

Research Paper

Periodic Temporal Super Resolution Based on Phase Registration and Manifold Reconstruction

YASUSHI MAKIHARA,^{†1} ATSUSHI MORI^{†1}
and YASUSHI YAGI^{†1}

This paper describes a method for periodic temporal super resolution, namely, reconstructing a one period image sequence with high frame-rate from a single quasi-periodic image sequence with low frame-rate. First, the periodic image sequence to be reconstructed is expressed as a manifold in the parametric eigenspace of the phase, namely, period-normalized time. Given an input image sequence, phase registration data in sub-frame order among multiple periods of the image sequence is estimated. The phase registration and manifold reconstruction are alternately executed iteratively within an energy minimization framework that considers data fitness and the smoothness of both the manifold and the phase evolution. The energy minimization problem is solved through three-step coarse-to-fine procedures to avoid local minima. The proposed periodic temporal super resolution is evaluated through the experiments using both simulated and real data in terms of phase noise, the number of input frames, frame-rate, spatial registration noise, and image noise, respectively.

1. Introduction

Image super resolution²⁸⁾ is one of the fundamental low level vision techniques for creating an image with high spatial resolution from images with low spatial resolution. Methods for image super resolution fall mainly into two categories: (1) reconstruction-based methods (RBM) using multiple images registered at sub-pixel order^{5),10),20),26)}, and (2) example-based methods using correspondence between low and high resolution image patches from training sets^{2),8),12)}. Furthermore, Glasner et al.⁹⁾ proposed a sophisticated framework, combining the two approaches by using the patch recurrence within and across scales of a single image. Most of the existing methods, however, are applicable to static scenes only.

Several methods deal with dynamic scenes in the context of near-real-time super resolution¹⁶⁾, taking advantage of hot-air optical turbulence¹⁵⁾ and image super resolution in the presence of motion deblurring^{4),22)}. Nevertheless, these methods still focus on spatial super resolution, and fail to address temporal super resolution.

Contrary to the above methods, Shechtman et al.²³⁾ proposed a space-time super resolution by combining information from multiple low-resolution video sequences of the same dynamic scene. Moreover, Agrawal et al.¹⁾ proposed an optimal multiplexing coded sampling for temporal super-resolution. These methods fall into the so-called RBMs and hence, require multiple video cameras to obtain multiple image sequences.

Inspired by the effective use of patch recurrence⁹⁾, it is noticeable that temporal recurrence in a single sequence can be used for temporal super resolution. In particular, given a periodic image sequence, multiple periods of subsequences with sub-frame order *phase displacement* can be used for the temporal super resolution. If a sequence can be segmented into period-based subsequences composed of more than a few frames, a spatio-temporal sequence registration⁷⁾ method is applicable. On the contrary, in cases where each subsequence contains only a few frames or a single frame, period segmentation and the spatio-temporal sequence registration may be error prone, since the method discards useful cues of inter-period phase continuity. Moreover, if the period fluctuates due to a non-uniform sampling rate or to target motion fluctuation appearing in human periodic actions (e.g., gait), phase registration becomes much more difficult.

Therefore, we propose a method to reconstruct a one period image sequence at a high frame-rate from a single low frame-rate quasi-periodic image sequence. Instead of segmenting the whole sequence into multiple subsequences, we solve the *phase registration* problem for the whole sequence of frames to take advantage of the inter-period phase continuity. Then, a one period image sequence is expressed as a *manifold in phase-parametric eigenspace* and the manifold is reconstructed based on the registered phase information. These two processes are repeated in an energy minimization framework by taking into consideration the following three aspects: (1) data fitness, (2) smoothness of the manifold, and (3) smoothness of the phase evolution.

^{†1} Osaka University

The remainder of this paper is organized as follows. Section 2 summarizes related works, and Section 3 describes a problem setting for this work. Section 4 addresses an energy minimization framework to solve phase registration and manifold reconstruction iteratively. Section 5 presents the experiments of periodic temporal super resolution, and shows the performance against phase noise, the number of input images, frame-rate, spatial registration noise, and image noise, respectively. Section 6 presents discussions on limitation of the proposed method, and finally, Section 7 concludes the paper.

2. Related Work

Temporal interpolation: In addition to spatial interpolation, temporal interpolation has been developed for video editing (e.g., retiming²¹⁾) and video compression based on motion compensation methods²⁹⁾. In particular, morphing approaches^{3),18),25)} based on optical flow¹³⁾ are regarded as some of the most promising methods for temporal interpolation, with these ideas having been extended to view and time interpolation²⁴⁾ and spatio-temporal interpolation from two sequences, one with high resolution and low frame rate and the other with low resolution and high frame rate³⁰⁾. Temporal interpolation does not, however, work well in cases where motion between two frames is relatively large, in other words, where the sampling interval is relatively long in fairly low frame-rate video.

Geometric fitting: Our manifold reconstruction step is similar to the fitting of geometric primitives, such as lines, ellipses, or conics^{11),27)}, in terms of parameter estimation from a series of sampling points. In geometric fitting, global parameters are used for the primitive representation (e.g., 5 parameters for an ellipse) and sampling points are treated independently. On the other hand, piecewise local parameters are used for manifold representation and sampling points are dependent on each other due to the phase evolution smoothness as described in the following sections.

3. Problem Setting

3.1 Assumptions

Several assumptions are used in this paper to allow us to focus only on temporal

super resolution based on a phase registration framework. The first assumption is that an image sequence is spatially registered in advance. This assumption is basically true in several scenes, such as those with periodic sign language such as a waving hand, periodic actions at the same position such as jumping jacks or walking on a treadmill, or a rotating object such as a fan. Even in other scenes with periodic actions such as walking, skipping, and running on the ground, spatial registration is possible once the target object has been accurately tracked to some extent.

A further assumption is that motion blurs are negligible. Although this assumption may not be true in scenes with fast moving objects, there are still several possible situations in which low frame-rate image sequences with less motion blur are stored. For example, consider the situation where a CCTV camera captures a person walking on the street. In this case, the image sequence is often stored at a low frame-rate due to limited communication band width or storage size, and besides motion blur is not significant since walking motion is relatively slow compared with normal shutter speeds. Therefore, gait recognition in low frame-rate video is one of the typical applications of the proposed method.

3.2 Quasi-periodic Image Sequence

Next, we define a quasi-periodic image sequence in conjunction with **Fig. 1**. An image drawn from the periodic image sequence at time t is denoted by the vector-form $\mathbf{x}(t)$, which satisfies

$$\mathbf{x}(t + P) = \mathbf{x}(t) \quad \forall t, \quad (1)$$

where P is a period. Then, two non-dimensional time parameters, *phase* s and *relative phase* \tilde{s} , are introduced as

$$s = s_P(t) = \frac{t}{P} \quad (2)$$

$$\tilde{s} = s - \lfloor s \rfloor, \quad (3)$$

where $s_P(\cdot)$ is a phase evolution function and $\lfloor \cdot \rfloor$ is a floor function. Now, the periodic image sequence is represented in the phase domain as

$$\mathbf{x}_s(s) = \mathbf{x}(s_P^{-1}(s)). \quad (4)$$

Note that the periodic image sequence constructs a manifold with respect to the relative phase $\tilde{s} \in [0, 1]$ and which satisfies $\mathbf{x}_s(1) = \mathbf{x}_s(0)$.

