IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 2, 2020, accepted November 11, 2020, date of publication December 1, 2020,
date of current version December 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3041796

Tracking Carotid Artery Wall Motion Using an Unscented Kalman Filter and Data Fusion

JAN DORAZIL[1,2], RENE REPP[3], THOMAS KROPFREITER[1], RICHARD PRÜLLER[1], KAMIL ŘÍHA[2], AND FRANZ HLAWATSCH[1], (Fellow, IEEE)

[1]Institute of Telecommunications, TU Wien, 1040 Vienna, Austria
[2]Department of Telecommunications, Brno University of Technology, 601 90 Brno, Czech Republic
[3]Acoustics Research Institute, Austrian Academy of Sciences, 1030 Vienna, Austria

Corresponding author: Jan Dorazil (doraz.jan@gmail.com)

This work was supported by the Czech Science Foundation (GAČR) under Grant 17-19638S.

ABSTRACT Analyzing the motion of the common carotid artery (CCA) wall yields effective indicators for atherosclerosis. In this work, we propose a state-space model and a tracking method for estimating the time-varying CCA wall radius from a B-mode ultrasound sequence of arbitrary length. We employ an unscented Kalman filter that fuses two sets of measurements produced by an optical flow algorithm and a CCA wall localization algorithm. This fusion-and-tracking approach ensures that feature drift, which tends to impair optical flow based methods, is compensated in a temporally consistent manner. Simulation results show that the proposed method outperforms a recently proposed optical flow based method.

INDEX TERMS Atherosclerosis, data fusion, unscented Kalman Filter, motion estimation, ultrasonography, carotid artery, medical imaging, ultrasound imaging.

I. INTRODUCTION

A. BACKGROUND, MOTIVATION, STATE OF THE ART

According to the World Health Organization (WHO), over 31% of all deaths in 2016 were caused by coronary events, most of which can be attributed to a single progressive disease called atherosclerosis [1], [2]. It is well known that the progression of atherosclerosis is associated with increasing arterial stiffness [3]. Several diagnostic methodologies exploit this association in order to estimate the risk of coronary events. These methodologies are generally based on measurements of quantities that are intrinsically associated with arterial stiffness [4], [5]. One of these quantities is the motion of the arterial wall in response to variations in blood pressure or blood flow, which occurs naturally in the human body.

This article presents a new methodology for the analysis of the motion of the wall of the common carotid artery (CCA) from a B-mode ultrasound sequence. In screening tests, methods based on the analysis of the motion of the CCA wall have been shown to produce effective indicators for atherosclerosis [6], [7]. The main challenge here is an accurate and reliable estimation of CCA wall motion from an ultrasound sequence. Typically, computer-aided methods for this task explicitly or implicitly involve a procedure called speckle tracking [8]. One of the first applications of speckle tracking to the estimation of CCA wall motion was based on block matching [9], [10]; however, the performance of block matching methods is negatively affected by a phenomenon known as speckle decorrelation [9]. Subsequent studies addressed this issue by using a state-space model for the evolution of the reference block [11], [12] or for the movement of the artery [7], [13].

As an alternative to block matching, several authors proposed speckle tracking methods based on optical flow [14]. For estimating CCA wall motion in particular, the Lucas-Kanade algorithm and its extensions were used [15], [16]. A comparative study of different speckle tracking methods in [17] demonstrated that a modified Lucas-Kanade algorithm outperforms other optical flow algorithms as well as block matching algorithms. In optical flow methods, speckle decorrelation translates into a phenomenon known as feature drift [18], which again negatively affects the performance. Indeed, the position of a selected point (feature) is tracked through cumulative summation of the estimated displacements between consecutive ultrasound frames, and hence the errors in the overall displacement estimates are accumulated as well. This may result—especially in the case of long ultrasound sequences—in a progressive divergence of the estimated point trajectory from the true trajectory, i.e., the estimated point "drifts away" from its true position. In [18], a Lucas-Kanade algorithm based method with explicit feature drift compensation was proposed; this method outperforms an earlier method described in [15]. All the mentioned studies

The associate editor coordinating the review of this manuscript and approving it for publication was Halil Ersin Soken.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

analyze longitudinal scans of the CCA, with the exception of [15] and [18], where transverse scans of the CCA are analyzed using the Lucas-Kanade algorithm. In a transverse scan, the shape of the CCA wall cross section can be approximated by a circle, which simplifies the modeling.

B. CONTRIBUTION

In this work, we propose a new model-based method for estimating CCA wall motion. The CCA wall cross section is modeled as a circle, and our goal is to estimate the time-varying radius of that circle from—possibly long—B-mode ultrasound sequences showing the CCA transverse section. Based on a new state-space model for CCA wall motion, and using the result of an optical flow algorithm, we employ the unscented Kalman filter (UKF) [19] to track the time-varying center point and radius of the CCA wall circle.

As mentioned earlier, the performance of optical flow methods is impaired by the phenomenon of feature drift. In order to counteract feature drift, our implementation of the UKF fuses measurements from two different sources. One source is the pyramidal implementation of the Lucas-Kanade optical flow algorithm [20], which tracks the positions of a set of *feature points* (FPs) around the CCA wall. Here, an FP is a point in the ultrasound frame that is assumed to represent the local tissue movement, as manifested by a high response of the Harris detector [21]. The other source is the circle localization algorithm proposed in [22], which provides preliminary estimates of the center point and radius of the CCA circle. In this new fusion approach, feature drift in the tangential direction (with respect to the CCA circle) is reflected by a deviation from the state-space model, while feature drift in the radial direction is reflected by a deviation from the estimates provided by the circle localization algorithm. In contrast to the method of [18], the proposed method compensates for the feature drift in a temporally consistent manner that takes into account the natural smoothness of CCA wall motion. This inherent temporally consistent feature drift compensation is a major advantage of our method.

We assess the accuracy and robustness of the proposed method in several different synthetic scenarios with various levels of additive and multiplicative noise and various degrees of speckle decorrelation. We also demonstrate experimentally that our method outperforms the method proposed in [18] in all considered scenarios. We chose the latter method as a reference method since, to the best of our knowledge, it is the only existing method performing an explicit feature drift compensation, which is required for good performance in the case of long ultrasound sequences. The reference method was already benchmarked to a conventional artery wall tracking method in [18]. Finally, we present results obtained for a real ultrasound sequence and discuss their physiological validity.

The proposed method differs from the method previously presented in our conference publication [23] in that it estimates (tracks) the fundamental frequency—corresponding to the heart rate—used in the underlying state-space model. Thus, in contrast to [23], the fundamental frequency does not have to be precisely known in advance and is allowed to vary with time. Furthermore, compared to [23], our validation of the performance of the proposed method is based on more realistic simulated sequences with a time-varying fundamental frequency, and we present a detailed description of the generation of these sequences. Finally, the technical presentation is more detailed.

C. PAPER ORGANIZATION

The remainder of this paper is organized as follows. An overview of the proposed method is presented in Section II. The calculation of the input to the UKF is described in Section III. The signal and state-space models on which the proposed method is based are developed in Section IV. The operation of the UKF is discussed in Section V. Finally, experimental results obtained for synthetic and real data are presented in Section VI.

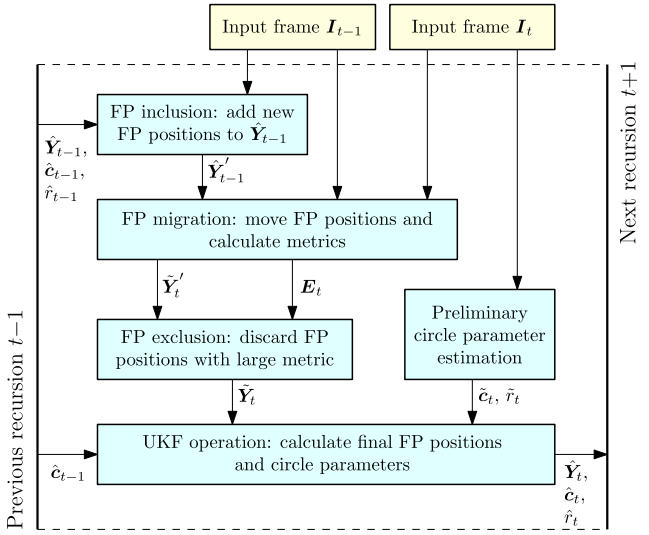


FIGURE 1. One frame recursion of the proposed method.

II. METHOD OVERVIEW

For tracking a circle representing the CCA wall cross section, the proposed method fuses the results produced by the pyramidal Lucas-Kanade optical flow algorithm of [20] and by the circle localization algorithm of [22]. This fusion is performed by an UKF [19], which tracks the circle parameters and related quantities in a frame-sequential, frame-recursive manner. The input to our method is a B-mode ultrasound video sequence consisting of frames $\boldsymbol{I}_t$ that show temporally successive transverse scans of the CCA. One recursion of the method, corresponding to frame (or time) index t, is visualized in Figure 1. The basic geometry of the CCA wall representation by a circle and of the surrounding FPs is depicted in Figure 2. We denote the center point of the circle by $\boldsymbol{c}_t$, the circle radius by r_t, and the two-dimensional Cartesian position vector of the n-th FP by $\boldsymbol{y}_t^{(n)}$. In Figure 3,

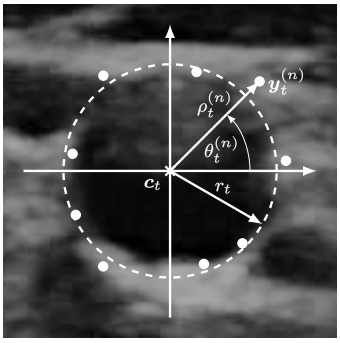


FIGURE 2. CCA wall circle (dashed) and FPs (bullets).

we visualize each step of the proposed method (as presented in Figure 1).

For each frame $\boldsymbol{I}_t$ with $t = 1, 2, \ldots$, our method calculates a list $\hat{\boldsymbol{Y}}_t$ of estimates of N_t FP positions, $\hat{\boldsymbol{y}}_t^{(n)}$ for $n = 1, 2, \ldots, N_t$, as well as an estimate of the CCA circle center point, $\hat{\boldsymbol{c}}_t$, and of the CCA circle radius, $\hat{r}_t$. For the first frame, $t = 1$, the FP position list $\hat{\boldsymbol{Y}}_1$ is empty and the CCA circle parameter estimates $\hat{\boldsymbol{c}}_1$ and $\hat{r}_1$ are calculated by applying the circle localization algorithm of [22] to $\boldsymbol{I}_1$.

For the further frames, $t = 2, 3, \ldots$, the FP position list $\hat{\boldsymbol{Y}}_t$ and the circle parameter estimates $\hat{\boldsymbol{c}}_t$ and $\hat{r}_t$ are calculated frame-sequentially as follows (see Figures 1 and 3). In a first stage, as explained in Section III, we determine quantities that serve as the input to the UKF. First, the previous FP position list $\hat{\boldsymbol{Y}}_{t-1}$ is updated. To this end, we perform an *FP inclusion* step in which new FPs are identified in an annular search region in frame $\boldsymbol{I}_{t-1}$ that depends on the previous circle parameter estimates $\hat{\boldsymbol{c}}_{t-1}$ and $\hat{r}_{t-1}$, and the positions of these new FPs are estimated and included in the list. The resulting augmented list will be denoted by $\hat{\boldsymbol{Y}}_{t-1}^{'}$. Next, in the *FP migration* step, the pyramidal Lucas-Kanade optical flow algorithm is used to move the FP positions in $\hat{\boldsymbol{Y}}_{t-1}^{'}$ such that they conform to the current frame $\boldsymbol{I}_t$; this results in an updated FP position list $\tilde{\boldsymbol{Y}}_t^{'}$ and a list of corresponding error metrics $\boldsymbol{E}_t$. Finally, in the *FP exclusion* step, FPs with a low confidence of successful localization by the optical flow algorithm, as determined from $\boldsymbol{E}_t$, are excluded from the FP position list. This yields a preliminary FP position list $\tilde{\boldsymbol{Y}}_t$ consisting of preliminary FP position estimates $\tilde{\boldsymbol{y}}_t^{(n)}$, $n = 1, 2, \ldots, N_t$. We note that the FP inclusion and exclusion steps are performed because of speckle decorrelation. Indeed, an FP may become progressively harder to localize (i.e., with less confidence) via the optical flow algorithm as the ultrasound image evolves. Such FPs need to be excluded from the tracking process and possibly replaced by new FPs.

