Preprocessing methods for quantitative phase image stitching

PIOTR STEPIEŃ, Wojciech Krauze, and Małgorzata Kujawińska

Warsaw University of Technology, Institute of Micromechanics and Photonics, ul. Sw. A. Boboli 8, Warsaw, 02-525, Poland

piotr.stepien.dokt@pw.edu.pl

Abstract: Quantitative phase imaging of cell cultures and histopathological slides often requires measurements in large fields of view which is realized through the stitching of multiple high resolution phase maps. Due to the characteristic properties of phase images, careful preprocessing is crucial for maintaining the metrological value of the stitched phase image. In this work, we present various methods that address those properties. Our efforts are focused on increasing robustness to minimize error propagation in consecutive preprocessing steps.

1. Introduction

Quantitative Phase Imaging (QPI) constitutes a wide range of imaging methods that are increasingly improved and incorporated into biomedical research [1]. Most significant advantage of QPI is that it allows for the investigation of highly transparent objects without staining or fluorescent labeling, preserving neutral environment throughout measurement. This also implies lower costs of investigation, both in terms of preparation time and money. Furthermore, QPI provides results that are more than visual images – each image is a quantitatively accurate direct representation of phase delay introduced by the sample, so called optical phase difference (OPD). The most widespread and reliable QPI technique, which provided datasets for this work, is digital holographic microscopy (DHM) which allows for high spatial resolution phase image retrieval [2] at the cost of decreased field of view (FoV). Some QPI techniques like lensless holographic microscopy [3], Fourier ptychography [4] provide results with increased space-bandwidth product (SBP) – they achieve high resolution while maintaining unconventionally wide fields of view (FoV). Deep neural networks have also been recently proposed to increase the SBP of a conventional QPI technique (DHM) beyond physical limits of the optical system [5]. However, to fully utilize the benefits of a variety of QPI techniques it might be beneficial to combine them with image stitching – a broad and general technique that aims at producing large field of view image by fusing multiple small FoV images [6–9]. Such images are required to have overlapping regions – small regions at the margins of FoV to achieve data redundancy. This allows for the adjustment of mutual location of neighboring images in the process of image registration. Additionally, in case of phase images the overlapping regions allow for necessary corrections to be applied. The acquisition with the overlaps is achieved by motorized stage incorporated into the microscope setup, with movement steps slightly smaller than the single FoV size. Unfortunately, if the quantitative nature of phase images is to be maintained in full FoV, the stitching should not be applied directly – it requires preprocessing to address properties typical to phase images, namely the presence of systematic and variable aberrations as well as lack of common baseline OPD. Otherwise quantitative analysis of the stitched image is invalid. Especially the lack of common baseline OPD is important to correct, since the problem is present in all QPI techniques, regardless of optical system or sample stability. This problem was briefly mentioned in the work of Lee et al. [8] and later discussed in our previous work [9]. Lee et al. did not discuss the details of methodology used to perform the corrections. Other works,
earlier to ours and following ours, did not consider this issue [10–13] although Min et al. [10] and Lambert [13] refer to custom software used to process the images without disclosing the applied procedures.

This paper discusses three methods that may be used to correct the mentioned features while maintaining the quantitative aspect of stitched images. One of the methods is a modification of the previously proposed averaged multiple exposure (AME) method used to correct systematic aberrations, while the other two, together with proposed variants, are, to the best of our knowledge, novel solutions. We focused our efforts on increasing the robustness and reliability of proposed methods in comparison to the previously developed methods [9].

2. Features of phase images and the correction pipeline

The problems related to phase image stitching have been discussed in our previous publication [9], but a brief summary is given here to outline the features of the images, together with the explanation of the processing pipeline (Fig. 1) to prepare images for stitching. Particular methods used for each step are described in Sec. 3.

Fig. 1. Preprocessing procedure pipeline: a) General pipeline used throughout this work, with distinguished alternative paths. b) Visualization for each performed correction step. Systematic error and trends are exaggerated by a factor of 4 for visualization purposes.

The effects of aberrations generally result in image blurring and are present in both, QPI and non-quantitative imaging modalities. The aberration influence on the quality of imaging may be ignored when the aberrations are sufficiently small. However in QPI such small aberrations may nonetheless result in other effect which is the additive term which is overlaid on the measured phasemap. Because the term is additive it may be simply subtracted, once it is properly estimated, such as with the proposed methods. In the scope of this work we use the term aberrations only in such context.

