

Photoplethysmography signal analysis to assess obesity, age group and hypertension

Hany Ferdinando, Matti Huotari, and Teemu Myllylä

Abstract—Photoplethysmography (PPG) provides a simple, convenient and noninvasive method to assess pulse oximetry. Several attempts have been made to use PPG also to estimate blood pressure and arterial stiffness. This paper attempts to assess obesity classes, age group, and hypertension classes using PPG measured from the finger. One set of features was derived from the normalized pulse width of PPG and the other from original PPG. The features were calculated based on the pulse decomposition analysis using five lognormal functions and the up-slope of the PPG pulse. Using kNN and SVM as classifiers, the results were validated using leave-one-out validation. Performances of both features sets have no significant difference, and the kNN outperformed the SVM. The best accuracies are 93%, 88%, and 92% for obesity (5 classes), age group (7 classes), and hypertension (4 classes) respectively. These three assessment targets have a strong relationship with arterial stiffness, therefore it also leads to a study about arterial stiffness using PPG. Width normalization to 1 second might affect some features points based on pulse decomposition analysis. This study also found that the up-slope analysis might give good indices when width normalization was employed. However, these findings still require more experiments to gain conclusions that are more comprehensive.

I. INTRODUCTION

Photoplethysmography (PPG) signals represent blood volumetric changes in peripheral circulations measured using a simple and low-cost optical method. It means cardiac activity drives this signal mainly. Therefore, for instance, PPG and ECG signals are always synchronized to each other. Some common locations to measure PPG signals are finger and earlobes. Each location has different peripheral circulation characteristic and properties, thus it results in the different shape of PPG signals. This shape also varies among different subjects with various physiological conditions, making the PPG signal a good technique to provide indices to assess different kinds of characteristic from subjects.

PPG is noninvasive, inexpensive, convenient, and an easy to use tool for measuring oxygen saturation. Further, it can provide other physiological signals such as heart rate variability (HRV), heart rate (HR), and respiration rate. However, recently more and more attention has been given to PPG for other possible applications.

Research supported by Academy of Finland Grants 314502 & 318347

H. Ferdinando is with the Research Unit of Medical Imaging, Physics and Technology, University of Oulu, Finland and Department of Electrical Engineering, Petra Christian University, Indonesia (phone: +358 50 564 3469; e-mail: hany.ferdinando@oulu.fi).

M. Huotari is with Biomimetics and Intelligent Systems Group, University of Oulu, Oulu, Finland (e-mail: matti.huotari@oulu.fi).

T. Myllylä is with the Research Unit of Medical Imaging, Physics and Technology and with the Optoelectronics and Measurement Techniques Unit, University of Oulu, Finland (e-mail: teemu.myllyla@oulu.fi).

Several attempts have been made to use PPG signals to estimate blood pressure using the PPG signals [1]–[5] and arterial stiffness [6]–[8]. Although they are not matured yet, these facts indicate that PPG signals carry comprehensive information on the human body. Elgendi [9] published an article about deriving features from PPG signal and its first and second derivative waveforms. It includes systolic amplitude, pulse width, pulse area, peak-to-peak interval, augmentation index, large artery stiffness index, etc. It shows that PPG contains useful information for further exploration. Moreover, obesity seems to have strong relationship with arterial stiffness. Tarnoki et al. [10] highlighted the complexity of the association between obesity and arterial stiffness. Momin et al. [11] concluded that increased BMI is a positive factor against the development of arterial stiffness in Chinese rural-dwelling adults with primary hypertension undergoing antihypertension treatments, after adjusting for confounding factors. Nordstrand et al. [12] observed general and abdominal obesity, and they were associated with arterial stiffness in morbid women.

The wall of large conduit arteries thicken and lose elasticity over time [13]. Arterial walls, assessed via the pulse wave velocity (PWV) and augmentation index, stiffen with age and they have a mostly linear relationship [14]. Using cardio-ankle vascular index, Wen et al. [15] found that a growth curve provides better explanation about the age-related progression of arterial stiffness than a straight line. Jayasree et al. [16] showed that the mean pulse shape of PPG varies among young, adult, and old people. Furthermore, analyzing the associations of blood pressure with 12 different presentations of cardiovascular disease using linked electronic health records from 1997 to 2010, Rapsomaniki et al. [17] found the lowest risk for cardiovascular disease was in people with a systolic blood pressure of 90-114 mmHg and diastolic blood pressure of 60-74 mm Hg in each age group. Pickering et al. [18] classified them as normal blood pressures. Consequently, PPG can be potentially used to assess obesity, age group, and hypertension.

