Analytical Methods

Spectral data quality assessment based on variability analysis: application to noninvasive hemoglobin measurement by Dynamic Spectrum

Journal:	*Analytical Methods*
Manuscript ID:	AY-ART-03-2015-000669.R1
Article Type:	Paper
Date Submitted by the Author:	25-May-2015

Complete List of Authors:	He, Wenqin; Tianjin University, State Key Laboratory of Precision Measurement Technology and Instruments; Tianjin University, Tianjin Key Laboratory of Biomedical Detecting Techniques & Instruments Li, Xiaoxia; Hebei University of Technology, School of Electrical Engineering Wang, Mengjun; Hebei University of Technology, School of Electrical Engineering Li, Gang; Tianjin University, State Key Laboratory of Precision Measurement Technology and Instruments; Tianjin University, Tianjin Key Laboratory of Biomedical Detecting Techniques & Instruments Lin, Ling; Tianjin University, State Key Laboratory of Precision Measurement Technology and Instruments; Tianjin University, Tianjin Key Laboratory of Biomedical Detecting Techniques & Instruments

Journal Name

RSCPublishing

ARTICLE

Spectral data quality assessment based on variability analysis: application to noninvasive hemoglobin measurement by Dynamic Spectrum

Wenqin He,[a,b] Xiaoxia Li,[c] Mengjun Wang[c] ,Gang Li[a,b] and Ling Lin*[a,b]

Cite this: DOI: 10.1039/xoxxoooooox

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/xoxxoooooox

www.rsc.org/

The quality of spectral data is crucial to the accuracy of quantitative spectral analysis, especially the noninvasive measurement of blood components. As an important part of research, establishing an effective and reliable quality assessment metric to select data before modelling is indispensable. According to the principle of Dynamic Spectrum (*DS*) and the characteristics of Photoplethysmogram (PPG), this study proposed a novel method to assess spectral quality ---- stability coefficient (*SC*) and we proposed an analytical formula. To verify the feasibility, in simulation analysis, we calculated the stability coefficient of simulated spectra and evaluated the performance of extraction by Root Mean Square Error (*RMSE*). The result shows a negative correlation between stability coefficient and *RMSE*. After simulation analysis we conducted control experiment based on data from 427 subjects by developing calibration models between DS data and hemoglobin concentration. The average correlation coefficient is 0.875 in the test set of experimental group, while that of the control group is only 0.715. The actual experimental result is consistent with the simulation analysis to demonstrate that the assessment method can evaluate the quality of spectral data efficiently and accurately. This new quantitative method provides a reliable way to assess and screen spectral data. It could not only be applied to noninvasive measurement of blood components but also to other related fields such as spectral analysis, signal measurement and processing.

Introduction

Noninvasive measurement of blood components is one of the frontiers of biomedical engineering [1-3], especially the measurement of hemoglobin concentration [4-8], and each study has its own uniqueness and advantages. Hemoglobin concentration is an important indicator of health condition. Continuous noninvasive monitoring of hemoglobin can assist in prevention and diagnosis of diseases. High hemoglobin level usually means poor heart or lung function which result in chronically low blood oxygen level, or Polycythaemia Vera. More commonly, low concentration of hemoglobin which is referred to as anemia, is now affecting approximately a quarter of the world population [9]. Also, in order to prevent excessive bleeding, monitoring of hemoglobin concentration during and after the surgical operation is necessary. Conventional methods for hemoglobin measurement always bring fear and suffering, therefore, accurate and convenient method for noninvasive hemoglobin measurement has long been everybody's aspiration.

Because of the optical properties of human tissues in visible and near-infrared (NIR) region, quantitative spectral analysis [10] is an effective, convenient and painless approach to monitor the blood components. In order to eliminate the effect of the "static tissues" and the skin pigment, dynamic spectrum (DS) [11-12] has developed rapidly, which measure the blood components based on PPG [13]. DS method can individually extract the absorption spectrum of arterial pulsatile blood. As a promising approach, it can dispel the influence of measurement conditions and individual variation in principle. Therefore it markedly shows great potential.

Improving the accuracy of measurement is the common goal of all studies, many methods have been proposed to enhance the performance of DS, such as FDLIA [14], double-sampling method [15]. But previous studies mainly focused on the data sampling and extraction. Modelling between the DS data and the blood component concentration is an efficient approach for measurement. So the quality of training data is vital to the performance of DS method. Therefore, the spectral quality assessment and data selection is quite important.

The term "spectral quality assessment" has different implications based on the application. *e.g.*, a narrow peak without secondary peaks is preferred in traditional spectroscopy, while the quality of PPG wave is the primary factor in DS. Generally, there are two types of approaches to assess the quality of dynamic spectral data: subjective evaluation and objective evaluation. In subjective method, researchers observe the PPG wave and spectral curve and make subjective judgements of the

quality. As for objective evaluation, which is more commonly used, is based on objective indicators and characteristics of the spectral data, *e.g.*, in time domain, the signal to noise ratio (*SNR*) in different wave band [16-17]: the background noise on the spectral curve is usually a serious interference factor. There is also assessment methods based on frequency domain analysis, *e.g.*, "stable wavelength number" [18]: assessing the spectral quality in frequency domain by analysing the differences among the fundamental frequencies of the PPG signal at different wavelengths by FFT. But the rationality of its definition may need to be further proven because the fundamental frequencies of one's PPG at different wavelengths are unlikely to be different at the same time unless the noise is extremely intense.