Systematic aberrations, compensated with averaging-derived error map

Aberrations of the optical system are present in the reconstructed phase image. They appear as a curvature or debris from imperfect calibration measurement [14] and are invariable for all constituent images. Systematic aberrations are estimated once per full stitching measurement or less often in case of timelapse measurements if the system and sample are stable in the long
Aberration estimation proposed in this work is filtered AME (FAME, see Sec. 3.2.1), which produces a filtered averaging-derived image that represents a systematic error map.

Residual linear aberrations, compensated with estimated trends

Besides the systematic error, the residual aberration is present, due to short term system or sample instabilities. Such aberration is very well approximated by a linear trend, generally different for each image. The trend within each sub-image needs to be estimated and subtracted, otherwise it is very disruptive due to baseline unification explained further. Popular least-squares plane fitting (LSQ-PF) technique often fails to provide good results due to strong object dependence. Two presented methods address this issue, either by being more object-independent in case of gradient-based plane fitting (grad-PF, see Sec. 3.2.2) or by reference to data in overlapping regions in case of optimization-based simultaneous baseline unification and trend correction (opt-BT, see Sec. 3.3.2).

Lack of common baseline OPD, compensated with offset OPD values

Due to the nature of phase imaging, the QPI measurements are not represented in an absolute scale. This means that for a given phase image, the OPD is represented without any specified baseline, so the phase differences may initially be compared only within a single FoV. For example, object-free regions are a natural baseline (0 rad OPD), but even there, directly after phase retrieval, the OPD values are generally different for each image. Before stitching a common baseline needs to be established for all images, especially the ones that do not include an object-free region. Such a baseline is obtained by analyzing differences in the overlapping regions and subtracting estimated offset values from individual images. A proposed optimization-based baseline unification (opt-B) method to perform this correction is given in Sec. 3.3.1, while the mentioned opt-BT is discussed in Sec. 3.3.2.

The importance of the corrections outlined above is best explained by analyzing Eq. (1) [15] which relates the dry mass $M$ of the analyzed sample based to the obtained OPD:

$$ M = \frac{1}{\alpha} \int_S \bar{OPD}_o(x, y) ds = \frac{S_o}{\alpha} \langle \bar{OPD}_o \rangle $$

with:

$$ \bar{OPD}_o(x, y) = OPD_o(x, y) + E(x, y) $$

where: $\langle \square \rangle$ indicates the average value, $OPD_o(x, y)$ is the spatial distribution of OPD of the object such as a cell, $S_o$ is the area occupied by the object and $\alpha$ is the refractive index increment, which is a scaling factor that links the integrated $OPD_o$ introduced by the object with its dry mass. $E$ is the error term composed of both corrected errors such as systematic errors, phase trends and baseline error, as well as errors that are not accounted for. $OPD_o$ should be given in reference to the object-free area, thus the baseline unification plays a crucial role. However, both systematic errors and phase trends strongly influence the baseline unification procedure, due to its dependence on similarity of overlapping regions. Therefore, $E$ both directly and indirectly distorts the calculation of $M$, hence the importance of proposed corrections.

3. Methods

In this section three methods and their variants are introduced and discussed in detail. They allow to mitigate, the problems described in the previous section.

Since the methods presented here are preprocessing methods, the results are demonstrated on images tiled next to each other, without registration of the overlapping regions to correct positions of the images. This could be considered a limitation of the study since the optimization-based methods that rely on overlapping regions comparison may benefit on performing registration beforehand, however it should be acknowledged that the methods perform well even without it, hence simplifying the overall stitching procedure.
The methods are demonstrated on phase images retrieved from off-axis image-plane holograms captured by digital holographic microscope [2]. The objects were HaCaT cell cultures measured in the wound healing scenario [16] – an object is a dense cell culture with artificially created gap, free of cells. The measurement consists of timelapse capture to investigate cell growth, filling the wound region over time. The scene is particularly difficult for phase image stitching, because the cells and background create a well defined and regular boundary – it corrupts AME performed at a single time point and skews the results when LSQ-PF is applied at numerous single FoV images.

