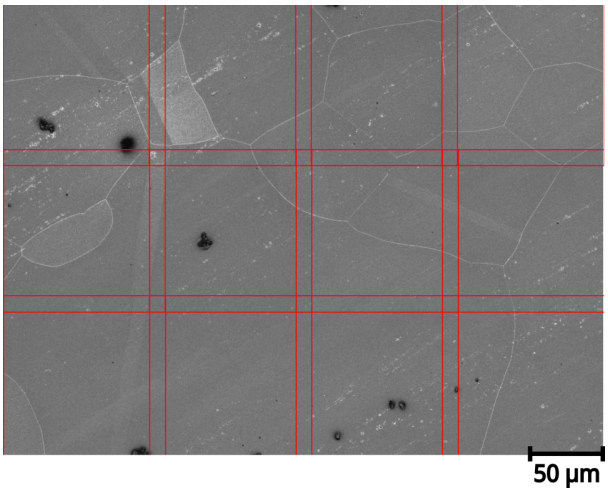

Figure 7: Calibration region. The image mosaic is composed of 3×4 images.

The first configuration consisted of conventional image stitching without correcting distortions and corresponded to the reference configuration - distortion corrections are supposed to improve the stitching operation. A sub-pixel registration between both images was then computed to identify local displacement related to errors due to the physical noise of the LSCM, scanning artifacts, and/or the stitching operation. ϵ_{xx}, ϵ_{yy}, and ϵ_{xy} strain fields are then calculated from displacement fields (Figs. 8(a)-(c)). Interestingly, intense strain values were found for strain maps along their principal direction, *i.e.*, horizontal red bands (y direction) for ϵ_{yy} and vertical bands (x direction) for ϵ_{xx}. The ϵ_{xy} maps are noisy, but artifacts are less pronounced compared to ϵ_{yy} and ϵ_{xx} fields.

Different sets of distortion modes were then tested and compared to the reference configuration. Selected distortion modes, along with related strain maps, were presented in Figs. 8(d)-(o). It is worth noting that high artifacts appear when considering the odd modes $\tilde{y}^3$ (in x direction) and $\tilde{x}^3$ (in y direction), especially on the ϵ_{xy} fields, as depicted in Figs. 8(j)-(o). Similar but less pronounced artifacts are found when adding the $\tilde{x}^2$ (in x direction) and $\tilde{y}^2$ (in y direction) distortion modes (Fig. 8(d)-(i)). Root-mean-square values of strain field components and gray-level residual fields are summarized in Table 2, and minimal values are observed for Case #1 and Case #2. These results confirm (i) that the present algorithm is capable of correcting distortion fields with sub-pixel accuracy to improve conventional stitching techniques, and (ii) that trying to over-correct images with non-necessary modes leads to higher gray-level residual values and stronger artifacts. Modes leading to the lowest root-mean-square values for both strain components and gray-level residual fields are the one to be selected for distortion

corrections of the whole dataset. These modes to be selected depend on the image acquisition configuration and are not generic. Interestingly, the modes that allow axi-symmetric distortion corrections exhibit strain fields with minimal artifacts and are defined by Eq. (17):

$$\begin{cases} \mathbf{N}^x(x,y) = \left(\tilde{x}\tilde{y}, \quad \tilde{y}^2, \quad \tilde{x}^2\tilde{y}, \quad \tilde{x}\tilde{y}^2, \quad \tilde{y}^3\right) \\ \mathbf{N}^y(x,y) = \left(\tilde{x}\tilde{y}, \quad \tilde{x}^2, \quad \tilde{x}^2\tilde{y}, \quad \tilde{x}\tilde{y}^2, \quad \tilde{x}^3\right) \end{cases} . \tag{17}$$

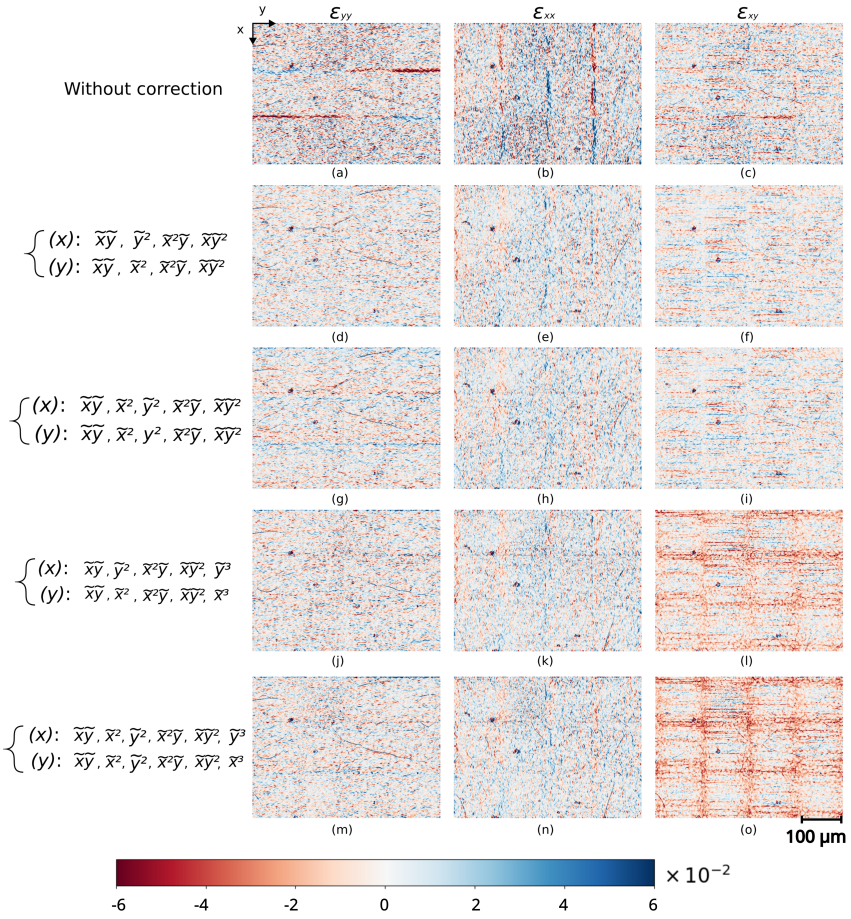

Figure 8: ϵ_{xx}, ϵ_{yy}, and ϵ_{xy} strain fields between the top two consecutive mosaics of images of the same region of interest, showing artifacts inherent to the selection of distortion modes.