On the other hand, an input image sequence is composed of N^{in} discretely

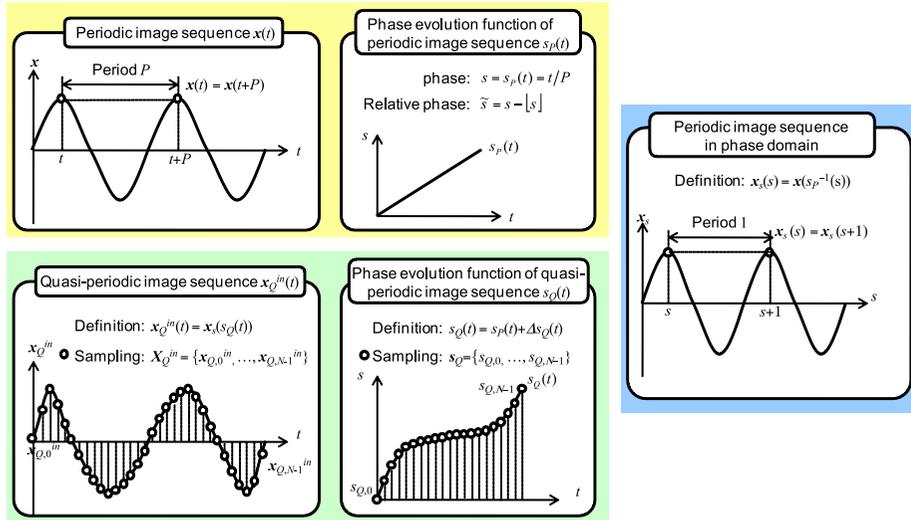


Fig. 1 Process of quasi-periodic image sequence.

observed images $\mathbf{X}^{in} = \{\mathbf{x}_i^{in}\} (i = 0, \dots, N^{in} - 1)$. In cases where the scene is completely periodic and where the frame-rate f is completely constant during the input image sequence, the input image sequence \mathbf{X}^{in} , time sequence $\mathbf{t} = \{t_i\}$, and phase sequence $\mathbf{s}_P = \{s_{P,i}\}$ are denoted as

$$\mathbf{x}_i^{in} = \mathbf{x}(t_{P,i}) = \mathbf{x}_s(s_{P,i}) \quad (5)$$

$$t_{P,i} = t_0 + \frac{i}{f} \quad (6)$$

$$s_{P,i} = s_P(t_{P,i}) = s_0 + \frac{i}{fP}. \quad (7)$$

This assumption is, however, often violated due to fluctuations in the frame-rate or in the timing of the periodic motion. Hence, the input image sequence is degraded to a *quasi-periodic image sequence* $\mathbf{X}_Q^{in} = \{\mathbf{x}_{Q,i}^{in}\}$ defined as

$$\mathbf{x}_{Q,i}^{in} = \mathbf{x}_s(s_{Q,i}) \quad (8)$$

$$s_{Q,i} = s_{P,i} + \Delta s_i, \quad (9)$$

where $\mathbf{s}_Q = \{s_{Q,i}\}$ is a quasi-periodic phase sequence.

In summary, the problem setting can be stated as a simultaneous estimation

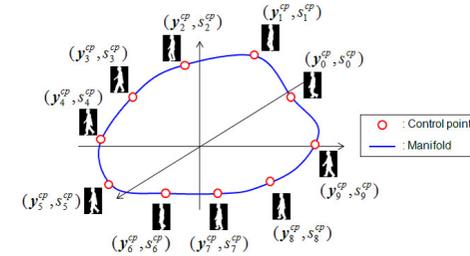


Fig. 2 Manifold representation of periodic image sequence.

problem of a periodic manifold \mathbf{x}_s and a phase sequence \mathbf{s}_Q from an input quasi-periodic image sequence \mathbf{X}_Q^{in} . Analogous to the spatial super resolution case, note that the phase sequence estimation process and periodic manifold estimation correspond to an image registration process and super-resolution process, respectively. In addition, the problem setting falls into the so-called reconstruction-based methods using multiple observations, although a single sequence is used in this problem setting. This is because *multiple periods* in the single sequence serve as multiple observations.

3.3 Manifold Representation

The periodic manifold \mathbf{x}_s is represented by a parametric eigenspace method¹⁷⁾. This is motivated by the property of the parametric eigenspace method that a phase for an input image can be estimated by projection to the manifold, which plays quite a significant role in the phase registration stage. More specifically, we use a cubic N-spline function parameterized by the phase in the eigenspace. Let us consider N^{cp} control points $\{\mathbf{y}_j^{cp}\}$ in the M dimensional eigenspace accompanied by corresponding phases $\{s_j^{cp} (= j/N^{cp})\}$, ($j = 0, \dots, N^{cp} - 1$) as shown in **Fig. 2**. Next, a spline parameter vector for a k power-term coefficient at the j th interval $[s_j^{cp}, s_{j+1}^{cp}]$ is denoted as $\mathbf{a}_{j,k}^{sp} \in \mathbb{R}^M$, ($k = 0, 1, 2, 3$ for a cubic N-spline) and subsequently, a submatrix A_j^{sp} at the j th interval and a total spline matrix A^{sp} are defined as $A_j^{sp} = [\mathbf{a}_{j,0}^{sp}, \dots, \mathbf{a}_{j,3}^{sp}]^T \in \mathbb{R}^{4 \times M}$ and $A^{sp} = [A_0^{spT}, \dots, A_{N^{cp}-1}^{spT}]^T \in \mathbb{R}^{4N^{cp} \times M}$, respectively. Then, an interpolated point $\hat{\mathbf{y}}(\tilde{s})$ in the eigenspace for a relative phase \tilde{s} at the j th interval is expressed as

$$\hat{\mathbf{y}}(\tilde{s}) = A^{spT} \mathbf{w}(\tilde{s}) \quad (10)$$

$$\mathbf{w}(\tilde{s}) = [0, \dots, 0, 1, w, w^2, w^3, 0, \dots, 0]^T \quad (11)$$

$$w = \frac{\tilde{s} - s_j^{cp}}{s_{j+1}^{cp} - s_j^{cp}}, \quad (s_j^{cp} \leq \tilde{s} \leq s_{j+1}^{cp}), \quad (12)$$

where $\mathbf{w}(\tilde{s})$ is an interpolation coefficient vector whose components from $4j$ to $(4j + 3)$ are $[1, w, w^2, w^3]$, and w is the interpolation ratio between the control points.

On the other hand, the relation between a control points matrix $Y^{cp} = [\mathbf{y}_0^{cp}, \dots, \mathbf{y}_{N^{cp}-1}^{cp}]^T$ and a spline parameter matrix A^{sp} is derived from the C2-continuous boundary conditions as

$$CA^{sp} = DY^{cp} \quad (13)$$

$$C = \begin{bmatrix} C_1 & C_2 & O & \cdots \\ \ddots & \ddots & \ddots & \ddots \\ C_2 & O & \cdots & C_1 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 6 \end{bmatrix}, \quad (14)$$

$$C_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 0 \end{bmatrix} \quad (14)$$

$$D = \begin{bmatrix} D_1 & D_2 & O & \cdots \\ \ddots & \ddots & \ddots & \ddots \\ D_2 & O & \cdots & D_1 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (15)$$

Hence, the spline parameter matrix A^{sp} is linearly solved as $A^{sp} = C^{-1}DY^{cp}$ given the control points Y^{cp} . This also indicates that interpolation $\hat{\mathbf{y}}(\tilde{s})$ for a relative phase \tilde{s} is obtained by Eqs. (10) and (13) once the control points Y^{cp} have been given. Then, the reconstruction problem of the periodic manifold $\mathbf{x}_s(s; Y^{cp})$ can be replaced by an estimation problem of the control points Y^{cp} as discussed in the following sections.

