Visual Stabilization of Beating Heart Motion by Model-Based Transformation of Image Sequences

Evgeniya Bogatyrenko and Uwe D. Hanebeck

Abstract— In order to assist a surgeon while operating on a beating heart, visual stabilization makes the beating heart appear still by providing the current heart view as stationary and non-moving. In this way, the surgeon is not disturbed during an operation by the motion of the heart and has the impression of performing conventional surgery. In contrast to existing methods for visual stabilization, the proposed approach involves a model-based transformation of image sequences provided by a camera system. This transformation incorporates the knowledge of physical characteristics of the heart in form of a mathematical model of the heart surface. The main advantage of this transformation is that the uncertainties of the model and measurements are considered. This occurs by estimating the parameters of the transformation. Furthermore, the quality of the visual stabilization is additionally improved by adapting the parameters of the underlying physical model. The performance of the proposed approach is evaluated in an experiment with a pressure-regulated artificial heart. In comparison to standard approaches, it provides superior results illustrating the high quality of the visual stabilization.

I. Introduction

Beating heart surgery may allow high-risk patients to undergo bypass surgery. In comparison with conventional surgery, it is advantageous because of a significantly lower mortality rate. However, beating heart surgery is more challenging for a surgeon. For synchronizing surgical instruments with the heart surface, the surgeon has to superpose the desired motion with the beating heart motion. This demands a high concentration, short reaction time, and accurate hand-eye coordination from the surgeon. To overcome these difficulties, a computer-assisted surgical system will facilitate operating on the beating heart [1]. When the surgical robot takes care of the synchronization, the provided visual stabilization will enable the presentation of the beating heart surface as if it were stationary and non-moving. In this way, the surgeon will not be disturbed during an operation by the motion of the heart and has the impression of performing conventional surgery.

When prediction of the heart motion for the synchronization of surgical instruments is investigated [1]–[4], only few methods are proposed for the visual stabilization of the beating heart. In [5], the use of electrocardiogram-triggered strobed light for making the heart appear still to a surgeon is explored. The presented experimental results show that this method is not able to compensate the motion of the heart. The application of strobed light leads to increasing demands on the surgeon's concentration and to fatigue. The

approach to visual stabilization proposed in [6] simulates the stabilized view of the heart surface by changing the extrinsic camera parameters. This study shows that the approach performs well if a very small and smooth area of the heart surface is observed. However, it fails in the presence of local distortions. In [7], the visual stabilization is based on mapping of the images acquired at different time steps. For this purpose, a linear interpolation is applied. As can be easily recognized, this method demands a fine discretization for achieving a high quality of the stabilization. Furthermore, the visual stabilization may become rough [8] since no smoothness constraints regarding the heart surface motion are considered.

In summary, the quality of the visual stabilization provided by all these approaches can deteriorate in case of inaccurate measurements. The main reason for this is that uncertainties are ignored. Furthermore, for the processing of measurement data, these approaches do not involve models of both physical phenomena and the measurement process, and neglect a priori information. As a result, the measurement disturbances, such as stochastic uncertainties caused by camera noise, environmental disturbances like smoke due to tissue cutting, and inaccuracies of feature extraction cannot be removed. Ignoring these disturbances can reduce the quality of the visual stabilization.

The specialty of the method proposed in this paper is the visual stabilization of the beating heart motion using a *model-based transformation* of image sequences. This method has the following three particular characteristics:

1) incorporating knowledge of the underlying physical characteristics of the beating heart in form of a mathematical model along with a measurement model,
2) considering the model uncertainties as well as accompanying measurement noise directly in the visual stabilization, and
3) improving the visual stabilization by parameter adaptation of the underlying mathematical model.

II. Problem Formulation

This section provides a brief introduction to the problem of visual stabilization, which is formulated as a transformation of image sequences. As common in computer-assisted surgery, we assume that an elastically deformable object such as the surface of the beating heart is observed by a multi-camera system. Every camera provides an image sequence containing information about the deformation of the object. We define every image as a set of quadratic pixels with the assigned color values. Every pixel is identified by two indices. For example, as illustrated in Fig. 1, the pixel p_k^i is identified by the indices $p_k^{x,i}$ and $p_k^{y,i}$ at time step t_k. The
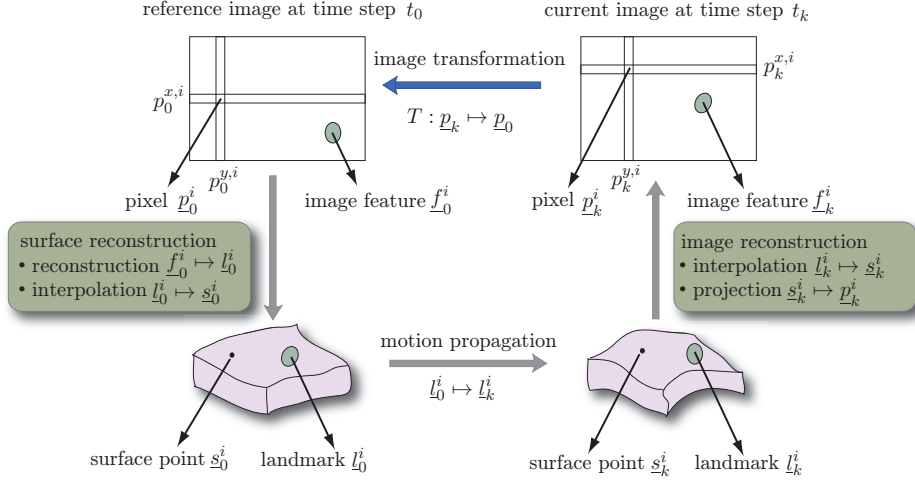
Fig. 1. Visual stabilization of the elastically deformable object incorporating *model-based image transformation*. This transformation is defined by a transformation function T that provides the mapping between the pixel indices of the reference $\underline{p}_0$ and current $\underline{p}_k$ images. These images represent the camera projections of the moving object at two different time steps t_0 and t_k. For better clarity, only a landmark $\underline{l}_0^i$ is placed on the surface of the object.