II. MATERIAL AND METHODS

Liang et al. [5] provide 657 short-recorded data segments of PPG along with sex, age, weight, height, BMI, and systolic and diastolic blood pressure from 219 subjects of 20-89 years old. Sampled at 1 kHz, the length of each data segment is 2.1 seconds. Thus, it contains 2-3 PPG pulse.

A. PPG Signal Processing

The raw signals contain many spikes that might come from ambient light at the photodetector. A 2nd-order of Chebyshev low-pass filter was used to remove high-frequency noise. However, the baseline wandering signals were still visible. The baseline-wandering signal was estimated using cubic spline interpolation based on the start of each PPG pulse [19]. The start of the pulse was detected based on the five-point digital differentiator [20]. The estimated baseline wandering was subtracted from the PPG to get a cleaned PPG signal ready for the further process.

B. Feature Extraction and Selection

For all feature calculation, the individual PPG pulse were normalized to [0,1]. One set of features were calculated based on normalized PPG pulse width to 1 second to avoid HR influence (feature set #1) and the other based on the original width (feature set #2). Both sets used the same feature extraction process. The first sub-set was based on the pulses wave decomposition; decomposing PPG pulse into 5 lognormal pulse, see (1), as suggested in [21]. The main idea lies on the theory of reflected pulses in the arteries. Therefore, a PPG pulse is a summation of several individual pulse. The pulse wave decomposition method aims to extract these pulses.

$$f(A, c, \sigma) = \sum_{n=1}^5 \frac{A_n}{x\sigma_n\sqrt{2\pi}} e^{-\frac{\left(\ln \frac{x}{c_n}\right)^2}{2\sigma_n^2}} \quad (1)$$

It used curve-fitting method using objective function

$$\min |f(A, c, \sigma) - PPG| \quad (2)$$

subject to

$$A_1 > A_3, A_3 > A_2, A_3 > A_4, A_3 > A_5 \quad (3a)$$

$$c_1 < c_2 < c_3 < c_4 < c_5 \quad (3b)$$

$$\sigma_1 > \sigma_2, \sigma_1 > \sigma_3, \sigma_1 > \sigma_4, \sigma_1 > \sigma_5 \quad (3c)$$

The Matlab implementation employed `fmincon` function because this function enables some constraints during optimization. All fifteen parameters of the decomposition results cannot but used directly for classification. Therefore, several features were calculated based on them:

- Peaks of lognormal function (5)
- Times at peaks of lognormal function (5)
- Ratio between peaks of lognormal function (10)
- Difference between peaks of lognormal function (10)
- Ratio between times at peaks of lognormal function (10)
- Difference between times at peaks of lognormal function (10)
- Amplitude (A) of lognormal function (5)
- Ratio between Amplitude (A) of lognormal function (10)

- Difference between Amplitude (A) of lognormal function (10)
- Width (σ) of lognormal function (5)
- Ratio between width (σ) of lognormal function (10)
- Difference between width (σ) of lognormal function (10)
- Center (c) of lognormal function (5)
- Ratio between center (c) of lognormal function (10)
- Difference between center (c) of lognormal function (10)
- Area under curve of lognormal function (5)
- $t_1 - 2t_2$ from Huotari et al. [21] (1)

The first up-slope of young people is steeper than the one of the elderly is. The arterial stiffness also influences pulse wave velocity and consequently blood pressure. Thus, the up-slope of the pulse might provide useful information.

The features of the second sub-set are based on the morphology of PPG pulse:

- Rise time of the PPG pulse (1)
- Statistical distribution of the first up slope gradient (14)
- Index based on second derivative proposed by Takazawa et al. [22] (1)

The whole process collected 147 features, but not all features were useful. Therefore, feature selection using the Sequential Forward Floating Selection (SFFS) was employed to select only features with high discriminant values for each assessment. Using BMI data provided by the database, six classes of obesity were made based on [23]: underweight, normal, overweight, obesity I, and obesity II. Unfortunately, obesity II was not found among the data. Age data was grouped as follows: less than 30, 30-39, 40-49, 50-59, 60-69, 70-80 and over 80 years. The database put four blood pressure classification: normal, prehypertension, hypertension I, and hypertension II.

