Combining finger and toe photoplethysmograms for the detection of atherosclerosis

Combining finger and toe photoplethysmograms for the detection of atherosclerosis

Mikko Peltokangas[1], Antti Vehkaoja[1], Matti Huotari[2], Jarmo Verho[1], Ville M. Mattila[3], Juha Röning[2], Pekka Romsi[4], Jukka Lekkala[1], and Niku Oksala[3]

[1] Department of Automation Science and Engineering, BioMediTech, Tampere University of Technology, Tampere, Finland
[2] Oulu University, Oulu, Finland
[3] Tampere University Hospital, Tampere, Finland, and University of Tampere (School of Medicine, Surgery), Tampere, Finland
[4]Oulu University Hospital, Oulu, Finland.

E-mail: mikko.peltokangas@tut.fi

October 2016

Abstract. *Objectives:* In this study, we propose and analyze a noninvasive method for detecting the atherosclerotic changes of vasculature based on the analysis photoplethysmographic (PPG) signals. *Methods:* The proposed method is called finger-toe (FT)-plot analysis that utilizes both finger and toe PPG signals. For the features extracted from the FT-plots, we implemented different linear discriminant analysis (LDA) based classifiers and analyzed seven promising ones in detail. We used the signals recorded from altogether 75 test subjects (categorized as 27 atherosclerotic patients and 48 control subjects based on ankle brachial pressure index, symptoms and disease history) in the training, and testing of the method. Besides leave one out cross validation (LOOCV), we tested the method by using training data independent signals recorded with two different PPG devices. The performance of the FT-plot is compared with other indicators related to the risk of cardiovascular diseases. *Results:* We found an average area under ROC (receiver operating characteristic) curve of 91.6%±5.6% (mean±standard deviation based on different datasets), sensitivity of 81.4%±4.9%, specificity of 87.4%±8.0%, accuracy of 86.1%±4.0%, performance of 84.5%±3.4% and positive and negative predictive values of 81.2%±7.8% and 88.8%±2.3%, respectively, for the different tested classifiers. *Conclusions:* The study shows that the FT-plot analysis could be a useful additional tool for detecting atherosclerotic changes. Our findings provide evidence for the utility of multi-channel pulse wave measurements and analysis for the detection of atherosclerosis. This may facilitate development of novel early diagnostic approaches and preventive strategies.

Keywords: Atherosclerosis, Body sensor networks, Classification, Photoplethysmography

1. Introduction

The degenerative changes of the arterial tree include stiffening of the arteries during aging, classically known as arteriosclerosis and thickening, stenosis or occlusion of the

arteries due to accumulation of cholesterol, i.e. atherosclerosis. In a clinical point of view, these conditions are considered as a continuum of degenerative changes, and they are indicators of increased risk for acute cardiovascular events, such as stroke and myocardial infarction.

Advanced atherosclerosis is relatively easy to detect by clinical examination with established detection techniques such as ankle-brachial pressure index (ABI) measurement or different imaging-modality based angiograms. However, the angiograms need expensive equipment and are not suitable for a quick screening studies. With the ABI measurement, a problem is that its sensitivity and specificity vary widely [1] especially in case of patients having early-stage atherosclerotic changes in vasculature. For these reasons, improvements to the standard procedure have been proposed e.g. by using the lowest ankle pressure in idenfication of the patient in risk [2], but there is also a need for alternative cost-effective methods to detect early-stage atherosclerosis. The analysis of arterial pulse waves (PW) may provide additional information for the vascular evaluation since the propagation of PWs depends on the properties of the arterial walls.

Commonly utilized noninvasively recorded signals for the PW analysis are collected as index finger photoplethysmogram (PPG) [3, 4, 5, 6] and radial or carotid artery tonometry [7]. Due to its ease of use, cost-effectiveness, and wide number of different applications, there has been lots of interest towards measurement and analysis methods related to the PPG signals [3]. Different kinds of indices related to the status of the vasculature have been proposed based on a single signal recorded from a certain measurement site, although combining the information obtained from several measurement points may provide a more comprehensive view on the subject's vascular status [8]. These indices include PPG based aging index (AGI) [5], amplitude ratios, such as reflection index (RI) [4] and augmentation indices (AIx) [9, 7] or time delays between the peaks of PW, such as stiffness index (SI) [4]. Besides direct PW-derived features, support vector machines utilizing different features extracted from index finger PPGs — originating both from the assumed physiological factors and purely mathematical properties of the data — have been proposed [10] for finding the subjects having a high risk for the cardiovascular events. Different kinds of PW decompositions modeling the reflections with highly non-linear basis functions have also been proposed for characterizing the vasculature [11, 6].

Even though the arteries at the upper limbs are easily accessible for noninvasive measurement, the arteries of lower limbs are often more prone to atherosclerotic changes and are thus more attractive measurement sites. Allen *et al* [12, 13] have proposed hallux PPG-based shape index in discriminating healthy subjects and the patients suffering from peripheral arterial occlusive disease. Their methods compare the patient's normalized toe-PPG PW with a normalized standard PW contour or bilateral differences between the left and right leg.

Due to wide variation in the performance of the ABI measurement [1], the authors feel that there is room for improvements in techniques and algorithms. In particular, by combining multiple signals, it should be possible to increase the validity of the results. In this study, we propose and analyze a method which we call finger-toe (FT) plot analysis which is based on the PPG signals recorded from both upper and lower limbs for detecting the subjects having atherosclerotic changes. The method compares the finger and toe PPGs at intra-subject level instead of comparing finger or toe PPGs directly between different subjects unlike the PW analysis is conventionally done. The original idea of the method is proposed in [11], but as far as the authors are aware

of, this is the first study that provides quantitative measures from the FT-plot and evaluates its performance. We concentrate on the description of the methodology and the analysis of its classification performance. We believe that combining the information obtained from multiple PPG measurements will provide an improved diagnostic tool for screening people having a potential risk for cardiovascular diseases and a more comprehensive view on the vascular status compared with traditional methods or individual PW measurement sites.

2. Materials and methods

2.1. Measurement hardware, sensor placement, and signal preprocessing

All the volunteer test subjects participating in the study were examined with the PPG probes connected into the synchronous wireless body sensor network (WBSN) presented in [14]. The WBSN utilizes time division multiple access as the mechanism for the band-width management and each data point of the sampled signals is labeled with a time stamp. The excitation wavelength of the PPG was at infrared range, being 905 nm, and the signals were quantized by using a 16-bit analog-to-digital converter (ADC) (ADS8325) with a sampling rate of 500 Hz. The slowly-varying DC-component of the PPG was high-pass filtered with a cut-off frequency of 0.15 Hz.

