Curve Approximation Models based on Statistical Distribution with Application to Photoplethysmography (PPG) Signal

Remya Raj¹, J.Selvakumar², Vivek Maik³
¹,²,³ Department of Electronics and Communication, SRM Institute of Science and Technology, Chennai, India

e-mail: ¹ remyatraj@gmail.com,² selvakuj@srmist.edu.in,³ vivekmad@srmist.edu.in

Abstract. Photoplethysmography (PPG) signal captures blood volume change in the arteries that carry blood. PPG signals are often used to check the cardiovascular health in patients. Health care automation have made more importance into study of PPG signal and its automatic prognosis and diagnosis. In this paper we aim to achieve motion artifact reduction using low rank minimization and PPG signal representation in mathematical form using statistical distribution models. The proposed approach has been tested for Gaussian, Bifurcated Gaussian, Exponentially broadened Gaussian and lognormal distributions. The accuracy of each distribution in modeling the PPG signal was also studied.

Keywords: Brain Computer Interface, Bi-Conjugate Gradient, Electroencephalography, Online Recursive Independent Component Analysis, Recursive Least Square, Wavelet decomposition

1. Introduction

Photoplethysmography (PPG) is a vital health care monitoring device used for monitoring the arterial blood oxygen saturation and blood volume change [1,2]. Usually the volume of blood through the artery can vary depending on the arterial contraction and expansion. PPG is one of the vital parameter that gives the cardiovascular health of a person. The PPG signal is detected by either transmittance or reflectance method. The PPG signal consist of two or more peaks as shown in figure.1

Fig. 1. Original PPG signal and its relation between incident and reflected wave to PPG signal
The first peak is the main wave or main peak formed by blood volume transmitted from left ventricle of heart to finger. The rest of the peaks are called reflected waves or reflected peaks formed by the blood flows to aorta due to resistance difference in the lower extremities [3-6]. Forward and reflected wave have equal importance for determining cardiovascular health. However, this paper focuses on Arterial Stiffness disorder and it can be obtained by studying reflected wave of the PPG. In older person and those who are suffering from high arterial stiffness, reflected wave returns during early systole and it returns during late systole in younger person.

Thus the study of PPG signals becomes vital for detecting and diagnosing the arterial stiffness condition. This paper proposed an automated approach to mathematically model the PPG signal using statistical distribution function. The novelty in the proposed approach is that the mathematical representation of PPG signal makes it quantitatable, thereby computational analysis can be carried out. PPG wave decomposition separates the PPG into forward and reflected wave. The modeling of the PPG wave using statistical distribution models has been adopted previously with good performances [7]. In most of the works, Gaussian forms a universal distribution function base. Rayleigh distribution provides another useful alternative to model PPG signal. The adaptive combination of Gaussian and Rayleigh can also be seen [8]. A mixture of Gaussian have also been used for fitting PPG signal for carotid and radial artery pressure waveform [9,10]. The lognormal basis distribution performed better than more generalized and global Gaussian functions [11]. Also adaptive hybrid algorithms which work with combination of two or more distribution functions are more effective than a single distribution function [12,13].

In this work we introduced a method of motion artifact reduction using low rank minimization algorithm and a comparative analysis of four different distribution functions namely Gaussian, bifurcated Gaussian, Exponentially broadened Gaussian and lognormal distribution has been analyzed. The main contribution of the proposed approach is the error estimation for each of these four distributions. The errors computed help us to identify the distribution which succeeds in a hybrid adaptive environment which forms base of the forthcoming research.