According to the basic theory of DS, what carries the major information of the absorption properties of pulsatile blood is the peak-to-peak values of the PPG wave. In ideal conditions (there is no noise or other disturbance), we can extract them accurately. But in practical situations, it's extremely difficult to get a stable PPG signal because of the influence of various disturbances such as motion artifacts, baseline drift. After analysing the characteristics of PPG and the principle of DS, in order to select good spectral data for modelling, we proposed a novel method based on variability analysis to evaluate the spectral quality: stability coefficient (*SC*). We focus our attention on the quality of measurement: stability and repeatability, which is the key issue of all kinds of dynamic measurements. Variability analysis is a practical common method for measurement reliability analysis [19-20] and signal analysis [21]. It is often used to assess the quality of measurements or assays [22]. In this study, we evaluated the extent of variability of acquired spectral data, *i.e.*, the quality of data. *SC* is an intuitive indicator which indicates the degree of original information being polluted by the noise of equipment, ambient and other factors.

We performed some simulations and the results demonstrate its feasibility. In the actual experiments, performances of BP artificial neural networks trained on different data set were compared, and the results illustrate that it can indeed be used to evaluate the spectral quality and can help to screen data for modelling.

Theory and analysis

Dynamic spectrum

Dynamic Spectrum (DS) is a spectral method for noninvasive measurement of arterial blood components. DS method is based on the measurement and analysis of photoplethysmography (PPG). As shown in Fig. 1, due to abundant blood capillaries under the skin, absorbance of the tissue cyclically vary with the perfusion of blood to the dermis and subcutaneous tissue. Supposing there is a light beam transmitting through the fingertip tissue. With each cardiac cycle the heart pumps blood to the periphery. During diastole, with the decreasing of blood flow in the arteries and arterioles, the transmitted light will reaches its maximum intensity. Then when heart contracts, absorbance of the tissue will increases, and at the end of systole we can get the lowest transmitted light intensity. Theoretically, when measuring the PPG, the subcutaneous tissue can be simply considered to be composed of two parts: pulsatile part and static part (*i.e.*, analyte and matrix). The absorbance of all the "static" tissues (mainly the venous blood and other tissues) remain constant during the measurement process. *i.e.*, the pulsatile part of arterial blood is the only factor which lead to the circular change of absorbance of tissues. Therefore, we can acquire the absorption information of arterial blood by measuring the alternating component of PPG signal: subtracting the background absorption from the total absorption. DS method which can eliminate the individual discrepancies of static tissues is more like a kind of *ex vivo* measurement of blood components.

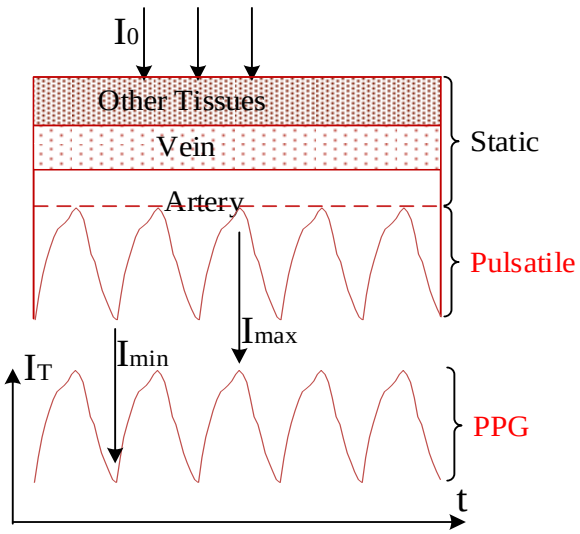

Fig. 1 The generation of Photoplethysmography (PPG) and a Simplified model of tissue. PPG is a simple optical method to detect blood volume changes in microvascular bed of tissue [23]. PPG waveform is consist of two part: AC (pulsatile) and DC (baseline), so we can regard the tissue as a combination of pulsatile part and "static" part.

DS theory is based on the modified Lambert-Beer's law (*MLBL*) [24]. As Eq. (1) and (2), I_0 and I_T represent the incident light intensity and the transmitted light intensity respectively. The maximum and minimum transmitted light intensities are represented by I_{max} and I_{min} which correspond respectively to the minimum and maximum arterial blood flow. ε_i^λ is the molar extinction coefficient of component i at wavelength λ, c_i is the concentration of the component, d represents the optical pathlength. L is differential pathlength factor (DPF), which represents the pathlength lengthening caused by scattering and G depicts the scattering loss. L and G are considered invariable throughout the measurement.

$$OD^\lambda = ln\left(\frac{I_0^\lambda}{I_T^\lambda}\right) = ln\left(I_0^\lambda\right) - ln\left(I_T^\lambda\right) = \sum_{i=1}^{n} \varepsilon_i^\lambda c_i dL + G \quad (1)$$

ΔOD^λ is the change of absorbance during one cardiac cycle at wavelength λ. As Eq. (2), by extracting the alternating component of the logarithmic PPG (PPG signal after logarithmic transformation), we can get the light absorption information of arterial pulsatile blood.

$$\Delta OD^{\lambda} = ln\left(I_{max}^{\lambda}\right) - ln\left(I_{min}^{\lambda}\right) = A^{\lambda} \quad (2)$$

A^{λ} represents the absorbance of arterial pulsatile blood. If we measure A at different wavelengths, we can get the Dynamic Spectrum (DS) which in essence is the optical density of the pulsatile blood. One of the advantages of DS method is eliminating the influence of the variation of light source or other measurement conditions, and the disturbance of static tissues.

After obtaining the logarithmic PPG, DS data can be extracted by existing methods such as frequency domain analysis [25], single trial estimation [26] which is more commonly used. Frequency domain analysis extracts the amplitude of fundamental of the logarithmic PPG by FFT, but its accuracy is low because PPG is not a single-frequency signal. As we all know, if a signal has a certain level of total energy, the wider the frequency spectrum width, the weaker the amplitude of fundamental. Single trial estimation has a strong anti-noise ability due to the elimination of abnormal data and the error correction by line-fitting.