In parallel to the FP inclusion, migration, and exclusion operations, preliminary circle parameter estimates $\tilde{\boldsymbol{c}}_t$ and $\tilde{r}_t$ are calculated by applying the circle localization algorithm of [22] to the current frame $\boldsymbol{I}_t$.

Then, in a second stage, the UKF fuses the information provided by the preliminary FP position list $\tilde{\boldsymbol{Y}}_t$ and by the preliminary circle parameter estimates $\tilde{\boldsymbol{c}}_t$ and $\tilde{r}_t$. This information constitutes the input (i.e., the "observed measurements") of the UKF. The UKF then calculates final FP position estimates $\hat{\boldsymbol{y}}_t^{(n)}, n = 1, 2, \ldots, N_t$, which make up the final FP position list $\hat{\boldsymbol{Y}}_t$, as well as final circle parameter estimates $\hat{\boldsymbol{c}}_t$ and $\hat{r}_t$. This is described in Section V. The UKF is based on a stochastic "system model" that describes the temporal variation of the CCA circle parameters, the evolution of the FP positions, and the extraction of the UKF input. This system model will be developed in Section IV.

III. CALCULATION OF THE UKF INPUT

The input to the UKF at frame time t consists of the preliminary FP position list $\tilde{\boldsymbol{Y}}_t$ comprising preliminary FP position estimates $\tilde{\boldsymbol{y}}_t^{(n)}$, $n = 1, 2, \ldots, N_t$, as well as of the preliminary circle parameter estimates $\tilde{\boldsymbol{c}}_t$ and $\tilde{r}_t$ (see Figure 1). These quantities are calculated as described next.

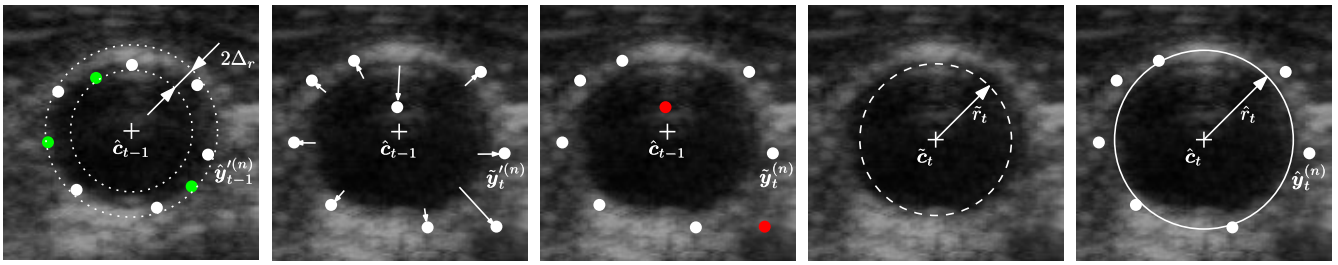
A. FP INCLUSION

First, the previous FP position list $\hat{\boldsymbol{Y}}_{t-1}$ is updated by performing the FP inclusion, migration, and exclusion steps. In the FP inclusion step, we use the Harris detector [21] to detect new FPs and estimate their positions within an annular search region in frame $\boldsymbol{I}_{t-1}$. The annular search region is defined by the center point $\hat{\boldsymbol{c}}_{t-1}$, the inner radius $\hat{r}_{t-1} - \Delta_r$, and the outer radius $\hat{r}_{t-1} + \Delta_r$, where $\hat{\boldsymbol{c}}_{t-1}$ and $\hat{r}_{t-1}$ are the previous circle parameter estimates and $\Delta_r > 0$ is a fixed parameter. From the set of detected new FPs, we discard all FPs whose Harris response [21] is lower than a threshold τ_{h}, and the surviving new FPs are ordered in the sense of decreasing Harris response. Let $\boldsymbol{Y}_t^{\mathrm{new}}$ denote the resulting ordered list of new FP positions.

Going through the list $\boldsymbol{Y}_t^{\mathrm{new}}$ in the chosen order, each new FP position is now included in the overall FP position list if it is sufficiently distant from all the FP positions in $\hat{\boldsymbol{Y}}_{t-1}$ and all the new FP positions included previously. More specifically, if the Euclidean distances of the first FP position in $\boldsymbol{Y}_t^{\mathrm{new}}$ (i.e., the one with the largest Harris response) from all the FP positions in $\hat{\boldsymbol{Y}}_{t-1}$ are larger than a threshold $d > 0$, then that FP position is added to $\hat{\boldsymbol{Y}}_{t-1}$, which results in an augmented list. A similar operation is done for the second, third, etc. FP position in $\boldsymbol{Y}_t^{\mathrm{new}}$: if the distances of the respective FP position from all the FP positions in the current augmented list are larger than d, then that FP position is added to the current augmented list. Processing all the FP positions in $\boldsymbol{Y}_t^{\mathrm{new}}$ in this manner yields a final augmented list of $N_{t-1}^{\prime}$ FP positions $\hat{\boldsymbol{y}}_{t-1}^{(n)\prime}$, $n = 1, 2, \ldots, N_{t-1}^{\prime}$, which will be denoted as $\hat{\boldsymbol{Y}}_{t-1}^{'}$. By the above construction, the FP positions $\hat{\boldsymbol{y}}_{t-1}^{(n)\prime}$ are approximately uniformly distributed in the annular search region, and the corresponding FPs have predominantly high Harris responses.

B. FP MIGRATION

The augmented FP position list $\hat{\boldsymbol{Y}}_{t-1}^{'}$ was calculated from frame $\boldsymbol{I}_{t-1}$, and thus it does not yet take into account the current frame $\boldsymbol{I}_t$. Therefore, in the next step of our method—the FP migration step—we adapt $\hat{\boldsymbol{Y}}_{t-1}^{'}$ to $\boldsymbol{I}_t$. To this end,



(a) FP inclusion (b) FP migration (c) FP exclusion (d) Preliminary circle parameter estimation (e) UKF operation

FIGURE 3. Visualization of the individual steps of the proposed method—as presented in Figure 1—for two successive ultrasound frames $\boldsymbol{I}_{t-1}$ (in (a)) and $\boldsymbol{I}_t$ (in (b)–(e)). Bullets represent FPs. In particular, the green bullets in (a) depict FPs newly included by the FP inclusion step, while the red bullets in (c) depict FPs to be excluded by the FP exclusion step. The dotted circles in (a) represent the boundary of the annular search region used in the FP inclusion step. The dashed circle in (d) represents the preliminary CCA circle calculated by the circle localization algorithm of [22]. The solid circle in (e) represents the final CCA circle estimate produced by the UKF. The width of the annular search region in (a) and the length of the FP motion vectors in (b) are exaggerated for better visualization.

we use the pyramidal implementation of the Lucas-Kanade optical flow algorithm [20] to move each FP position $\hat{\boldsymbol{y}}_{t-1}^{(n)\prime}$ in $\hat{\boldsymbol{Y}}_{t-1}^{\prime}$ to a new position, denoted as $\tilde{\boldsymbol{y}}_t^{(n)\prime}$, that depends on both $\hat{\boldsymbol{y}}_{t-1}^{(n)\prime}$ and $\boldsymbol{I}_t$. The resulting list of $N_{t-1}^{\prime}$ FP positions $\tilde{\boldsymbol{y}}_t^{(n)\prime}$, $n = 1, 2, \ldots, N_{t-1}^{\prime}$ is denoted as $\tilde{\boldsymbol{Y}}_t^{\prime}$.

For each FP position $\tilde{\boldsymbol{y}}_t^{(n)\prime}$ in $\tilde{\boldsymbol{Y}}_t^{\prime}$, we furthermore determine an error metric $\varepsilon_t^{(n)}$ by comparing the $J \times J$ patch of frame $\boldsymbol{I}_t$ centered at position $\tilde{\boldsymbol{y}}_t^{(n)\prime}$ to the $J \times J$ patch of frame $\boldsymbol{I}_{t-1}$ centered at position $\hat{\boldsymbol{y}}_{t-1}^{(n)\prime}$. More specifically, $\varepsilon_t^{(n)}$ is calculated as the ℓ_1 metric between the two patches, i.e.,

$$\varepsilon_t^{(n)} \triangleq \sum_{k=1}^{J}\sum_{l=1}^{J} \left| I_t^{(\tilde{\boldsymbol{y}}_t^{\prime})}(k,l) - I_{t-1}^{(\hat{\boldsymbol{y}}_{t-1}^{\prime})}(k,l) \right|, \tag{1}$$

where $I_t^{(\tilde{\boldsymbol{y}}_t^{\prime})}(k,l)$ and $I_{t-1}^{(\hat{\boldsymbol{y}}_{t-1}^{\prime})}(k,l)$ denote corresponding pixels of the two patches. The list of error metrics $\varepsilon_t^{(n)}$, $n = 1, 2, \ldots, N_{t-1}^{\prime}$ will be denoted as $\boldsymbol{E}_t$. The use of this metric is based on the assumption that the motion of the examined tissue in the neighborhood of an FP can be approximated sufficiently well by a simple translation. This assumption can be relaxed, as discussed in Section VII.

C. FP EXCLUSION

The FP migration step described above is followed by the FP exclusion step. An FP position $\tilde{\boldsymbol{y}}_t^{(n)\prime}$ is excluded from the FP position list $\tilde{\boldsymbol{Y}}_t^{\prime}$ if the corresponding error metric $\varepsilon_t^{(n)}$ in list $\boldsymbol{E}_t$ satisfies $\varepsilon_t^{(n)} > \alpha\overline{\varepsilon}_t$, where $\overline{\varepsilon}_t$ is the arithmetic mean of all the error metrics in $\boldsymbol{E}_t$ and $\alpha > 1$ is a fixed parameter. This rule is motivated by our empirical observation that there is usually a small proportion of FPs whose error is much higher than the average, and excluding these "outliers" improves the tracking accuracy. The FP positions surviving this exclusion step are denoted as $\tilde{\boldsymbol{y}}_t^{(n)}$, $n = 1, 2, \ldots, N_t$, and the corresponding list as $\tilde{\boldsymbol{Y}}_t$. This list constitutes one of the inputs to the UKF.

D. PRELIMINARY CIRCLE PARAMETER ESTIMATION

The second input to the UKF is obtained by applying the circle localization algorithm of [22] to the current frame $\boldsymbol{I}_t$. This results in "preliminary" circle parameter estimates $\tilde{\boldsymbol{c}}_t$ and $\tilde{r}_t$. The circle localization algorithm of [22] differs from standard algorithms such as the Hough transform [24] in that it utilizes the pixel intensity directly rather than the image gradient. This is beneficial in our context because the blurred appearance and speckled texture of ultrasound images make accurate gradient determination challenging. The preliminary circle parameter estimates $\tilde{\boldsymbol{c}}_t$ and $\tilde{r}_t$ provide a complementary "absolute" position information that is used in the UKF to compensate for feature drift. Without this absolute position information—i.e., relying only on the relative optical flow information underlying the FP migration step—the FP position estimates and circle radius estimates produced by the UKF would drift away from the CCA wall (as explained earlier in Section I-A).

IV. SYSTEM MODEL

A major contribution of our work is a new state space model that constitutes a stochastic description of the dynamics of CCA wall motion (within the transverse cross section considered) and of a related measurement process. This model is inspired by the state-space model of [25], which uses a Fourier series with time-varying coefficients whose temporal evolution is modeled by a random walk. The proposed state space model consists of a *state evolution model* and a *measurement model*, and it provides the basis for the operation of the UKF. We note that although this model is designed for the CCA transverse section, the underlying approach can be used in many other speckle-tracking scenarios where a quasi-periodic motion is observed.

