RESEARCH ARTICLE

PAAQ: Paired Alternating AcQuisitions for Virtual High Frame Rate Multichannel Cardiac Fluorescence Microscopy

François Marelli[1,2] ⓘ *, Alexander Ernst[3] ⓘ, Nadia Mercader[3,4] ⓘ and Michael Liebling[1,5] ⓘ *

[1]Computational Bioimaging, Idiap Research Institute, Martigny, CH-1920, Switzerland
[2]Electrical Engineering, École Polytechnique Fédérale de Lausanne, Lausanne, CH-1015, Switzerland
[3]Institute of Anatomy, University of Bern, Bern, CH-3000, Switzerland
[4]Cardiovascular Regeneration Program, Centro Nacional de Investigaciones Cardiovasculares, Madrid, ES-28029, Spain
[5]Electrical & Computer Engineering, University of California, Santa Barbara, CA 93106, USA
*Corresponding author. Email: francois.marelli@idiap.ch; michael.liebling@idiap.ch

Abstract

In vivo fluorescence microscopy is a powerful tool to image the beating heart in its early development stages. A high acquisition frame rate is necessary to study its fast contractions, but the limited fluorescence intensity requires sensitive cameras that are often too slow. Moreover, the problem is even more complex when imaging distinct tissues in the same sample using different fluorophores. We present Paired Alternating AcQuisitions (PAAQ), a method to image cyclic processes in multiple channels, which requires only a single (possibly slow) camera.
We generate variable temporal illumination patterns in each frame, alternating between channel-specific illuminations (fluorescence) in odd frames and a motion-encoding brightfield pattern as a common reference in even frames. Starting from the image pairs, we find the position of each reference frame in the cardiac cycle through a combination of image-based sorting and regularized curve fitting. Thanks to these estimated reference positions, we assemble multichannel videos whose frame rate is virtually increased.
We characterize our method on synthetic and experimental images collected in zebrafish embryos, showing quantitative and visual improvements in the reconstructed videos over existing non-gated sorting-based alternatives. Using a 15 Hz camera, we showcase a reconstructed video containing 2 fluorescence channels at 100 fps.

Impact Statement

Our method enables imaging fast repeating processes like the heartbeat and reconstructing multichannel videos with a high temporal resolution without the need for a high-speed camera. It uses a custom hardware controller to modulate and synchronize multiple illumination and imaging devices. We share the code (for simulation and reconstruction) and hardware schematics to implement PAAQ, which allows replication and implementation in a wide array of existing lab microscopes.

1. Introduction

In vivo fluorescence microscopy is an essential tool to study the early stages of organ development in animal embryos thanks to its intrinsically selective contrast[22,40,46]. When used to image the zebrafish, an excellent animal model for cardiovascular research[2,30], it allows studying the development stages of the heart[6,38,42] and understanding the progression of early cardiac defects and diseases[8,16,41]. Correctly imaging the dynamics of the heartbeat requires high-speed acquisition due to the rapid beating of the heart. This becomes a limitation when the low fluorescence intensity emitted by the labelled structures requires a longer exposure time or extremely sensitive cameras that are either too slow or prohibitively expensive. In addition to this, the use of several fluorophores to label multiple tissues usually requires that their emission be recorded sequentially in different channels.

Some acquisition platforms are able to perform simultaneous high-speed acquisition in multiple fluorescent channels. This is usually achieved by splitting the emitted light into multiple beams that are redirected to different cameras, through the use of mirrors[13,29] and prisms[7]. However, these come at the cost of an increased complexity in the optics due to the addition of the splitting components. These optics and the use of multiple high-speed cameras make these systems very expensive and could make them difficult to integrate into existing imaging platforms.

As an alternative, when such parallel high-speed imaging is not available, images acquired from different channels at low speed could still be registered to their position in the heartbeat period (phase). Approaches to solve this problem fall into two categories: prospective gating and retrospective gating (reviewed in[43]).

In prospective gating, the acquisition of images is triggered at precise timings that correspond to desired sampling phases in the heartbeat cycle. On big samples, this trigger can be extracted from cardiac probes that precisely measure the start and duration of a heartbeat[5,18,39]. Such signals would be difficult to access in zebrafish embryos due to their small size. Instead, the trigger can be obtained by processing (in real-time) a video signal captured by a dedicated second camera[44,45]. These prospective gating methods give very good results even with slow acquisition devices, but they also require a dedicated optical setup with real-time processing capabilities for triggering.

Retrospective gating methods acquire images at arbitrary phases and then attempt to estimate these phases via post-acquisition algorithms. Their results are usually not as precise as those of prospective gating, but as they do not require additional triggering hardware they can be easier to implement with existing imaging setups. For example, it is possible to sort images acquired at random times in a period based only on image distance to reconstruct a singe-channel video with a virtually increased frame rate[26,27,47]. In addition to their limited accuracy, these methods can struggle with multichannel data. In some cases, it is possible to perform time registration between multiple channels by using a mutual information criterion[24,25] or by registering them with a common reference channel[33]. Nevertheless, these channel alignment methods require that high-speed sequences be available in each channel to work.

When imaging with slower cameras, one possibility would be to first generate virtual high frame rate videos of each channel separately using the sorting methods above[26,27]. These *virtual* high-speed sequences could then be synchronized using a mutual information-based algorithm designed for raw high-speed movies[23] to obtain the final a multichannel video. We see two main drawbacks to this potential method. Firstly, the virtual high frame rate sequences obtained with sorting are not uniformly sampled, and this sampling is different for each channel. This means that a perfect frame-to-frame pairing from one channel to another does not exist, which will negatively impact the performance of channel alignment. Secondly, high mutual information between the different channels is not guaranteed, and it can be difficult to obtain correct channel registration depending on the imaged region or the fluorophores used. These two drawbacks limit the expected performance of such a method, and could even make it unusable in some scenarios involving highly uncorrelated signals.

Instead of solving the virtual frame rate increase and channel registration problems sequentially, we propose a method that considers them jointly. In a previous work, we introduced a method to disambiguate images that appeared similar despite being captured in different phases of the heartbeat. We paired sharp images with blurred images that encode the motion at the time of acquisition [28]. In this paper, we extend this technique into a general approach centred around Paired Alternating AcQuisitions (PAAQ), which result in image sequences whose frames alternate between a common reference modality and other channels (e.g., fluorescence channels). We then use this common reference to sort images from all the channels at once, achieving simultaneously a virtual high frame rate and multichannel registration. In order to address the lack of precision of naive frame sorting approaches, we also propose a phase estimation algorithm. In addition to improving phase accuracy, it also overcomes two central limitations of plain image sorting: the resulting movies can be inverted in time and their phases cannot be assigned quantitative time units to measure durations.

2. Problem definition and metrics

We define f_c as the intensity of the signal obtained by imaging a periodic phenomenon (the embryo heartbeat) using fluorescence microscopy, with each channel number $c = 0, \ldots, C - 1$ referring to a distinct combination of illumination and emission wavelengths. The measured signal varies according to its spatial location (x, y) and its phase θ. The intensity f_c has a period of 2π, which is expressed as:

$$f_c(x, y, \theta) = f_c(x, y, \theta + k2\pi), \ \forall k \in \mathbb{Z}. \tag{1}$$

The phase itself is a function of time expressed as $\theta(t)$. In ideal time-periodic systems, it increases linearly over time t following $\theta = \omega t$, where ω is the angular frequency. In practice, such a relation is too simplistic for the cardiac cycle. Indeed, some variability stems from both the biological nature of the phenomenon and from environmental changes (e.g., temperature rising due to prolonged exposure to light leading to an increase in heart rate [32]). Although the exact relationship between time and phase $\theta(t)$ is unknown, it still follows a trend that is roughly linear $\theta(t) \simeq \bar{\omega} t$, where $\bar{\omega}$ is the average angular frequency calculated over multiple periods.

We consider C sequences of N images $\mathbf{f}_c[:, :, n]$ (where $:, :$ is a shorthand notation to describe all rows and columns of the image [12], and $n = 0, \ldots, N - 1$) obtained by sampling the signal f_c at unknown phases $\theta_{c,n}$ with a frame rate F_{acq}, using a procedure that we detail below. Our objective is to retrieve the phases $\theta_{c,n}$, which contain all the information required to reconstruct synchronized image sequences across all channels with an increased temporal resolution, as the phases give the position of each frame in the cycle.

More specifically, we want to compute estimates $\tilde{\theta}_{c,n}$ as close as possible to the ground-truth phases $\theta_{c,n}$, which are available only during simulation-based evaluation. We define our evaluation criterion as the following error function:

$$\mathcal{E} = \min_{\Theta} \frac{1}{NC} \sum_{c=0}^{C-1} \sum_{n=0}^{N-1} \left| \theta_{c,n} \ominus \left(\tilde{\theta}_{c,n} + \Theta \right) \right|, \tag{2}$$

where Θ is a phase shift common to all estimations, and $\ominus$ is a phase difference operator:

$$\alpha \ominus \beta = \alpha - \beta + k2\pi \ \text{ with } \ k \in \mathbb{Z} \ \text{ s.t. } \ \alpha \ominus \beta \in (-\pi, \pi]. \tag{3}$$

The global phase shift Θ in the error signifies that we do not care about the absolute phases of our measurements, but only about their position relative to each other in the period. Indeed, the phase of

the first frame in our signal is unknown, but this information is not necessary to derive meaningful information such as the time elapsed between events in the period, or to synchronize the channels.

3. Methods: acquisition and reconstruction

In order to solve the previously defined problem, we introduce a two-stage method that combines PAAQ imaging and post-acquisition processing to virtually increase the frame rate. The purpose of PAAQ is to create a reference signal common to all fluorescence channels that allows synchronizing the sequences with high precision. We then use the image-processing step to refine the estimate of the phase of each frame, as well as to compute the main properties of the imaged signal. We illustrate the whole pipeline in Fig. 1. Sections 3.1 to 3.4 detail the different stages of our method.
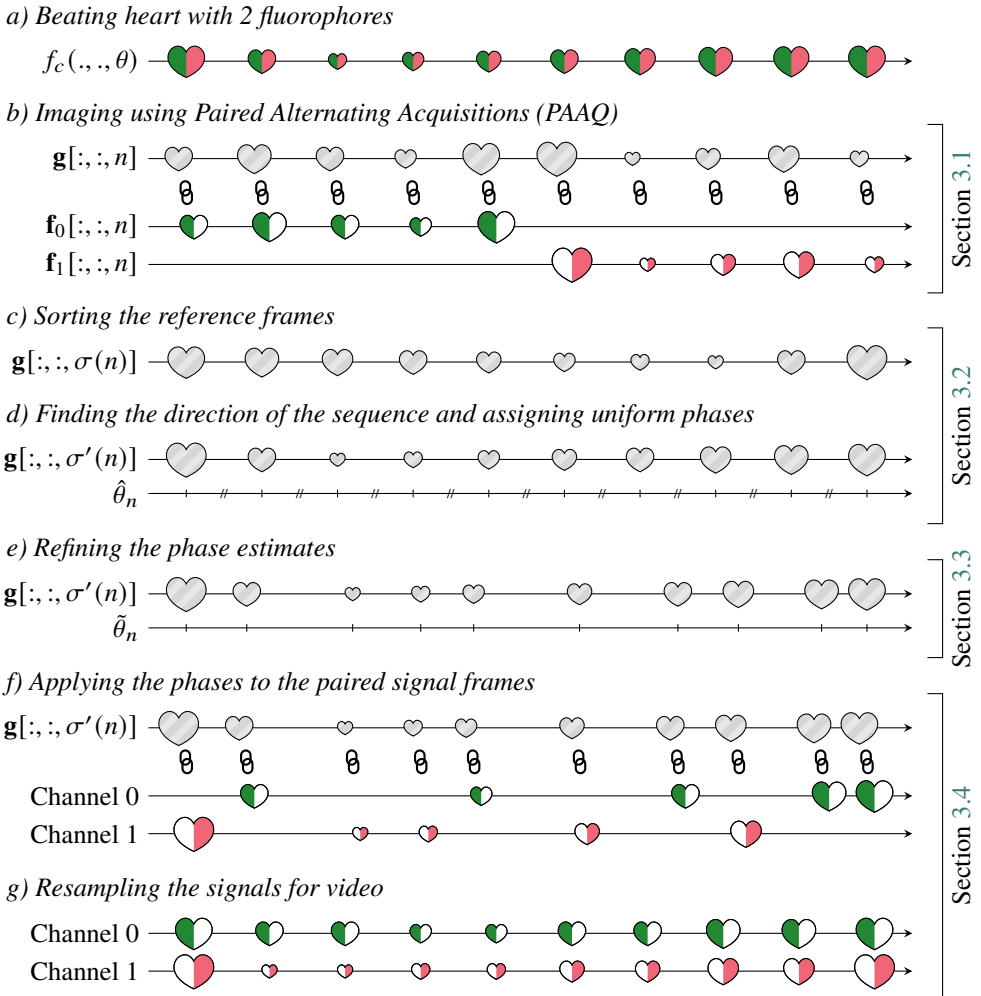