The PPG signals were recorded from all study subjects in supine position from the index finger and second toe. In order to show that the results are not fully dependent on the measurement hardware, a part of the study population (15 persons) was examined also with a photoplethysmograph that used phase-sensitive PPG (PSP) technique and is presented in [11]. The excitation wavelength of the PSP was 920 nm and the signals with the DC-component were quantized with a 16-bit ADC having a sampling frequency of 1000 Hz.

All the signal processing was done offline in MATLAB (version R2014b). Before FT-plot analyss, the signals were lowpass-filtered with a finite impulse response (FIR) filter having a cut-off-frequency of 10 Hz, transition band of 10 Hz–12 Hz, pass band ripple of 0.05 dB and stop band attenuation of 100 dB, as proposed in [4] and implemented in [8].

2.2. Study subjects and datasets

The clinical patient measurements were conducted in two Finnish university hospitals (Tampere and Oulu). The study subjects were divided into different groups: group A as atherosclerotic patients having abnormal ABI (ABI$<$ 0.9 or ABI$>$ 1.3) and several cardiovascular risk factors and group B as control patients having normal ABI ($0.9 \leq$ ABI ≤ 1.3). The atherosclerotic patients were not divided into two groups based on the ABI value since there were only 5 instances having ABI$>$ 1.3. In addition to the normal ABI, the control subjects had no history with the following symptoms or diagnosed diseases: cerebrovascular disease (amaurosis fugax, transient ischemic attack, ishcemic stroke), coronary artery disease (angina pectoris, myocardial infarction) or peripheral arterial disease (intermittent claudication, critical limb ischemia, acute limb ischemia). All subjects in groups A and B were at least 65-year-old.

Third group, group C, consists of younger test subjects with no cardiovascular symptoms or diagnosed disorders. Similar measurements as in the hospitals were conducted in Tampere University of Technology (Tampere, Finland) for the test

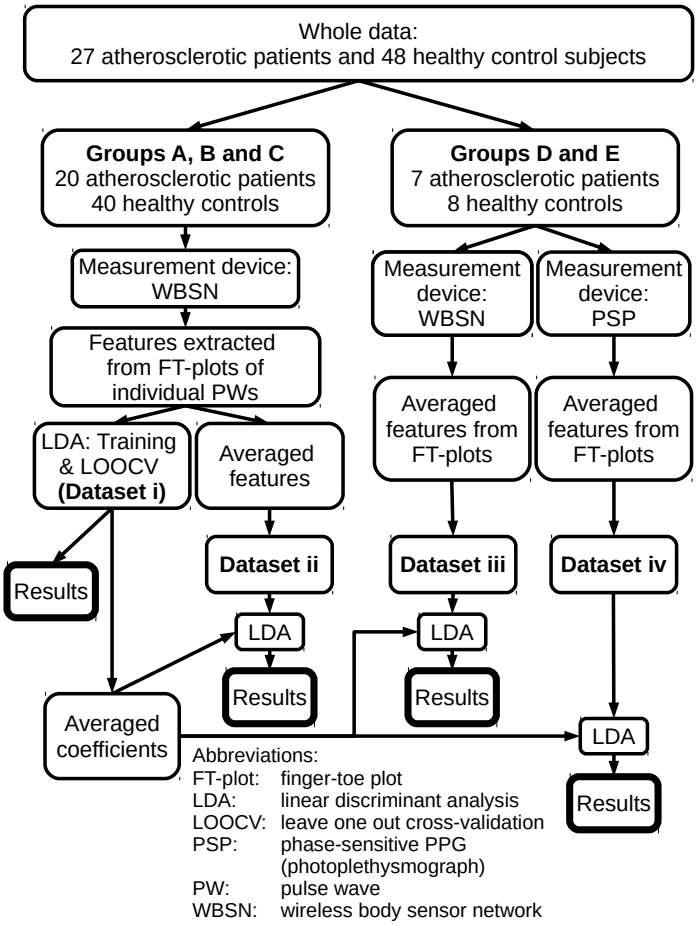


Figure 1. The whole data is divided into different study groups A–E (see Table 1) that form different datasets i–iv for different purposes. Groups A, B, and C form datasets i and ii. Dataset i consist of the features extracted from FT-plots of the individual PWs which are recorded with the WBSN. The processing of the dataset i results in LOOCV based metrics for the classifier performance as well as averaged weights utilized in the LOOCV based classifiers implemented for the other datasets. Dataset ii is the testing dataset and consists of averaged FT-plot features based on dataset i. Additional datasets iii and iv are independent of datasets i and ii (different study subjects) and they are recorded with two different devices, WBSN and PSP.

subjects in group C. Seven atherosclerotic patients and eight control subjects were examined also with the PSP device in Oulu University Hospital and they form groups D and E. The inclusion and exclusion criteria with these groups were similar as with groups A and B. A more detailed information on the study subjects is presented in Table 1 and Fig. 1 illustrates how different groups form different datasets i-iv for different purposes. A starting point is the training dataset i which is composed of the subjects in groups A (atherosclerotic patients), B (old healthy subjects) and C (young healthy subjects) and features extracted from individual PWs. The training results in leave one out cross-validated (LOOCV) classifier performance metrics and

Table 1. Different study groups and the number and proportion of test subjects having different cardiovascular risk factors.

Group	A (Athero-sclerotic)	B (Old healthy)	C (Young healthy)	B+C (Old and young healthy)	D (Athero-sclerotic)	E (Old healthy)	A+B+C+D+E (All)
Age (mean± std)	76.2 ± 7.8	74.6 ± 7.0	42.2 ± 15.3	56.8 ± 20.4	68.6 ± 1.9	71.0 ± 5.0	64.6 ± 17.7
Subjects	20	18	22	40	7	8	75
Males	14 (70.0%)	6 (33.3%)	22 (100.0%)	28 (70.0%)	6 (85.7%)	1 (12.5%)	49 (65.3%)
Smoking	14 (70.0%)	4 (22.2%)	0 (0.0%)	4 (10.0%)	7 (100.0%)	3 (37.5%)	28 (37.3%)
Dyslipidemia	12 (60.0%)	2 (11.1%)	0 (0.0%)	2 (5.0%)	6 (85.7%)	3 (37.5%)	23 (30.7%)
Diabetes	9 (45.0%)	1 (5.6%)	1 (4.6%)	2 (5.0%)	5 (71.4%)	0 (0.0%)	16 (21.3%)
Rheumatoid arthritis	1 (5.0%)	0 (0.0%)	1 (4.6%)	1 (2.5%)	0 (0.0%)	3 (37.5%)	5 (6.7%)
Hypertension	13 (65.0%)	2 (11.1%)	0 (0.0%)	2 (5.0%)	7 (100.0%)	3 (37.5%)	25 (33.3%)

averaged weights for further evaluation. With small sample sizes, like in this study, the outputs of LOOCV procedure are typical results. However, we also tested the resulting LOOCV-based classifier weights with dataset ii which is composed of the averaged FT-plot features of the study subjects in groups A, B and C. In addition to this, two datasets completely independent of the subjects in groups A–C were utilized in the evaluation of the performance of the proposed method: datasets iii and iv are based on study groups D and E and the averaged FT-plot features extracted from the signals recorded with two different devices WBSN and PSP, respectively.