	ε_{yy}	ε_{xx}	ε_{xy}	Residual RMS
Without correction	1.56×10^{-3}	1.41×10^{-3}	1.05×10^{-3}	0.76
Case 1 (Row 2-Fig. 8)	8.94×10^{-4}	9.63×10^{-4}	7.27×10^{-4}	0.74
Case 2 (Row 3-Fig. 8)	8.87×10^{-4}	9.79×10^{-4}	7.34×10^{-4}	0.74
Case 3 (Row 4-Fig. 8)	9.62×10^{-4}	1×10^{-3}	9.93×10^{-4}	0.77
Case 4 (Row 5-Fig. 8)	9.55×10^{-4}	1.05×10^{-3}	1.14×10^{-3}	0.77

Table 2: Root mean square of the strain field (RMS) components and the root mean square of the gray-level residual field. RMS quantities are computed only on the overlapping regions.

6.3. Identified distortion fields from subset of images and sensitivity analysis

Once the appropriate distortion modes were identified, the optimization algorithm was applied to the subset of images and the whole mosaic of images to compare the errors in displacement magnitude directly. An example of measured distortion fields, *i.e.* horizontal and vertical distortion fields, is illustrated in Fig. 9. As aforementioned, the correction field is defined by $\mathbf{P}$. To visualize the distorted image, the transformation $\mathbf{P}^{-1} - \mathbf{id}$ has to be applied to the initial field of view, as shown in Fig. 9. Interestingly, the distortion amplitude differs for the horizontal and vertical distortion, one with negative distortion parameters and the other with positive ones. The distortion function of the present optical setup is thus a combination of a positive and negative radial distortion of the Brown-Conrady model due to the complex optical path with mirrors and scanning galvanometers.

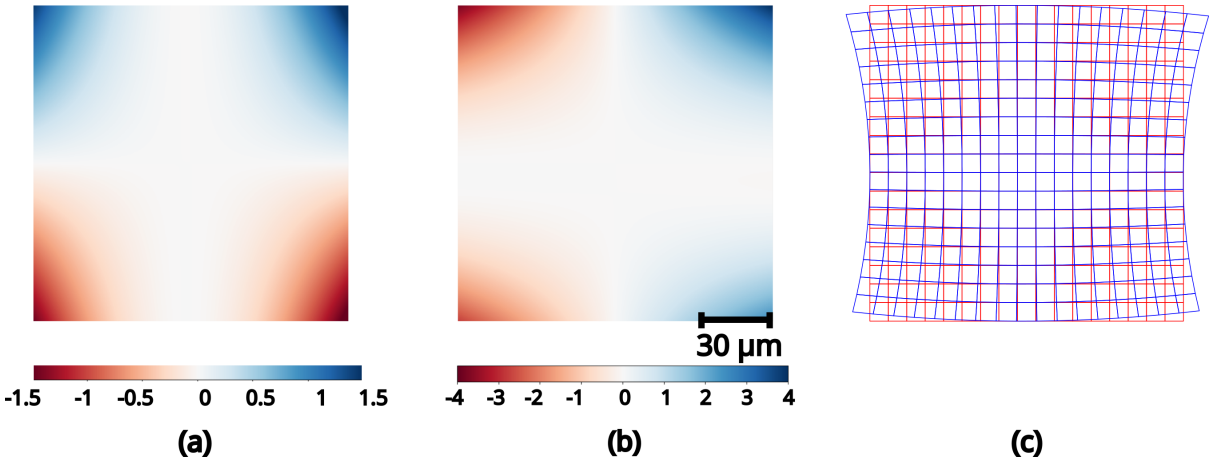

Figure 9: Identified distortion field defined by $\mathbf{P}^{-1} - \mathbf{id}$. (**a**) Vertical distortion. (**b**) Horizontal distortion. Deformed grid with the amplified field (factor of 20).

The influence of the calibration size for the distortion function identification was studied by measuring

the distortion fields for regions #1 to #6 (see Fig. 6). Interestingly, the measured distortion fields for regions #3, #4, #5, and #6 are very similar, with a maximum Root Mean Square Error (RMSE) of 10^{-1} pixels between the regions #1 and #6. The identified distortion differs from region #1 to region #2 with a maximal error of 0.5 pixels. Sensitivity parameters for the different region sizes are summarized in Table 3. While the optimization on all the images shows better results, one can pragmatically consider a calibration on a region size of 3×3 images to save computational time when needed.

Size	2×2	2×3	3×3	4×3	4×4
RMSE	1.47×10^{-3}	8.38×10^{-4}	4.2×10^{-4}	4.17×10^{-4}	3.72×10^{-4}

Table 3: Evolution of the RMSE (relative to the distortion field obtained from calibrating the 11×11 images) with respect to the size of the calibration target.