Moreover, once a manifold is reconstructed, high frame-rate periodic image sequence is generated by back projection of the manifold in the eigenspace. Because

we typically project back the manifold control points in the eigenspace, we can regard the number of the control points N^{cp} as one of the parameters to adjust the reconstructed frame rate.

4. Energy Minimization Framework

4.1 Energy Function

Similar to the spatial super resolution case, we adopt an energy minimization approach. First, suppose that an input quasi-periodic image sequence in the eigenspace is expressed as $\mathbf{Y}_Q^{in} = \{\mathbf{y}_{Q,i}^{in}\}$ and recall that the accompanying phase sequence $\mathbf{s}_Q = \{s_{Q,i}\}$ is unknown. Subsequently the interpolation coefficient vector for the i th phase $s_{Q,i}$ is defined as $\mathbf{w}(s_{Q,i})$, in the same way as Eq. (11), and then projection onto the periodic manifold by the i th phase $s_{Q,i}$ is

$$\hat{\mathbf{Y}}(Y^{cp}, s_{Q,i}) = A^{spT} \mathbf{w}(s_{Q,i}) = Y^{cpT} (C^{-1}D)^T \mathbf{w}(s_{Q,i}). \quad (16)$$

The energy function is constructed by considering the following three aspects: (1) data fitness between the interpolation $\hat{\mathbf{y}}(Y^{cp}, s_{Q,i})$ and the input $\mathbf{y}_{Q,i}^{in}$, (2) smoothness of the periodic manifold $\mathbf{y}_s(s; Y^{cp})$ in the eigenspace, and (3) smoothness of the phase evolution \mathbf{s}_Q .

The manifold smoothness term is defined as an integral of squared manifold curvatures and it suppresses abrupt changes on the manifold. Note that the cubic N-spline itself only guarantees piece-wise C2-continuity within each interval and hence it cannot guarantee global smoothness of the manifold. For example, let's assume that a control point (y_0^{cp}, s_0^{cp}) is swapped for another control point (y_4^{cp}, s_4^{cp}) in Fig. 2. The cubic N-spline just reconstruct a piece-wise smooth manifold for the swapped control points and does not give any penalties for large curvatures around the swapped control points. On the other hand, the manifold smoothness term gives penalties for such large curvatures around the swapped control points, and hence it enforces global smoothness of the reconstructed manifold.

The phase evolution smoothness term is regarded as a kind of soft constraint on the phase relationship that the phase difference between two adjacent frames is equal to an inverse of a global period P' in the frame domain, which is defined as the product of the frame-rate f and period P in the time domain, that is,

$P' = fP$. Note that the order of an input image sequence has meaning in the phase evolution smoothness term due to the above phase relationship. On the other hand, without the phase evolution smoothness term, phases for each frame $\mathbf{s}_Q = \{s_{Q,i}\}$ appear only in the data term and hence they are estimated independently of each other. In such a case, because order information of the input image sequence is not exploited, phase estimation from the image sequence with order information is equivalent to phase estimation from a set of images without order information, and it may result in unstable phase estimation. Although some researchers tackle such problems, namely, reconstruction of a periodic image sequence from a set of images without the order information, it was reported that the reconstructed periodic image sequences sometimes contain errors such as shortcuts³¹⁾. Therefore, the phase evolution smoothness term plays an important role to suppress such phase estimation errors.

Consequently, the actual form of the energy function is

$$\begin{aligned}
E(Y^{cp}, \mathbf{s}_Q) &= \frac{1}{N^{in}} \sum_{i=0}^{N^{in}-1} \|Y^{cpT}(C^{-1}D)^T \mathbf{w}(s_{Q,i}) - \mathbf{y}_{Q,i}^{in}\|^2 \\
&+ \lambda_m \frac{1}{N^{cp}} \int_0^1 \left\| \frac{d^2 \mathbf{y}_s(s; Y^{cp})}{ds^2} \right\|^2 ds \\
&+ \lambda_s \frac{1}{N^{in}} \sum_{i=1}^{N^{in}-1} \left(s_{Q,i+1} - s_{Q,i} - \frac{1}{P'} \right)^2, \quad (17)
\end{aligned}$$

where the first, second, and third terms are the data term, and the smoothness terms for the periodic manifold and the phase evolution, respectively. Since the integration in the second term is calculated in advance with respect to the domain of the interpolation ratio w for each interval, it is rearranged as a quadratic form of Y^{cp} as

$$\int_0^1 \left\| \frac{d^2 \mathbf{y}_s(s; Y^{cp})}{ds^2} \right\|^2 ds = Y^{cpT}(C^{-1}D)^T B(C^{-1}D)Y^{cp} \quad (18)$$

$$B = \begin{bmatrix} B_{sub} & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & B_{sub} \end{bmatrix}, \quad B_{sub} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 6 \\ 0 & 0 & 6 & 12 \end{bmatrix} \quad (19)$$

We can see that the objective function $E(Y^{cp}, \mathbf{s}_Q)$ is a quadratic form with respect to the manifold control points Y^{cp} and therefore, that a linear solution of the manifold control points Y^{cp} is provided under the fixed phase \mathbf{s}_Q . On the other hand, the phase \mathbf{s}_Q is a complex form, since the spline curves are switched piecewise based on phase \mathbf{s}_Q and the interpolation ratio w appears as a sixth-order polynomial in the data term. To solve this highly nonlinear optimization problem, three-step coarse-to-fine solutions are provided in the following sections.

4.2 Solution by Linear Approximation

In the first step, the phase sequence \mathbf{s}_Q is limited to a completely periodic domain, that is, linear phase evolution \mathbf{s}_P . Because the initial phase s_0 is not significant in this problem setting, it is set to zero without loss of generality. Hence, the objective function in the first step is rewritten as

$$E_{init}(P) = E(Y_P^{cp}, \mathbf{s}_P) \quad (20)$$

$$Y_P^{cp} = \arg \min_{Y^{cp}} E(Y^{cp}, \mathbf{s}_P) \quad (21)$$

$$s_{P,i} = \frac{i}{fP}. \quad (22)$$

Note that manifold control points Y_P^{cp} for each hypothesis of period P are linearly solved as described in Section 4.1. On the other hand, the objective function $E_{init}(P)$ has many local minima, so a gradient descent method from multiple initial solutions is applied. Optimal solutions of the period, phase and manifold control points in the periodic-domain are denoted as P^* , \mathbf{s}_{P^*} , and $Y_{P^*}^{cp}$, respectively.

4.3 Dynamic Programming Solution

In the second step, in order to extend from a periodic domain to a quasi-periodic domain, continuous Dynamic Programming (DP)¹⁹⁾ is applied within a so-called corridor, that is, the neighboring search area from the periodic-domain phase solution \mathbf{s}_{P^*} as $R_i^{cdr} = \{s | s_{P^*,i} - s^{cdr} \leq s \leq s_{P^*,i} + s^{cdr}\}$ under fixed

manifold control points $Y_{P^*}^{cp}$. First, the phase space is quantized by the same interval as that of the manifold control points ($1/N^{cp}$). This means that the j th-phase interpolation coincides with the j th control point $\mathbf{y}_{P^*,j}^{cp}$ and that the phase difference between the j th and k th phases is equal to $(j-k)/N^{cp}$. Next, a cumulative cost and the optimal transition from a previous step at the i th input frame and j th phase are denoted as $c(i, j)$ and $p(i, j)$, respectively, and the optimal phase is provided by the following DP steps.