A. CCA CIRCLE RADIUS AND FP RADII

Our state space model is based on models for the time-dependence of the CCA circle radius and of the FP radii.

The regularity of the heart beat manifests itself in an approximate periodicity of the movement of the CCA wall. This approximate periodicity suggests the use of a Fourier series to model the CCA wall movement [26, Ch. 11]. Concretely, we model the time-dependence of the CCA circle radius r_t by the sum of a constant R that describes the mean CCA circle radius, a quasi-periodic function that is constructed as the superposition of M harmonic components with time-varying Fourier coefficients $a_t^{(m)}$ and $b_t^{(m)}$, $m = 1, 2, \ldots, M$ and time-varying fundamental frequency, and a slowly varying component γ_t that represents the effect of breathing [18], [27]. That is,

$$r_t = R + \sum_{m=1}^{M} \left(a_t^{(m)} \cos(m\varphi_t) + b_t^{(m)} \sin(m\varphi_t) \right) + \gamma_t, \quad (2)$$

where $m\varphi_t$ is the instantaneous phase of the m-th component and φ_t can be considered as the "fundamental" instantaneous phase (whose temporal derivative—if t were continuous—would be a time-varying fundamental frequency).

We consider N_t FPs at each frame time t. The position of the n-th FP at frame time t, relative to the CCA circle center point $\boldsymbol{c}_t$, can be expressed as

$$\boldsymbol{y}_t^{(n)} = \begin{pmatrix} \rho_t^{(n)} \cos\theta_t^{(n)} \\ \rho_t^{(n)} \sin\theta_t^{(n)} \end{pmatrix}, \quad n = 1, 2, \ldots, N_t. \quad (3)$$

Here, $\rho_t^{(n)}$ (referred to as the "FP radius") and $\theta_t^{(n)}$ represent the polar coordinates of FP n with respect to the center $\boldsymbol{c}_t$, as illustrated in Figure 2. The FP radius is modeled as

$$\rho_t^{(n)} = r_t + \tilde{\delta}^{(n)} + \delta_t^{(n)}, \quad n = 1, 2, \ldots, N_t, \quad (4)$$

where r_t is the CCA circle radius, $\tilde{\delta}^{(n)}$ denotes an initial approximation of the radial deviation of the FP from the CCA circle (which was calculated at the frame time when the FP was included in the FP position list, see Section V-A), and $\delta_t^{(n)}$ denotes a time-varying radial deviation.

B. STATE

The UKF estimates a time-varying state vector $\boldsymbol{x}_t$ in a frame-recursive manner, using at each frame time t the results of the previous frame time $t-1$ and a measurement vector $\boldsymbol{z}_t$. In our method, the state is a vector of dimension $L_t \triangleq 2M + 2N_t + 7$ defined as

$$\boldsymbol{x}_t \triangleq \left(\boldsymbol{a}_t^{\mathrm{T}} \ \boldsymbol{b}_t^{\mathrm{T}} \ \boldsymbol{\varphi}_t^{\mathrm{T}} \ \boldsymbol{\gamma}_t^{\mathrm{T}} \ R_t \ \boldsymbol{c}_t^{\mathrm{T}} \ \boldsymbol{p}_t^{(1)\mathrm{T}} \cdots \boldsymbol{p}_t^{(N_t)\mathrm{T}} \right)^{\mathrm{T}}. \quad (5)$$

Here, related to the CCA circle radius model in (2), $\boldsymbol{a}_t \triangleq (a_t^{(1)} \cdots a_t^{(M)})^{\mathrm{T}}$ and $\boldsymbol{b}_t \triangleq (b_t^{(1)} \cdots b_t^{(M)})^{\mathrm{T}}$ are the vectors of time-varying Fourier coefficients of the quasi-periodic component of the CCA circle radius, $\boldsymbol{\varphi}_t \triangleq (\varphi_t \ \ \dot{\varphi}_t)^{\mathrm{T}}$ comprises the fundamental instantaneous phase and its derivative, i.e., the time-varying fundamental frequency, $\boldsymbol{\gamma}_t \triangleq (\gamma_t \ \ \gamma_{t-1})^{\mathrm{T}}$ comprises the current and previous samples of the breathing component of the CCA circle radius, and R_t is a (formally) time-varying version of the mean CCA circle radius R. Furthermore, related to the FP model in (3) and (4), $\boldsymbol{c}_t \triangleq (c_t^{(1)} \ c_t^{(2)})^{\mathrm{T}}$ is the CCA circle center point in Cartesian coordinates, and the *FP position parameter vector*

$$\boldsymbol{p}_t^{(n)} \triangleq \begin{pmatrix} \delta_t^{(n)} \\ \theta_t^{(n)} \end{pmatrix} \quad (6)$$

characterizes the position of FP n in polar coordinates. Note that from the state $\boldsymbol{x}_t$, it is possible to determine the CCA circle center point $\boldsymbol{c}_t$ and radius r_t as well as all the FP positions (up to the initial radial deviation approximations $\tilde{\delta}^{(n)}$).

C. STATE EVOLUTION MODEL

The UKF relies on a state evolution model, which is a stochastic model for the one-frame evolution of the state $\boldsymbol{x}_t$, i.e., the transition from $\boldsymbol{x}_{t-1}$ to $\boldsymbol{x}_t$. For a basic formulation of the state evolution model, we temporarily assume that $N_t = N$ and $L_t = 2M + 2N + 7 = L$ are constant; an extension to time-varying N_t and L_t will be given in Section V-A. We choose a linear state evolution model given by

$$\boldsymbol{x}_t = \boldsymbol{\Phi}\boldsymbol{x}_{t-1} + \boldsymbol{u}_t, \quad t = 1, 2, \ldots. \quad (7)$$

Here, $\boldsymbol{\Phi}$ is an $L \times L$ block-diagonal matrix defined as

$$\boldsymbol{\Phi} \triangleq \operatorname{diag}\{\mathbf{I}_{2M}, \boldsymbol{\Phi}_\varphi, \boldsymbol{\Phi}_\gamma, 1, \mathbf{I}_2, \underbrace{\phi_\delta, 1, \ldots, \phi_\delta, 1}_{N \text{ times } ``\phi_\delta, 1"}\}, \quad (8)$$

where

$$\boldsymbol{\Phi}_\varphi \triangleq \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\Phi}_\gamma \triangleq \begin{pmatrix} 1+\phi_\gamma & -\phi_\gamma \\ 1 & 0 \end{pmatrix},$$

and $\phi_\gamma, \phi_\delta \in (0, 1)$. Furthermore, the driving process $\boldsymbol{u}_t$ is an independent and identically distributed (iid) L-dimensional zero-mean process with covariance matrix

$$\boldsymbol{C}_u \triangleq \operatorname{diag}\{\sigma_{ab}^2\mathbf{I}_{2M}, \boldsymbol{\Sigma}_\varphi, \boldsymbol{\Sigma}_\gamma, 0, \sigma_c^2\mathbf{I}_2, \underbrace{\sigma_\delta^2, \sigma_\theta^2, \ldots, \sigma_\delta^2, \sigma_\theta^2}_{N \text{ times } ``\sigma_\delta^2, \sigma_\theta^2"}\}, \quad (9)$$

where

$$\boldsymbol{\Sigma}_\varphi \triangleq \sigma_\varphi^2 \begin{pmatrix} 1/4 & 1/2 \\ 1/2 & 1 \end{pmatrix}, \quad \boldsymbol{\Sigma}_\gamma \triangleq \sigma_\gamma^2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

This model summarizes the following individual state evolution models:

$$\begin{aligned}
a_t^{(m)} &= a_{t-1}^{(m)} + (\boldsymbol{u}_t)_m, \quad m = 1, 2, \ldots, M, \\
b_t^{(m)} &= b_{t-1}^{(m)} + (\boldsymbol{u}_t)_{M+m}, \quad m = 1, 2, \ldots, M, \\
\boldsymbol{\varphi}_t &= \boldsymbol{\Phi}_\varphi \boldsymbol{\varphi}_{t-1} + (\boldsymbol{u}_t)_{2M+1,\,2M+2}, \\
\boldsymbol{\gamma}_t &= \boldsymbol{\Phi}_\gamma \boldsymbol{\gamma}_{t-1} + (\boldsymbol{u}_t)_{2M+3,\,2M+4}, \\
R_t &= R_{t-1}, \\
\boldsymbol{c}_t &= \boldsymbol{c}_{t-1} + (\boldsymbol{u}_t)_{2M+6,\,2M+7}, \\
\delta_t^{(n)} &= \phi_\delta \delta_{t-1}^{(n)} + (\boldsymbol{u}_t)_{2M+6+2n}, \quad n = 1, 2, \ldots, N, \\
\theta_t^{(n)} &= \theta_{t-1}^{(n)} + (\boldsymbol{u}_t)_{2M+7+2n}, \quad n = 1, 2, \ldots, N.
\end{aligned}$$

Here $(\cdot)_i$ denotes the i-th element of the vector in parentheses and $(\cdot)_{i,j}$ denotes the vector comprising the i-th and j-th elements of the vector in parentheses. Note that the state evolution models for $a_t^{(m)}$ and $b_t^{(m)}$ $(m = 1, 2, \ldots, M)$, $\boldsymbol{c}_t$, and $\theta_t^{(n)}$ $(n = 1, 2, \ldots, N)$ are random walk models, $\boldsymbol{\varphi}_t$ is modeled by a constant velocity model, the state evolution model for γ_t is a second-order auto-regressive (AR(2)) model, R_t is modeled as constant (because the variance of $(\boldsymbol{u}_t)_{2M+5}$ is zero), and $\delta_t^{(n)}$ is modeled by a first-order auto-regressive (AR(1)) model. We use AR models, rather than simple random walk models, for γ_t and $\delta_t^{(n)}$ because the AR model effectively limits the range of feasible values. This helps prevent the FPs from drifting away from the CCA wall, and is one reason why our method is able to compensate the feature drift resulting from the optical flow algorithm. (The other reason is the use of the results of the circle localization algorithm of [22] as an additional input to the UKF, as explained in Section III-D.) For a larger variance of $(\boldsymbol{u}_t)_l$, the respective state component $(\boldsymbol{x}_t)_l$ tends to change more rapidly and less smoothly.

D. MEASUREMENT MODEL

The UKF also relies on a measurement model in addition to the state evolution model. The measurement model is a stochastic model for the dependence of the observed measurements on the state $\boldsymbol{x}_t$. We recall from Sections II and III that our measurements are given by the preliminary FP position estimates $\tilde{\boldsymbol{y}}_t^{(n)}$, $n = 1, 2, \ldots, N$ contained in the preliminary FP position list $\tilde{\boldsymbol{Y}}_t$ and by the preliminary CCA circle parameter estimates $\tilde{\boldsymbol{c}}_t$ and $\tilde{r}_t$.

We model the preliminary FP position estimates $\tilde{\boldsymbol{y}}_t^{(n)}$ as the CCA circle center point $\boldsymbol{c}_t$ plus noisy versions of the true FP position vectors $\boldsymbol{y}_t^{(n)} = \left(\rho_t^{(n)} \cos\theta_t^{(n)} \;\; \rho_t^{(n)} \sin\theta_t^{(n)}\right)^{\mathrm{T}}$ (cf. (3)), i.e.,

$$\tilde{\boldsymbol{y}}_t^{(n)} = \boldsymbol{c}_t + \boldsymbol{y}_t^{(n)} + \boldsymbol{v}_t^{(n)}, \quad n = 1, 2, \ldots, N. \tag{10}$$

Here, $\rho_t^{(n)}$ is given by Equation (4), i.e., $\rho_t^{(n)} = r_t + \tilde{\delta}^{(n)} + \delta_t^{(n)}$, in which, in turn, r_t is given by Equation (2) with R formally replaced by R_t. Thus, expression (10) involves the state components (cf. (5)) $\boldsymbol{a}_t$, $\boldsymbol{b}_t$, φ_t, γ_t, R_t, $\boldsymbol{c}_t$, and $\boldsymbol{p}_t^{(n)} = (\delta_t^{(n)} \;\; \theta_t^{(n)})^{\mathrm{T}}$ for $n = 1, 2, \ldots, N$; note that φ_t and $\theta_t^{(n)}$ enter in a nonlinear manner. Furthermore, the measurement noise processes $\boldsymbol{v}_t^{(n)}$ are mutually independent, iid, zero-mean, two-dimensional processes with covariance matrix $\boldsymbol{C}_{v^{(n)}} \triangleq \mathrm{diag}\{\sigma_{v,1}^2, \sigma_{v,2}^2\}$.

