Optical flow-based vascular respiratory motion compensation

Keke Yang, Zheng Zhang, Meng Li, Tuoyu Cao, Maani Ghaffari, and Jingwei Song*

arXiv:2308.16451v1 [cs.RO] 31 Aug 2023

Abstract—This paper develops a new vascular respiratory motion compensation algorithm, Motion-Related Compensation (MRC), to conduct vascular respiratory motion compensation by extrapolating the correlation between invisible vascular and visible non-vascular. Robot-assisted vascular intervention can significantly reduce the radiation exposure of surgeons. In robot-assisted image-guided intervention, blood vessels are constantly moving/deforming due to respiration, and they are invisible in the X-ray images unless contrast agents are injected. The vascular respiratory motion compensation technique predicts 2D vascular roadmaps in live X-ray images. When blood vessels are visible after contrast agents injection, vascular respiratory motion compensation is conducted based on the sparse Lucas-Kanade feature tracker. An MRC model is trained to learn the correlation between vascular and non-vascular motions. During the intervention, the invisible blood vessels are predicted with visible tissues and the trained MRC model. Moreover, a Gaussian-based outlier filter is adopted for refinement. Experiments on in-vivo data sets show that the proposed method can yield vascular respiratory motion compensation in $0.032\,\sec$, with an average error $1.086\,\text{mm}$. Our real-time and accurate vascular respiratory motion compensation approach contributes to modern vascular intervention and surgical robots.

Index Terms—robot-assisted vascular interventions, vascular respiratory motion compensation, dynamic roadmapping, optical flow

I. INTRODUCTION

Robot-assisted vascular interventional therapy is a rapidly developing technology in the field of cardiovascular disease treatment [1–3]. Its value in peripheral blood vessels, particularly in tumor embolization therapy, is gaining increasing attention. Among them, robot-assisted vascular intervention reduces radiation and has drawn more attention recently [4]. In the vascular intervention procedure, roadmapping is the process of superimposing 2D vascular on live fluoroscopic images and plays a key role in surgeries [5].

Vascular respiratory motion compensation (or dynamic roadmapping) technique pushes conventional **static** roadmaps to **dynamic** roadmaps and helps the interventionists/robots manipulate catheters and guidewires or place stents by visualization of the map and devices on one screen [5]. In typical interventions, catheters and guidewires are guided under live fluoroscopic images, which contain 2D device information [6].

K. Yang, Z. Zhang, M. Li, and T. Cao are with United Imaging, Shanghai, China. {keke.yang, zheng.zhang, meng.li02, tuoyu.cao}@united-imaging.com.

M. Ghaffari is with the University of Michigan, Ann Arbor, MI 48109, USA. maanigj@umich.edu.

J. Song* (Corresponding author) is with United Imaging Research Institute of Intelligent Imaging, Beijing 100144, China jingweisong.eng@outlook.com
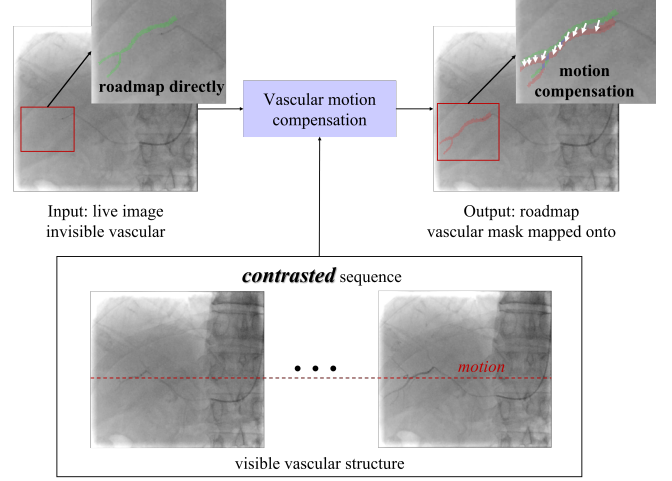

Fig. 1: The scenario of vascular respiratory motion compensation: input live image with invisible vascular, output roadmap of the live image with mapped red vascular mask. The roadmap can be obtained by contrasted sequences and mapped onto the live image directly, as indicated by the green mask. Vascular respiratory motion compensation utilizes contrasted sequences to predict motion between the live image and roadmap, as white arrows indicate. The red dotted line is for easy visualization of motion.

During the process, contrast agents are injected to provide clear vascular structures [7]. However, contrast agents flow quickly, and the vascular no longer develops after the contrast agents disappear [3]. Moreover, deforming/moving roadmapping brings large errors in organs like the liver due to respiration motion and requires additional compensation. Fig. 1 shows a typical vascular respiratory motion compensation for handling the two issues in conventional static roadmapping. The green mask obtained from the contrasted image (the fluoroscopic image with contrast agent injection as shown in Fig. 1 contrasted sequences) is mapped onto the live image directly. The error is evident from the guidewire due to breathing motion. Vascular respiratory motion compensation can dynamically move 2D roadmaps to properly match the live fluoroscopic images, especially after the contrast agent disappears and vascular structures are not visible from fluoroscopic images. The red mask in Fig. 1 is the prediction of vascular respiratory motion compensation, which can provide immediate feedback to robots/physicians during surgeries.

Existing vascular respiratory motion compensation methods can be categorized as respiratory state-driven and motion modeling [7]. Compensation based on respiratory state extraction connects vascular motion with the respiratory state, which can be extracted through external devices or images. External devices include respiratory belts [8], surface markers [9, 10],

and electromagnetic sensors [11]. Although estimating the respiratory state with external devices-based is straightforward and applicable, it disrupts the clinical workflow and is not robust due to unfixed respiratory rate and amplitude [7]. Unlike using additional devices, image-based approaches estimate respiratory state with anatomical landmarks such as the diaphragm to obtain respiratory state [7, 12–14]. [7] estimated respiratory state based on diaphragm location in the non-contrasted images and fitted between vascular affine transformation and respiratory state with a linear function. For non-contrasted live images, an affine transformation of vascular was estimated by respiratory state, which is computationally efficient. [7] reported an average one-frame processing time of $17\ ms$, which is the fastest. Image-based respiratory state extraction requires no additional clinical workflow but is limited by the Field-of-View (FoV). As [15] pointed out, although respiratory state-based methods are time-efficient, their accuracy and robustness are limited because motions over the respiratory cycle are generally less reproducible.

