Physiological System Identification with the Kalman Filter in Diffuse Optical Tomography

Solomon Gilbert Diamond[1], Theodore J. Huppert[1], Ville Kolehmainen[3], Maria Angela Franceschini[1], Jari P. Kaipio[3], Simon R. Arridge[2], and David A. Boas[1]

[1] Massachusetts General Hospital, Martinos Center for Biomedical Imaging*, Charlestown MA 02129, USA
sdiamond@nmr.mgh.harvard.edu
http://www.nmr.mgh.harvard.edu/PMI
[2] Department of Computer Science, University College London, London, UK
[3] Department of Applied Physics, University of Kuopio, Kuopio, Finland

Abstract. Diffuse optical tomography (DOT) is a noninvasive imaging technology that is sensitive to local concentration changes in oxy- and deoxyhemoglobin. When applied to functional neuroimaging, DOT measures hemodynamics in the scalp and brain that reflect competing metabolic demands and cardiovascular dynamics. Separating the effects of systemic cardiovascular regulation from the local dynamics is vitally important in DOT analysis. In this paper, we use auxiliary physiological measurements such as blood pressure and heart rate within a Kalman filter framework to model physiological components in DOT. We validate the method on data from a human subject with simulated local hemodynamic responses added to the baseline physiology. The proposed method significantly improved estimates of the local hemodynamics in this test case. Cardiovascular dynamics also affect the blood oxygen dependent (BOLD) signal in functional magnetic resonance imaging (fMRI). This Kalman filter framework for DOT may be adapted for BOLD fMRI analysis and multimodal studies.

1 Introduction

Diffuse optical tomography (DOT) is a noninvasive imaging technology that uses near infrared (IR) light to image biological tissue. The dominant chromophores in this spectrum are oxyhemoglobin (HbO), deoxyhemoglobin (HbR), lipids and water. The basis of DOT is *in vivo* near infrared spectroscopy of these dominant chromophores in the tissue. Tomographic images in DOT are constructed by simultaneously measuring from many local regions that cover a larger volume of tissue. The in-plane resolution limit of DOT increases rapidly with depth because biological tissue is a highly scattering medium for near infrared light. This diffuse property of the light also limits the penetration depth in adult human

* This research was funded by NIH T32-CA09502, P41-RR14075, R01-EB001954 and the MIND Institute.

J. Duncan and G. Gerig (Eds.): MICCAI 2005, LNCS 3750, pp. 649–656, 2005.

brain imaging to about 3 cm, which is sufficient to study most of the cerebral cortex. See Gibson et al. for a complete description of DOT [1]. Clinical and research applications of DOT arise due to is its specificity to the physiologically relevant chromophores HbO and HbR. Potential clinical and research applications for DOT abound in brain injury, degenerative neurovascular diseases and in cognitive neuroscience. Other research areas for DOT include fetal and neonatal monitoring and breast cancer detection. DOT is particularly suitable for *in situ* monitoring and multi-modal imaging [2].

The dynamics measured with DOT in the functional neuroimaging application are caused by local changes in blood volume and oxygenation in the scalp and in the brain. Due to the physical constraints of noninvasive imaging with DOT, the scalp and brain effects are combined in the measurements. The measured hemodynamics are caused by blood pressure regulation, cerebral blood flow autoregulation, local vasomotion and the vascular response to neuronal activity. Complexity arises because of interactions between these factors. The primary aim of DOT functional neuroimaging is to separate the stimulus related brain function signal from the background physiology related signal. The main problem is that the latter of these two is much stronger. A method to help resolve the physiological components in DOT is to include noninvasive auxiliary physiological measurements in the analysis. Many instruments can be used during DOT experiments. Examples are the blood pressure monitor, pulse oximeter, electrocardiogram (ECG), chest band respirometer, spirometer and capnograph. A further complexity of DOT analysis is that even when auxiliary physiology is included in the analysis, their effects do not appear to be stationary in time or space. We commonly observe that signal dynamics that are correlated with respiration, for example, will vary significantly in amplitude and relative phase angle at different measurement locations even when breathing rate and depth are held constant. The present objective is to separate the physiological components of DOT with a dynamical model.