indices of all pixels are collected in the vector

$$\underline{p}_k := \left[\underline{p}_k^{x\,\mathrm{T}}, \underline{p}_k^{y\,\mathrm{T}}\right]^{\mathrm{T}} .$$

For visual stabilization, every image of the image sequence should be transformed to a certain reference image with the aim to compensate the shift of the pixels, i.e., shift of their indices. In other words, this transformation is described by a image transformation function T that maps the pixels of the current image to the pixels of the reference image. The pixels of the reference image are described by the vector of indices

$$\underline{p}_0 := \left[\underline{p}_0^{x\,\mathrm{T}}, \underline{p}_0^{y\,\mathrm{T}}\right]^{\mathrm{T}} .$$

The definition of this transformation function deserves careful attention. We are convinced that it should incorporate not only geometric but also physical constraints on the heart motion for achieving highly accurate visual stabilization. Therefore, the parameters of the image transformation function should depend on the physical characteristics of the heart. Furthermore, the uncertainties of the camera measurements and inaccuracies of the transformation function should be considered. Most existing image transformation methods, such as warping [8], [9] or morphing [10] ignore the uncertainties by assuming that the measurements and the transformation function are exactly known. However, the real-world measurements are corrupted by disturbances. Also the transformation function may not exactly reproduce the real transformation and involves stochastic and systematic uncertainties. Ignoring these errors can lead to an inaccurate visual stabilization.

III. KEY IDEA

The key idea of the proposed approach is the design of a three-dimensional model-based image transformation. As shown in Fig. 1, the current and the reference images represent the camera projections of the moving heart surface at two different time steps t_0 and t_k. Therefore, the shift of the indices of the pixels $\underline{p}_k \mapsto \underline{p}_0$ can be described by the displacements of the heart surface. For example, in this

figure, the shift of the pixel $\underline{p}_0^i$ in the reference image to the pixel $\underline{p}_k^i$ in the current image is determined by the changing position of the corresponding heart surface point from $\underline{s}_0^i$ to $\underline{s}_k^i$. Unfortunately, the measurements of the heart surface displacements are available only at certain measurement points, which are represented by the landmarks, whose three-dimensional coordinates are collected in the vector

$$\underline{l}_k := \left[\underline{l}_k^{x\,\mathrm{T}}, \underline{l}_k^{y\,\mathrm{T}}, \underline{l}_k^{z\,\mathrm{T}}\right]^{\mathrm{T}} . \tag{1}$$

In this paper, for achieving a high accuracy of measurements, artificial landmarks, such as circular markers, are attached to the heart surface.

For estimating the displacements of all surface points, their three-dimensional coordinates are assembled in the vector

$$\underline{s}_0 := \left[\underline{s}_0^{x\,\mathrm{T}}, \underline{s}_0^{y\,\mathrm{T}}, \underline{s}_0^{z\,\mathrm{T}}\right]^{\mathrm{T}} . \tag{2}$$

We exploit a physical heart surface model, which is initialized based on the three-dimensional reference positions of the landmarks, described by the vector

$$\underline{l}_0 := \left[\underline{l}_0^{x\,\mathrm{T}}, \underline{l}_0^{y\,\mathrm{T}}, \underline{l}_0^{z\,\mathrm{T}}\right]^{\mathrm{T}} . \tag{3}$$

As illustrated in Fig. 1 with the example of landmark l_0^i, the positions of the landmarks are reconstructed from two-dimensional positions of the image features, assembled in the vector

$$\underline{f}_0 := \left[\underline{f}_0^{x\,\mathrm{T}}, \underline{f}_0^{y\,\mathrm{T}}\right]^{\mathrm{T}} . \tag{4}$$

In this figure, only a single landmark is depicted to simplify the explanation.