In practice, the concentration of blood component such as hemoglobin is analysed by establishing calibration model. The precondition of high accuracy is that the spectral data we used for training are representative and high-quality, so the quality assessment of spectral data is a very key issue.

Spectral data quality assessment

Stability Coefficient

As mentioned, the core idea of DS method is to acquire high quality spectral data at an early stage to obtain abundant accurate information, so as to set up model with good predictive ability and robustness. Because the nature of DS is dynamic measurement, the spectral data is easily to be disturbed by noises, such as environmental factor (ambient light), instrumental error and man-made disturbance (hand or finger tremor, coughing, and changes of the breathing pattern). These factors damage the signal and decrease the final quality of spectrum, and in awful circumstances the useful signal may be even submerged in noises.

The peak-to-peak value of the logarithmic PPG, which contains the absorption characteristics of arterial pulsatile blood [26], is the primary information source in DS method. In practical situations, more often than not, the acquired signal is noisy and the peak-to-peak values are uneven because of the noise interference. As for common measurement or signal acquisition systems, compared with the multiplicative noise caused by inaccurate calibration or unsatisfactory transmission characteristics of the system, it is the additive noise that do the most harm to the data quality. Although we can improve the *SNR* by averaging the results of repeated measurements on condition that there is only white noise. In reality, the influence of the most pernicious noises such as random impulsive noise and burst noise can hardly be mitigated. Due to the complicated model, it's hard to attenuate the noises in measurement and transmission. Once the spectral data acquiring is finished, the only thing we can do to improve the accuracy is to pick out high quality data from the existing results to build model. Spectral data of a subject is regarded valuable if the PPG wave is steady and noiseless, namely, the distribution of the measurement results is highly concentrated. But if the data is seriously polluted by noises and the amplitudes of PPG fluctuate drastically, we just discard the spectrum of that subject because its absorption information can hardly be extracted correctly, and such spectral data is useless or even detrimental to the modelling. In other words, assessment and screening spectral data is of great significance to noninvasive hemoglobin measurement.

To assess the spectral quality accurately and effectively, we analysed the stability of PPG signal at different wavelengths and assess its quality, as follows:

(1) Smooth and denoise the original spectrum, and remove the saturated spectral data.
(2) Generate the PPG model by adding up all the logarithmic PPG waves over a certain wavelength range and remove the signal segments with gross error (In this study, we use Grubbs' test method with significance level of 0.05 for gross error elimination).
(3) Over the available wavelengths i, fit the rising and falling to the corresponding edges of the model, and the slopes are A_1^i, A_2^i,…, A_n^i.
(4) Calculate the Stability Coefficient (SC) over the available m wavelengths in which we are interested, as Eq. (3):

$$SC = \sum_{i=1}^{m} \eta_i \cdot \frac{1}{CV_i} = \sum_{i=1}^{m} \eta_i \cdot \frac{\overline{A^i}}{S(A^i)} \quad (3)$$

CV_i is the coefficient of variation of A_1^i, A_2^i,…, A_n^i, which expresses the dispersion of signal at wavelength *i*. Coefficient of variation is a precise description that can help to compare the extent of variation from one data series to another, even if the averages of each series vary drastically. The most common approach for variability analysis is calculating CV, which is also common used in queuing Theory and reliability theory. η_i is the weighting factor given to the corresponding wavelength, which can be adjusted according to specific condition (*e.g.*, we can weighting the wavelengths depending on the importance of each independent variable in modelling). In this study, all independent variables are equally weighted ($\eta = 1/m$). *SC* is a standardized indicator of the measurement quality, and its value is independent of the scale of sampling values over different wavelengths.

Simulation Analysis

To analyse the feasibility and effectiveness of the *SC* evaluation method, a series of computer simulations were implemented in *MATLAB*. A computer-generated sinusoidal signal S_0 with unit amplitude was substituted for the PPG wave. The frequency *f* was randomly selected from the range between 0.8 *Hz* and 1.5 *Hz*.

$$S_0 = sin(2\pi \cdot f \cdot t) \quad (4)$$

In order to simulate the transmission spectrum of human tissue, the original virtual PPG (S_0) was modulated to different levels, as Eq. (5).

$$S^i = A^i \cdot S_0 = A^i \cdot sin(2\pi \cdot f \cdot t),\ i = 1, 2 \ldots m \quad (5)$$

Supposing the transmission spectrum is a sine curve between 0 and π, as described in Eq. (6),

$$A^i = sin\left(\frac{i}{m}\cdot\pi\right),\ i\ =\ 1\,,2\ \ldots\ m \qquad (6)$$

where m expresses the number of wavelengths, and we set m= 200 . All the modulated signals together constituted the "mock spectrum". Moreover, the sampling rate of PPG was 50 Hz, and the number of sampling points was 1000.