At our disposal were six datasets captured with the interval of 10 minutes over total time of approximately 24 hours, resulting in up to 144 time points of $5 \times 5$ images in each time point. Size of each FoV was $378 \times 456$ pixels or $359 \times 299 \, \mu m$. The size of the nominal overlapping region was 7% margin from each side.

3.1. Simulated datasets and quantitative assessment

In order to quantify the influence of the proposed corrections, we developed a simulated scene, composed of spherical objects randomly located within the big FoV. The example of this generated dataset and its components is presented in Fig. 2. Diminishing occurrence of the spheres simulates the cell gap present in experimental data. This phase image is divided into a $6 \times 6$ set of smaller images with 10% overlap. Each of the images is corrupted with systematic aberration term (third order Legendre polynomial) as well as random components: linear trend, baseline error. Legendre coefficients, slopes and baseline errors were drawn from uniform probability distribution between -1 and 1. The spatial dimensions of the images were also simulated in the range between -1 and 1. Such a dataset is then corrected with different sets of methods, and the process is repeated 50 times (each time with new random components) in order to statistically quantify the performance. It should be noted that due to non-uniform gradient of simulated spheres and their purely random location (the spheres do not collide) the character of the simulation is different compared with the experimental data, however this is not relevant for the comparative analysis of both previously known and the proposed methods.

![Fig. 2. Example of the simulated dataset preparation, with the initial data on the left, followed by added components of systematic aberrations, linear aberrations and offsets, with the resulting dataset on the right.](image)

Having full knowledge about the simulated object, we investigated two error metrics. The first one, $A$, is the root mean squared error (RMSE) of the objects OPD, divided by its maximum value and averaged over 50 simulations. Such an error summarizes local error to be expected at any given point of the stitched FoV.

$$A = \frac{\text{RMSE} \left( \text{OPD}_0 - \overline{\text{OPD}}_0 \right)}{\max(\text{OPD}_0)}_{50}$$  (2)
with \( \mathcal{H} \{\square\} \) indicating the operator of a given set of correction methods. The second metric, \( B \), is the absolute value of the relative cumulative OPD error, averaged over 50 simulations.

\[
B = \frac{\sum \text{OPD}_o - \sum \mathcal{H} \{\text{OPD}_o\}}{\sum \text{OPD}_o} \tag{3}
\]

Such an error estimates overall error to be expected while calculating the dry mass of all objects within the stitched FoV. We present these two metrics, expressed in percentages, at the end of each section describing the proposed methods.

Note that some combinations of the compared methods repeat between sections, for example the set of experiments corrected with FAME, grad-PF and opt-B is used separately for comparison with AME (labeled as FAME), LSQ-PF and opt-BT (in both cases labeled as grad-PF).

### 3.2. Single step methods

In this section we provide description and examples of the single step correction methods. In contrast to optimization-based methods from section 3.3, these methods require a single pass through the frames in order to estimate the correction term.

#### 3.2.1. Systematic aberration correction with filtered averaged multiple exposure

AME [9] is the systematic aberration estimation method where multiple phase images within the dataset are averaged. With a typical cell culture with cells randomly distributed over the FoV AME typically creates an accurate representation of the error map within a single time point (STP) of a timelapse measurement. However in some cases the estimated error term is corrupted due to strong phase contrast within the images – in our datasets captured in a wound healing scenario such regions are the edges of the wound. This problem is solved by extending AME on multiple time points (MTPs) by averaging measurements from different time points. However, such approach might cause different problem due the evaporation of the cell medium, resulting in slow movement of the parasitic fringes that are a virtually invariable within a single time point but vary significantly over a larger time range, causing unwanted features in the estimated error map.

The proposed solution is FAME, which consists of AME followed by image filtration of the obtained error term. Any filtration may be applied, although in the analyzed datasets the error had unwanted high-frequency features, thus the polynomial fitting approach has been used. This allows for the close approximation of the low frequency information captured by AME, while removing the parasitic fringe pattern that shifts in subsequent time points (see Fig. 3 a) vs b)). A possible drawback is that some high frequency information, such as debris from reference measurement [14], is often desired for removal. However, this may be addressed by carefully choosing the filtration method and properly selecting the time range where the averaging is applied.

Figure 3 shows the results of AME and FAME at MTPs and STP, where the FAME MTPs (Fig. 3(a)) provides best results – the background is smoother compared to Fig. 3(b), which is visible in the cross-sections. STP approaches (Fig. 3(c) and 3(d)) show very strong corruption by the high contrast wound edge. However, FAME and variable time range should be considered an additional tool to deal with systematic aberrations, not universally better than AME or STP estimation.

