An Adaptive Optimal Control Design for a Bolus Chasing Computed Tomography Angiography

Er-Wei Bai, Fellow, IEEE, Zhijun Cai, Member, IEEE, Robert McCabe, and Ge Wang, Fellow, IEEE

Abstract—To improve imaging quality and to reduce contrast dose and radiation exposure, an optimal adaptive bolus chasing controller is proposed and tested based on actual patient data. The controller estimates and predicts the unknown two dimensional bolus density on line and then determines the optimal control actions. Tracking errors are mathematically quantified in terms of estimation errors. The test results not only support the analytical analysis and exhibit its superior performance over the current constant-speed controller, but also demonstrate the clinical feasibility.

Index Terms—CT Angiography, adaptive control, synchronization, tracking, estimation, polynomial approximation.

I. INTRODUCTION

Today, Computed Tomography Angiography (CTA) is becoming increasingly important for Peripheral Vascular Disease (PAD) with the advent of multi-row CT [15], [18]. To acquire higher quality vascular images, a certain amount of contrast medium (bolus) is injected into a vein before scanning. During the scanning, contrast medium propagates in the artery, and meanwhile the patient is fed into the CT gantry by translating the CT table. Scanning the blood vessels with the higher density contrast medium inside results in a greater Contrast to background Noise Ratio (CNR), which is desirable. Therefore, synchronizing the contrast bolus peak density position and CT gantry window is crucial to the CT performance. Unfortunately, the bolus dynamics are quite different and complicated for diverse patients and injection methods. Even for the same patient and injection method, bolus dynamics may not be the same under different situations. To that end, a pre-set tracking profile would not work for all patients at all times, and methods that ensure synchronization with the maximum density is most desirable from the standpoint of an optimal image quality [2], [9].

Currently, the rudimentary CT control techniques are triggered by a pre-set threshold or by a timing bolus. As for the former method, a threshold is set for the bolus density at a specific position, e.g., aorta, which is being monitored after the injection. When the observed density reaches the threshold, the CT table is automatically started [9]. However, it is very difficult for operators to choose the correct threshold [23]. If it is set too low, the CT table will be started too early, which results in bad tracking. If it is set too high, the CT table might not be started at all. The other method, referred to as the timing (test) bolus technique [16], calls for an injection of a small amount contrast prior to the main injection used for diagnosis. The arrival time of the timing bolus to the Region Of Interest (ROI) will be taken as the delay time to start the CT table for the actual contrast injection. Obviously, it is assumed that the timing bolus and actual bolus have the same peak arrival time, which may not be true in most cases [9]. Furthermore, it requires an additional injection and test, which increases the contrast dose and the diagnosis time. The operator selects a pre-set constant speed according to CT scanning protocol, which comes from experience. Clearly, it is impossible for the constant-speed method to follow the bolus peak as it traverses in the arteries. One method used to increase the CNR in the medical image is to modify the bolus injection as to make the resulting bolus profile more conducive to the pre-set constant-speed tracking method. This method is called bolus geometry optimization [1], [3], [4], [13], [19] and requires a priori knowledge of the patient’s vascular system to achieve the desirable results, which cannot be known before a preliminary injection. Therefore, the geometry optimization method potentially increases the overall contrast dose.

In this paper, we propose an adaptive-optimal tracking approach based on real clinical bolus data. Several steps are involved in the method. First, a mathematical model is developed. The purpose of the model is not to design a specific control law but to provide some insights that could guide the design. Then, a sequence of local approximations are estimated on line and the optimal control laws are developed based on these local approximations. The idea of bolus chasing is not new and much research has been done in applications of Magnetic Resonance Angiography (MRA) [20], [17] and Digital Subtraction Angiography (DSA) [25]. However, little work has been done in the field of CTA due to the following difficulties:

• It is not clear how to design a controller based on available contrast bolus characteristics. Although some works have reported on the bolus model, most of them have concentrated on the time-density (temporal) curve and the spatial dependence was seldom discussed. For control purposes, time, distance, and density bolus characteristics...
are all required.

- CT has a small Field of View (FOV) and short scanning time (in the range of half a minute). This means that only local information is available for the tracking scheme. Consequently, the ability to track the peak is severely compromised without a well designed control algorithm.

The work reported in this paper is a continuation of our preliminary studies [5], [6], [10], [11], [21]. By exploring the properties of the bolus dynamics, an adaptive-optimal controller is developed and validated. The controller possesses the following preferable features.

- Robustness. The controller is not completely model based, although it needs some basic information about the bolus model. It works for a variety of patients, in which bolus dynamics are not the same.
- Optimality. The controller is optimal in terms of the maximizing the contrast to background noise ratio at least in the ideal scenario.
- Simplicity. The controller does not take a large amount of on-line computation time. Thus, it is fairly simple and can be easily installed on a CT machine with few parameters to be adjusted.
- Physical constraints. Physical constraints, e.g., patient’s comfort, mechanical constraints, reduction of dose, diagnosis time etc. can be easily incorporated into the controller.