2.3. Ethics and patient safety

The study was conducted in accordance with the Helsinki declaration and approved by the local ethical committees of the hospital districts (R14096 (Pirkanmaa Hospital District) and 147/2014 (Northern Ostrobothnia Hospital District)), the Finnish National Supervisory Authority of Health and Welfare (Valvira) (IDs 272 (WBSN) and 276 (PSP)) and the technical departments of the hospitals. An informed consent was obtained from all the volunteer test subjects.

2.4. Statistical methods

Due to relatively small number of test subjects in datasets iii and iv, non-parametric statistical tests (two-tailed Mann-Whitney U-test) were implemented to find if there are statistically significant differences in the features extracted from the FT-plots between the atheroscleric patients and control subjects. The p-values less than 0.05 were considered as significant.

2.5. FT-plot analysis

Examples of time domain presentations of the PWs are shown in Figs. 2a and 3a. In the FT-plot analysis, PPG-signal measured from the toe ($\mathbf{y} \in \mathbb{R}^n$, a column vector) is drawn as a function of a simultaneous PPG-signal from the finger ($\mathbf{x} \in \mathbb{R}^n$, a column vector) (Figs. 2b and 3b). Even a visual comparison of FT-plots (Figs. 2b and 3b) drawn for atherosclerotic patient and healthy subject reveals clear differences, but for providing objective results, quantitative analysis is needed. We observed the major differences in FT-plots (Figs. 2b and 3b) are especially in the region starting from the upper right corner and ending to the lower left corner. This region corresponds to the falling parts after the peak values in the time domain presentations in Figs. 2a and 3a. These regions from Figs. 2b and 3b are extracted into Figs. 2c and 3c and were selected for further analysis since the reflections that provide information on the arterial elasticity arrive after the peak value of the PW. To enable curve fitting to the

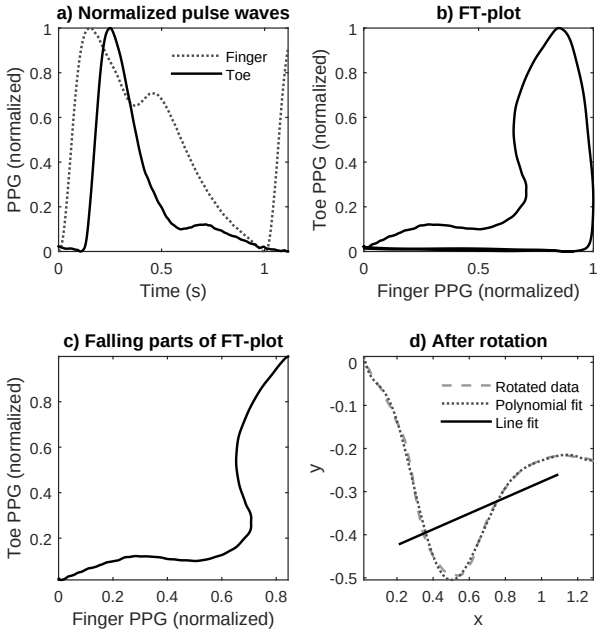


Figure 2. FT-plot signal processing, a young healthy subject.

data in cartesian coordinates, a rotation operation was implemented for Figs. 2c and 3c as

$$\begin{bmatrix} \mathbf{x'}^{\mathrm{T}} \\ \mathbf{y'}^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \mathbf{x}^{\mathrm{T}} \\ \mathbf{y}^{\mathrm{T}} \end{bmatrix} \tag{1}$$

where the coefficient matrix is a rotation matrix and rotation angle $\theta = -60°$ which provided sufficient results from person to person. Example of the results after the rotation are show in Figs. 2d and 3d.

After the rotation, a 9^{th}-order polynomial p was fitted to the rotated curve by using least mean square (LMS) algorithm. The 9^{th}-order polynomial follows sufficiently the original curve but does not suffer from overfitting. In addition to the polynomial, a line l was fitted to middle-region (i.e. to the region where $x_1 < x < x_2$, $x_1 = 0.15\Delta x$ and $x_2 = 0.85\Delta x$ where Δx is the width of rotated the FT-plot curve) of the rotated part of the FT-plot by using LMS algorithm. The middle-region was selected for further analysis since the biggest differences between the patient and the control groups were found from this section. The endpoints of the falling parts had to be excluded because even the atherosclerotic patients have significant nonlinear regions close the endpoints.

After preprocessing the FT plot, 10 different parameters that are potentially capable for discriminating healthty subjects from atherosclerotic patients were derived from the resulting curves and 11^{th} one directly based on the the PWs from finger and toe:

1. Maximum difference of the slopes in the middle-region, i.e. $\max(p') - \min(p')$.

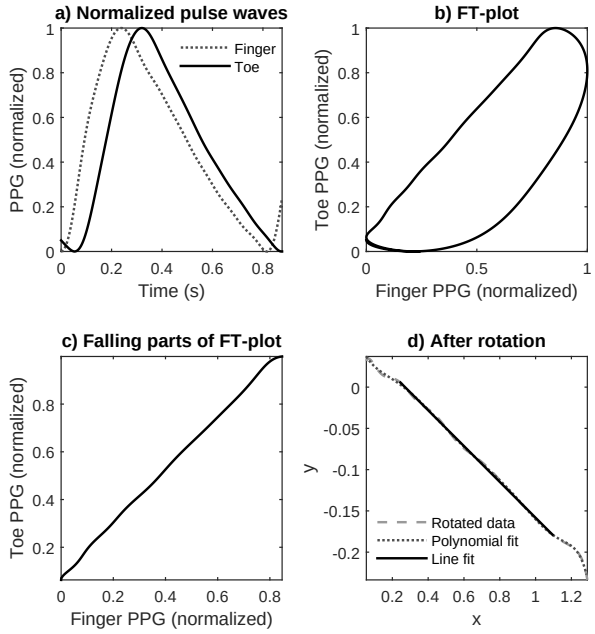


Figure 3. FT-plot signal processing, an atherosclerotic patient.