Then we added three types of noises to the mock spectrum, *viz.*, (i) Low-frequency sine wave N_L with frequency f_L randomly drawn from standard uniform distribution on 0.2 Hz - 0.4 Hz and amplitude A_L randomly drawn from standard uniform distribution on (0.5, 3.5), for simulation of drifted baseline caused by breathing, *etc.* (ii) Step signals N_S with random height and location, for simulation of motion artifacts. As Eq. (8), A_H was randomly drawn from standard uniform distribution on (-2, 2) and $\tau \in (0, 20)$ represented the location of step signal. (iii) White Gaussian noise N_{wg} whose power relative to the original signal was randomly drawn from the range between -5 dB and -20 dB. Furthermore, all the simulated signals were superposed onto a certain DC offset D.

$$N_L = A_L \cdot sin(2\pi \cdot f_L \cdot t) \qquad (7)$$

$$N_s = A_H \cdot H(t-\tau) \qquad (8)$$

So far we got a noisy simulated spectrum as Eq. (9), which was similar to the real sampled spectral data, as shown in Fig. 2. The waveforms of this simulated PPG are depicted in Fig. 3 (m=120).

$$S^i_{Noisy} = A^i \cdot S_0 + N_L + \sum_{n=1}^{4} N_{S(n)} + N_{WG} + D,$$
$$i\ =\ 1\,,2\ \ldots\ m \qquad (9)$$

Subsequently, the *SC* of the simulated spectrum was calculated according to Eq. (3), and we extracted the "spectrum" by single trial estimation. The extraction accuracy was evaluated by the root-mean-square error (*RMSE*) between the normalized extracted signal and unit-amplitude sinusoidal signal.

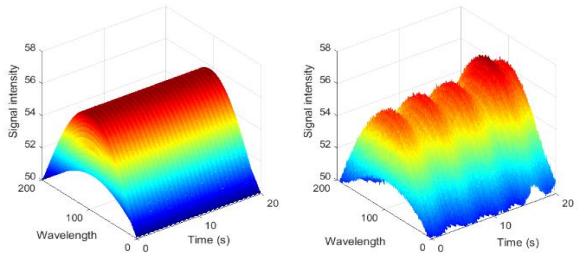

Fig. 2 Simulated spectrum (the simulated absorbance curve is sine wave). The left part shows the original signal and the right part shows the signal with noises. At any moment, the absorbance curve can be seen from the wavelength-axis.

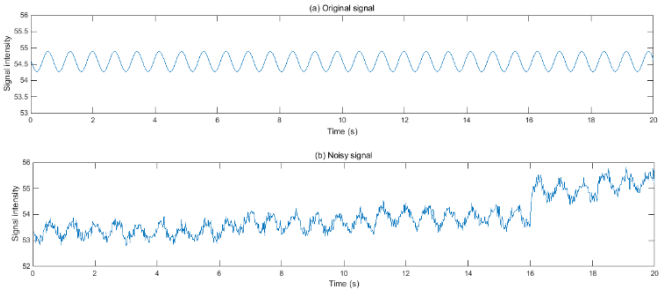

Fig. 3 Simulated signal at a certain wavelength (m=120) versus time.

The above steps were repeated 10 times, then we analysed their data quality (*SC*) and extraction errors (*RMSE*). As illustrated in Table 1, the results are sorted by SC value. It shows a negative correlation between *RMSE* and *SC*. In other words, the higher the *SC* value of a spectrum, the more accurately we can restore the spectrum. That means we can evaluate the quality of spectral data by its *SC* value. In addition, the data type we used in simulation was double-precision floating point (64 bits).

Table 1 Simulation results. The original simulated absorbance curve was a sine wave with unit amplitude and we could extract it perfectly. After adding noises to the signal, the original information was distorted, so the extracted signal can't be identical with the original sine wave and the extraction error will certainly increase as the noise intensity raise.

Simulation *No.*	5	10	6	1	2	8	3	9	4	7
SC	5.29	6.75	6.82	7.37	8.26	8.75	9.20	10.28	10.51	12.86
Extraction *RMSE*	0.0379	0.0216	0.0236	0.0171	0.0155	0.0117	0.0092	0.0084	0.0065	0.0063

Materials and methods

Experimental section

The spectral data acquiring system consisted of four major components --- a test bench with light source and optical collimators and optical fiber bundle, a power supply, a spectrometer, and a PC for controlling and data saving. As illustrated in Fig. 4, broadband light source was a tungsten-halogen lamp (*PHILIPS*, 50 *W*) and a stabilized DC power source supplied power for it. Collimated light beam made by lenses and diaphragms fell on the fingertip and the transmitted light was piped to the spectrometer through optical fiber bundle which contains several fibers with numerical aperture (*NA*) of 0.22. The spectrometer (*AvaSpec-HS1024x58TEC-USB2*, manufactured by *Avantes*, Holland), which has an adjustable

spectral resolution (1.5 *nm* – 20 *nm*), continuously scanned the transmission spectrum of fingertip tissue over visible and NIR waveband (200 *nm*-1160 *nm*) for a certain amount of times and the scans are stored in internal memory in real time. The spectrometer was controlled by PC, and after the last measurement had been made the spectrometer sent the data through USB to the PC.

We conducted a series of experiments in the Physical Examination Center of Tianjin People Hospital, from June 13, 2014 to July 1, 2014. Subjects of the experiments were recruited from the people who were going to accept a medical examination in the hospital. Before participating our study, the potential subjects were provided the general information about our research and were notified of the right to withdraw the consent to participate at any time, to ensure they could make an informed choice.