When the heart deforms, the model represents the current deformations of the heart surface and thus, provides estimates of current three-dimensional positions of all surface points

$$\underline{s}_k := \left[\underline{s}_k^{x\,\mathrm{T}}, \underline{s}_k^{y\,\mathrm{T}}, \underline{s}_k^{z\,\mathrm{T}}\right]^{\mathrm{T}} . \tag{5}$$

The positions of these points are determined by interpolating between the current positions of the landmarks $\underline{l}_k \mapsto \underline{s}_k$. Finally, by projecting the deformed model to the current image $\underline{s}_k \mapsto \underline{p}_k$, the correspondences between the pixel indices and also positions of the image features in the reference and current images are established. Given this
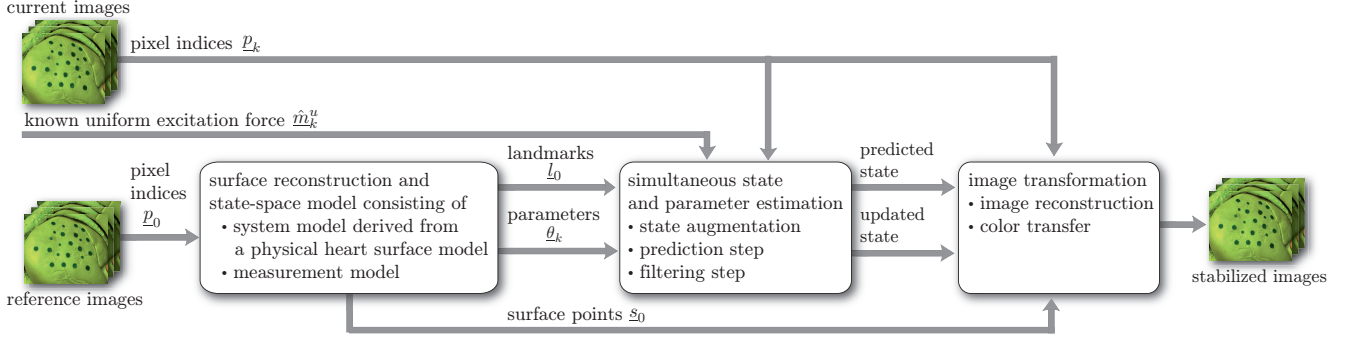
Fig. 2. Model-based transformation of the current images to the stabilized images incorporating the state-space model, which considers the physical characteristics of the heart surface. The simultaneous state and parameter estimation allows for considering measurement and model uncertainties and adapting model parameters.

mapping, stabilized images can be obtained by transforming the color information of the current pixels to reference pixels.

For improving the visual stabilization, the differences between the estimated positions of the image features, collected in the vector

$$\underline{f}_k := \left[\underline{f}_k^{x\mathrm{T}}, \underline{f}_k^{y\mathrm{T}} \right]^{\mathrm{T}} \tag{6}$$

and the corresponding measured positions

$$\hat{\underline{f}}_k := \left[\hat{\underline{f}}_k^{x\,\mathrm{T}}, \hat{\underline{f}}_k^{y\,\mathrm{T}} \right]^{\mathrm{T}} \tag{7}$$

are traced back and used to adapt the underlying physical model and thus, the image transformation.

IV. VISUAL STABILIZATION

The purpose of this section is the detailed explanation of the proposed approach for the visual stabilization, which is depicted in Fig. 2. First, the image transformation function T is derived in Section IV-A. Then, the underlying state-space model incorporating the physical heart surface model is introduced in Section IV-B. Finally, the parameters of the physical model are estimated using a simultaneous state and parameter estimation, which will be described in Section IV-C.

A. Image Transformation Function

By recalling that the image transformation T is defined by the projecting the observed object to the camera images at the reference time step and current time step, this function can be written as

$$\begin{aligned} \underline{p}_k^x - \underline{p}_x^0 &= \frac{r_{12}\underline{u}_k^x + r_{12}\underline{u}_k^y + r_{12}\underline{u}_k^z + r_{14}}{r_{31}\underline{u}_k^x + r_{32}\underline{u}_k^y + r_{33}\underline{u}_k^z + r_{34}} , \\ \underline{p}_k^y - \underline{p}_y^0 &= \frac{r_{21}\underline{u}_k^x + r_{22}\underline{u}_k^y + r_{23}\underline{u}_k^z + r_{24}}{r_{31}\underline{u}_k^x + r_{32}\underline{u}_k^y + r_{33}\underline{u}_k^z + r_{34}}, \end{aligned} \tag{8}$$

where the r_{ij}, $i = 1, \ldots, 3$, $j = 1, \ldots, 4$ represents the elements of the projection matrix $\mathbf{P}$ provided by a calibration of the cameras [11]. The vector

$$\underline{u}_k = \left[\underline{u}_k^{x\mathrm{T}}, \underline{u}_k^{y\mathrm{T}}, \underline{u}_k^{z\mathrm{T}} \right]^{\mathrm{T}}$$

collects the three-dimensional displacements of the heart surface. By using (8), we can establish the correspondences between the pixels of the current image and reference image. Finally, the color information of the current image can be written to the reference image.