6.4. Stitching quality on the overlapping regions

To stitch the different sub-images of the global mosaic, a standard blending approach (using average or linear blending) was used as detailed in Appendix A. Overlapping regions of the blended image with and without distortion correction are shown in Fig. 11. The magnification on the horizontal and vertical overlapping regions demonstrates the improvement of the distortion correction prior to stitching, with blended regions being particularly blurred in the absence of distortion correction. As the distortion magnitude is greater for distortions in the x direction, the distortion correction particularly improves the blended regions. One can also visualize the gray-level residual field, which is the difference between two consecutive images on the overlapping regions. The improvement is highlighted in Fig. 10 as the residual field is brighter for the corrected images (see Fig. 10(a) against Fig. 10(b)).

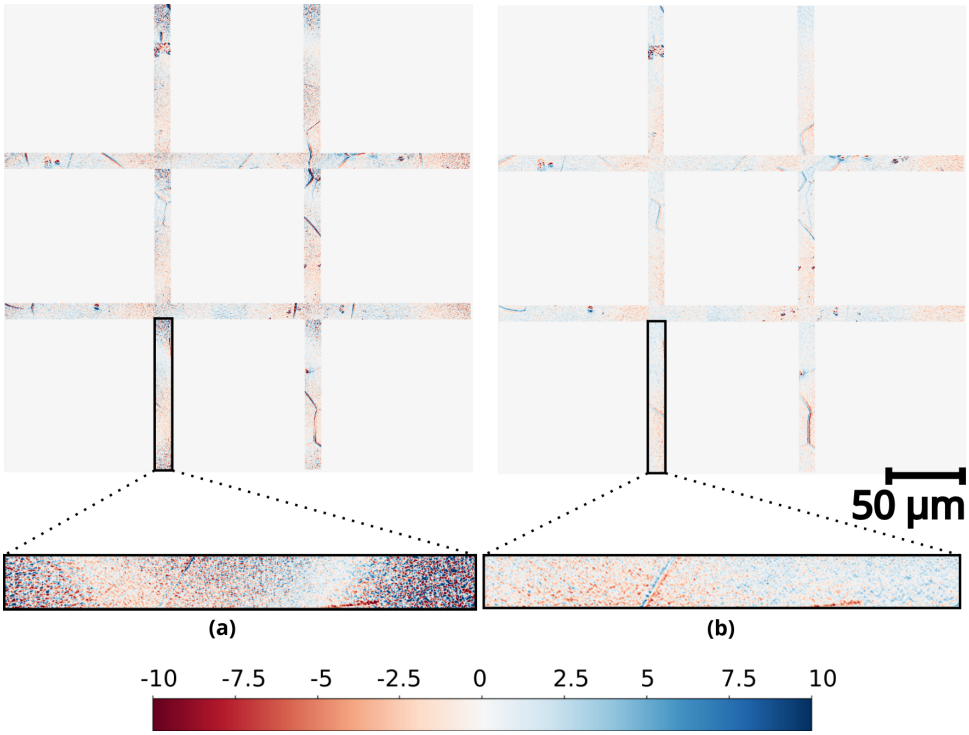

Figure 10: Residual field on overlapping regions. (a) Zoom on a region without correction. (b) Zoom on a region with correction.

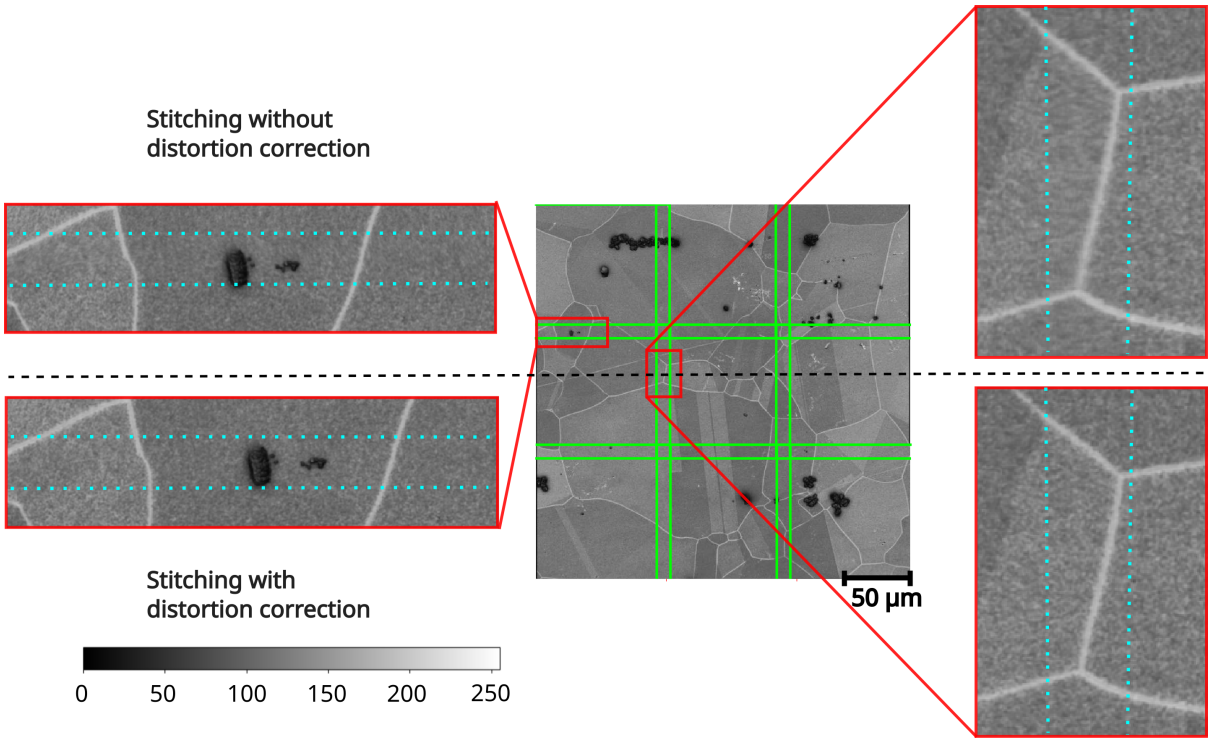

Figure 11: Effect of distortion on image stitching

The Laplacian of the corrected and uncorrected stitched images was compared to quantify the distortion blur coming from the stitching of distorted/corrected images. The Laplacian of the corrected