In a similar manner, the preliminary CCA circle parameter estimates $\tilde{\boldsymbol{c}}_t$ and $\tilde{r}_t$ are modeled as noisy versions of the true CCA circle parameters $\boldsymbol{c}_t$ and r_t, respectively, i.e.,

$$\tilde{\boldsymbol{c}}_t = \boldsymbol{c}_t + \boldsymbol{v}_t^{(\mathrm{c})}, \quad \tilde{r}_t = r_t + v_t^{(\mathrm{r})}, \tag{11}$$

where $\boldsymbol{v}_t^{(\mathrm{c})}$ and $v_t^{(\mathrm{r})}$ are mutually independent, iid, zero-mean processes with covariance $\boldsymbol{C}_{v^{(\mathrm{c})}} \triangleq \sigma_{v^{(\mathrm{c})}}^2 \mathbf{I}_2$ and variance $\sigma_{v^{(\mathrm{r})}}^2$, respectively. The expressions (11) involve the state components (cf. (5)) $\boldsymbol{a}_t$, $\boldsymbol{b}_t$, φ_t, γ_t, R_t, and $\boldsymbol{c}_t$.

Finally, the overall measurement vector $\boldsymbol{z}_t$ comprises all the measurements, i.e.,

$$\boldsymbol{z}_t \triangleq \left(\tilde{\boldsymbol{y}}_t^{(1)\mathrm{T}} \; \tilde{\boldsymbol{y}}_t^{(2)\mathrm{T}} \cdots \tilde{\boldsymbol{y}}_t^{(N)\mathrm{T}} \; \tilde{\boldsymbol{c}}_t^{\mathrm{T}} \;\; \tilde{r}_t\right)^{\mathrm{T}}.$$

We can now summarize the measurement models (10) and (11) into an overall measurement model

$$\boldsymbol{z}_t = \boldsymbol{g}(\boldsymbol{x}_t) + \boldsymbol{v}_t, \tag{12}$$

where

$$\boldsymbol{g}(\boldsymbol{x}_t) \triangleq \left(\boldsymbol{c}_t^{\mathrm{T}} + \boldsymbol{y}_t^{(1)\mathrm{T}} \; \boldsymbol{c}_t^{\mathrm{T}} + \boldsymbol{y}_t^{(2)\mathrm{T}} \cdots \boldsymbol{c}_t^{\mathrm{T}} + \boldsymbol{y}_t^{(N)\mathrm{T}} \; \boldsymbol{c}_t^{\mathrm{T}} \;\; r_t\right)^{\mathrm{T}} \tag{13}$$

and $\boldsymbol{v}_t \triangleq \left(\boldsymbol{v}_t^{(1)\mathrm{T}} \cdots \boldsymbol{v}_t^{(N)\mathrm{T}} \; \boldsymbol{v}_t^{(\mathrm{c})\mathrm{T}} \; v_t^{(\mathrm{r})}\right)^{\mathrm{T}}$. The covariance matrix of $\boldsymbol{v}_t$ is given by

$$\boldsymbol{C}_v \triangleq \mathrm{diag}\big\{\underbrace{\sigma_{v,1}^2, \sigma_{v,2}^2, \ldots, \sigma_{v,1}^2, \sigma_{v,2}^2}_{N \text{ times } ``\sigma_{v,1}^2, \sigma_{v,2}^2"}, \sigma_{v^{(\mathrm{c})}}^2, \sigma_{v^{(\mathrm{c})}}^2, \sigma_{v^{(\mathrm{r})}}^2\big\}.$$

V. UKF OPERATION

The UKF [19] is a suboptimal sequential Bayes filter that calculates at each frame time $t = 2, 3, \ldots$ estimates of the posterior mean $\mathrm{E}\{\boldsymbol{x}_t|\boldsymbol{z}_{1:t}\}$ and posterior covariance $\mathrm{Cov}\{\boldsymbol{x}_t|\boldsymbol{z}_{1:t}\}$ of the state $\boldsymbol{x}_t$, where $\boldsymbol{z}_{1:t} \triangleq (\boldsymbol{z}_1^{\mathrm{T}} \cdots \boldsymbol{z}_t^{\mathrm{T}})^{\mathrm{T}}$. These estimates will be denoted by $\hat{\boldsymbol{x}}_t$ and $\boldsymbol{P}_t$, respectively; note that $\hat{\boldsymbol{x}}_t \approx \mathrm{E}\{\boldsymbol{x}_t|\boldsymbol{z}_{1:t}\}$ provides an estimate of $\boldsymbol{x}_t$. We use a UKF, rather than simply a Kalman filter [28, Ch. 2], because of the nonlinearity of our measurement model (12). Also, in comparison to the extended Kalman filter [28, Ch. 2], the UKF usually provides a better estimation accuracy while its computational complexity is of the same order [19]. We use a slightly modified version of the UKF algorithm that includes an FP allocation/deallocation step to account for the time-varying number of state variables. Each UKF recursion thus consists of the following steps, which are described further below: FP allocation/deallocation, state prediction, calculation of sigma-points, measurement prediction, update, and estimation. These steps are performed at each frame time $t = 2, 3, \ldots$.

A. FP ALLOCATION/DEALLOCATION

According to (5), the inclusion of a new FP n in the previous FP position list $\hat{\boldsymbol{Y}}_{t-1}$—as described in Section III-A—entails a corresponding allocation (insertion) of the FP position parameter vector $\boldsymbol{p}_{t-1}^{(n)}$ in the state vector $\boldsymbol{x}_{t-1}$. This implies an analogous allocation of an estimate of $\boldsymbol{p}_{t'}^{(n)}$ in the state estimate vector $\hat{\boldsymbol{x}}_{t'}$ for $t' = t-1, t, \ldots$ We initialize this estimate as (cf. (6)) $\hat{\boldsymbol{p}}_{t-1}^{(n)} \triangleq (\hat{\delta}_{t-1}^{(n)} \;\; \hat{\theta}_{t-1}^{(n)})^{\mathrm{T}} = \left(0 \;\; \tan^{-1}(\xi_{t-1}^{(2)}/\xi_{t-1}^{(1)})\right)^{\mathrm{T}}$, where $\xi_{t-1}^{(1)}$ and $\xi_{t-1}^{(2)}$ are the Cartesian coordinates of $\boldsymbol{\xi}_{t-1} \triangleq \hat{\boldsymbol{y}}_{t-1}^{(n)} - \hat{\boldsymbol{c}}_{t-1}$. Here, $\hat{\boldsymbol{y}}_{t-1}^{(n)}$ is the initial FP position obtained in the FP inclusion step (see Section III-A), and $\hat{\boldsymbol{c}}_{t-1}$ is the previous estimate of the circle center. Furthermore, corresponding variances are allocated as two additional elements on the diagonal of the approximate covariance matrix $\boldsymbol{P}_{t'}$ for $t' = t-1, t, \ldots$. These variances are initialized in $\boldsymbol{P}_{t-1}$ as s_δ^2 and s_θ^2, which are parameters expressing our a priori

uncertainty about $\delta_{t-1}^{(n)}$ and $\theta_{t-1}^{(n)}$, respectively. Finally, we calculate an initial approximation of the radial deviation of the new FP from the previous estimate of the CCA circle as $\tilde{\delta}^{(n)} = \|\boldsymbol{\xi}_{t-1}\|_2 - \hat{r}_{t-1}$, where $\|\cdot\|_2$ denotes the ℓ_2 (Euclidean) norm; we recall that $\tilde{\delta}^{(n)}$ is used in the FP radius model (4) for $t' = t-1, t, \ldots$.

Similar allocation operations are applied to the transition matrix $\boldsymbol{\Phi}$ in (8) and the covariance matrix $\boldsymbol{C}_u$ in (9), which thereby become time-varying. More specifically, the values ϕ_δ and 1 are inserted as additional diagonal elements of $\boldsymbol{\Phi}_{t'}$, and the values σ_δ^2 and σ_θ^2 are inserted as additional diagonal elements of $\boldsymbol{C}_{u,t'}$, in both cases for $t' = t-1, t, \ldots$.

Conversely, the exclusion of FP n at time t—as described in Section III-C—entails a deallocation (removal) of the corresponding FP position parameter vector $\boldsymbol{p}_{t-1}^{(n)}$ in the state vector $\boldsymbol{x}_{t-1}$. This implies analogous deallocation operations in $\hat{\boldsymbol{x}}_{t'}$, $\boldsymbol{P}_{t'}$, $\boldsymbol{\Phi}_{t'}$, and $\boldsymbol{C}_{u,t'}$ for $t' = t-1, t, \ldots$.

B. STATE PREDICTION

In the state prediction step of the UKF recursion at frame time t, a "predicted" state mean $\hat{\boldsymbol{x}}_{t|t-1}$ and covariance $\boldsymbol{P}_{t|t-1}$ are calculated by propagating the previous mean $\hat{\boldsymbol{x}}_{t-1}$ and covariance $\boldsymbol{P}_{t-1}$ through the linear state evolution model (7). This results in

$$\hat{\boldsymbol{x}}_{t|t-1} = \boldsymbol{\Phi}_t \hat{\boldsymbol{x}}_{t-1}, \qquad \boldsymbol{P}_{t|t-1} = \boldsymbol{\Phi}_t \boldsymbol{P}_{t-1} \boldsymbol{\Phi}_t^{\mathrm{T}} + \boldsymbol{C}_{u,t}.$$

C. CALCULATION OF SIGMA-POINTS

Next, so-called sigma-points are used to approximate the propagation of the predicted mean $\hat{\boldsymbol{x}}_{t|t-1}$ and covariance $\boldsymbol{P}_{t|t-1}$ through the nonlinear measurement model (12). According to [19], for state vector dimension L_t, there are $2L_t + 1$ sigma-points $\hat{\boldsymbol{x}}_t^{(l)}$ and corresponding weights $w_t^{(l)}$, $l = 0, 1, \ldots, 2L_t$, which are calculated as

$$\hat{\boldsymbol{x}}_t^{(l)} = \begin{cases} \hat{\boldsymbol{x}}_{t|t-1}, & l = 0, \\ \hat{\boldsymbol{x}}_{t|t-1} + (\boldsymbol{\Sigma}_t^{1/2})_l, & l = 1, 2, \ldots, L_t, \\ \hat{\boldsymbol{x}}_{t|t-1} - (\boldsymbol{\Sigma}_t^{1/2})_{l-L_t}, & l = L_t{+}1, L_t{+}2, \ldots, 2L_t, \end{cases}$$

$$w_t^{(l)} = \begin{cases} 1/3, & l = 0, \\ \dfrac{1 - w_t^{(0)}}{2L_t}, & l = 1, 2, \ldots, 2L_t. \end{cases}$$

Here, $(\boldsymbol{\Sigma}_t^{1/2})_l$ denotes the l-th column of $\boldsymbol{\Sigma}_t^{1/2}$, which is a square root of the matrix $\boldsymbol{\Sigma}_t \triangleq \frac{L_t}{1-w_t^{(0)}} \boldsymbol{P}_{t|t-1}$, i.e., any square matrix satisfying $\left(\boldsymbol{\Sigma}_t^{1/2}\right)^{\mathrm{T}} \boldsymbol{\Sigma}_t^{1/2} = \boldsymbol{\Sigma}_t$. We note that $w_t^{(0)}$ can be chosen differently from 1/3, as discussed in [19].