However, since the camera measurements are available only at discrete points and corrupted by uncertainties, the displacements of the heart surface between these two time steps should be estimated.

B. Model-Based Description

For estimating the heart surface displacement, a state-space model consisting of system and measurement equations, also called *system and measurement models*, will be derived in this section.

The specialty of the proposed method is that here the *system model* incorporates knowledge of the underlying physical characteristics of the heart. The reason for this is that the system model is derived from a mathematical model that describes the heart surface as a physical elastic body. It is important to note that only the intervention area is considered. Because of small size of this area, the assumption of an isotropic material structure can be made. Furthermore, this area is stabilized by a mechanical stabilizer during a heart operation. The stabilizer damps the heart motion so that the heart surface deformations are small [12]. This allows to assume a linear elastic behavior of the heart surface. Moreover, the model makes use of the fact that the force generated by the myocardium can be represented as force volume relation or pressure as stated in medical experimental studies [13]. Therefore, based on the linear constitutive law, the heart motion is approximately characterized by a dynamic distributed parameter system in form of partial differential equations, according to [14]. Applying the meshless collocation method [15] and using the implicit Euler method for temporal discretization, the heart surface model is converted to a discrete-time system equation in state-space form

$$\underline{x}_{k+1} = \mathbf{A}_k(\underline{\psi}_k)\,\underline{x}_k + \mathbf{B}_k(\underline{\psi}_k)\,(\hat{\underline{m}}_k^u + \underline{w}_k) \ , \tag{9}$$

where the vector $\hat{\underline{m}}_k^u = [\hat{\underline{m}}_k^{u,x\mathrm{T}}, \hat{\underline{m}}_k^{u,y\mathrm{T}}, \hat{\underline{m}}_k^{u,z\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{3M}$ describes the known model excitation, which is assumed to be uniformly distributed. The modeling errors denoted by $\underline{w}_k$ are assumed to be white zero-mean Gaussian $\underline{w}_k \sim \mathcal{N}(\underline{0}, \mathbf{C}_k^w)$ with covariance matrix $\mathbf{C}_k^w \in \mathbb{R}^{6N \times 6N}$. The matrix $\mathbf{A}_k \in \mathbb{R}^{6N \times 6N}$ represents the system matrix and the matrix $\mathbf{B}_k \in \mathbb{R}^{6N \times 3M}$ denotes the input matrix. These matrices depend on the physical parameters of the heart surface

$$\underline{\psi}_k = \left[\boldsymbol{\rho}_k, \boldsymbol{\eta}_k^1, \boldsymbol{\eta}_k^2, \boldsymbol{E}_k, \boldsymbol{\nu}_k \right]^{\mathrm{T}},$$

such as material density $\boldsymbol{\rho}_k$, Rayleigh damping coefficients $\boldsymbol{\eta}_k^1$, $\boldsymbol{\eta}_k^2$, Young's modulus $\boldsymbol{E}_k$, and Poisson's ratio $\boldsymbol{\nu}_k$. It should be noted that the state vector in equation (9)

$$\underline{\boldsymbol{x}}_k = \left[\underline{\boldsymbol{c}}_k^{d\,\mathrm{T}}, \underline{\boldsymbol{c}}_k^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{10}$$

includes the finite approximation coefficients

$$\underline{\boldsymbol{c}}_k = \left[\underline{\boldsymbol{c}}_k^{x\,\mathrm{T}}, \underline{\boldsymbol{c}}_k^{y\,\mathrm{T}}, \underline{\boldsymbol{c}}_k^{z\,\mathrm{T}}\right]^{\mathrm{T}} \tag{11}$$

and their discrete derivatives

$$\underline{\boldsymbol{c}}_k^d = (\underline{\boldsymbol{c}}_k - \underline{\boldsymbol{c}}_{k-1})/\Delta t\,, \tag{12}$$

where the Δt is the time step and the vector $\underline{\boldsymbol{c}}_{k-1}$ denotes the approximation coefficients from the previous time step. It is important to note that the approximation coefficients $\underline{\boldsymbol{c}}_k$ represent the parameters of the transformation function. They are used for approximating the three-dimensional velocities and displacements of the heart surface. For example, the displacements of the heart surface are calculated by

$$\underline{\boldsymbol{u}}_k \approx \boldsymbol{\Phi}_k \underline{\boldsymbol{c}}_k\,, \tag{13}$$

where the approximation matrix

$$\boldsymbol{\Phi}_k := \mathrm{diag}\left\{\boldsymbol{\Phi}_k^i, \boldsymbol{\Phi}_k^i, \boldsymbol{\Phi}_k^i\right\}$$

is assembled from the matrix

$$\boldsymbol{\Phi}_k^i := \begin{pmatrix} \varphi_k(\underline{s}_0^1, l_0^1) & \cdots & \varphi_k(\underline{s}_0^1, l_0^N) \\ \vdots & \ddots & \vdots \\ \varphi_k(\underline{s}_0^L, l_0^1) & \cdots & \varphi_k(\underline{s}_0^L, l_0^N) \end{pmatrix} \tag{14}$$

with elements defined by the moving least-squares meshless function φ_k defined in [16]. It should be noted that the approximation incorporates N landmarks, whose three-dimensional positions $\underline{l}_0$ are averaged over the period of the heart motion. The positions of these landmarks are reconstructed according to [14]. The reconstruction is based on the corresponding image features $\underline{f}_0$ extracted from all images of the camera system. The reference positions of the L surface points $\underline{s}_0$ are obtained by interpolating between the landmarks $\underline{l}_0$ according to [17].