Figure 1. *Overview of proposed PAAQ method for virtual high frame rate cardiac imaging. Chain symbols represent images that we consider as paired because they were acquired in rapid succession.*

3.1. Implementing PAAQ with active illumination

In previous work [28], we have introduced a technique that uses active illumination to acquire pairs of images in quick succession by alternating between temporal illumination patterns. The first image in a pair encodes movement information for unequivocally sorting the sequence in order of increasing phase, while the second uses a short light flash that produces a sharp image used for analysis and display of the reconstructed sequence. We extend this method to achieve synchronization of multiple fluorescent channels by switching between imaging modalities at the same time as illumination patterns.

We start by defining a reference signal $g(x, y, \theta)$ that is obtained by imaging the beating heart using brightfield microscopy. Any modality is suitable to acquire g, and even one of the f_c could be the reference. However, since it is a sacrificial signal that will be discarded after imaging (as the time-encoding illumination results in blur), we choose to use brightfield as it causes little to no photobleaching. The intensity of this reference varies over the same space and phase as f_c:

$$g(x, y, \theta) = g(x, y, \theta + k2\pi), \ \forall k \in \mathbb{Z}. \tag{4}$$

For each channel, we use PAAQ to capture a sequence of N image pairs $(\mathbf{g}_c[:, :, n], \mathbf{f}_c[:, :, n])$, consisting of a reference image and a signal image captured in quick succession by switching from brightfield to fluorescence illumination every other frame:

$$\text{even frames:} \qquad \mathbf{g}_c[l, m, n] = \int_{t_c+2n\Delta_T}^{t_c+2n\Delta_T+\Delta_E} r(t - (t_c + 2n\Delta_T)) \, g\left(l\Delta x, m\Delta y, \theta(t)\right) \, dt \tag{5}$$

$$\text{odd frames:} \qquad \mathbf{f}_c[l, m, n] = \int_{t_c+(2n+1)\Delta_T}^{t_c+(2n+1)\Delta_T+\Delta_P} f_c\left(l\Delta x, m\Delta y, \theta(t)\right) \, dt, \tag{6}$$

where Δ_x and Δ_y are the pixel width and height, $l = 0, \ldots, L - 1$ and $m = 0, \ldots, M - 1$ the row and column index pairs, and $n = 0, \ldots, N - 1$ the time frame index. Δ_T is the time interval between two consecutive frames, and t_c the arbitrary time at which the first reference image associated to channel c is acquired. Since we capture the images without interruption, Δ_T directly relates to the acquisition frame rate of the camera $F_{\text{acq}} = 1/\Delta_T$. The phase corresponding to each image $\mathbf{f}_c[:, :, n]$ is $\theta_{c,n} = \theta(t_c + (2n + 1)\Delta_T)$ and $r(t)$ is the temporal illumination pattern used to capture images of the reference signal. It is defined over the interval $[0, \Delta_E)$ where Δ_E is the exposure time of the camera. When acquiring images of the fluorescent signals, a short light pulse of duration Δ_P illuminates the sample, with $\Delta_P \leq \Delta_E$. By using a short pulse, we can ensure that the fluorescent images are sharp even if the frame rate of the camera is low. Figure 2 illustrates PAAQ imaging, corresponding to Step b) in Fig. 1.

The PAAQ method is based on the assumption that the delay between a reference image and its associated signal image is small. By acquiring the reference image at the very end of the exposure time of the camera, we make this delay entirely independent of the frame rate of the device. Indeed, this delay becomes $\Delta_C = \Delta_T - \Delta_E$ the transfer time of the camera, which is the time between two exposure periods during which the device reads the pixel information and gets ready to capture a new frame. Even on devices with low frame rate, this transfer time is typically very low (on the order of a few ms). Therefore, we shape $r(t)$ to capture most information at the end of the exposure period, minimizing the delay between images in a pair. Instead of the continuous ramp used in the original method [28] we use a series of pulses of increasing amplitude for $r(t)$, as this pattern is easier to generate while retaining the same motion-encoding properties as the ramp.

Given that the delay between a signal frame and its associated reference is very small, the phase difference between the two will also be small due to the linear trend between time and phase. Over such a low duration, the variability in the heartbeat cycle is negligible, and we can consider the phase
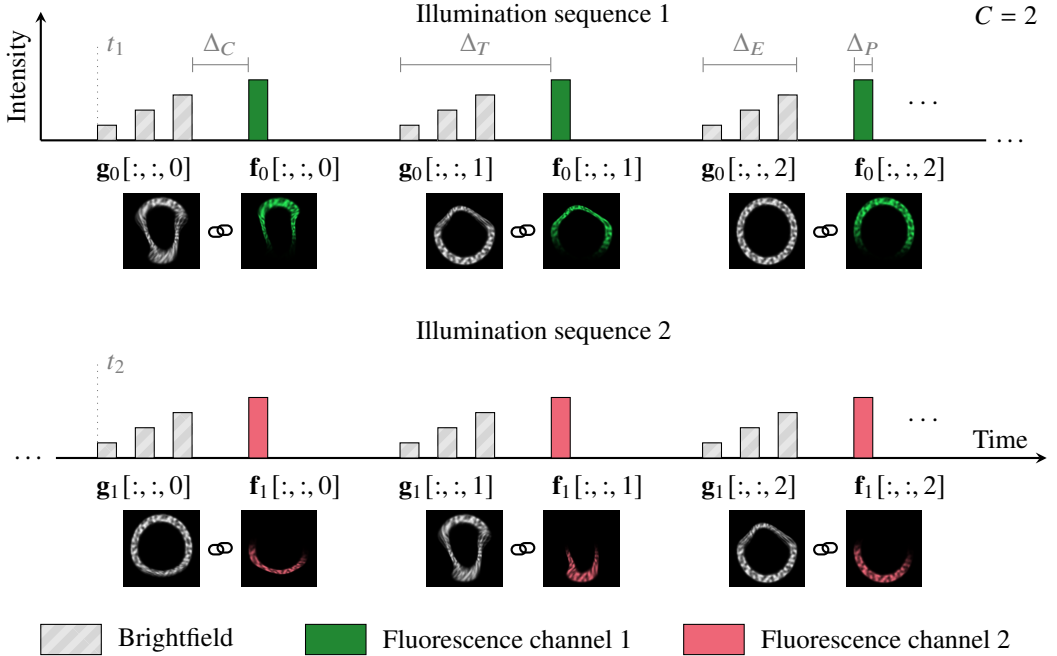
Figure 2. *Illumination modulation for PAAQ imaging. Quickly switching between different channels and modalities allows associating fluorescence frames to a common brightfield reference. Chain symbols represent images that we consider as paired because they were acquired in rapid succession.*

delay between each reference and its paired signal frame to be a constant. Therefore, since a constant phase shift does not impact the global error defined in Eq. (2), finding the phase of each signal frame is equivalent to finding the phase of its associated reference. This is a much easier problem to solve, as all references are of the same modality, and the results obtained on reference sequences will be directly applicable to their associated signals from different channels. We are thus able to extend the PAAQ technique to generate synchronized virtual high frame rate sequences from multiple fluorescent channels at no additional computational cost.

3.2. Sorting and orienting the references

We combine all the reference images into a single sequence $\mathbf{g}[:, :, n]$ by concatenating the acquired images $\mathbf{g}_c$ for each channel, containing a total of $N_g = NC$ reference images. Similarly, we concatenate signal frames from all channels into a single sequence $\mathbf{f}[:, :, n]$. We can then apply a phase-sorting algorithm to the reference sequence $\mathbf{g}$ to obtain a permutation $\sigma : \{0, \ldots, N_g - 1\} \rightarrow \{0, \ldots, N_g - 1\} : n \rightarrow \sigma(n)$ that sorts the references by increasing phase. For this we use a travelling salesman method introduced in previous works[26] that finds the permutation σ by minimizing the total absolute image difference between consecutive images. The obtained sorted sequence $\mathbf{g}[:, :, \sigma(n)]$ contains all the reference images reordered to reproduce a single period of the heartbeat, with a virtually increased frame rate that depends only on the amount of images N_g. This corresponds to Step c) in Fig. 1.

This sorting-based method assumes that the images in the reordered sequence correspond to a uniform sampling of the signal. It assigns a linear phase estimate based on the position of each frame in

the sorted sequence:

$$\hat{\theta}_n = \frac{\sigma^{-1}(n)}{N_g} 2\pi. \tag{7}$$

The travelling salesman solution does not contain directional information. Indeed, the permutation $\sigma'(n) = N_g - \sigma(n) + 1$ corresponds to a tour with the exact same cost. However, direction is important for the reconstruction of a faithful video that follows the chronological order of events. If we assume that the sampling frequency of the reference frames (i.e., half the frame rate of the camera) is at least twice as big as the frequency of the imaged phenomenon (similarly to the Nyquist criterion), we can retrieve the direction of the sequence by computing the average phase distance between consecutively acquired frames:

$$\bar{\Delta}_{\hat{\theta}} = \frac{1}{C(N-1)} \sum_{c=0}^{C-1} \sum_{n=0}^{N-2} \left(\hat{\theta}_{c,n+1} \ominus \hat{\theta}_{c,n} \right). \tag{8}$$

When the sampling frequency assumption is met, the phase distance between consecutive frames must be smaller than π (half a period), and $\bar{\Delta}_{\hat{\theta}}$ must be positive. If $\bar{\Delta}_{\hat{\theta}}$ is negative, the permutation obtained using sorting is in the wrong direction, and we must use σ' instead to sort the sequence and compute the uniform phase estimate $\hat{\theta}$. This is shown as Step d) in Fig. 1.