Different from respiratory state-driven approaches, model-based approaches build models to learn and predict the motion in fluoroscopic frames [3, 16–19]. Model-based methods can be categorized as **catheter-based** and **catheter-free**. **Catheter-based** methods track the catheter tip to conduct vascular respiratory motion compensation. [16] learned displacements of the catheter tip by a Convolutional Neural Network (CNN). In their works, vascular motion was factorized into respiratory and heart motion. Heart motion compensation was done by ECG, and respiratory motion was predicted by CNN. [17] conducted cardiac and respiratory-related vascular motion compensation by ECG alignments and catheter tip tracking in X-ray fluoroscopy, respectively. In particular, to realize accurate and robust tracking of the catheter tip, they proposed a new deep Bayesian filtering method that integrated the detection outcome of a CNN and the sequential motion estimation using a particle filtering framework. **Catheter-free** methods conduct vascular respiratory motion compensation by soft tissue motion. [20, 21] utilized soft tissue around the heart to model vascular motion. These works are based on the assumption that vascular motion and soft tissue motion followed the same affine transformation, and Lucas-Kanade (LK) tracker was improved to estimate soft tissue affine transformation to conduct vascular respiratory motion compensation. [20] proposed special handling of static structures to recover soft tissue motion. [20] also applied the LK tracker on multiple observed fluoroscopic images to gain robustness. However, the LK tracker on the entire X-ray image needs heavy computation. Meanwhile, vascular motion and soft tissue motion are not always consensus. To sum up, methods based on motion modeling can yield high accuracy and robust compensations without additional input, but their computations are large. To our knowledge, no study can achieve both real-time implementation and high accuracy in hepatic vascular respiratory motion compensation.

In this paper, we propose Motion-Related Compensation (MRC) algorithm to achieve real-time and accurate vascular respiratory motion compensation. To enable fast vascular respiratory motion compensation, feature points are extracted from the observed fluoroscopic image; deforming vascular motions are estimated based on the tracked feature points. Furthermore, a correlation model is built between vascular points motion and non-vascular points motion when the contrast agent develops in X-ray images. After the disappearance of the contrast agent, our trained model can predict the invisible motion of the vascular points based on the motion of the non-vascular points. Moreover, a Gaussian-based Outlier Filtering (GOF) technique is adopted to refine the correlation model's prediction. In summary, our main contributions are:

- To our knowledge, the proposed MRC is the first applicable method that is both real-time and accurate for respiratory motion compensation. It achieves $31Hz$ and $1.086mm$ for the typical fluoroscopic image size 512×512 on a modern desktop with Intel Core i5-10500 CPU at 3.10GHz.
- We propose a novel vascular MRC method without assuming that vascular and non-vascular motion are identical, which uses multi-frame contrasted images to learn the model of vascular-nonvascular motion correlation.
- GOF is adopted to improve the accuracy of our MRC predictions.

II. PRELIMINARIES

Fig. 1 describes the vascular respiratory motion compensation process. Blood vessels are visible from a small sequence with the contrast agent. After the contrast agent flows, the vascular is not visible in the live X-ray images. Vascular respiratory motion compensation on live images significantly benefits surgeons and surgical robots. Thus, the purpose of this research is to implement vascular motion compensation on live X-ray images based on the limited images with visible vascular structures.

Denote the sequence with and without the contrast agent as 2D images $\mathbf{I} = \{\mathbf{I}_1, \mathbf{I}_2, ..., \mathbf{I}_k\}$ and $\mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_q, ...\}$. Reference frame taken from contrasted images is specially defined as $\mathbf{I}_r$. The binary vascular mask of reference frame $\mathbf{I}_r$ is denoted as $\mathbf{M}_r$. The vascular corners and non-vascular corners extracted from reference frame $\mathbf{I}_r$ by Shi-Tomasi approach [22] are denoted as $\mathbf{C}_r^v = [\mathbf{c}_r^{v(1)}, ..., \mathbf{c}_r^{v(N_v)}]^\top$ and $\mathbf{C}_r^n = [\mathbf{c}_r^{n(1)}, ..., \mathbf{c}_r^{n(N_n)}]^\top$, where $\mathbf{c}_r^{v(i)}, \mathbf{c}_r^{n(i)} \in \mathbb{R}^{2 \times 1}$ and N_v, N_n are the number of vascular corners and non-vascular corners. The vascular corners motion flow between $\mathbf{I}_i$ and $\mathbf{I}_r$ is denoted as $\mathbf{D}_i^v = [\mathbf{d}_i^{v(1)}, ..., \mathbf{d}_i^{v(N_v)}]^\top \in \mathbb{R}^{N_v \times 2}$, where $\mathbf{d}_i^{v(j)} \in \mathbb{R}^{2 \times 1}$. The non-vascular motion flow is denoted as $\mathbf{D}_i^n = [\mathbf{d}_i^{n(1)}, ..., \mathbf{d}_i^{n(N_n)}]^\top \in \mathbb{R}^{N_n \times 2}$, where $\mathbf{d}_i^{n(j)} \in \mathbb{R}^{2 \times 1}$. The vascular corners motion flow between $\mathbf{R}_q$ and $\mathbf{I}_r$ is denoted as $\mathbf{F}_q^v = [\mathbf{f}_q^{v(1)}, ..., \mathbf{f}_q^{v(N_v)}]^\top \in \mathbb{R}^{N_v \times 2}$, where $\mathbf{f}_q^{v(j)} \in \mathbb{R}^{2 \times 1}$. The non-vascular motion flow is denoted as $\mathbf{F}_q^n = [\mathbf{f}_q^{n(1)}, ..., \mathbf{f}_q^{n(N_n)}]^\top \in \mathbb{R}^{N_n \times 2}$, where $\mathbf{f}_q^{n(j)} \in \mathbb{R}^{2 \times 1}$.

III. METHOD

A. System Overview

Fig. 2 shows the pipeline of our proposed MRC algorithm consisting of three modules: **sparse corners alignment**, **motion-related model**, and **GOF**. **Sparse corners alignment**
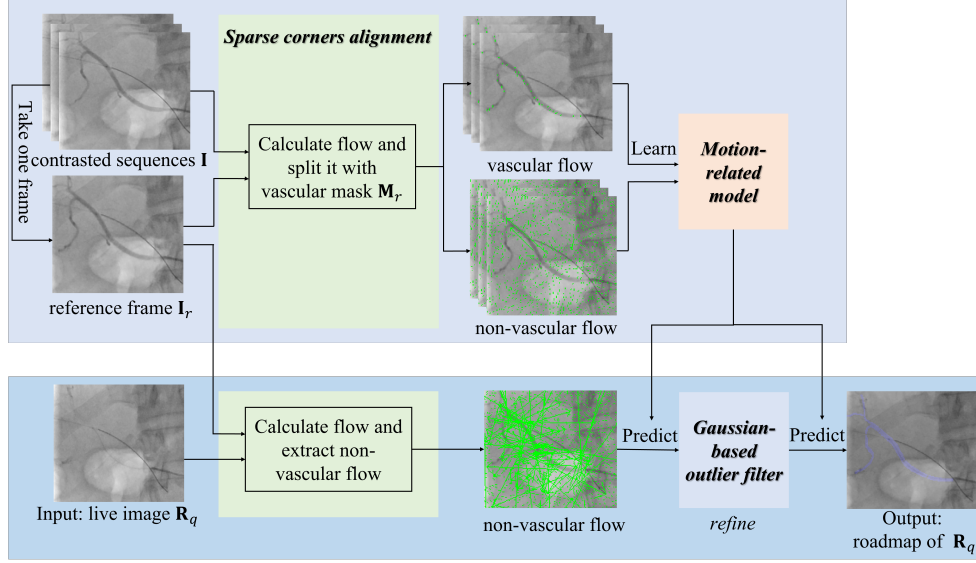
Fig. 2: Overview of the proposed MRC algorithm.

module calculates motion flow between reference frame $\mathbf{I}_r$ and live moving frames and splits motion flow into vascular motion flow and non-vascular motion flow with vascular mask $\mathbf{M}_r$ extracted from the reference frame. **Motion-related model** builds the correlation between vascular motion flow and non-vascular motion flow. **GOF** module filters outliers with the obtained non-vascular motion flow to refine prediction.