(1) Initialize cost matrix

$$c(0, j) = \|\mathbf{y}_{P^*,j}^{cp} - \mathbf{y}_0^{in}\|^2, \quad \forall j \in R_0^{cdr} \quad (23)$$

(2) Update cumulative cost and transition path

$$p(i, j) = \arg \min_{k \in R_{i-1}^{cdr}} \{c(i-1, k) + \lambda_s g_{P^*}(j, k)\} \quad (24)$$

$$c(i, j) = c(i-1, p(i, j)) + \lambda_s g_{P^*}(j, p(i, j)) + \|\mathbf{y}_{P^*,j}^{cp} - \mathbf{y}_i^{in}\|^2, \quad \forall j \in R_i^{cdr} \quad (25)$$

$$g_{P^*}(j, k) = \left(\frac{\min\{|j-k|, |j-k+N^{cp}|\}}{N^{cp}} - \frac{1}{fP^*} \right)^2 \quad (26)$$

(3) Optimize the terminal phase

$$p^*(N^{in} - 1) = \arg \min_j c(N^{in} - 1, j), \quad j \in R_{N^{in}-1}^{cdr} \quad (27)$$

(4) Back track

$$p^*(i-1) = p(i, p^*(i)), \quad \forall 1 \leq i \leq N^{cp} - 1 \quad (28)$$

Finally, the optimal phase \mathbf{s}_{DP}^* is set based on the optimal path $\{p^*(i)\}$ and the manifold control points are updated as

$$Y_{DP}^{cp} = \arg \min_{Y^{cp}} E(Y^{cp}, \mathbf{s}_{DP}^*) \quad (29)$$

Strictly speaking, one can bypass the DP procedure and proceed to the following step where there is a sufficiently small fluctuation in phase from the periodic solution. On the other hand, in cases with a substantial fluctuation in phase, the following step may result in local optima only and therefore, the DP procedure is still essential for finding the global optimum.

4.4 Iterative Solution by Quadratic Approximation

A crucial procedure in the third step is quadratic approximation of the data

term with respect to the phase within a narrow search range from the DP solution. First, the DP solution, \mathbf{s}_{DP}^* and Y_{DP}^{cp} , is set as the initial solution for this step, \mathbf{s}_Q^0 and $Y^{cp,0}$, respectively. Then, the narrow search region of the phase at the r th iteration is set as $R_i^r = \{s | s_{Q,i}^{r-1} - s^{tol} \leq s \leq s_{Q,i}^{r-1} + s^{tol}\}$. The data term minimum for the i th phase $s_{Q,i}$ is found by the Newton method within the intervals that include at least a part of the search region R_i^r . After the phase minimizing the data term has been obtained as $s_{Q,i}^{data,r*}$, the i th data term in Eq. (17) is approximated by a Taylor expansion up to the second-order terms as

$$\begin{aligned} \hat{E}_i^{data,r}(s_{Q,i}) &= E_i^{data,r}(s_{Q,i}^{data,r*}) + \left. \frac{dE_i^{data,r}}{ds_{Q,i}} \right|_{s_{Q,i}=s_{Q,i}^{data,r*}} (s_{Q,i} - s_{Q,i}^{data,r*}) \\ &\quad + \frac{1}{2} \left. \frac{d^2E_i^{data,r}}{ds_{Q,i}^2} \right|_{s_{Q,i}=s_{Q,i}^{data,r*}} (s_{Q,i} - s_{Q,i}^{data,r*})^2. \end{aligned} \quad (30)$$

Now, the total energy function is a quadratic form with respect to the phase \mathbf{s}_Q and thus the optimal phase \mathbf{s}_Q is given as

$$\mathbf{s}_Q^{r*} = \arg \min_{\mathbf{s}_Q} \left\{ \frac{1}{N^{in}} \sum_{i=0}^{N^{in}-1} \hat{E}_i^{data,r}(s_{Q,i}) + \lambda_s \frac{1}{N^{in}} \sum_{i=1}^{N^{in}-1} \left(s_{Q,i+1} - s_{Q,i} - \frac{1}{fP^*} \right)^2 \right\} \quad (31)$$

$$\text{s.t. } s_{Q,i}^{r-1} - s^{tol} \leq s_{Q,i} \leq s_{Q,i}^{r-1} + s^{tol} \quad (32)$$

$$s_{Q,i+1} \geq s_{Q,i}, \quad (33)$$

where Eqs. (32) and (33) are the lower and upper limit constraints and the monotonically increasing constraints, respectively. As a result, the problem is formulated as a convex quadratic programming one and is solved by the active set method. Then the manifold control points at the r th iteration are updated in the same way as in the previous steps.

$$Y^{cp,r} = \arg \min_{Y^{cp}} E(Y^{cp}, \mathbf{s}_Q^{r*}) \quad (34)$$

These procedures are iterated until convergence by gradually relaxing the manifold smoothness constraint so that the manifold fits the data.

4.5 Connection to Spatial Super Resolution

Conventional spatial super resolution typically contains two steps: acquisition

of image registration data and spatial super resolution based on the image registration data. Considering relationship between the proposed periodic temporal super resolution and the spatial super resolution, while phase estimation step in the periodic temporal super resolution corresponds to image registration step in the spatial super resolution, manifold reconstruction in the periodic temporal super resolution corresponds to reconstruction of a super-resoluted image in the spatial super resolution. More specifically, the manifold fitness data term and the manifold smoothness term, namely, the first and the second terms in Eq. (17) in the periodic temporal super resolution, correspond to so-called data term and smoothness term in the spatial super resolution, respectively.

5. Experiments

5.1 Experimental Setup

In these experiments, the proposed temporal super resolution method is applied to low frame-rate quasi-periodic image sequences. We used two different types of image sequences: the first is CG data of a conical pendulum viewed from an oblique direction (**Fig. 3** (top)), while the other is real data in the form of silhouette and color textured image sequences of a person walking on a treadmill (**Fig. 3** (middle, bottom)). Regarding the real data, size normalization and spatial registration are done in advance. Moreover, images comprising one period are manually extracted as a subsequence and a completely periodic image sequence is constructed by repeating this subsequence. The image sizes, original frame rates, and periods of the two types of image sequences are listed in **Table 1**.

Then, low frame-rate periodic image sequences are down sampled from the

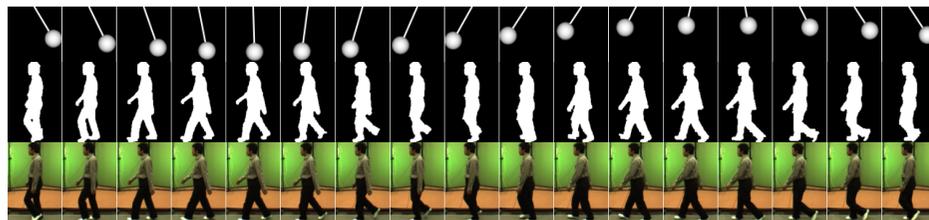


Fig. 3 Periodic image sequences of conical pendulum (top), gait silhouette (middle), and gait color textured image (bottom) used in the experiments.

original image sequences. Note that the original high frame-rate image sequences were naturally not used in the temporal super resolution process at all and that but used as ground truth data in quantitative evaluation in the following section. In addition, quasi-periodic image sequences are produced by randomly changing the sampling interval up to a predefined phase noise level from the linear phase evolution. **Figure 4** shows examples of quasi-periodic image sequences of the conical pendulum at 6 fps and 1 fps, respectively.