An important point to emphasize is that in [14] the heart excitation is assumed to be generated by a uniformly distributed pressure inside the cardiac chambers. However, the heart motion is initiated with internal forces generated in a distributed fashion within the muscles. In our paper, the considered heart surface model is extended so that we do not need an assumption on the uniformly distributed heart excitation. For this purpose, an additional unknown non-uniformly distributed excitation force $\underline{\boldsymbol{m}}_k^n$ is introduced in the system model (9) so that this model takes the form

$$\underline{\boldsymbol{x}}_{k+1} = \mathbf{A}_k(\underline{\boldsymbol{\psi}}_k)\,\underline{\boldsymbol{x}}_k + \mathbf{B}_k(\underline{\boldsymbol{\psi}}_k)\,(\hat{\underline{m}}_k^u + \underline{\boldsymbol{m}}_k^n + \underline{\boldsymbol{w}}_k)\,. \tag{15}$$

Furthermore, the unknown model parameters are augmented in the parameter vector

$$\underline{\boldsymbol{\theta}}_k = [\underline{\boldsymbol{\psi}}_k^{\mathrm{T}}, \underline{\boldsymbol{m}}_k^{n\,\mathrm{T}}]^{\mathrm{T}}, \tag{16}$$

where the vector $\underline{\boldsymbol{\psi}}_k$ is initialized by a nonlinear least-squares method, described in [14].

The *measurement process model* describes the relationship between the propagated state vector $\underline{\boldsymbol{x}}_k$ and the measurements. To simplify the explanation, we focus on the derivation of the measurement equation when the measurements

are provided by a single camera. By assuming that the positions of F image features are currently measured, we define the measurements by the vector $\hat{\underline{f}}_k \in \mathbb{R}^{2F}$. It should be noted that the number F of measured features can vary at every time step, e.g., when these measurement points, i.e., landmarks, are occluded or features are not detected. Furthermore, in this paper, the lens distortion of the cameras is assumed negligible.

For getting the measurement equation, first, the three-dimensional heart surface displacements are approximated using the state vector and moving least-squares meshless functions, according to (13). The approximation matrix $\boldsymbol{\Phi}_k$ is then computed according to equation (14) by substituting the vector $\underline{l}_0$ into $\underline{s}_0$. Finally, the current positions of the measurement points are projected to the current camera view at every time step k

$$\begin{aligned} \hat{\underline{f}}_k^x - \underline{f}_0^x &= \frac{r_{12}\underline{\boldsymbol{u}}_k^x + r_{12}\underline{\boldsymbol{u}}_k^y + r_{12}\underline{\boldsymbol{u}}_k^z + r_{14}}{r_{31}\underline{\boldsymbol{u}}_k^x + r_{32}\underline{\boldsymbol{u}}_k^y + r_{33}\underline{\boldsymbol{u}}_k^z + r_{34}} + \underline{\boldsymbol{v}}_k^x\,, \\ \hat{\underline{f}}_k^y - \underline{f}_0^y &= \frac{r_{21}\underline{\boldsymbol{u}}_k^x + r_{22}\underline{\boldsymbol{u}}_k^y + r_{23}\underline{\boldsymbol{u}}_k^z + r_{24}}{r_{31}\underline{\boldsymbol{u}}_k^x + r_{32}\underline{\boldsymbol{u}}_k^y + r_{33}\underline{\boldsymbol{u}}_k^z + r_{34}} + \underline{\boldsymbol{v}}_k^y\,. \end{aligned} \tag{17}$$

The error terms $\underline{\boldsymbol{v}}_k^x$ and $\underline{\boldsymbol{v}}_k^y$ denote the measurement disturbances $\underline{\boldsymbol{v}}_k = \left[\underline{\boldsymbol{v}}_k^{x\,\mathrm{T}}, \underline{\boldsymbol{v}}_k^{y\,\mathrm{T}}\right]^{\mathrm{T}}$ due to image noise and inaccurate projection. These uncertainties are assumed to be white zero-mean Gaussian $\underline{\boldsymbol{v}}_k \sim \mathcal{N}(\underline{0}, \mathbf{C}_k^v)$ with covariance $\mathbf{C}_k^v \in \mathbb{R}^{2F \times 2F}$.

For processing the measurements provided by a camera system, the equation (17) is constructed for every camera.