2. Materials And Methods

PPG Datasets

A cohort of 120 subjects from the local area of southern Tamil Nadu was recruited. First group involved, Sixty-eight stable non-smoking subjects between the ages of 16 and 35 were included (36 men and 32 women) without any known hypertension, diabetes mellitus and no medications affecting the cardiovascular system. Second group involved a total of fifty-two subjects (32 males and 20 females) between the ages of 40 and 82 with hypertension and no cardiovascular medication. The signal sampling rate was taken as 125 samples per second. Subjects were required to be in a supine position for 10 minutes before fingertip signal acquisition. HRM-2115E, which functions in the transmission mode, is the optical sensor used here. For the analysis, the first 10 pulses were taken from each subject and the mean value of the parameters were taken into account. A written consent form was given to each participant to participate in the study, which was approved by the SRM Medical College and Research Center Ethical Committee, Chennai, Tamil Nadu.

Low Rank Minimization based Artifact removal algorithm

Using the optimized low rank minimization we can change the complex signal to minimum directions. The PPG signals are reduced to minimum dimensions because it has less number of rank. This method offers constraints on the envelope and any deformities of the signal. The method can be experimentally proven to reduce the dimension and make it simple without losing the information. The low rank optimization algorithm helps to reduce the PPG signal with lesser dimensions. The MA contaminated PPG signal can be written as

\[ 3(Y) = 3(\Phi)3(X) + V \]  

(1)
where Y is the signal with noise Φ. The aim is to estimate Φ and X from Y and it is illustrated as an ill-posed problem. The ill posed problem states that the function that are available is less than functions not available. On of the advantage of the study is that the algorithm works good for estimation motion artifact free signal. \( \mathcal{I} \) represents the algorithm in frequency domain where the convolution operation in time domain can be converted to multiplication operation, for the simplest way of separation of X and Y. This is more easier because the motion artifacts has higher frequencies. The input PPG signal Y can be reduced to fewer dimension by reducing the number of rows, hence this can reduce the rank of the matrix also.

The ill posed problem can be solved by reducing the redundancy that relates to the rank. The number of examinations is greater than or equal to the degree of freedom. The degree of freedom is given by \( SP^+Q+T \) that denotes number of rows, column and rank of the matrix. The low rank matrix method used here is Augmented lagrangian method.

The optimization algorithm used here is SVD and it is expressed as

\[
Q = U\Sigma V^T
\]  

\[Q = \text{non-singular values and U, V are the singular matrix.} \]

When the signal consists of low rank and less observation, the signal is decomposed into \( N \) atoms and all the atoms form an atomic set B. The atomic set B formed by \( N \) atoms is

\[
B = \{ \pm b_1, \pm b_2, \pm b_3, ..., \pm b_N \}
\]

The SVD applied to atomic set B is given by

\[
B = \{ uv^T : \|u\|_2 = 1, \|v\|_2 = 1, u \in \mathbb{R}^M, v \in \mathbb{R}^N \}
\]

The convex optimization for motion artifact removal for the atomic set B is given as

\[
\min \|X\|_*, \text{subject to } Y = \Phi X
\]

The low rank optimization problem with SVD is given by

\[
B_{k+1} = UQ \frac{\lambda}{\mu_k} V^T
\]

\[
H_{k+1} = Q \frac{\lambda}{\mu_k} \left[ D - B_{k+1} + H_{k+1} \right]
\]

where \( \|H\|_F \) represent the Frobenius norm and H is the motion artifact. The low rank optimization problem with SVD is given by

\[
\min_{B,H} \|H\|_F \text{ s.t. rank}(B) \leq r \& D = B + H
\]

The iterative optimization step introduced in this work is given by

\[
\min(Y - \Phi X) = \|B_{k+1}\| + \lambda\|H_k\| + \frac{1}{\mu_k}\|Y_{k+1}\|_F^2
\]

\[
A_{k+1} = A_k - \gamma \nabla q
\]

This method varies from other methods that uses single rank matrix optimization. The circular form of MA corrupted PPG is given as

\[
Y = \begin{bmatrix}
    y_1 & y_2 & \cdots & y_k \\
    y_{2} & y_{3} & \cdots & y_{k+1} \\
    \vdots & \vdots & \ddots & \vdots \\
    y_{L} & y_{L+1} & \cdots & y_{M}
\end{bmatrix}
\]  

