Real-Time Sensitivity Analysis of Blood Flow Simulations to Lumen Segmentation Uncertainty

Sethuraman Sankaran, Leo J. Grady, and Charles A. Taylor

HeartFlow Inc., 1400 Seaport Blvd. Bldg B, Redwood City, CA 94063, USA

Abstract. Patient-specific modeling of blood flow combining CT image data and computational fluid dynamics has significant potential for assessing the functional significance of coronary artery disease. An accurate segmentation of the coronary arteries, an essential ingredient for blood flow modeling methods, is currently attained by a combination of automated algorithms with human review and editing. However, not all portions of the coronary artery tree affect blood flow and pressure equally, and it is of significant importance to direct human review and editing towards regions that will most affect the subsequent simulations. We present a data-driven approach for real-time estimation of sensitivity of blood-flow simulations to uncertainty in lumen segmentation. A machine learning method is used to map patient-specific features to a sensitivity value, using a large database of patients with precomputed sensitivities. We validate the results of the machine learning algorithm using direct 3D blood flow simulations and demonstrate that the algorithm can predict sensitivities in real time with only a small reduction in accuracy as compared to the 3D solutions. This approach can also be applied to other medical applications where physiologic simulations are performed using patient-specific models created from image data.

Keywords: sensitivity analysis, blood flow simulations, machine learning, segmentation accuracy.

1 Introduction

Coronary artery disease (CAD) affects hundreds of millions of people worldwide. Current noninvasive tests for CAD either provide data on anatomic narrowing (e.g. via coronary computed tomography angiography (cCTA)) or data on myocardial dysfunction (e.g. single photon emission computed tomography (SPECT)), but cannot identify whether specific coronary lesions are functionally significant. Recently, fractional flow reserve (FFR) has emerged as the invasive gold-standard for determining which lesions in the coronary arteries are flow-limiting and should be stented, and which patients should be treated medically. FFR is defined as the ratio of blood flow rate under conditions of maximal hyperemia (reduced myocardial bed resistance) at a given location to the hypothetical value if no disease were present. Under modest assumptions, the FFR can be shown to be equal to the ratio of local coronary artery blood pressure

P. Golland et al. (Eds.): MICCAI 2014, Part II, LNCS 8674, pp. 1–8, 2014.

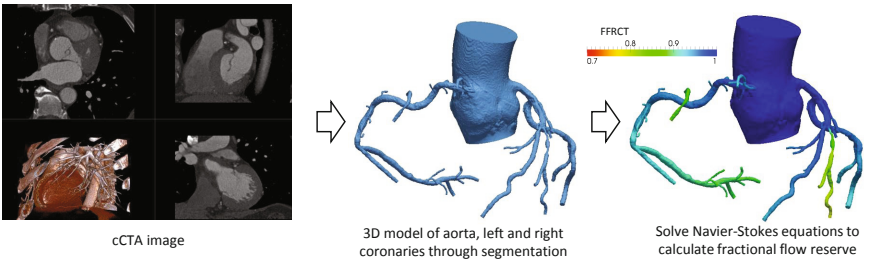

Fig. 1. An overview of the process of calculating FFR_{CT} from cCTA images. The coronary arteries of interest as well as a portion of the aorta are segmented using standard algorithms [5,6]. The Navier-Stokes equations are solved to calculate FFR_{CT}, which has been shown to have high accuracy in prediction of measured FFR [4].

to aortic blood pressure measured invasively using pressure wires under maximal hyperemic conditions. Recent developments in patient-specific modeling and blood flow simulations have enabled the computation of FFR non-invasively from cCTA data, referred to as FFR_{CT} [1] (refer Figure. 1). Data from three multi-center clinical trials indicates that prediction of FFR_{CT} using blood flow simulations significantly improves the noninvasive assessment of coronary artery disease [2,3,4].

Blood flow simulations rely on (i) accurate segmentation of the coronary artery geometry [7] from cCTA image data, (ii) accurate representation of the system properties such as blood viscosity, (iii) accurate enforcement of boundary conditions such as aortic pressure and the microvascular coronary bed resistance. Recent evidence suggests that the impact of inaccuracies in lumen segmentation is likely more significant than the other variables [8]. Specifically, errors in the minimum lumen diameter (MLD) can lead to significant errors in the FFR_{CT}. Hence, it is important to understand the relationship between geometry and FFR_{CT}, and how uncertainty in segmentation propagates to uncertainty in blood-flow simulations. This can help focus human review process on regions where improved segmentation accuracy is required based on a sensitivity cutoff.