D. MEASUREMENT PREDICTION

The sigma-points are now propagated through the (noiseless) nonlinear measurement model (12), i.e.,

$$\hat{\boldsymbol{z}}_t^{(l)} = \boldsymbol{g}(\hat{\boldsymbol{x}}_t^{(l)}), \quad l = 0, 1, \ldots, 2L_t.$$

Then, a "predicted" measurement $\hat{\boldsymbol{z}}_{t|t-1}$ and an innovation covariance matrix $\boldsymbol{S}_{t|t-1}$ are calculated as

$$\hat{\boldsymbol{z}}_{t|t-1} = \sum_{l=0}^{2L_t} w_t^{(l)} \hat{\boldsymbol{z}}_t^{(l)},$$

$$\boldsymbol{S}_{t|t-1} = \sum_{l=0}^{2L_t} w_t^{(l)} (\hat{\boldsymbol{z}}_t^{(l)} - \hat{\boldsymbol{z}}_{t|t-1})(\hat{\boldsymbol{z}}_t^{(l)} - \hat{\boldsymbol{z}}_{t|t-1})^{\mathrm{T}} + \boldsymbol{C}_v.$$

E. UPDATE

Next, we perform the update step of the ordinary Kalman filter [28, Ch. 2], i.e.,

$$\hat{\boldsymbol{x}}_t = \hat{\boldsymbol{x}}_{t|t-1} + \boldsymbol{G}_t(\boldsymbol{z}_t - \hat{\boldsymbol{z}}_{t|t-1}),$$
$$\boldsymbol{P}_t = \boldsymbol{P}_{t|t-1} - \boldsymbol{G}_t \boldsymbol{S}_{t|t-1} \boldsymbol{G}_t^{\mathrm{T}}.$$

Here $\boldsymbol{G}_t$ is a sigma-point approximation of the Kalman gain matrix, which is given by

$$\boldsymbol{G}_t = \left(\sum_{l=0}^{2L_t} w_t^{(l)} (\hat{\boldsymbol{x}}_t^{(l)} - \hat{\boldsymbol{x}}_{t|t-1})(\hat{\boldsymbol{z}}_t^{(l)} - \hat{\boldsymbol{z}}_{t|t-1})^{\mathrm{T}} \right) \boldsymbol{S}_{t|t-1}^{-1}.$$

F. ESTIMATION

Evaluating $\boldsymbol{g}(\cdot)$ in (13) for the state estimate $\hat{\boldsymbol{x}}_t$ yields

$$\boldsymbol{g}(\hat{\boldsymbol{x}}_t) = \left(\hat{\boldsymbol{c}}_t^{\mathrm{T}} + \hat{\boldsymbol{y}}_t^{(1)\mathrm{T}} \; \hat{\boldsymbol{c}}_t^{\mathrm{T}} + \hat{\boldsymbol{y}}_t^{(2)\mathrm{T}} \cdots \; \hat{\boldsymbol{c}}_t^{\mathrm{T}} + \hat{\boldsymbol{y}}_t^{(N_t)\mathrm{T}} \; \hat{\boldsymbol{c}}_t^{\mathrm{T}} \; \hat{r}_t\right)^{\mathrm{T}}$$

We use the vectors $\hat{\boldsymbol{y}}_t^{(n)}$ as the final estimates of the FP positions $\boldsymbol{y}_t^{(n)}$, for $n = 1, 2, \ldots, N_t$; these estimates constitute the final FP position estimate list $\hat{\boldsymbol{Y}}_t$. Similarly, we use $\hat{\boldsymbol{c}}_t$ and $\hat{r}_t$ as the final estimates of the CCA circle parameters $\boldsymbol{c}_t$ and r_t. Note that the FP position estimate list $\hat{\boldsymbol{Y}}_t$ and the center point estimate $\hat{\boldsymbol{c}}_t$ are used only internally, namely, as an input to the FP inclusion step at the next frame time $t{+}1$, as explained in Sections II and III-A. On the other hand, $\hat{r}_t$ forms the output of the overall method.

G. INITIALIZATION

At frame time $t = 1$, no FPs are allocated yet, i.e., $N_1 = 0$. We initialize the mean $\hat{\boldsymbol{x}}_1$ and covariance matrix $\boldsymbol{P}_1$ as (cf. (5))

$$\hat{\boldsymbol{x}}_1 = \big(\underbrace{0 \; 0 \cdots 0}_{2M+1 \text{ times "0"}} \; \dot{\varphi}_1 \; 0 \; 0 \; \hat{r}_1 \; \hat{\boldsymbol{c}}_1^{\mathrm{T}} \big)^{\mathrm{T}},$$
$$\boldsymbol{P}_1 = \mathrm{diag}\big\{ s_{ab}^2 \mathbf{I}_{2M}, s_\varphi^2 \mathbf{I}_2, s_\gamma^2 \mathbf{I}_2, s_R^2, s_c^2 \mathbf{I}_2 \big\}.$$

Here, $\dot{\varphi}_1$ is an initial value[1] of the time-varying fundamental frequency $\dot{\varphi}_t$, and $\hat{r}_1$ and $\hat{\boldsymbol{c}}_1$ are initial CCA circle parameter estimates that are calculated by the circle localization algorithm of [22]. The variances s_{ab}^2, s_φ^2, s_γ^2, s_R^2, and s_c^2 are parameters expressing our a priori uncertainty about the Fourier coefficients $a_1^{(m)}$ and $b_1^{(m)}$ ($m = 1, 2, \ldots, M$) as well as about $\boldsymbol{\varphi}_1$, $\boldsymbol{\gamma}_1$, R_1, and $\boldsymbol{c}_1$, respectively.

[1] The fundamental instantaneous phase φ_t and the time-varying fundamental frequency $\dot{\varphi}_t$ are part of the state $\boldsymbol{x}_t$—see (5)—and thus are estimated (tracked) along with the other state components. Therefore, in contrast to [23], the fundamental frequency does not have to be precisely known in advance. However, a reasonable initialization of $\dot{\varphi}_t$ at $t = 1$ is required for the method to converge.

TABLE 1. Parameters of the proposed method that were used in the experiments (common for all ultrasound sequences). One pixel (px) corresponds to 70 μm.

Parameter	Δ_r	d	J	α	M	ϕ_γ	ϕ_δ	σ_{ab}^2
Value	1 px	5 px	21 px	1.5	9	0.998	0.6	10^{-7} px^2
Parameter	σ_c^2	σ_δ^2	σ_θ^2	σ_φ^2	σ_γ^2	$\sigma_{v(c)}^2$	$\sigma_{v(r)}^2$	s_δ^2
Value	10^{-2} px^2	10^{-7} px^2	10^{-8} rad^2	10^{-6} rad^2	10^{-5} px^2	$2 \cdot 10^{-2}$ px^2	$2 \cdot 10^{-2}$ px^2	10^{-2} px^2
Parameter	s_θ^2	s_{ab}^2	s_φ^2	s_γ^2	s_R^2	s_c^2		
Value	10^{-2} rad^2	10^{-1} px^2	10^{-5} rad^2	1 px^2	1 px^2	0.1 px^2		

VI. NUMERICAL RESULTS AND DISCUSSION

In this section, we present numerical results demonstrating the performance of the proposed method in comparison to an improved version of the state-of-the-art reference method described in [18]. The latter method was chosen as a reference method because, similarly to the proposed method, it performs an explicit feature drift compensation and is thus suited to the case of long ultrasound sequences.

A. BASIC SIMULATION SETUP

We considered the following types of experimental data. *Data set S* was created synthetically through a time-varying transformation of a real ultrasound image, and was used to evaluate the robustness of the proposed method to additive and multiplicative noise. *Data set F* was created synthetically by means of the Field II simulation program [29], [30], and was used to determine the method's performance in the presence of speckle decorrelation. Finally, we validated our method on a real ultrasound sequence. More detailed descriptions of these experimental data will be provided in later subsections.

TABLE 2. Parameters of the proposed method that were used in the experiments (not common for all ultrasound sequences). One pixel (px) corresponds to 70 μm.

	Parameter	τ_h (-)	$\sigma_{v,1}^2$ (px^2)	$\sigma_{v,2}^2$ (px^2)	$\dot{\varphi}_1$ (rad)
Sequence	S_{25}^{+}	0.1	$5 \cdot 10^{-3}$	10^{-3}	0.29
	S_{15}^{+}	0.1	$5 \cdot 10^{-3}$	10^{-3}	0.29
	S_{25}^{*}	0.1	$5 \cdot 10^{-3}$	10^{-3}	0.29
	S_{15}^{*}	0.1	$5 \cdot 10^{-3}$	$5 \cdot 10^{-3}$	0.29
	$\mathrm{F}_{0,2,5}$	0.3	$4 \cdot 10^{-2}$	$1 \cdot 10^{-2}$	0.29
	Real	0.1	10^{-3}	10^{-3}	0.31

The parameters of the proposed method that we used in our experiments are listed in Tables 1 and 2. In the course of our experiments, we observed that the circle localization algorithm used in the reference method [18] performs poorly on our data. Therefore, for a fair comparison of tracking accuracy, we modified the reference method such that it uses the circle localization algorithm used in our method (see Section III-D). Furthermore, in both the proposed method and the reference method, we included a correction of the radius estimation bias exhibited by the circle localization algorithm. More specifically, the preliminary circle radius $\tilde{r}_t$, used subsequently by the UKF, was obtained by subtracting a constant from the radius estimate computed by the circle localization algorithm. For data sets S and F, we estimated this constant from sequence S_∞ and F_0, respectively, by averaging over all frames the error of the radius estimates produced by the circle localization method. Note that we were able to determine this error because for data sets S and F the ground truth is available.

The proposed method was implemented in MATLAB on an Intel(R) Core(TM) i5-7500 CPU with a base frequency of 3.40 GHz, without the use of multi-threading. The processing of one frame always took less than 40 ms. In particular, the processing of a single frame of 348 $\times$ 280 px from dataset S (discussed later) took on average 26 ms.

B. MOTION MODEL

To generate the simulated sequences, we modeled the CCA wall transverse section by a circle with a time-varying radius r_t that is the sum of a pulse wave component and a breathing component. For the pulse wave component, we employed the parametric model proposed in [31], whose parameters we chose such that they conformed to radius estimates obtained from real ultrasound sequences. In order to introduce smooth variations in the pulse rate, we frequency-modulated the pulse wave by a sine function with a period of half a minute. During one period of the modulating sine function, the heart rate varies between 70 and 110 beats per minute. The breathing component was modeled as the absolute value of a sine function whose frequency is much lower than the assumed heart rate. Further details of the motion model and its parameters are provided in [18].

C. PERFORMANCE METRICS

For measuring the accuracy of radius estimation in those cases where a ground truth r_t is available (data sets S and F), we use the root mean square error (RMSE). Assuming K simulation runs, each producing an estimated radius sequence $\hat{r}_{t,k}$ ($t = 1, 2, \ldots, T$; $k = 1, 2, \ldots, K$), the RMSE is defined as the square root of the average of the squared error $(\hat{r}_{t,k} - r_t)^2$ taken over all frames t and all simulation

runs k, i.e.,

$$\mathrm{RMSE} \triangleq \sqrt{\frac{1}{K}\sum_{k=1}^{K}\mathrm{RMSE}_k^2},$$

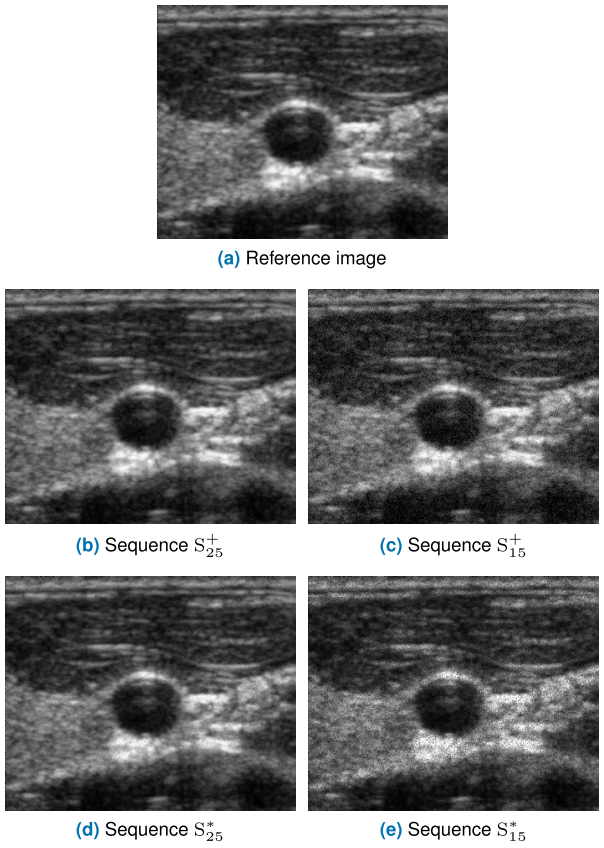
where $\mathrm{RMSE}_k \triangleq \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{r}_{t,k}-r_t)^2}$. We also consider the inter-realization empirical standard deviation of RMSE_k, which is denoted SD and defined as

$$\mathrm{SD} \triangleq \sqrt{\frac{1}{K}\sum_{k=1}^{K}(\mathrm{RMSE}_k-\overline{\mathrm{RMSE}_k})^2},$$

where $\overline{\mathrm{RMSE}_k} \triangleq \frac{1}{K}\sum_{k=1}^{K}\mathrm{RMSE}_k$.

