Research Article

Reliable and Efficient Approach of BOLD Signal with Dual Kalman Filtering

Cong Liu and Zhenghui Hu

State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou 310027, China

Correspondence should be addressed to Zhenghui Hu, zhenghui@zju.edu.cn

By introducing the conflicting effects of dynamic changes in blood flow, volume, and blood oxygenation, Balloon model provides a biomechanical compelling interpretation of the BOLD signal. In order to obtain optimal estimates for both the states and parameters involved in this model, a joint filtering (estimate) method has been widely used. However, it is flawed in several aspects (i) Correlation or interaction between the states and parameters is incorporated despite its nonexistence in biophysical reality. (ii) A joint representation for states and parameters necessarily means the large dimension of state space and will in turn lead to huge numerical cost in implementation. Given this knowledge, a dual filtering approach is proposed and demonstrated in this paper as a highly competent alternative, which can not only provide more reliable estimates, but also in a more efficient way. The two approaches in our discussion will be based on unscented Kalman filter, which has become the algorithm of choice in numerous nonlinear estimation and machine learning applications.

1. Introduction

A thorough understanding of the dynamic relationship between cerebral blood flow (CBF), cerebral blood volume (CBV), and the blood oxygenation level dependent (BOLD) signal is essential for the physiological interpretation of fMRI activation data. The Balloon model described by Buxton et al. (1998) [1] is the first biomechanical plausible model to expound this relationship: increasing the flow (or perfusion rate) generally leads to dilution of venous deoxyhemoglobin (dHb), reducing the tendency of the blood to attenuate the magnetic resonance signal. The resultant increase in signal intensity is referred to as the BOLD response [2]. It is by extending this model to cover the dynamic coupling between CBF and synaptic activity, more sophisticated physiological realities are incorporated, for example, oxygen metabolism dynamics, both intra- and extravascular signal [3, 4], and more intricate models obtained.

The Balloon model is an input-state-output model with three state variables: blood flow, and volume, deoxyhemoglobin content and several biologically reasonable parameters. The problems of state estimation and parameter estimation (sometimes referred to as system identification or

machine learning) associated with Balloon model are often formulated in a state-space representation, where Balloon model serves as a set of continuous-time system equations to describe the hemodynamic process. The equations are nonlinear, corresponding to the fact that Balloon model is one of the numerous nonlinear approaches to characterizing evoked hemodynamic response in fMRI.

Several work has utilized the Balloon model or its enhanced versions in the analysis of fMRI response. Some approaches, including expectation maximization (EM) [5, 6] and maximum likelihood [7], model the BOLD observation as deterministic hemodynamic process. However, a limitation of these methods is that they can only deal with measurement noise. Many promising approaches to the dual estimation problem belong to filtering algorithm that is able to account for both the physiological and measurement noise. Riera et al. [8] addressed the data assimilation problem in an extended Kalman filter (EKF) strategy. As EKF might lead to the problem of divergence due to linearized approximation, Johnston et al. proposed particle filter [9, 10] to avoid the flaw of linearization. Moreover, Hu et al. [11–13] employed unscented Kalman filter (UKF) that also outperforms EKF in terms of estimation error but with

roughly the same computational cost. Most recently, Friston et al. described variational filtering to optimize the approximation of posterior density on hidden model variables, while accumulating sufficient statistics to optimize the conditional densities of parameters and precision [14, 15].

The approaches mentioned above have greatly improved our ability to explore, and above all, to quantify the physiological mechanism involved in neural activation. However, they still have palpable defects. It is noteworthy that many of them actually are in the spirit of joint filtering, in which the underlying states and parameters are concatenated into a single higher dimensional joint state space, a filter runs for estimating both the states and parameters. Despite its straightforwardness in theory and convenience in implementation, the weakness of joint filtering is obvious. The objective of this paper is to introduce and develop an estimator equally concise but with higher performance—dual filtering. We will demonstrate the advantage of dual filter from two aspects, by the example of dual UKF versus joint UKF. (1) In terms of Balloon model, there is no inherent biophysical correlation between the states and parameters. By treating them separately, dual filtering can avoid undesired transaction between them. (2) Larger dimension of state-space vector implies much more computational expense. Specifically, computational complexity for general state-space problems is $\mathcal{O}(L^3)$ [16]. Although the frequency of predict-update cycle required by dual filter is the twice of that required by joint filter, dual estimate is much more computational efficient.