B. Sparse Corners Alignment

Affine motion parameterization is used in [20, 21] to model non-vascular motion incurred by respiration, which is regarded as vascular motion directly. Affine motion parameterization is also used in [7] to model vascular motion. However, the affine motion model may not simulate real vascular motion due to a low degree of freedom. Moreover, it is not accurate to treat vascular motion and non-vascular motion equally; for example, the amplitude of respiratory motion at the top of the liver is larger than at the bottom of the liver. Therefore, we adopt sparse optical flow as a tracker to obtain motion at the pixel level, which does not limit the freedom of motion nor assume the consistency of vascular motion and non-vascular motion. Sparse flow selects a sparse feature set of pixels (e.g., interesting features such as edges and corners) to track its motion vector, which only tracks much fewer points but achieves fast speed and high accuracy. Inspired by [23], Shi-Tomasi corner detection method [22] is used to extract sparse features, and LK [24] is used to track corners' motion sequence in this paper. And experiments in Section IV validate their efficiencies.

Sparse corner alignment calculates vascular and non-vascular motion flow separately. Firstly, the vascular corners $\mathbf{C}_r^v \in \mathbb{R}^{N_v \times 2}$ and the non-vascular corners $\mathbf{C}_r^n \in \mathbb{R}^{N_n \times 2}$ in reference frame $\mathbf{I}_r$ are extracted using Shi-Tomasi [22]. Then, sparse vascular motion flow $\mathbf{D}_i^v \in \mathbb{R}^{N_v \times 2}$ and non-vascular motion flow $\mathbf{D}_i^n \in \mathbb{R}^{N_n \times 2}$ between other frame $\mathbf{I}_i$ and $\mathbf{I}_r$ can be estimated by LK [24]. Similarly, for any live image $\mathbf{R}_q$, its non-vascular motion flow $\mathbf{F}_q^n$ can be obtained by LK [24].

Although dense flow can also obtain motion at the pixel level, it computes the optical flow sequence for every pixel, which produces accurate results but runs much slower. In addition, ORB and SIFT are more efficient in aligning sparse features with scale and orientation differences. Experiments show that ORB and SIFT are less robust in vascular respiratory motion compensation.

C. Motion-Related Model

The motion-related model first formulates a correlation between vascular and non-vascular motion and then predicts vascular motion after contrast agents disappear. We assume the motion of soft tissues caused by breathing is smooth, for example, the hepatic artery follows the diaphragm to make the correlated motion. Inspired by this, we assume that vascular motion is caused by the motion of the surrounding non-vascular tissues. Enforcing the smoothness presumption, our motion-related model retrieves vascular motion and non-vascular motion based on the sequence with visible vascular structures. It should be noted that the model parameters need to be updated for each injection of the contrast agent for each patient. Our model can only carry out vascular respiratory motion compensation for the blood vessels that has been observed. It is reasonable to do vascular respiratory motion compensation by the peripheral non-vascular motion that is elastic and connected to blood vessels [25]. It should be noted that contrast agents flow quickly, and our algorithm can reduce the use of contrast agents.

[20] retrieves vascular respiratory motion considering vascular and non-vascular motion as the same. It is not accurate in practice due to different motion magnitudes incurred by respiration. To establish a relation between vascular and non-vascular motion, a linear and a non-linear regression can be used to fit the relation. Its slow speed [26], hyperparameters sensitivity, and similar accuracy to linear regression drive us to select a linear model. In addition, [25] used the linear elastic

model, which indicates a linear regression. And experiments also verify the effectiveness of the linear regression model.

Specifically, the Pearson coefficient is adopted to quantify the correlation [27]. The ith vascular corner motion flows on k frames are denoted as $\mathbf{Y}_i = [\mathbf{d}_1^{v(i)}, ..., \mathbf{d}_k^{v(i)}] \in \mathbb{R}^{2 \times k}$. The jth non-vascular motion flows on k frames are denoted as $\mathbf{X}_j = [\mathbf{d}_1^{n(j)}, ..., \mathbf{d}_k^{n(j)}] \in \mathbb{R}^{2 \times k}$. Pearson coefficient between the ith vascular corner motion flows, and the jth non-vascular motion flows can be calculated by

$$\rho_{i,j} = \frac{\mathrm{cov}(\mathbf{X}_j|_x, \mathbf{Y}_i|_x)}{\sigma_{\mathbf{X}_j|_x} \sigma_{\mathbf{Y}_i|_x}} \cdot \frac{\mathrm{cov}(\mathbf{X}_j|_y, \mathbf{Y}_i|_y)}{\sigma_{\mathbf{X}_j|_y} \sigma_{\mathbf{Y}_i|_y}}, \qquad (1)$$

where $|_x$ represents its first component and $|_y$ represents its second component, $\mathrm{cov}(\cdot, \cdot)$ calculates the covariance between two vectors. The jth non-vascular corner is used to predict the ith vascular corner motion flow if $\rho_{i,j} > \rho_{th}$ where ρ_{th} is a predefined threshold. Then, least square serves as linear regressor between $\mathbf{X}_j$ and $\mathbf{Y}_i$ as (2) shows. And ρ and $\hat{\mathbf{Q}}$ are correlation model parameters, which can be used to predict vascular corner motion flow by non-vascular motion flow.

$$\hat{\mathbf{Q}}_{i,j} = \begin{bmatrix} \hat{\mathrm{a}}_x & \hat{\mathrm{a}}_y \\ \hat{\mathrm{b}}_x & \hat{\mathrm{b}}_y \end{bmatrix}$$

$$\hat{\mathrm{a}}_x, \hat{\mathrm{b}}_x = \arg\min_{\mathrm{a,b}} \sum_{m=1}^{k} \|\mathrm{a} \cdot \mathbf{d}_m^{n(j)}|_x + \mathrm{b} - \mathbf{d}_m^{v(i)}|_x\|^2, \quad (2)$$

$$\hat{\mathrm{a}}_y, \hat{\mathrm{b}}_y = \arg\min_{\mathrm{a,b}} \sum_{m=1}^{k} \|\mathrm{a} \cdot \mathbf{d}_m^{n(j)}|_y + \mathrm{b} - \mathbf{d}_m^{v(i)}|_y\|^2$$

In summary, for each vascular corner, $\mathbf{c}_r^{v(i)}$ in $\mathbf{I}_r$, non-vascular points with high correlation are selected based on pre-defined threshold ρ_{th}. Then, fitting parameters between this vascular corner and each selected non-vascular corner point are calculated by (2). The procedure of establishing the motion-related model is shown in Algorithm 1.