D. SYNTHETIC DATA SET S

Data set S consists of a synthetic ultrasound sequence, referred to as S_∞, as well as $K = 100$ realizations of four "noisy" synthetic ultrasound sequences, referred to as S_{25}^{+}, S_{15}^{+}, S_{25}^{*}, and S_{15}^{*}. Each sequence comprises $T = 960$ ultrasound frames. With a frame rate of 32 fps, this corresponds to 30 seconds of continuous recording.



(a) Reference image

(b) Sequence S_{25}^{+} (c) Sequence S_{15}^{+}

(d) Sequence S_{25}^{*} (e) Sequence S_{15}^{*}

FIGURE 4. (a) Reference image (simultaneously first frame of sequence S_∞), (b)–(e) first frame of realizations of the synthetic sequences S_{25}^{+}, S_{15}^{+}, S_{25}^{*}, and S_{15}^{*}, respectively.

Sequence S_∞ was derived from a reference image, shown in Figure 4(a), through a time-varying deformation. The reference image was chosen as a real ultrasound image of the CCA transverse section of a healthy subject, which was acquired as described in Section VI-F. The frames of S_∞ were then obtained by deforming the reference image in accordance with the motion model described in Section VI-B. More specifically, the CCA circle radius for each frame was calculated from the motion model, a displacement field (describing the tissue displacement at each pixel of the reference image) was calculated from the CCA circle radius, and the reference

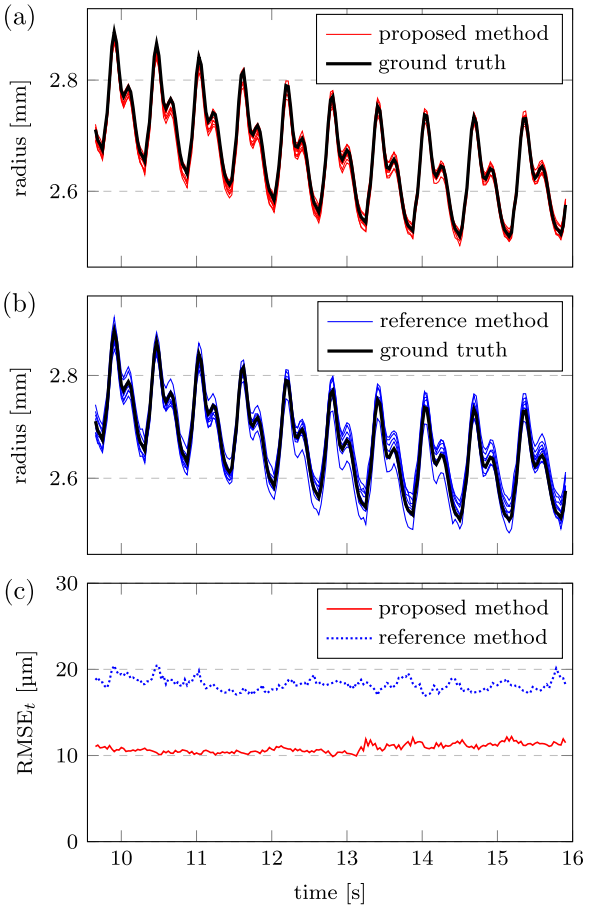


FIGURE 5. Results of the proposed method and the reference method for ten realizations of S_{15}^{+} during 200 frames: (a) Radius estimates obtained with the proposed method, (b) radius estimates obtained with the reference method, (c) corresponding time-varying RMSEs.

TABLE 3. RMSE and SD (both in μm) for data set S. Lower values are better.

Sequence	RMSE		SD	
	Proposed method	Reference method	Proposed method	Reference method
S_∞	12.01	28.03	—	—
S_{25}^{+}	13.72	17.67	2.13	9.78
S_{15}^{+}	13.77	18.70	5.01	5.56
S_{25}^{*}	14.59	40.68	1.03	13.86
S_{15}^{*}	27.41	47.51	3.42	11.30
average	16.30	30.52	2.90	10.13

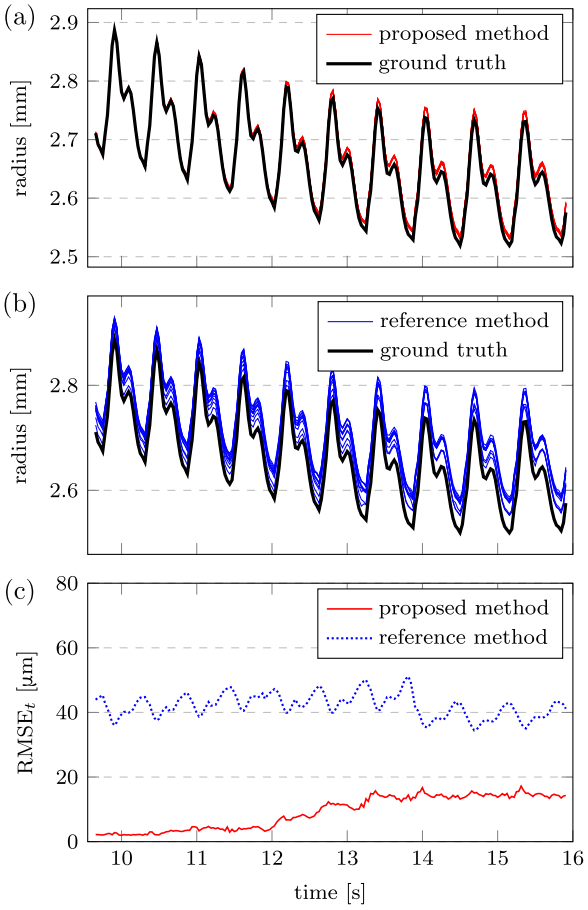


FIGURE 6. Results of the proposed method and the reference method for ten realizations of S^*_{15} during 200 frames: (a) Radius estimates obtained with the proposed method, (b) radius estimates obtained with the reference method, (c) corresponding time-varying RMSEs.

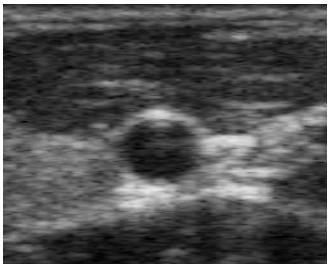


FIGURE 7. First frame of the synthetic sequence F_0.

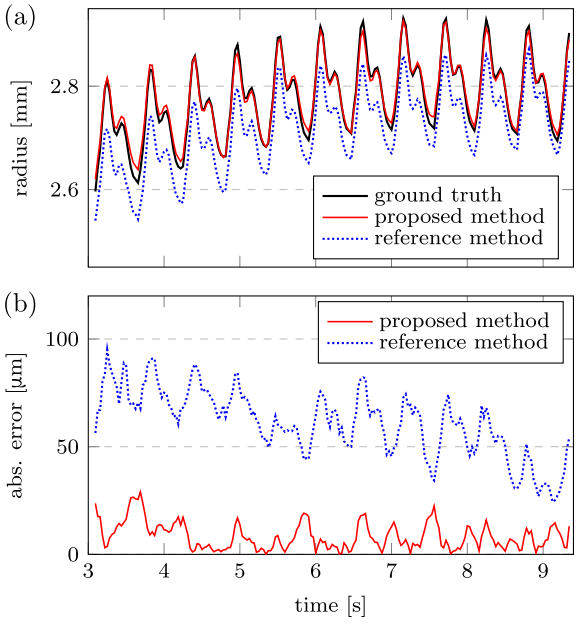


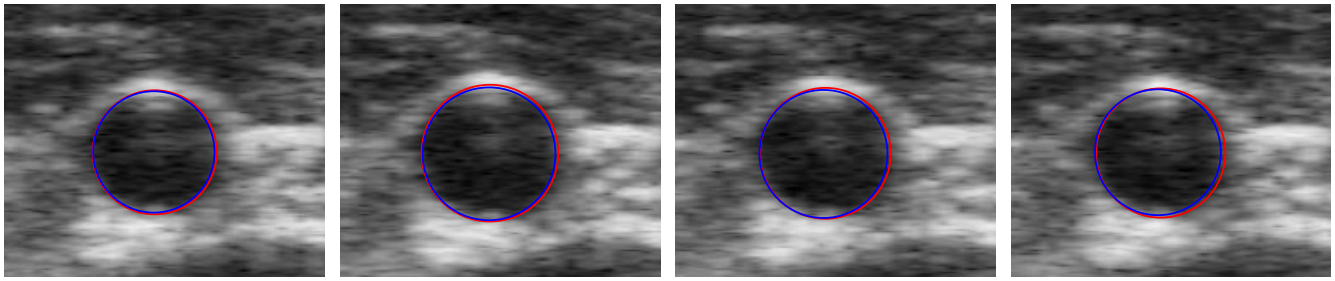
FIGURE 8. Results of the proposed method and the reference method for sequence F_2 during 200 frames: (a) Radius estimates, (b) absolute value of the corresponding estimation errors.

TABLE 4. RMSE (in μm) for data set F. Lower values are better.

Sequence	Proposed method	Reference method
F_0	17.76	65.44
F_2	17.15	53.22
F_5	16.90	33.93
average	17.27	50.86

image was transformed according to the displacement field. A more detailed description of this procedure can be found in [18]. Finally, the noisy sequences of data set S were obtained by corrupting S_∞ by additive Gaussian noise (in the case of S^+_{25} and S^+_{15}) or by multiplicative Gaussian noise (in the case of S^*_{25} and S^*_{15}). The subscript, 25 or 15, indicates the signal-to-noise ratio in decibels. Similarly to [17], we use noisy sequences to evaluate the robustness of the methods within a simple, controlled, and reproducible setting. In real B-mode (log-compressed) images, the noise characteristics would be more complex than the additive and multiplicative Gaussian characteristics employed here, although additive noise would still be the most prominent component [32]. We created $K = 100$ realizations of each of the four sequence types S^+_{25}, S^+_{15}, S^*_{25}, and S^*_{15} by using 100 realizations of the respective noise process (i.e., the same ultrasound sequence S_∞ was corrupted by different noise realizations). Note that for S_∞, formally, $K = 1$. For each of the noisy sequences S^+_{25}, S^+_{15}, S^*_{25}, and S^*_{15}, the first frame of one realization is shown in Figures 4(b)–(e), respectively.

For each sequence or sequence type, we calculated RMSE and SD for the radius estimates that we obtained with the proposed method and with the reference method [18]. Except for S_∞, this calculation was based on the $K = 100$ realizations of the respective noise process. The results are presented in Table 3, along with the average RMSE and SD taken over all sequences. One can see that the proposed method always outperforms the reference method in terms of RMSE. In addition, the SD results show that the estimates produced by the proposed method are much more consistent across multiple realizations.