Finally, the parameters for the energy minimization framework are determined experimentally as follows. PCA is exploited for eigenspace projection and the information loss rate is set to 1%, in other words, the cumulative contribution ratios of the eigen values is 99%. The number of manifold control points is 100, and the search range for the dynamic programming s^{cdr} and quadratic approximation s^{tol} are 0.25 and 0.02, respectively. The smoothness term coefficients λ_z and λ_s are set to 50.0 and 1.0, respectively, with λ_z reduced by half, down to a minimum of 1.0, in the iterative process of quadratic approximation.

5.2 Refinement Process

In this subsection, to explain the refinement process throughout the three steps, i.e., Linear Approximation (LA), Dynamic Programming (DP), and Quadratic Approximation (QA), we at first focus on a specific example of the conical pendulum image sequence, with frame-rate, number of frames, and phase noise set to 3 fps, 67 frames, and 20%, respectively.

First, phase estimation errors in the Ground Truth (GT) are shown in

Table 1 Image sequence properties.

Image sequence	Image size [pixel]	Original frame-rate [fps]	Period [sec]
Conical pendulum	64×64	- (Arbitrary)	1.17
Gait	88×128	60	1.17

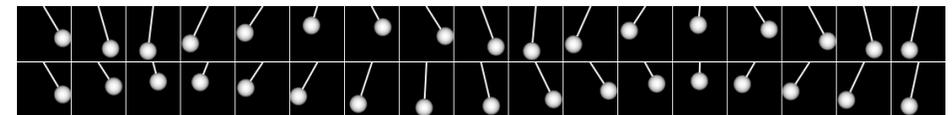


Fig. 4 Examples of input quasi-periodic conical pendulum image sequences with 10% phase noise at 6 fps (top) and 1 fps (bottom).

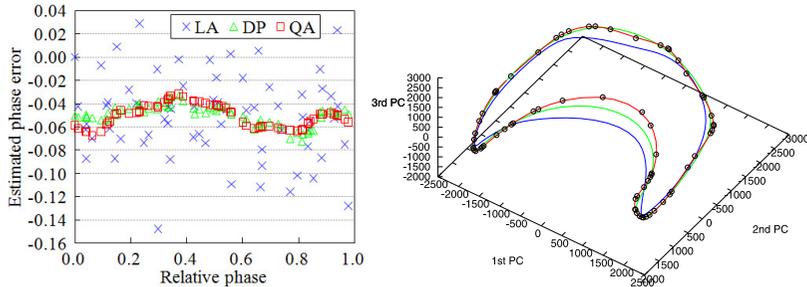


Fig. 5 Phase estimation error on relative phase (left) and manifold in three principal components of eigenspace (right).

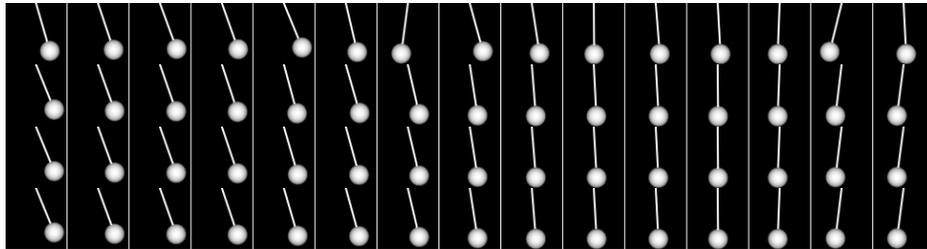


Fig. 6 A section of the conical pendulum image sequences sorted by estimated relative phase. Rows from the top to the bottom indicate LA, DP, QA, and GT, respectively.

Fig. 5(a). While the estimated phases in LA deviate widely, since LA cannot absorb non-linear phase noise, those in DP and QA converge within a small range and their deviation is approximately periodic. Although phase error biases from GT (approx. -0.05) are observed in DP and QA, they are not necessarily significant since phase registration up to a relative relation among all the input frames is sufficient for temporal super resolution. In other words, only the standard deviation of the phase estimation errors should be evaluated.

Next, to demonstrate the impact of the phase estimation errors in a more visual way, image sequences sorted by estimated relative phases are shown in **Fig. 6**. Based on these results, we can see several significant “phase backs” in LA and hence, a sufficiently large manifold smoothness coefficient λ_z is essential for the LA step, otherwise a manifold disturbed by the phase backs is reconstructed. On



Fig. 7 A section of the gait color textured image sequences. The first row: Input low frame-rate image sequence (3 fps). The second to fourth rows: Super resolved image sequence by back projection of the manifold control points by LA, DP, and QA, respectively.

the other hand, relative phase orders of DP and QA are almost consistent with that of the GT and hence, the smoothness coefficient λ_z can be relaxed.

Subsequently, manifolds in the three principal components of the eigenspace are shown in **Fig. 5(b)**. We can see that the manifolds’ fitness to the input points improves throughout the three-step phase estimation refinement. In particular, the fitness of QA is the best, since the smoothness coefficient λ_z is relaxed in the iteration process.

Although our energy minimization framework does not guarantee convergence in theory, it does not induce any divergences throughout our experiments. In practice, at most 10 iterations are sufficient for obtaining satisfactory accuracy of manifold and phase registration in QA step.

Finally, to demonstrate the effectiveness of the propose method, super resolved image sequences for the gait color textured image sequences are generated by back projection of manifold control points as shown in **Fig. 7**. As a result, we can see that the reconstructed image sequences become clearer throughout the three-step refinement and that sufficient visual quality is obtained in QA step (**Fig. 7** (bottom)).