2. Materials and Methods

2.1. Hemodynamic Model. Balloon model describes the coupled kinetic changes from synaptic activity to the fMRI BOLD signal at a given region. This model has been extended by Friston et al. (2000) [5] to include the effects of external inputs to an autoregulated vasodilatory signal, assuming that the relationship between evoked neural activity and blood flow is linear. The subsequent work added different variations to this model, several of them were reviewed and integrated in Stephan et al. (2004) [18] and Buxton (2004) [19]. Based on fundamental physiology, rather than empirical approaches, these enhanced models are able to unify existing literature and provide insight into how the underlying physiological mechanisms result in stable or/and transit BOLD response. However, the original model proposed by Buxton et al. and completed by Friston et al. is sufficient to account for the nonlinear behaviors observed in real-time series [5]. Too many state variables and parameters will not serve our purpose here better.

The dynamic intertwinement between multiple physiological variables, the cerebral blood flow (CBF) f, blood venous volume v, and veins deoxyghemoglobin content q, can be given as a set of nonlinear nondimensional differential equations [1, 20]:

$$\ddot{f} = \epsilon u(t) - \frac{\dot{f}}{\tau_s} - \frac{f-1}{\tau_f},$$

$$\dot{v} = \frac{1}{\tau_0}\left(f - v^{1/\alpha}\right),$$

$$\dot{q} = \frac{1}{\tau_0}\left(f\frac{1-(1-E_0)^{1/f}}{E_0} - v^{1/\alpha}\frac{q}{v}\right), \tag{1}$$

where ϵ is neuronal efficacy, reflecting the significance of neuronal activity evoked by experimental event, hence it varies with trial event; τ_s and τ_f represent time constant for signal decay and autoregulatory feedback from blood flow, respectively. The existence of feedback term can be inferred from the poststimulus undershoots in CBF [21]. The degree of nonlinearity of the BOLD signal is largely determined by the stiffness parameter α, which characterizes the balloon-like capacity of the venous compartment to expel blood at a greater rate when distended [22]. E_0 is resting net oxygen extraction fraction. All variables are expressed in normalized form, relative to resting values.

Noticing that the first equation has a second-order time derivative, so we can write this input-state-output system as a set of first-order ordinary differential equations by introducing another variable $s = \dot{f}$. By defining the state vector as $\mathbf{x}(t) = [f, s, q, v]^T$, the system dynamic equation can be constructed from (1):

$$\dot{x} = f(\mathbf{x}, \theta, \mathbf{u}, \mathbf{v}) \quad v \sim N(0, \mathbf{Q}), \tag{2}$$

where $\theta = \{\epsilon, \tau_s, \tau_f, \tau_0, \alpha, E_0, V_0\} \in \mathbb{R}^l$ is system parameter, the neuronal input $\mathbf{u}$ represents system input, and $\mathbf{v}$ is to account for the process noise.

The observed signal can be taken as a nonlinear function of volume v and deoxyghemoglobin q that comprises a volume-weighted sum of intra- and extravascular signal:

$$y(t) = V_0\left(k_1(1-q) + k_2\left(1 - \frac{q}{v}\right) + k_3(1-v)\right),$$

$$k_1 = 7E_0, \qquad k_2 = 2, \qquad k_3 = 2E_0 - 0.2, \tag{3}$$

appropriate for a 1.5 tesla magnet [1]. V_0 is the resting blood volume fraction, which generally varies across brain regions and subjects. All parameters are independent of each other. Their physiological definitions and probability distributions are given in Table 1 [23].

The actual observation $\mathbf{y}$ is then composed of a deterministic part $h(x, \theta, t)$ and a stochastic part $\mathbf{w}$:

$$\mathbf{y} = h(\mathbf{x}, \beta, \mathbf{w}) \quad w \sim N(0, \mathbf{R}), \tag{4}$$

where $\mathbf{y}$ is the observation vector, $\mathbf{w}$ is measurement noise, and β consists of k_1, k_2, and k_3. Simultaneous estimation of V_0 and other parameters would be impossible, since their product is settled for each sampled measurement y_k. The stiffness parameter α is a nominal factor to BOLD contrast, it can be fixed to any value with its reasonable range in system identification [17].

Equation (2) describes a continuous-time hemodynamic process, and (4) models fMRI measurement as discrete sampling of the continuous system states, together they have formed a standard state-space representation for fMRI data assimilation. Given y_k, the physiological states $\mathbf{x}$ and the

TABLE 1: Hemodynamic model parameters and their probability distribution.

