Motion Artifact Reduction in PPG Signals from Face : Face Tracking & Stochastic State Space Modeling Approach

Chankyu Park[1] and Ho-jin Choi[2]

Abstract— **The Photoplethymography(PPG) is generally measured on a finger or an ear using contact sensors. The recent several studies using non-contact sensor such as CCD camera and web-cam to measure PPG have been introduced. However the motion artifact issue is also emerging in non-contact camera sensing similar to contact-type one because it is sensitive to artifacts generated by subject's head and body motion. In this paper, the two sequential approaches for a motion artifact reduction algorithm are presented; the one is a face tracking method that detects and tracks the skin region of face which is containing PPG signals, the other is the reduction method of motion artifact due to various head & face movement such as roll, yaw, pitch, translation and scale. Results of the proposed KF are compared to these of the FIR band pass filter(BPF).**

I. INTRODUCTION

These days, with accelerating evolution of wearable devices, the applicable area of the PPG has been extended to that of sports and daily life as well as affective computing area. In recent years, several methods using non-contact sensor such as CCD camera and web-cam have been introduced under the desktop or mobile computing environment[9][12][11][7].

The studies showed the possibilities of extracting the PPG signal in the face remotely[9][12]. Currently, everyone can use several smart-mobile apps to implement the extraction the PPG signal using its built-in camera from user's face and to show his/her heart rate as basic function and are launched as two OS version such as the iOS and the Android[1].

However, the motion artifact issues are also emerging in non-contact camera sensing similar to contact-type measurement. In detail, motion artifacts are accompanied by inherent causes; variation of face region caused by roll, pitch, yaw, translation and arbitrary motion of head, changes of background and illumination around face and a blurring effect by high speed motion as well as changes of face expression. The recent studies and smart phone apps using camera sensors only have focused on proving the validation of extracted PPG signals from faces that are not moving[11][7]. Thus, a robust artifact canceling method needs to be developed for using desktop or mobile devices while users' face and head are moving naturally and unconsciously.

In this paper, We regard PPG signals as stochastic times series. PPG time series are modeled by state space modeling(SSM) approach and its system parameters are estimated by subspace system identification[3][4][10][6]. Finally, the Kalman filter(KF) built by these parameters are applied to predict and correct distorted PPG signals. Results of the proposed KF are compared to these of the FIR band pass filter(BPF). The paper is organized as follows; Section II describes the proposed face tracking method and the SSM model for the PPG signal from face; Section III reports the performance results and discussion.

II. METHODS

A. Motion Artifacts and PPG

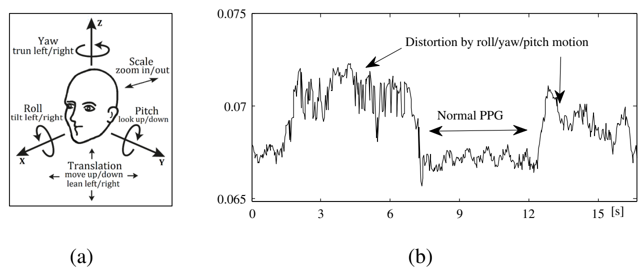

Fig. 1. (a) Head motion, (b) an example of distorted PPG signals

One of basic and strong conditions for the PPG measurement should be generally free body motion for acquisition of the normal PPG signal. In comparison with using contact sensors[5], the type of sensor, the measurement spot, and the causes of motion artifacts are changed[8][2]. Fig. 1 (a), (b) shows various types of head movements and the example of real distorted PPG signal due to them. We can clearly identify the difference of signal's distortion due to motion artifacts by head compared to period of normal PPG in the Fig. 1 (b).

B. Face Tracker

For experimental setup, we used Logitech C905 model as a front web-cam camera to record the videos. Videos were recorded at 30.00 fps and 24-bit RGB color with 640x480 pixel. 10 volunteers(8 males, 2 females) were participated in this experiment with consent. Their ages are between 26-51 years. Subjects were seated at a desk and looking at the monitor with the web-cam camera. Subjects are asked to elicit various head posture: roll, yaw, pitch, scale, translation.