The limitations and alternatives for AME, and by extension FAME have been discussed in the previous work [9].

**AME versus FAME in simulation** For the simulated assessment of the methods we focused solely on the STP experiment. We applied the same set of correction methods, that is grad-PF
Fig. 3. Comparison of FoVs tiling showing the results of FAME and AME corrections estimated with MTPs and STP. In the first row the error term is shown. Below is the FoVs tiling with the error term removed. Besides the tested systematic error correction method, the grad-PF and opt-B methods were used. In third row the zoomed-in single FoV is shown. The cross section through the single FoV is shown in fourth row, with baseline indicated by the black dotted line.

and opt-B, so that the AME and FAME were the only differing steps in the comparison. The results are summarized in Table 1. The FAME performs better compared with AME due to high contrast and non-uniform distribution of spheres. These factors cause that the residual object features are present in AME-obtained aberration component, hence causing disruption in further corrections. FAME is more resilient to those factors due to the filtering step.

| Table 1. Comparison of the AME and FAME methods applied to simulated data. |
|---------------------------------|--------|--------|
| Error metrics      | AME    | FAME   |
| A                  | 7% ± 1%| 5% ±1% |
| B                  | 12% ± 4%| 8% ± 4%|

3.2.2. Residual linear aberration correction – trend correction with plane fitting based on image gradients

Least squares (LSQ) based surface fitting, especially plane fitting (PF) is a popular solution for removing trends in 2D images [17]. LSQ-PF provides a plane that fits the data with minimized least squares error. The approach delivers good results in specific cases where object centroid is close to geometrical center, but otherwise strongly corrupts the preprocessing. If this condition is not met, like in the case of the wound healing experiment where the edge of the wound constitutes a step-like function, LSQ-PF results in images where neither the object nor the background are level. The higher the phase-contrast, the problem is more pronounced – see LSQ-PF result in
Fig. 4(b). Additionally, the error introduced by LSQ-PF is propagating, causing whole regions to be drastically underestimated in terms of OPD, which makes the dry mass calculation unfounded (see Eq. (1)). This effect may be observed in Sec. 3.3.2.

![Fig. 4. First row shows comparison of FoV tiling with the results of grad-PF and LSQ-PF trend corrections. Below is a zoomed-in single FoV. Besides the compared plane fitting algorithm, the AME MTPs and opt-B were used.](image)

Visual inspection suggests that better result may be achieved if only the slope of the data is considered instead of the data itself. A natural tool to investigate the slope is the derivative, which provides a local slope of a given function. To find the slope that dominates the image, we analyze the image gradient, which is composed of partial derivatives of the image. Furthermore, the estimation should be performed with a non-linear function such as median in order to prevent the distortion due to outliers such as edges of the cells.

The proposed algorithm (grad-PF) is to first calculate the image gradient, find the median value for partial derivatives in $x$ and $y$ directions and then subtract planes with the slopes equal to consecutive median slopes. The equation that represents this procedure is:

$$\tilde{I}(x, y) = I(x, y) - \text{med}\left\{\frac{\partial I(x, y)}{\partial x}\right\} x - \text{med}\left\{\frac{\partial I(x, y)}{\partial y}\right\} y,$$

where $I$ and $\tilde{I}$ are respectively image before and after correction, $x$ and $y$ are the spatial coordinates of the images and med (□) denotes the median of the value inside the bracket.

Comparison of grad-PF and LSQ-PF methods in Fig. 4 shows how crucial is the trend correction step in the preprocessing pipeline, with whole left side of tiled image underestimated due to LSQ-PF performance in the wound edge region (Fig. 4(b)). The cross-sections through both single FoVs indicate that grad-PF provides results with more level background. Wound edge on the right is less affected, because it was coincidentally located near the boundaries of individual FoVs at this given time point. This shows that in similar cases grad-PF is more robust in terms of object independence. Contrary to LSQ-PF, grad-PF is independent of the image centroid location.

The limitation of the method is that it assumes that the image is dominated by the trend that is not intrinsic to the object. This is not always the case – the arbitrary object is not guaranteed to
be planar. Such problem is visible in Sec. 3.3.2 (Fig. 6(b)), where the cell profile is not a step but a curve.