State-space estimation has previously been applied to DOT without physiological regressors [3]. Prince et al. [4] fit the amplitude and phase angle of three non-stationary sinusoids to DOT time-series data using the Kalman filter. While supporting the principle of using the Kalman filter in DOT analysis, the three-sinusoid model does not allow for the most commonly used event related experimental designs nor can it use readily available physiological measurements such as blood pressure as a regressor. Zhang et al. used principal component analysis (PCA) to reduce the background physiological variance in functional neuroimaging experiments [5]. Anecdotal evidence was presented that certain principal components correlate with blood pressure and respiratory dynamics. This observation of statistically uncorrelated blood pressure and respiratory dynamics contradicts known respiratory interactions in blood pressure regulation [6]. Due to physiological interactions, the orthogonal projections in PCA are more likely to be mixtures of physiological effects. Standard linear regression methods in fMRI analysis [7] accept multiple regressors that could easily include auxiliary physiological measurements but will not accommodate temporal non-

stationarity of the linear models. Dynamical system identification for fMRI is at the forefront of new analysis methods [8]. These advances in fMRI have not extended to DOT thus far mainly because the DOT inverse problem is typically ill posed and requires a more complicated physical model. In this paper, we present a framework that employs the Kalman filter for dynamical modeling of the physiological components of DOT.

2 Methods

The Kalman filter is a recursive solution to discrete linear filtering and prediction problems [9]. The objective of the Kalman filter is to estimate the time-varying states of a discrete-time process that is described by stochastic equations for updating the states and measurements over time. There are many ways to model the same physical system within the generality of the Kalman filter. Our proposed Kalman filter model of the DOT system begins by naming the constants in table 1 and naming the model variables in table 2. The respective sizes of each variable are indicated with a parenthetical subscript notation and those that vary with time are indicated.

The inputs $\mathbf{u}$ in the DOT model are the Boolean stimulus time vector and time-series physiological measurements such as blood pressure and heart rate variability. The states $\mathbf{x}$ in the DOT model are the discrete finite impulse response (FIR) functions that are convolved with the inputs to yield local concen-

Table 1. Length constant names

n_u inputs (regressors)	n_w wavelengths
n_d source-detector pairs	n_c chromophores
n_s voxels	n_g spatial basis functions
n_r regression time points	n_h temporal basis functions
n_x states ($n_x = n_u n_c n_g n_h$)	n_y measurements ($n_y = n_w n_d$)
n_z auxiliary states ($n_z = n_u n_c n_g n_r$)	n_k total time points

Table 2. Variable names and sizes

$k(t)$	time index	$\mathbf{u}_{(n_u,1)}(t)$	input vector
$\mathbf{x}_{(n_x,1)}(t)$	state vector	$\mathbf{V}_{(n_x,n_x)}(t)$	state covariance
$\mathbf{w}_{(n_x,1)}(t)$	process noise	$\mathbf{Q}_{(n_x,n_x)}$	process noise covariance
$\mathbf{z}_{(n_z,1)}(t)$	auxiliary state vector	$\mathbf{y}_{(n_y,1)}(t)$	measurement vector
$\mathbf{v}_{(n_y,1)}(t)$	measurement noise	$\mathbf{R}_{(n_y,n_y)}$	meas. noise covariance
$\mathbf{A}_{(n_x,n_x)}$	state update model	$\mathbf{B}_{(n_z,n_z)}$	auxiliary update model
$\mathbf{C}_{(n_z,n_u)}$	auxiliary input model	$\mathbf{D}_{(n_y,n_x)}(t)$	measurement model
$\mathbf{K}_{(n_x,n_y)}(t)$	Kalman gain matrix	$\mathbf{S}_{(n_y,n_u n_y)}$	summing matrix
$\mathbf{U}_{(n_u n_c n_g,n_z)}(t)$	input matrix	$\mathbf{M}_{(n_u n_c n_g,n_z)}$	input mask matrix
$\mathbf{L}_{(n_u n_y,n_u n_w n_s)}$	pathlength matrix	$\mathbf{L}_{0(n_w n_d,n_w n_s)}$	pathlength submatrix
$\mathbf{G}_{(n_u n_w n_s,n_u n_w n_g)}$	spatial basis set	$\mathbf{G}_{0(n_s,n_g)}$	spatial basis submatrix
$\mathbf{E}_{(n_u n_w n_g,n_u n_c n_g)}$	extinction matrix	$\mathbf{E}_{0(n_w,n_c)}$	extinction submatrix
$\mathbf{H}_{(n_z,n_x)}$	temporal basis set	$\mathbf{H}_{0(n_r,n_h)}$	temporal basis submatrix