Notation	Definition	Distribution
ϵ	Neuronal efficacy	$\epsilon \sim N(0.54, 0.1^2)$
τ_s	Signal decay	$\tau_s \sim N(1.54, 0.25^2)$
τ_f	Autoregulation	$\tau_f \sim N(2.46, 0.25^2)$
τ_0	Transit time	$\tau_0 \sim N(0.98, 0.25^2)$
α	Stiffness parameter	$\alpha \sim N(0.33, 0.045^2)$
E_0	Resting oxygen extraction	$E_0 \sim N(0.34, 0.1^2)$
V_0	Resting blood volume fraction	$V_0 \sim N(0.02, 0.005^2)$

optimal parameters for a certain voxel can be estimated by use of UKF—a recursive minimum mean-square-error (MMSE) estimator.

2.2. Dual UKF and Joint UKF. The unscented Kalman filter has been applied extensively to the field of nonlinear estimation for both states and parameters. The basic framework of UKF involves estimation of the states of a discrete-time nonlinear dynamic system:

$$\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{v}_k),$$
$$\mathbf{y}_k = \mathbf{H}(\mathbf{x}_k, \mathbf{n}_k), \tag{5}$$

where $\mathbf{x}_k$ represents the unknown system states, the system is driven by a known exogenous input $\mathbf{u}_k$ and process noise $\mathbf{v}_k$. The observation noise is given by $\mathbf{n}_k$.

The UKF generally involves recursive utilization of a deterministic "sampling" approach. The sampled points (sigma points) completely capture the true mean and covariance of the variables, and when propagated through the nonlinear system ($\mathbf{F}$ in this case), they are able to capture the posterior mean and covariance accurately to the 2nd order of Taylor series expansion [16].

Parameter estimation, or machine learning, on the other hand, involves determining a nonlinear mapping:

$$\mathbf{y}_k = \mathbf{G}(\mathbf{x}_k, \mathbf{w}), \tag{6}$$

where $\mathbf{x}_k$ is the input, $\mathbf{y}_k$ is the output, and the nonlinear map $\mathbf{G}(\cdot)$ is parameterized by the vector $\mathbf{w}$. Typically, a training set is provided by sample pairs consisting of known input and desired output, $\{\mathbf{x}_k, \mathbf{d}_k\}$. The goal of the learning can be expressed to some degree as solving for $\mathbf{w}$ which minimizes the error of the machine: $\mathbf{e}_k = \mathbf{d}_k - \mathbf{G}(\mathbf{x}_k, \mathbf{w})$. In order to estimate the parameters by utilizing UKF, a new state-space representation can be written:

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \mathbf{r}_k,$$
$$\mathbf{d}_k = \mathbf{G}(\mathbf{x}_k, \mathbf{w}_k) + \mathbf{e}_k, \tag{7}$$

where the parameters $\mathbf{w}_k$ correspond to a stationary process with identity state transition matrix, driven by process noise $\mathbf{r}_k$.

Given that BOLD signal is the only output and observation of the system in terms of Balloon model, the dual estimation problem, in which the system states and model parameters must be estimated simultaneously, can be given as follows:

$$\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{v}_k, \mathbf{w}_k),$$
$$\mathbf{w}_{k+1} = \mathbf{w}_k + \mathbf{r}_k, \tag{8}$$
$$\mathbf{y}_k = \mathbf{G}(\mathbf{x}_k, \mathbf{w}_k) + \mathbf{e}_k.$$

Since standard UKF cannot be applied to this system immediately, dual UKF and joint UKF have been proposed as two alternatives. In the dual filtering method, two UKFs—one for state estimation, the other for parameter estimation—run in an alternate way. At each time step, the current estimate of parameters $\hat{\mathbf{w}}_k$ is used in the state filter as given input, and likewise the current states estimate $\hat{\mathbf{x}}_k$ is used in the parameter filter. On the contrary, a single UKF runs for both state and parameter estimation in the joint filtering. A higher dimensional joint state vector is defined: $\tilde{\mathbf{x}}_k = [\mathbf{x}_k{}^T \mathbf{w}_k{}^T]^T$, and the state-space model is reformed as follows:

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{F}}(\hat{\mathbf{x}}_k, \mathbf{u}_k, \hat{\mathbf{v}}_k),$$
$$\mathbf{y}_k = \mathbf{G}(\hat{\mathbf{x}}_k, \mathbf{n}_k). \tag{9}$$

The dual UKF and joint UKF approaches are illustrated in Figure 1.