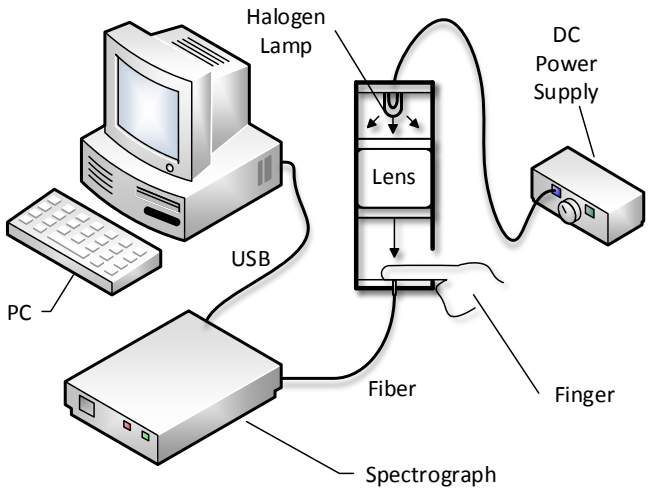

Fig. 4 Schematic diagram of the spectral data acquiring system

In every early morning, after on-site recruitment, volunteers had their blood drawn for a routine blood test which examined the main blood cells such as hemoglobin concentration, total WBC count, platelet count. The blood samples were tested by a fully automated hematology analyser (*ABX Pentra 60*, Manufactured by *HORIBA ABX SAS*, Japan) in the hospital. After blood sampling, the spectral data acquiring started immediately: we had the volunteer sit down quietly, and then the volunteer placed his or her right hand on the test bench with the middle finger pressing on the interface of optical fiber bundle, lightly and steadily. When the volunteer felt comfortable we activated the measurement of the spectrometer, as shown in Fig. 4, the narrow light beam was cast on the fingertip and the light transmitted by tissues was measured by the spectrometer. The Integration time of spectrometer was set to 20ms, which means that the sampling rate was 50 *Hz*. The number of scans for each subject was 1400, so the whole process lasted 28 seconds. As shown in Fig. 5, that is the PPG signal (20 *s*) of *No.* 32 volunteer. It's a time-varying signal which contains the transmission spectrum over a fixed wavelength range but at different moments. At the waveband of high signal intensity, we can identify the pulse wave clearly. After the collection of spectral data, we could extract DS from the raw data by existing method. Fig. 6 shows the DS of *NO.* 32 volunteer, which was extracted by single trial estimation. Obviously, the DS curve is consistent with the absorption spectrum of oxy-hemoglobin (HBO_2) at the same band [27], as painted in Fig.7 (spectrum curve in the dotted box).

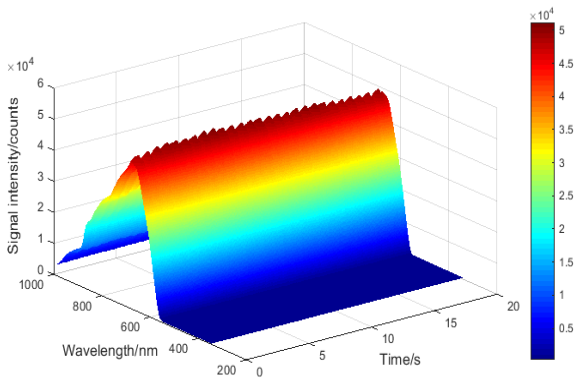

Fig. 5 One volunteer's 20s PPG signal (NO. 32)

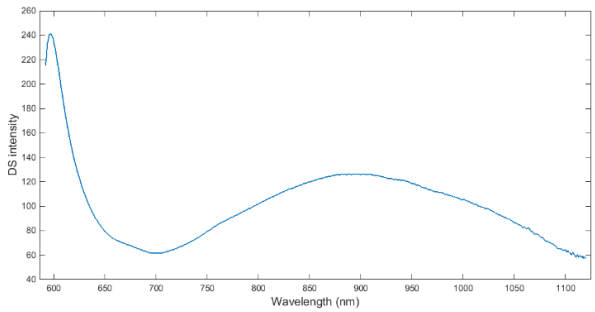

Fig. 6 Dynamic spectrum (DS) extracted by single trial estimation (*No.* 32 volunteer) over the wavelength range of 591.8nm-1120nm (altogether 586 wavelengths)

Altogether we had collected data from 609 subjects during the experiments. The general information of these volunteers are listed in Table 2. All the subjects were adults aged between 19 and 96, and all of them volunteered for this study.

All these experiments were conducted in compliance with the relevant laws as well as the guidelines issued by the Ethical Committee of Tianjin University and Tianjin people hospital.

Table 2 General information of the subjects

Gender	Maximum age	Minimum age	Average age	Sample size
Male	96	21	48	574
Female	91	19	36	35
Overall	96	19	45	609

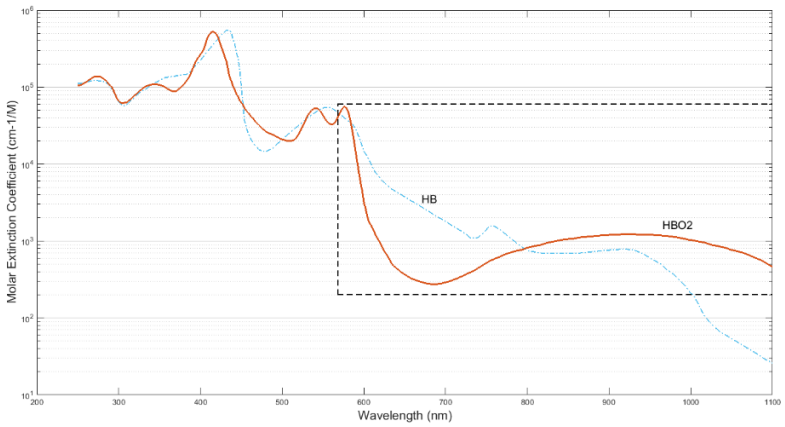

Fig. 7 Molar extinction coefficients of hemoglobin versus wavelength (250 *nm* – 1100 *nm*). It's easy to see that the DS curve (Fig. 6) is almost identical with the absorption coefficient curves of HBO2. Adapted from [27].

Data processing and analysis

After a preliminary screening, we discarded those invalid (saturated) spectra and picked out valid data of 427 subjects for further analysis.