C. Classifiers and Validation

The k-nearest neighbor (kNN) was used as a classifier for these assessments. It provides a simple but intuitive system with fast computation and requires no training. The experiments also employed the support vector machine (SVM). The results were validated using leave-one-out cross-validation. All methods were implemented in Matlab 2018b and ran using Intel(R) Core(TM) i5-7500 2.70 GHz and 8 GB RAM.

III. RESULTS

The average time required in pulse decomposition analysis using curve fitting method was 0.08 seconds. It means the method can be used almost in real-time analysis. The worst elapsed time was 0.9 seconds while the best one 0.03 seconds. The elapsed time variation depends on the problem and the initial values at the beginning of the curve fitting iteration. Obviously, this elapsed time can be influenced by many factors such as CPU speed, memory, and tasks running by the computer.

Fig. 1 shows the results of the decomposition analysis from both young and elderly people on normalized width. The pulses appeared in order as desired. The correlation

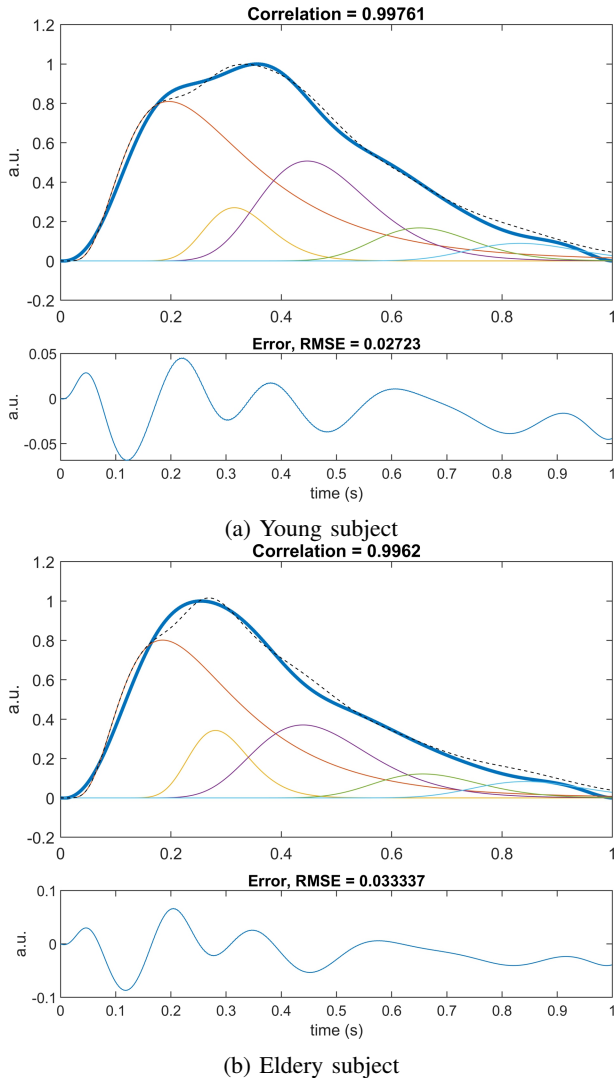


Fig. 1: Example of pulse decomposition results from young (a) and elderly (b) people. Both figures display the original and approximated signals in solid blue and dotted lines respectively. The pulses appear from pulse #1 to #5 from left to right.

coefficients also indicate that the estimated parameters could fit the pulse almost perfect. There is no significant difference between results from normalized and original width. Table I and II show the results from the whole experiments for both feature sets using kNN and SVM respectively. All performances were above the corresponding chance level for each case. With the lowest performance at 79%, both feature sets offer suitable indices for these assessments. The performances from feature set #2 looked slightly better than the one from feature set #1 using both classifiers, but the differences were not statistically significant.

IV. DISCUSSIONS

The SFFS selected different features from feature set #1 and #2. All selected features from feature set #1 were related to up-slope characteristics, while features set #2

TABLE I: ACCURACY OF TWO SETS OF FEATURES FROM THE WHOLE ASSESSMENTS USING KNN

Assessment	feature set #1	feature set #2
Obesity (5 classes)	91.9%	93.4%
Age group (7 classes)	88.2%	88.1%
Hypertension (4 classes)	90.8%	92.2%

TABLE II: ACCURACY OF TWO SETS OF FEATURES FROM THE WHOLE ASSESSMENTS USING SVM

Assessment	feature set #1	feature set #2
Obesity (5 classes)	89.9%	90.5%
Age group (7 classes)	83.7%	83.0%
Hypertension (4 classes)	82.1%	87.4%

also took features based on pulse decomposition analysis. It looks like the width normalization weakens the features based on the pulse decomposition analysis. Most probably, width normalization to 1 second transforms some useful points to other positions and makes them less useful. It is also possible that the up-slope analysis provides good indices with width normalization. This process enhances the characteristic of the up-slope and makes the indices stronger than using the original width. However, these two analyses still require further exploration using various databases for more comprehensive conclusions.