To the best of our knowledge, this is the first time an adaptive optimal controller is developed for the control of CTA. The paper is organized as follows. Section 2 is devoted to the modeling of the bolus and some critical properties of the bolus dynamics are observed. Control designs are provided in Section 3 along with performance analysis. Detailed discussions on real time estimation is given in Section 4. Tests of the proposed controller on real patient data are carried out in Section 5. Some possible improvements are discussed in Section 6 and Section 7 concludes the paper.

II. MODEL OF DENSITY FUNCTION

Bolus dynamics are extremely complicated and are influenced by patient characteristics, injection methods and injection sites etc [2], [9]. Many researchers have tried different methods to describe bolus dynamics. In [1], [24], a physiologic model of contrast enhancement is developed, which incorporates injection methods, physiologic data and contrast pharmakinetics. The model is too complicated for practical use and is sensitive to many unknown parameters which are impossible to estimate on line. Several reports regarded the bolus dynamics as a black box, and assumed it is a linear system. By examining the input (injection) and output (enhancement) of this black box [13], the bolus dynamics are represented by a linear system. Both of the above two methods can be used to predict bolus density at a time and position. However, they require an extensive amount of previous data to construct the model or to estimate the system parameters, which will take extra time before performing CTA. Additionally, these models have been shown to be inaccurate at times. The other, more efficient method is to use a certain function to fit the bolus time-density curves at certain regions of interest. In [8], a gamma variate function is used to fit the bolus time-attenuation curve in the aorta, which shows the ability to capture the bolus dynamics. In [7], a lagged normal density model is proposed as

\[
B(t, z) = C(z) \frac{1}{\sigma(z)\sqrt{2\pi}} e^{-\frac{(t-t_c(z))^2}{2\sigma(z)^2}} \ast e^{-\frac{1}{\tau(z)}} f(z)
\]  

(1)

where * denotes the convolution operation. This is actually a linear system of infinite order. The coefficient \(C(z)\) relies on the injection (input) pattern. At a fixed location \(z\), \(C, \sigma, t_c\) and \(\tau\) are constants.

To further validate this bolus model and facilitate our controller design, we have collected more than thirty clinical MRA timing bolus data sets, which are collected from: 1) University of Iowa Hospital and Clinics (UIHC), and 2) Northwestern University (NU). UIHC studies are performed on a 1.5 Tesla GE CV/i scanner using a 2D gradient echo sequence (TR = 5.7 ms, TE = 1.5 ms, 40 deg flip angle), and a 3 cc contrast bolus (1.5 mmole gadodiamide) is injected into a peripheral vein followed by a 25 cc saline flush, both at a rate of 3 cc/sec. NU studies are performed on a 3 Tesla SIEMENS MR scanner using a 2D gradient echo sequence (TR = 4.2 ms, TE = 1.11 ms, 8 deg flip angle, and a 2 cc contrast bolus is injected into a peripheral vein followed by a 10 cc saline flush, both at a rate of 2 cc/sec. All MR images are acquired in a sagittal or oblique sagittal plane through a long segment of the thoracoabdominal aorta. A linear relationship between Signal Intensity (SI) and contrast density is assumed for all MRA data. Throughout the paper, when we talk about the bolus density of MRA data, we mean the signal intensity. MRA data is used because it has a larger field of view so that we can observe the bolus density in a long artery at several time instants. Figure 1 and Figure 2 show the real data and the fitted data by the model (1).

In the figures, (a) is in the time-distance density profile, (b) is the contour of the bolus with respect to time and distance, where each curve represents the same bolus density level, and the inner curves have a higher density level, (c) is bolus time-density curve at certain position, and (d) is the bolus distance-density curve at certain time.

It is very interesting to observe that both the data and the model reveal intrinsic properties of the contrast bolus dynamics. As far as the control design is concerned, the most important property is:

At any given position \(z\), the bolus density \(B(t, z)\) is monotonically (not necessarily strictly) increasing in time \(t\) before it reaches the maximum and is monotonically (not necessarily strictly) decreasing after that. This includes asymmetric curves with a more or less flat slopes in the vicinity of the maximum. In other words, the maximum is not necessarily unique but all the maximum points are connected.

Since this property is frequently used in the paper, we define such a function as a class \(\mathcal{A}\) function.

Definition II.1: A differentiable function \(f(t)\) is said to belong to the class \(\mathcal{A}\) if

1) There exists a maximum point \(t^*\) such that \(t^* = \arg \max_t f(t)\),
2) \( f(t) \) is monotonically (not necessarily strictly) increasing for \( t \in (-\infty, t^*) \) and monotonically (not necessarily strictly) decreasing for \( t \in [t^*, \infty) \).