Quantifying the effect of segmentation error is challenging because the nonlinear partial differential equations governing blood flow for a given arterial geometry are computationally complex. Traditional Monte-Carlo methods are infeasible due to computational expense. Even recently developed methods based on stochastic space representation [9] could take up to a few days with typical parallel computing resources. If sensitivity information is to be useful for user review, it must be real time.

We describe herein a machine learning method for calculating the sensitivity of FFR_{CT} computations to uncertainty in the lumen segmentation by training on patient-specific geometries with precomputed sensitivities. We pick an exhaustive set of features encompassing geometric, clinical and analytical model based features, such as analytical solutions for pipe flow parameterized by vessel radii and length, flow rate, and blood viscosity. The machine learning predictor

encodes the stochastic collocation algorithm in addition to the Navier-Stokes equation. Results show promise in using this method for real-time estimation of sensitivity.

2 Method

Our analysis starts with a three-dimensional geometric model encompassing ascending aorta and the coronary arteries segmented from cCTA data.

2.1 Review of Local Sensitivity Analysis

We define geometric sensitivity as the standard deviation in $\mathrm{FFR}_{\mathtt{CT}}$, $\sigma_{\mathrm{FFR}_{\mathtt{CT}}}$, due to uncertainty in the segmentation. We are interested in the impact of inaccuracies in a section of the segmentation on the simulation. Hence, we associate the maximum value of $\sigma_{\mathrm{FFR}_{\mathtt{CT}}}$ across the coronary tree with a chosen section. This value indicates the maximum impact of changing the cross sectional area of the given section, which could be either upstream or downstream of the given section.

Calculating $\sigma_{\mathrm{FFR}_{\mathtt{CT}}}$ involves (a) parameterizing geometry of the coronary artery tree and dividing it into sections, and (b) sampling the geometries among the full family of geometries where $\mathrm{FFR}_{\mathtt{CT}}$ will be estimated. We describe each of these steps below.

Parameterizing Geometry. While there are many ways to parameterize the family of geometries due to uncertainty in lumen segmentation, we use a scheme in which all points between two bifurcations, a bifurcation and an outlet, or ostium and a bifurcation are considered a single section.

Hence, the variability in all points that constitute a section can be described by a single underlying random variable. The number of independent sources of uncertainty depends on the size of coronary tree, number of bifurcations and outlets. The initial segmentation is only used as a convenient mechanism to define the family of geometries. If a lesion is missed, the sensitivity estimate will capture its impact as long as the family of geometries encompasses the new minimum lumen diameter. Our results do not depend strongly on the initial segmentation since we use a uniform probability distribution function in stochastic space.

Sampling Geometries. The distribution above represents an infinite set of possible geometries. The family of geometries is described using a stochastic space, where each axes of the stochastic space is an independent uniform variable that represents each section. The sensitivities (refer Fig. 2) are calculated by first computing the $\mathrm{FFR}_{\mathtt{CT}}$ across this stochastic space, and then computing standard deviation in $\mathrm{FFR}_{\mathtt{CT}}$ by sampling the stochastic space using Smolyak quadrature algorithm [9]. The stochastic collocation method shows significant improvement over Monte-Carlo method. However, the computational time is still of the order of few days, and hence not feasible to use in practice.

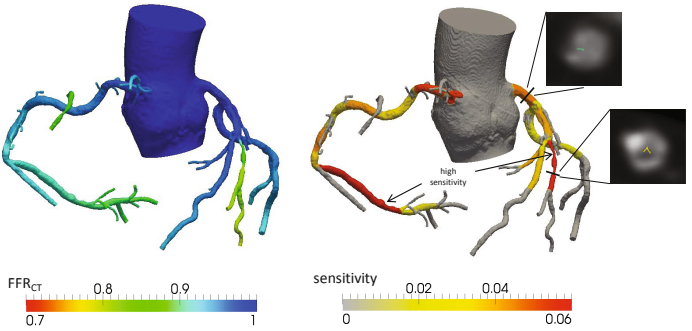

Fig. 2. FFR_{CT} map for a patient-specific coronary model (left) and the corresponding sensitivity map (right) highlighting regional impact of lumen area on FFR_{CT} map. Sensitivity information can save computational effort for user review of segmentation by focusing only on a small region of coronary tree. Unfortunately, existing methods for calculating sensitivity information could take hours to days, and the motivation is to accelerate computation of sensitivities. The figure also shows high sensitivities in regions of severe disease, though disease by itself is not a good predictor of sensitivity.