3.3. Non-uniform phase estimation using image distance

The assumption that the sorted sequence corresponds to a uniform sampling of a single period of the signal is imprecise and leads to artefacts in the reconstructed sequence. Due in part to the ratio between the acquisition frequency and the frequency of the imaged signal, and in part to the variations in the periodicity of the phenomenon, different regions of the period will contain more or less sampling points, as shown in Fig. 3 b). When reconstructing a video with the assumption that the sampling is uniform will generate distortions: stretching and shrinking will appear in low and high sampling rate areas, respectively, as illustrated in Fig. 3 c). It is therefore important to refine the phase estimate in order to reconstruct a faithful signal that correctly represents the dynamics of the imaged system.
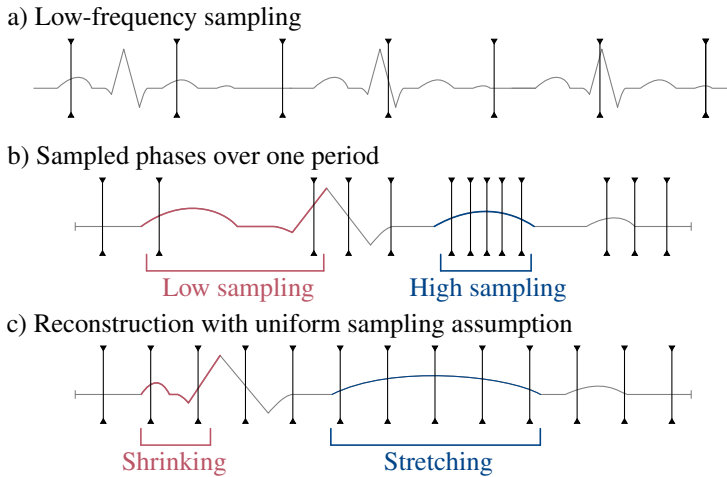

Figure 3. *Reconstruction artefacts introduced by the uniform sampling assumption. a) Fast repeating process sampled with a low frequency. b) Rearranged over a single period, the sampled points are not uniformly spaced. c) If assuming a uniform phase sampling the reconstructed signal is deformed.*

In order to correct distortions in volumetric electron microscopy imaging caused by inhomogeneous sampling in the axial direction, Hanslovsky et al. have developed an image-based method that uses image-to-image similarity to estimate the position of slices in the volume[15]. Although their application is concerned with re-sampling the depth dimension in static volumes rather than the temporal dimension in videos of the beating heart, the underlying ideas are highly relevant to our cardiac application problem and we have implemented an algorithm that is inspired by theirs. For the sake of completeness, we provide the derivation in full and point to similarities and differences.

The gist of the method is the following. If one were to fix an origin at the arbitrary time point with index n in the image sequence and plotted an image-distance between the image at the origin and images located at continuous-phase positions before or after the origin, one would obtain a function $\mathcal{D}_n(u)$, with u the phase relative to the origin. If that function was known, one could deduce the phase of any image relative to the image at the origin simply by computing their image-distance and inverting $\mathcal{D}_n(u)$. Since this function is neither known in practice nor unique—it depends on the image contrast, the field of view, and the cardiac phase of the image taken as the origin, etc.—the key idea is to approximate it by a model. Two assumptions allow building the model and estimating its parameters.

The first assumption is that the pixel-based distance between images in the signal increases with their phase difference. Indeed, images that are temporally close to one another in the signal should be similar, while images further away are more likely to look different. This echoes the hypothesis on which the sorting-based virtual high frame rate technique is built[26]. We use the Minkowski distance of order 1 as the distance metric d between two images $\mathbf{a}$ and $\mathbf{b}$ of size $L \times M$:

$$d(\mathbf{a}, \mathbf{b}) = \sum_{l=0}^{L-1} \sum_{m=0}^{M-1} |\mathbf{a}[l, m] - \mathbf{b}[l, m]| \,. \tag{9}$$

The above assumption means that the distance curve at any n, $\mathcal{D}_n(u)$, monotonically increases from 0 at $u = 0$ to a maximum. Yet since the distance curves also inherit the periodicity of the imaged signal:

$$\mathcal{D}_n(u) = \mathcal{D}_n(u + k2\pi) \,, \quad \forall k \in \mathbb{Z}, \tag{10}$$

the $\mathcal{D}_n(u)$ must first increase, then monotonically decrease from their maximum back down to 0 at $u = 2\pi$.

The second assumption of our method is that two distance curves corresponding to two nearby origin points (whose phase is close) should be near identical. This is referred to as *local constancy of shape* by the authors of the volumetric imaging method[15]. This assumption seems reasonable for our problem as the motion of the heart is continuous, and over any arbitrarily short timespan its speed is near constant.

In comparison to the method introduced for volumetric imaging[15] our assumptions are similar, except that we do not assume that the distance curves are monotonic over the whole domain as it could not account for the periodicity of the cardiac cycle. Instead, we impose that it can be split in two monotonic parts, one increasing and one decreasing, which repeat periodically. It is noteworthy that our derivation uses a distance metric (which decreases with the phase difference) rather than a similarity measure (which increases with the phase difference). Beyond the definition of $\mathcal{D}_n$, we will note further differences in Sections 3.3.1 and 3.3.2.

Based on our assumptions above, we can now derive a two-step algorithm to estimate the non-uniform positions of the samples: first we estimate the $\mathcal{D}_n(u)$ corresponding to each sample in our sequence, and then we use these to compute new phase estimates, as illustrated in Fig. 4. We iterate over these two steps until the phase estimation converges, which corresponds to Step e) in Fig. 1. The following subsections describe each step of the algorithm in more detail.
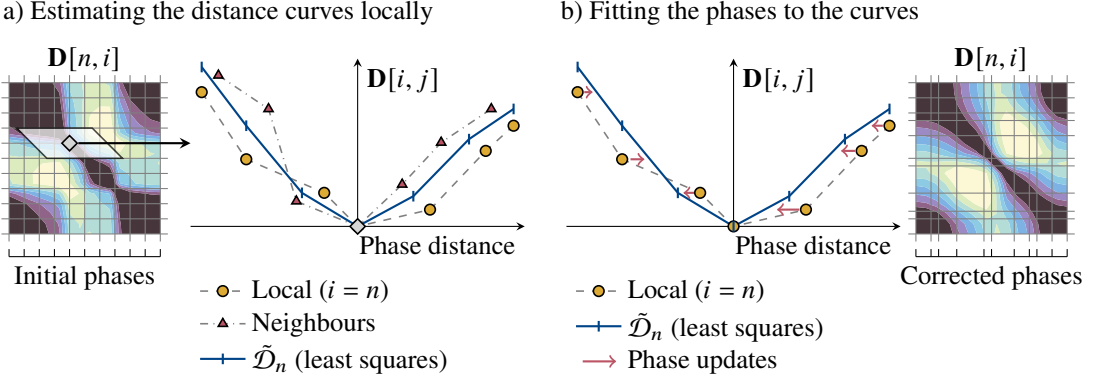
a) Estimating the distance curves locally

b) Fitting the phases to the curves

Figure 4. *Image-based phase estimation algorithm. a) We compute a local average distance curve $\tilde{\mathcal{D}}_n$ for each origin position n via least-squares fitting on the measurements $\mathbf{D}[:,:]$ b) We update the phase estimates by minimizing the distance between local measurements and the computed average distance curves. We iterate over these steps until convergence. The distance matrix $\mathbf{D}[:,:]$ gets smoother as the phase estimates improve.*

3.3.1. Estimating the distance curves locally

We model the distance curves $\mathcal{D}_n$ using piecewise-linear functions $\tilde{\mathcal{D}}_n$ characterized by values $\tilde{\mathbf{d}}_n[l]$, $l = 0, \ldots, L-2$ which we will compute below, defining L intervals over $u \in [0, 2\pi)$:

$$
\tilde{\mathcal{D}}_n(u) = \begin{cases} L\frac{u}{2\pi} \, \tilde{\mathbf{d}}_n[0] & \text{if } \frac{u}{2\pi} < \frac{1}{L} \\ \left(L\frac{u}{2\pi} - l\right)\tilde{\mathbf{d}}_n[l] + \left(l + 1 - L\frac{u}{2\pi}\right)\tilde{\mathbf{d}}_n[l-1] & \text{if } \frac{l}{L} \leq \frac{u}{2\pi} < \frac{l+1}{L}, \ l \in \{1, \ldots, L-2\} \\ \left(L - L\frac{u}{2\pi}\right)\tilde{\mathbf{d}}_n[L-2] & \text{if } \frac{L-1}{L} \leq \frac{u}{2\pi} < 1, \end{cases} \tag{11}
$$

$$
\tilde{\mathcal{D}}_n(u) = \tilde{\mathcal{D}}_n(u + k2\pi), \ \forall k \in \mathbb{Z}. \tag{12}
$$

We impose constraints on the values $\tilde{\mathbf{d}}_n[:]$ to fulfil the monotonicity assumption on $\mathcal{D}_n$ (stemming from the hypothesis that images farther away in phase look less similar):

$$
\tilde{\mathbf{d}}_n[0] \geq 0 \tag{13}
$$
$$
\tilde{\mathbf{d}}_n[l] \geq \tilde{\mathbf{d}}_n[l-1] \qquad \forall l \in \{1, \ldots, L_n^{\max}\} \tag{14}
$$
$$
\tilde{\mathbf{d}}_n[l] \leq \tilde{\mathbf{d}}_n[l-1] \qquad \forall l \in \{L_n^{\max} + 1, \ldots, L-2\} \tag{15}
$$
$$
\tilde{\mathbf{d}}_n[L-2] \geq 0, \tag{16}
$$

where $L_n^{\max}$ is the position of the maximum of the function $\tilde{\mathcal{D}}_n$, which is unknown a priori.

The image distance between any arbitrary origin and any other image at indexes n and i in our sequence $d(\mathbf{g}[:,:,n], \mathbf{g}[:,:,i])$ should match the value of the associated unknown distance curve evaluated at the corresponding phase distance: $\mathcal{D}_n(\tilde{\theta}_i - \tilde{\theta}_n)$. Taking advantage of the local constancy of shape of $\mathcal{D}_n$, we can consider all image distances in a local neighbourhood around any origin sample n as measurements of the same distance curve. This allows us to find the values $\tilde{\mathbf{d}}_n[:]$ for each n by

solving a weighted linear least squares fitting problem that minimizes:

$$\mathcal{L}_{\text{dist}}(\tilde{\mathbf{d}}_n) = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \left(\mathbf{D}[i,j] - \tilde{\mathcal{D}}_n(\tilde{\theta}_j \ominus \tilde{\theta}_i)\right)^2 w_\sigma\left(\tilde{\theta}_i \ominus \tilde{\theta}_n\right), \tag{17}$$

where $\mathbf{D}[i,j] = d(\mathbf{g}[:,:,i], \mathbf{g}[:,:,j])$ is the matrix containing all pairwise image distances between reference frames, and $w_\sigma(u)$ is a Gaussian windowing function of given standard deviation σ that gives more importance to measurements in a close neighbourhood. This step is illustrated in Fig. 4 a).

If the phase estimates $\tilde{\theta}_n$ are close enough to the real phases, then the modelled $\tilde{\mathcal{D}}_n$ will be good approximations of the distance curves $\mathcal{D}_n$. When solving this step for the first iteration, we initialize $\tilde{\theta}_n$ with the uniform sampling approximation $\hat{\theta}_n$ defined in Eq. (7) which gives a mean error of less than 10% of the period if the sorted sequence contains at least 10 images [27].

Taking into account the monotonicity constraints in Eqs. (13) and (16), we use an efficient quadratic cone programming algorithm [31] to minimize Eq. (17). However, writing the constraints requires knowing $L_n^{\max}$, which we find by solving the problem without monotonicity constraints beforehand. We detail the expression of this quadratic programming problem in Appendix A.