tration changes in HbO and HbR. In order to perform this convolution within a Kalman filter framework, it is convenient to define auxiliary states $\mathbf{z}$ that merely store a regression length n_r of most recent inputs. The measurements $\mathbf{y}$ are the time-series of changes in optical density (ΔOD) for each source-detector pair and wavelength. This discrete-time process can be described as

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} \tag{1}$$
$$\mathbf{z}_k = \mathbf{B}\mathbf{z}_{k-1} + \mathbf{C}\mathbf{u}_k \tag{2}$$
$$\mathbf{y}_k = \mathbf{D}(\mathbf{z}_k)\mathbf{x}_k + \mathbf{v}_k \ . \tag{3}$$

In order to include a physical model for the DOT inverse problem, the model elements $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are defined as

$$\mathbf{A} = \mathbf{I}_{(n_x)} \tag{4}$$
$$\mathbf{B} = \mathbf{I}_{(n_u n_c n_g)} \otimes \begin{bmatrix} \mathbf{0}_{(1,n_r-1)} & 0 \\ \mathbf{I}_{(n_r-1)} & \mathbf{0}_{(n_r-1,1)} \end{bmatrix} \tag{5}$$
$$\mathbf{C} = \mathbf{I}_{(n_u)} \otimes \mathbf{1}_{(n_c n_g,1)} \otimes \begin{bmatrix} 1 & \mathbf{0}_{(n_r-1,1)} \end{bmatrix} \tag{6}$$
$$\mathbf{S} = \mathbf{1}_{(1,n_u)} \otimes \mathbf{I}_{(n_y)} \tag{7}$$
$$\mathbf{L} = \mathbf{I}_{(n_u)} \otimes \mathbf{L}_0 \tag{8}$$
$$\mathbf{G} = \mathbf{I}_{(n_u n_w)} \otimes \mathbf{G}_0 \tag{9}$$
$$\mathbf{E} = \mathbf{I}_{(n_u)} \otimes \mathbf{E}_0 \otimes \mathbf{I}_{(n_g)} \tag{10}$$
$$\mathbf{M} = \mathbf{I}_{(n_u n_c n_g)} \otimes \mathbf{1}_{(1,n_r)} \tag{11}$$
$$\mathbf{H} = \mathbf{I}_{(n_u n_c n_g)} \otimes \mathbf{H}_0 \tag{12}$$
$$\mathbf{U}(\mathbf{z}_k) = \mathbf{1}_{(n_u n_c n_g,1)} \mathbf{z}_k^T \odot \mathbf{M} \tag{13}$$
$$\mathbf{D}(\mathbf{z}_k) = \mathbf{SLGEU}(\mathbf{z_k})\mathbf{H} \ , \tag{14}$$

where T is the transpose operator, $\otimes$ is the Kronecker tensor product, $\odot$ is term-by-term array multiplication, $\mathbf{I}$ is the identity matrix, $\mathbf{1}$ is a matrix of ones, $\mathbf{0}$ is a matrix of zeros and matrix sizes are indicated with parenthetical subscripts. The submatrix $\mathbf{L}_0$ is a block diagonal matrix formed from meassurement by voxel average effective pathlengths for each wavelength as described by [10]. The columns of $\mathbf{G}_0$ contain a set of spatial basis functions that can be used to reduce the number of states and/or impose spatial smoothing of the state estimates. Known optical extinction coefficients are contained in the wavelength by chromophore submatrix $\mathbf{E}_0$. The columns of $\mathbf{H}_0$ contain temporal basis functions to reduce the number of states and/or impose temporal smoothing.