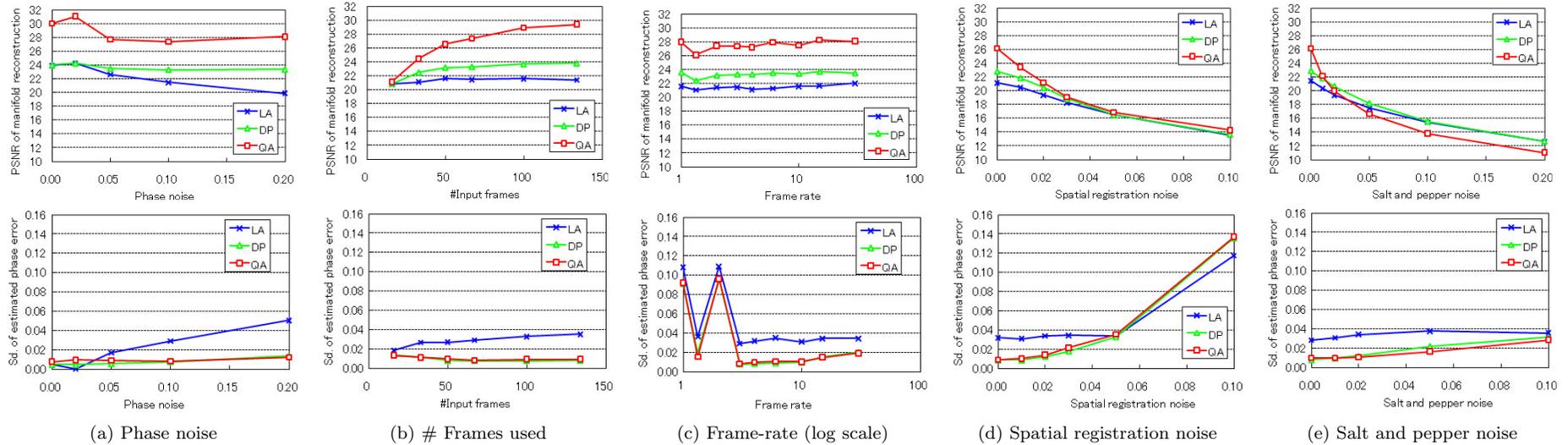


Fig. 8 Performance evaluation of manifold PSNR (top row) and phase error SD (bottom row) using conical pendulum image sequence. Each column denotes each factor: (a) phase noise, (b) the number of frames used, (c) frame-rate, (d) spatial registration noise, and (e) salt and pepper noise. The vertical axes in the top and bottom rows indicate the PSNR between the reconstructed manifold and ground truth and the SD of the phase estimation error, respectively.

5.3 Evaluation

In this section, manifold reconstruction errors and phase estimation errors are evaluated. First, as for manifold reconstruction, a single period of ground truth images are initially projected to a manifold reconstructed from a low frame-rate quasi-periodic input image sequence and then they are back projected to the image domain to calculate Peak Signal to Noise Ratio (PSNR) between the ground truth images and reprojected images. As for phase estimation, errors between the ground truth phase and estimated phases of the input image sequence are calculated and then their Standard Deviation (SD) is evaluated, since their biases are not significant as discussed in the previous section.

The above two items are evaluated with respect to five factors: (1) phase noise, (2) the number of frames used, (3) frame-rate, (4) spatial registration noise, and (5) image noise (salt and pepper noise) as shown in **Fig. 8**. The baseline settings for the three factors are 10%, 67 frames, 3 fps, 0%, and 0%, respectively.

As for the phase noise (Fig. 8 (a)), although LA achieves the lowest phase error SD in the case of lower phase noise, both the manifold PSNR and the phase error SD in LA deteriorate in proportion to an increase in phase noise. On the contrary, such degradations in DP and QA are suppressed within a certain range.

Regarding the number of frames used (Fig. 8 (b)), the performance naturally improves as the number of frames increases in DP and QA. On the other hand, the performance in LP does not improve in proportion to the frames used. This is because LP does not solve non-linear phase noise, and besides, an increase in the phase-error input points in the eigenspace makes the manifold reconstruction worse.

Regarding frame-rate (Fig. 8 (c)), the manifold reconstruction PSNRs improve slightly as the frame-rate increases. On the other hand, critical phase errors are observed in an extremely low frame-rate (1 fps and 2 fps) and this is discussed further in the next section.

Regarding spatial registration noise (Fig. 8(d)), the manifold reconstruction PSNRs and phase error SDs significantly degrade as the spatial registration noise. This is because the proposed method assumes that the spatial registration has been done in advance and hence the spatial registration process would be included in the energy minimization framework as well as phase registration process if the input image sequence is not spatially registered.

Regarding image noise (Fig. 8(e)), while the phase error SDs are suppressed within a certain range, the manifold reconstruction PSNRs are degraded as the image noise increases. If this kind of noisy image sequence is given, image denoising process (e.g., bilateral filtering, non-local mean filtering, or graph cut-based denoising) should be done before applying the proposed method.

Moreover, we have made another experiment to check the sensitivity of the proposed method in terms of the two significant parameters: manifold smoothness coefficient λ_z and phase evolution smoothness coefficient λ_s as shown in Fig. 9. As a result, the performances of both manifold reconstruction and phase estimation in QA step are almost unchanged for a wide range of the smoothness coefficients except for a case where the initial manifold smoothness term is extremely small (e.g., $\lambda_z = 1$) or large (e.g., $\lambda_z = 10,000$). Consequently, we can see that the proposed method is insensitive to the smoothness coefficients if only the manifold smoothness coefficient is set within a middle range (e.g., $10 \leq \lambda_z \leq 1,000$).

5.4 Comparison with the Other Methods

In this section, the proposed method is compared with two shape morphing techniques: signed distance field morphing⁶⁾ and Earth Mover's Morphing¹⁴⁾ as temporal interpolation. The morphing techniques only cannot deal with phase noise and then we provide low frame-rate gait image sequence without phase noise, which serve as key shapes for morphing (see the first row of Fig. 10). The other factors are set to the same baseline setting as in the previous section.

The morphing techniques work well for interpolating relatively smooth image sequence, namely, not so low frame-rate image sequence. They, however, suffer from abrupt image changes existing in low frame-rate image sequence (see the third and the fourth rows in Fig. 10).

Moreover, the proposed method is compared with a graph-based periodic mo-

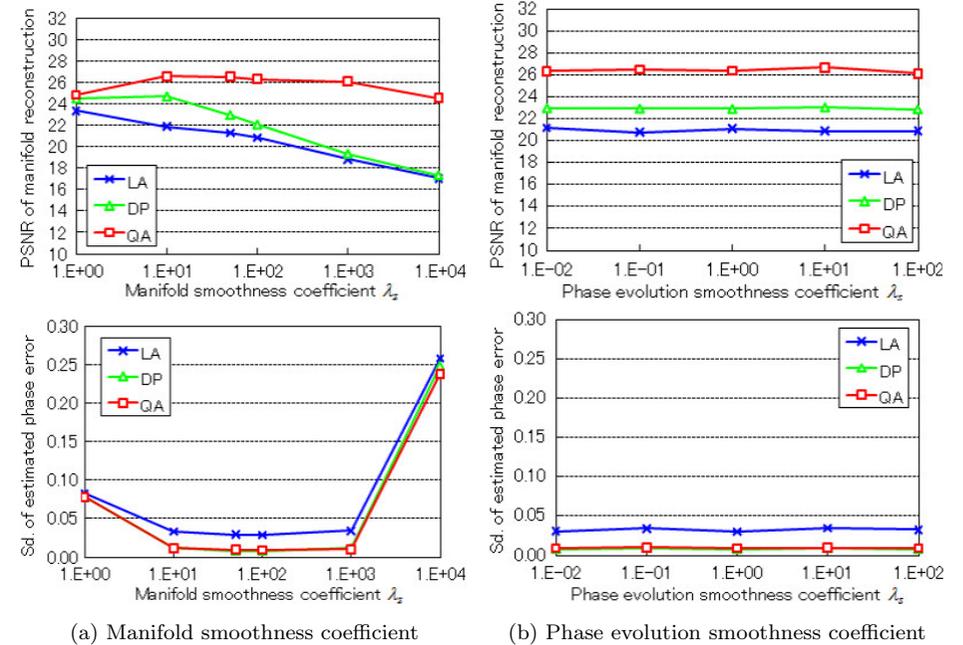


Fig. 9 Sensitivity of manifold smoothness coefficient and phase evolution smoothness coefficient using a gait silhouette sequence. Top: manifold PSNR, bottom: phase error SD.

tion path reconstruction technique³¹⁾. Although the method originally uses a contour-based shape feature, we substitute silhouette-based shape feature because a single closed curve cannot represent silhouette holes between both legs as shown in the first and second frames in Fig. 10.

