Claims

1. A system for estimating vitals of a person from a short sequence of RGB and NIR images, no more than 5 seconds, and up-sampling the raw signal to interpolate it and estimate vitals rapidly, comprising of~~comprising~~:

an image sensor adapted to capture a sequence of color (RGB) and near infrared (NIR) images;

a processing unit comprising a processor, memory and instructions stored on the memory and executable by the processor to:

estimating vitals of the person from the sequence of RGB and NIR images~~.~~

~~2. The system according to claim 1 ,~~wherein the instruction to estimating vitals of the person from the sequence of RGB and NIR images comprises instructions to:

time synchronising the sequence of RGB and NIR images;

detecting region of interest within the sequence of RGB and NIR images;

extracting raw signals from the detected region of interest to form a time series of predetermined length L for each frequency band;

applying a moving average filter to a pixel intensity value in time series for each frequency band, the moving average filter has a window length of $\alpha_1 * L$;

up-sampling the pixel intensity value in the time series for each frequency band using cubic spline interpolation method to form a new length, L_2, where $L_2 = \alpha_2 * L$;

applying a detrending algorithm based on smoothness priors' formulation for the upsampled pixel intensity value in the time series for each frequency band wherein a smoothness factor, λ, of the detrending algorithm is set based on the following expression, $\lambda = \alpha_3 * L_2$;

applying Independent Component Analysis (ICA) to extract PPG signal from the detrended pixel intensity value in the time series for each frequency band, wherein values of $\alpha_1, \alpha_2, \alpha_3$ & L are calibrated based on the imaging sensor

extracting photoplethysmography (PPG) signals from the raw signals; and

estimating vitals of the person from the extracted PPG signals.

23. The system according to claim 12 wherein the instruction to detecting region of interest within the sequence of RGB and NIR images comprises instructions to:

detecting a face of the person within each of the sequence of NIR images;

in response to detecting the face of the person, detecting at least one region of interest within the detected face of the person; and

mapping the detected region of interest in each of the sequence of NIR images to the corresponding sequence of RGB images.

34 The system according to claim 23 wherein the instruction to detecting at least one region of interest within the detected face of the person comprises instructions to:

applying facial landmark detection where 81 prominent feature points on the face the person are detected; and

overlaying a box at a top of the detected 81 prominent feature points between facial edge and eyebrows, wherein the box is one region of interest.

45 The system according to claim 34 wherein the instruction to detecting at least one region of interest within the detected face of the person further comprises instructions to:

overlaying two boxes between facial edge and nose, wherein the two boxes are two other region of interests.

5~~6~~. The system according to any one of claims 2-4~~3-5~~ wherein the instruction to detecting region of interest within the sequence of RGB and NIR images comprises instructions to:

computing an average pixel intensity of the detected region of interest for each of a plurality of frequency bands including Red, Green and Blue from the sequence of RGB images and NIR frequency band from the sequence of NIR images; and

appending the average pixel intensity value of each frequency band to the pixel intensity value of corresponding frequency band in respective RGB and NIR images to form a time series for each frequency band.

6~~7~~. The system according to claim 5~~6~~ wherein the instruction to detecting region of interest within the sequence of RGB and NIR images further comprises instructions to:

determining if a length of the time series is equal to a predetermined length, *L*;

in response to the length of the time series being equal to the predetermined length, *L*, normalising the pixel intensity value in the time series for each frequency band.

7~~8~~. The system according to claim 6~~7~~ wherein *L* is no more than 5 seconds~~the instruction to extracting PPG signals from the raw signals comprises instructions to:~~

~~applying a moving average filter to the pixel intensity value in the time series for each frequency band, the moving average filter has a window length of $\alpha_1 * L$;~~

~~up-sampling the pixel intensity value in the time series for each frequency band using cubic spline interpolation method to form a new length, L_2, where $L_2 = \alpha_2 * L$;~~

~~applying a detrending algorithm based on smoothness priors' formulation for the upsampled pixel intensity value in the time series for each frequency band wherein a smoothness factor, λ, of the detrending algorithm is set based on the following expression, $\lambda = \alpha_3 * L_2$;~~ ~~and~~

~~applying Independent Component Analysis (ICA) to extract PPG signal from the detrended pixel intensity value in the time series for each frequency band, wherein values of $\alpha_1, \alpha_2, \alpha_3$ & L are calibrated based on the imaging sensor~~.