2. Mean absolute slope of the middle-region, i.e. mean($|p'|$)
3. Standard deviation of the slope of the middle-region, i.e. std(p')
4. Standard deviation of the absolute slope of the middle-region, i.e. std($|p'|$)
5. Integral absolute error between the fitted line and polynomial fit in the middle-region, i.e. $\int_{x_1}^{x_2} |l(x) - p(x)| \mathrm{d}x$
6. Integral square error between the fitted line and polynomial fit in the middle-region, i.e. $\int_{x_1}^{x_2} (l(x) - p(x))^2 \mathrm{d}x$
7. The arc length of the middle-region normalized by the distance between its endpoints, i.e. $\frac{1}{\sqrt{(x_2-x_1)^2+(p(x_2)-p(x_1))^2}} \int_{x_1}^{x_2} \sqrt{1+[p'(x)]^2} \mathrm{d}x$
8. The arc length of the middle-region normalized by the distance between the endpoints of the rotated curve, i.e. $\frac{1}{\sqrt{(\Delta x)^2+(\Delta y)^2}} \int_{x_1}^{x_2} \sqrt{1+[p'(x)]^2} \mathrm{d}x$
9. The maximum absolute difference between the slope of the polynomial fit and the slope of the fitted line, i.e. max($|p' - l'|$)
10. The slope of the fitted line, l'.
11. Ratio of the areas under the finger and toe PPG, i.e. $\frac{A_{finger}}{A_{toe}}$.

Table 2. Mean and standard deviation (Std) values utilized in the normalization of data (columns 2–3). Averaged weights with 4 decimals based on LOOCV for different FT-plot derived features 1–11 in different LDA classifers I–VII (columns 4–10). Empty cell indicates that the particular feature i is not utilized in that classifier.

	Normalization		Classifier						
i	Mean	Std	I	II	III	IV	V	VI	VII
1	1.347	1.060		0.3267				0.3563	
2	0.388	0.232			0.3856				
3	0.417	0.313			-2.1653			0.4245	
4	0.258	0.177							0.3264
5	3.110	2.306	0.3267	0.4542	2.6791	0.3278	0.3230	-2.3065	0.4254
6	0.240	0.346	0.4530	-2.2534	1.0544	0.4241	0.4173	2.9664	0.0441
7	1.101	0.125				0.0822			-2.2834
8	0.757	0.088							
9	0.889	0.780					0.0658		
10	-0.060	0.224	-2.2533	3.2087		-2.3080	-2.2576	1.0711	3.1778
11	1.165	0.209	3.2066	-0.0028	-0.4845	3.1958	3.1790	-0.8258	0.0346

2.6. Classification

In order to see if the combining the information provided by different FT-derived features improves the discrimination capability and diagnostic accuracy, linear discriminant analysis (LDA) classifier and LOOCV method were implemented for the features extracted from the FT-plots. In the LDA analysis, the classifying variable z is defined as

$$z = \mathbf{w}^{\mathrm{T}}\mathbf{p} \tag{2}$$

in which $\mathbf{p} \in \mathbb{R}^{\mathbf{n}}$ is a set of normalized features. For each parameter, Z-score normalization is performed first by subtracting the sample mean and then dividing with sample standard deviation based on the 44236 individual pulse waves in training data set i. The means and standard deviations used in the normalization are shown in Table 2.

The weights $\mathbf{w} \in \mathbb{R}^{\mathbf{n}}$ are defined so that the inter-class variance of the training data is maximized with respect to its intra-class variance when projected onto $\mathbf{w}$. This holds when

$$\mathbf{w} = (\mathbf{C_1} + \mathbf{C_2})^{-\mathbf{1}}(\mu_{\mathbf{1}} - \mu_{\mathbf{2}}) \tag{3}$$

in which $\mathbf{C_1}, \mathbf{C_2} \in \mathbb{R}^{n\times n}$ are covariance matrices and $\mu_{\mathbf{1}}, \mu_{\mathbf{2}} \in \mathbb{R}^n$ are the averages of the datapoints in classes 1 and 2.

In the LOOCV method, the data from one test subject at a time is left outside the training data and used as validation data for the coefficients $\mathbf{w}$. All possible combinations (2047 combinations) of the features extracted from FT-plots were tested as LDA classifier inputs, but in this study we concentrate on the classifiers having the input parameters and weights presented in Table. 2.

The presented classifiers (Table. 2) were selected based on their high overall discrimination capability which was evaluated with four datasets i–iv as explained in 2.2 by using the following indicators: the area under the ROC (receiver operating characteristic) curve (AUC), sensitivity (SE $= a/(a+c)$), specificity (SP $= d/(b+d)$),

Table 3. Statistics for AUC. Columns 1 and 2: Number of classifiers having equal or higher AUC than the presented AUC with all the datasets i–iv. Columns 3 and 4: Number of classifiers with AUC>0.9 in each dataset.

AUC_i	$AUC>AUC_i$	Dataset	AUC>0.9
0.800	1626	i	1397
0.850	720	ii	1707
0.860	639	iii	1858
0.870	541	iv	443
0.880	417		
0.890	197		
0.900	116		
0.905	27		
0.910	9		
0.915	0		

AUC: area under receiver operating characteristic curve

performance (PE = (SE + SP)/2), accuracy (AC = $(a + d)/(a + b + c + d)$), positive predictive value (+PV = $a/(a + b)$) and negative predictive value (−PV = $d/(c + d)$) in which a is the number of true positive, b is the number of false positive, c is the number of false negative, and d is the number of true negative cases. The proposed classifiers are examples of classifiers having high overall performance, having all the listed indicators higher than 0.729 with all the datasets. This limitation was set in order to keep the amount of results manageable: there are e.g. 246 different classifiers having all the aforementioned indicators higher than 0.70 with all the datasets.