### III. CONTROL DESIGNS

#### A. Control with no constraints

The goal of bolus chasing CTA is to achieve the optimal image quality. Specifically, the control objective is to synchronize the bolus peak with the X-ray source by translating the patient table, that is, to maximize contrast to background noise ratio at every position \( z \). Here, we clarify that “bolus tracking” is different from tracking a solid object with respect to time. Contrast bolus is a liquid that dilutes in the artery as time advances. At each location \( z \), there is a time \( t(z) \) that the bolus density at \( z \) achieves the maximum and an image is preferably be taken at that time. Note this is a two dimensional problem. Now, let the bolus function be given by \( B(t, z) \), where \( t \) represents time \( (t = 0 \) corresponds to the injection instant) and \( z \) represents distance \( (z = 0 \) corresponds to the initial monitored position). Then the control objective is to find a function of \( t(z) \) in terms of \( z \) which maximizes the following quantity.

\[
t^*(z) = \arg \max_{t(z)} J_c = \arg \max_{t(z)} \int_0^{z_c} B(t(z), z) \, dz \tag{2}
\]

where \( z_c \) denotes the scan length. The motivation behind (2) is to maximize the average density over the scan length.

Note for implementation purposes, it is desired to obtain \( z(t) \) as a function of \( t \) not \( t(z) \) as a function of \( z \). However, if we constrain the derivative of \( t(z) \) to be greater than zero, \( z(t) \) can be uniquely calculated. The control scheme can be illustrated in Figure 3.

In the case that \( B(t, z) \) is known, the problem of maximizing \( J_c \) is a variational problem and can be easily solved.

**Theorem 3.1:** The solution to the unconstrained optimization problem (2) is given by

\[
t^*(z) = \arg \max_{t(z)} B(t(z), z)
\]

except possibly on a set of measure zero.

**Proof:** See APPENDIX A.

It is interesting to observe that the solution can also be derived from the well known Euler equation [12] of (2),

\[
\frac{\partial B}{\partial t} - \frac{d}{dz} \frac{\partial B}{\partial t} = 0 \iff \frac{\partial B}{\partial t} = 0.
\]

The control goal and its solution in continuous time domain are derived and provide much insight. Because of the inherent
discrete-time nature of CT, however, a discrete time formulation is more appropriate. To this end, we segment the scan range into \( N \) sections. Let \( z_k = k \bar{z} \), \( k = 0 \ldots N \), where \( z_e \) is the scan distance. Now our discrete objective reduces to finding the sequence, \( \{ t^*(z_k) \} \) for \( k = 1, \ldots, N \), such that
\[
\{ t^*(z_k) \}^N_{k=1} = \arg \max \left\{ \sum_{i=1}^{N} B(t(z_i), z_i) \right\}.
\]

Clearly, if the bolus profile \( B(t, z) \) is known, the solution is given by
\[
t^*(z_k) = \arg \max_{t(z_k)} B(t(z_k), z_k), \quad k = 1, \ldots, N.
\]

As discussed before, however, the exact profile \( B(t, z) \) is not known a priori or during the scanning. Determination of the optimal \( t^*(z_k) \)'s has to rely on a sequence of local approximations \( \{ \hat{B}_k(t, z_{k+1}) \}^{N-1}_{k=0} \) and
\[
\tilde{t}(z_k) = \arg \max_{t(z_k)} \hat{B}_{k-1}(t(z_k), z_k), \quad k = 1, \ldots, N
\]

There are two questions. The first is how to find such approximations \( \{ \hat{B}_k(t, z_{k+1}) \}^{N-1}_{k=0} \) which is the topic of next section. The second question is how small the error \( t^*(z_k) - \tilde{t}(z_k) \) is and more importantly, how close \( B(t^*(z_k), z_k) \) and \( B(\tilde{t}(z_k), z_k) \) is that we are really after. Intuitively, if \( \hat{B}_{k-1}(t, z_k) \) approximates \( B(t, z_k) \) well in the vicinity of \( t^*(z_k) \), we expect a small error. The following lemma verifies this intuition.

**Lemma 3.1:** Let \( f(t) \) be a class \( A \) function with \( t^* = \max_t f(t) \). Given \( \delta > 0 \), denote \( t_1 < t^* \) as the maximum value such that \( f(t_1) = f(t^*) - 2\delta \) and denote \( t_2 > t^* \) as the minimum value such that \( f(t_2) = f(t^*) - 2\delta \), see Figure 4. Define \( d = \max \{ |t_1 - t^*|, |t_2 - t^*| \} \). Let \( f(t) \) be a class \( A \) function with \( \bar{t} = \arg \max_t f(t) \) and not-constant for \( t \in [t_1, t_2] \). Assume \( |f(t) - f(t)| \leq \delta \) for \( t \in [t^* - d, t^* + d] \). Then, we have
\[
|\bar{t} - t^*| \leq d \quad \text{and} \quad |f(\bar{t}) - f(t^*)| \leq \delta.
\]

**Proof:** See APPENDIX B.

![Fig. 4. Diagram for lemma 3.1.](image)

The interpretation of Lemma 3.1 is that given an approximation error \( \delta \) at each iteration \( |\hat{B}_{k-1}(t, z_k) - B(t, z_k)| < \delta \) for \( t \in [t^*(z_k) - d, t^*(z_k) + d] \), then
\[
|\tilde{t}(z_k) - t^*(z_k)| < d \quad \text{and} \quad |B(\tilde{t}(z_k), z_k) - B(t^*(z_k), z_k)| \leq 2\delta.
\]