In this section, the framework of UKF is briefly reviewed, a dual estimation problem with two approaches have been presented. In the next section, we will focus on examining the different performances of dual UKF and joint UKF, and all of our discussion will be in the context of Balloon model.

3. Results and Discussion

3.1. Biophysical Interpretation. One of the most prominent bifurcations between dual estimate and joint estimate is whether to incorporate interaction or correlation between states and parameters into filtering. As discussed earlier, the joint filter concatenates the state and parameter random variables into a single augmented state $(\hat{\tilde{x}}_k)$, so the cross-covariance between states and parameters is effectively modeled, that is,

$$E\left[\left(\hat{\tilde{x}} - E\left[\hat{\tilde{x}}\right]\right) \left(\hat{\tilde{x}} - E\left[\hat{\tilde{x}}\right]\right)^T \right] = \begin{bmatrix} \mathbf{P}_{x_k x_k} & \mathbf{P}_{x_k w_k} \\ \mathbf{P}_{w_k x_k} & \mathbf{P}_{w_k w_k} \end{bmatrix}. \tag{10}$$

Dual filtering, on the other hand, decouples (in a statistical sense) the dual estimation problem by treating states and parameters separately, which means $\mathbf{P}_{x_k w_k} = \mathbf{P}_{w_k x_k} = 0$. For states and parameters involved in Balloon model, no dynamic interaction or biophysical correlation between them has been observed (they are uncorrelated variables), therefore, it is reasonable to expect dual filtering to exhibit more biophysical accuracy. Thus the fMRI experiments substantiated our assumption.

The real fMRI data was acquired from 8 health subjects. 136 acquisitions in total were made (RT = 2s), in block of 8, giving 16 16-second blocks. The condition for successive
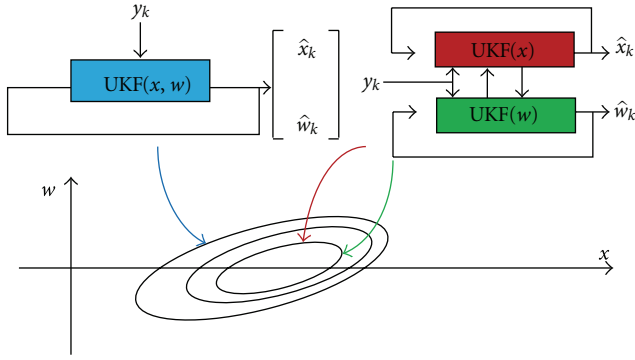
FIGURE 1: Schematic diagrams of joint filter (left) and dual filter (right).
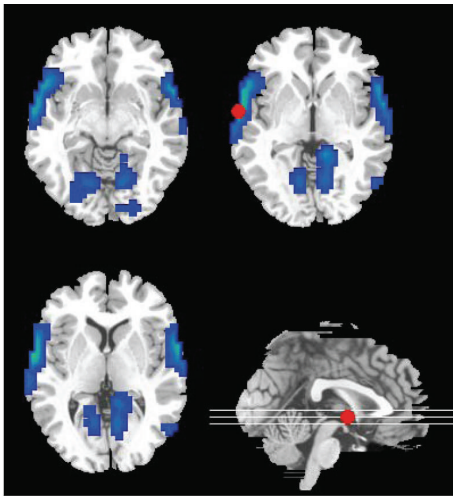
TABLE 2: The fact that joint filtering requires half of the iterations that are required by dual filtering has been taken into account.

Algorithm	Dimension of state vector	Total flops
Joint UKF	$L = L_x + L_w = 9$	285
Dual UKF	$L_x = 4, L_w = 5$	170

FIGURE 2: The greatest activated area of the group in the superior temporal gyrus (GT) for data assimilation.

overshoot, followed by gradual return to reduced plateau, and ending with a strong poststimulus undershoot. On the other hand, joint UKF fails to some degree to reconstruct a clean BOLD signal: the plateau is missing, neither does the evolving pattern of the signal show itself in a stable way in each block. The overestimated transit time (τ_0) leads to a reduction in amplitude of the BOLD peak; the less intense poststimulus undershoot can be explained by the underestimated signal decay (τ_s). Given the fact that dual UKF and joint UKF have very similar performance for state estimation, which is made clear in **Figure 3**, it is safe to attribute the failure to the undesired fluctuations of parameters (especially E_0, which affects the estimated signal directly by (3)), or more precisely, the undesired interaction between parameters and states.