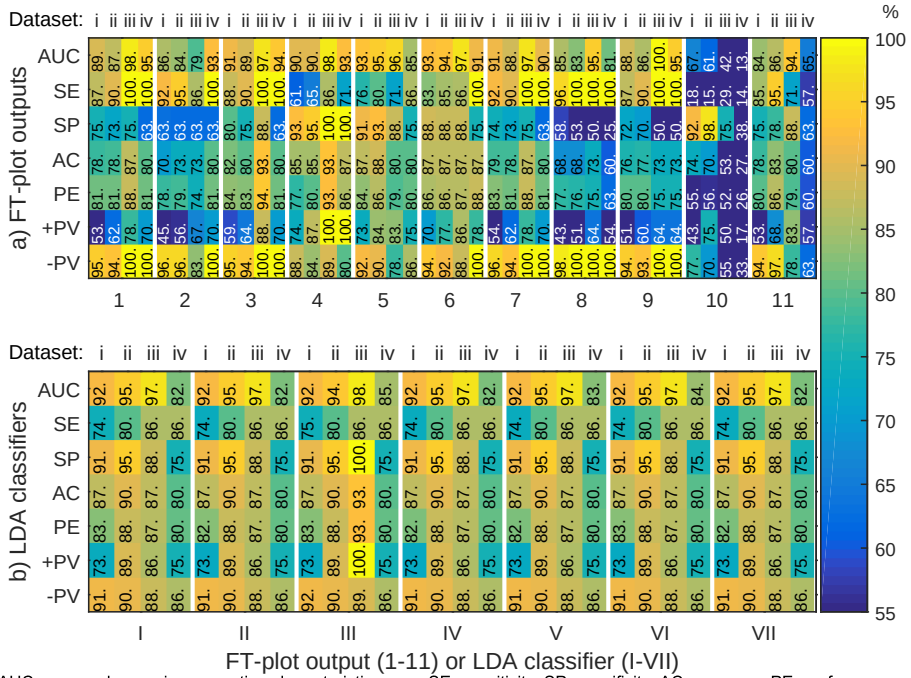
3. Results

As seen in Table 3, excellent AUC values were found for high number of classifiers based on features extracted from the FT-plots. The performance analysis of individual FT-plot features 1–11 and the selected classifiers I-VII are shown as heatmaps in Figs. 4a) and 4b), respectively. All the presented classifiers in Fig 4a) have all the following indicators higher than 0.729 with all the datasets i–iv: AUC, SE, SP, AC, PE, +PV, and -PV. As seen in Figs. 4a and 4b, the best performing individual FT-plot derived features 5 and 6 are almost at the same level with the LDA classifiers I–VII in terms of classification capability. For this reason, clearly worse-performing individual FT-plot features are left outside of the presented results.

The averages and sample standard deviations of the performance metrics computed over different classifiers I–VI (i.e. the data presented as a heatmap in Fig. 4b) are shown in Table 4 for the different datasets are and their joint distribution. In addition, Table 4 contains also the averages for the individual FT-plot features 5 and 6 computed over different datasets.

In addition to the classifier metrics, the ROC-curves are shown in Fig. 5 for the classifiers I–VII and the individual FT-plot features 5–6 for the for each dataset. the differences in the results obtained with different devices are often similar except for one subject. The locations of incorrectly classified instances with respect to the partition value and the distributions of the testing data (dataset ii) are shown in Fig. 6.

Two-tailed Mann-Whitney U-test returns p-values $p < 0.05$ for all the proposed



AUC: area under receiver operating characteristic curve, SE: sensitivity, SP: specificity, AC: accuracy, PE: performance, +PV: positive predictive value, -PV: negative predictive value, FT-plot: finger toe plot, LDA: linear discriminant analysis

Figure 4. Classifier metrics in heatmap form, including percentage numbers. Panel a): Results for individual parameters 1–11, i.e. the outputs of FT-plot. Panel b): Results for different classifiers I–VII. In each panel, each 4-pixel-width column separated by white lines represents the results for 4 datasets i–iv, labels are shown on top of the panel. The values are shown in each pixel and the scale of the color coding (%) is shown at right. The column order from left to right: i: training data (LOOCV, individual PWs), ii: averaged parameters based on the training data i.e. testing data, iii: WBSN and iv: PSP.

Table 4. Average (mean±std, %) performance metrics computed over different classifiers for each dataset and their joint distribution. For individual FT-plot features 5 and 6, the numbers are computed over the different datasets i–iv.

Dataset	i (training) (LOOCV)	ii (aver. training data)	iii (WBSN)	iv (PSP)	Joint distribution	Feature 5	Feature 6
AUC	91.7 ± 0.1	94.5 ± 0.3	97.4 ± 0.3	82.9 ± 1.1	91.6 ± 5.6	92.2 ± 5.1	93.6 ± 2.7
SE	74.0 ± 0.4	80.0 ± 0.0	85.7 ± 0.0	85.7 ± 0.0	81.4 ± 4.9	78.3 ± 6.1	88.4 ± 7.9
SP	91.0 ± 0.1	95.0 ± 0.0	89.3 ± 4.7	75.0 ± 0.0	87.6 ± 8.0	86.5 ± 7.9	84.6 ± 6.4
AC	86.8 ± 0.1	90.0 ± 0.0	87.6 ± 2.5	80.0 ± 0.0	86.1 ± 4.0	83.9 ± 4.5	86.7 ± 0.1
PE	82.5 ± 0.2	87.5 ± 0.0	87.5 ± 2.4	80.4 ± 0.0	84.5 ± 3.4	82.4 ± 3.1	86.5 ± 0.8
+PV	73.1 ± 0.1	88.9 ± 0.0	87.8 ± 5.4	75.0 ± 0.0	81.2 ± 7.8	79.0 ± 5.6	77.7 ± 6.4
-PV	91.4 ± 0.1	90.5 ± 0.0	87.7 ± 0.5	85.7 ± 0.0	88.8 ± 2.3	86.4 ± 6.3	93.4 ± 5.2

LOOCV: leave one out cross-validation, WBSN: wireless body sensor network, PSP: phase-sensitive PPG (photoplethysmogram), AUC: area under receiver operating characteristic curve, SE: sensitivity, SP: specificity, AC: accuracy, PE: performance, +PV: positive predictive value, -PV: negative predictive value.