8~~9~~. The system according to claim 6~~8~~ wherein the instruction to applying Independent Component Analysis (ICA) to extract PPG signal from the detrended pixel intensity value in the time series for each frequency band comprises instructions to:

grouping the detrended pixel intensity value in the time series for each frequency band into a set of Red PPG signal and a set of NIR PPG signals wherein the Red PPG signal is derived from the frequency bands of Red, Green and Blue and the NIR PPG signal is derived from the frequency bands of NIR, Green and Blue;

applying a band pass filter with cutoff frequency of between 0.16 and 2 Hz to the Red PPG signal; and

applying a band pass filter with cutoff frequency of between 0.16 and 2 Hz to the NIR PPG signal.

9~~10~~. The system according to claim 8~~9~~ wherein the instruction to estimating vitals of the person from the extracted PPG signals comprises instructions to use a machine learning model based on a linear regression model.

10~~1~~. The system according to claim 8~~9~~ wherein the instruction to estimating vitals of the person from the extracted PPG signals comprises instructions to:

computing Ratio (R) of time-domain range for NIR PPG signal and Red PPG signal wherein the Ratio (R) can be expressed with the following expression,

$$Ratio\ (R) = \frac{Amplitude\ range\ of\ Red\ PPG\ signal \Big/ Arithmetic\ mean\ of\ Red\ PPG\ signal}{Amplitude\ range\ of\ NIR\ PPG\ signal \Big/ Arithmetic\ mean\ of\ NIR\ PPG\ signal}$$

where amplitude range of Red PPG signal relates to difference between minimum and maximum values of the amplitude of the Red PPG signal in time domain, arithmetic mean of Red PPG signal relates to arithmetic mean of the Red PPG signal in between minimum and maximum values of the amplitude of the Red PPG signal in time domain, amplitude range of NIR PPG signal relates to difference between minimum and maximum values of the amplitude of the NIR PPG signal in time domain and arithmetic mean of NIR PPG signal relates to arithmetic mean of the NIR PPG signal in between minimum and maximum values of the amplitude of the NIR PPG signal in time domain; and

estimating the Blood Oxygen Saturation with the following expression,

$$Blood\ Oxygen\ Saturation\ (SPO_2) = \beta_1 + \beta_2 * R$$

where β_1 refers to off-set parameter and β_2 refers to slope parameter of the linear regression curve which maps the Ratio (R) to the blood oxygen saturation (SPO_2).

11~~2~~. The system according to claim 8~~9~~ or 9~~10~~ wherein the instruction to estimating vitals of the person from the extracted PPG signals comprises instructions to:

detecting Systolic Peak (P_s), Diastolic peak (P_d), and global minima (P_m) from the time-domain NIR PPG signal;

estimating the Systolic and Diastolic Blood Pressure with the following expressions,

$$\text{Systolic Blood Pressure (SBP)} = e^{(\gamma_1 * \log(HR) + \gamma_2 * \log(P_s - P_m) + \gamma_3)}$$

$$\text{Diastolic Blood Pressure (DBP)} = e^{(\gamma_4 * \log(HR) + \gamma_5 * \log(P_d - P_m) + \gamma_6)}$$

where $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6$ are parameters of the regression curve which maps the Heart Rate and NIR PPG signal amplitude features to the systolic and diastolic blood pressure, $\gamma_1\ and\ \gamma_4$ are frequency-slope parameters for systolic and diastolic blood pressures respectively, $\gamma_2\ and\ \gamma_5$ are amplitude-slope parameters for systolic and diastolic blood pressures respectively, $\gamma_3\ and\ \gamma_6$ are off-set parameters for systolic and diastolic blood pressures respectively.

12~~3~~. The system according to claim 11~~2~~ wherein the instruction to detecting Systolic Peak (P_s), Diastolic peak (P_d), and global minima (P_m) from the time-domain NIR PPG signal comprises instructions to:

applying a peak detection algorithm on the NIR PPG signal to obtain a second order differential;

identifying a global maxima in the NIR PPG signal as the Systolic Peak (P_s), a peak in the NIR PPG signal corresponding to a minima following the Systolic Peak (P_s) in the second order derivative as Diastolic peak (P_d), a maxima immediately after the systolic peak in the second order derivative as the Diatric notch, a minima of the entire NIR PPG signal as the global minima (P_m).