First and foremost, we analysed the stability coefficient (*SC*) of these spectra and watched the waveforms, to explore the relation between *SC* value and the spectral quality. The initial observation is that higher *SC* value equate to better stability of PPG waveform and less noise. Generally, if the *SC* value equals 5 or more the signal is steady and the pulse wave is strong, while the quality of spectrum become very inferior if *SC* value is 2 or less, because such low *SC* value indicates that the fluctuation extent of PPG amplitudes outstripped half of its average level. Fig. 8 (a) shows a volunteer's PPG at 620 *nm*, which should have had a higher *SNR* than other wavelengths due to high absorbance of HGB, was affected by noise seriously and we can hardly recognize the pulse wave. Because there were much impulsive noise and strong high-frequency noise, almost the signal of every cardiac cycle became disparate and the calculation shows its *SC* value only equals 1.02. We extracted its DS, as depicted in Fig. 8 (b), due to the unsteadiness of original signal, the spectrum curve changes so rapidly even at adjacent wavelengths and the real information was distorted badly. The noisy spectrum didn't match the absorption properties of oxy-hemoglobin over the NIR range in Fig. 7 (especially at wavelengths of greater than 1000 *nm*, it distorted seriously). In contrast, Fig. 9 exhibits the corresponding data of a subject whose SC equals 10.7. We can see the pulse wave clearly, and the smoother DS curve of the later subject was attributed to less noise during the measurement process. Surely it contains information of more authentic and the net analyte signal [28] of hemoglobin in the spectrum is more accurate than the former.

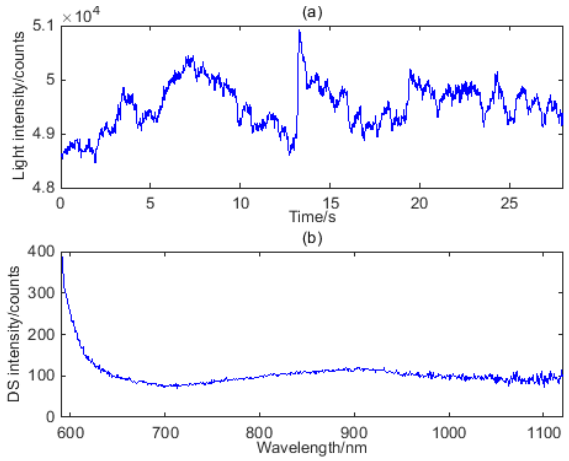

Fig. 8 One subject's PPG and DS at 620 *nm* (*SC*=1.02)

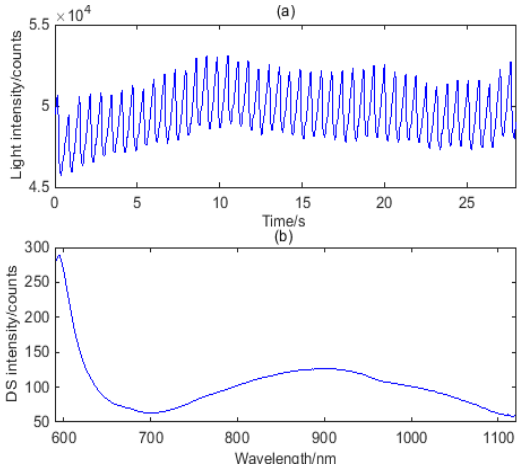

Fig. 9 One subject's PPG and DS at 620 *nm* (*SC*=10.7)

We analysed all the 427 spectra by calculating the stability coefficients of them and we found the *SC* was distributed between 0.429 and 16.535. A rough estimation of the probability density can be made by constructing a histogram of that 427 values, as shown in Fig. 10, the entire range of *SC* values was divided into 16 small intervals whose length is 1. The greater the abscissa of histogram, the better the data quality. The mass of the distribution is concentrated slightly on the left of the figure, and a somewhat positive-skewed distribution can be found. According to the above analysis and the distribution of *SC*, we divided all the 427 spectra into three groups: good data ($SC > 5$), average data ($2 < SC \leq 5$), and inferior data ($SC \leq 2$). The number of subjects in each group are 163, 230 and 34, respectively, and the percentages of them are 38%, 54%, and 8%. Statistics shows that the average data is in the majority, which is consistent with practical experiences.

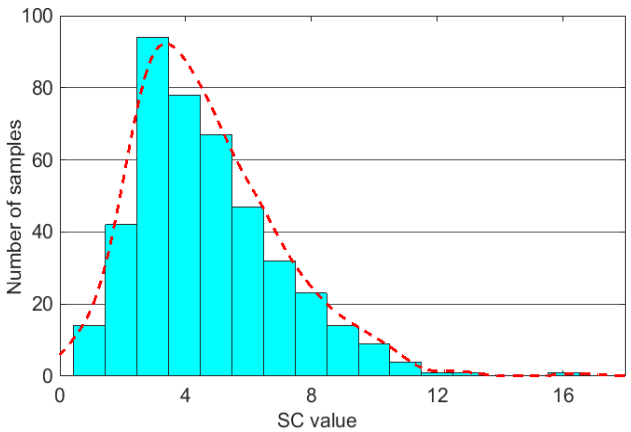

Fig. 10 Quality distribution of all the valid spectra

In order to further demonstrate the application performance of the SC assessment method, we conducted a control experiment. We established two calibration models with same structure, but the data for modelling were sampled by different ways. Because prediction performance is the most persuasive indicator to evaluate the data quality of not just the training set, but also the data in test set, the performance of the two predictive models was analysed to quantitatively compare the content of useful information in the data set, *i.e.*, the spectral quality. We extracted DS of all the 427 spectra over the wavelength range of 591.8 *nm*-1120 *nm* (altogether 586 wavelengths) by single trial estimation. Then we established the calibration models between DS data and the hemoglobin concentration measured by the hematology analyser.