C. Simultaneous State and Parameter Estimation

The proposed image transformation is characterized by considering transformation uncertainties. This is achieved by involving a simultaneous state and parameter estimation in the transformation process.

For that purpose, the system state (10) is augmented with the parameter vector (16) in the vector $\underline{\boldsymbol{z}}_k = \left[\underline{\boldsymbol{x}}_k^{\mathrm{T}}, \underline{\boldsymbol{\theta}}_k^{\mathrm{T}}\right]^{\mathrm{T}}$. It can be easily shown that the substitution of this vector in the linear system model (15) leads to a nonlinear estimation problem that can be solved by different Bayesian nonlinear estimators. In this paper, the Gaussian filter, first proposed in [18], is applied for simultaneous state and parameter estimation. This linear regression Kalman filter characterizes the state estimate by its first two moments, i.e., mean and covariance matrix. These moments are captured in an approximate fashion by regression points, which are determined using deterministic sampling approach. Propagating these points through the nonlinear system and measurement equations (15, 17) implicitly linearizes these equations. Assuming the process $\underline{\boldsymbol{w}}_k$ and the measurement noise $\underline{\boldsymbol{v}}_k$ as uncorrelated and by representing the augmented state as a Gaussian random vector $\underline{\boldsymbol{z}}_k \sim \mathcal{N}\left(\underline{\mu}_k^e, \mathbf{C}_k^e\right)$, the Gaussian filter recursively updates the mean and covariance of the augmented state

$$\underline{\mu}_k^e = \left[\hat{\underline{x}}_k^{e\,\mathrm{T}}, \hat{\underline{\theta}}_k^{e\,\mathrm{T}}\right]^{\mathrm{T}}, \quad \mathbf{C}_k^e := \mathrm{diag}\left\{\mathbf{C}_k^{x,e}, \mathbf{C}_k^{\theta,e}\right\}$$

in two steps: prediction step and filter step. The prediction step provides the a priori density of the predicted state

Fig. 4. The measured posions of the small green markers are used for evaluating the visual stabilization and therefore, they are not considered by filtering. The measured positions of the large green markers are included in the measurement vector $\underline{\hat{f}}_k$.

Fig. 3. Experimental setup for evaluation of the image stabilization performance. The motion of the pressure-regulated artificial heart is measured by a trinocular camera system.

$f^p(\underline{z}_{k+1}) \sim \mathcal{N}\left(\underline{\mu}_{k+1}^p, \mathbf{C}_{k+1}^p\right)$. Since the system model is conditionally linear, similar to [19] and [20], the system equation can be decomposed into linear and nonlinear substructures by means of Rao-Blackwellization. For efficiency of the estimation, only the nonlinear part is processed by the Gaussian filter in an approximate fashion. Furthermore, the filter step provides the a posteriori density of the updated state $f^e(\underline{z}_k) \sim \mathcal{N}\left(\underline{\mu}_k^e, \mathbf{C}_k^e\right)$. Since the measurement model (17) is nonlinear, the standard filter step provided by the Gaussian filter [18] is applied. In this step, the deviation between the predicted position of the image features and corresponding measurements $\underline{\hat{f}}_k$ is used for correcting the state and thus adapting the parameters of the transformation function and of the underlying physical model, as shown in Fig. 2.

Using the mean values of the updated state, the mean value of the three-dimensional heart surface displacements $\underline{\hat{u}}_k = \left[\underline{\hat{u}}_k^{x\mathrm{T}}, \underline{\hat{u}}_k^{y\mathrm{T}}, \underline{\hat{u}}_k^{z\mathrm{T}}\right]^{\mathrm{T}}$ is estimated by equation

$$\underline{\hat{u}}_k = \mathbf{H}_k \underline{\hat{x}}_k^e \ , \tag{18}$$

where the matrix $\mathbf{H}_k := \mathrm{diag}\{\mathbf{0}, \mathbf{\Phi}_k\}$ contains the approximation matrix (13), depending on the positions of the landmarks $\underline{l}_0$ and the positions of the measurement points $\underline{f}_0$. As a result, by substituting $\underline{\hat{u}}_k$ into $\underline{u}_k$ of (8), we obtain the mapping between the pixels of the reference image and current image.

In contrast to the existing approaches, this model-based image transformation function allows to transform the actual image to the reference image also when no measurement information is available. In this case, for reconstructing the heart surface displacements (18), the predicted state $\underline{\hat{x}}_k^p$ is used. As a result, in spite of the loss of measurement information loss, the correspondences between the pixels of the current image and the reference images are estimated and the color information of the current image can be written to the reference image.