With a chance level of 20% for obesity case, the selected features offer good performances from both feature sets using kNN and SVM. The confusion matrices revealed that both features sets make the classifiers struggled to classify obesity I. It was found that less than 2% of the samples belong to obesity I. However, it seemed this proportion had nothing to do with the accuracies. Underweight, for example, was about 9% of the whole sample, but it achieved the second highest result. It was even higher than overweight that represents 27% of the whole samples.

Age group assessment has a chance level at 14%, and the overall performance for both feature sets were good, but they were slightly below the obesity assessment. The confusion matrices showed that both feature sets make the classifiers struggled to classify age older than 80 years old. The confusion matrices also confirmed that the proportion of the class had no influence on the accuracy individually. The highest performance was achieved by the 30-39 years old class, which was less than 3%. Hypertension cases have a chance level at 25% and the performances are good for both sets. Based on the confusion matrices, the most struggling class was the hypertension I. We also did not find any correlation between the class proportion and the individual accuracy from the confusion matrices of both classifiers.

This study found that a single pulse PPG can be used to assess obesity, age group, and hypertension with accuracies up to 93%. These three assessment targets have a strong relationship with arterial stiffness. Thus, it also confirms the previous finding that PPG signals can be used to assess arterial stiffness [22]. With average processing time about 80

ms for each pulse, pulse decomposition analysis can be used in a real-time application. Future works consist of using PPG signals measured from different parts of the body to assess the same target and exploring pulse decomposition analysis as a marker in arterial stiffness study.