We note here a second set of differences between our method and the volumetric imaging one [15], which lie in how we write and solve the curve fitting problem. Indeed, we do not resample the distance matrix $\mathbf{D}[:,:]$ to compute the loss in Eq. (17). This gives a solution that corresponds more closely to the measurements but makes the least squares problem more complex to write in matrix form. Moreover, using a constrained solver allows us to guarantee the monotonicity of the estimate, thus removing the need of an additional truncating step to ensure the validity of the solution. We also choose to use all frames of the sequence to compute each curve estimate instead of using an additional windowing that would exclude distant points from the computations. This allows us to obtain distance curves defined over the whole period, which we will use in the following step to update the phase estimates.

3.3.2. Fitting the phases to the curves

Using the results of the previous step, we can now update the phases $\tilde{\theta}_n$ in order to make the measurements $\mathbf{D}[:,:]$ match the estimated distance curves $\tilde{\mathcal{D}}_n$ as closely as possible. We write this as a weighted least squares cost function:

$$C\left(\tilde{\theta}\right) = \frac{\sum_{n=0}^{N_g-1} \sum_{i=0}^{N_g-1} \left(\mathbf{D}[n,i] - \tilde{\mathcal{D}}_n\left(\tilde{\theta}_i \ominus \tilde{\theta}_n\right)\right)^2 w_\sigma\left(\tilde{\theta}_i \ominus \tilde{\theta}_n\right)}{\sum_{n=0}^{N_g-1} \sum_{i=0}^{N_g-1} w_\sigma\left(\tilde{\theta}_i \ominus \tilde{\theta}_n\right)}. \tag{18}$$

We use the same Gaussian windowing function w_σ as in Eq. (17) to give less importance to samples far from the origin point of the distance curves. This addresses the fact that these samples are more noisy and less reliable.

We introduce a regularization term derived from the fact that we acquire images consecutively for each channel, at a constant sampling rate. We consider the neighbourhood containing any three reference frames acquired successively. Over this region, we approximate the evolution of the phase over time using a partial Taylor sum of degree 1. As the time delay between the acquisition of consecutive reference frames is constant, according to this local model the phase step between these frames should be nearly constant. We express this as the following regularization cost:

$$\mathcal{R}\left(\tilde{\theta}\right) = \frac{1}{C(N-2)} \sum_{c=0}^{C-1} \sum_{n=1}^{N-2} \left(\left(\tilde{\theta}_{c,n+1} \ominus \tilde{\theta}_{c,n}\right) - \left(\tilde{\theta}_{c,n} \ominus \tilde{\theta}_{c,n-1}\right)\right)^2. \tag{19}$$

We combine the two cost functions using a regularization strength λ:

$$\mathcal{L}_{\text{phase}}\left(\tilde{\theta}_n\right) = (1-\lambda)\,C\left(\tilde{\theta}_n\right) + \lambda\,\mathcal{R}\left(\tilde{\theta}_n\right). \tag{20}$$

We minimize this objective function using a gradient descent approach. More specifically, we use the Adam adaptive learning rate optimizer[20] to find an optimal estimate for $\tilde{\theta}_n$.

Since we update the phase estimates in this step, the modelled distance curves $\tilde{\mathcal{D}}_n$ are no longer valid and we must recompute them with the new phases. We iteratively minimize Eqs. (17) and (20) until convergence of the phase estimates. We describe a reliable way to choose the values of the different hyperparameters for minimization in Section 3.6.

We note here the third main set of differences between the volumetric imaging method[15] and ours, which concerns the loss function $\mathcal{L}_{\text{phase}}$ used for phase estimation. First, we use the distance metrics rather than the phases to compute the cost in Eq. (18). This circumvents the need to use the inverse of $\tilde{\mathcal{D}}_n$, which is not clearly defined as the distance curves are not bijective functions. Then, we do not use the scaling factor introduced in the volumetric imaging method, as its purpose is to correct for varying degrees of noise between images, and localized artefacts on some frames. Given the imaging method we use, the noise level is comparable in all images, and using this scaling factor would only add more complexity to the algorithm with no expected performance improvement. Finally, we introduce a new regularization term in Eq. (19), tailored to our imaging procedure to ensure that the estimated phases are consistent with the acquisition sampling rate.

3.4. Reconstructing signal sequences

Our assumption that the phase delay between reference and signal images is constant allows us to use the reference phase estimates $\tilde{\theta}_n$ to reconstruct synchronized virtual high frame rate sequences from the signal frames, as shown in Fig. 1 f). To generate videos, we resample the sorted sequences at an arbitrary fixed rate Each video frame contains the signal frame of the corresponding channel with the closest possible phase, as shown in Step g) of Fig. 1.

By combining the phase estimates with the known acquisition times, we can compute the average frequency of the imaged signal over any region of the sequence. Given that the phase estimates are wrapped over a single period, we first apply a simple temporal phase unwrapping algorithm[17]:

$$\tilde{\theta}'_{c,n} = \tilde{\theta}_{c,n} + \phi_{c,n} \tag{21}$$

$$\phi_{c,1} = 0 \qquad\qquad \forall c \in \{0,\ldots,C-1\} \tag{22}$$

$$\phi_{c,n} = \begin{cases} \phi_{c,n-1} + 2\pi & \text{if } \tilde{\theta}_{c,n} - \tilde{\theta}_{c,n-1} < -\frac{\pi}{2} \\ \phi_{c,n-1} & \text{otherwise} \end{cases} \qquad \forall n \in \{1,\ldots,N-1\}, \tag{23}$$

where $\tilde{\theta}'_{c,n}$ are the unfolded phases for each channel, and $\phi_{c,n}$ represents how many periods have elapsed before its associated frame. It increases by a full period when the difference between two consecutive frames is lower than a threshold of $-\frac{\pi}{2}$, which prevents counting too many periods because of small errors in the phase estimates. We can then compute the average angular frequency over a chosen region of the signal by fitting a linear model on the points $(t_{c,n}, \tilde{\theta}'_{c,n})$ within that interval using linear least squares. The slope of the model directly gives $\bar{\omega}$ over the analysed region.

The frequency information, which was missing in previous sorting-based virtual high frame rate methods[26,28,47], allows to study the heart rate of the sample. It also allows converting phases into time units, which is better suited for analysing the dynamics of the heartbeat.

3.5. Theoretical analysis of the error caused by heart rate variability

In practice, as the heart rate varies, so does the phase delay between paired frames. This introduces errors in the reconstruction step described in the previous section, as it is based on the assumption that these phase variations are negligible. In order to find in which range this approximation is valid, we want to quantify the impact of the cardiac frequency variability on the performance of our method.

Using Eq. (2), we derived (see Appendix B) an upper bound for the error $\mathcal{E}_\epsilon$ that results from the variability in the heart rate:

$$\mathcal{E}_\epsilon \leq \frac{\Delta_C}{NC} \sum_{c=0}^{C-1} \sum_{n=0}^{N-1} \left| \epsilon_{c,n} \right|, \tag{24}$$

where Δ_C is the time delay between paired images, and $\epsilon_{c,n}$ is the deviation between the instantaneous frequency and its average: $\omega_{c,n} = \bar{\omega} + \epsilon_{c,n}$.

Thus, we can compute the upper bound of the error introduced by heart rate variability if we know the statistical distribution of the cardiac frequency during the acquisition. For practical applications, it is easier to describe it by using $\omega_{\min}$ and $\omega_{\max}$, the minimal and maximal values (respectively) that this frequency can take over the imaging process.

In the worst possible case, the heart rate is maximal during half of the acquisition and jumps without transition to its minimal value for the other half of the acquisition. The bound in Eq. (24) then becomes:

$$\mathcal{E}_\epsilon \leq \Delta_C \frac{\omega_{\max} - \omega_{\min}}{2}. \tag{25}$$

If the cardiac frequency is uniformly distributed over its range during the acquisition (e.g., if the heart linearly accelerates at a constant rate), the bound gets smaller:

$$\mathcal{E}_\epsilon \leq \Delta_C \frac{\omega_{\max} - \omega_{\min}}{4}. \tag{26}$$

And if its distribution is a normal (e.g., if the heart rate oscillates around its average value), using the three-sigma rule to set the interval, the bound tends towards an even lower value:

$$\mathcal{E}_\epsilon \leq \Delta_C \sqrt{\frac{2}{\pi}} \frac{\omega_{\max} - \omega_{\min}}{6}. \tag{27}$$

In the case of a healthy heart that smoothly accelerates or decelerates, and may settle around a given cardiac frequency, the distribution lies between the uniform and the normal.

For example, we compute the upper bound in a realistic scenario with a delay of $\Delta_C = 10\,\text{ms}$ between reference and signal images, and an average heart rate of 2.5 beats per second, considering the cardiac frequency distribution to be at worst uniform. If the heart rate stays between 0.5 and 4.5 beats per second, the error due to cardiac frequency variability will be lower than 1% of the period: $\mathcal{E}_\epsilon \leq 10^{-2} \frac{4.5-0.5}{4} 2\pi$.

We consider a different scenario with all parameters equal except the interval $\Delta_C = 0.033\,\text{s}$. This corresponds for example to using a camera with a frame rate of 30 fps in a setup where triggering the reference acquisition at the end of the exposure time is not possible. To guarantee the same 1% upper bound as above in this scenario, the heart would have to stay between 1.9 and 3.1 beats per second: $\mathcal{E}_\epsilon \leq 0.033 \frac{3.1-1.9}{4} 2\pi$.

We can use these bounds to predict what performance to expect, based on rough estimates of the range of heart rates that are expected or measured. This error induced by frequency variability will then be combined with the precision of the phase estimation algorithm when measuring the overall error of our method.

3.6. Reliable computation of the hyperparameter values

The performance of our phase estimation algorithm depends on the values of multiple hyperparameters. In order to be able to apply our method in practice, we propose a reliable way to compute optimal values for these parameters based on the acquired images. This guarantees the reproducibility of our results, and allows to apply our method in new scenarios without requiring manual hyperparameter tuning.

The regularization strength λ is the parameter with the most impact on performance. We find its optimal value using an L-curve[14]. This consists in plotting the values of the residual cost from Eq. (18) against the final regularization cost from Eq. (19) for multiple values of λ. The curve resembles the shape of an L, as seen in Fig. 6 c). To compute the optimal λ, we first apply min-max normalization to both costs then multiply the regularization cost by a factor 10. This is because in our case, the regularization corresponds to a strong prior about the system, and must play a bigger role in selecting λ. We then choose the point in the rescaled L-curve that is the closest to the origin as the optimal λ.

The width of the Gaussian window σ used in Eqs. (17) and (18), as well as the number of points L used to characterize the distance curves also impact the performance. To guarantee good results, σ must be small enough to capture local details, but large enough to avoid amplifying noise. The precision overall improves with larger L, but at the cost of a slower convergence. In Appendix C, we detail our method to find appropriate values for these parameters and for the gradient descent update step.

To enhance the performance of phase estimation, we can restart our iterative algorithm after convergence, using the obtained phase estimates as the initial guess for the new run. We call this additional step a precision pass. When restarting the minimization, we halve the width of the Gaussian window σ and double L to get finer phase estimates. Even with the higher L and smaller σ, the minimization converges as it starts from a better initial guess. This allows getting better precision using bigger values of L that would have caused convergence issues if used directly on the initial uniform phase estimates.

4. Methods: fluorescence cardiac imaging simulation

Previous works used simplified simulation models of the heart for performance characterization, consisting of a circle (or cylinder) textured with a sinusoidal pattern, with a radius varying over time to represent beating[23,26–28]. These models lack properties such as irregular texture, asymmetric contraction or small variations in the periodicity that we find in experimental cardiac images. In order to test the different methods and characterize their performance as reliably as possible, we designed a simulation tool that reproduces the relevant properties of fluorescent beating heart images. We aim to synthesize a smooth organic-like textured object that deforms following a nonlinear contraction wave that propagates periodically through time and space asymmetrically. We also want to generate multiple channels that contain variable amounts of mutual information, and to emulate the small variations in the periodicity of the heartbeat.