images (Fig. 12(b)) is homogeneous compared to the one of the uncorrected stitched image (Fig. 12(a)), showing vertical and horizontal brighter bands on the overlapped regions. This result demonstrates the improvement of the present stitching considering optical distortions. Finally, the mean of gray-level residual dispersion over all the regions of the 11×11 mosaic clearly shows that the disparity between the overlapping images is reduced when correcting the images (see *e.g.* Table 4).

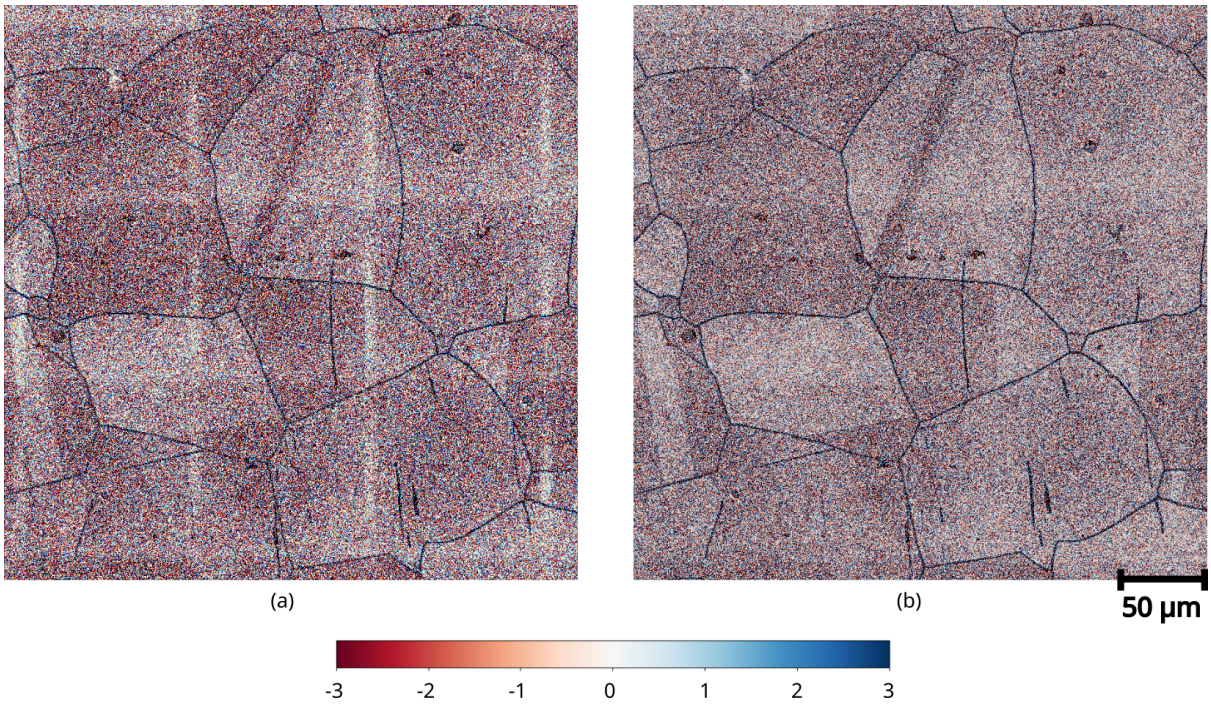

Figure 12: Laplacian of the stitched image of Fig. 11. Left (**a**) Without correction. Right (**b**) With correction.

Indicators	Max	Min	Mean	Std
Without correction	6.65	3.01	4.39	0.66
With correction	4.35	1.55	2.12	0.35

Table 4: Quantification of the stitching quality by considering statistics of the indicator (16) over all the overlapping regions of the mosaic.

6.5. Improvement of strain measurement accuracy

The present numerical development improved the gray-level contrast and sharpness of the blended images from the mosaic of images. However, this analysis is at the pixel level, while applying sub-pixel distortion during the image warping. To validate the robustness of the present numerical development, digital image correlation (DIC) techniques were applied to the blended images before and after small irreversible deformation to distinguish microplasticity in a metallic polycrystalline material from overlap artifacts. DIC is an efficient technique that allows identifying sub-pixel displacements (down to one-tenth of a pixel) when comparing two images together. In the present case, corrected and uncorrected stitched images before and after mechanical deformation were computed using the Dense Inverse Search optical