REFERENCES

- [1] I. C. Jeong, J. I. Ko, S. O. Hwang, and H. R. Yoon, "A new method to estimate arterial blood pressure using photoplethysmographic signal," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug 2006, pp. 4667–4670.
- [2] X. Xing and M. Sun, "Optical blood pressure estimation with photoplethysmography and fft-based neural networks," *Biomedical Optics Express*, vol. 7, no. 8, pp. 3007–3020, 2016.
- [3] S. S. Mousavi, M. Firouzmand, M. Charmi, M. Hemmati, M. Moghadam, and Y. Ghorbani, "Blood pressure estimation from appropriate and inappropriate ppg signals using a whole-based method," *Biomedical Signal Processing and Control*, vol. 47, pp. 196 – 206, 2019.
- [4] S. Song, J. Cho, H. Oh, J. Lee, and I. Kim, "Estimation of blood pressure using photoplethysmography on the wrist," in *2009 36th Annual Computers in Cardiology Conference (CinC)*, Sep. 2009, pp. 741–744.
- [5] Y. Liang, Z. Chen, G. Liu, and M. Elgendi, "A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in china," *Scientific Data*, vol. 5, no. 180020, pp. 1–7, 2019.
- [6] E. v. Wowern, G. Östling, P. M. Nilsson, and P. Olofsson, "Digital photoplethysmography for assessment of arterial stiffness: Repeatability and comparison with applanation tonometry," *Plos ONE*, vol. 10, no. 8, pp. 1–15, 2015.
- [7] K. Pilt, K. Meigas, R. Ferenets, K. Temitski, and M. Viigimaa, "Photoplethysmographic signal waveform index for detection of increased arterial stiffness," *Physiol. Meas.*, vol. 35, no. 10, pp. 2027–2036, 2015.
- [8] C. Wei, "Developing an effective arterial stiffness monitoring system using the spring constant method and photoplethysmography," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 1, pp. 151–154, Jan 2013.
- [9] M. Elgendi, "On the analysis of fingertip photoplethysmogram signals," *Current Cardiology Reviews*, vol. 8, pp. 14–25, 2012.
- [10] A. D. Tarnoki, D. L. Tarnoki, L. H. Bogl, E. Medda, C. Fagnani, L. Nisticò, M. A. Stazi, S. Brescianini, P. Lucatelli, E. Boatta, C. Zini, F. Fanelli, C. Baracchini, G. Meneghetti, J. Osztoivits, G. Jermendy, R. G. Kiss, I. Preda, K. Karlinger, A. Lannert, A. A. Molnar, L. Littvay, Z. Garami, V. Berczi, G. Pucci, G. Baffy, G. Schillaci, and K. H. Pietiläinen, "Association of body mass index with arterial stiffness and blood pressure components: A twin study," *Atherosclerosis*, vol. 229, no. 2, pp. 388 – 395, 2013.
- [11] M. Momin, J. Li, Y. Zhang, F. Fan, X. Xu, X. Xu, B. Wang, W. Yang, L. Xie, and Y. Huo, "Body mass index is inversely associated with arterial stiffness in chinese adults with primary hypertension: results from the china stroke primary prevention trial (csppt)," *Clinical and Experimental Hypertension*, vol. 39, no. 5, pp. 394–401, 2017.
- [12] N. Nordstrand, E. Gjevestad, K. Dinh, D. Hofso, J. Røislien, E. Saltvedt, I. Os, and J. Hjelmeæth, "The relationship between various measures of obesity and arterial stiffness in morbidly obese patients," *BMC Cardiovascular Disorders*, vol. 11, pp. 1–7, 2011.
- [13] Z. Sun, "Aging, arterial stiffness, and hypertension," *Hypertension*, vol. 65, no. 2, pp. 252–256, 2015.
- [14] H. Lee and B. Oh, "Aging and arterial stiffness," *Circulation Journal*, vol. 74, no. 1, pp. 2257–2262, 2010.
- [15] W. Wen, R. Luo, X. Tang, L. Tang, H. X. Huang, X. Wen, S. Hu, and B. Peng, "Age-related progression of arterial stiffness and its elevated positive association with blood pressure in healthy people," *Atherosclerosis*, vol. 238, no. 1, pp. 147 – 152, 2015.
- [16] T. S. V. Jayasree and P. Radhakrishnan, "Non-invasive studies on age related parameters using a blood volume pulse sensor," *Measurement Science Review*, vol. 8, no. 4, pp. 82–86, 2008.
- [17] E. Rapsomaniki, A. Timmis, J. George, M. Pujades-Rodriguez, A. D. Shah, S. Denaxas, I. R. White, M. J. Caulfield, J. E. Deanfield, L. Smeeth, B. Williams, A. Hingorani, and H. Hemingway, "Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 125 million people," *The Lancet*, vol. 383, no. 9932, pp. 1899 – 1911, 2014.
- [18] T. G. Pickering, J. E. Hall, L. J. Appel, B. E. Falkner, J. W. Graves, M. N. Hill, D. W. Jones, T. Kurtz, S. G. Sheps, and E. J. Rocella, "Recommendations for blood pressure measurement in humans: An aha scientific statement from the council on high blood pressure research professional and public education subcommittee," *The Journal of Clinical Hypertension*, vol. 7, no. 2, pp. 102–109, 2005. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1524-6175.2005.04377.x>
- [19] L. Yang, S. Zhang, X. Li, and Y. Yang, "Removal of pulse waveform baseline drift using cubic spline interpolation," in *2010 4th International Conference on Bioinformatics and Biomedical Engineering*, June 2010, pp. 1–3.
- [20] R. Couceiro, P. Carvalho, R. P. Paiva, J. Henriques, I. Quintal, M. Antunes, J. Muehlsteff, C. Eickholt, C. Brinkmeyer, M. Kelm, and C. Meyer, "Assessment of cardiovascular function from multi-gaussian fitting of a finger photoplethysmogram," *Physiological Measurement*, vol. 36, no. 9, pp. 1801–1825, aug 2015.
- [21] M. Huotari, A. Vehkaoja, K. Mitt, and J. Kostamovaara, "Photoplethysmography and its detailed pulse waveform analysis for arterial stiffness," *Journal of Mechanics of Materials and Structures*, vol. 44, pp. 345–362, 2011.
- [22] K. Takazawa, N. Tanaka, M. Fujita, O. Matsuoka, T. Saiki, M. Aikawa, S. Tamura, and C. Ibukiyama, "Assessment of vasoactive agents and vascular aging by the second derivative of photoplethysmogram waveform," *Hypertension*, vol. 32, no. 2, pp. 365–370, 1998.
- [23] J. Q. Purnell, *Endotext*, L. D. Groot, G. Chrousos, K. Dungan, K. Feingold, A. Grossman, J. Hershman, C. Koch, M. Korbonits, R. McLachlan, M. New, J. Purnell, R. Rebar, F. Singer, and A. Vinik, Eds. South Dartmouth (MA): MDText.com, Inc., 2018.