*This work was supported by the IT R&D program of MSIP/KEIT. [10041826 , Development of emotional features sensing, diagnostics and distribution s/w platform for measurement of multiple intelligence from young children]

[1]Chankyu Park is with Senior Engineering Staff of Dept. of Intelligent Cognitive Research , Electronics Telecommunication Research Institute,161, gajeong-dong yusunggu, Daejeon, KOREA `parkck at etri.re.kr`

[2]Ho-jin Choi is with the Department of Computer Science, Korea Advanced Institute of Science and Technology(KAIST), 291, daehakro yusung, Daejeon, KOREA `hojinc at kaist.ac.kr`

In order to constantly keep track of the region of interest (ROI) as face, the tracker need to assign the initial region to track. First, we conduct two step to initialize ROI as follows. Repeat face detection until a face is detected by the face detector using OpenCV library. Once a face is detected, repeat finding a nose by the nose detector in the face until a nose is found. The found nose $R(0)$ is finally set to initial ROI as input of continuously adaptive mean shift(CAMShift) tracker. The $R(0)$ is converted to hue channel of the HSV color space. Whenever the CAMShift tracker repeats the stage of tracking face, updates the rectangular region $R^{'}(t)$ including a face and background around it. Fig. 2 shows that the system does the next step to extract face region from $R^{'}(t)$ in every frame. The face region $S(t)$ used to calculate PPG components is defined by

$$S(i,j,t) = \begin{cases} I_G(i,j,t) & \text{if } R^{'}(i,j,t) < th \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where $(i,j) \in R^{'}(t)$, $I_G(i,j,t) =$ G channel of I_{RGB} image, $th =$ threshold value to exclude non-skin region. We set $th = 0.1$ in this study. $S(i,j,t)$ plays a role in extracting a skin region of face as a mask. We used only a green channel $I_G(i,j,t)$of I_{RGB} image to extract PPG signals, because we examined that it contains the largest amount of PPG component in various color models; RGB, YCbCr, HSV, and Grey. The recent researches about camera-based PPG extraction [3][4][5] support it to use $I_G(i,j,t)$.

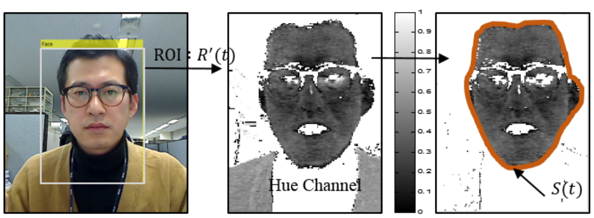

Fig. 2. Extration the region of face $S\hat{(}t)$ from ROI $R^{'}(t)$

The face region $S(i,j,t)$ in (1) finally is converted into a type of 1 dimensional PPG signal in time domain. We define the PPG signal y_t as follows

$$y_t = \frac{1}{N_1} \sum_{i=1}^{N} \left(\sum_{j=1}^{M} \frac{S(i,j,t)}{M_j} \right) \tag{2}$$

where $(M,N) =$ size of $S(i,j,t)$, $M_j =$ count of all nonzero number in jth column vector and $N_1 =$ count of all nonzero number in a $1st$ row vector. The PPG y_t is derived by averaging over all green channel pixels in the $S(t)$.

C. State space Model

One of the appealing features of state space models is that many traditional models, such as AR and ARIMA, can be expressed in a linear state space system[4][10]. For linear Gaussian state space models, the KF can be used to compute predictors of the state variables and one-step-ahead predictors of the observations.

