Information Fusion 35 (2017) 58–67

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Motion prediction via online instantaneous frequency estimation for vision-based beating heart tracking

Bo Yang[a,*], Chao Liu[b], Wenfeng Zheng[a], Shan Liu[a]

[a] School of Automation Engineering, University of Electronic Science and Technology of China, ChengDu 611731, China
[b] LIRMM, UMR 5506, CNRS-UM, Montpellier 34095, France

ARTICLE INFO

Article history:
Received 20 September 2015
Revised 24 June 2016
Accepted 12 September 2016
Available online 13 September 2016

Keywords:
Motion prediction
Beating heart tracking
Instantaneous frequency
Motion compensation
Kalman filter

ABSTRACT

The beating heart tracking based on stereo endoscope remains challenging due to highly dynamic scenes and poor imaging conditions in minimally invasive surgery. This paper proposes a new prediction method for robust tracking of heart motion. The dual time-varying Fourier series is used for modeling the motion of points of interest (POI) on heart surfaces, which is driven jointly by breathing and heartbeat motion. A dual Kalman filtering scheme is used to estimate the frequencies and Fourier coefficients of the model respectively. A novel orthogonal decomposition algorithm is developed to measure the instantaneous frequencies of breathing and heartbeat motion online from the 3D trajectory of the POI. The difference in direction between breathing and heartbeat motion is exploited by using principal component analysis on the past trajectory, and optimal 1D principal component signals are extracted for measuring the corresponding frequencies. The frequencies calculated from the orthogonal subbands are fused based on an additive noise model for optimal frequency measurement. The proposed method is evaluated and compared with other available prediction methods based on the simulated data and the real-measured signals from the videos recorded by the daVinci® surgical robot. The prediction algorithm is finally incorporated into a well-established visual tracking method to handle long-term occlusions.

1. Introduction

Over the past several decades, minimally invasive surgery (MIS) has been widely adopted, due to its advantages over traditional surgery, such as smaller incisions, less post-operative pain and shorter hospital stay. Recently, robotic technologies have further widened the scope of MIS by enhancing surgeon's dexterity and coordination and improving operating stability and accuracy [1]. Robotic-assisted surgeries replace the conventional MIS tools with robotic instruments (slave), which are inserted into patients body through tiny incisions and execute control command from surgeons through teleoperation. The surgeon views real-time images of the surgical site captured by an endoscopy to manipulate the master device and close the control loop.

The robotic surgery system enables surgeons to perform complex and delicate procedures, e.g. off-pump *coronary artery bypass graft* (CABG), more efficiently. In comparison with traditional on-pump CABG, the off-pump procedure reduces the risk of complications associated with stopping heart and replacing it with a heart-lung machine (on pump). However, operating on beating heart is very challenging for surgeons, because the rapid heart motion is difficult to handle by hand effectively. As indicated by Trejos [2], the heart motion exceeds human maximum reliable tracking bandwidth. Although mechanical stabilizer can be used to constrain heart motion, residual motion is still remarkable, and installing stabilizer is laborious in MIS.

Active motion compensation (AMC) with the assistance of surgical robot has been explored to overcome this limitation. In [3], Nakamura et al. firstly indicated that a virtually stable operating environment is possible to be achieved by tracking physiological motion and actively synchronizing robotic instruments with the motion. Following Nakamura's concept, various compensation techniques and schemes based on the existing robotic systems and devices have been investigated, which mainly focused on the two fundamental issues: measurement (or tracking) and control. In order to measure the motion of points of interest (POI) on a heart surface, tracking algorithms based on various sensor systems are studied, such as 2D visual systems [4], laser-scan [5], whisker sensor [6], sonomicrometry [7,8], and stereo endoscopes [9–14]. Controllers and controlling schemes are also studied in the literature to quickly and accurately synchronize the movement of robotic arms. These studies have justified the feasibility of the AMC technique in the beating heart surgery, though there are still some po-

* Corresponding author.
E-mail address: boyang@uestc.edu.cn (B. Yang).

http://dx.doi.org/10.1016/j.inffus.2016.09.004
1566-2535/

tential problems in safety, reliability, accuracy and ergonomics for clinical use.

The motion prediction of POI is an important step for both the measurement and control purposes, though their requirements on prediction performance are different. For efficient measurement, recursive tracking schemes are usually adopted. Since the measurement of current frame depends on the correct result measured from the previous frame, as an initial estimation, recursive schemes normally cannot recover themselves when the tracking chain is interrupted. In a highly dynamic MIS, the interruption is inevitable, especially for a stereo-endoscope based tracking, which as indicated by Mountney et al. [1] maybe the most practical solution for the AMC because no additional instrument port besides the endoscope is required. Unfortunately, endoscope images are often affected by dynamic effects, such as motion blurring, specular reflection, and occlusions (arising from surgical instruments, smoke or blood, etc.), which will result in tracking failure. A well-designed prediction algorithm can supply missing information in case of tracking failure and restore normal tracking after disturbances. In addition, motion prediction can improve the efficiency and robustness of recursive tracking algorithms by providing closer initialization for the forthcoming frame.