4.1. Simulating a beating heart section

We define a function $h(x, y)$ over a two-dimensional space to simulate a heart section:

$$h(x, y) = e(x, y)\, \sigma(x, y), \tag{28}$$

where $e(x, y)$ represents an ellipse with a thick outline and $\sigma(x, y)$ is a two-dimensional simplex noise[34,35] used for texturing. Simplex noise is a gradient noise function that generates visually isotropic and continuous textures, and is commonly used in computer-generated graphics to procedurally create natural looking images. Its properties make it suitable for creating synthetic textures with realistic features in biomedical imaging[1,3,9,10], and we use it here to generate an organic-like texture for the heart walls.

For later use in our simulations, we introduce a periodic continuously smooth pulse function obtained by composing a sigmoid and a sine function:

$$\text{sigsin}(\theta) = \frac{\text{sig}\left(\left(\sin(\theta) - b\right) s\right)}{\text{sig}\left(\left(1 - b\right) s\right)} \tag{29}$$

$$\text{sig}(x) = \frac{1}{1 + \exp(-x)}, \tag{30}$$

where b is the sigsin bias and s its slope. b controls the width of the pulse, with higher bias generating shorter pulses, while s affects the overall slope of the rise and fall of the function, with higher slope creating a fast rising pulse.

In order to simulate the propagation of heart contraction through the section, we create a continuous space transformation that preserves topology. This transformation varies on x and θ in order to model the propagation of the contraction along one spatial dimension in time. It consists locally of a small rotation followed by a scaling, representing the twisting and shrinking of the heart walls that happen during contraction. This transformation can be expressed in matrix form:

$$\mathbf{T}(x, \theta) = \mathbf{R}(x, \theta) \cdot \mathbf{S}(x, \theta) \tag{31}$$

where the rotation and scaling matrices $\mathbf{R}(x, \theta)$ and $\mathbf{S}(x, \theta)$ are defined as follows:

$$\mathbf{R}(x, \theta) = \begin{bmatrix} \cos(\phi(x, \theta)) & -\sin(\phi(x, \theta)) \\ \sin(\phi(x, \theta)) & \cos(\phi(x, \theta)) \end{bmatrix} \qquad \mathbf{S}(x, \theta) = \begin{bmatrix} 1 + S_x(x, \theta) & 0 \\ 0 & 1 + S_y(x, \theta) \end{bmatrix}. \tag{32}$$

The rotation angle $\phi(x, \theta)$ and shrinking factors $S_x(x, \theta)$ and $S_y(x, \theta)$ are governed by a sigsin pulse:

$$\xi(x, \theta) = A \, \text{sigsin}\left(\theta + \frac{x}{\lambda} + \theta_0\right), \tag{33}$$

where A is the amplitude of the pulse, θ_0 is its initial phase shift, and λ is the spatial wavelength of the contraction pulse. The parameters A, θ_0, b and s are set independently for the functions ϕ, S_x and S_y, allowing to adjust the model to obtain a more realistic contraction.

Using the defined transform, we then compute the simulated reference signal as:

$$g^{\text{sim}}(x, y, \theta) = h\left(\mathbf{T}(x, \theta) \cdot \begin{bmatrix} x \\ y \end{bmatrix}\right). \tag{34}$$

In order to generate multiple fluorescent channels, we multiply this reference signal with two-dimensional masking functions $m_c(x, y)$:

$$f_c^{\text{sim}}(x, y, \theta) = h\left(\mathbf{T}(x, \theta) \cdot \begin{bmatrix} x \\ y \end{bmatrix}\right) m_c\left(\mathbf{T}(x, \theta) \cdot \begin{bmatrix} x \\ y \end{bmatrix}\right). \tag{35}$$

This method to simulate multiple fluorescent channels is not biologically plausible, but serves only as a means to easily control the amount of overlap between channels through the shape of the masks. This is sufficient to characterize the performance of our algorithms, and a biologically more accurate model is unnecessary in this scope. Simulated heart sections in different positions are visible in Fig. 2.

4.2. Modelling the heart rate variability

We model the variability of the heartbeat as the combination of two contributions: a random jitter that affects the phase locally, and a random acceleration that affects the phenomenon over a longer time span. The first is motivated by the imperfect nature of a biological process that creates local variations within a cycle, while the latter represents a slow variation of the average angular frequency over time. Using this model, we express the simulated phase at time t as:

$$\theta^{sim}(t) = \theta_0^{sim} + \int_0^t \omega_0^{sim} + \sigma_\omega W_\omega(t) \, dt + \sigma_\theta W_\theta(t), \tag{36}$$

with θ_0^{sim} the starting phase and ω_0^{sim} the initial angular frequency. We use Wiener stochastic processes to model the variability in phase $W_\theta(t)$ and in frequency $W_\omega(t)$ of the system, defined by:

$$W(0) = 0 \qquad W(t + u) - W(t) \sim \mathcal{N}(0, u), \; \forall t, u > 0, \tag{37}$$

with $\mathcal{N}(0, u)$ is a normal distribution of mean 0 and variance u. Their amplitudes σ_θ and σ_ω allow to control the overall amount of randomness in the simulation, as well as the strength ratio between the two sources of stochasticity. We chose Wiener stochastic processes because of their property to generate continuous paths, which is essential for modelling a heartbeat cycle as the motion of the heart is itself continuous. We illustrate the evolution of phase variability in multiple scenarios using different strengths for phase and frequency deviations in Fig. 5.
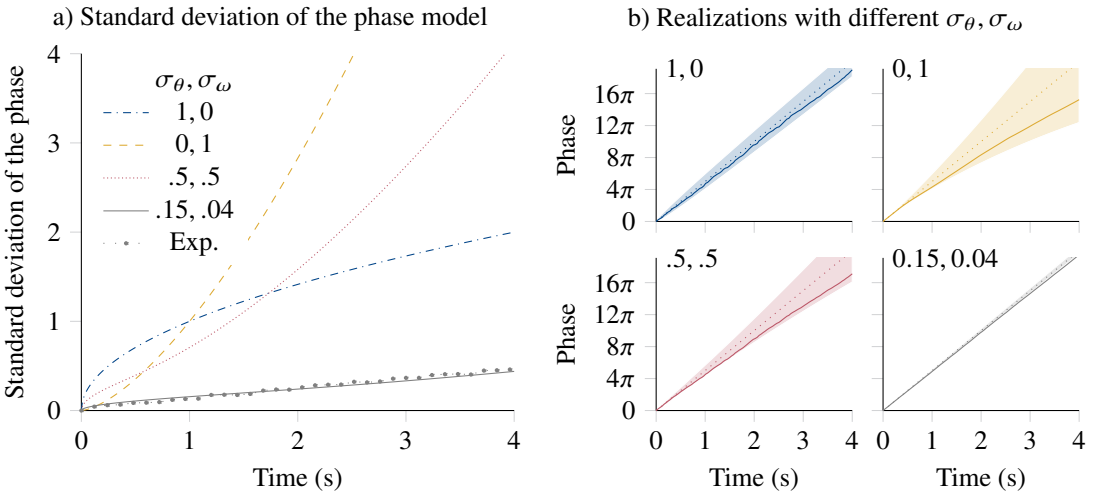
Figure 5. *Stochastic phase model for simulating heart rate variability. Our model uses two parameters to represent variability: a phase deviation σ_θ and a frequency deviation σ_ω. Plots illustrate the model for a simulated 2.5 beats per second signal. a) Varying the ratio of these contributions changes how fast the phase uncertainty increases, with σ_ω generating a rapidly increasing variance. Measurements on experimental data match the simulation model. b) Realizations of our model in different scenarios show that σ_θ generates local noise, while σ_ω contributes to bigger smoother variations. The bottom-right panel illustrates values matching experimental measurements.*

Using the phase $\theta^{sim}(t)$, we sample the simulated signals g^{sim} and f_c^{sim} as defined in Eqs. (5) and (6) to obtain simulated image pairs $\left(\mathbf{g}_c^{\mathrm{sim}}[:,:,n], \mathbf{f}_c^{\mathrm{sim}}[:,:,n]\right)$. Thanks to the biologically inspired simulation models, these images share the relevant properties allowing us to use them to characterize the methods exposed in this work. A full simulated heartbeat is visible in supplementary Video 1, where we also show the simulation of PAAQ imaging generated with our phase model.

5. Experiments and results

We characterized our method on both synthetic and real images. The following sections detail how we conducted these experiments, and discuss the obtained results. For fast computations, we implemented our gradient descent algorithm using JAX[4], a high-performance numerical computing framework accelerated with parallel computing. When running experiments, we noticed that keeping the state of the Adam optimizer from one iteration of the phase estimation to the next instead of resetting its parameters greatly speeds up convergence.

5.1. Validation of the simulation model

To confirm the plausibility of our phase simulation model, we measured the heart rate variability in a high-speed dataset that was acquired for previous works[26]. It consists in a 55 hpf wild type zebrafish embryo imaged on a Leica DMR microscope equipped with a FASTCAM SA3 camera using brightfield at 1000 fps. We randomly isolated a subsequence of 4 seconds in the data, and identified its first period using image similarity. We assigned linear phases from 0 to 2π to the frames of this first period, and then estimated the phase of all the other images in the subsequence as the phase of the most similar frame in the reference period.

We used these estimations to measure the phase deviation with respect to a perfectly periodic repetition of the first period. We repeated this measurement many times with different starting frames to build a relevant statistical representation of the experimental data. As shown in Fig. 5 a), the measurements closely match our model when using the values $\sigma_\theta = 0.15$ and $\sigma_\omega = 0.04$. This corresponds to a relatively stable heart rate, as illustrated in the bottom-right panel of Fig. 5 b).

5.2. Characterization on synthetic data

We simulated the imaging of a heart with an average of 2.44 beats per second in two fluorescent channels ($C = 2$) using a camera with a frame rate of 20 frames per second (fps). We generated images of size 256×256, and added Poisson noise to each pixel before quantizing their values to 12 bits. We used $\sigma_\omega = 0.04$ and $\sigma_\theta = 0.15$ to model the variability of the phase with comparable dynamics to what we measured on experimental data. We used a camera transfer time of 5 ms as the gap between reference and signal frames for PAAQ. We simulated imaging sequences of variable length, repeating each experiment with 10 different random seeds to generate confidence intervals.

We used $L = 50$ for the phase estimation, and set the other parameters using the strategy described in Section 3.6. Before applying the method we downsampled the images by a factor 2 with spatial averaging, to reduce the computation time and noise.