2.2 Estimation Method

The most expensive step in the estimation of local sensitivities is the blood flow simulations performed at the quadrature points. We use a surrogate model [8], which predicts FFR_{CT}, to generate training sensitivity data. We then extract salient features from each patient-specific model (refer Fig. 3). The feature set is enriched if necessary, by analyzing cases with large error.

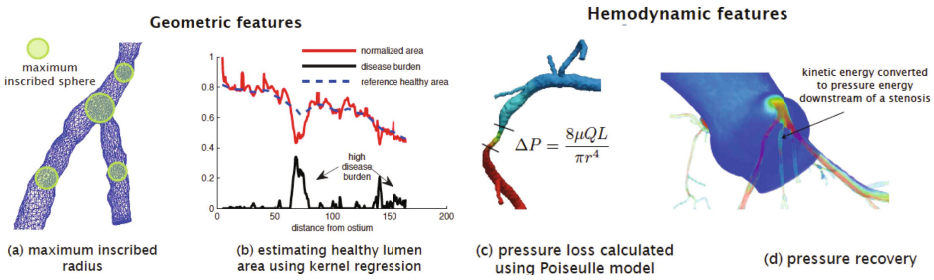

Fig. 3. We train a machine learning algorithm to accelerate sensitivity calculation shown in Fig. 2 using data from clinical trials. Salient features of the machine learning algorithm are shown. Geometric features used - (a) radius of maximum inscribed sphere, (b) disease burden score using a robust kernel regression fit, and hemodynamic features - (c) pressure loss and (d) pressure recovery due to drop in kinetic energy.

Geometric Features. We use a combination of local, upstream, downstream, and disease burden measures as geometric features associated with a centerline point. We use the radius of the maximum inscribed sphere as representative of

lumen radius. Upstream features used are volume of blood upstream, distance to nearest bifurcation, distance to ostium (location where aorta meets coronary artery), minimum upstream diameter, distance to minimum upstream diameter, number of upstream bifurcations, average upstream diameter, area of nearest upstream bifurcation and area of previous centerline point. Similar features are calculated for downstream vasculature, and in addition net downstream boundary resistance is calculated. In addition to these, a health index score is calculated which is defined as

$$\kappa(x) = \frac{r(x)}{r_{\mathtt{h}}(x)} \tag{1}$$

where $r_{\mathtt{h}}(x)$ is the theoretical healthy radius of the lumen and $r(x)$ is the radius of the maximum inscribed sphere. Motivated from Shahzad et.al. [10], we use three different regressors in this paper to calculate $r_{\mathtt{h}}$ given by the general form:

$$r_{\mathtt{h}}(x) = \frac{\sum_{x'=1}^{n} \mathcal{N}(x'|x, \sigma_x) S(x', x) w_{x'} r(x')}{\sum_{x'=1}^{n} \mathcal{N}(x'|x, \sigma_x) S(x', x) w_{x'}} \tag{2}$$

where $w_x = \mathcal{N}(r(x)|r_{x,\mathtt{max}}, \sigma_{\mathtt{max}})$ are Gaussian weighting functions, and n is the number of points used in the regression. The following regressors are used.

- global kernel fit, defined for each path from the root (ostium) to the leaves, where the function $S(x', x)$ is uniformly unity.
- segmental kernel fit, where the function $S(x', x)$ is unity only within the current section.
- anisotropic kernel fit where $S(x', x)$ is a sigmoidal function centered at the nearest branch designed to minimize the effect of sharp radius variation at the branch,

$$S(x', x) = \frac{1}{1 + \alpha e^{-kd_{\mathtt{offset}}(x', x)}}, \mathtt{and} \tag{3}$$

$$d_{\mathtt{offset}}(x', x) = d(x', x_{\mathtt{ostium}}) - d(x, x_{\mathtt{ostium}}) - d(x, x_{\mathtt{up}}), \tag{4}$$

  , where $x_{\mathtt{up}}$ is the location of the nearest upstream branch to x, and $d(.,.)$ is the euclidean distance.

Five parameters for σ_x, $\sigma_{\mathtt{max}}$, α and $r_{x,\mathtt{max}}$ were chosen for each of these regressors, making a total of 15 health index scores.