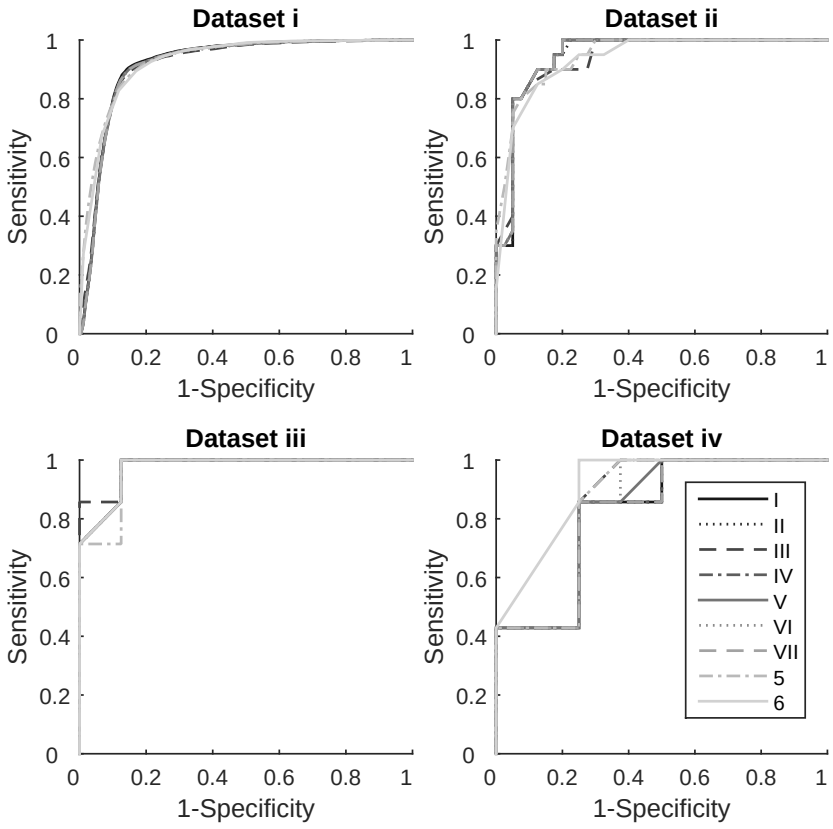


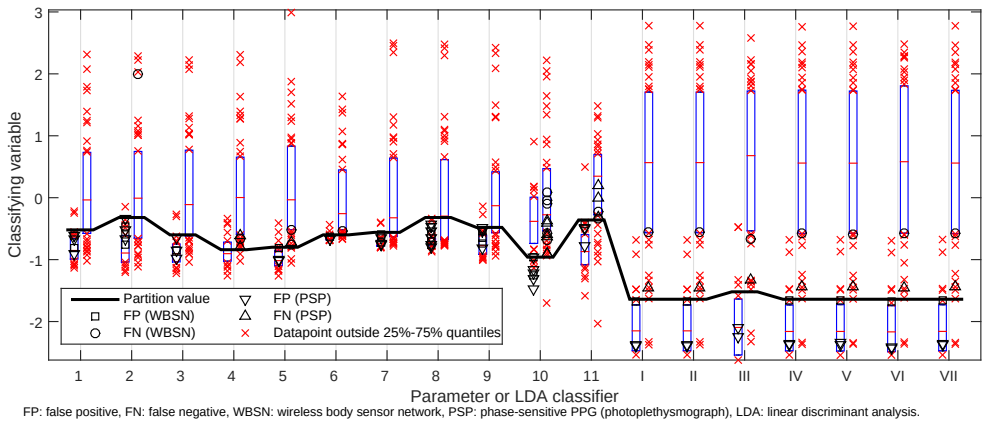
Figure 5. ROC curves for datasets i–iv and presented classifiers I-VII (solid lines) and individual FT-plot features 5 and 6 (dashed lines).

LDA classifiers with all the datasets and $p < 0.002$ for all the classifiers with datasets i–iii. The individual input parameters of the classifiers have more deviation in the p-values: only 6 out of 11 have $p < 0.05$ and only 5 out of 11 have $p < 0.01$ for all the datasets.

4. Discussion

According to the results, FT-plot-based method provides a well-performing tool for detecting the markers of atherosclerotic changes. In this study, we found (see Table 4) the average AUC of 91.6%, sensitivity of 81.4%, specificity of 87.6%, accuracy of 86.1%, performance of 84.5% and positive and negative predictive values of 81.2% and 88.8%, respectively, for the presented LDA-classifiers.

The distributions of the PW parameters extracted for different study groups from individual measurement sites have often wide overlapping ranges close to the partition value. Although this often results statistically significant differences between the groups and even a good value for the optimized one of the following two indicators, sensitivity or specificity, a poorer value is often found for the non-optimized one as e.g. with bilateral differences related parameters extracted from lower-limb PPG signals



FP: false positive, FN: false negative, WBSN: wireless body sensor network, PSP: phase-sensitive PPG (photoplethysmograph), LDA: linear discriminant analysis.

Figure 6. Distributions of the dataset ii as boxplots for both individual FT-plot derived features 1–11 and LDA classifiers I-VII. The left-hand-side boxes of each boxplot pair are for atherosclerotic patients and right-hand-side boxes for control subjects. The data points of dataset ii outside the 25%–75% quantiles are shown as red crosses. The y axis is for the classifying variable. The solid black line illustrates the partition value used for drawing Figs. 4a)–b). False positive (FP) and negative (FN) for the results obtained with datasets iii (WBSN) and iv (PSP) are also shown.

in [13], ABI results in [1] and radial artery tonometry-based AIx in [15]. In our results with the proposed method, AUC values higher than 0.90 are found, leading to the typical sensitivities and specificities around 0.8–0.9.

The proposed method is not biased with any assumptions on normal or typical PW contour as the basic idea behind the proposed methos is to compare PW signals from two different measurement sites only in the intra-subject level. The feature detection procedures are also simpler compared with the direct PW based parameters (such as AIx, AGI, SI, and RI). The problem with direct PW based parameters is that the detection of the PW features needed for computing the parameters is often everything but clear and obvious, and at least the automated detection of the required feature points often leads to implementation of noise-sensitive high-order derivative analysis [8, 9]. A wider comparative study between the proposed method and direct PW based parameters is in our future interests.

The results show that the similarity of the falling parts of the PWs recorded from finger and toe indicate atherosclerotic changes in the arterial tree. This similarity is caused by a lack of a clear incisura point as a borderline between systolic and diastolic parts of the PW. The physiological background behind this is assumed to be related to the loss of arterial elasticity: the stiffer the arteries are due to the atherosclerotic changes, the less the arteries are capable of storing the energy during systole provided by the heartbeat-induced blood pulse and gradually releasing it during the diastole, the more the percussion wave and its reflections are overlapping and the more similar PW is observed at different recording sites due to the increased propagation velocity and thus decreased time delays between the percussion wave and its reflections. These are especially related to the pressure PW, but the pressure PW and the volume PW measured with PPG have non-linear dependence on each other [6]. This combined with

the stenoses or occlusions in the arterial pathway as a result of advanced atherosclerosis decrease the peripheral blood perfusion and modify the observed PPG.