This method eliminates the noises in PPG signal without disturbing its fundamental and harmonics frequencies. Figure 2. Illustrates the motion artifact eliminated PPG after low rank matrix minimization.
Statistical Modeling of Distribution

In this section we will see how any given PPG signal can be represented using statistical distribution function. The statistical modeling of PPG signals were applied to these database signals. The PPG signal is first sampled and these sample values are used to derive the distribution model function in an iterative manner. The iterative optimization converges when the error between the function values and sample values becomes close to zero. In this paper we modeled the PPG with four distribution function as shown in Table I.

Table I: Distribution Functions And Its Parameters

| Distribution Function | Formula | Parameters computed in iteration |
|------------------------|---------|---------------------------------|
| Gaussian               | $f_k g(x) = H \cdot \exp \left( -\frac{(x - \mu)^2}{2\sigma^2} \right)$ | Center position($\mu$), Width($\sigma$), Height($H$) |
| Bifurcated Gaussian    | $f_k b(x) = H \cdot \exp \left( -\frac{(x - \mu)^2}{2\sigma(x - \mu)^2} \right)$ | Center position($\mu$), Width($\sigma$), Height($H$) |
| Exponentally Broadened Gaussian | $f_k e(x) = \text{Exp Broaden}(f g', \tau)$ | Time constant($\tau$), Gaussian function($g$) |
| Lognormal Distribution function | $f_k l(x) = H \cdot \exp \left( -\frac{(\ln(x) - \mu)^2}{2\sigma^2} \right)$ | Center position($\mu$), Width($\sigma$) |
3. Results And Discussion

Fig.3 to fig.6 represents the estimated distribution function and plots of residual errors. The residual error is calculated as the average difference between the sample points of the actual PPG and the estimated PPG through iterative optimization. Table II represents the residual errors of all distributions. The main goal of this method was to calculate the ideal mixture of different statistical distribution functions to model PPG signal. From fig.3 to fig.5 we can see that all the Gaussian and lognormal distributions fitted PPG perfectly. In the proposed method we found that the accuracy of lognormal distribution has improved to 0.3 percent with comparison with all other methods [10].

The position, width, height and area of all the distributions are given in distribution figures. These parameters are iteratively optimized according to the curve fitting with least square. As the baseline of signals is varying from peak to peak, we used linear baseline subtraction to correct the signal to baseline.

The accuracy of PPG waveform fitting using different distribution functions were calculated by comparing the curve fitted function \( f(n) \) with given original PPG \( F(n) \). The Mean Absolute Error (MAE) is given by

\[
MAE = \frac{\sum_{n=1}^{N} |F(n) - f(n)|}{N}
\]  

(12)

Where \( N \) is the total number of sample points for a complete PPG wave. The residual error is given as

\[
Residual\ Error = F(n) - f(n)
\]  

(13)

The value of residual error depends on the parameters computed during each iteration.

![Fig. 3: Gaussian function and the corresponding residual error](image_url)
Fig. 4: Bifurcated Gaussian function and the corresponding residual error

Fig. 5: Exponentially broadened Gaussian function and the corresponding residual error

Fig. 6: Lognormal function and the corresponding residual error
Table II: Residual Errors For Estimated Ppg Signal

| DISTRIBUTION FUNCTION     | RESIDUAL ERROR(IN PERCENTAGE) |
|---------------------------|-------------------------------|
| GAUSSIAN                  | 1.38                          |
| BIFURCATED GAUSSIAN       | 1.29                          |
| EXPONENTIALLY BROADENED GAUSSIAN | 1.34                      |
| LOGNORMAL                 | 1.07                          |