V. EVALUATION

After a detailed description of the experimental setup, the accuracy of the visual stabilization will be analyzed in a series of experiments.
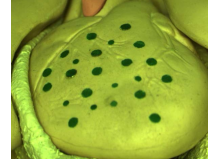
A. Experimental Setup

The performance of the proposed approach for visual stabilization is evaluated in an experiment with a pressure-regulated artificial heart. The experimental setup is shown in Fig. 3. The motion of the stabilized beating heart, according to the medical study performed in [12], is $0.59\,\mathrm{mm}$ in the lateral plane (x-y-axis) and $2.1\,\mathrm{mm}$ out-of-plane (z-axis). The obtained amplitude of the artificial heart motion is $1.3\,\mathrm{mm}$ and $1.2\,\mathrm{mm}$ in x- and y-direction, respectively, and $2.5\,\mathrm{mm}$ out-of-plane. It is generated by a pressure signal with amplitude of $100\,\mathrm{hPa}$ and a frequency of $1.2\,\mathrm{Hz}$. Therefore, the motion of the artificial heart is assumed to be comparable to the real heart deformations. For measuring the heart surface motion, a trinocular camera system is installed at a distance of $50\,\mathrm{cm}$ above the monitored operation area. The three cameras PIKE F-210C with resolution $1920\,\mathrm{pixel} \times 1080\,\mathrm{pixel}$ provide an accuracy of $0.2\,\mathrm{mm/pixel}$ in y-direction and $0.11\,\mathrm{mm/pixel}$ in x-direction of the image coordinate system. The camera baselines are about $57\,\mathrm{cm}$, their focal length is about $35\,\mathrm{mm}$, and the field of view is $12.8\,\mathrm{cm} \times 17.02\,\mathrm{cm}$. The image size of every camera is reduced by cutting out the defined regions of interest.

For evaluating the performance of the visual stabilization, additional evaluation points are introduced. These points are presented in Fig. 4 by small green markers. Their measured position is not considered by the estimator in the filter step. For the evaluation, six image sequences are collected, each of these consists of 300 images. The frame rate is about $22\,\mathrm{frames/\,sec}$. These sequences are finally transformed to the reference image. For every image sequence, the distances between the stabilized and the reference position averaged over the number of the evaluation points is defined as a transformation error at every time step.

B. Experimental Results

For demonstrating the high quality of the visual stabilization, the proposed model-based image transformation is compared with the geometric image transformation based on the direct processing of the measurement data. This transformation is introduced in [9]. In Fig. 5(a), the transformation error of both approaches averaged over 6 image sequences is shown over time. In order to emphasize the challenges of the visual stabilization, the transformation error of the non-stabilized motion of the evaluation points is depicted in Fig. 5(b). According to the presented results, the maximum error of the geometric image transformation is $3.2\,\mathrm{pixel}$, where the maximum error of the model-based image transformation is $2.6\,\mathrm{pixel}$. Therefore, the proposed physics-based image transformation provides about 19% higher accuracy
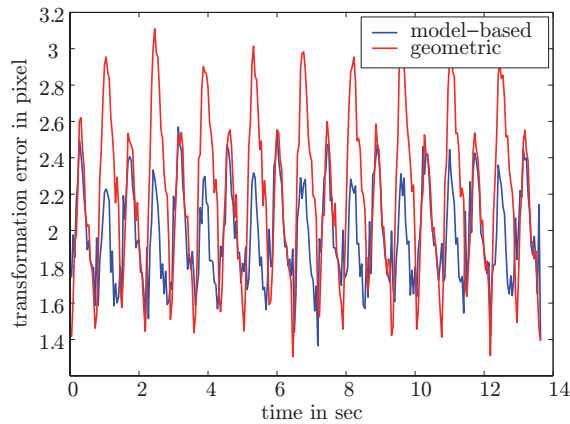
(a) The transformation errors of the geometric and model-based image transformation methods.
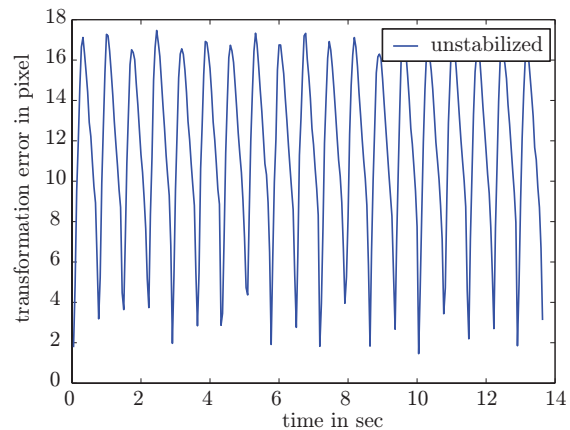
(b) The transformation error of the non-stabilized image sequence for representing the challenges of the image transformation.

Fig. 5. Comparison of the results provided by the geometric and model-based visual stabilization. The transformation error is averaged over six image sequences containing 300 images. According to the maximum transformation error, the model-based image transformation provides about 19% higher accuracy then the geometric image transformation.

then the geometric approach. It is especially efficient in compensating strong displacements of the evaluation points, which range up to 18 pixel.

VI. CONCLUSION

The main idea of the visual stabilization is assisting a surgeon during an operation on a beating heart by presenting him a beating heart as non-moving. In this way, the surgeon can operate without being disturbed by the heart motion. This enhances the surgeons capabilities.