Hemodynamic Features. Using the downstream resistances and coronary flow rate, we approximate the flow rate through each section of the model. At each branch, flow splits based on the inverse of net downstream resistances (a series of resistances in parallel). The geometric resistance is obtained using Poiseulle's law as $R = \frac{8\mu L}{\pi r_{\mathtt{avg}}^4}$ and corresponding pressure loss as $\Delta P_i = RQ_i$. We also introduce a pressure recovery factor, based on observations in a few patients, where pressure can recover distal to a disease if the radius returns back to higher than the original healthy value and remains there. We define a cutoff of 20% (chosen empirically), i.e. if the radius increases to more than 20% of

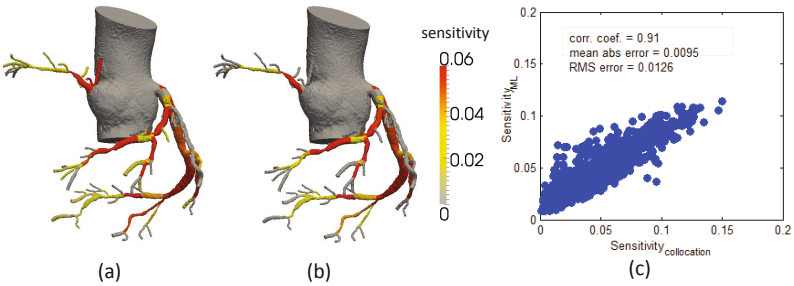

Fig. 4. (left) Comparison of sensitivities calculated using (a) simulation which takes a couple of days and (b) machine learning which takes around 10 seconds, and (c) performance of the machine learning regressor on testing set where each point represents a section of a vessel. Correlation coefficients quantify how well the sensitivity calculated using 3D simulations compare with those using machine learning, and the prediction error is quantified using RMS error and mean absolute error. There are around 50 sections per patient, and the graph was constructed over the testing set of 82 patients.

its pre-stenotic healthy value, then we define the pressure recovery factor to be $P_{\texttt{recovery}} = \frac{r_{\texttt{post}}}{r_{\texttt{pre}}}$.

2.3 Gather Data from Multiple Patients and Compute Sensitivities

Sensitivities for all the patients in the database were first computed by performing multiple simulations using the stochastic collocation algorithm [8]. Geometric and hemodynamic parameters are calculated for each centerline point within each geometric section, which are then aggregated using (i) mean, (ii) standard deviation, (iii) minimum and (iv) maximum values for each section.Standard transformations (square, square root, cube, cube root, etc.) on each of these variables is then performed. Two full sweeps of the coronary tree are needed, once from the root to the leaves to calculate all upstream features and once from the leaves to the root to calculate all downstream features. Two additional sweeps starting from the root are needed, one to calculate the flow rate and another to calculate pressure drop, since they involve a combination of upstream and downstream indices.

2.4 Use Machine Learning to Estimate Sensitivity

The goal is to find an optimal map between parameters identified in the previous section and the sensitivities. The optimal machine learning regressor was chosen using trial-and-error of various candidate regressors. Bootstrap aggregating (bagging) is performed to avoid over-fitting and reduce variance in test set. All the attributes listed above are aggregated and written to a training database, and analyzed using *Weka* [11].

3 Results

3.1 Training and Testing Set Description

We use 240 patient-specific models from previous clinical trials of FFR_{CT}. We split this into 158 patients comprising the training set and the rest comprising the testing set. The datasets comprised of patients with varying levels of coronary disease, distributed throughout the coronary arteries. Sensitivity analysis is performed on each of these patients first. Then, we traverse through the coronary tree to calculate all the geometric and hemodynamic features.

3.2 Analysis and Results

A comparison of the performance on the testing set indicated that a bagged decision tree regressor was optimal. We achieved a correlation coefficient of 0.910, mean absolute error of 0.010 and root mean square error of 0.013. A square transformation with five bootstrap aggregates yielded the optimal result. We used an inclusion threshold of 15 (minimum number of instances a relationship is detected). Low inclusion thresholds resulted in over-fitting to the training data and high inclusion thresholds resulted in an inadequate decision tree. Figure 4 shows performance of the machine learning regressor on the testing set where each point represents a section of the vessel. Figure 4 also shows a comparison of sensitivity map calculated using simulation and machine learning. The correlation coefficient between disease burden and sensitivity was 0.4, implying that disease burden alone is not a good predictor of sensitivity. **Speed:** A summary of the speed of the algorithm on the full dataset is shown in Fig. 5. For the sensitivity simulation, solving 3D Navier-Stokes equations takes multiple days on 90 cores of a computing cluster, while a surrogate model could reduce it to less than an hour and machine learning can further reduce it to a few seconds in a single processor workstation. This demonstrates that machine learning enables real time computation of sensitivity, which could then be used for user review of segmentation.