As a baseline for performance comparison, we considered the ungated imaging of the fluorescent channels at low frame rate without using PAAQ, in addition to the acquisition of the brightfield images without illumination patterns to serve as registration reference. We sorted each channel separately for virtual high frame rate, and synchronized the obtained sequences using a temporal registration procedure based on maximizing the mutual information between the fluorescent channels and the brightfield sequence. This corresponds to a sequential approach to solving virtual high frame rate and multichannel registration, as discussed in Section 1. As the direction of the sequences is not given by the sorting,

we applied the registration to all possible combinations of orientations for the channels, and kept only the solution with the maximum mutual information. We used a simple temporal upsampling of the sequences to allow for sub-frame precision registration.
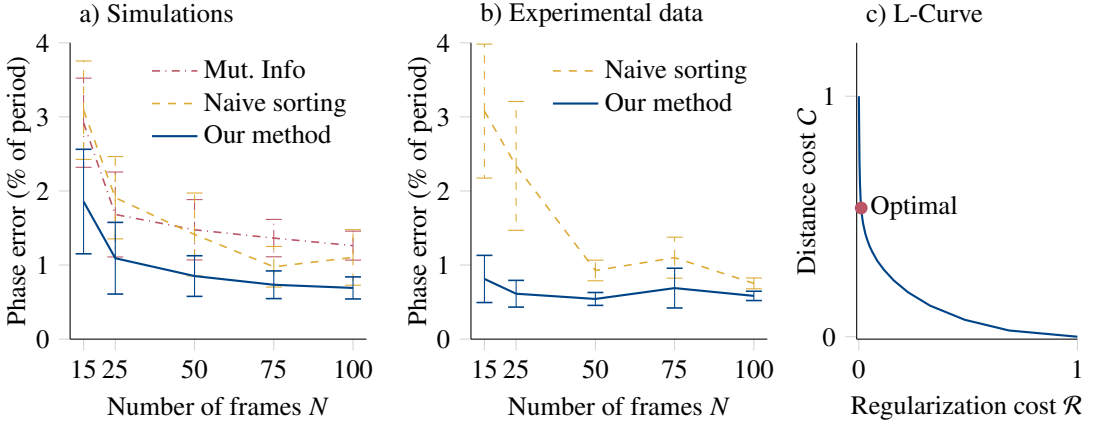

Figure 6. *Quantitative results on synthetic and experimental data. Error bars are 68% confidence intervals computed over 10 repetitions. a) On synthetic data, our method yields a relative precision improvement of 30% over competing methods, independently of the number of frames acquired. b) On experimental data, our method performs similarly, with a relative precision improvement of up to 50% (for $N \leq 25$). c) We choose the regularization strength for phase estimation using an L-curve, emphasizing the regularization cost due to strong priors.*

Figure 6 a) compares how the phase estimation performance of the methods varies with the number of frames acquired. The naive sorting method refers to the solution given by using PAAQ imaging and estimating the phases with the uniform sampling assumption (Steps b-c in Fig. 1, i.e., no phase correction algorithm). The mutual information-based method and the naive sorting obtain a similar performance over the whole range of experiments, which is expected as they are both limited by the uniform sampling assumption for phase estimation. Our method consistently improves on this performance, giving a relative 30% reduction of the phase error overall. Using Eq. (24), we computed that the error caused by heart rate variability is bounded to a maximum of 0.13% of one period on our simulated data. Therefore, the plotted phase error mostly characterizes the precision of our phase estimates, showing the benefit of using a non-uniform phase estimation algorithm.

We computed the initial average frequency of the simulated signal over the 2 first acquired periods using the method detailed in Section 3.4. The obtained relative error is directly tied to the error on phase estimation, dropping from 2% to below 1% when $N \geq 50$. We repeated these experiments with increasing Poisson noise strength, with no noticeable performance impact on any of the methods.

5.3. Validation on high-speed experimental data

In order to confirm the applicability of our method on real data, we used the same high-speed dataset as in Section 5.1. To simulate slow PAAQ imaging, we randomly picked a frame as the start, then sampled the dataset every 100 frames to obtain the reference sequence, corresponding to a simulated camera frame rate of 20 fps. To emulate the paired signal frames, we sampled the data similarly starting 5 frames after the initial reference frame, corresponding to a simulated camera transfer time of 5 ms. We repeated this process with a different random starting frame to simulate the acquisition of a second

channel. In this experiment, the signal sequences are all acquired using brightfield, but it does not affect the conclusions that we can draw on the precision of our phase estimation, as we use only the references for computation.

In order to generate the reference phases, we isolated a full period in the high-speed data starting from the first reference frame of the first channel, using image similarity to automatically find the start of the next period. We assigned linear phases from 0 to 2π to the frames in this initial period. We then estimated the ground-truth phase of all signal frames in all channels as the phase of the most similar frame (by image distance) in the reference period. We cannot know the real ground-truth phase of our samples, but this estimate is a sufficiently precise approximation given the high acquisition speed of this dataset. Indeed, since it is an order of magnitude higher than the frame rates we are simulating, the error on the ground-truth estimate has a negligible impact on our observations.

The high-speed dataset was acquired without the use of a motion-encoding illumination pattern, therefore it is missing in our simulated reference images. However, these data do not contain similar frames at different phases that would require motion disambiguation for sorting. Thus, the lack of patterned reference illumination does not impact the results of our method in this case, since it only affects the sorting of ambiguous frames. In a practical application, assessing the necessity of a motion-encoding pattern is only possible if high-speed reference images are available.

We simulated the slow imaging with varying sequence lengths, repeating each experiment 10 times with different random seeds to compute confidence intervals. For the phase estimation, we used $L = 50$ and set the other hyperparameters according to Section 3.6, running 2 additional precision passes for best performance.

Figure 6 b) compares the phase estimation error obtained with naive sorting (see Section 5.2) and our method. Due to the absence of separate fluorescent channels in the data, we did not evaluate the mutual information approach. The performance of naive sorting is consistent with what we measured on the simulated dataset. Our method achieves even better performance here, obtaining a phase error below 1% of the period for all sequence lengths. This represents a significant relative error reduction of more than 50% for sequences with few acquired frames ($N \leq 25$).

The accuracy of the uniform phase sampling assumption increases with the number of acquired frames in both Fig. 6 a) and b). As the sampling becomes denser, the average distance between samples decreases, leading to a similar drop in phase error. However, even with a high number of frames, the sampling is non-uniform and our algorithm still provides a better accuracy.

Figure 6 c) illustrates the L-curve obtained for an experiment with $N = 50$. The highlighted point corresponds to the selected value of regularization strength λ used for the phase estimation. It lies close to the vertical axis, as the regularization corresponds to a strong prior on the data.

5.4. Results on low-frame rate experimental data

We performed experiments with zebrafish (Danio rerio) embryos from animals held at Institute of Anatomy (National Licence Number 35) from the University of Bern. Adult fish needed for breeding were raised and maintained at maximal 5 fish/L with the following environmental conditions: 27.5 to 28 °C, with 14 h of light and 10 h of dark, 650 to 700 μs/cm, pH 7.5 and 10% of water exchange daily. Adult Tg(*fli1a:GFP*)^{y1Tg} [21] and Tg(*myl7:mRFP*)^{ko08Tg} [37] double transgenic zebrafish were crossed to obtain homozygous embryos. Eggs were collected within 30 min and kept in E3 medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl2, 0.33 mM MgSO4) with Methylene Blue (10^{-5}%). 30 min after egg collection, those that did not transition to two-cell stage were removed. The next day, at 24 h after removing the screens, chorions were removed by incubating 2 mg/mL Pronase in E3 medium (about 3 min) until gentle shaking of the Petri dish freed the larvae from the chorion. Subsequently, embryos of the same developmental stage were selected and the most frequent stage was estimated [19]. Before 24 hpf, we added 0.003% 1-phenyl-2-thiourea (PTU, Sigma-Aldrich) to avoid pigmentation.

We anaesthetized the embryos with Tricaine at 0.08 mg/mL, pH 7 and embedded the embryos with the anterior side (head) up in a fluorinated ethylene propylene tube in 1% low melting agarose. We imaged the sample on an implementation of the OpenSPIM microscope[36] with two lasers (Vortran Stradus, 488 and 561 nm), a UMPlanFL N 20×/0.50 W semi-apochromat water dipping objective lens (Olympus), and a sCMOS camera (Andor Zyla 4.2) mounted on a U-TV 0.5×C-3 adapter (Olympus).

We designed a low-cost hardware controller to generate modulated illumination patterns in 4 different channels based on an Arduino microcontroller. We used standard BNC connexions to make it compatible with most imaging and illumination hardware found on the market. We implemented drivers and plugins to integrate it into the μManager control software[11] and provide a user-friendly interface to use the device. Supplementary Figure 1 shows the controller and its schematics. We open-sourced our custom hardware and software at https://github.com/idiap/CBI-MMTools. This provides a low-cost and easy way to implement our method on devices that cannot easily generate the required temporal illumination patterns.

We used this controller to generate the PAAQ illumination patterns depicted in Fig. 2. We acquired images with a frame rate of 14.42 Hz, corresponding to a 60 ms exposure time Δ_E and a camera transfer time of 10 ms. We set a pulse width Δ_P of 10 ms for the signal illumination, and a linear ramp containing 3 pulses of 10 ms separated by 15 ms gaps for the reference pattern. We used $L = 50$ for the phase estimation, and used Section 3.6 to set the other hyperparameters, running one additional precision pass for performance.
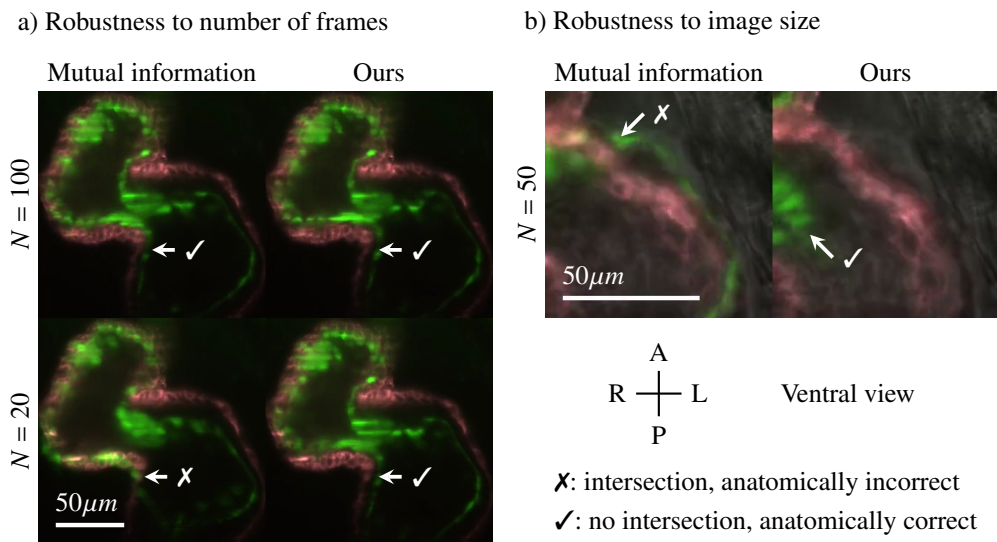
Figure 7. *Comparison of the reconstructed videos using our method and a mutual information approach, 4 dpf transgenic Fli1V/Myl7mR zebrafish heartbeat. Cross marks highlight anatomically implausible spots where myocardium (red) intersects with endocardium (green). a) With high number of frames both methods give similar results, but mutual information gives artefacts with N = 20, while our method is not impacted. b) When working on small images, mutual information gives incorrect channel synchronization between red and green, while our method generates a plausible solution. See also supplementary Videos 2–4 .*

Our first experiment compares the performance of our algorithm to the mutual information-based method when processing 2 fluorescent channels with varying number of images. We extract a region

of size 512×512 centred on the heart and downsample the image by a factor 2 with spatial averaging before using the algorithms. Figure 7 a) shows a frame in each of the reconstructed videos for $N = 100$ and $N = 20$, corresponding to a single estimated phase in the signal. With a high number of frames, the two methods give visually similar results, however with fewer frames the mutual information video contains artefacts in the synchronization of the channels. This is highlighted at the cross mark, where the myocardium (red) appears to intersect with the endocardium (green), which is anatomically implausible. With the same number of frames, our method does not generate this artefact, and stays consistent with the solutions computed with more frames. This illustrates the better robustness of our algorithm with respect to the amount of images available. Moreover, the solution obtained using mutual information contains anatomical artefacts even when using a high number of frames, as visible in supplementary Video 2. In the same situation, our method generates a video that does not contain visible anatomical artefacts, benefitting from the better phase estimates.