3.2. Computational Interpretation. One of the most computationally expensive operations in UKF corresponds to calculating the new set of sigma points at each time update. This requires taking a matrix squareroot of the state covariance matrix, $\mathbf{P} \in \mathbb{R}^{L \times L}$, given by $\mathbf{SS}^T = \mathbf{P}$ [16]. An efficient implementation using a Cholesky factorization requires in general $\mathcal{O}(L^3/6)$ computations [26]. Therefore enlarging the dimension of state-space vector will dramatically increase the computational complexity (also can be referred to as time complexity). In this subsection we will introduce two criteria for evaluating the property of dual UKF and joint UKF in time complexity.

Number of Floating-Point Operations (flops). In computing, floating point can be thought of as a computer realization of scientific notation, which is able to represent a wide range of values. Flops number required by a given algorithm or computer program is independent of the computing platform, although its precise value may differ under different counting rules. MATLAB (version before 6.0) has provided us with a useful function *flops* to specify the cumulative number of flops. For instance, if A and B are N-by-N matrixes, then the output of *flops* $(A + B)$ and *flops* (AB) will approximately be N^2 and $2N^3$.

Since Cholesky factorization is the only operation within UKF whose time complexity is proportional to N^3, it is appropriate to consider that the flops number of Cholesky factorization is sufficient to determine the flops of the whole algorithm. For Balloon model, the dual estimation problem is about determining four state variables ($L_x = 4$) and five parameters (all the parameters except V_0 and α, $L_w = 5$) at each time step. Table 2 shows the total flops number of Cholesky factorization involved in each predict-update cycle of UKF.

blocks alternated between rest and auditory stimulation, starting with rest. Auditory stimulation was emotionally neutral words presented at a rate of 60 per minute. We selected the largest activated voxels in superior temporal gyrus (GT) to implement data assimilation [24] (**Figure 2**). Bias correction was performed using the method in [25]. The two algorithms were initialized in identical way on experimental data and parameters.

Figure 3 shows the hemodynamic states given by joint UKF and dual UKF. The lower peak of blood flow inferred from joint UKF corresponds to the smaller neuronal efficacy (ϵ) in **Figure 4**.

Parameters estimated are shown in **Figure 4**. Signal decay, autoregulation et al. remain unchanged (almost) during dual filtering. While for joint UKF, the parameters do not converge to their final values until the 4th $\sim$ 5th block (60 $\sim$ 80 s after the first stimulation). This phenomenon is a strong indicator for the introduced interaction between states and parameters.

Real and estimated fMRI signals are plotted in **Figure 5**. The simulated BOLD signal given by dual UKF shows a slight