Algorithm 1 Motion-related model estimation process

Input: reference frame $\mathbf{I}_r$, contrasted sequences $\mathbf{I}$
Output: weight matrix $\mathbf{W} \in \mathbb{R}^{N_v \times N_n}$ and fitting parameters matrix $\mathbf{L}^A, \mathbf{L}^B \in \mathbb{R}^{N_v \times N_n \times 2}$
1: Initialize $\mathbf{W}$ and $\mathbf{L}^A, \mathbf{L}^B$ with zeros matrix
2: Calculate training vascular motion flows $\{\mathbf{D}_1^v, \mathbf{D}_2^v, ..., \mathbf{D}_k^v\}$ and training non-vascular motion flows $\{\mathbf{D}_1^n, \mathbf{D}_2^n, ..., \mathbf{D}_k^n\}$
3: **for** each vascular corner $i \in [1, N_v]$ **do**
4:     **for** each non-vascular corner $j \in [1, N_n]$ **do**
5:         Calculate Pearson coefficient $\rho_{i,j}$ by (1)
6:         **if** $\rho_{i,j} > \rho_{th}$ **then**
7:             $\mathbf{W}_{i,j} = \rho_{i,j}$
8:             Calculate fitting parameters $\hat{\mathbf{Q}}_{i,j}$ by (2)
9:             $\mathbf{L}_{i,j,.}^A = [\hat{\mathrm{a}}_x, \hat{\mathrm{a}}_y]$, $\mathbf{L}_{i,j,.}^B = [\hat{\mathrm{b}}_x, \hat{\mathrm{b}}_y]$
10:         **end if**
11:     **end for**
12:     Normalize weight vector $\mathbf{W}_{i,.} = \frac{\mathbf{w}_{i,.}}{\sum_{j=1}^{N_n} \mathbf{w}_{i,j}}$
13: **end for**

Notion: the for loop is implemented in vectorized form for efficiency.

The ith vascular corner motion flow is predicted by the weighted average of all selected non-vascular motion flows according to

$$\hat{\mathbf{f}}_q^{v(i)} = (\mathbf{W}_{i,.} \cdot (\mathbf{L}_{i,.,.}^A \odot \mathbf{F}_q^n + \mathbf{L}_{i,.,.}^B))^\top, \qquad (3)$$

where $\odot$ denotes Hadamard product (element-wise multiplication), $\hat{\mathbf{f}}_q^{v(i)} \in \mathbb{R}^{2 \times 1}$ is the predicted ith vascular corner motion flow between $\mathbf{R}_q$ and $\mathbf{I}_r$. However, sparse motion flows of some non-vascular corner points between $\mathbf{R}_q$ and $\mathbf{I}_r$ may have large errors, which makes the predicted vascular motion flow $\hat{\mathbf{f}}_q^{v(i)}$ sometimes inaccurate. To refine vascular motion flow prediction, we propose GOF-based flow motion prediction in the following subsection.

D. GOF-based Motion Flow Predicting

GOF-based process refines vascular motion prediction $\hat{\mathbf{f}}_q^{v(i)}$ from the motion-related model and deletes outliers based on Gaussian distribution. In order to reduce the influence of large errors between $\mathbf{R}_q$ and $\mathbf{I}_r$, we assume the predictions are i.i.d and follow Gaussian distribution.

For the ith vascular corner, with fitting coefficients $\mathbf{L}^A$ and $\mathbf{L}^B$, its non-vascular prediction $\hat{\mathbf{P}}_i = [\hat{\mathbf{p}}_i^{(1)}, \hat{\mathbf{p}}_i^{(2)}, ..., \hat{\mathbf{p}}_i^{(N_n)}]^\top \in \mathbb{R}^{N_n \times 2}$ is inferred by

$$\hat{\mathbf{P}}_i = \mathbf{L}_{i,.,.}^A \odot \mathbf{F}_q^n + \mathbf{L}_{i,.,.}^B, \qquad (4)$$

where $\hat{\mathbf{p}}_i^{(j)} \in \mathbb{R}^{2 \times 1}$ represents the jth non-vascular prediction for the ith vascular corner. Ideally, for any $j \in [1, N_n]$, the value of $\hat{\mathbf{p}}_i^{(j)}$ should be equivalent since they belong to the same vascular point. Therefore, each element in $\mathbf{P}_i$ should obey Gaussian distribution in both directions. That is variable $\mathbf{P}_i|_x \sim \mathcal{N}(\mu_i|_x, \sigma_i^2|_x)$, $\mathbf{P}_i|_y \sim \mathcal{N}(\mu_i|_y, \sigma_i^2|_y)$ where $\mu_i, \sigma_i \in \mathbb{R}^2$ are mean and standard deviation of the ith vascular corner predicting $\hat{\mathbf{P}}_i$. μ_i and σ_i can be statistically calculated by (5). For a random variable $Z \sim \mathcal{N}(\mu, \sigma^2)$, it has a probability 0.9974 within the range of $(\mu - 3\sigma, \mu + 3\sigma)$. Therefore, the outlier in $\hat{\mathbf{P}}_i$ caused by non-vascular motion flow estimating error can be deleted based on the 3σ bound, which updates motion-related model $\mathbf{W}$ as shown in Algorithm 2. Then each vascular corner motion flow prediction can be refined by (6) utilizing updated weight $\tilde{\mathbf{W}}$. Algorithm 2 describes how to predict vascular motion flows based on GOF, which outputs vascular motion flows $\tilde{\mathbf{F}}_q^v \in \mathbb{R}^{N_v \times 2}$. Finally, the reference frame vascular mask $\mathbf{M}_r$ is mapped onto the live image $\mathbf{R}_q$ according to the predicted vascular motion flows $\tilde{\mathbf{F}}_q^v$.

$$\mu_i|_x = \frac{\sum_{j=1}^{N_n} \hat{\mathbf{p}}_i^{(j)}|_x}{N_n}, \sigma_i|_x = \sqrt{\frac{\sum_{j=1}^{N_n} (\hat{\mathbf{p}}_i^{(j)}|_x - \mu_i|_x)^2}{N_n}},$$
$$\mu_i|_y = \frac{\sum_{j=1}^{N_n} \hat{\mathbf{p}}_i^{(j)}|_y}{N_n}, \sigma_i|_y = \sqrt{\frac{\sum_{j=1}^{N_n} (\hat{\mathbf{p}}_i^{(j)}|_y - \mu_i|_y)^2}{N_n}}, \qquad (5)$$

$$\tilde{\mathbf{f}}_q^{v(i)} = (\tilde{\mathbf{W}}_{i,.} \cdot \hat{\mathbf{P}}_i)^\top. \qquad (6)$$

Algorithm 2 Predicting vascular motion flows based on GOF.