Our second experiment compares the performance of the two methods when the amount of information available is limited. We achieve this by restricting the field of view to only a smaller part of the heart in a region of interest of size 256×256. We also downsample by a factor 2 before applying both methods. For this experiment, we acquired $N = 50$ images in $C = 3$ channels, treating the brightfield modality as an additional signal channel. As shown in Fig. 7 b), the mutual information reconstruction fails to correctly synchronize the red and green channels, even though the synchronization between red fluorescence and brightfield is correct. In the highlighted area, the myocardium (red) and endocardium (green) visibly intersect in an impossible way in the solution given by mutual information, while in the corresponding frame obtained with our method the channels are correctly synchronized. The incorrect solution obtained with mutual information is visible in supplementary Video 3, where the green channel follows the movement of the red one with an obvious delay. In the smaller image, the amount of mutual information between channels is lower, partly because there are not enough relevant data available to build a meaningful statistical representation of the sequences. Our method does not require mutual information between the channels to give a correct solution, making it suitable to work on data with limited size but also in cases where the different channels contain entirely uncorrelated signals.

Supplementary Video 4 illustrates the whole reconstruction process, showing the acquired images and the virtual high frame rate multichannel video reconstructed using our method. It corresponds to the top-right frame in Fig. 7 a).

6. Discussion

Our method is able to generate multichannel videos with a virtual high frame rate with high temporal precision. As derived in Section 3.5, the channel synchronization error caused by heart rate variability is bounded and depends only on the transfer time of the imaging device. Therefore, the overall accuracy of our method mainly depends on the total number of images acquired, which can be chosen as high as needed to achieve a desired temporal resolution. The frame rate of the acquisition device has no direct impact on the precision of the reconstruction. Nevertheless, a faster frame rate will speed up the overall imaging time. In practice, imaging over a shorter period means that the heartbeat is more likely to remain regular.

Although we used a sensitive camera with low noise to obtain the experimental results presented in Section 5.4 we expect our method to perform similarly well on more noisy images. Indeed, the spatial averaging and downsampling step applied before sorting reduces the strength of noise present in the images used for phase estimation. This is supported by the results in Section 5.2, where we observed that increasing the noise in our simulations did not noticeably impact the performance of the algorithm.

Prospective optical gating methods[44,45] provide a high accuracy, with a better control on when the cardiac cycle is sampled. They can also be used for 3D imaging, by using the same trigger to synchronize videos at different depths. However, our technique has lower hardware requirements, making it more cost-effective and easier to integrate into existing imaging platforms. Our method is currently limited to 2D imaging, although it might be generalized by implementing depth sweeping during imaging, combined with a method that removes the induced scan aberrations[27]. Another possibility would be to use an imaging modality with a large depth of field for the reference frames, allowing to accurately synchronize neighbouring sections of the volume through a temporal registration step.

Methods that do not use PAAQ imaging (such as the mutual information synchronization used for comparison) have no hardware requirements at all. Nevertheless, we have shown that our technique is more robust to the size of the data available, both in terms of amount of images and their size. At the cost of a more complex acquisition procedure requiring illumination switching, it guarantees a better temporal resolution. Therefore, our method offers a compromise, giving high accuracy while still keeping the hardware cost and complexity low.

The virtual high frame rate reconstructions obtained with our method can still contain artefacts where the sample is moving aperiodically, or according to a period different from the cardiac cycle. This is visible in Supplementary Video 4, where a trabecula at the top of the reconstructed video appears to have an aperiodic motion. This limitation is inherent to slow frame rate imaging and only a fast acquisition device would be suitable to capture such aperiodic phenomena accurately.

Similarly, registration between channels is not possible in regions of the sample if they have an aperiodic motion and channels would need to be acquired simultaneously. In a scenario requiring exact registration in all regions despite local aperiodicity, one could use a hybrid method using beam-splitting optics for simultaneous acquisition of the channels, paired with our phase estimation to enhance the frame rate. This would however not solve the limitation expressed above regarding temporal resolution in aperiodically moving regions.

Although we only described the use of fluorescence imaging paired with a brightfield reference, PAAQ could exploit other combination of imaging modalities, provided that it is possible to switch quickly enough between them. The reference signal could be any modality, but one that induces minimal phototoxicity, like brithgfield, is clearly preferable.

7. Conclusion

We have introduced a method to reconstruct virtual high frame rate multichannel videos from slow single-channel acquisition of periodic processes. We have validated our method on experimental data, and shown that it does not contain the anatomical artefacts visible when using a state-of-the-art technique based on mutual information.

PAAQ is based on alternating acquisitions between a common reference modality and fluorescence channels, allowing to reconstruct synchronized videos of multiple modalities without requiring mutual information between channels. We have shown that the synchronization error caused by variability in the signal periodicity is small, even with big frequency variations if the transfer time of the imaging device is small enough.

We have proposed an image-only phase estimation algorithm that improves on existing sorting-based virtual high frame rate methods thanks to a non-uniform sampling assumption. We have shown on both synthetic and experimental data that our algorithm brings an overall relative reduction of the phase error by 30% compared to that of techniques that use a uniform sampling assumption. For synthetic experiments, we designed a simulation framework that reproduces the most relevant properties of cardiac fluorescence microscopy, and can model the heart rate variability with statistical properties

matching experimental observations. We have proposed a reproducible way to compute the optimal hyperparameters for our method that guarantees good performance without manual tuning.

We released the schematics and drivers for a hardware controller that can modulate light sources and synchronize active illumination with imaging devices. Similarly, we released the simulation framework and the implementation of our algorithms.

Although we developed our method for the specific case of cardiac microscopy, our algorithm is generic and we foresee that it could be applied in other scenarios involving the imaging of a cyclic phenomenon with a slow device.

A. Constrained weighted linear least squares expression for distance curve fitting

In Section 3.3.1, we compute the parameters $\tilde{\mathbf{d}}_n$ of the estimated distance curves by solving a linear least squares problem. It is expressed as follows:

$$\left(\mathbf{X}_n^T \mathbf{W}_n \mathbf{X}_n\right) \tilde{\mathbf{d}}_n = \mathbf{X}_n^T \mathbf{W}_n \mathbf{y}_n, \tag{A.1}$$

where the matrices are defined as:

$$\mathbf{y}_n[N_g i + j] = \mathbf{D}[i, j] \tag{A.2}$$

$$\mathbf{W}_n[N_g i + j, N_g i + j] = w_\sigma\left(\tilde{\theta}_i \ominus \tilde{\theta}_n\right) \tag{A.3}$$

$$\mathbf{X}_n[N_g i + j, l] = \begin{cases} l + 2 - \Delta_{ij} & \text{if } l+1 \le \Delta_{ij} < l+2 \\ \Delta_{ij} - l & \text{if } l \le \Delta_{ij} < l+1 \\ 0 & \text{otherwise}, \end{cases} \tag{A.4}$$

with $i = 0, \ldots, N_g$, $j = 0, \ldots, N_g$ and $l = 0, \ldots, L-2$, and

$$\Delta_{ij} = L \frac{\tilde{\theta}_j - \tilde{\theta}_i}{2\pi} + kL, \ k \in \mathbb{Z} \ \text{ s.t. } \Delta_{ij} \in [0, L). \tag{A.5}$$

The quadratic programming problem is then:

$$\mathcal{L}_{\text{dist}}(\tilde{\mathbf{d}}_n) = \frac{1}{2}\tilde{\mathbf{d}}_n \mathbf{Q}_n \tilde{\mathbf{d}}_n + \mathbf{c}_n^T \tilde{\mathbf{d}}_n \qquad \text{with } \mathbf{A}_n \tilde{\mathbf{d}}_n \le \mathbf{b}_n \tag{A.6}$$

$$\mathbf{Q}_n = 2\mathbf{X}_n^T \mathbf{W}_n \mathbf{X}_n \qquad\qquad \mathbf{c}_n = -2\mathbf{y}_n^T \mathbf{W}_n \mathbf{X}_n \qquad\qquad \mathbf{b}_n = \mathbf{0}, \tag{A.7}$$

with the monotonicity constraint matrix:

$$\mathbf{A}_n = \begin{bmatrix} -1 & & & & & & & \\ 1 & -1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & -1 & & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & -1 & 1 & & \\ & & & & & \ddots & & \\ & & & & & & -1 & 1 \\ & & & & & & & -1 \end{bmatrix} \tag{A.8}$$

of size $(L \times L - 1)$, where the change of sign on the diagonal (highlighted) happens after line $L_n^{\max}$.

B. Derivation of the upper bound for the heart rate variability error

In Section 3.5, we analyse the phase estimation error caused by heart rate variability using its upper bound. To derive it, we express the phase delay between a reference frame and its associated signal frame as the product of the time delay between the images Δ_C and the instantaneous frequency of the heartbeat during that delay $\omega_{c,n}$. We use $\epsilon_{c,n}$ to describe the deviation between this instantaneous frequency and its average: $\omega_{c,n} = \bar{\omega} + \epsilon_{c,n}$. In order to measure only the error caused by heart rate variability, we imagine that our phase estimation algorithm finds the exact phases of each reference frame: $\tilde{\theta}_{c,n} = \theta_{c,n}^g$.

We can then compute the error $\mathcal{E}_\epsilon$ that results from the variability in the heart rate using Eq. (2):

$$\mathcal{E}_\epsilon = \min_\Theta \frac{1}{NC} \sum_{c=0}^{C-1} \sum_{n=0}^{N-1} \left| \theta_{c,n}^g + (\bar{\omega} + \epsilon_{c,n})\Delta_C \ominus \left(\theta_{c,n}^g + \Theta \right) \right|. \tag{A.9}$$

From the definition of the phase operator in Eq. (3), it stems that $|\alpha \ominus \beta| \le |\alpha - \beta|$, and we can find an upper bound for our error:

$$\mathcal{E}_\epsilon \le \min_\Theta \frac{1}{NC} \sum_{c=0}^{C-1} \sum_{n=0}^{N-1} \left| \theta_{c,n}^g + (\bar{\omega} + \epsilon_{c,n})\Delta_C - \left(\theta_{c,n}^g + \Theta \right) \right| \tag{A.10}$$

$$\le \min_\Theta \frac{1}{NC} \sum_{c=0}^{C-1} \sum_{n=0}^{N-1} \left| \bar{\omega}\Delta_C - \Theta + \epsilon_{c,n}\Delta_C \right|. \tag{A.11}$$

Given that $\epsilon_{c,n}$ is sampled from a distribution with a zero mean, the right term is minimized when $\Theta = \bar{\omega}\Delta_C$, leaving the final expression of the upper bound:

$$\mathcal{E}_\epsilon \le \frac{1}{NC} \sum_{c=0}^{C-1} \sum_{n=0}^{N-1} \left| \epsilon_{c,n}\Delta_C \right| = \frac{\Delta_C}{NC} \sum_{c=0}^{C-1} \sum_{n=0}^{N-1} \left| \epsilon_{c,n} \right|. \tag{A.12}$$

C. Hyperparameters selection method

In Section 3.6, we highlighted that selecting optimal values for the hyperparameters is crucial to guarantee the performance of the phase estimation algorighm.

The width of the Gaussian window σ used in Eqs. (17) and (18) influences precision significantly. Indeed, a window too large will decrease precision due to too much averaging, while a window too narrow will be too sensitive to noise due to not including enough measurements. To find good value for σ, we plot the residual error obtained when minimizing Eq. (17) with the initial phase estimates, using different window widths up to 2π. We select the last local minimum in that curve as the best value for σ. This represents the best fitting performance, ignoring the minima obtained with very small σ that correspond to overfitting the noise in the data.