As a result, while the graph-based technique works well for several parts of a sequence, it suffers from a significant *shortcut* of a forward-step sequence which should originally appear between the fourth and fifth frames in the bottom row of Fig. 10. This type of failure has been already reported in Ref. 31) and one of reasons for this failure is that the method does not consider the order information of an input image sequence when reconstructing a periodic image sequence from a set of images.

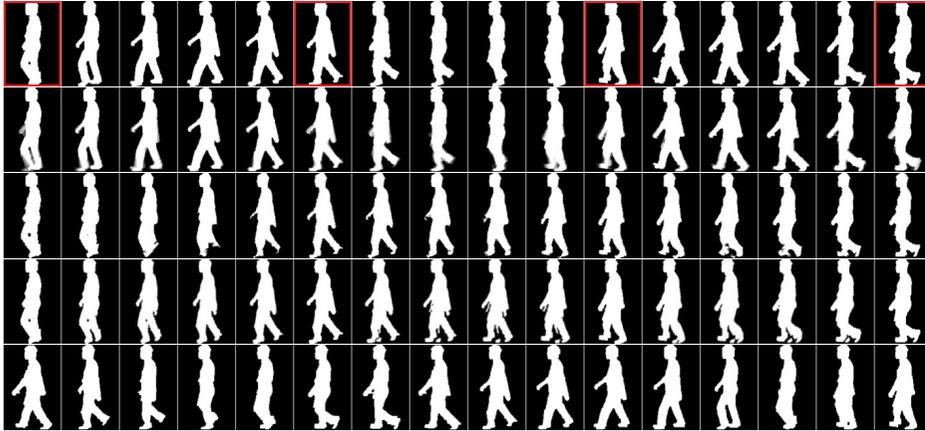


Fig. 10 Comparison with the other methods. Rows from the top to the bottom indicate ground truth, the proposed temporal super resolution, signed distance field morphing⁶⁾, Earth Mover's Morphing¹⁴⁾, and graph-based periodic motion path reconstruction³¹⁾, respectively. Key frames are bounded by red box in ground truth image sequence (top row).

On the other hand, because the proposed method integrates various phase images from multiple periods and also considers phase evolution smoothness together with the order information of an input image sequence, it can reconstruct better one period of gait image as shown in the second row in Fig. 10.

6. Discussion

In this section, prerequisites and limitations of the proposed temporal super resolution are discussed. In cases where a periodic image sequence is observed at a coarse sampling interval compared with its period, we need to consider two main issues: (1) the *stroboscopic effect* (*temporal aliasing*), and (2) the *wagon wheel effect* as reported in Ref. 23).

The *stroboscopic effect* typically occurs when the sampling interval coincides with the period of a moving object. In such cases, the observed image sequence appears to be standing still because the observed images are always the same even though the object is actually moving periodically.

From this observation, we can introduce a theoretical upper bound of the tem-

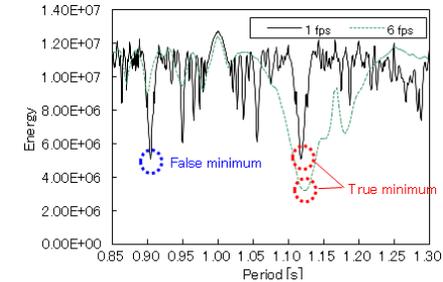


Fig. 11 False period detection by “Wagon wheel effects” in the LA step. The horizontal and vertical axes denote period and energy, respectively. False minimum by backward play (blue broken circle) competes with true minimum by forward play (red broken circle) at 1 fps, while true minimum is obviously the global minimum at 6 fps.

poral super resolution from a single period image sequence. Intuitively speaking, temporal super resolution is impossible when each phase in one period is exactly the same as the corresponding phase in the other periods, that is, the period in the frame domain is not a sub-frame order, but merely a frame order defined as an integer. Let us consider a low frame-rate periodic image sequence and denote its period in the frame domain P' [frame]. Then, assume that the period P' is expressed as a fraction by relative prime numbers, m and n ,

$$P' = \frac{m}{n} \in \mathbb{Q}, m, n \in \mathbb{N}. \quad (35)$$

Now, if $n = 1$, the period P' is degraded from a fraction to an integer, that is, the period is just a frame order and the temporal super resolution is impossible. On the other hand, if $n > 1$, the period P' is a sub-frame order and hence different phase images can be observed among multiple periods and the temporal super resolution is possible. In summary, the low frame-rate image sequence can be up-converted to a n -times frame-rate, if m input frames are given.

The *wagon wheel effect* typically occurs when the sampling interval is slightly smaller than the period of a moving object as shown in the bottom row of Fig. 4, where the sampling interval and period are 1.0 sec and 1.17 sec, respectively. In this case, the false minimum by backward play competes with the true minimum by forward play, and then the false minimum is often adopted in the LA step as shown in **Fig. 11**. This is the reason that the critical phase errors occur in the

extremely low frame-rate in Fig. 8 (c).

7. Conclusion

This paper described a method for temporal super resolution from a single quasi-periodic image sequence. The temporal super resolution was formulated as simultaneous phase registration and reconstruction of a manifold of the periodic image sequence in a phase parametric eigenspace. An energy minimization framework considering data fitness, and the smoothness of both the manifold and the phase evolution, was introduced and solved through three-step coarse-to-fine procedures to avoid local minima. Experiments using synthesized conical pendulum and real gait silhouette and color textured image sequences were conducted to evaluate the effects of phase noise, the number of frames used, frame-rate, spatial registration noise, and salt and pepper noise on the PSNR of the manifold reconstruction and the standard deviation of the phase estimation errors.

In this paper, we focused mainly on phase registration while ignoring spatial registration process and blur effects, and hence, it is necessary, in future, to include the spatial registration and blur effects in the manifold reconstruction framework. Moreover, not only phase fluctuation, but also inter-period image deformation (e.g., view or speed transition in gait scenes) should be considered for real applications.

Acknowledgments This work was supported by Grant-in-Aid for Scientific Research (S) 21220003.