In [4], Ortmaier et al. developed a robust predication algorithm by using Takens' theorem and coupling with electrocardiogram (ECG) and respiration pressure signal (RPS) data to predict the position of landmarks on a heart surface when visual tracking was lost. Wong et al. [13] proposed a quasi-spherical triangle model for 3D tissue tracking. To handle occlusion, they improved upon the prediction algorithm developed in [4] by detecting the peaks and valleys of motion signals. Bader et al. [15] represented a partition of the heart surface with a pulsating membrane model, whose motion behavior was described using a partial differential equation. A standard Kalman filter was employed to improve state estimation and identify corresponding (or retrieve missed) markers in consecutive image frames for robust tracking. Yuen et al. [16] developed a 1-DOF ultrasound-guided motion compensation system for beating heart surgery. To counteract the delays and noise during 3D ultrasound imaging, the extended Kalman filter (EKF) based on a time-vary Fourier series (TVFS) model was developed. Two autoregressive (AR) filters estimated by the standard and exponentially weighted recursive least squares (RLS) were also evaluated for comparison. Richa et al. [17] studied and compared the vector autoregressive (VAR) model and the TVFS model to predict immediate future POI motion for efficient visual tracking. The EKF was employed to estimate the model parameters. In [18], a dual time-varying Fourier series (DTVFS), which models the POI motion as a combination of heartbeat component and breathing component explicitly, was further incorporated into the EKF scheme to bridge tracking failure. Mountney and Yang [19] described breathing cycles with a asymmetrically periodic model and proposed an EKF-based probabilistic framework to predict and estimate endoscope, breathing and map motion simultaneously in a liver MIS.

Besides, motion prediction is also investigated for improving control performance of the AMC. Direct error feedback control alone is unable to handle high-bandwidth heart motion with measurement delays and disturbances. A predictive controller in the feedforward path is necessary for the AMC with high accuracy. Ginhoux et al. [20] separated heart motion into the heartbeat component of five harmonics and the breathing component using adaptive frequency-cancellation filters. A generalized predictive controller for AMC was deigned based on the periodicity of the components. In [7], Bebek and Cavusoglu developed an intelligent AMC algorithm for the CABG surgery, which employed a receding horizon model-predictive controller to cancel the heartbeat motion. The future reference signal was created using the last heartbeat cycle and the period of heartbeat was adjusted online using ECG data. In [21], Bachata et al. reviewed existing prediction methods and proposed a model-based prediction algorithm to improve the AMC control. They expressed the coupling between the breathing and heartbeat components using amplitude-modulated Fourier series with known heartbeat and breathing frequencies. The model parameters were estimated by the RLS algorithm with a forgetting factor. Tuna et al. [22] studied and incorporated two VAR-modeled prediction algorithms, originally described by Franke et al. [23,24], into the AMC controller designed in [7], and tested them on a 3-DOF robotic system. The exponentially weighted RLS was employed for adaptive estimation of the model parameters.

The primary goal of this study is to develop an efficient prediction method to copy with occlusions during visual tracking. A relatively long prediction horizon (approximately 1–3 s) is required in this case, rather than the immediate short accuracy which is desired by the predictive controllers. The POI motion is modeled using the DTVFS. A dual Kalman filter (DKF) scheme [25] is employed to estimate the frequency parameters and the Fourier coefficients of the DTVFS model, which overcomes the flaws of the EKF in linearizing nonlinear models. The instantaneous frequencies of heartbeat and breathing motion are measured independently from the 3D motion signals and estimated using one Kalman filter. The directionality of breathing and heartbeat motion is studied for efficient frequency measurement, which to the author's knowledge has not been addressed in the literature. The estimated frequencies transform the DTVFS into a linear model, whose parameters are estimated by another Kalman filter. The proposed prediction method is robust to irregularity of POI motion and measurement noise introduced by visual tracking, due to the separated dual Kalman filtering and the optimized frequency measurement based on an additive noise model. It is more efficient than the EKF method [18] because linear and separable state-space models with lower dimensions are involved. The proposed method is evaluated with various datasets, generated by simulation modeling and measured from phantom heart videos and in vivo videos. The feasibility of the application of the prediction algorithm to visual tracking is studied by incorporating it into a well-established tracking method [12].

The rest of the paper is organized as follows. The DKF prediction scheme based on the DTVFS model is presented in Section 2. Section 3 describes the online instantaneous frequency measurement based on an orthogonal decomposition fusion algorithm. Section 4 discusses simulation and experimental results. Finally, conclusions are given in Section 5.

2. Dual Kalman filtering prediction scheme

The main properties of heart motion are firstly reviewed in this section by analyzing the POI motion data obtained by visual tracking from in vivo videos. The DTVFS is used to model the POI motion based on these analyses, and the DKF is developed to compute the frequency parameters and Fourier coefficients of the model.

2.1. Heart motion analysis and modeling

The motion of heart surfaces has been already studied in [1,4,7,12,16]. These studies indicate that the heart motion is primarily induced by breathing motion and heartbeat motion. The POI motion on a heart surface can be considered as a combination of the breathing and heartbeat components with different frequency ranges. These components are quasi-periodic in nature. The statistics such as frequencies and amplitudes are slowly varying during surgery, making accurate prediction challenging.

The 3D positions of a POI are measured offline from a stereo video of in vivo heart by vision tracking [12] to illustrate the