The proposed approach for visual stabilization of the beating heart is characterized by three new ideas. First, in contrast to existing methods based on direct processing of measurement data, it exploits information from a stochastic physical heart surface model. This allows visual stabilization also when no measurement information is available. Second, this approach copes with uncertainties of the measurements and the image transformation by using nonlinear estimation. Third, the quality of visual stabilization is continuously improved during a beating heart operation by adapting the parameters of the transformation function.

In future work, the extension of the proposed approach to three-dimensional visual stabilization is planned. This is possible by projecting two stabilized stereo views to a head-mounted display of the surgeon or transformation of the stabilization in three-dimensional space.

REFERENCES

[1] Y. Nakamura, K. Kishi, and H. Kawakami, "Heartbeat Synchronization for Robotic Cardiac Surgery," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2001)*, Seoul, Korea, May 2001, pp. 2014–2019.

[2] R. Richa, P. Poignet, and C. Liu, "Three-dimensional Motion Tracking for Beating Heart Surgery Using a Thin-plate Spline Deformable Model," *International Journal of Robotics Research*, vol. 20, no. 2-3, pp. 218–230, Feb. 2010.

[3] S. G. Yuen, P. M. Novotny, and R. D. Howe, "Quasiperiodic Predictive Filtering for Robot-Assisted Beating Heart Surgery," in *Proceeding of the IEEE International Conference on Robotics and Automation (ICRA 2008)*, Pasadena, CA, USA, May 2008, pp. 3875–3880.

[4] E. Bogatyrenko, U. D. Hanebeck, and G. Szabo, "Heart Surface Motion Estimation Framework for Robotic Surgery Employing Meshless Methods," in *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2009)*, October 2009.

[5] T. J. Gilhuly, S. E. Salcudean, and S. V. Lichtenstein, "Evaluating Optical Stabilization of the Beating Heart," *IEEE Engineering in Medicine and Biology Magazine*, pp. 133–140, 2003.

[6] D. Stoyanov and G. Yang, "Stabilization of Image Motion for Robotic Assisted Beating Heart Surgery," *Medical Image Computing and Computer-Assisted Intervention*, vol. 10, pp. 417–424, 2007.

[7] M. Gröger and G. Herzinger, "Image Stabilisation of the Beating Heart by Local Linear Interpolation," in *Medical Imagaing: visualization, image-guided procedures and display*, 2006.

[8] C. A. Glasbey and K. V. Mardia, "A Review of Image Warping Methods," *Journal of Applied Statistics*, vol. 25, pp. 155–171, 1998.

[9] N. Arad and D. Reisfeld, "Image Warping Using few Anchor Points and Radial Funcions," *CVGIP: Graphical Models and Image Processing*, vol. 56, pp. 161–172, 1994.

[10] G. Wolberg, "Image Morphing: a Survey," *The Visual Computer*, vol. 14, pp. 360–372, 1998.

[11] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge: Cambridge University Press, 2000.

[12] P. Cattina, H. Daveb, J. Gruenenfelder, G. Szekelya, M. Turin, and G. Zuend, "Trajectory of Coronary Motion and Its Significance in Robotic Motion Cancellation," *European Journal of Cardio-thoracic Surgery*, vol. 25, pp. 786–790, 2004.

[13] R. M. Shoucri, "The End-Systolic Pressure-Volume Relation and Its Application to the Study of the Contractility of the Cardiac Muscle," *Computers in Cardiology*, vol. 33, pp. 297–300, 2006.

[14] E. Bogatyrenko, P. Pompey, and U. D. Hanebeck, "Efficient Physics-Based Tracking of Heart Surface Motion for Beating Heart Surgery Robotic Systems," *International Journal of Computer Assisted Radiology and Surgery (IJCARS 2010)*, 2010.

[15] Atluri, *The meshless Method (MLPG) for Domain and BIE Discretisation*. Tech Science Press, 2004.

[16] G.-R. Liu, *Mesh Free Methods: Moving Beyond the Finite Element Methods*. CRC Press, 2003.

[17] S. Lee, G. Wolberg, and S. Shin, "Scattered data interpolation with multilevel B-splines," *Visualization and Computer Graphics, IEEE Transactions*, vol. 3, pp. 228 – 244, 1997.

[18] M. F. Huber and U. D. Hanebeck, "Gaussian Filter based on Deterministic Sampling for High Quality Nonlinear Estimation," in *Proceedings of the 17th IFAC World Congress (IFAC 2008)*, vol. 17, no. 2, Seoul, Korea, Jul. 2008.

[19] F. Beutler, M. F. Huber, and U. D. Hanebeck, "Gaussian Filtering using State Decomposition Methods," in *Proceedings of the 12th International Conference on Information Fusion (Fusion 2009)*, Seattle, Washington, Jul. 2009.

[20] E. Bogatyrenko and U. D. Hanebeck, "Simultaneous State and Parameter Estimation for Physics-Based Tracking of Heart Surface Motion," in *Proceedings of the 2010 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2010)*, Salt Lake City, Utah, Sep. 2010.