flow algorithm [54] available in the OpenCV library. Practically, registrations were performed with the default "DISOPTICAL FLOWPRESET MEDIUM" parameters. Artifacts on the overlapping regions are very intense, sometimes as intense as the deformation relative to micro-plasticity (inclined bands), as shown in Fig. 13. These artifacts prevent the visualization and identification of microplasticity (sub-pixel displacement jump due to slip activity) after deformation [13]. Interestingly, artifacts on the overlapping regions are more and more prominent when going from the top-left corner to the bottom-right corner of the strain fields. This artifact is mainly due to the mechanical alignment of the mosaic of images for the image acquisition before and after deformation, this image alignment being done on the images in the top-left corner. More precisely, in the present study, the first set of images corresponds to the images before mechanical deformation and the second set of images corresponds to the images after plastic deformation. The mechanical test has been done *ex-situ* using a conventional electromechanical tensile testing machine. Therefore, the sample was physically removed from the microscope between both the mosaic of images. The image alignment procedure was done on the top-left corner image. The distance of registered images far from this top-left corner images increased due to both rigid alignment errors and cumulated displacement due to plastic deformation. This leads to increased deformation artifacts on strain maps on the non-distorted images. For the strain fields of corrected stitched images, artifacts on the overlapping regions were nearly suppressed, thus enabling a clear visualization of the microplasticity on the different strain fields. Some artifacts are more pronounced for some overlapped regions potentially due to (i) the scanning technique, (ii) slight tilt/rotation on the motorized stage due to XY-stage stroke, (iii) Z-data integration to provide the reconstructed image. Therefore, this analysis confirms that the present numerical development improves the gray-level contrast, sharpness, and local distortion of the blended images at the sub-pixel level.

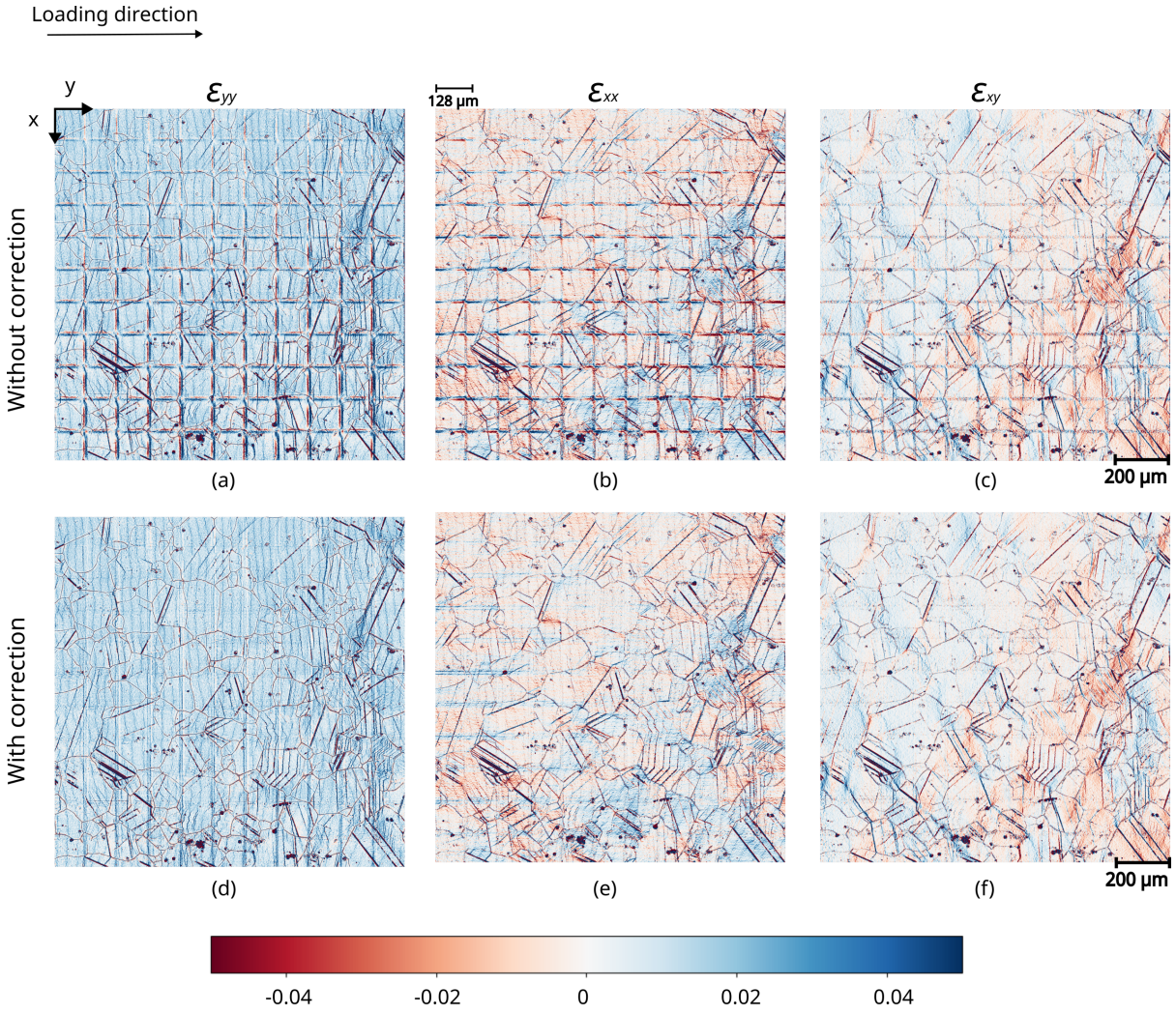

Figure 13: Result on a real high-resolution mosaic (121 sub-images of 1 Mpix). Comparison of the strain fields of the tensile test with rigid stitching (top) and the proposed non-rigid stitching (bottom). ε_{xx} (b, e), ε_{yy} (a, d) and ε_{xy} (d, f) stand for the components of the symmetric strain tensor defined by $\boldsymbol{\epsilon}(\mathbf{u}) = (\nabla\mathbf{u} + \nabla\mathbf{u}^T)/2$.

7. Conclusions and perspectives

Classic stitching procedures applied to images corrupted with geometric distortions generally lead to inhomogeneous blurring and affect image registration results. This study raises the challenges of accurately stitching distorted images solely from raw images, including those with a low degree of overlap. To achieve sub-pixel resolution measurements, a new approach using a self-calibrated registration was developed to dissociate translations from the intrinsic distortions of the imaging system. With this approach, each local image of the mosaic is corrected, and the final stitching is performed with standard blending approaches. Overall, this work represents an enhancement of 2D image stitching for the specific use of mosaic of images acquired using automated microscopy. It aims to improve the accuracy and quality of high-resolution image registration at the sub-pixel level.