**B. Control with constraints**

In equations (4) and (5), it is apparent that \( \tilde{t}(z_k) \) and \( t^*(z_k) \) need not be strictly increasing sequences which is problematic for real time implementation. Also, the table’s speed is bounded by mechanical factors and patient comfort, i.e., the table speed should be limited by lower and upper bounds
\[
0 < \frac{z_e/N}{\Delta_b} \leq \frac{dz}{dt} \leq \frac{z_e/N}{\Delta_s} < \infty
\]

for some \( \Delta_b \) and \( \Delta_s > 0 \), where \( z_e/N \) is the length of each section. Or equivalently,
\[
0 < \frac{\Delta_s}{z_e/N} \leq \frac{dt}{dz} \leq \frac{\Delta_b}{z_e/N} < \infty.
\]

This translates into the constraint
\[
0 < \frac{\Delta_s}{z_e/N} \leq \frac{\Delta_b}{z_e/N} < \infty, \quad k = 0, \ldots, N-1
\]

Therefore, the control laws have to be modified considering the above constraint. If \( B(t, z) \) is known, the solution is now given by
\[
t^*_k = \begin{cases} 
\tilde{t}^*_{k-1} + \Delta_s, & \text{if } t^*(z_k) \leq \tilde{t}^*_{k-1} + \Delta_s \\
\tilde{t}^*_k + \Delta_b, & \text{if } t^*(z_k) \geq \tilde{t}^*_{k-1} + \Delta_b \\
\tilde{t}^*_k, & \text{otherwise}
\end{cases}
\]

where \( t^*(z_k) = \max_{t(z_k)} B(t(z_k), z_k) \), \( k = 1, \ldots, N \).

Similarly, if \( B(t, z) \) is unavailable, the control based on its estimates \( \hat{B}_{k-1}(t, z_k) \)'s is given by
\[
\tilde{t}_k = \begin{cases} 
\tilde{t}_k + \Delta_s, & \text{if } \tilde{t}(z_k) \leq \tilde{t}_{k-1} + \Delta_s \\
\tilde{t}_k, & \text{if } \tilde{t}(z_k) \geq \tilde{t}_{k-1} + \Delta_b \\
\tilde{t}_k, & \text{otherwise}
\end{cases}
\]

where \( \tilde{t}(z_k) = \arg \max_t \hat{B}_{k-1}(t, z_k), \quad k = 1, \ldots, N \).

We now quantify the errors between \( t^*_k - \tilde{t}_k \) and \( B(t^*_k, z_k) - B(\tilde{t}_k, z_k) \). The following lemma is useful in quantifying these errors.

**Lemma 3.2:** Let \( a_1, a_2, b_1, b_2 > 0 \). Then, \( |a_1 - a_2| \leq d \) and \( |b_1 - b_2| \leq d \) imply
\[
1) \quad |\max(a_1, b_1) - \max(a_2, b_2)| \leq d,
2) \quad |\min(a_1, b_1) - \min(a_2, b_2)| \leq d.
\]

**Proof:** See APPENDIX C.

We are now in a position to quantify the error in the presence of control constraints.

**Theorem 3.2:** Let \( B(t, z_k) \) and \( \hat{B}_{k-1}(t, z_k) \) belong to class \( A \) and \( \left| \frac{\Delta B(t^*_k, z_k)}{\Delta t} \right| \leq \epsilon, \forall t, k \). Suppose there exists a positive \( \delta > 0 \), and
\[
|B(t, z_k) - \hat{B}_{k-1}(t, z_k)| \leq \delta, \quad t \in [t^*(z_k) - d_k, t^*(z_k) + d_k],
\]

where \( d_k = \max \{ |t^*(z_k) - t_1(z_k)|, |t^*(z_k) - t_2(z_k)| \} \) and \( t_1(z_k) < t^*(z_k) \), \( t_2(z_k) > t^*(z_k) \) is the maximum (minimum) value such that \( B(t_1(z_k), z_k) \leq B(t^*(z_k), z_k) - 2\delta \) and \( B(t_2(z_k), z_k) \leq B(t^*(z_k), z_k) - 2\delta \). Let \( d = \max_k \{ d_k \} \) and assume that \( \hat{B}_{k-1}(t, z_k) \) is not constant in \( t \in [t^*_k, t^*_k + \Delta_b] \). If \( t^*_k \) and \( \tilde{t}_k \) denote the sequences resulting from (7) and (8), respectively, then we have
\[
|t^*_k - \tilde{t}_k| \leq d \quad \text{and} \quad |B(t^*_k, z_k) - B(\tilde{t}_k, z_k)| \leq d \cdot \epsilon
\]
for $k = 1, \ldots, N$, if $|t_k^5 - \bar{t}_k| \leq \delta$.

Proof: See APPENDIX D.