The number of points L used to represent the distance curves has an impact on both accuracy and performance. The precision of the method tends to increase with L, as the curve approximation gets finer, but the convergence time of the method increases as well due to the added complexity. Generally

speaking, we have found that our algorithm fails to converge when L is too big (typically bigger than $2N_g$), and that using more than $L = 200$ does not seem to improve performance further.

The update step size of the gradient descent optimizer does not impact the precision much (thanks to the adaptive scaling of the Adam optimizer), but choosing a good value can speed convergence up. We have found that a default value of 10^{-3} works well in all scenarios.

Competing Interests. The authors declare no competing interests associated with this manuscript.

Data Availability Statement. The data used in this paper are publicly available: https://www.idiap.ch/en/dataset/paaq-heart. Replication code can be found in the Idiap CBI Toolbox library: https://github.com/idiap/cbi_toolbox.

Ethical Standards. Zebrafish breeding and procedures were carried out within the scope of an institutional protocol approved by the Veterinary Office of the Canton of Bern, Switzerland (fluorescent) and the Institutional Animal Care and Use Committee at the University of California, Santa Barbara (wild type).

Author Contributions. Conceptualization, methodology: F.M., M.L.; Data curation, investigation, software, visualization: F.M.; Resources: A.E., N.M.; Supervision: M.L.; Writing: F.M., A.E., N.M., M.L.; All authors approved the final submitted draft.

Supplementary Material. Supplementary figures and videos are supplied with the paper.

References

1. Abdolhoseini, M., Kluge, M. G., Walker, F. R., and Johnson, S. J. (2019). Neuron image synthesizer via Gaussian mixture model and Perlin noise. In *IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 530–533.
2. Bakkers, J. (2011). Zebrafish as a model to study cardiac development and human cardiac disease. *Cardiovascular research*, 91(2):279–288.
3. Barufaldi, B., Abbey, C. K., Lago, M. A., Vent, T. L., Acciavatti, R. J., Bakic, P. R., and Maidment, A. D. A. (2021). Computational breast anatomy simulation using multi-scale Perlin noise. *IEEE Transactions on Medical Imaging*, 40(12):3436–3445.
4. Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., Vander-Plas, J., Wanderman-Milne, S., and Zhang, Q. (2018). JAX: composable transformations of Python+NumPy programs, http://github.com/google/jax.
5. Brau, A. C., Wheeler, C. T., Hedlund, L. W., and Johnson, G. A. (2002). Fiber-optic stethoscope: a cardiac monitoring and gating system for magnetic resonance microscopy. *Magnetic Resonance in Medicine*, 47(2):314–321.
6. Bussmann, J., Bakkers, J., and Schulte-Merker, S. (2007). Early endocardial morphogenesis requires Scl/Tal1. *PLOS Genetics*, 3(8):1–13.
7. Cai, H., Wang, Y. L., Wainner, R. T., Iftimia, N. V., Gabel, C. V., and Chung, S. H. (2019). Wedge prism approach for simultaneous multichannel microscopy. *Scientific reports*, 9(1):17795.
8. Chico, T. J., Ingham, P. W., and Crossman, D. C. (2008). Modeling cardiovascular disease in the zebrafish. *Trends in Cardiovascular Medicine*, 18(4):150–155.
9. Dustler, M., Bakic, P., Petersson, H., Timberg, P., Tingberg, A., and Zackrisson, S. (2015). Application of the fractal Perlin noise algorithm for the generation of simulated breast tissue. In *Medical Imaging 2015: Physics of Medical Imaging*, volume 9412, page 94123E. International Society for Optics and Photonics, SPIE.
10. Dustler, M., Förnvik, H., and Lång, K. (2018). Binary implementation of fractal Perlin noise to simulate fibroglandular breast tissue. In *Medical Imaging 2018: Physics of Medical Imaging*, volume 10573, page 1057357. International Society for Optics and Photonics, SPIE.
11. Edelstein, A., Amodaj, N., Hoover, K., Vale, R., and Stuurman, N. (2010). Computer control of microscopes using μManager. *Current Protocols in Molecular Biology*, 92(1):14.20.1–14.20.17.
12. Golub, G. H. and Van Loan, C. F. (2013). *Matrix computations*. JHU press.
13. Gregor, I., Butkevich, E., Enderlein, J., and Mojiri, S. (2021). Instant three-color multiplane fluorescence microscopy. *Biophysical Reports*, 1(1):100001.
14. Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34(4):561–580.
15. Hanslovsky, P., Bogovic, J. A., and Saalfeld, S. (2016). Image-based correction of continuous and discontinuous non-planar axial distortion in serial section microscopy. *Bioinformatics*, 33(9):1379–1386.

16. Huang, C.-J., Tu, C.-T., Hsiao, C.-D., Hsieh, F.-J., and Tsai, H.-J. (2003). Germ-line transmission of a myocardium-specific GFP transgene reveals critical regulatory elements in the cardiac myosin light chain 2 promoter of zebrafish. *Developmental Dynamics*, 228(1):30–40.

17. Huntley, J. M. and Saldner, H. (1993). Temporal phase-unwrapping algorithm for automated interferogram analysis. *Appl. Opt.*, 32(17):3047–3052.

18. Jenkins, M. W., Chughtai, O. Q., Basavanhally, A. N., Watanabe, M., and Rollins, A. M. (2007). In vivo gated 4D imaging of the embryonic heart using optical coherence tomography. *Journal of Biomedical Optics*, 12(3):030505.

19. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., and Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Developmental Dynamics*, 203(3):253–310.

20. Kingma, D. P. and Ba, J. (2015). Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR 2015)*.

21. Lawson, N. D. and Weinstein, B. M. (2002). In vivo imaging of embryonic vascular development using transgenic zebrafish. *Developmental Biology*, 248(2):307–318.

22. Lichtman, J. W. and Conchello, J.-A. (2005). Fluorescence microscopy. *Nature methods*, 2(12):910–919.

23. Liebling, M., Forouhar, A. S., Gharib, M., Fraser, S. E., and Dickinson, M. E. (2005). Four-dimensional cardiac imaging in living embryos via postacquisition synchronization of nongated slice sequences. *Journal of biomedical optics*, 10(5):054001.

24. Liebling, M. and Ranganathan, H. (2009). Wavelet domain mutual information synchronization of multi-modal cardiac microscopy image sequences. In *Wavelets: Applications in Signal and Image Processing XIII*, page 744602, San Diego CA, USA. SPIE.

25. Liebling, M., Vermot, J., Forouhar, A. S., Gharib, M., Dickinson, M. E., and Fraser, S. E. (2006). Nonuniform temporal alignment of slice sequences for four-dimensional imaging of cyclically deforming embryonic structures. In *Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging: Macro to Nano (ISBI 2006)*, pages 1156–1159, Arlington, VA, USA.

26. Mariani, O., Chan, K., Ernst, A., Mercader, N., and Liebling, M. (2019). Virtual high-framerate microscopy of the beating heart via sorting of still images. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 312–315.

27. Mariani, O., Ernst, A., Mercader, N., and Liebling, M. (2020). Reconstruction of image sequences from ungated and scanning-aberrated laser scanning microscopy images of the beating heart. *IEEE Transactions on Computational Imaging*, 6:385–395.

28. Mariani, O., Marelli, F., Jaques, C., Ernst, A., and Liebling, M. (2021). Unequivocal cardiac phase sorting from alternating ramp-and pulse-illuminated microscopy image sequences. In *IEEE 18th International Symposium on Biomedical Imaging (ISBI 2021)*, pages 868–872.

29. Mickoleit, M., Schmid, B., Weber, M., Fahrbach, F. O., Hombach, S., Reischauer, S., and Huisken, J. (2014). High-resolution reconstruction of the beating zebrafish heart. *Nature methods*, 11(9):919–922.

30. Nguyen, C. T., Lu, Q., Wang, Y., and Chen, J.-N. (2008). Zebrafish as a model for cardiovascular development and disease. *Drug Discovery Today: Disease Models*, 5(3):135–140. Cardiovascular disease models.

31. O'Donoghue, B. (2021). Operator splitting for a homogeneous embedding of the linear complementarity problem. *SIAM Journal on Optimization*, 31:1999–2023.

32. Ohn, J. and Liebling, M. (2011). In vivo, high-throughput imaging for functional characterization of the embryonic zebrafish heart. In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1549–1552.

33. Ohn, J., Yang, J., Fraser, S. E., Lansford, R., and Liebling, M. (2011). High-speed multicolor microscopy of repeating dynamic processes. *genesis*, 49(7):514–521.

34. Perlin, K. (1985). An image synthesizer. *SIGGRAPH Comput. Graph.*, 19(3):287–296.

35. Perlin, K. (2002). Improving noise. In *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '02, pages 681–682, New York, NY, USA. Association for Computing Machinery.

36. Pitrone, P. G., Schindelin, J., Stuyvenberg, L., Preibisch, S., Weber, M., Eliceiri, K. W., Huisken, J., and Tomancak, P. (2013). OpenSPIM: an open-access light-sheet microscopy platform. *Nature methods*, 10(7):598–599.

37. Rohr, S., Bit-Avragim, N., and Abdelilah-Seyfried, S. (2006). Heart and soul/PRKCi and nagie oko/Mpp5 regulate myocardial coherence and remodeling during cardiac morphogenesis. *Development*, 133(1):107–115.

38. Rohr, S., Otten, C., and Abdelilah-Seyfried, S. (2008). Asymmetric involution of the myocardial field drives heart tube formation in zebrafish. *Circulation Research*, 102(2):e12–e19.

39. Sablong, R., Rengle, A., Ramgolam, A., Saint-Jalmes, H., and Beuf, O. (2014). An optical fiber-based gating device for prospective mouse cardiac MRI. *IEEE Transactions on Biomedical Engineering*, 61(1):162–170.

40. Sanderson, M. J., Smith, I., Parker, I., and Bootman, M. D. (2014). Fluorescence microscopy. *Cold Spring Harbor Protocols*, 2014(10):1042–1065.

41. Scherz, P. J., Huisken, J., Sahai-Hernandez, P., and Stainier, D. Y. R. (2008). High-speed imaging of developing heart valves reveals interplay of morphogenesis and function. *Development*, 135(6):1179–1187.

42. Stainier, D. Y. and Fishman, M. C. (1994). The zebrafish as a model system to study cardiovascular development. *Trends in Cardiovascular Medicine*, 4(5):207–212.

43.   Taylor, J. M. (2014). Optically gated beating-heart imaging. *Frontiers in Physiology*, 5.

44.   Taylor, J. M., Girkin, J. M., and Love, G. D. (2012). High-resolution 3D optical microscopy inside the beating zebrafish heart using prospective optical gating. *Biomed. Opt. Express*, 3(12):3043–3053.

45.   Taylor, J. M., Saunter, C. D., Love, G. D., Girkin, J. M., Henderson, D. J., and Chaudhry, B. (2011). Real-time optical gating for three-dimensional beating heart imaging. *Journal of Biomedical Optics*, 16(11):116021.

46.   Vonesch, C., Aguet, F., Vonesch, J.-L., and Unser, M. (2006). The colored revolution of bioimaging. *IEEE Signal Processing Magazine*, 23(3):20–31.

47.   Zhang, Q. and Pless, R. (2005). Segmenting cardiopulmonary images using manifold learning with level sets. In *Computer Vision for Biomedical Image Applications*, pages 479–488. Springer Berlin Heidelberg.