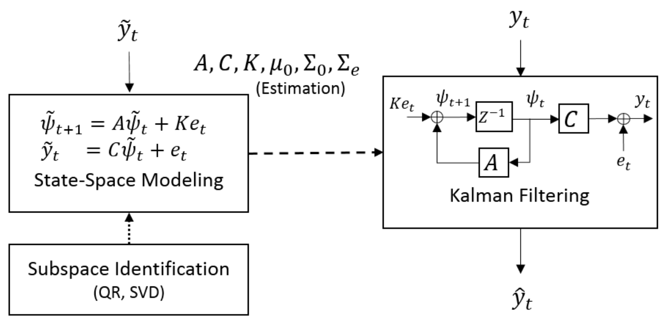

Fig. 3. The unified approach of motion artifact reduction : state space modeling and Kalman filtering for reducing of motion artifact

Fig. 3 shows the unified approach considering two aspect of estimations; system identification as a modeling, Kalman filtering as a deployment. In other words, the former is related to an estimation of unknown parameters in the system matrices: maximum likelihood estimation, subspace state space systems identification(4SID)[10]. The latter is related to an estimation of the unobservable state in the system: prediction, filtering and smoothing. In this study, we exploit subspace identification method to estimate unknown system parameters for building a SSM which is not sensitive to motion artifacts, because it has advantages to other methods in terms of model complexity, computational efficiency and simplicity which means building models directly from the input/output data. Thus, in terms of subspace identification, given $\tilde{y}_t$ as a only output data without input, then the method estimates system parameters of $\tilde{x}_t$, A, C, K, μ_0, Σ_0, Σ_e in Fig. 3. In terms of Kalman filtering, given y_t as observation data, A, C, K, μ_0, Σ_0, then the KF estimates system's state variable of $\hat{x}_{t+1}$, Σ_e in Fig. 3.

The definition and modeling of $\tilde{y}_t$ used as output variable in system identification of Fig. 3 is first step to build up the SSM. A basic model for representing a PPG time series is the additive model in this study

$$y_t = \alpha_t + \psi_t + \varepsilon_t \qquad\qquad \varepsilon_t \sim N(0, \Sigma_\varepsilon) \tag{3}$$

where the classical decomposition is $y_t =$ actual PPG observation, $\alpha_t =$ trend component including motion artifacts, $\psi_t =$ periodic PPG component, $\varepsilon_t =$ irregular component. The behavior of y_t can be generally explained by expressing a regression model whose explanatory variables were a deterministic trend and a set of periodic variables[4]. If these components are not stable then this formulation would be not proper and it would be necessary for the regression coefficients to change over time. This flexibility is possible with structural models like SSM[4]. The above (3) can be relaxed in order to allow the series' level to change through time, leading to the local level model, in which the level at each t is the sum of the previous period value and a random element. In order to consider motion artifacts, we also define these as members of trend component as well as respiration, because we have found that these components was belonged to low frequency band ($0 \sim 1$ Hz) through iteration of experiments. The model (3) can be defined as:

$$\alpha_t \simeq m_{t-1} + r_{t-1} + \eta_t \qquad \eta_t \sim N(0, \Sigma_\eta) \tag{4}$$

$$\psi_t \simeq \psi_{t-1} + \zeta_t \qquad \zeta_t \sim N(0, \Sigma_\zeta) \tag{5}$$

where m_t = motion artifacts, r_t = respiration components, η_t = irregular component in this trend local level, and ζ_t = irregular component in this period local level.

Second, we exclude the trend component α_t in y_t beforehand for letting the system not to know motion artifacts in the modeling phase, so that the system can try to fit as closely as possible the measured data of only $\psi_t + \varepsilon_t$. Therefore, we take a simple model $\tilde{y}_t$ in which trend ψ_t, with no all trend α_t.This model is simplified as:

$$\tilde{y}_t = \psi_t + \varepsilon_t \tag{6}$$

$$\psi_{t+1} = \psi_t + \zeta_t \tag{7}$$

Finally, we convert the system (6), (7) to a state space model of a stochastic system in an innovation form.

$$\tilde{\psi}_{t+1} = A\tilde{\psi}_t + Ke_t \tag{8}$$

$$\tilde{y}_t = C\tilde{\psi}_t + e_t \tag{9}$$

where K is the Kalman gain and e_k is an unknown innovation with $e_t \sim N(0, \Sigma_e)$. The system (6), (7) has good physical concept, however, the innovation model is more suitable for 4SID methods, because it has a noise model with less degrees freedom, which can be more appropriately identified from the given $\tilde{y}_t$ output data[10].