IV. ON LINE ESTIMATION OF $B(t, z)$

Since $B(t, z)$ is unavailable, calculation of the control law relies on the estimates $\hat{B}_{k-1}(t, z_k)$. As shown in the previous section, if the estimate $\hat{B}_{k-1}(t, z_k)$ is close to the true but unknown $B(t, z)$ locally, the effect of approximation is negligible. However, $\hat{B}_{k-1}(t, z_k)$ has to be estimated in real time solely based on observed local bolus information. On one hand, the approximation error depends on how rich the structure of the approximation $\hat{B}_{k-1}(t, z_k)$ is that incorporates well the local bolus information into its representation. On the other hand, the structure should be simple enough so that it can be easily estimated on line. There is a trade off between approximation ability and estimation accuracy. A natural choice of $\hat{B}_{k-1}(t, z_k)$ is a polynomial. The order of the polynomial balances the ability to approximate $B(t, z)$ and the estimation simplicity. We consider a second order polynomial,

$$B(t, z) \approx B(t_k, z_k) + \nabla_t B|_{(t_k, z_k)} (t - t_k) + \nabla_z B|_{(t_k, z_k)} (z - z_k) + \frac{1}{2} \left( \frac{t - t_k}{z - z_k} \right)^T \nabla^2 B|_{(t_k, z_k)} \left( \frac{t - t_k}{z - z_k} \right)$$

$$= a_0 + a_1 t + a_2 z + a_3 t^2 + a_4 z t + a_5 z^2,$$

for some $a_0, a_1, a_2, a_3, a_4$ and $a_5$, which will be estimated on line. Another important factor is the selection of approximation data. Obviously, the smaller the region, the better a second order polynomial can approximate $B(t, z)$. On the other hand, we would like the approximation function to give us some information about the bolus away from the observation points. This implies that the approximation region cannot be too small. In addition, realistic CT specifications have to be considered.

- CT is assumed to have multiple rows of detectors. This is very reasonable and most modern CT have more than four rows of detectors. Thus, at a given time $t_k$, we observe the density information at the current position $z_k$ as well as at $z_k \pm \delta z$, where $\delta z = 2.5 \text{ mm}$ [13].
- CT gantry rotation speed is set to $\Delta T = 1/3$ second per rotation, a standard in modern CT.
- CT image reconstruction time is assumed to be very small considering that the real CT scanner has a very fast computer and a low resolution is needed during the tracking procedure. The bolus density information is taken as the average HU value in the ROI of main blood vessel in the reconstructed CT slice.
- The maximum patient table speed in a modern CT is about 10 cm/sec [15], which sets the lower bound $\Delta_s$ in the constraint.
- The minimum speed is set to 0 cm/sec which ensures that the patient table does not go back, a standard practice. This sets the upper bound $\Delta_b$.

With these constraints, data points of $B(t, z)$ at $\bar{t}_k, t_{k-1}, t_{k-2}$ and $z = z_k, z_k \pm \delta z$ are collected for each rotation of the gantry. These data are used to identify the $a_i$’s in the second order approximation in the least squares sense

$$\hat{a}_i = \arg \min_{a_i} \sum_{z = z_k, z_k \pm \delta z, \ t=\bar{t}_k, t_{k-1}, t_{k-2}} \{ B(t, z) - (a_0 + a_1 t + a_2 z + a_3 t^2 + a_4 z t + a_5 z^2) \}^2.$$ (10)

Further, the next $\bar{t}_{k+1}$ is determined from (8) with

$$\bar{t}_{k+1} = \hat{a}_0 + \hat{a}_1 t + \hat{a}_2 z + \hat{a}_3 t^2 + \hat{a}_4 z t + \hat{a}_5 z^2.$$ (11)

Figure 5 shows a $B(t, z)$ (dashed) as a function of $t$ at 4 different $z_k$’s and 4 corresponding second order polynomial approximations (solid) derived from on line estimation as discussed above. The maximum error is less than $2.5 \times 10^{-3}$ mm. Obviously, a simple second order approximation is powerful enough to approximate the unknown $B(t, z)$ locally. We emphasize the word “locally”. If $B(t, z)$ is needed to be approximated in a large area, a very high order of polynomial is required that would make on line estimation infeasible.

![Fig. 5. $B(t, z_k)$ (dashed) as a function of $t$ for fixed $z_k$ and the second order approximations $B(t, z_k)$ (solid) in the vicinity of $t^*(z_k)$.

V. EXPERIMENTAL SIMULATION

We have tested the proposed adaptive-optimal control scheme on the actual clinical data (see Section 2). In the tests, the upper and lower bounds $\Delta_b, \Delta_s$ are set so the minimum and maximum table speeds correspond to 0 cm/sec and 10 cm/sec, respectively. $z_c/N$ is set to be 5 mm. To show its superiority, the proposed adaptive-optimal control is compared to the performance of existing technology, a constant speed control of the patient table, i.e., $t_k = cz_k + t_0$ for some constant $c$. A common practice in clinic is to set $c$ so that the table speed is at 3 cm/sec [14].