Although the method was capable of improving the stitching quality by reducing the distortion

blur, there are still a number of areas for improvement. One common approach to image distortion compensation in computer vision is to use a calibration pattern. In stereo-vision, checkerboard patterns are commonly used to determine the distortions in the image and apply the appropriate correction. One of the perspectives of image compensation in microscopy is to use or manufacture dedicated calibration patterns in order to determine the most complete and accurate distortion fields, *i.e.* the non-affine modes but also the affine ones that cannot be identified with the present numerical development.

In addition, some geometrical tolerances have to be considered here for the optical system. In microscopy, the tilt, which is the deviation of the sample from the ideal perpendicular position relative to the optical axis of the microscope, can also lead to distortions, blur, and depth-dependent effects. These effects were implicitly compensated in the 2D polynomial model but were not dissociated from the true lens distortion. Therefore, it would be interesting to take this 3D effect into account for further developments.

From a computational point of view, the most time-consuming computation operation of the present algorithm is the successive correction/update of the corrected overlapping gray-levels. Even though the algorithm is parallelized over the overlapping region, the sub-pixel evaluation of the images (by bilinear or cubic spline interpolation) is still costly.

Appendix A. Image stitching

Image fusion or blending is a common and well-known problem in computer vision. In this appendix, we simply recall how one can perform image blending when considering the distortion correction transformation $\mathbf{P}$ (see *e.g.* Sections 3-4). For an identified distortion correction, the fused image using average blending is defined as follows:

$$\forall \mathbf{x} \in \Omega, \quad f(\mathbf{x}) = \frac{\sum_{i=1}^{m} \omega_i(\mathbf{x}) f_i \circ \mathbf{P}(\mathbf{x} - \mathbf{t}_i)}{\sum_{i=1}^{m} \omega_i(\mathbf{x})} \tag{A.1}$$

where ω_i is the blending function of each image. It is defined as follows:

$$\omega_i(\mathbf{x}) = \begin{cases} 1 & \text{if} \quad \mathbf{P}(\mathbf{x} - \mathbf{t}_i) \in [0, s_x - 1] \times [0, s_y - 1] \\ 0 & \text{else} \end{cases} . \tag{A.2}$$

s_x and s_y stand for the size of one sub-image of the grid of images. A second approach would be to consider the linear blending. In this case, one can consider a weighting defined as follows:

$$\forall \mathbf{x} \in \Omega, \quad f(\mathbf{x}) = \frac{\sum_{i=1}^{n} \sum_{j=0}^{1} \omega_i^j(\mathbf{x}) f_i^j \circ \mathbf{P}\left(\mathbf{x} - \mathbf{t}_{f_i^j}\right)}{\sum_{i=1}^{n} \sum_{j=0}^{1} \omega_i^j(\mathbf{x})} \tag{A.3}$$

where ω_i^j is the blending function of each overlapping region. It is defined as follows:

- If Ω_i is vertical, then the left and right image weightings are defined respectively as follows:

$$\omega_i^0(\mathbf{x}) = \begin{cases} 1 \quad \text{if} \quad \mathbf{P}_y\left(\mathbf{x} - \mathbf{t}_{f_i^0}\right) \in [0, s_x - 1] \times [0, s_y - 1 - h_y^i] \\ \dfrac{\mathbf{P}_y\left(\mathbf{x} - \mathbf{t}_{f_i^0}\right) - (s_y - 1)}{-h_y^i} \qquad \text{else} \end{cases} . \tag{A.4}$$

$$\omega_i^1(\mathbf{x}) = \begin{cases} \dfrac{\mathbf{P}_y\left(\mathbf{x} - \mathbf{t}_{f_i^1}\right)}{h_y^i} \quad \text{if} \quad \mathbf{P}_y\left(\mathbf{x} - \mathbf{t}_{f_i^1}\right) \in [0, s_x - 1] \times [0, h_y^i] \\ 1 \quad \text{else} \end{cases} . \tag{A.5}$$

- If Ω_i is horizontal, then the top and bottom image weightings are defined as it follows:

$$\omega_i^0(\mathbf{x}) = \begin{cases} 1 \quad \text{if} \quad \mathbf{P}_x\left(\mathbf{x} - \mathbf{t}_{f_i^0}\right) \in [0, s_x - 1 - h_x^i] \times [0, s_y - 1] \\ \dfrac{\mathbf{P}_y\left(\mathbf{x} - \mathbf{t}_{f_i^0}\right) - (s_y - 1)}{-h_y^i} \qquad \text{else} \end{cases} . \tag{A.6}$$

$$\omega_i^1(\mathbf{x}) = \begin{cases} \dfrac{\mathbf{P}_x\left(\mathbf{x} - \mathbf{t}_{f_i^1}\right)}{h_y^i} \quad \text{if} \quad \mathbf{P}_x\left(\mathbf{x} - \mathbf{t}_{f_i^1}\right) \in [0, h_x^i] \times [0, s_y - 1] \\ 1 \quad \text{else} \end{cases} . \tag{A.7}$$

Funding acknowledgments

This work was supported by the European Research Council [project HT-S_4DefOx - Grant number 948007]. A [CC-BY public copyright license] has been applied by the authors to the present document and will be applied to all subsequent versions up to the Author Accepted Manuscript arising from this submission, in accordance with the grant's open access conditions. The authors are grateful to the *Centre National de la Recherche Scientifique (CNRS)* for the mobility grant with the International Research Project denoted "CIN&MAT"

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. Farsiu, M. D. Robinson, M. Elad, P. Milanfar, Fast and robust multiframe super resolution, IEEE Transactions on Image Processing 13 (2004) 1327–1344.