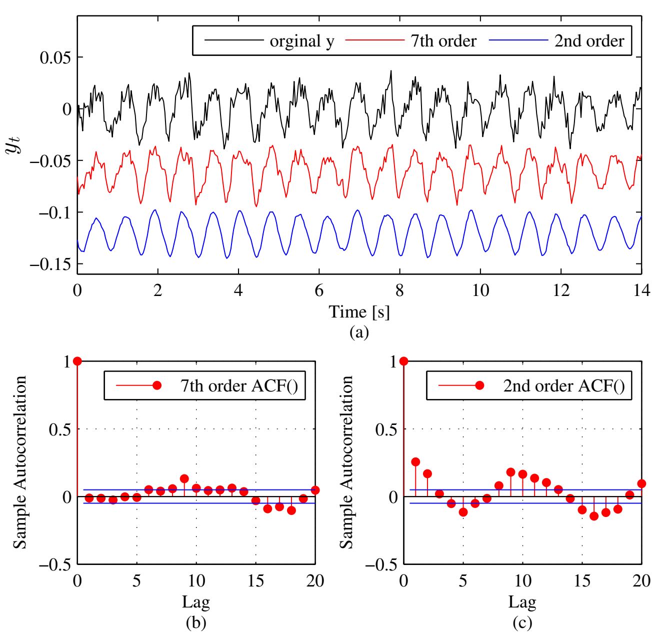

Fig. 4. Modeling result of $\tilde{y}_t$ by 4SID : validation and autocorrelation test

Experiments are performed on this web-cam based PPG system so as to acquire $\tilde{y}_t$ time series, which do not contain motion artifacts in accordance with our (6) definition. Fig. 4 shows modeling results by subspace identification method, which are validation results according to the order of the fitted model by $\tilde{y}_t$ and autocorrelation results of residuals which are differences between the one-step-predicted output from the model and $\tilde{y}_t$ the validation data. Residuals mean the part of the validation data not described by the model. The chosen $\tilde{y}_t$ for modeling has properties of sampling frequency = 30 Hz, heart rate frequency = 1.38 Hz = 82.8 bpm, and the length = 3 min.

The validation signal in Fig. 4(a) is getting to fit the original $\tilde{y}_t$. Each autocorrelation function(ACF) of residuals in Fig. 4(b), (c) is getting low with increasing order of the model. However, with the opposite case, the signal is gradually smoothing and ACF is getting high. In terms of the regression, it is preferable to increase the order of the model, but it is rather advantageous to decrease it in this study, because we can consider to remove noises of high frequency band due to the low resolution of sampling frequency 30 Hz and low dynamic range of flesh tone color. Finally, we adopt the order of the model to 2nd order to satisfy both morphology of PPG and model complexity simultaneously. This model is as follows:

$$A = \begin{bmatrix} 0.9511 & 0.2936 \\ -0.2812 & 0.9571 \end{bmatrix}, K = \begin{bmatrix} 9.657 \\ -3.917 \end{bmatrix} \tag{10}$$

$$C = \begin{bmatrix} 0.01351 & -0.002839 \end{bmatrix} \tag{11}$$

III. Results And Discussion

In this section, we demonstrate usefulness and performance of our proposed stochastic state space model for PPG signals from face and the Kalman filtering for reducing motion artifacts. The two general criterion of a waveform analysis in time domain and power spectral density(PSD) analysis in frequency domain are used to compare the proposed KF with the conventional finite impulse response(FIR) band pass filter (BPF). The KF is configured to 2nd order system parameters (10), (11) generated by subspace system identification: 4SID. The BPF is configured to bandwidth = 0.3 Hz ∼ 2.5 Hz, sampling frequency = 30 Hz, filter type = equiripple BF, and filter order = 50th. The Welch method was used to estimate PSD and all experiments in this study are done under the environment of the Matlab.