Two comparisons are made. The first is the achieved bolus density. To this end, we define

$$I_a = \frac{\sum B(t^*_k, z_k)}{\sum B(t^*(z_k), z_k)}, \quad I_c = \frac{\sum B(cz_k + t_0, z_k)}{\sum B(t^*(z_k), z_k)}$$

where $c$ is chosen so the constant speed control is at 3 cm/sec. $\sum B(t^*(z_k), z_k)$ is the maximum achievable bolus density. $\sum B(t^*_k, z_k)$ and $\sum B(cz_k + t_0, z_k)$ are the actually
achieved bolus densities for the adaptive-optimal control and the constant speed control respectively. Typical results are shown in Figure 6 and 7, where the top plots show the bolus contour (solid thin curve), the bolus peak trajectory (dashed), adaptive-optimal trajectory (thick solid) and constant-speed trajectory (dash dot); and the bottom plots show the maximum achievable bolus density at each position $z_k$ (dashed), the achieved bolus density by the adaptive-optimal control (solid) and the achieved bolus density by the constant speed control (dash dot). Recall that the table movement is initially triggered by a pre-set threshold. Table 1 summarizes the performance index for ten patients, where the threshold for UIHC patient and NU patient are set to 30 and 5, respectively. Clearly, in all cases the proposed adaptive-optimal control outperforms the current constant speed control significantly.

### Table 1

<table>
<thead>
<tr>
<th>Patient</th>
<th>$I_a$</th>
<th>$I_c$</th>
<th>Patient</th>
<th>$I_a$</th>
<th>$I_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>UIHC 1</td>
<td>0.98</td>
<td>0.90</td>
<td>NU 1</td>
<td>0.99</td>
<td>0.93</td>
</tr>
<tr>
<td>patient 2</td>
<td>0.95</td>
<td>0.78</td>
<td>patient 2</td>
<td>0.98</td>
<td>0.85</td>
</tr>
<tr>
<td>patient 3</td>
<td>0.93</td>
<td>0.74</td>
<td>patient 3</td>
<td>0.97</td>
<td>0.81</td>
</tr>
<tr>
<td>patient 4</td>
<td>0.95</td>
<td>0.83</td>
<td>patient 4</td>
<td>0.99</td>
<td>0.90</td>
</tr>
<tr>
<td>patient 5</td>
<td>0.93</td>
<td>0.84</td>
<td>patient 5</td>
<td>0.98</td>
<td>0.58</td>
</tr>
</tbody>
</table>

The second performance to be compared is the robustness with respect to the threshold variation. When the observed density reaches the threshold, the CT table is automatically started [9]. If the threshold is set too low, the CT table is started too early and if the threshold is set too high, the CT table is started too late. In both cases, the performance deteriorates. Figures 8-10 show the results of NU Patient 2 using thresholds 4, 7 and 10, respectively. Figure 11 summarizes the robustness with respect to threshold variations. Again, the proposed method outperforms the constant speed control. We also make an important observation here. Figure 11 seems to suggest that a large threshold is more appropriate for the constant speed control. Unfortunately, this is not always true and is patient dependent. Figure 12 shows the threshold variation result for Patient 5. The performance deteriorates if the threshold is set too high. This poses a difficult problem for the constant speed control. On the other hand, the adaptive-optimal control performs uniformly well independent of various threshold settings.

### VI. DISCUSSION

In this paper, we proposed an adaptive-optimal control that substantially outperforms the existing technology. There are
a number of modifications that could be incorporated into the control scheme either to improve the performance or to simplify implementation.

- Optimality of the control with constraints: Without constraints (6), the obtained solution $t^*(z_k)$ is optimal in terms of the objective function (3). With the constraints (6), the solution $t_k$ of (8) may not be optimal. In theory, an optimal solution could be developed if all the future values of $B(t, z)$ are known. This is not possible in reality. $B(t, z)$ varies from one patient to another and has to be estimated on line. To estimate all future values of $B(t, z)$, a long term estimation has to be made. Combining the fact that $B(t, z)$ is so complicated, a long term estimation would make prediction so unreliable and result in a compromised control scheme.

The solution of (8) derived in the paper performs well and is the same or very close to the true optimal solution had $B(t, z)$ been known a priori. This is due to two properties of the solution (8). The first one is the one-step ahead optimality. Though it may not be globally optimal, the solution is one-step ahead optimal. Specifically, the solution (8) finds a time $t_{k+1}$ that provides the maximum $B(t, z_{k+1})$ at $z_{k+1}$, given the constraints (6) and the previous $t_k, t_{k-1}, ..., t_1$. Secondly, $t^*(z_k)$ is reachable from $t_{k-1}$ independent of $t_{k-1}, ..., t_1$. In other words, the solution (8) has a self-recovery ability to find the globally optimal $t^*(z_k)$ even it misses the previous one $t^*(z_{k-1})$. One possible improvement over the one step-ahead optimal controller is to estimate $B(t, z)$ several steps ahead and then to determine the optimal times several steps ahead. This conceptually would improve the one-step ahead controller at a price of increased computational complexity. Clearly, too many steps ahead is not a good idea because it brings too much estimation errors into the control. How to balance these two conflict objectives is an interesting topic.