The Kalman filter is a recursive solution to the state estimation problem for the discrete-time process described by equations 1, 2 and 3. The recursions require initialization of the state estimate $\hat{\mathbf{x}}_0$ and estimated state covariance $\hat{\mathbf{V}}_0$ and then proceed with the following prediction-correction algorithm

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} \tag{15}$$
$$\mathbf{z}_{k|k-1} = \mathbf{B}\mathbf{z}_{k-1|k-1} + \mathbf{C}\mathbf{u}_k \tag{16}$$

$$\hat{\mathbf{V}}_{k|k-1} = \mathbf{A}_t \hat{\mathbf{V}}_{k-1|k-1} \mathbf{A}^T + \mathbf{Q} \tag{17}$$

$$\mathbf{U}_k = \mathbf{1}_{(n_u n_c n_g, 1)} \mathbf{z}_{k|k-1}^T \odot \mathbf{M} \tag{18}$$

$$\mathbf{D}_k = \mathbf{SLGEU}_k \mathbf{H} \tag{19}$$

$$\mathbf{K}_k = \hat{\mathbf{V}}_{k|k-1} \mathbf{D}_k^T \left(\mathbf{D}_k \hat{\mathbf{V}}_{k|k-1} \mathbf{D}_k^T + \mathbf{R} \right)^{-1} \tag{20}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_t \left(\mathbf{y}_k - \mathbf{D}_k \hat{\mathbf{x}}_{k|k-1} \right) \tag{21}$$

$$\mathbf{z}_{k|k} = \mathbf{z}_{k|k-1} \tag{22}$$

$$\hat{\mathbf{V}}_{k|k} = \hat{\mathbf{V}}_{k|k-1} - \mathbf{K}_k \mathbf{D}_k \hat{\mathbf{V}}_{k|k-1} \ . \tag{23}$$

We designed an experiment to test the basic functionality of this proposed method. We combine real DOT data with a simulated functional response and then analyze the result with static deconvolution and the proposed Kalman filter method. This experiment allows us to compare the estimated responses with the "true" response. Only a single trial is examined so the results are mainly illustrative. Data was collected from a human subject who was instructed to sit quietly and breath freely. Measurements were taken with a continuous wave DOT instrument [11] then demodulated and down sampled to 1 Hz. The measurements were high pass filtered in a forward then reverse direction with a 6th order IIR Butterworth filter with a cutoff frequency of 0.05 Hz and zero phase distortion. This filtering removes slow physiology that is sufficiently outside the frequency range of interest for the hemodynamic response that it can be ignored. Short-term variability including the respiratory sinus arrhythmia, Mayer waves and vasomotion remain after filtering. The photon fluence $\Phi(t, \lambda)$ was then converted to a change in optical density $\Delta\mathrm{OD}$

$$\Delta\mathrm{OD}(t, \lambda) = \ln \left(\frac{\Phi(t, \lambda)}{\Phi_0(\lambda)} \right) \ , \tag{24}$$

where Φ_0 is average detected photon fluence, $\Delta\mathrm{OD}(t, \lambda)$ are measurements $\mathbf{y}$ as a function of time t and wavelength λ. The three model inputs contained in $\mathbf{u}$ were the Boolean stimulus time vector, the blood pressure (BP) and heart rate variability (HRV) with normalized variances.

Data from only a single source fiber and three detector locations were included in the analysis. The three detectors were arranged about 2 cm apart in a row and the source was placed 3 cm away from the center detector and equidistant from the other two. Three voxels defined the tissue volume under the optical probes. Voxel 1 represented the scalp and was common to all the detectors. Voxels 2 and 3 represented two regions of the brain located under the scalp voxel. A simulated functional response was added into the baseline hemodynamics in voxel 2. The stimulus paradigm was event related with a 12 to 18 second inter-stimulus interval over the 300 second trial. The model used to simulate the hemodynamics was one period of a raised cosine with a delay and amplitude set differently for the HbO and HbR functional responses. The simulated waveforms can be seen in the results figures.