References

- 1) Agrawal, A., Gupta, M., Veeraraghavan, A. and Narasimhan, S.: Optimal Coded Sampling for Temporal Super-Resolution, *23rd IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, San Francisco, CA, USA, pp.599–606 (2010).
- 2) Baker, S. and Kanade, T.: Limits on Super-resolution and How to Break Them, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.24, No.9, pp.1167–1183 (2002).
- 3) Beymer, D. and Poggio, T.: Image Representations for Visual Learning, *Science*, Vol.272, No.28, pp.1905–1909 (1996).
- 4) Blake, A., Basclé, B., and Zisserman, A.: Motion Deblurring and Super-Resolution from an Image Sequence, *Proc. European Conf. Computer Vision*, pp.312–320 (1996).
- 5) Borman, S. and Stevenson, R.: Spatial Resolution Enhancement of Low-Resolution Image Sequences: A Comprehensive Review with Directions for Future Research, Technical report, University of Notre Dame (1998).
- 6) Breen, D. and Whitaker, R.: A Level-set Approach for the Metamorphosis of Solid Models, *IEEE Trans. Visualization and Computer Graphics*, Vol.7, No.2, pp.172–192 (2001).
- 7) Caspi, Y. and Irani, M.: Spatio-Temporal Alignment of Sequences, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.24, No.11, pp.1409–1425 (2002).
- 8) Freeman, W., Jones, T. and Pasztor, E.: Example-Based Super-Resolution, *IEEE Trans. Computer Graphics and Applications*, Vol.22, No.2, pp.56–65 (2002).
- 9) Glasner, D., Bagon, S. and Irani, M.: Super-Resolution from a Single Image, *Proc. 12th Int. Conf. Computer Vision* (2009).
- 10) Irani, M. and Peleg, S.: Improving Resolution by Image Registration, *Computer Vision, Graphics, and Image Processing*, Vol.53, No.3, pp.231–239 (1991).
- 11) Kanatani, K.: Ellipse Fitting with Hyperaccuracy, *IEICE Trans. Inf. Syst.*, Vol.E89-D, No.10, pp.2653–2660 (2006).
- 12) Liu, C., Shum, H. and Zhang, C.: A Two-Step Approach to Hallucinating Faces: Global Parametric Model and Local Non-Parametric Model, *IEEE Int. Conf. Computer Vision and Pattern Recognition*, pp.192–198 (2001).
- 13) Lucas, B. and Kanade, T.: An Iterative Image Registration Technique with an Application to Stereo Vision, *Proc. 7th Int. Joint Conf. Artificial Intelligence*, pp.674–679 (1981).
- 14) Makihara, Y. and Yagi, Y.: Earth Mover’s Morphing: Topology-Free Shape Morphing Using Cluster-based EMD Flows, *Proc. 10th Asian Conf. Computer Vision*, Queenstown, New Zealand (2010) (to appear).
- 15) Shimizu, M., Yoshimura, S., Tanaka, M. and Okutomi, M.: Super-Resolution from Image Sequence under Influence of Hot-Air Optical Turbulence, *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1–8 (2008).
- 16) Tanaka, M. and Okutomi, M.: Near-real-time Video-to-video Super-resolution, *Proc. 8th Asian Conference on Computer Vision (ACCV2007), Demo-A*, p.1 (2007).
- 17) Murase, H. and Nayar, S.K.: Parametric Eigenspace Representation for Visual Learning and Recognition, *Proc. SPIE, 2031* (1993).
- 18) Nagahara, H., Matsunobu, T., Iwai, Y., Yachida, M. and Suzuki, T.: High-resolution Video Generation Using Morphing, *Proc. IAPR Int. Conf. Pattern Recognition (ICPR2006)*, Hong Kong (2006).
- 19) Oka, R.: Spotting method for classification of real world data, *Comput. J.*, Vol.41, No.8, pp.559–565 (1998).
- 20) Protter, M., Elad, M., Takeda, H. and Milanfar, P.: Generalizing the Nonlocal-means to Super-resolution Reconstruction, *IEEE Trans. Image Processing*, Vol.18, No.1, pp.36–51 (2009).

- 21) REALVIZ2: Retimer, available from www.realviz.com/products/rt (2000).
- 22) Sezan, M., Patti, A. and Tekalp, A.: Superresolution Video Reconstruction with Arbitrary Sampling Lattices and Nonzero Aperture Time, *IEEE Trans. Image Processing*, Vol.6, No.8, pp.1064–1076 (1997).
- 23) Shechtman, E., Caspi, Y. and Irani, M.: Space-Time Super-Resolution, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.27, No.4, pp.531–545 (2005).
- 24) Stich, T., Linz, C., Albuquerque, G. and Magnor, M.: View and Time Interpolation in Image Space, *Computer Graphics Forum (Proc. Pacific Graphics)*, Vol.27, No.7, pp.1781–1787 (2008).
- 25) Stich, T. and Magnor, M.: Image morphing for space-time interpolation, *SIGGRAPH '07: ACM SIGGRAPH 2007 sketches*, New York, NY, USA, ACM, p.87 (2007).
- 26) Tanaka, M. and Okutomi, M.: A Fast MAP-Based Super-Resolution Algorithm for General Motion, *Proc. SPIE-IS & T Electronic Imaging 2006, Computational Imaging IV, 6065, 1B*, pp.1–12 (2006).
- 27) Taubin, G.: Estimation of Planar Curves, Surfaces, and Nonplanar Space Curves Defined by Implicit Equations with Applications to Edge and Range Image Segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.13, No.11, pp.1115–1138 (1991).
- 28) van Ouwerkerk, J.: Image super-resolution survey, *Image and Vision Computing*, Vol.24, No.10, pp.1039–1052 (2006).
- 29) Wang, Y., Ostermann, J. and Zhang, Y.-Q.: *Video Processing and Communications*, Prentice Hall (2002).
- 30) Watanabe, K., Iwai, Y., Nagahara, H., Yachida, M. and Suzuki, T.: Video Synthesis with High Spatio-Temporal Resolution Using Spectral Fusion, *Proc. Int. Workshop Multimedia Content Representation, Classification and Security*, Istanbul, Turkey, pp.683–690 (2006).
- 31) Xu, X., Wan, L., Liu, X., Wong, T.-T., Wang, L. and Leung, C.-S.: Animating animal motion from still, *ACM Trans. Graphics*, Vol.27, No.5, pp.117:1–8 (2008).

(Received November 7, 2010)

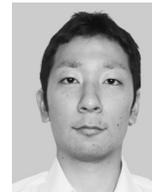
(Accepted May 12, 2011)

(Released December 28, 2011)

(Communicated by Cha Zhang)



Yasushi Makihara received his B.S., M.S., and Ph.D. degrees in engineering from Osaka University in 2001, 2002, and 2005, respectively. He is currently an Assistant Professor of the Institute of Scientific and Industrial Research, Osaka University. His research interests are gait recognition, morphing, and temporal super resolution. He is a member of IPSJ, RJS, and JSME.



Atsushi Mori received his B.S. and M.S. degrees in information science from Osaka University in 2007 and 2009. His research interest is gait recognition.



Yasushi Yagi is the Professor of Intelligent Media and Computer Science, and the Assistant Director of the Institute of Scientific and Industrial Research, Osaka university, Ibaraki, Japan. He received his Ph.D. degrees from Osaka University in 1991. In 1985, he joined the Product Development Laboratory, Mitsubishi Electric Corporation, where he worked on robotics and inspections. He became a Research Associate in 1990, a Lecturer in 1993, an Associate Professor in 1996, and a Professor in 2003 at Osaka University. International conferences for which he has served as Chair include: FG1998 (Financial Chair), OMINVIS2003 (Organizing chair), ROBIO2006 (Program co-chair), ACCV2007 (Program chair), PSVIT2009 (Financial chair), ICRA2009 (Technical Visit Chair) and ACCV2009 (General chair). He has also served as the Editor of IEEE ICRA Conference Editorial Board (2007, 2008). He is the Editor-in-Chief of IPSJ Transactions on Computer Vision & Image Media and the Associate Editor-in-Chief of IPSJ Transactions on Computer Vision & Applications. He was awarded ACM VRST2003 Honorable Mention Award, IEEE ROBIO2006 Finalist of T.J. Tan Best Paper in Robotics, IEEE ICRA2008 Finalist for Best Vision Paper, MIRU2008 Nagao Award. His research interests are computer vision, medical engineering and robotics. He is a member of IEICE, RSJ, and IEEE.