- Alternative objective functions: In the paper, the objective function is defined as

$$\max_t \sum B(t, z_k)$$

The meaning is to achieve the maximum bolus density at each location $z_k$. An alternative objective function can be used

$$\max_z \sum B(t_k, z)$$

and the optimal control could be developed accordingly. The meaning is to maximum average density with respect to time. In general, these two objective functions are different, e.g., the bolus peak stays in a region and does not traverse the entire region, and thus, the performances of the controls would be different. In many cases, however, these two objective functions produce more or less the same control laws. One advantage of the optimal control with respect to the later objection function lies in the solution form $z(t)$ instead of $t(z)$ which makes implementation a little easier.

- On line estimation: Obviously, the performance of the adaptive-optimal control derived from the discrete version of the objection function (3) depends on the number $N$ of the partitions. Theoretically, the larger $N$, the better the performance. Practically, there is an ill-conditioning problem in the estimation if $N$ is too large. Note that a large $N$ implies a small section $z_k - z_{k-1} = z_0/N$ which translates into a small time interval $t_k - t_{k-1}$. This small time interval makes the least squares estimation problem of (10) ill-conditioned. To avoid this problem, a threshold $\Delta T$ could be set. If $t_k - t_{k-1} < \Delta T$, no new estimation is carried out since $t_k \approx t_{k-1}$ and the estimate

![Fig. 10. Tracking results for NU Patient 2, threshold is set to 10.](image)

![Fig. 11. Performance on NU Patient 2 of adaptive-optimal method (solid diamond) and constant-speed method with different threshold (dashed cross).](image)
obtained for \( \hat{t}_{k-1} \) should be also valid at \( \hat{t}_k \). Only when \( \hat{t}_{k+1} - \hat{t}_{k-1} \geq \Delta T \), for some \( p \geq 0 \), a new estimate is called for. This avoids the ill-conditioning problem.

- Pitch consideration: Bolus chasing CTA requires an adaptive change of the patient table velocity during a scan. This will generally result in a varying pitch helical scan, which may in turn lead to a variation in image noise and radiation dose to the patient. Nevertheless, this type of non-uniformity can be effectively managed by selection of an optimal source rotation speed, modulation of the tube current as a function of the helical pitch and adaptation of the multi-slice/cone-beam reconstruction. As reported in [22], the patient dose is independent of pitch if the tube current is appropriately modulated.

VII. Concluding Remarks and Future Work

Atherosclerosis of the aorta, iliac and lower extremity arteries is common, and may have hemodynamically significant effects on vessel lumen diameter and blood flow. Up to 2% of patients over age 45 have symptoms related to occlusive aorto-iliac or lower extremity atherosclerotic disease including pain, tissue loss and tissue infarction. Autopsy series have shown that over 2% of the population has aorto-iliac or lower extremity aneurysms. Aortic aneurysm rupture may lead to sudden death from exsanguination. CT angiography (CTA) has become a popular alternative due to its acquisition speed, volumetric nature, and non-invasiveness. Bolus dynamics is complex and influenced by contrast administration protocol and patient characteristics (age, sex, weight, height, cardiovascular status, renal function, etc.). The intravascular contrast bolus travels fast in the torso and slowly in the legs. There can be substantial differences in flow velocity between the legs when asymmetric peripheral vascular disease exists. Scanning too early may result in over-estimation of stenosis, while scanning too late may result in overlap of venous structures. With a pre-set scanning speed, it is difficult and often impossible to synchronize the imaging aperture with the moving bolus peak. Misalignment may be even more problematic when scanning speed is fast, contrast volume is small and/or injection rate is high (leading to reduced peak duration) or there are large or small capacity vessels, either from aneurysm formation or occlusive diseases. Therefore, we underline that our work represents a major refinement relative to the state of the art of CT angiography and may have a major and lasting impact on patient care.

In the paper, we have shown that the proposed adaptive-optimal algorithm has a great potential in tracking the bolus with the highest density. The proposed algorithm is built on a solid theoretical foundation and is simple enough for easy implementation. We believe that control concepts in biomedical applications opens a new area that will have an unprecedented impact on medicine. The future work is two-stage. In the first stage, we will test the proposed control algorithm on the vascular phantom using the real CT scanner, which is being carried out. The second stage will be focused on clinical trials.

Appendix A

Proof of Theorem 3.1

Proof: This is proved by contradiction. Suppose that \( t(z) \) is the solution to (2) and \( t(z) \neq t^*(z) \) on \( S \), where \( S \) is a set which is not measure zero. This implies

\[
\int_0^\infty B(t^*(z), z)dz - \int_0^\infty B(t(z), z)dz = \int_S (B(t^*(z), z) - B(t(z), z))dz > 0
\]

by the definition of \( t^*(z) \). This contradicts the assumption that \( t(z) \) is the solution. Hence, \( t^*(z) \) is the solution to (2).