The pathlength submatrix $\mathbf{L}_0$ was computed with a diffusion approximation to the transport equation for a semi-infinite medium [10]. An identity matrix

was used for the spatial basis set $\mathbf{G}_0$ and a normalized Gaussian function was used for the temporal basis set $\mathbf{H}_0$. The standard deviation for the Gaussian function was fixed at 1.5 seconds and the means were separated by 1.5 seconds over the regression time. The same temporal basis set was used for the static deconvolution. The state update noise covariance $\mathbf{Q}$ only contained nonzero terms on the diagonal elements. Diagonal terms related to the functional response were set to 3×10^{-6} and those related to BP and HRV were set to 10^{-5}. This imbalance in state update noise caused the functional response model to evolve more slowly than the systemic physiological models. The measurement noise covariance matrix $\mathbf{R}$ was set to an identity scaled by 10^{-3}. These variances act as regularization and were adjusted to stabilize the estimation scheme.

3 Results

The state estimates from the Kalman filter were propagated through the forward model to calculate the component of the measurements that relates to each input. An example result of this signal separation for the 890 nm measurement from detector 1 is shown in figure 1. The functional response to the stimulus only accounts for 2.8% of the variance in the measurement whereas BP and HRV account for 11.3% and 77.9% respectively. The sum of the modeled components accounts for over 99.9% of the variance in the measurement.

The results of the static deconvolution analysis to recover the functional hemodynamic response is shown in figure 2. The functional responses are clearly present in the estimates but are distorted by large physiological noise artifacts. Compared to the true hemodynamics, the HbO estimate resulted in $R^2 = 0.78$, which is reasonably good considering that the physiological noise dominates the measurement. For the smaller HbR signal, $R^2 = 0.57$ with the true hemodynam-

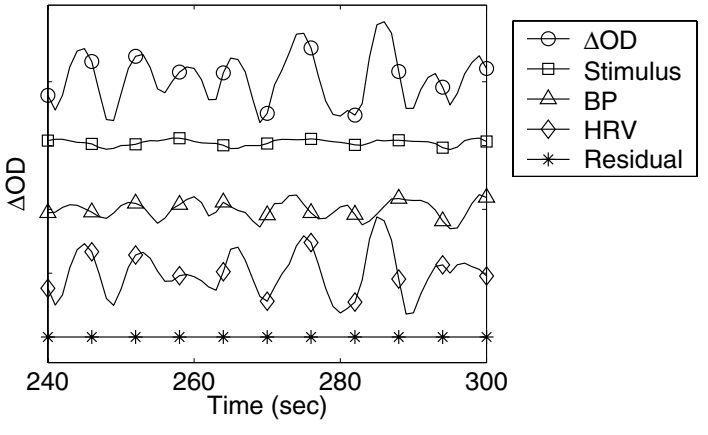

Fig. 1. Separating a ΔOD measurement into components related to each model input. The scale for each component of ΔOD was shifted for visual comparison.

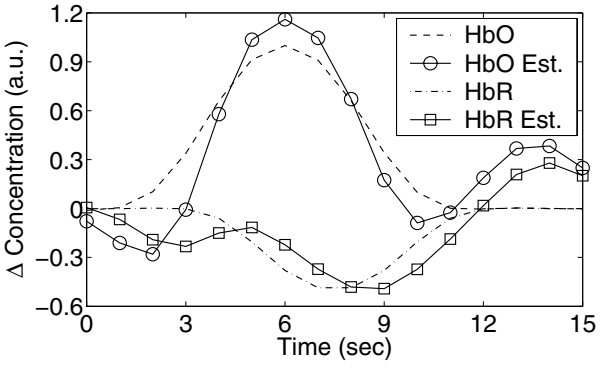

Fig. 2. Result for deconvolution of functional response from hemodynamics in voxel 2