Input: weight matrix $\mathbf{W}$, linear fitting coefficient $\mathbf{L}^A, \mathbf{L}^B$, live image $\mathbf{R}_q$, reference frame $\mathbf{I}_r$

Output: predicted vascular motion flows $\tilde{\mathbf{F}}_q^v \in \mathbb{R}^{N_v \times 2}$ between $\mathbf{R}_q$ and $\mathbf{I}_r$

1: Calculate non-vascular motion flows $\mathbf{F}_q^n \in \mathbb{R}^{N_n \times 2}$ between $\mathbf{R}_q$ and $\mathbf{I}_r$

2: **for** each vascular corner $i \in [1, N_v]$ **do**

3:     Calculate non-vascular predicting $\hat{\mathbf{P}}_i$ by (4)

4:     /* Delete outlier */

5:     Calculate μ_i, σ_i by (5)

6:     **for** each non-vascular predicting $j \in [1, N_n]$ **do**

7:         **if** $\hat{\mathbf{p}}_i^{(j)}|_x \notin (\mu_i|_x - 3\sigma_i|_x, \mu_i|_x + 3\sigma_i|_x)$ or $\hat{\mathbf{p}}_i^{(j)}|_y \notin (\mu_i|_y - 3\sigma_i|_y, \mu_i|_y + 3\sigma_i|_y)$ **then**

8:            $\mathbf{W}_{i,j} = 0$

9:         **end if**

10:     **end for**

11:     Re-normalize weight $\tilde{\mathbf{W}}_{i,\cdot} = \frac{\mathbf{W}_{i,\cdot}}{\sum_{j=1}^{N_n} \mathbf{W}_{i,j}}$

12:     Predict vascular motion flow $\tilde{\mathbf{f}}_q^{v(i)}$ by (6)

13: **end for**

TABLE I: Number of images used to compensate vascular motion. Seq.A-Seq.M represents different sequences.

	Seq.A	Seq.B	Seq.C	Seq.D	Seq.E	Seq.F	Seq.G	Seq.H	Seq.I	Seq.J	Seq.K	Seq.L	Seq.M
Training	18	12	14	11	11	16	12	10	11	8	11	7	11
Testing	42	118	11	103	44	40	83	20	15	7	20	8	9
Labeled	22	22	11	22	22	22	22	20	15	7	20	8	9

IV. EXPERIMENT RESULTS

A. Experimental Setup

To validate our proposed MRC algorithm, X-ray sequences generated during hepatic vascular surgeries TACE and TIPS from Zhongshan Hospital were collected, and we also obtained X-ray image sequences generated from a male porcine. 13 image sequences are screened with contrast agents. These image sequences include frames with and without contrast agents. During the process, the patient breathed freely. The images are with the size either in 512×512 or in 1024×1024 and the pixel resolution either $0.746mm$, $0.308mm$, or $0.390mm$. The Region of Interest of the image ranges from 216×231 to 562×726. For the images without the contrast agent, physicians [1] manually labeled the vascular centerlines for each sequence, which was used as *the reference* for validation. We performed the proposed MRC on a total of 520 frames and quantitative evaluation on 222 labeled frames. Table I lists detailed information on 13 sequences used for MRC. We also collected X-ray image sequences generated during coronary artery surgeries, but we did not count them due to poor image quality and large deformation.

Our MRC algorithm was compared with two other state-of-the-art algorithms, WSSD [20] and CRD [7]. To show that linear regression is suitable for robot application, we compare it with a typical nonlinear regression method named Gaussian Process Regression (GPR)[2][28]. All the hyperparameters were pre-tuned to guarantee the best performance. All experiments were conducted on a commercial desktop with Intel Core

---

[1]The second and third authors obtained M.D. degrees and did the score.

[2]Details can be found in the supplementary material.

---

i5-10500 CPU at 3.10GHz with 16Gb memory. The hyperparameter of our proposed MRC ρ_{th} was set as 0.9 in our experiments.

B. Evaluation criteria

Vascular MRC algorithms can be evaluated in terms of time and accuracy. The accuracy can be evaluated qualitatively by experienced physicians and quantitatively with the average Euclidean distance between predicted and *reference* frames. In this paper, we adopt ratio R and mean Euclidean distance MD to quantitatively evaluate algorithms accuracy, which are calculated by

$$R = \frac{|\mathbf{M}_{gt} \cap \mathcal{M}(\mathbf{M}_r, \tilde{\mathbf{F}}_q^v)|}{|\mathbf{M}_{gt}|}, \tag{7}$$

$$MD = \frac{1}{N_p} \sum_{m=1}^{N_p} \|g(\mathbf{M}_{gt})(m) - g(\mathcal{M}(\mathbf{M}_r, \tilde{\mathbf{F}}_q^v))(m)\|_2, \tag{8}$$

where $\mathbf{M}_{gt}$ is the labeled *reference* centerline, function $\mathcal{M}$ maps $\mathbf{M}_r$ onto live image $\mathbf{R}_q$ using motion $\tilde{\mathbf{F}}_q^v$, g extracts centerline point coordinates of mask, N_p is the number of centerline points in $\mathbf{M}_{gt}$.

C. Experiment Results

We conducted vascular compensation experiments on 13 sequences. It should be noted that CRD [7] was tested on 4 sequences because data input required the image to contain the liver's top. For Seq.L and Seq.M, GPR is not conducted because of the very long training time caused by large data. What's more, our MRC's error is low enough (mean $MD = 0.652mm$ and mean $MD = 0.523mm$). Even if GPR's error on Seq.L and Seq.M is zero, the average accuracy of GPR is lower than our MRC.

1) Accuracy: To assess the compensation accuracy, two clinicians scored them using the compensated roadmaps. Table II shows the mean score, which indicates that MRC possesses the best visualization for clinicians. Fig. 3 shows some sample results.[3] It should be noted that in order to provide better visualization, only the image of the vascular region is shown. As Fig. 3 indicates, there are cases where the results of WSSD, GPR, and MRC look the same, such as for sequences A, C, D, E, H, I, and M. In some cases, MRC works better than WSSD and GPR, such as in sequences B, F, and J. For sequences G, K, L, and M, MRC and GPR look the same and better than WSSD. CRD is worse than WSSD and MRC in sequences H, L, and M. According to the frame image shown in the figure, the result of CRD is the same as WSSD and MRC for sequences E. However, in some frames of sequence E, the visualization results are worse than that of WSSD and MRC. And its overall results vary considerably, which can be concluded from Fig. 4 and Fig. 5. Although the error variance of CRD can be decreased by increasing the number of training X-ray images, patients suffer from higher doses of X-ray.