[2] S. S. Olivier, L. Seppala, K. Gilmore, Optical design of the LSST camera, in: E. Atad-Ettedgui, D. Lemke (Eds.), Advanced Optical and Mechanical Technologies in Telescopes and Instrumentation, volume 7018, International Society for Optics and Photonics, SPIE, 2008, p. 70182G.

[3] O. S. Cossairt, D. Miau, S. K. Nayar, Gigapixel computational imaging, in: 2011 IEEE International Conference on Computational Photography (ICCP), 2011, pp. 1–8.

[4] D. J. Brady, M. E. Gehm, R. A. Stack, D. L. Marks, D. S. Kittle, D. R. Golish, E. M. Vera, S. D. Feller, Multiscale gigapixel photography, Nature 486 (2012) 386–389.

[5] E. E. Thomson, M. Harfouche, K. Kim, P. C. Konda, C. Seitz, C. Cooke, S. Xu, W. S. Jacobs, R. Blazing, Y. Chen, S. Sharma, T. W. Dunn, J. Park, R. Horstmeyer, E. A. Naumann, Gigapixel imaging with a novel multi-camera array microscope, eLife 11 (2022) 1–47.

[6] T. Cohen Hyams, K. Mam, M. C. Killingsworth, Scanning electron microscopy as a new tool for diagnostic pathology and cell biology, Micron 130 (2020) 102797.

[7] ZEISS Atlas 5 Master Your Multi-scale Challenge, 2023. URL: `https://www.zeiss.com/microscopy/en/products/software/zeiss-atlas-5.html`.

[8] A. L. Keller, D. Zeidler, T. Kemen, High throughput data acquisition with a multi-beam SEM, in: M. T. Postek, D. E. Newbury, S. F. Platek, T. K. Maugel (Eds.), SPIE Proceedings, SPIE, 2014.

[9] R. Black, T. Garbowski, C. Bean, A. Eberle, S. Nickell, D. Texier, V. Valle, J. Stinville, High-throughput high-resolution digital image correlation measurement by Multi-Beam SEM imaging, Experimental mechanics (2023).

[10] M. Sutton, N. Li, D. Joy, A. Reynolds, X. Li, Scanning electron microscopy for quantitative small and large deformation measurements Part I: SEM imaging at magnifications from 200 to 10,000, Experimental Mechanics 47 (2007) 775–787.

[11] H. Teyssedre, S. Roux, G. Régnier, A. Tracz, Filtering out slow-scan drifts in atomic force microscopy images, The Journal of Strain Analysis for Engineering Design 46 (2011) 361–367.

[12] A. D. Kammers, S. Daly, Digital image correlation under scanning electron microscopy: methodology and validation, Experimental Mechanics 53 (2013) 1743–1761.

[13] J. Stinville, M. Echlin, D. Texier, F. Bridier, P. Bocher, T. Pollock, Sub-grain scale digital image correlation by electron microscopy for polycrystalline materials during elastic and plastic deformation, Experimental mechanics 56 (2016) 197–216.

[14] S. Maraghechi, J. P. Hoefnagels, R. H. Peerlings, M. G. Geers, Correction of scan line shift artifacts in scanning electron microscopy: An extended digital image correlation framework, Ultramicroscopy 187 (2018) 144–163.

[15] T. Vermeij, J. Verstijnen, T. Ramirez y Cantador, B. Blaysat, J. Neggers, J. Hoefnagels, A nanomechanical testing framework yielding front&rear-sided, high-resolution, microstructure-correlated SEM-DIC strain fields, Experimental Mechanics 62 (2022) 1625–1646.

[16] W. Qian, J. Li, J. Zhu, W. Hao, L. Chen, Distortion correction of a microscopy lens system for deformation measurements based on speckle pattern and grating, Optics and Lasers in Engineering 124 (2020) 105804.

[17] S. Maraghechi, J. Hoefnagels, R. Peerlings, O. Rokos, M. Geers, Correction of scanning electron microscope imaging artifacts in a novel digital image correlation framework, Experimental mechanics 59 (2019) 489–516.

[18] J. Hastings, F. Zhang, H. Smith, Nanometer level stitching in raster scanning electron beam lithography using spatial phase locking, J Vac Sci Technol, B, Microelectron Nanometer Struct 21 (2003) 2650–2656.

[19] D. Gledhill, G. Y. Tian, D. Taylor, D. Clarke, Panoramic imaging - A review, Computers and Graphics (Pergamon) 27 (2003) 435–445.

[20] R. Szeliski, Image alignment and stitching: A tutorial, Foundations and Trends® in Computer Graphics and Vision 2 (2007) 1–104.

[21] E. Adel, M. Elmogy, H. Elbakry, Image stitching based on feature extraction techniques: A Survey, International Journal of Computer Applications 99 (2014) 1–8.

[22] J. Buckman, Use of automated image acquisition and stitching in scanning electron microscopy: Imaging of large scale areas of materials at high resolution, Microsc. Anal 28 (2014) 13–15.

[23] A. Rouwane, D. Texier, S. Hémery, J.-C. Passieux, Q. Sirvin, J. Genée, A. Proietti, J.-C. Stinville, Strain localization in Ti and Ti-allys using three-dimensional topographic imaging, in: Titanium 2023, 2023.

[24] F. Nielsen, Randomized adaptive algorithms for mosaicing systems, IEICE Transactions on Information and Systems E83-D (2000) 1386–1394.