Fig. 5 shows results of reducing motion artifact due to the representative cases of head & body motions. Videos used for this test are chosen from the portion of 10 subjects records. We can easily grasp how much four kinds of PPG signals in Fig. 5 are distorted whenever various head related motions are happening. We found that distorted PPG signals have a common feature in that they are steeply ascent or descent whenever head motions are tried through repeated experiments and inspection.

The waveform results of the 50th order BPF are generally not cope with steep variation and wave trends of a low frequency band about all cases in Fig. 5. Whereas, in terms of PSD, the location of two estimates of heart rate frequency= f_{HR} of the BPF in Fig. 5 (a), (d) is nearly the same with those of the proposed KF. Fig. 5 (b) shows that many portion of distorted PPG is still not recovered in case of the BPF. It estimates the wrong f_{HR} compared to the KF. Fig. 5 shows that the proposed KF is efficiently removing motion artifacts

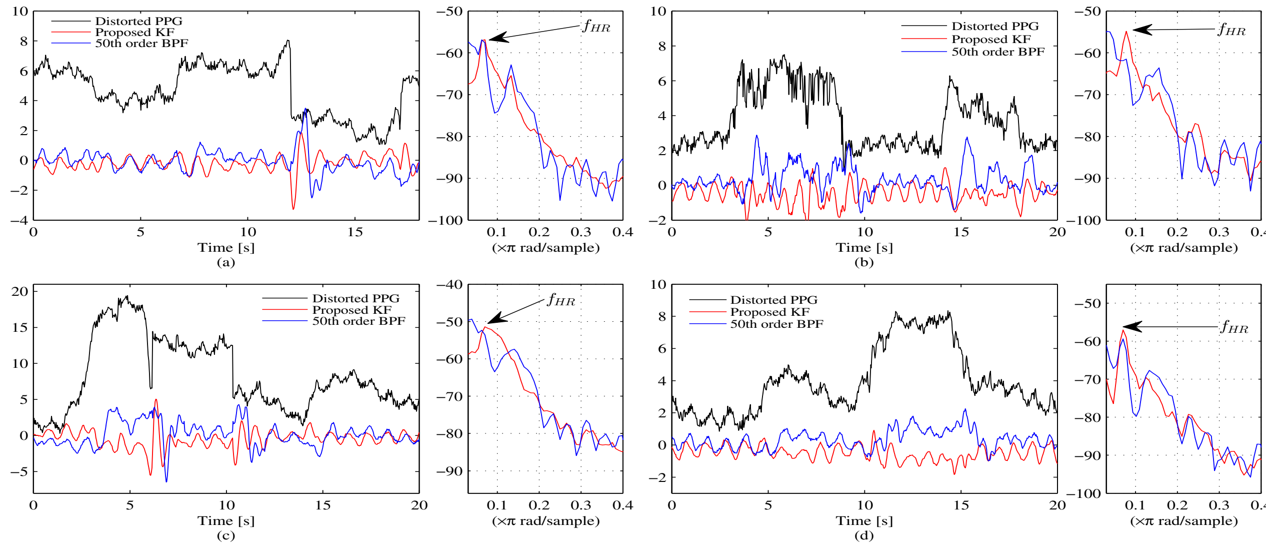

Fig. 5. Reduction results of motion artifact due to head and body motion : the prosed KF with system parameters generated by SSM vs. conventional 50th order BPF (a) roll motion of head, (b) yaw motion of head, (c) pitch motion of head, (d) scale and translation motion of head & body

rather then the BPF with high filter order. We summarized quantitative values of estimated f_{HR} to compare reference heart rates(bpm) as ground truth in Table I.