(a) Frame 1 (b) Frame 200 (c) Frame 400 (d) Frame 600

FIGURE 9. Estimates of the CCA wall circle obtained for four frames of sequence F_2 with the proposed method (red) and the reference method (blue): (a) frame 1 (0 s), (b) frame 200 (6.25 s), (c) frame 400 (12.50 s), and (d) frame 600 (18.75 s). The ground truth is not shown since it typically overlaps with the estimates produced by the proposed method.

In Figure 5, we show the individual estimated radius sequences $\hat{r}_{t,k}$ obtained with either method for ten realizations of S_{15}^{+}. The figure also shows the time-varying RMSEs, i.e., $\text{RMSE}_t \triangleq \sqrt{\frac{1}{K}\sum_{k=1}^{K}(\hat{r}_{t,k} - r_t)^2}$. It can be observed that the estimates obtained with the proposed method are almost always closer to the ground truth. The large variability of the results of the reference method (across the individual realizations) is primarily caused by feature drift [18].

Figure 6 shows analogous results obtained for S_{15}^{*}. One can observe that the proposed method exhibits an even larger performance gain relative to the reference method. This is due to an even stronger feature drift, which is effectively compensated by the proposed method but less well by the reference method. We note that the increase in the RMSE of the proposed method that is observed in Figure 6(c) between 12 s and 14 s is caused by a temporary circle localization error; the RMSE starts decreasing again around 22 s, which is however not shown in Figure 6(c).

E. SYNTHETIC DATA SET F

Data set F consists of three synthetic ultrasound sequences referred to as F_0, F_2, and F_5. Each sequence again comprises $T = 960$ ultrasound frames or 30 seconds of continuous recording (assuming a frame rate of 32 fps).

Sequence F_0 was created by means of the Field II simulation program [29], [30]. First, a three-dimensional tissue model (reference phantom) was generated. This reference phantom consists of infinitesimally small objects called scatterers, whose positions are uniformly distributed in the phantom with a density of about $50\,\text{mm}^{-3}$. With the resolution cell volume given by $0.2\,\text{mm}^3$, about ten scatterers occupy one resolution cell. Each scatterer has an amplitude that is randomly drawn from a zero-mean Gaussian distribution. The standard deviation of that distribution was chosen equal to the intensity of the real ultrasound image that was used to create data set S (see Figure 4(a)), at the position corresponding to the scatterer's lateral and axial coordinates relative to a simulated ultrasound probe. The simulated ultrasound probe consisted of 192 elements, 72 of which were active during each single-line scan. The center frequency of the probe was

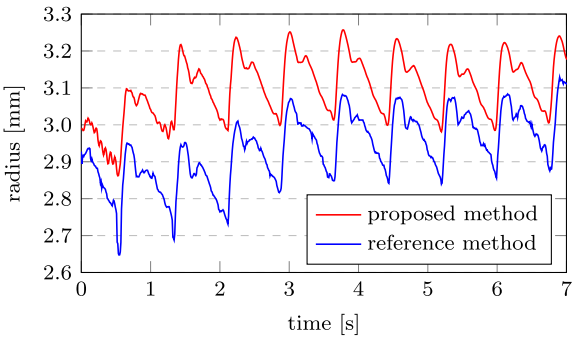


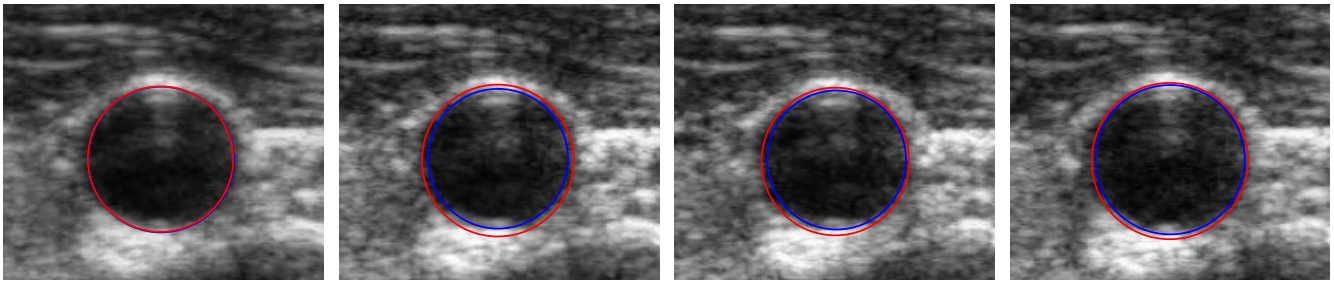
FIGURE 10. Estimates of the CCA circle radius obtained for the real ultrasound sequence with the proposed method and the reference method.

set to 10 MHz. In each frame, 64 lines were scanned. With this set of parameters, we achieved a lateral resolution of 1.6 mm and an axial resolution of 0.32 mm.

An individual phantom was then created for each frame by displacing the scatterers of the reference phantom according to the motion model described in Section VI-B. The resulting sequence of phantoms was passed to the Field II simulation program, which produced the ultrasound sequence F_0. The first frame of F_0 is shown in Figure 7.

To simulate speckle decorrelation, we then created sequences F_2 and F_5 by replacing in each frame of F_0 2% and 5% of the scatterers, respectively, with new scatterers.[2] The new scatterers were again uniformly distributed in the phantom of the respective frame, and their amplitudes were chosen as explained above. A more detailed description of this generation process is provided in [18]. Because of the high runtime of the Field II simulations, we generated only

[2]Speckle decorrelation is more commonly associated with a nonuniform motion of the scatterers [9]. However, our approach of randomly replacing a proportion of the scatterers with new ones can be motivated by the fact that the complex ultrasound image (before envelope detection) can be modeled as the sum of complex phasors whose amplitudes and phases are determined by the scatterer's reflectivity and distance to the transducer, respectively [32]. Due to phase wrapping, it is irrelevant whether a scatterer travels, e.g., half or one and a half of the ultrasound wavelength. Thus, the effect of a nonuniform scatterer motion is similar to that of randomly drawing new scatterers.



(a) Frame 1 (b) Frame 200 (c) Frame 400 (d) Frame 600

FIGURE 11. Estimates of the CCA wall circle obtained for four frames of the real ultrasound sequence with the proposed method (red) and the reference method (blue): (a) frame 1 (0 s), (b) frame 200 (1.77 s), (c) frame 400 (3.56 s), and (d) frame 600 (5.35 s). Note that for frame 1, the two estimates coincide visually.

a single realization for each of the sequences F_0, F_2, and F_5, i.e., the number of realizations for data set F is $K = 1$. The frames of sequences F_2 and F_5 are similar visually to those of F_0 (see Figure 7).

In Table 4, we list the RMSE obtained with the proposed method and the reference method for data set F. As in the case of data set S, the proposed method is seen to consistently outperform the reference method. Furthermore, Figure 8 shows the estimated radius sequences and the absolute estimation errors obtained for sequence F_2. Again, the proposed method is seen to estimate the radius much more accurately than the reference method. Finally, still for sequence F_2, Figure 9 shows the CCA circle estimates obtained with the proposed method and the reference method at four different frames times. It is seen that, while the estimates of the two methods are generally similar, the estimates of the reference method are slightly biased towards lower values of the radius. This confirms the results shown in Figure 8(a).

F. REAL ULTRASOUND SEQUENCE

To complement the results obtained for the synthetic data sets S and F, we consider a real ultrasound video sequence that shows the CCA transverse section of a healthy test subject (sex: male, age: 27, weight: 64 kg). This sequence was measured using an Ultrasonix OP device with linear probe L145/38 (Ultrasonix Medical, Richmond, BC, Canada). The sequence length is 787 frames, and the frame rate is 112 fps. The sequence was obtained with the subject's informed consent and with approval from the ethics committee.

Figure 10 shows the estimated radius waveform, while Figure 11 shows the CCA circle estimates at four different frame times. From Figure 10, one can conclude that the estimated radius waveform obtained with either method exhibits the typical characteristics described for example in [26], where the waveform during a cardiac cycle consists of two consecutive peaks (local maxima), the first associated with the forward wave and the second associated with the reflected wave. One can furthermore see in Figure 10 that after an initial convergence period—which spans approximately two cardiac cycles—the waveform produced by the proposed method is smoother than that produced by the reference method. Both Figure 10 and Figure 11 show that the proposed method generally yields larger radius estimates than the reference method; however, no conclusion regarding estimation accuracy can be drawn from this observation as no ground truth is available. From the radius waveform obtained with the proposed method, we can derive the following approximate values of diagnostically relevant parameters: the heart rate is obtained as 80 beats per minute, the diastolic artery diameter D_d as 6.00 mm, the systolic artery diameter D_s as 6.48 mm, the pulse diameter $D_\mathrm{s} - D_\mathrm{d}$ as 0.48 mm, and the circumferential strain $(D_\mathrm{s} - D_\mathrm{d})/D_\mathrm{d}$ as 0.08. All these values are in the typical range for a healthy male subject [26, Ch. 9]. If measurements of the pulse pressure were available, it would also be possible to calculate various arterial stiffness indices such as arterial distensibility and compliance [26, Ch. 9].

VII. CONCLUSION

The method we proposed in this paper is able to continuously and accurately track a circular approximation of the common carotid artery (CCA) wall, based on an observed B-mode ultrasound sequence of arbitrary length. We used an unscented Kalman filter (UKF) to track a composite state characterizing CCA wall motion. The operation of the UKF relies on a new state-space model that describes the dynamics of CCA wall motion and a related measurement process in a stochastic manner. The UKF fuses two different sets of measurements, of which one is produced by the pyramidal Lucas-Kanade optical flow algorithm and the other by a recently proposed CCA wall localization method. This fusion enables an effective, temporally consistent compensation of feature drift, which normally impairs the performance of optical flow based methods. The temporally consistent compensation of feature drift is a major advantage of our method over state-of-the-art methods.

We performed a quantitative evaluation of the accuracy of the proposed method for synthetic data sets containing additive and multiplicative noise and emulating speckle decorrelation. Our results demonstrate significant performance advantages relative to the state-of-the-art method presented in [18]. We also validated our method for a real ultrasound sequence.

The CCA circle radius waveform estimated by the proposed method can be used to calculate various indices that provide information about the health of the cardiovascular system [6], [7]. Moreover, because by its effective feature drift compensation the proposed method can be applied to arbitrarily long ultrasound sequences, it is possible to average clinically relevant indices over many cardiac cycles, thereby improving accuracy and reliability. In addition, the Fourier series coefficients calculated by our method can be used for a frequency analysis of the CCA radius waveform sequence, which is potentially useful for the diagnosis of cardiovascular diseases [33]. Developing an extension of our method that allows the estimation of other clinically important parameters of the CCA, such as radial and longitudinal strain [26], is an interesting direction for future research.

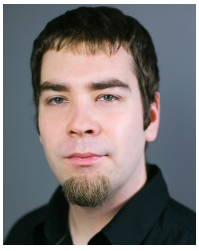
The proposed method can be extended in various other ways. For example, some or all of the parameters of the method can be included in the state and estimated online. The CCA wall can be characterized by a more sophisticated family of contours (instead of circles). Finally, an improved optical flow algorithm can be employed. Indeed, the pyramidal Lucas-Kanade algorithm used in our method is based on the assumption that tissue motion in the neighborhood of an FP can be approximated sufficiently well by a simple translation. The same assumption underlies our definition of the error metric $\varepsilon_t^{(n)}$ in (1). To better account for nonuniform tissue motion, the translation assumption can be relaxed by using an affine model for the velocity field [34] or a global optical flow algorithm such as [35]. The image patch $I_t^{(\tilde{\mathbf{y}}_t')}$ in (1) can then be replaced by an appropriately warped and translated version of the original image patch.