Overall, the results (Figs. 4, 6 and Table 4) indicate that the proportion of false positive instances is higher than the proportion of false negative instances. Possible natural reasons for this observation are the subclinical or latent atherosclerotic changes that cannot be observed with the ABI measurement [1, 2] or do not cause any symptoms yet. However, the proportions of false positive and negative results are dependent on the selected partition value. The most optimal partition value is subject of further studies and depends on the application of the diagnostic test, such as if it is particularly important to detect atherosclerotic changes or exclude the disease. In this study, we weighted all the partition value dependent performance metrics (SE, SP, AC, PE, +PV and -PV) equally. One option is to utilize different classifiers in different applications in possible clinical use of the proposed method. Also the utilization of individual FT-plot based features is a potential option since e.g. FT-plot features 5 and 6 have all the tested metrics of 0.70 or higher. and As seen e.g in Fig. 4, the individual FT-plot derived features are often more sensitive but less specific than the presented LDA classifiers.

Early detection of atherosclerotic changes serves preventive strategies: the patient can be medicated and monitored properly or encouraged towards healthier diet and lifestyle. Postponed or prevented surgical operations and hospitalization periods will save (public) health care costs in a world with increasing aging population. In addition, a false positive result in the detection of asymptomatic atherosclerosis does not lead to any treatments with high risk, but a false negative result may lead into severe consequences if the patient is not regularly monitored.

Testing with datasets iii and iv suggests that the developed method works independently on the measurement hardware although the classification results with different datasets are not exactly equal. However, as seen in Fig. 6, the differences are often a result of one differently classified instance. With some of the classifiers, the same two asymptomatic subjects measured with WBSN and PSP and having several risk factors are falsely classified with respect to our reference value (ABI measurement) as atherosclerotic patient. In addition, none of the test subjects in datasets iii and iv was under hospital treatment during the measurements whereas the data used in LOOCV based training contained both patients being under hospital treatment due to atherosclerosis, as well as young healthy subjects and patients being under hospital treatment for other reasons than cardiovascular diseases. In this point of view, as none of the test subjects in the groups D and E (Table 1) and thus in datasets iii and iv was suffering from acute manifestations atherosclerosis (e.g. myocardial infarction, stroke or acute limb ischemia), the groups are expected to be more homogenenous in terms of the status of the vasculature. The differences may be caused also by the different frequency responses of the different photoplethysmograms: all the training data was recorded with WBSN, but dataset iv was recorded with PSP. On the other hand, physiological reasons such as the changes in the blood pressure levels or the activity of autonomous nervous system between the recordings with the different devices are also possible reasons. A notable issue is also the posture of upper limb between the measurements: the upper limbs were always next to torso in the measurements with WBSN, whereas the upper limbs were bent and placed on the abdominal region with part of the test subjects in the measurements with PSP due to slightly too short measurement cables, and this may affect to the results since PPG waveform depends on the height of the limb with respect to heart level[16]. With respect to these facts,

the results obtained with the datasets iii and iv but recorded with different systems are excellent.

A direct comparison between the developed method and the current gold standard and quick test, ABI, is not straigthforward. ABI and the results of risk factor questionaire are considered as ground truth in evaluating the developed method although the atherosclerotic changes may be latent and asymptomatic for relatively long time before the diagnosis, and this is one limitation of our study. As found in literature, even the performance of a diagnostic test ABI$<$ 0.9 has wide variation from study to study (SE: 15%–79%, SP: 83.3%–99.0%, AC: 72.1%–89.2%, +PV: 36.4%–99.0% and -PV: 40.7%–98.0%) [1]. Due to the wide variation in the performance of the ABI measurement, the developed technique could be a well comparable alternative in daily practices along with the ABI, although the current methods in clinical use have a long history.

Related to the toe-PPG measurements, Allen *et al* have reported better results for the diagnosis of atherosclerosis from a leg having ABI$<$ 0.9: they have found sensitivities of 88.9%–92.7%, specificities of 89.3%–90.6% and accuracies of 90.2%–90.5% for a parameter called shape index [12, 13]. Their method compares the measured and normalized PW contour to a reference PW contour which is considered as normal PW. This differs from our method where the features utilized in the diagnosis of atherosclerosis are found in the comparison of the PPGs recorded from different locations from the subject themself.

For the bilateral differences between the shape indices from different limbs, a sensitivity of 58.3%–59.4%, specificity of 88.9%–92.0% and accuracy of 82.2%–75.7% are reported [12, 13]. Related to this, the performance of a bilateral comparison of any PW-related quantity may decrease if there are atherosclerotic changes in both legs. Other parameters related to include bilateral delays between foot or peak points of PWs [17, 12, 13], rise time of the PW and calibrated amplitude ratio. For these quantities, sensitivities of 31.5%–84.2%, specificities of 82.1%–97.3% and accuracies of 65.8%–87.9% have been reported.

For comparison, the aging indices calculated from the same data and based only on finger PPG and only on toe PPG had AUCs 0.7 and 0.79, respectively, whereas AUCs higher than 0.90 were typical results with selected the FT-plot derived parameters. When deciding whether the subject has high or low risk for cardiovascular diseases, diagnostic accuracy (AC), sensitivity (SE) and specificity (SP) of 87.5% have been reported in [10] for the best-performing classifier utilizing support vector machine and index finger PPG. The diagnostic accuracy of the method in [10] is approximately equal to the proposed FT-plot based classifiers, but the refence value in [10] for the classification result is the carotid femoral pulse wave velocity. This limits straightforward comparison with the presented results having the ABI as a reference value.

The proposed method has a single technical limitation: the measurement of the PPG signals, toe PPG in particular, is not necessarily possible with a subject having extremely occluded arteries and thus decreased perfusion. However, such patients have distinctively atherosclerosis, so the FT-plot analysis does not provide additional information.

A primary goal of this study was to develop diagnostic tools for atherosclerosis and to show the power of multichannel PPG measurement. Despite the promising results, there are potential limitations. The first one is the limited number of study subjects consisting of 75 caucasians which is inadequate for generalization of the results to the

whole population. A separate validation study with larger number of study subjects and more reliable reference methods, such as angiogram, is needed for the validation of the results. With respect to analysis methodology, the number of classifiers with promising results is high (see Table 3), and we present here only the performance of very limited number of all possible combinations. In larger studies, different combinations may turn out to be the best ones as there are numerous promising combinations of FT-plot outputs (see Table 3). For further studies, it is also important to standardize the measurement conditions, including the posture. Also a comparison between existing PW analysis methods and repeatability related research questions are in our future interests.