TABLE I

COMPARISON OF f_{HR}, BPM, REFERENCE IN FIG. 5

	Fig. 5 (a) roll		Fig. 5 (b) yaw		Fig. 5 (c) pitch		Fig. 5 (d) scale	
Filter	KF	BPF	KF	BPF	KF	BPF	KF	BPF
f_{HR}	1.054	0.937	1.171	0.468	1.053	0.585	1.055	1.055
bpm	63.27	56.25	70.31	28.12	63.18	25.15	63.31	63.31
Reference	63		70		63		64	

Through two aspect of time and frequency analysis, the proposed approach has advantages to filter motion artifacts cause by head & body rather than the conventional BPF filter. One of the merits of this method is the fact that in advance the system can reflect the special properties of observable $\tilde{y}_t$ using maximum likelihood or subspace identification at the stage of modeling[3]. Motion artifacts can be also removed in distorted PPG, since not only motion artifact components are excluded in modeling stage in (6) but these are regarded as ε_t error term in the filtering state.

Heart rate variability(HRV) analysis which is an important analysis method of PPG signals totally depends on the quality of pulse shape in time domain. If distorted PPG signals are processed to RR signals as peak-to-peak of pulse component, all HRV parameters obtained from both time and frequency domain analysis are not properly generated. Therefore, it is important for a filter to recover original pulse's shape of the PPG as closely as possible in HRV analysis.

The quasi periodic component f_{HR} is varying over time according to individual's physiological and psychological condition. The performance consistency between high f_{HR} and low f_{HR} not satisfied while the BPF is applied to motion artifacts, so that it does not adaptively cover all available bandwidth of PPG because of fixed bandwidth of the BPF. On the other hand, the proposed KF fairly tracks time varying f_{HR} and corrects errors by motion artifacts due to various head & body motions. The design of a filter trades off complexity and performance. The complexity of the proposed KF only using 2nd order parameters is very low, while the BPF uses high 50th filter order.

REFERENCES

[1] "What's my heart rate - measure your heart rate by just looking at your screen." [Online]. Available: http://facion.net/WhatsMyHeartRate/

[2] R. Cooper, J. Selb, L. Gagnon, D. Phillip, H. W. Schytz, H. K. Iversen, M. Ashina, and D. A. Boas, "A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy," *Brain Imaging Methods*, vol. 6, p. 147, 2012.

[3] R. A. Davis, G. Rodriguez-Yam, and W. Dunsmuir, "Estimation for state-space models; an approximate likelihood approach," *Draft Paper, Colorado State University*, 2003.

[4] J. T. Jalles, "Structural time series models and the kalman filter: a concise review," *FEUNL Work Pap*, vol. 541, 2009.

[5] J.-W. Lee, "Design of kalman filter to estimate heart rate variability from PPG signal for mobile healthcare," *Journal of information and communication convergence engineering*, vol. 8, pp. 201–204, 2010.

[6] V. P. Oikonomou, A. T. Tzallas, S. Konitsiotis, D. G. Tsalikakis, and D. I. Fotiadis, "The use of kalman filter in biomedical signal processing."

[7] M. Poh, D. McDuff, and R. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," 2010.

[8] K. T. Sweeney, T. E. Ward, and S. F. McLoone, "Artifact removal in physiological signals: Practices and possibilities," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 488–500, May 2012.

[9] C. Takano and Y. Ohta, "Heart rate measurement based on a time-lapse image," *Medical Engineering & Physics*, vol. 29, no. 8, pp. 853–857, Oct. 2007.

[10] P. Trnka, "Subspace identification methods," Ph.D. dissertation, Ph. D. dissertation, Czech Technical University in Prague, 2007.

[11] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Optics express*, vol. 16, no. 26, pp. 21 434–21 445, Dec. 2008. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2717852/

[12] J. Zheng and S. Hu, "The preliminary investigation of imaging photoplethysmographic system," *Journal of Physics: Conference Series*, vol. 85, p. 012031, Oct. 2007.