---

[3]We strongly recommend readers watch the video to appreciate the results.
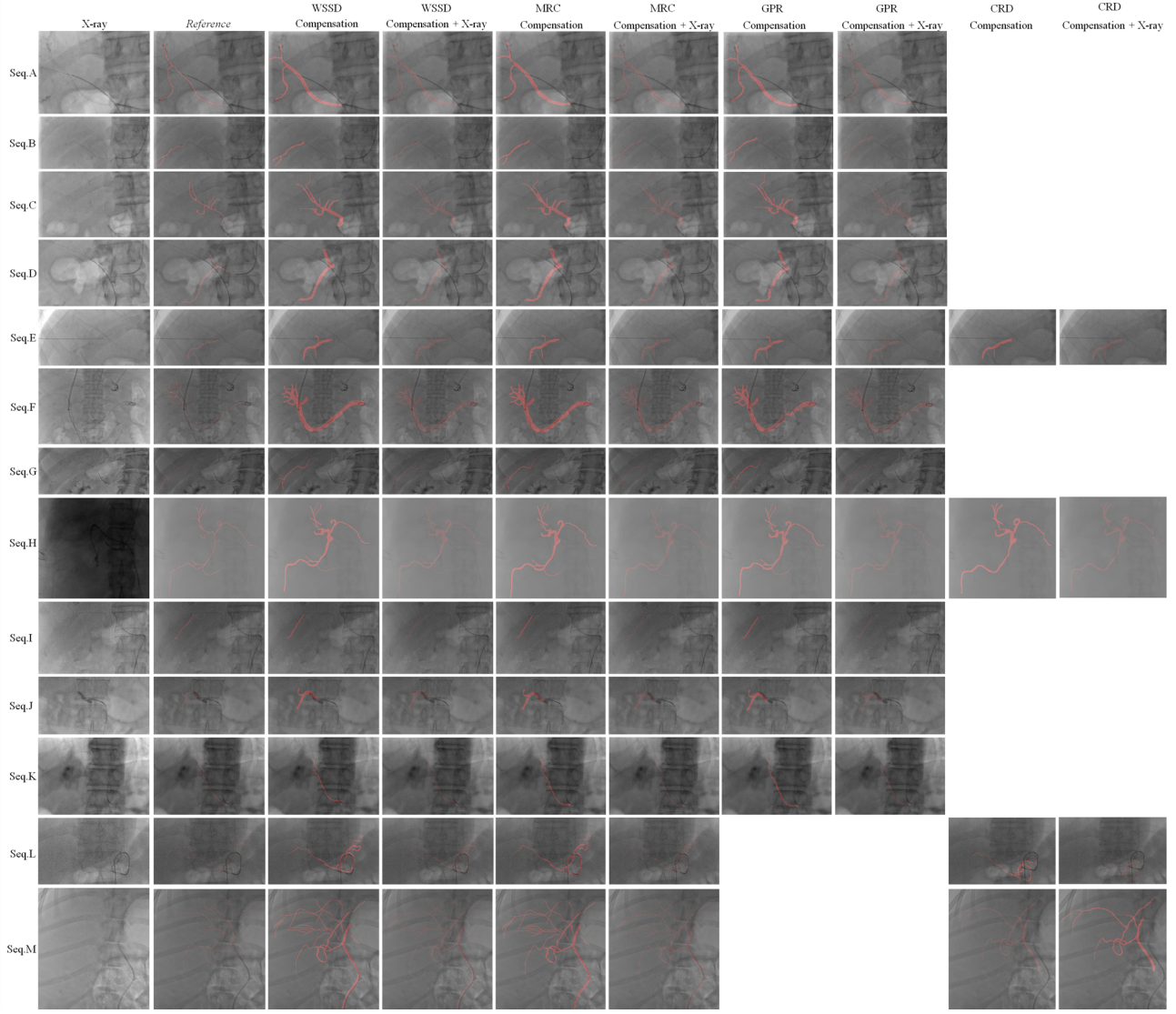
Fig. 3: Visualization of vascular respiratory motion compensation. Each row represents one sample frame. The first column is the original image, the second column is the *reference* labeled by experts, and the remaining eight columns are the results of WSSD, MRC, GPR, and CRD algorithms. For every algorithm, the compensation column uses bright red to emphasize the vascular mask, and the compensation + X-ray column uses dark red to observe accuracy.

TABLE II: The mean score of two clinicians grading according to the visualization roadmap. 5 is a perfect score. Seq.A-Seq.M represents different sequences.

	Seq.A	Seq.B	Seq.C	Seq.D	Seq.E	Seq.F	Seq.G	Seq.H	Seq.I	Seq.J	Seq.K	Seq.L	Seq.M	mean
WSSD	3	2	3.5	3.5	4.5	2	2	3.5	3	2.5	3	4	3.5	3.077
MRC	4	3	4	3.5	4.5	4	3	4.5	4	4.5	4	4.5	4	3.962
GPR	2.5	2	3	2	4	1.5	1.5	4	3.5	3	3.5	-	-	2.773
CRD	-	-	-	-	2.5	-	-	0.1	-	-	-	0.1	0.1	0.7

Results of the mean Euclidean distance on 13 sequences are shown in Fig. 4. It can be seen that MD of CRD is dispersed, and its mean and median are higher. For sequences A, B, F, G, H, I, K, L, and M MRC performs best on MD. For sequences C, D, and E, GPR possesses minimum error on MD. For sequences J, WSSD performs best on MD. Fig. 4 (b) shows that MRC achieves the best accuracy.

Fig. 5 shows the results of ratio R. 1 means the best prediction, and 0 indicates failure. It can be seen that R of CRD is lower than the other three algorithms. For sequences B, D, F, G, H, I, K, and L, MRC possesses the highest R.

For sequences A, C, E, J, and M, WSSD performs best on R than the other three algorithms. In general, our proposed MRC performs best on R as shown in Fig. 5 (b). The error of the sparse corners alignment algorithm varies with the image and can not be estimated, which makes GPR's accuracy vary with the accuracy of the additive Gaussian noise prior. To sum up, our MRC performs best on accuracy, which benefits from a reasonable motion model in that we consider the difference of motion in different positions incurred by respiration.

2) Time Consumption: Real-time performance is critical for clinical usage. Both learning time consumption and motion-

TABLE III: The model learning time[s] for different sequences and different methods. Seq.A-Seq.M represents different sequences.

	Seq.A	Seq.B	Seq.C	Seq.D	Seq.E	Seq.F	Seq.G	Seq.H	Seq.I	Seq.J	Seq.K	Seq.L	Seq.M
WSSD	110.127	82.646	121.894	68.250	272.435	376.654	593.059	232.944	262.118	117.140	150.492	152.328	235.327
MRC	**0.077**	**0.029**	**0.041**	**0.029**	**0.060**	**0.404**	**0.221**	**0.071**	**0.076**	**0.044**	**0.039**	**0.044**	**0.089**
GPR	10935.236	4978.496	13922.305	7642.906	28351.750	268172.254	194248.282	29759.035	5347.611	15609.621	10876.458	-	-
CRD	-	-	-	-	12.641	-	-	10.755	-	-	-	5.105	18.140

TABLE IV: The mean of time[s] of each frame to predict vascular motion on 13 sequences. Seq.A-Seq.M represents different sequences.