REFERENCES

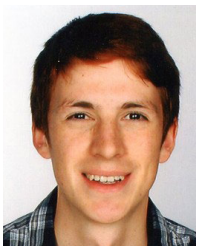
[1] *World Health Statistics 2018: Monitoring Health for the SDGs, Sustainable Development Goals*, World Health Org., Geneva, Switzerland, 2018.
[2] J. N. Cohn, "Arterial stiffness, vascular disease, and risk of cardiovascular events," *Circulation*, vol. 113, no. 5, pp. 601–603, Feb. 2006.
[3] Y.-X. Wang and R. Fitch, "Vascular stiffness: Measurements, mechanisms and implications," *Current Vascular Pharmacol.*, vol. 2, no. 4, pp. 379–384, Oct. 2004.
[4] S. Laurent, J. Cockcroft, L. Van Bortel, P. Boutouyrie, C. Giannattasio, D. Hayoz, B. Pannier, C. Vlachopoulos, I. Wilkinson, and H. Struijker-Boudier, "Expert consensus document on arterial stiffness: Methodological issues and clinical applications," *Eur. Heart J.*, vol. 27, no. 21, pp. 2588–2605, Sep. 2006.
[5] D. A. Duprez and J. N. Cohn, "Arterial stiffness as a risk factor for coronary atherosclerosis," *Current Atherosclerosis Rep.*, vol. 9, no. 2, pp. 139–144, Apr. 2007.
[6] M. Mokhtari-Dizaji, M. Montazeri, and H. Saberi, "Differentiation of mild and severe stenosis with motion estimation in ultrasound images," *Ultrasound Med. Biol.*, vol. 32, no. 10, pp. 1493–1498, Oct. 2006.
[7] Z. Gao, Y. Li, Y. Sun, J. Yang, H. Xiong, H. Zhang, X. Liu, W. Wu, D. Liang, and S. Li, "Motion tracking of the carotid artery wall from ultrasound image sequences: A nonlinear state-space approach," *IEEE Trans. Med. Imag.*, vol. 37, no. 1, pp. 273–283, Jan. 2018.
[8] E. J. Chen, R. S. Adler, P. L. Carson, W. K. Jenkins, and W. D. O'Brien, "Ultrasound tissue displacement imaging with application to breast cancer," *Ultrasound Med. Biol.*, vol. 21, no. 9, pp. 1153–1162, Jan. 1995.
[9] S. Golemati, A. Sassano, M. J. Lever, A. A. Bharath, S. Dhanjil, and A. N. Nicolaides, "Carotid artery wall motion estimated from b-mode ultrasound using region tracking and block matching," *Ultrasound Med. Biol.*, vol. 29, no. 3, pp. 387–399, Mar. 2003.
[10] J. Tat, J. S. Au, P. J. Keir, and M. J. MacDonald, "Reduced common carotid artery longitudinal wall motion and intramural shear strain in individuals with elevated cardiovascular disease risk using speckle tracking," *Clin. Physiol. Funct. Imag.*, vol. 37, no. 2, pp. 106–116, Mar. 2017.
[11] A. Gastounioti, S. Golemati, J. Stoitsis, and K. S. Nikita, "Kalman-filter-based block matching for arterial wall motion estimation from B-mode ultrasound," in *Proc. IEEE Int. Conf. Imag. Syst. Techn.*, Thessaloniki, Greece, Jul. 2010, pp. 234–239.
[12] G. Zahnd, M. Orkisz, A. Sérusclat, P. Moulin, and D. Vray, "Evaluation of a Kalman-based block matching method to assess the bi-dimensional motion of the carotid artery wall in B-mode ultrasound sequences," *Med. Image Anal.*, vol. 17, no. 5, pp. 573–585, Jul. 2013.
[13] A. Gastounioti, S. Golemati, J. Stoitsis, and K. Nikita, "Comparison of Kalman-filter-based approaches for block matching in arterial wall motion analysis from B-mode ultrasound," *Meas. Sci. Technol.*, vol. 22, no. 11, pp. 1–9, Oct. 2011.
[14] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Int. Joint Conf. Artif. Intell.*, Vancouver, BC, Canada, Aug. 1981, pp. 674–679.
[15] K. Říha and I. Potúček, "The sequential detection of artery sectional area using optical flow technique," in *Proc. Int. Conf. Circ. Syst. Electron. Contr. Signal Process. (WSEAS)*, Stevens Point, WI, USA, Dec. 2009, pp. 222–226.
[16] A. Gastounioti, N. N. Tsiaparas, S. Golemati, J. S. Stoitsis, and K. S. Nikita, "Affine optical flow combined with multiscale image analysis for motion estimation of the arterial wall from B-mode ultrasound," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Boston, MA, USA, Aug. 2011, pp. 559–562.
[17] S. Golemati, J. S. Stoitsis, A. Gastounioti, A. C. Dimopoulos, V. Koropouli, and K. S. Nikita, "Comparison of block matching and differential methods for motion analysis of the carotid artery wall from ultrasound images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 5, pp. 852–858, Sep. 2012.
[18] K. Říha, M. Zukal, and F. Hlawatsch, "Analysis of carotid artery transverse sections in long ultrasound video sequences," *Ultrasound Med. Biol.*, vol. 44, no. 1, pp. 153–167, Jan. 2018.
[19] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.
[20] J.-Y. Bouguet, "Pyramidal implementation of the Lucas Kanade feature tracker," Microprocessor Res. Labs, Intel Corp., Mountain View, CA, USA, Tech. Rep., 2000.
[21] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Alvey Vis. Conf.*, Manchester, U.K., 1988, pp. 147–152.
[22] J. Dorazil, K. Řiha, and M. K. Dutta, "Common carotid artery wall localization in B-mode ultrasound images for initialization of artery wall tracking methods," in *Proc. 42nd Int. Conf. Telecommun. Signal Process. (TSP)*, Budapest, Hungary, Jul. 2019, pp. 605–608.
[23] J. Dorazil, R. Repp, T. Kropfreiter, R. Prüller, K. Říha, and F. Hlawatsch, "Feature drift resilient tracking of the carotid artery wall using unscented Kalman filtering with data fusion," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Barcelona, Spain, May 2020, pp. 1095–1099.
[24] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Commun. ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972.
[25] S. Kim, A. S. Paul, E. A. Wan, and J. McNames, "Multiharmonic tracking using sigma-point Kalman filter," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Vancouver, BC, Canada, Aug. 2008, pp. 2648–2652.
[26] W. Nichols, M. O'Rourke, and C. Vlachopoulos, *McDonald's Blood Flow in Arteries: Theoretical, Experimental and Clinical Principles*, 6th ed. London, U.K.: CRC Press, 2011.
[27] M. Schlaikjer, S. T. Petersen, J. A. Jensen, and P. F. Stetson, "Tissue motion in blood velocity estimation and its simulation," in *Proc. IEEE Ultrason. Symp.*, Sendai, Japan, Oct. 1998, pp. 1495–1499.
[28] S. Challa, M. R. Morelande, D. Mušicky, and R. J. Evans, *Fundamentals of Object Tracking*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[29] J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 39, no. 2, pp. 262–267, Mar. 1992.
[30] J. A. Jensen, "Field: A program for simulating ultrasound systems," in *Proc. Nordic-Baltic Conf. Biomed. Eng.*, Tampere, Finland, Jun. 1996, pp. 351–353.

[31] J. Stoitsis, S. Golemati, E. Bastouni, and K. S. Nikita, "A mathematical model of the mechanical deformation of the carotid artery wall and its application to clinical data," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Lyon, France, Aug. 2007, pp. 2163–2166.
[32] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound B-scans," *IEEE Trans. Sonics Ultrason.*, vol. 30, no. 3, pp. 156–163, May 1983.
[33] A. T. Yokobori, Jr., K. Watanabe, Y. Saiki, Y. Nishikawa, K. Kudo, T. Ohmi, Y. Hayatsu, M. Suzuki, and K. Sasaki, "Frequency and chaotic analysis of pulsatile motion of blood vessel wall related to aneurysm," *Bio-Medical Mater. Eng.*, vol. 30, no. 2, pp. 243–253, Mar. 2019.
[34] J.-Y. Bouguet, "Pyramidal implementation of the affine Lucas Kanade feature tracker description of the algorithm," Microprocessor Res. Labs, Intel Corp., Mountain View, CA, USA, Tech. Rep., 2001.
[35] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, nos. 1–3, pp. 185–203, Aug. 1981.

JAN DORAZIL received the B.Sc. degree in teleinformatics and the M.Sc. degree in telecommunication and information technology from Brno University of Technology, Brno, Czech Republic, in 2015 and 2017, respectively. He is currently pursuing the Ph.D. degree in teleinformatics with the Faculty of Electrical Engineering and Communication, Brno University of Technology. From 2017 to 2019, he was with Brno University of Technology. Since 2020, he has been with TU Wien, Vienna, Austria. His research interests include biomedical image processing and statistical signal processing with a focus on Bayesian inference.



RENE REPP received the B.Sc. degree in electrical engineering and the Dipl.-Ing. degree (M.Sc. equivalent) in telecommunication engineering from TU Wien, Vienna, Austria, in 2012 and 2016, respectively. He is currently pursuing the Ph.D. degree with the Institute of Telecommunications, TU Wien. Since June 2020, he has been with the Acoustics Research Institute, Austrian Academy of Sciences, Vienna. His research interests include distributed multi-object tracking in sensor networks, point processes, Bayesian nonparametric models, and machine learning.



THOMAS KROPFREITER received the B.Sc. degree in electrical engineering and the Dipl.-Ing degree (M.Sc. equivalent) in telecommunication engineering from TU Wien, Vienna, Austria, in 2012 and 2014, respectively. He is currently pursuing the Ph.D. degree with the Institute of Telecommunications, TU Wien. From 2013 to 2015, he was a member of the Mobile Communications Group, TU Wien, where he worked on beamforming in multi-cell scenarios. His research interests include multi-object tracking, distributed signal processing in wireless sensor networks, message passing algorithms, and finite set statistics.

RICHARD PRÜLLER received the B.Sc. degree in electrical engineering and the Dipl.-Ing. degree (M.Sc.) in telecommunications from TU Wien, Vienna, Austria, in 2016 and 2019, respectively. He is currently pursuing the Ph.D. degree in telecommunications with the Institute of Telecommunications, TU Wien. From 2016 to 2019, he was a Student Assistant with TU Wien, where he is currently a University Assistant. His research interests include stochastic graphical models, random matrix theory, and wireless channel models.

KAMIL ŘÍHA received the M.Sc. degree in electronics and communication, the Ph.D. degree in 3D scene acquisition for auto-stereoscopic display, and the Associate Professor degree in the analysis of bloodstream geometry in image sequences from the Brno University of Technology, Brno, Czech Republic, in 2003, 2007, and 2014, respectively. Since 2006, he has been with the Faculty of Electrical Engineering, Brno University of Technology. His research interests include digital image and video processing.

FRANZ HLAWATSCH (Fellow, IEEE) received the Diplom-Ingenieur, Dr.Techn., and Univ.-Dozent (Habilitation) degrees in electrical engineering/signal processing from TU Wien, Vienna, Austria, in 1983, 1988, and 1996, respectively.

Since 1983, he has been with the Institute of Telecommunications, TU Wien, where he is currently an Associate Professor. From 1991 to 1992, funded by an Erwin Schrödinger Fellowship, he spent a sabbatical year with the Department of Electrical Engineering, University of Rhode Island, Kingston, RI, USA. In 1999, 2000, and 2001, he held one-month Visiting Professor positions at INP/ENSEEIHT, Toulouse, France and at IRCCyN, Nantes, France. He has (co)authored a book, three review articles that appeared in the *IEEE Signal Processing Magazine*, about 200 refereed scientific articles and book chapters, and three patents. He coedited three books. His research interests include statistical and compressive signal processing methods and their application to inference and learning problems. He is a EURASIP Fellow. He was a member of the IEEE SPCOM Technical Committee from 2004 to 2009. He was a Technical Program Co-Chair of EUSIPCO 2004 and served on the technical committees of numerous IEEE conferences. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2003 to 2007, the IEEE TRANSACTIONS ON INFORMATION THEORY from 2008 to 2011, and the IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS from 2014 to 2017. He coauthored articles that won an IEEE Signal Processing Society Young Author Best Paper Award, a Best Student Paper Award at IEEE ICASSP 2011, and a Student Paper Award at FUSION 2020.