Appendix B

Proof of Lemma 3.1

Proof: First, we claim \( \hat{t} \in [t_1, t_2] \). If not, say \( \hat{t} < t_1 \), then by the decreasing and non-constant properties of \( f(t) \), \( t \in [t_1, t_2] \),

\[
|\hat{f}(t_1) - f(t_1)| \leq \delta \quad \text{and} \quad |\hat{f}(t_2) - f(t_2)| \leq \delta
\]

can not be simultaneously satisfied. Similar argument shows that \( \hat{t} > t_2 \) is impossible. Thus, \( \hat{t} \in (t_1, t_2) \). Now,

\[
f(t^*) + \delta \geq f(\hat{t}) + \delta \geq \hat{f}(\hat{t}) \geq f(t^*) - \delta.
\]

This completes the proof.

Appendix C

Proof of Lemma 3.2

Proof: For Part 1, there are four cases.

1) \( \max(a_1, b_1) = a_1 \) and \( \max(a_2, b_2) = a_2 \), \( \max(a_1, b_1) - \max(a_2, b_2) = |a_1 - a_2| \leq d \).

2) \( \max(a_1, b_1) = b_1 \) and \( \max(a_2, b_2) = b_2 \), \( \max(a_1, b_1) - \max(a_2, b_2) = |b_1 - b_2| \leq d \).

3) \( \max(a_1, b_1) = a_1 \) and \( \max(a_2, b_2) = b_2 \), if \( a_1 > b_2, |a_1 - b_2| < |a_1 - a_2| \leq d \), since \( a_1 > b_2 > a_2 \), else if \( a_1 < b_2, |a_1 - b_2| < |b_1 - b_2| \leq d \), since \( b_1 < a_1 < a_2 \).

4) \( \max(a_1, b_1) = b_1 \) and \( \max(a_2, b_2) = a_2 \), if \( b_1 > a_2, |b_1 - a_2| < |b_1 - b_2| \leq d \), since \( b_1 > b_2 > a_2 \), else if \( b_1 < a_2, |b_1 - a_2| < |a_1 - a_2| \leq d \), since \( a_1 < a_2 \).

By the same token, there are again four cases for Part 2.

1) \( \min(a_1, b_1) = a_1 \) and \( \min(a_2, b_2) = a_2 \), \( \min(a_1, b_1) - \min(a_2, b_2) = |a_1 - a_2| \leq d \).

2) \( \min(a_1, b_1) = b_1 \) and \( \min(a_2, b_2) = b_2 \), \( \min(a_1, b_1) - \min(a_2, b_2) = |b_1 - b_2| \leq d \).

3) \( \min(a_1, b_1) = a_1 \) and \( \min(a_2, b_2) = b_2 \), if \( a_1 > b_2, |a_1 - b_2| < |b_1 - b_2| \leq d \), since \( b_1 > a_1 > b_2 \), else if \( a_1 < b_2, |a_1 - b_2| < |a_1 - a_2| \leq d \), since \( a_1 < b_2 < a_2 \).

4) \( \min(a_1, b_1) = b_1 \) and \( \min(a_2, b_2) = a_2 \), if \( b_1 > a_2, |b_1 - a_2| < |a_1 - a_2| \leq d \), since \( a_1 > b_1 > a_2 \), else if \( b_1 < a_2, |b_1 - a_2| < |b_1 - b_2| \leq d \), since \( b_2 > a_2 > b_1 \).

This completes the proof.
APPENDIX D

PROOF OF THEOREM 3.2

Proof: The proof is by induction. Since for \( k = 0 \), it is true. We assume that \(|t_k^* - \tilde{t}_k| \leq d\) and want to show \(|t_{k-1}^* - \tilde{t}_{k-1}| \leq d\). From (7) and (8), we write \( t_k^* \) and \( \tilde{t}_k \) in the following compact form,

\[
t_k^* = \min \left\{ t_{k-1}^* + \Delta_b, \max\{t^*(z_k), t_{k-1}^* + \Delta_s\} \right\}. \tag{11}
\]

\[
\tilde{t}_k = \min \left\{ \tilde{t}_{k-1} + \Delta_b, \max\{f(z_k), \tilde{t}_{k-1} + \Delta_s\} \right\}. \tag{12}
\]

Since \(|t_{k-1}^* + \Delta_s - (\tilde{t}_{k-1} + \Delta_s)| = |t_{k-1}^* - \tilde{t}_{k-1}| \leq d\) and \(|t^*(z_k) - f(z_k)| \leq d\), we have by using the first part of Lemma 3.2,

\[
\left| \max\{t^*(z_k), t_{k-1}^* + \Delta_s\} - \max\{f(z_k), \tilde{t}_{k-1} + \Delta_s\} \right| \leq d.
\]

Similarly, by applying the second part of the Lemma 3.2 , \(|t_k^* - \tilde{t}_k| \leq d\). This shows, by induction, for each \( k \),

\[
|t_k^* - \tilde{t}_k| \leq d.
\]

Furthermore, because \( \frac{\partial B(t,z_k)}{\partial t} \leq \epsilon \) for all \( k \) we have

\[
|B(t_k^*, z_k) - B(\tilde{t}_k, z_k)| \leq d \cdot \epsilon.
\]

This completes the proof. ■

REFERENCES