Artificial Neural Network (*ANN*) is an excellent statistical learning algorithm which is good at regression or classification. It can be deemed as a system of connected "neurons" that have independent transfer-functions and can compute values from the input vector. The network has a strong capability of approximating or estimating the non-linear relation between inputs and outputs.

We established neural networks with four layers between DS and HGB (an input layer with 427 nodes, two hidden layers with 87 and 10 nodes respectively, and an output layer which has only one node corresponding to HGB concentration). The networks were trained by backpropagation algorithm in which the errors propagate backwards from the output layer to the input layer. The transfer function of our network is Tan-sigmoid function (the two hidden layers), linear transfer function (the output layer). We use MATLAB for network's establishing and training, and the training function of network was *traingdx*, which updates weight and bias values according to gradient descent momentum and an adaptive learning rate. In our networks the momentum constant was 0.9. The Learning rate was set to 0.01, and the ratio to increase and decrease learning rate was 1.05 and 0.7, respectively. The maximum number of epochs to train was 1000. The experiment was made up of two parts: experimental group and control group. As for experimental group in which we screened the data by *SC* value, 100 spectra with the corresponding HGB concentration were randomly sampled only from the "good data" group (*SC*>5), then 85 samples of them were used as calibration set while the other 15 samples were used for test (prediction). In control group, we randomly picked out 100 samples from all the valid data, regardless of the *SC* value. Similarly, the data were partitioned into 2 parts (training set and test set) in proportion to the experimental group.

In order to demonstrate the effectiveness, and more important, the reliability of the metric, we randomly sampled data and performed this control experiment for 10 times in total. Then we evaluated the networks by the key indicators: *RMSE* and the correlation coefficient.

Results and Discussion

The Results of these networks are listed in Table 3, in which R_T and R_P represent respectively the correlation coefficient of training set and test set. *RMSE* is the root-mean-square error of test set. The average values of their *RMSE* and correlation coefficients are listed in Table 4.

The average *SC* of the spectra of experimental groups is 7.124 and that of control group is 4.559. The experiment result shows that the $\overline{R_P}$ in control group was 0.715 while that in experimental group was 0.875, which was much greater than control group. The significant difference between the two groups shows that spectra with higher *SC* value did improve the performance of calibration model with same topology. It is because on the one hand, higher data quality of training set can make the generalization ability stronger and provide more authentic information for calibrating, which can be seen from the difference of the correlation coefficient of training set (The disparity of $\overline{R_T}$ is not as remarkable as $\overline{R_P}$, because the artificial neural network algorithm have a strong ability of detecting the complex nonlinear relationship between the inputs and outputs,

i.e., even if the training set is not that ideal, it can get a high R_T , but in such situations the calibration model try to approximate to random errors or noises rather than the underlying relationship between the inputs and outputs). On the other hand, predictive samples with higher quality make the model to calculate dependent variable from the independent variables more accurately.

Table 3 Result of control experiment

Experiment *No.*	Experimental Group				Control Group			
	$\overline{SC}$	R_T	R_p	*RMSE* ($g\ L^{-1}$)	$\overline{SC}$	R_T	R_p	*RMSE* ($g\ L^{-1}$)
#1	7.895	0.968	0.931	3.10	3.968	0.925	0.690	5.78
#2	7.408	0.950	0.900	3.78	3.514	0.944	0.637	6.13
#3	5.982	0.989	0.801	4.68	4.623	0.934	0.700	5.24
#4	7.854	0.981	0.820	4.56	4.211	0.917	0.705	5.01
#5	8.048	0.987	0.942	3.17	5.650	0.961	0.600	5.96
#6	6.461	0.931	0.859	4.47	4.169	0.901	0.640	5.97
#7	5.673	0.961	0.780	4.92	6.333	0.965	0.861	3.23
#8	7.009	0.981	0.925	3.65	4.423	0.932	0.792	4.99
#9	6.803	0.946	0.860	3.56	3.781	0.929	0.721	5.12
#10	8.108	0.980	0.927	3.11	4.913	0.930	0.800	4.84

Table 4 Comparison of experimental and control group

Group	$\overline{SC}$	$\overline{R_T}$	$\overline{R_P}$	*RMSE* ($g\ L^{-1}$)
Experimental	7.124	0.967	0.875	3.90
Control	4.559	0.934	0.715	5.23

The above-mentioned simulation as well as the control experiment jointly demonstrate that the stability coefficient of spectral data can characterize the quality of them, because in essence, *SC* represent the concentration of distribution of PPG amplitudes. Stable PPG wave is a critical element in numerous studies of related areas, *e.g.*, heart rate measurement, oximetry [29-30], blood pressure estimation based on pulse transit time (PTT) [31]. Actually, variability analysis is a practical and effective approach to evaluate the quality of periodic signal or repeated measurement, because a good repeatability of measurement is always good for us (unless the variation of signal or measurement results just happen to be the object that we study on, like the variation of pulse-to-pulse intervals or "R-R intervals" in ECG is an important indicator of the patients' health condition [32-33], which is also a hotspot in biomedical research). And the method we proposed in this paper is applicable not only to spectral data assessment, but also many related areas. For instance, if we intend to evaluate the stability of continuous scanning of a spectrometer, or time characteristic of a dynamic sampling instrument, this method may provide references.