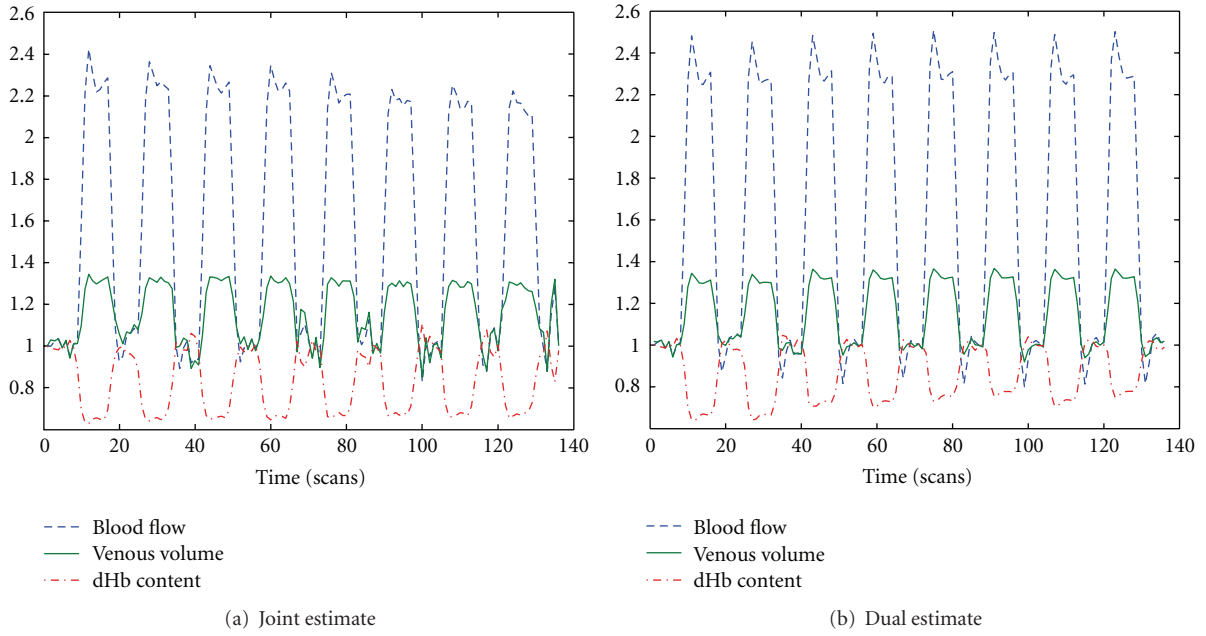(a) Joint estimate

(b) Dual estimate

FIGURE 3: States estimated by dual and joint UKF. The dotted line corresponds to change of blood flow, the solid line shows venous volume and dashed line depicts dHb content.
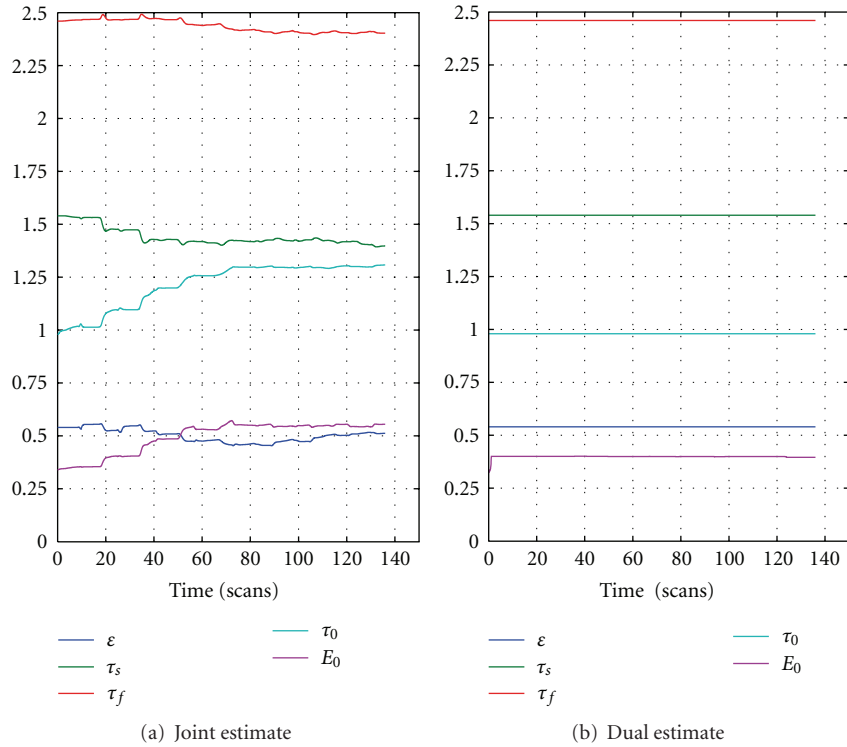

(a) Joint estimate

(b) Dual estimate

FIGURE 4: Parameters estimated by joint UKF and dual UKF. The mean values in Table 1 are used as initial values in our simulation. Apart from E_0, which grows sharply at the beginning and does not change afterwards, all the parameters obtained from dual UKF can be seen as constants. Resting oxygen extraction V_0 is the most important parameter in driving the model uncertainty [17], but simultaneous estimation of V_0 and other parameters would be impossible, as stated earlier. α is a nominal mechanism to BOLD signal. Therefore V_0 and α does not enter filtering.
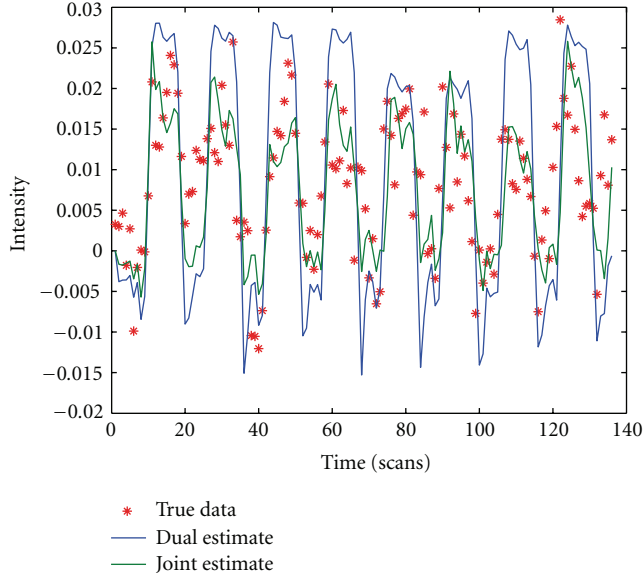
Figure 5: Estimated and real fMRI signal.