5. Conclusions

The proposed noninvasive PPG-based FT-plot technique has unexploited potential in vascular screening for revealing the latent disorders of early-stage atherosclerotic changes. Based on LOOCV-validated training data, testing data and two other datasets recorded with two different PPG devices, an average AUC of 91.6%, sensitivity of 81.4%, specificity of 87.6%, accuracy of 86.1%, performance of 84.5% and positive and negative predictive values of 81.2% and 88.8%, respectively, were achieved for simple linear discriminant analysis based classifiers. The present results indicate that the developed methods are slightly more efficient in exclusion than diagnosis of atherosclerosis. However, depending on the application and requirement, the partition value can be adjusted or different tests can be selected for different purposes.

The present results encourage us for further studies with noninvasive multichannel PW measurement and analysis methods. This approach may lead into development of accurate and easy-to-use diagnostic tools for the detection of atherosclerosis and subsequently facilitate preventive strategies.

6. Acknowledgment

We would like to thank all the volunteer test subjects for their valuable contribution to the study. Research nurse Ritva Heikkinen from Oulu University Hospital is especially acknowledged for her contribution in recruiting test subjects and collecting the reference values.

The work was funded by the Doctoral Programme of the President of Tampere University of Technology, Finnish Funding Agency for Technology and Innovation (TEKES) as a part of project VitalSens (decision ID 40103/14), and grants from Finnish Cultural Foundation/Pirkanmaa Regional Fund/Elli and Elvi Oksanen's Fund, Tekniikan edistämissäätiö and Emil Aaltonen Foundation.

References

[1] Dachun Xu, Jue Li, Liling Zou, Yawei Xu, Dayi Hu, Sherry L Pagoto, and Yunsheng Ma. Sensitivity and specificity of the ankle-brachial index to diagnose peripheral artery disease: a structured review. *Vascular Medicine*, 15(5):361–369, 2010.

[2] N.K.J. Oksala, J. Viljamaa, E. Saimanen, and M. Venermo. Modified ankle-brachial index detects more patients at risk in a finnish primary health care. *European Journal of Vascular and Endovascular Surgery*, 39(2):227 – 233, 2010.

[3] Mohamed Elgendi. On the analysis of fingertip photoplethysmogram signals. *Current cardiology reviews*, 8(1):14–25, 2012.

[4] Sandrine C Millasseau, Ronan P Kelly, James M Ritter, and Philip J Chowienczyk. The vascular impact of aging and vasoactive drugs: comparison of two digital volume pulse measurements. *American journal of hypertension*, 16(6):467–472, 2003.

[5] Kenji Takazawa, Nobuhiro Tanaka, Masami Fujita, Osamu Matsuoka, Tokuyu Saiki, Masaru Aikawa, Sinobu Tamura, and Chiharu Ibukiyama. Assessment of vasoactive agents and vascular aging by the second derivative of photoplethysmogram waveform. *Hypertension*, 32(2):365–370, 1998.

[6] Uldis Rubins. Finger and ear photoplethysmogram waveform analysis by fitting with gaussians. *Medical & biological engineering & computing*, 46(12):1271–1276, 2008.

[7] Vojtech Melenovsky, Barry A Borlaug, Barry Fetics, Kristy Kessler, Laura Shively, and David A Kass. Estimation of central pressure augmentation using automated radial artery tonometry. *Journal of hypertension*, 25(7):1403–1409, 2007.

[8] M. Peltokangas, A. Vehkaoja, J. Verho, V-M. Mattila, P. Romsi, J. Lekkala, and N. Oksala. Age dependence of arterial pulse wave parameters extracted from dynamic blood pressure and blood volume pulse waves. *Biomedical and Health Informatics, IEEE Journal of*, PP(99):1–1, 2015.

[9] Kenji Takazawa, Nobuhiro Tanaka, Kazuhiro Takeda, Fujio Kurosu, and Chiharu Ibukiyama. Underestimation of vasodilator effects of nitroglycerin by upper limb blood pressure. *Hypertension*, 26(3):520–523, 1995.

[10] Stephen R Alty, Natalia Angarita-Jaimes, Sandrine C Millasseau, and Philip J Chowienczyk. Predicting arterial stiffness from the digital volume pulse waveform. *IEEE transactions on biomedical engineering*, 54(12):2268–2275, 2007.

[11] Matti Huotari, Antti Vehkaoja, Kari Määttä, and Juha Kostamovaara. Photoplethysmography and its detailed pulse waveform analysis for arterial stiffness. *Journal of Structural Mechanics*, 44(4):345–362, 2011.

[12] John Allen, Crispian P Oates, Timothy A Lees, and Alan Murray. Photoplethysmography detection of lower limb peripheral arterial occlusive disease: a comparison of pulse timing, amplitude and shape characteristics. *Physiological measurement*, 26(5):811, 2005.

[13] John Allen, Klaus Overbeck, Alexander F Nath, Alan Murray, and Gerard Stansby. A prospective comparison of bilateral photoplethysmography versus the ankle-brachial pressure index for detecting and quantifying lower limb peripheral arterial disease. *Journal of vascular surgery*, 47(4):794–802, 2008.

[14] M. Peltokangas, A. Vehkaoja, J. Verho, M. Huotari, J. Röning, and J. Lekkala. Monitoring arterial pulse waves with synchronous body sensor network. *Biomedical and Health Informatics, IEEE Journal of*, 18(6):1781–1787, Nov 2014.

[15] Eshan Patvardhan, Kevin S Heffernan, Jenny Ruan, Michael Hession, Patrick Warner, Richard H Karas, and Jeffrey T Kuvin. Augmentation index derived from peripheral arterial tonometry correlates with cardiovascular risk factors. *Cardiology research and practice*, 2011, 2011.

[16] M Hickey, J P Phillips, and P A Kyriacou. The effect of vascular changes on the photoplethysmographic signal at different hand elevations. *Physiological Measurement*, 36(3):425, 2015.

[17] Renars Erts, Janis Spigulis, Indulis Kukulis, and Maris Ozols. Bilateral photoplethysmography studies of the leg arterial stenosis. *Physiological measurement*, 26(5):865, 2005.