It should be noted that this study has only analysed the variation of amplitudes because the amplitude of PPG is the critical information in DS method, with no consideration of the variation in time dimension, such as the periodicity, which can be further analysed in other studies. Moreover, the number of wavelengths used for calculation should be adjusted according to the specific contents of research, for instance, if the data quality of pulse oximetry need to be evaluated, it's sufficient for researchers to analyse at only two or three wavelengths (depending on the combination of wavelengths).

Conclusion

A novel method based on variability analysis to evaluate the spectral quality in noninvasive hemoglobin measurement has been proposed. Through simulation and actual experiment, its feasibility and effectiveness have been proved: a prominent positive correlation between stability coefficient and extraction accuracy and the contrast between the results of two predictive models demonstrate that from different angles. In related studies which needs dynamic measurement, this method can be applied to assess the spectral quality and screen data before establishing calibration model, and the screening criteria can be changed to adapt to specific application, because the indicator is quantitative. This study has given a new way to improve the accuracy of spectral analysis, and provided a new idea for the spectral measurements and analysis which demands a stable time characteristic.

Acknowledgements

We are grateful to the help from the Physical Examination Center of Tianjin People Hospital. This study was supported by the National Natural Science Foundation of China (*No*. 30973964), and the Tianjin Application Basis & Front Technology Study Programs (*No*. 14JCZDJC33100).

Notes and References

[a] State Key Laboratory of Precision Measurement Technology and Instruments, Tianjin University, Tianjin 300072
[b] Tianjin Key Laboratory of Biomedical Detecting Techniques & Instruments, Tianjin University, Tianjin 300072
[c] School of Electrical Engineering, Hebei University of Technology, Tianjin, 300130

[1] Kraitl J, Klinger D, Advancement in Sensing Technology. 2013: 237-262.
[2] Zanon M, Sparacino G, Facchinetti A, et al. Medical & biological engineering & computing, 2012, 50(10): 1047-1057.
[3] Vashist S K. Analytica chimica acta, 2012, 750: 16-27.
[4] Kraitl J, Ewald H, Gehring H. Journal of Optics A: Pure and Applied Optics, 2005, 7(6): S318.
[5] McMurdy J, Jay G, Suner S, et al. Journal of biophotonics, 2009, 2(5): 277-287.
[6] Sun M X, Zhang L Y, Jiang C Y, et al. Optik-International Journal for Light and Electron Optics, 2015, 126(4): 460-463.
[7] Chen P, Fernald B, Lin W. Physics in medicine and biology, 2011, 56(13): 3985.
[8] Bender J E, Shang A B, Moretti E W, et al. Optics express, 2009, 17(26): 23396-23409.
[9] Janz T G, Johnson R L, Rubenstein S D. Emergency medicine practice, 2013, 15(11): 1-15.
[10] He Y, Tang L, Wu X, et al. Applied Spectroscopy Reviews, 2007, 42(2): 119-138.
[11] Wang H, Li G, Zhao Z, et al. Transactions of the Institute of Measurement and Control, 2011: 0142331211411965.
[12] Li G, Wang Y, Li Q X, et al. Journal of Infrared and Millimeter Waves, 2006, 25(5): 345-348.
[13] Allen J. Physiological measurement, 2007, 28(3): R1.
[14] Zhou M, Li G, Lin L. Journal of biomedical optics, 2013, 18(5): 057003-057003.
[15] Li G, Zhou M, Lin L. Optical and Quantum Electronics, 2014, 46(5): 691-698.
[16] LI G, ZHANG H, LIN L, et al. Spectroscopy and Spectral Analysis, 2012, 32(2): 486-490.
[17] LI G, ZHAO Z, LIN L. Increasing the Precision of the Noninvasive Blood Components Measurement Based on DS Method Using Harmonic Waves[J]. Spectroscopy and Spectral Analysis, 2010, 30(9): 2385-2389.
[18] LI G, ZHAO Z, LIN L, et al. Spectroscopy and Spectral Analysis, 2010, 30(10): 2802-2806.
[19] Atkinson G, Nevill A M. Sports medicine, 1998, 26(4): 217-238.
[20] Shechtman O. Methods of Clinical Epidemiology. Springer Berlin Heidelberg, 2013: 39-49.
[21] Wang B, Goodpaster A M, Kennedy M A. Chemometrics and Intelligent Laboratory Systems, 2013, 128: 9-16.
[22] Reed G F, Lynn F, Meade B D. Clinical and diagnostic laboratory immunology, 2002, 9(6): 1235-1239.
[23] Allen J. Physiological measurement, 2007, 28(3): R1.
[24] Delpy D T, Cope M, Van der Zee P, et al. Physics in medicine and biology, 1988, 33(12): 1433.
[25] Li G, Li Q X, Lin L, et al. Spectroscopy and Spectral Analysis, 2006, 26(12): 2177-2180.
[26] LI G, XIONG C, LIN L, et al. Spectroscopy and Spectral Analysis, 2011, 31(7): 1857-1861.
[27] Delpy D T, Cope M. Philosophical Transactions of the Royal Society B: Biological Sciences, 1997, 352(1354): 649-659.
[28] Lorber A. Analytical Chemistry, 1986, 58(6): 1167-1172.
[29] Pedersen T, Nicholson A, Hovhannisyan K, et al. The Cochrane Library, 2014.
[30] Silva S M L, Castilla M L D, Martin J P S. Journal of Biomedical Optics, 2003, 8(3): 525-533.
[31] Chang C, Metzger C D, Glover G H, et al. Neuroimage, 2013, 68: 93-104.
[32] Berntson G G, Bigger J T, Eckberg D L, et al. Psychophysiology, 1997 (34): 623-48.
[33] Thayer J F, Åhs F, Fredrikson M, et al. Neuroscience & Biobehavioral Reviews, 2012, 36(2): 747-756.