**LSQ-PF versus grad-PF in simulation** Simulated comparison of LSQ-PF and grad-PF has been obtained with other corrections performed by FAME and opt-B. The results are presented in Table 2.

| Error metrics | LSQ-PF | grad-PF |
|---------------|--------|---------|
| A             | 9% ± 1% | 5% ± 1% |
| B             | 22% ± 4% | 8% ± 4% |

The benefits of grad-PF are clear and directly correlate with our experience, that LSQ-PF should not be used for linear aberrations correction without previous segmentation of the background.

### 3.3. Corrections with numerical optimization

In this section we present a versatile optimization-based approach to corrections. Depending on the formulation, this approach may be used either to unify baselines between individual FoVs or perform simultaneous baseline unification and trend correction (see Fig. 1(a)). The core of the method is discussed in this section, whereas detailed formulations are presented in Sec. 3.3.1 (baseline unification) and Sec. 3.3.2 (simultaneous baseline unification and trend correction).

The approach is inspired by the iterative closest points (ICP) algorithm [18], which is implemented in BigStitcher for volumetric registration [7]. The difference is that the ICP algorithm is designed to correct discrepancies between two point clouds by shifting one point cloud relative to the other, while in our case only the pixel values of all images are modified such that the differences in overlapping regions are minimized, while the locations of images are fixed.

The general approach presented here is based on the observation that overlapping parts of the images should be minimally different overall, that is with all images corrected simultaneously instead of sequentially. We solve this problem using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm [19,20] which is an iterative local optimization algorithm.

The BFGS algorithm is given a set of parameters $P$ to modify and to find their optimal values. The parameters modify individual images in each iteration and the result is assessed in terms of some chosen loss function that is minimized throughout iterations. Approach tested here utilized the Euclidean distance ($L^2$) as the loss function. It has been applied to the difference between the overlapping regions. We considered only the side-by-side overlaps, which constitute the majority of the overlapping pixels – the diagonal overlaps have been ignored.

The images are arranged in a regular grid of $k$ rows and $l$ columns. Let $I_{ij}$ be the image from $i$-th row and $j$-th column and $\tilde{I}$ be the image $I$ modified according to an array of parameters $P$. Additionally let $B$, $T$, $L$ and $R$ be side margins of $\tilde{I}$, respectively from bottom, top, left or right side of the image, cropped according to nominal overlap percentage. With such convention the procedure may be expressed in general as:

$$\minimize_{P} f(P) = \sum_{i=1}^{k-1} \sum_{j=1}^{l-1} \|B_{ij} - T_{i+1j}\|_2 + \|R_{ij} - L_{i,j+1}\|_2$$  \hspace{1cm} (5)$$

The loss function operates on the difference of nominally overlapping parts of neighboring images – the mutual misalignment of images position has been ignored. Due to already broad scope of this work we consider this possible improvement beyond the scope of this study.

Once the method is implemented, it is easy to add a regularization term, which we attempt in Sec. 3.3.2.
3.3.1. Optimization approach to baseline unification

In case of pure baseline unification (opt-B) the modified image is given as:

\[ \hat{I}_{ij} = I_{ij} - p_{ij} \]  

(6)

Here, every parameter \( p_{ij} \) in the set \( P \) is simply the offset value that should be subtracted from corresponding image to give all images the same baseline OPD. Therefore the number of tuned parameters is the number of images \( n = k \cdot l \).

The procedure is very robust and reliable – during our study there was no case where the procedure failed. In Fig. 5 you may find the results of this approach compared to the sequential baseline unification, in a case where no major trend correction error is noticeable. As expected, if the trend correction procedure fails systematically the error propagates (see Fig. 4), however the faulty correction of a single image does not tend to accumulate. This may be observed in Sec. 3.3.2 (Fig. 6(b)). In our dataset the algorithm provides results within 30 seconds per time point (5 × 5 images, 378 × 456 pixels each, 7% nominal overlap), computed on the AMD Ryzen 7 1700 processor.

**Fig. 5.** Comparison of images tiling showing the results of optimization-based baseline unification and sequential baseline unification. Besides compared baseline unification methods the AME MTPs and grad-PF were used.