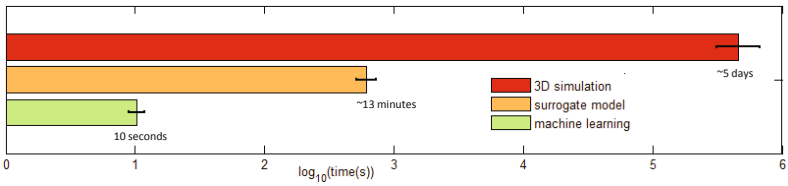

Fig. 5. Comparison of computational time to calculate sensitivities using full 3D simulations, reduced order model, and machine learning demonstrating that we can calculate sensitivities real time. Average time for 240 patients reduced from more than 3 days using full 3D simulations to 10 seconds per patient.

4 Conclusion

We demonstrated the use of a machine learning method for real-time calculation of geometric sensitivity information, since the relationship between local geometry and blood pressure and velocities governed by the Navier-Stokes equations is expensive to compute. The sensitivity map quantifies how much the $\mathrm{FFR}_{\mathtt{CT}}$ changes if the local lumen area is changed. We showed that the machine learning algorithm accelerates sensitivity computation from over a day using the 3D Navier-Stokes equations to just a few seconds. The sensitivity map can be used to focus, check and ensure lumen fidelity in the highly sensitive regions. This can also be used to pick parameters of geometry reconstruction which minimize error in the highly sensitive regions most relevant to the computation of $\mathrm{FFR}_{\mathtt{CT}}$. In the future, we could rank parameters in terms of their importance and perform feature selection. We can also explore performing machine learning without prediction of $\mathrm{FFR}_{\mathtt{CT}}$ by directly mapping features to sensitivities.

References

1. Taylor, C.A., Fonte, T.A., Min, J.K.: Computational fluid dynamics applied to cardiac computed tomography for noninvasive quantification of fractional flow reserve: scientific basis. JACC 61(22), 2233–2241 (2013)
2. Koo, B.K., et al.: Diagnosis of Ischemia-Causing Coronary Stenoses by Noninvasive Fractional Flow Reserve Computed From Coronary Computed Tomographic Angiograms Results From the Prospective Multicenter DISCOVER-FLOW Study. JACC 58(19), 1989–1997 (2011)
3. Min, J.K., et al.: Diagnostic accuracy of fractional flow reserve from anatomic CT angiography. Journal of American Medical Association 308(12), 1237–1245 (2012)
4. Norgaard, B.L., et al.: Diagnostic performance of non-invasive fractional flow reserve derived from coronary CT angiography in suspected coronary artery disease: The NXT trial. Journal of American College of Cardiology (2014), doi:10.1016.11.043
5. Schaap, M., et al.: Standardized evaluation methodology and reference database for evaluating coronary artery centerline extraction algorithms. Medical Image Analysis 13(5), 701–714 (2009)
6. Lesage, D., et al.: A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes. Medical Image Analysis 13(6), 819–845 (2009)
7. Karmonik, C., Brown, A., Debus, K., Bismuth, J., Lumsden, A.B.: CFD Challenge: Predicting Patient-Specific Hemodynamics at Rest and Stress through an Aortic Coarctation. In: Camara, O., Mansi, T., Pop, M., Rhode, K., Sermesant, M., Young, A. (eds.) STACOM 2013. LNCS, vol. 8330, pp. 94–101. Springer, Heidelberg (2014)
8. Sankaran, S., Grady, L.J., Taylor, C.A.: Fast geometric sensitivity analysis in hemodynamic simulations using machine learning. Journal of Computational Physics (2014) (under review)
9. Sankaran, S., Marsden, A.L.: A stochastic collocation method for uncertainty quantification and propagation in cardiovascular simulations. Journal of Biomechanical Engineering 133, 031001-1 (2011)
10. Shahzad, R., et al.: Detection and Quantification of Coronary Artery Stenoses on CTA. International Journal of Cardiovascular Imaging 29, 1–13 (2013)
11. Hall, M., et al.: The WEKA Data Mining Software: An Update. SIGKDD Explorations 11(1) (2009)