Table 3: Flops number for augmented state vector.

Algorithm	Dimension of state vector	Total flops
Joint UKF	$L = L_x + L_w = 19$	2470
Dual UKF	$L_x = 9, L_w = 11$	1582

However, in practice by slightly restructuring the state vector, the process and observation models, we may introduce the noise with the same order of accuracy as the uncertainty in the state. First, the state vector is augmented to give a $L^\alpha = L_x + L_v + L_n$ dimensional vector;

$$\mathbf{x}_k^\alpha = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{v}_k \\ \mathbf{n}_k \end{bmatrix}. \tag{11}$$

Then the process model is rewritten as a function of $\mathbf{x}_k^\alpha$, $\mathbf{x}_{k+1}^\alpha = \mathbf{F}(\mathbf{x}_k^\alpha, \mathbf{u}_k)$; the unscented transform uses sigma points that are drawn from

$$\mathbf{P}^\alpha = \begin{bmatrix} \mathbf{P}_x & 0 & 0 \\ 0 & \mathbf{R}_v & 0 \\ 0 & 0 & \mathbf{R}_n \end{bmatrix}, \tag{12}$$

where $\mathbf{R}_v$ and $\mathbf{R}_n$ are the process and observation noise covariance. In this situation, similarly we can derive Table 3.

For either case mentioned above, the flops number for joint filtering is at least 56% larger than that for dual filtering.

Comparing to flops count, overall execution time is a more tangible and practical criterion. We have tested our programs on several computers and collected their execution time data. Normally, a difference over 90% can be observed (16 s and 30 s for dual UKF and joint UKF, resp.). This result is even more impressive than that related to flops analysis, indicating that flops number is not the only factor influencing the operation time. Even if we take the rapid improvements in processing speed and memory into consideration, this variance is significant and should not be ignored.

4. Conclusions

In this paper we brought forward the dual Kalman filter as a reliable and efficient approach to estimating the states and parameters involved in balloon model. Comparing to the commonly used joint Kalman filter, its principle is in better conformity with the physiological reality, and by decoupling the dual estimation problem, it is much more calculational efficient to implement. The result of experiments showed good agreement with our conclusion.

Acknowledgments

References

[1] R. B. Buxton, E. C. Wong, and L. R. Frank, "Dynamics of blood flow and oxygenation changes during brain activation: the balloon model," *Magnetic Resonance in Medicine*, vol. 39, no. 6, pp. 855–864, 1998.

[2] R. D. Hoge, J. Atkinson, B. Gill, G. R. Crelier, S. Marrett, and G. Bruce Pike, "Investigation of bold signal dependence on cerebral blood flow and oxygen consumption: the deoxyhemoglobin dilution model," *Magnetic Resonance in Medicine*, vol. 42, pp. 849–863, 1999.

[3] Y. Zheng, J. Martindale, D. Johnston, M. Jones, J. Berwick, and J. Mayhew, "A model of the hemodynamic response and oxygen delivery to brain," *NeuroImage*, vol. 16, no. 3, pp. 617–637, 2002.

[4] T. Obata, T. T. Liu, K. L. Miller et al., "Discrepancies between BOLD and flow dynamics in primary and supplementary motor areas: application of the balloon model to the interpretation of BOLD transients," *NeuroImage*, vol. 21, no. 1, pp. 144–153, 2004.

[5] K. J. Friston, A. Mechelli, R. Turner, and C. J. Price, "Nonlinear responses in fMRI: the balloon model, volterra kernels, and other hemodynamics," *NeuroImage*, vol. 12, no. 4, pp. 466–477, 2000.

[6] K. J. Friston, "Bayesian estimation of dynamical systems: an application to fMRI," *NeuroImage*, vol. 16, no. 2, pp. 513–530, 2002.