4. Conclusion
In this paper, we have studied and compared how to model a PPG signal mathematically with best optimization-error trade off. This work will help us to computationally analyze the PPG signal for various medical conditions such as arterial stiffness. Three types of Gaussian functions are used for modeling and out of that bifurcated Gaussian gives less residual error when compared to fixed width Gaussian and exponentially broadened Gaussian. The better results obtained in bifurcated Gaussian is due to the variation in width of the wave with respect to the center position for each iteration as in PPG waveform. From the shape of PPG it is obvious that it has long tail features. Thus lognormal function works better than Gaussian function for the modeling of PPG pulse waveform as it has occurrences far from the head and center position. Also the proposed algorithm significantly reduces the cost of the PPG device in market expanding reach to more people. The drawback of the existing methodology which includes lack of adaptivity will be analyzed in future work along with use of multi-distribution function mode. In future work, the reflected wave parameters of the modelled PPG will also be studied for different age groups.

5. Acknowledgment
This work was supported by the grant from Visvesvaraya Ph.D scheme, Meity (Ministry of Electronics and Information Technology), Government of India with unique identity number <587>. The authors also thank SRM Medical College Hospital and Research Centre for providing the Ethical Clearance (1098/IEC/2017) and helping in acquiring the PPG signal

References
[1] J. Allen., “Photoplethysmography and its application in clinical physiological measurement,” Physiol Meas., vol. 28, pp. 1–39, 2007.
[2] Elgendi.M, “On the Analysis of Fingertip Photoplethysmogram Signals,” Curr Cardiol Rev, vol. 8, pp. 14–25, 2012.
[3] P. J. Chowienczyk, R. P. Kelly, and H. MacCallum, “Photoplethysmographic assessment of pulse wave reflection: Blunted response to endothelium-dependent beta2-adrenergic vasodilation in type II diabetes,” Journal of the American College of Cardiology, vol. 34, no. 7, pp. 2007–2014, 1999.
[4] B. E. Westerhof, I. G., N. W., and J. M. Karemaker, “Quantification of wave reflection in the human aorta from pressure alone: a proof of principle.,” Hypertension, vol. 48, no. 4, pp. 595–601, 2006.
[5] N. Stergiopulos., Y. Tardy., and J. J. Meister, “Nonlinear separation of forward and backward running waves in elastic conduits.,” J Biomech, vol. 26, no. 2, pp. 201–209, 1993.
[6] K. H. Parker and C. J. Jones, “Forward and backward running waves in the arteries: analysis using the method of characteristics.,” J Biomech, vol. 112, no. 3, pp. 322-326, 1990.
[7] U. Rubins, “Finger and ear photoplethysmogram waveform analysis by fitting with Gaussians.,” Med Biol Eng Comput., vol. 46, no. 12, pp. 1271–1276, 2008.
[8] D. Goswami, K. Chaudhuri, and J. Mukherjee, “A New Two-Pulse Synthesis Model for Digital Volume Pulse Signal Analysis,” Cardiovascular Engineering, vol. 10, no. 3, pp. 109–117, 2010.

[9] L. Wang, L. Xu, D. Zhao, M. Q. Meng, and K. Wang, “Multi-Gaussian fitting for pulse waveform using Weighted Least Squares and multicriteria decision making method,” Comput Biol Med., vol. 43, no. 11, pp. 1161–1172, 2013.

[10] Liu, Chengyu, et al. “Modeling carotid and radial artery pulse pressure waveforms by curve fitting with Gaussian functions.” Biomedical Signal Processing and Control, vol. 8, no. 5, 2013, pp. 449–454.

[11] L. Yun, . “A New Modeling Method of Photoplethysmography Signal Based on Lognormal Basis.” International Conference on Internet and Distributed Computing Systems, 2016, pp. 12–21.

[12] H. Sheng, . “Evaluation of Decomposition Analysis on Multi-Models for Digital Volume Pulse Signal.” World Congress on Medical Physics and Biomedical Engineering, 2015, pp. 1731–1734.

[13] S .Andrius, . “Modeling of the photoplethysmogram during atrial fibrillation.” Computers in Biology and Medicine, vol. 81, no. 1, 2016, pp. 130–138