[25] M. Z. Bonny, M. S. Uddin, Feature-based image stitching algorithms, in: 2016 International workshop on computational intelligence (IWCI), IEEE, 2016, pp. 198–203.

[26] G. Lowe, Sift-the scale invariant feature transform, Int. J 2 (2004) 2.

[27] K. G. Derpanis, The harris corner detector, York University 2 (2004) 1–2.

[28] D. Mistry, A. Banerjee, Comparison of feature detection and matching approaches: Sift and surf, GRD Journals-Global Research and Development Journal for Engineering 2 (2017) 7–13.

[29] X. Jiang, J. Ma, G. Xiao, Z. Shao, X. Guo, A review of multimodal image matching: Methods and applications, Information Fusion 73 (2021) 22–71.

[30] L. Chen, F. Rottensteiner, C. Heipke, Feature detection and description for image matching: from hand-crafted design to deep learning, Geo-spatial Information Science 24 (2021) 58–74.

[31] M. Brown, D. G. Lowe, Automatic panoramic image stitching using invariant features, International journal of computer vision 74 (2007) 59–73.

[32] M. A. Fischler, R. C. Bolles, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (1981) 381–395.

[33] D. Garcia, Mesure de formes et de champs de déplacements tridimensionnels par stéréo-corrélation d'images, Ph.D. thesis, 2001. URL: https://theses.hal.science/tel-00166476.

[34] J.-E. Pierré, J.-C. Passieux, J.-N. Périé, F. Bugarin, L. Robert, Unstructured finite element-based digital image correlation with enhanced management of quadrature and lens distortions, Optics and Lasers in Engineering 77 (2016) 44–53.

[35] A. Can, C. Stewart, B. Roysam, H. Tanenbaum, A feature-based technique for joint, linear estimation of high-order image-to-mosaic transformations: application to mosaicing the curved human retina, in: Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662), volume 2, 2000, pp. 585–591 vol.2.

[36] R. H. Castanheira de Souza, M. Okutomi, A. Torii, Real-time image mosaicing using non-rigid registration, in: Advances in Image and Video Technology, Springer Berlin Heidelberg, 2012, pp. 311–322.

[37] A. Miettinen, I. V. Oikonomidis, A. Bonnin, M. Stampanoni, NRStitcher: non-rigid stitching of terapixel-scale volumetric images, Bioinformatics 35 (2019) 5290–5297.

[38] J. Weng, P. Cohen, M. Herniou, et al., Camera calibration with distortion models and accuracy evaluation, IEEE Transactions on pattern analysis and machine intelligence 14 (1992) 965–980.

[39] Daniel Wagner and Dieter Schmalstieg, Artoolkitplus for pose tracking on mobile devices, ???? URL: https://github.com/artoolkitx/artoolkitx.

[40] J. Lavest, M. Viala, M. Dhome, Do we need an accurate calibration target?, Traitement du Signal 16 (1999) 241–254.

[41] J.-E. Dufour, F. Hild, S. Roux, Integrated digital image correlation for the evaluation and correction of optical distortions, Optics and Lasers in Engineering 56 (2014) 121–133.

[42] S. Yoneyama, H. Kikuta, A. Kitagawa, K. Kitamura, Lens distortion correction for digital image correlation by measuring rigid body displacement, Optical engineering 45 (2006) 023602–023602.

[43] H. Foroosh, J. B. Zerubia, M. Berthod, Extension of phase correlation to subpixel registration, IEEE transactions on image processing 11 (2002) 188–200.

[44] V. Rankov, R. J. Locke, R. J. Edens, P. R. Barber, B. Vojnovic, An algorithm for image stitching and blending, in: Three-dimensional and multidimensional microscopy: image acquisition and processing XII, volume 5701, SPIE, 2005, pp. 190–199.

[45] S. Preibisch, S. Saalfeld, P. Tomancak, Globally optimal stitching of tiled 3d microscopic image acquisitions, Bioinformatics 25 (2009) 1463–1465.

[46] V. Kaynig, B. Fischer, E. Müller, J. M. Buhmann, Fully automatic stitching and distortion correction of transmission electron microscope images, Journal of structural biology 171 (2010) 163–173.

[47] A. Rouwane, D. Texier, J.-C. Passieux, Distortion correction for image stitching - application to lscm microscopy, 20. URL: 10.5281/zenodo.11035480. doi:doi.org/10.5281/zenodo.11035480.

[48] D. P. Bertsekas, Nonlinear programming, Journal of the Operational Research Society 48 (1997) 334–334.

[49] J.-C. Passieux, R. Bouclier, Classic and inverse compositional Gauss-Newton in global DIC, International Journal for Numerical Methods in Engineering 119 (2019) 453–468.

[50] P. J. Burt, E. H. Adelson, The laplacian pyramid as a compact image code, in: Readings in computer vision, Elsevier, 1987, pp. 671–679.

[51] N. D. Ellingwood, Y. Yin, M. Smith, C.-L. Lin, Efficient methods for implementation of multi-level nonrigid mass-preserving image registration on gpus and multi-threaded cpus, Computer Methods and Programs in Biomedicine 127 (2016) 290–300.

[52] J. Liu, N. Vanderesse, J.-C. Stinville, T. Pollock, P. Bocher, D. Texier, In-plane and out-of-plane deformation at the sub-grain scale in polycrystalline materials assessed by confocal microscopy, Acta Materialia 169 (2019) 260–274.

[53] D. Texier, D. Monceau, J.-C. Salabura, R. Mainguy, E. Andrieu, Micromechanical testing of ultrathin layered material specimens at elevated temperature, Materials at high temperatures 33 (2016) 325–337.

[54] G. Bradski, The opencv library, Dr. Dobb's Journal of Software Tools (2000).