**Sequential baseline correction versus opt-B in simulation** Comparison of the performance of sequential baseline correction and opt-B was achieved with other corrections performed by FAME and grad-PF.

The results are presented in Table 3. In the regular case the differences are in favour of opt-B method, but by a small margin. This is most likely due to predictable behaviour of the simulated scene. Because the systematic aberration correction and linear aberration correction perform consistently well, the opt-B does not display its robustness. For this reason we have prepared a distorted scenario where we simulate the PF algorithm failure – after correcting the linear aberrations of all images one of the images in the center of the grid is again distorted with linear phase term. The slope has been drawn from uniform distribution in the range between -10 to 10. In such case, the benefits of opt-B are clearly shown in A and B metrics. The score for opt-B does not deteriorate, because the random tilt has been accounted for in the comparison with the reference in order to quantify the robustness against error propagation.

**Table 3.** The results comparing the sequential baseline correction and opt-B methods.

| Case    | Error metrics | Sequential baseline correction | opt-B     |
|---------|---------------|--------------------------------|-----------|
| Regular | A             | 6% ± 2%                         | 5% ± 1%   |
|         | B             | 8% ± 4%                         | 8% ± 4%   |
| Distorted | A             | 18% ± 6%                        | 4% ± 2%   |
|         | B             | 15% ± 11%                       | 7% ± 5%   |
3.3.2. Optimization approach to simultaneous baseline unification and trend correction

If both baseline unification and trend correction (opt-BT) are to be performed at the same time the modified image is given by equation

\[ \hat{I}_{ij}(x,y) = I_{ij}(x,y) - p_{ij,1} - p_{ij,2}x - p_{ij,3}y \]  

(7)

As indicated by a third subscript, in this case the number of parameters in \( P \) equals 3, with \( p_{ij,1} \) being the offset values; \( p_{ij,2} \) and \( p_{ij,3} \) being the slopes of the correction planes.

The example result of this approach is presented in Fig. 6(a). It shows that the method is capable of providing results better than the combination of plane fitting algorithms for trend correction combined with opt-B unification (Fig. 6(b) and (c)). The background region is more uniform and the OPD of cells regions does not tend to decrease as shown in the cross section plots. However the method failed in some cases, presumably due to local non-linear residual aberrations. We were able to alleviate such problem by introducing additional regularization term inspired by the grad-PF approach:

\[
\begin{align*}
\text{minimize } f'(P) &= f(P) + \lambda \sum_{i=1}^{k} \sum_{j=1}^{l} \left[ \text{med}^2 \left\{ \frac{\partial \hat{I}_{ij}(x,y)}{\partial x} \right\} + \text{med}^2 \left\{ \frac{\partial \hat{I}_{ij}(x,y)}{\partial y} \right\} \right] \\
\text{with: } \\
\text{med} \left\{ \frac{\partial \hat{I}_{ij}(x,y)}{\partial x} \right\} &= \text{med} \left\{ \frac{\partial I_{ij}(x,y)}{\partial x} \right\} - p_{ij,2} \\
\text{med} \left\{ \frac{\partial \hat{I}_{ij}(x,y)}{\partial y} \right\} &= \text{med} \left\{ \frac{\partial I_{ij}(x,y)}{\partial y} \right\} - p_{ij,3}
\end{align*}
\]

(8)

The strength of the regularization is controlled by the importance parameter \( \lambda \). The regularization is the penalty for slopes present in all individual FoVs. Notice that the median slope is estimated only once, since additional slope is equal to parameters \( p_{ij,2} \) and \( p_{ij,3} \). This means that there is virtually no computational overhead for the usage of the regularization. The result of such regularization is shown in Fig. 7(b). The low frequency artifacts present in the basic case (Fig. 7(a)) are removed, yet the differences at the overlapping regions are lower in the regularized opt-BT compared to grad-PF with opt-B approach (see zoomed-in region in Fig. 7(b) vs 7(c)). This is because regularized opt-BT has more freedom to minimize the differences, and uses the regularization only to assist the optimization procedure. We used \( \lambda = 0.35 \) without the need to re-adjust.

The biggest benefit of the method is the usage of clear quantitative metric to correct the measurements without reliance on any a priori assumptions about the object. As such, the metric may be a subject of future research, as an estimate correlated to the quality of the full preprocessing procedure that impacts the OPD accuracy in the stitched image. Another benefit is that once the basic approach fails the method may be simply tuned by the regularized approach and the \( \lambda \) parameter.