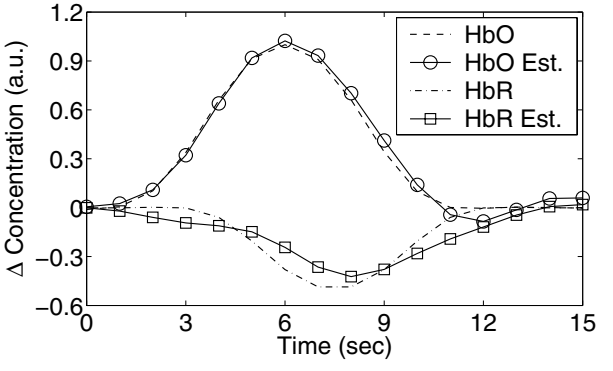

Fig. 3. Result for proposed Kalman filter estimate functional response in voxel 2

ics, indicating that the physiological noise artifacts are of comparable magnitude to the actual response.

The result for the Kalman filter was taken to be the last state estimate computed during a forward pass through the data. This result, shown in figure 3, appears to be a significant improvement over the deconvolution approach. The HbO estimate improved to $R^2 = 0.99$ and the HbR estimate jumped to $R^2 = 0.89$. There is some lag in the Kalman filter result which may have been caused by only using a forward pass through the data.

4 Discussion

We successfully implemented the Kalman filter for system identification in DOT. Based on the preliminary results described, the proposed analysis framework may help to improve estimates of functional hemodynamics in DOT neuroimaging. This result is potentially significant because improved hemodynamic estimates could make a broader range of brain activation paradigms possible with DOT.

The ability to separate signals into physiological components may also reveal new information about the local regulatory physiology and may be useful in identifying certain vascular pathologies. Unlike the prior work with the Kalman filter for DOT, the present formulation has the flexibility to be applied to any experimental design and for problems of reasonably large spatial and temporal dimension. The proposed Kalman filter formulation may also be useful for other imaging modalities such as fMRI, MEG and EEG or when multiple modalities are combined with a single state-space model of the underlying physiology.

References

1. A.P. Gibson, J.H., Arridge, S.: Recent advances in diffuse optical imaging. Phys. Med. Biol. **50** (2005) R1–R43
2. Strangman, G., Culver, J., Thompson, J., Boas, D.: A quantitative comparison of simultaneous bold fMRI and NIRS recordings during functional brain activation. NeuroImage **17** (2002) 719–731
3. Kolehmainen, V., Prince, S., Arridge, S., Kaipio, J.: State-estimation approach to the nonstationary optical tomography problem. J. Opt. Soc. Am. A **20** (2003) 876–889
4. Prince, S., Kolehmainen, V., Kaipio, J., Franceschini, M., Boas, D., Arridge, S.: Time-series estimation of biological factors in optical diffusion tomography. Phys. Med. Biol. **48** (2003) 1491–1504
5. Zhang, Y., Brooks, D., Franceschini, M., Boas, D.: Eigenvector-based spatial filtering for reduction of physiological interference in diffuse optical imaging. Journal of Biomedical Optics **10** (2005) 011014–1–11
6. Cohen, M.A., Taylor, J.A.: Short-term cardiovascular oscillations in man: measuring and modelling the physiologies. Journal of Physiology **542.3** (2002) 669–683
7. Frackowiak, R., Friston, K., Frith, C., Dolan, R., Price, C., Zeki, S., Ashburner, J., Penny, W., eds.: Human Brain Function. 2nd edn. Academic Press (2003)
8. Roche, A., Pinel, P., Dehaene, S., Poline, J.: Solving incrementally the fitting and detection problems in fMRI time series. Medical Image Computing and Computer-Assisted Intervention (MICCAI'04), Part 2, Lecture Notes in Computer Science **3217** (2004) 719–726
9. Kalman, R.: A new approach to linear filtering and prediction problems. Trans. of the ASME–Journal of Basic Engineering **82** (1960) 35–45
10. Arridge, S., Cope, M., Delpy, D.: The theoretical basis for the determination of optical pathlengths in tissue: temporal and frequency analysis. Phys. Med. Biol. **37** (1992) 1531–1560
11. Franceschini, M., Fantini, S., Thompson, J., Culver, J., Boas, D.: Hemodynamic evoked response of the sensorimotor cortex measured non-invasively with near-infrared optical imaging. Psychophysiology **40** (2003) 548–560