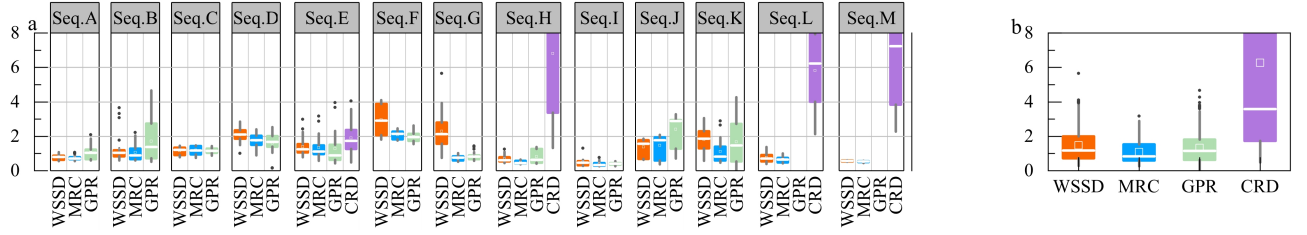|	Seq.A	Seq.B	Seq.C	Seq.D	Seq.E	Seq.F	Seq.G	Seq.H	Seq.I	Seq.J	Seq.K	Seq.L	Seq.M	Mean
WSSD	0.028	0.050	0.041	0.033	0.134	0.130	0.355	0.091	0.124	0.090	0.093	0.156	0.143	0.116
MRC	**0.012**	**0.019**	**0.008**	**0.005**	0.030	**0.124**	**0.046**	0.016	**0.011**	**0.035**	**0.029**	**0.040**	0.155	0.032
GPR	17.265	7.785	22.405	12.852	47.526	520.735	328.753	50.823	8.835	27.107	17.027	-	-	109.543
CRD	-	-	-	-	0.002	-	-	0.001	-	-	-	0.002	0.002	0.002

Fig. 4: The box chart of mean Euclidean distance[mm] on 13 sequences. (a) shows MD on 13 sequences respectively, Seq.A-Seq.M represents different sequences. (b) shows the overall results on all 13 sequences. It should be noted that the overall results of CRD are on 4 sequences.
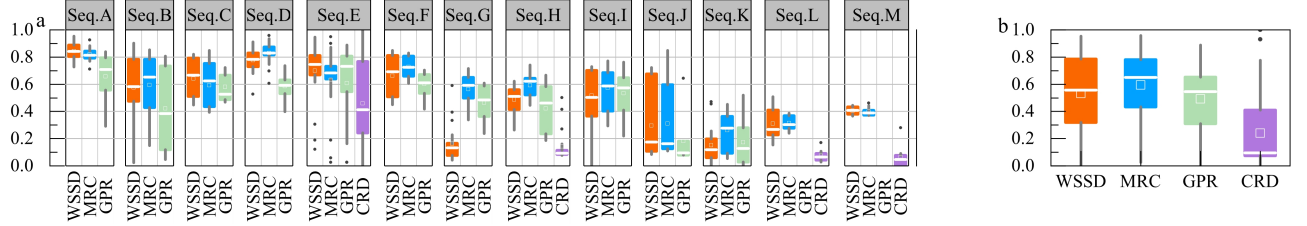

Fig. 5: The box chart of value of ratio R on 13 sequences. (a) shows R on 13 sequences respectively, Seq.A-Seq.M represents different sequences. (b) shows the overall results on all 13 sequences. It should be noted that the overall results of CRD are on 4 sequences.

prediction time consumption are tested. The model learning time is shown in Table III. The proposed MRC performs well on model learning time. The mean of each frame predicting vascular motion time is shown in Table IV. It demonstrates that the prediction time of MRC is less than WSSD and GPR on all 13 sequences. Although GPR can be accelerated [26, 29], its time cannot exceed linear regression. For sequences E, H, L, and M, the prediction time of CRD is less than MRC, but its robustness and accuracy are limited. In the meantime, application scenarios of CRD need the FOV of the image including the top of the liver. In summary, our MRC performs best on learning time and second best on prediction time because of relatively simple modeling and sparse point tracker. CRD is the fastest approach since it estimates the respiratory state from the image to predict motion. However, it is not robust if the respiration pattern changes.

3) Ablation Study: Ablation experiments were conducted to verify the efficiency of our GOF and sparse optical flow. The results are shown in Table V. Our MRC is tested with and without GOF. We also test the choice of dense and sparse optical flow. Results indicate that sparse optical flow can significantly reduce time consumption. GOF can improve the accuracy of sparse optical flow prediction with some time

TABLE V: Results of ablation experiments in accuracy and time. Sparse and dense represent sparse motion flow and dense motion flow.

	MD[mm]	R	prediction time[s]	Learning time[s]
sparse with GOF	**1.086**	**0.595**	0.032	0.092
sparse without GOF	1.565	0.432	**0.012**	**0.091**
dense with GOF	1.192	0.585	10.587	26.691
dense without GOF	1.192	0.587	4.842	21.524

consumption. Still, GOF is basically ineffective for dense optical flow prediction due to the more accurate dense optical flow algorithm. Thus, sparse motion flow with GOF possesses compromised accuracy and time.

D. Limitations & Future Works

Our proposed MRC comes with three drawbacks. First, only one frame contrasted image vascular mask is used to map onto the live image, the mapped vascular mask size may be small. Future works will conduct the fusion multi-frames contrasted images of vascular information to enrich the vascular mask. Secondly, the predicted vascular motion flow is not smooth enough due to the sparse flow of vascular. Although the local region is not smooth, it does not affect physicians' overall judgment. Lastly, it cannot handle the scenario with heart-beat

because of poor image quality, leading to the error in optical flow[4]. We plan to utilize optical flow based on deep learning to compute flow motion.

V. Conclusion

We propose MRC to conduct vascular respiratory motion compensation in real-time to predict vascular on the live fluoroscopic image with invisible vascular. Based on the linear correlation between vascular motion flow and non-vascular motion flow, multi-frame contrasted images are used to train a motion-related model. In the prediction stage, predictions from non-vascular points are refined with GOF. The proposed method was tested on 13 in-vivo vascular intervention fluoroscopic sequences. Results show that the proposed method achieves a compensation accuracy of $1.086\ mm$ in $0.032\ s$. Our approach provides a practical real-time solution to assist physicians/robots in vascular interventions. This work is in the process of commercialization by the company United Imaging of Health Co., Ltd.

References

[1] M. Rössle, "TIPS: 25 years later," *Journal of hepatology*, vol. 59, no. 5, pp. 1081–1093, 2013.