The downside of the method is its computational complexity – in our dataset the correction of a single time point (5 x 5 FoVs, 378 x 456 pixels each, 7% nominal overlap) took approximately 5 minutes on the AMD Ryzen 7 1700 processor. Although the computational time is beyond the scope of this work, faster performance might be achieved either by numerical optimizations or by using downsampled copies of the overlapping regions.

**Comparison of grad-PF + opt-B with opt-BT and regularized opt-BT in simulation**

The assessment of simultaneous optimization-based corrections has been performed by comparison to the combination of grad-PF + opt-B. For systematic aberration correction FAME has been used in all cases. The results are presented in Table 4.
Fig. 6. Tiled FoVs comparing opt-BT, grad-PF and LSQ-PF. All sets were treated with FAME MTPs while the two latter sets were also treated with opt-B, for fair comparison with opt-BT, which includes baseline unification.

Fig. 7. FoV tiling showing the results of regularized opt-BT (in center) compared with basic opt-BT (left) and grad-PF (right). The low-frequency error of basic opt-BT is better visible in the top row with full FoV tiling, while the trend correction error introduced by grad-PF is better visible in the zoomed-in region in bottom row. All cases were treated with FAME MTPs, with the last case additionally treated with opt-B for fair comparison.
Table 4. The results comparing the separate linear aberration correction and baseline correction with grad-PF + opt-B with opt-BT and regularized opt-BT methods.

| Error metrics | grad-PF + opt-B | opt-BT | opt-BT reg. |
|---------------|----------------|--------|-------------|
| A             | 5% ± 1%        | 39% ± 17% | 5% ± 1%     |
| B             | 8% ± 4%        | 42% ± 29% | 4% ± 3%     |

The massive error obtained with the opt-BT without regularization shows the disruptive effect of imperfect systematic aberration correction. Because the opt-BT is targeted only at removing the differences in the overlapping regions it is expected that it may propagate the errors in those regions to other images. Simple addition of regularization term resolves the issue, and provides best results in terms of B metric and similar results compared to the grad-PF + opt-B.

4. Conclusions

We present three novel methods that proved to be useful while working with multiple timelapse stitched measurements, comprised of thousands of single FoVs. The methods allowed to automate the stitching process to a manageable degree, which is the only feasible approach in case of large datasets. Manual input is required first at the beginning of the process, while choosing the systematic error correction method and later to review the results.

The methods are demonstrated both on experimental measurements, as well as in simulations. The latter allows to quantify the performance of the methods, showing the benefits of novel solutions presented in this work. However it should be stressed that the methods are object-dependent and experimental results and simulations are structurally different. Hence, none of the methods presented here should be considered a universal solution – both AME and FAME estimate systematic errors and both have positive and negative features depending on a given case. Similarly the combination of grad-PF and opt-B versus opt-BT address the same issues – first approach is faster but requires that the assumptions for the PF algorithm are fulfilled, whereas the opt-BT requires minimal to none assumptions, but is computationally expensive. For this reason we recommend opt-BT in difficult cases, such as measurements where high cell confluence disrupts the grad-PF approach. Nonetheless, those assumptions should be considered in case the tools presented here are used to provide quantitative results.

The influence of the size of the overlapping region was not tested, as the datasets at our disposal did not vary in this regard. The limitation applies to optimization-based methods that consider the differences in the overlapping regions.

Although the stitching technique has been long present in the context of QPI, the necessary robust processing has not been well reported. We believe that the developments presented here provide general approach to reliably increase the SBP of an arbitrary QPI technique by synthesizing multiple FoVs.

Funding. Narodowe Centrum Badań i Rozwoju (PL-TW/V/5/2018).

Acknowledgements. Authors would like to thank Sebastian Lejman who acquired the data that was analyzed here. The research leading to the described results was carried out within the Polish-Taiwanese Joint Research Project (PL-TW/V/5/2018) financed by the National Centre for Research and Development, Poland.