[7] T. Deneux and O. Faugeras, "Using nonlinear models in fMRI data analysis: model selection and activation detection," *NeuroImage*, vol. 32, no. 4, pp. 1669–1689, 2006.

[8] J. J. Riera, J. Watanabe, I. Kazuki et al., "A state-space model of the hemodynamic approach: nonlinear filtering of BOLD signals," *NeuroImage*, vol. 21, no. 2, pp. 547–567, 2004.

[9] L. A. Johnston, E. Duff, and G. F. Egan, "Particle filtering for nonlinear BOLD signal analysis," *Lecture Notes in Computer Science*, vol. 4191, pp. 292–299, 2006.

[10] L. A. Johnston, E. Duff, I. Mareels, and G. F. Egan, "Nonlinear estimation of the BOLD signal," *NeuroImage*, vol. 40, no. 2, pp. 504–514, 2008.

[11] Z. H. Hu and P. C. Shi, "Nonlinear analysis of bold signal: biophysical modeling, physiological states, and functional activation," in *Proceedings of the 10th Internation Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI '07)*, pp. 734–741, Brisbane, Australia, 2007.

[12] P. Shi, Z. Hu, X. Zhao, and H. Liu, "Nonlinear analysis of the BOLD signal," *Eurasip Journal on Advances in Signal Processing*, vol. 2009, Article ID 215409, 2009.

[13] Z. Hu, C. Liu, P. Shi, and H. Liu, "Exploiting magnetic resonance angiography imaging improves model estimation of bold signal," *PloS One*, vol. 7, no. 2, Article ID e31612, 2012.

[14] K. J. Friston, N. Trujillo-Barreto, and J. Daunizeau, "DEM: a variational treatment of dynamic systems," *NeuroImage*, vol. 41, no. 3, pp. 849–885, 2008.

[15] J. Daunizeau, K. J. Friston, and S. J. Kiebel, "Variational bayesian identification and prediction of stochastic nonlinear dynamic causal models," *Physica D*, vol. 238, no. 21, pp. 2089–2118, 2009.

[16] R. van der Merwe and E. A. Wan, "The square-root unscented Kalman filter for state and parameter-estimation," in *Proceedings of the IEEE Interntional Conference on Acoustics, Speech, and Signal Processing (ICASSP '01)*, vol. 6, pp. 3461–3464, May 2001.

[17] Z. H. Hu and P. C. Shi, "Sensitivity analysis for biomedical models," *IEEE Transactions on Medical Imaging*, vol. 29, no. 11, pp. 1870–1881, 2010.

[18] K. E. Stephan, L. M. Harrison, W. D. Penny, and K. J. Friston, "Biophysical models of fMRI responses," *Current Opinion in Neurobiology*, vol. 14, no. 5, pp. 629–635, 2004.

[19] K. Uludağ, R. B. Buxton, D. J. Dubowitz, and T. T. Liu, "Modeling the hemodynamic response to brain activation," *NeuroImage*, vol. 23, supplement 1, pp. S220–S233, 2004.

[20] R. B. Buxton and L. R. Frank, "A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation," *Journal of Cerebral Blood Flow & Metabolism*, vol. 17, no. 1, pp. 64–72, 1997.

[21] K. Irikura, K. I. Maynard, and M. A. Moskowitz, "Importance of nitric oxide synthase inhibition to the attenuated vascular responses induced by topical L-nitroarginine during vibrissal stimulation," *Journal of Cerebral Blood Flow & Metabolism*, vol. 14, no. 1, pp. 45–48, 1994.

[22] D. E. Glaser, K. J. Friston, A. Mechelli, R. Turner, and C. J. Price, "Haemodynamic modelling," *Human Brain Function*, Elsevier Press, London, UK, 2003.

[23] R. L. Bellaire, E. W. Kamen, and S. M. Zabin, "A new nonlinear interated filter with application to target tracking," in *Proceedings of the AeroSense: 8th International Symposium on Aerospace/Defense Sensing, Simulation, and Controls*, vol. 2561, pp. 240–251, 1995.

[24] Z. Hu, P. Ni, C. Liu, X. Zhao, H. Liu, and P. Shi, "Quantitative evaluation of activation state in functional brain imaging," *Brain Topography*. In press.

[25] Z. Hu, H. Liu, and P. Shi, "Concurrent bias correction in hemodynamic data assimilation," *Medical Image Analysis*. In press.

[26] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, 2 edition, 1992.