[2] M. Tsurusaki and T. Murakami, "Surgical and locoregional therapy of HCC: TACE," *Liver Cancer*, vol. 4, no. 3, pp. 165–175, 2015.

[3] P. Ambrosini, D. Ruijters, W. J. Niessen, A. Moelker, and T. van Walsum, "Continuous roadmapping in liver tace procedures using 2D–3D catheter-based registration," *International journal of computer assisted radiology and surgery*, vol. 10, pp. 1357–1370, 2015.

[4] W. Duan, T. Akinyemi, W. Du, J. Ma, X. Chen, F. Wang, O. Omisore, J. Luo, H. Wang, and L. Wang, "Technical and clinical progress on robot-assisted endovascular interventions: A review," *Micromachines*, vol. 14, no. 1, p. 197, 2023.

[5] C. N. Riviere, J. Gangloff, and M. De Mathelin, "Robotic compensation of biological motion to enhance surgical accuracy," *Proceedings of the IEEE*, vol. 94, no. 9, pp. 1705–1716, 2006.

[6] C. Unger, M. Groher, and N. Navab, "Image based rendering for motion compensation in angiographic roadmapping," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.* IEEE, 2008, pp. 1–8.

[7] M. G. Wagner, S. Periyasamy, C. Longhurst, M. J. McLachlan, J. F. Whitehead, M. A. Speidel, and P. F. Laeseke, "Real-time respiratory motion compensated roadmaps for hepatic arterial interventions," *Medical physics*, vol. 48, no. 10, pp. 5661–5673, 2021.

[8] P. Giraud and A. Houle, "Respiratory gating for radiotherapy: main technical aspects and clinical benefits," *ISRN Pulmonology*, vol. 2013, pp. 1–13, 2013.

[9] D. Spinczyk, A. Karwan, and M. Copik, "Methods for abdominal respiratory motion tracking," *Computer Aided Surgery*, vol. 19, no. 1-3, pp. 34–47, 2014.

[10] S. Sayeh, J. Wang, W. T. Main, W. Kilby, and C. R. Maurer Jr, "Respiratory motion tracking for robotic radiosurgery," in *Treating tumors that move with respiration.* Springer, 2007, pp. 15–29.

[11] P. M. Loschak, A. Degirmenci, C. M. Tschabrunn, E. Anter, and R. D. Howe, "Automatically steering cardiac catheters in vivo with respiratory motion compensation," *Int. J. Robot. Res.*, vol. 39, no. 5, pp. 586–597, 2020.

[12] M. G. Wagner, P. F. Laeseke, T. Schubert, J. M. Slagowski, M. A. Speidel, and C. A. Mistretta, "Feature-based respiratory motion tracking in native fluoroscopic sequences for dynamic roadmaps during minimally invasive procedures in the thorax and abdomen," in *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 10135. SPIE, 2017, pp. 382–390.

[13] A. P. King, R. Boubertakh, K. S. Rhode, Y. Ma, P. Chinchapatnam, G. Gao, T. Tangcharoen, M. Ginks, M. Cooklin, J. S. Gill *et al.*, "A subject-specific technique for respiratory motion correction in image-guided cardiac catheterisation procedures," *Med. Img. Ana.*, vol. 13, no. 3, pp. 419–431, 2009.

[14] P. Fischer, T. Pohl, A. Faranesh, A. Maier, and J. Hornegger, "Unsupervised learning for robust respiratory signal estimation from X-ray fluoroscopy," *IEEE Trans. Med. Imag.*, vol. 36, no. 4, pp. 865–877, 2016.

[15] M. Henningsson and R. M. Botnar, "Advanced respiratory motion compensation for coronary MR angiography," *Sensors*, vol. 13, no. 6, pp. 6882–6899, 2013.

[16] I. Vernikouskaya, D. Bertsche, W. Rottbauer, and V. Rasche, "Deep learning-based framework for motion-compensated image fusion in catheterization procedures," *Computerized Medical Imaging and Graphics*, vol. 98, p. 102069, 2022.

[17] H. Ma, I. Smal, J. Daemen, and T. van Walsum, "Dynamic coronary roadmapping via catheter tip tracking in X-ray fluoroscopy with deep learning based Bayesian filtering," *Medical image analysis*, vol. 61, p. 101634, 2020.

[18] S. Atasoy, M. Groher, D. Zikic, B. Glocker, T. Waggershauser, M. Pfister, and N. Navab, "Real-time respiratory motion tracking: roadmap correction for hepatic artery catheterizations," in *Medical Imaging 2008: Visualization, Image-Guided Procedures, and Modeling*, vol. 6918. SPIE, 2008, pp. 408–416.

[19] M.-C. V. Orozco, S. Gorges, and J. Pescatore, "Respiratory liver motion tracking during transcatheter procedures using guidewire detection," *International Journal of Computer Assisted Radiology and Surgery*, vol. 3, no. 1-2, pp. 79–83, 2008.

[20] Y. Zhu, Y. Tsin, H. Sundar, and F. Sauer, "Image-based respiratory motion compensation for fluoroscopic coronary roadmapping," in *Int. Conf. on Med. Image Comput. and Comput. Assist. Interv.* Springer, 2010, pp. 287–294.

[21] M. Manhart, Y. Zhu, and D. Vitanovski, "Self-assessing image-based respiratory motion compensation for fluoroscopic coronary roadmapping," in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro.* IEEE, 2011, pp. 1065–1069.

[22] J. Shi *et al.*, "Good features to track," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.* IEEE, 1994, pp. 593–600.

[23] J. Liu, X. Li, Y. Liu, and H. Chen, "RGB-D inertial odometry for a resource-restricted robot in dynamic environments," *IEEE Robot. and Autom. Lett.*, vol. 7, no. 4, pp. 9573–9580, 2022.

[24] J.-Y. Bouguet *et al.*, "Pyramidal implementation of the affine Lucas Kanade feature tracker description of the algorithm," *Intel corporation*, vol. 5, no. 1-10, p. 4, 2001.

[25] K. Fujimoto, T. Shiinoki, Y. Yuasa, and H. Tanaka, "Estimation of liver elasticity using the finite element method and four-dimensional computed tomography images as a biomarker of liver fibrosis," *Medical Physics*, vol. 48, no. 3, pp. 1286–1298, 2021.

[26] J. Kabzan, L. Hewing, A. Liniger, and M. N. Zeilinger, "Learning-based model predictive control for autonomous racing," *IEEE Robot. and Autom. Lett.*, vol. 4, no. 4, pp. 3363–3370, 2019.

[27] R. P. David Freedman, Robert Pisani, *Statistics*. W. W. Norton & Company, 2007.

[28] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.

[29] A. Ranganathan, M.-H. Yang, and J. Ho, "Online sparse gaussian process regression and its applications," *IEEE Transactions on Image Processing*, vol. 20, no. 2, pp. 391–404, 2010.

---

[4]The last video clip shows the failure case.