Disclosures. The authors declare that there are no conflicts of interest related to this article.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References
1. Y. Park, C. Depeursinge, and G. Popescu, “Quantitative phase imaging in biomedicine,” Nat. Photonics 12(10), 578–589 (2018).
2. M. K. Kim, “Principles and techniques of digital holographic microscopy,” J. Photonics Energy 1(1), 018005 (2010).
3. W. Bishara, T.-W. Su, A. F. Coskun, and A. Ozcan, “Lensfree on-chip microscopy over a wide field-of-view using pixel super-resolution,” Opt. Express 18(11), 11181–11191 (2010).

4. G. Zheng, R. Horstmeyer, and C. Yang, “Wide-field, high-resolution Fourier ptychographic microscopy,” Nat. Photonics 7(9), 739–745 (2013).

5. A. Butola, S. R. Kanade, S. Bhatt, V. K. Dubey, A. Kumar, A. Ahmad, D. K. Prasad, P. Senthilkumaran, B. S. Ahluwalia, and D. S. Mehta, “High space-bandwidth in quantitative phase imaging using partially spatially coherent digital holographic microscopy and a deep neural network,” Opt. Express 28(24), 36229–36244 (2020).

6. S. Preibisch, S. Saalfeld, and P. Tomancak, “Globally optimal stitching of tiled 3d microscopic image acquisitions,” Bioinformatics 25(11), 1463–1465 (2009).

7. D. Hörl, F. R. Rusak, F. Preusser, P. Tillberg, N. Randle, R. K. Chhetri, A. Cardona, P. J. Keller, H. Harz, H. Leonhardt, M. Treier, and S. Preibisch, “Densestitcher: reconstructing high-resolution image datasets of cleared and expanded samples,” Nat. Methods 16(9), 870–874 (2019).

8. M. Lee, E. Lee, J. Jung, H. Yu, K. Kim, J. Yoon, S. Lee, Y. Jeong, and Y. Park, “Label-free optical quantification of structural alterations in Alzheimer’s disease,” Sci. Rep. 6(1), 31034 (2016).

9. P. Stepień, D. Korbuszewski, and M. Kujawińska, “Digital holographic microscopy with extended field of view using tool for generic image stitching,” ETRI J. 41(1), 73–83 (2019).

10. E. Min, M. E. Kandel, C. J. Ko, G. Popescu, W. Jung, and C. Best-Popescu, “Label-free, multi-scale imaging of ex-vivo mouse brain using spatial light interference microscopy,” Sci. Rep. 6(1), 39667–12 (2016).

11. E. Kwee, A. Peterson, J. Stinson, M. Halter, L. Yu, M. Majurski, J. Chalfoun, P. Bajcsy, and J. Elliott, “Large field of view quantitative phase imaging of induced pluripotent stem cells and optical pathlength reference materials,” Proc. SPIE 10503, 105032F (2018).

12. M. K. Kim, “Wide area quantitative phase microscopy by spatial phase scanning digital holography,” Opt. Lett. 45(3), 784–786 (2020).

13. A. Lambert, “Live cell imaging with holotomography and fluorescence,” Microsc. Today 28(1), 18–23 (2020).

14. P. Ferraro, S. De Nicola, A. Finizio, G. Coppola, S. Grilli, C. Magro, and G. Pierattini, “Compensation of the inherent wave front curvature in digital holographic coherent microscopy for quantitative phase-contrast imaging,” Appl. Opt. 42(11), 1938–1946 (2003).

15. P. Girshovitz and N. T. Shaked, “Generalized cell morphological parameters based on interferometric phase microscopy and their application to cell life cycle characterization,” Biomed. Opt. Express 3(8), 1757–1773 (2012).

16. L. G. Rodriguez, X. Wu, and J.-L. Guan, “Wound-healing assay,” in Cell Migration (Springer, 2005), pp. 23–29.

17. J. Di, J. Zhao, W. Sun, H. Jiang, and X. Yan, “Phase aberration compensation of digital holographic microscopy based on least squares surface fitting,” Opt. Commun. 282(19), 3873–3877 (2009).

18. M. A. Audette, F. P. Ferrie, and T. M. Peters, “An algorithmic overview of surface registration techniques for medical imaging,” Med. Image Anal. 4(3), 201–217 (2000).

19. J. Nocedal and S. Wright, Numerical Optimization (Springer Science & Business Media, 2006).

20. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burosvk, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, I. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, and P. van Mulbregt, SciPy 1.0 Contributors, “SciPy 1.0: Fundamental algorithms for scientific computing in Python,” Nat. Methods 17(3), 261–272 (2020).