13~~4~~. The system according to any one of claims 6-12~~7-13~~ wherein the instruction to estimating vitals of the person from the extracted PPG signals comprises instructions to:

applying fast fourier transform to the NIR PPG signal to transform the NIR PPG signal from time-domain to frequency domain;

computing a plurality of local maxima of a power spectral density (PSD) of the NIR PPG signal in frequency domain;

in response to a local maxima between the frequency range of 0.8 and 2 Hz, estimating a Heart Rate (HR) with the following expression,

$$Heart\ Rate\ (HR) = 60 * f_{hr}$$

where f_{hr} refers to the local maxima that is between the frequency range of 0.8 and 2 Hz; and

in response to a local maxima between the frequency range of 0.1 and 0.5 Hz, estimating a Respiration Rate (RR) with the following expression,

$$Respiration\ Rate\ (RR) = 60 * f_{rr}$$

where f_{rr} refers to the local maxima that is between the frequency range of 0.1 and 0.5 Hz.

14~~5~~. A method of estimating vitals of a person from a short sequence of RGB and NIR images, no more than 5 seconds, and up-sampling the raw signal to interpolate it and estimate vitals rapidly, comprising of: ~~from a sequence of RGB and NIR images~~ comprising:

time synchronising the sequence of RGB and NIR images;

detecting region of interest within the sequence of RGB and NIR images;

extracting raw signals from the detected region of interest;

~~extracting photoplethysmography (PPG) signals from the raw signals~~applying a moving average filter to a pixel intensity value in time series for each frequency band, the moving average filter has a window length of $\alpha_1 * L$;

up-sampling the pixel intensity value in the time series for each frequency band using cubic spline interpolation method to form a new length, L_2, where $L_2 = \alpha_2 * L$;

applying a detrending algorithm based on smoothness priors' formulation for the upsampled pixel intensity value in the time series for each frequency band wherein a smoothness factor, λ, of the detrending algorithm is set based on the following expression, $\lambda = \alpha_3 * L_2$;

applying Independent Component Analysis (ICA) to extract PPG signal from the detrended pixel intensity value in the time series for each frequency band, wherein values of $\alpha_1, \alpha_2, \alpha_3$ & L are calibrated based on the imaging sensor; and

estimating vitals of the person from the extracted PPG signals.

15~~6~~. The method according to claim 14~~5~~ wherein the step of detecting region of interest within the sequence of RGB and NIR images comprises:

detecting a face of the person within each of the sequence of NIR images;

in response to detecting the face of the person, detecting at least one region of interest within the detected face of the person; and

mapping the detected region of interest in each of the sequence of NIR images to the corresponding sequence of RGB images.

16~~7~~. The method according to claim 15~~6~~ wherein the step of detecting at least one region of interest within the detected face of the person comprises:

applying facial landmark detection where 81 prominent feature points on the face the person are detected; and

overlaying a box at a top of the detected 81 prominent feature points between facial edge and eyebrows, wherein the box is one region of interest.

17~~8~~. The method according to claim 16~~7~~ wherein the step of detecting at least one region of interest within the detected face of the person further comprises:

overlaying two boxes between facial edge and nose, wherein the two box are two other region of interests.

18~~9~~. The method according to any one of claims 15-17~~16-18~~ wherein the step of detecting region of interest within the sequence of RGB and NIR images comprises:

computing an average pixel intensity of the detected region of interest for each of a plurality of frequency bands including Red, Green and Blue from the sequence of RGB images and NIR frequency band from the sequence of NIR images; and

appending the average pixel intensity value of each frequency band to the pixel intensity value of corresponding frequency band in respective RGB and NIR images to form a time series for each frequency band.

19~~20~~. The method according to claim 18~~9~~ wherein the step of detecting region of interest within the sequence of RGB and NIR images further comprises:

determining if a length of the time series is equal to a predetermined length, *L*;

in response to the length of the time series being equal to the predetermined length, *L*, normalising the pixel intensity value in the time series for each frequency band.

20~~1~~. The method according to claim 19~~20~~ wherein *L* is no more than 5 seconds~~the step of extracting PPG signals from the raw signals comprises:~~

~~applying a moving average filter to the pixel intensity value in the time series for each frequency band, the moving average filter has a window length of $\alpha_1 * L$;~~

~~up-sampling the pixel intensity value in the time series for each frequency band using cubic spline interpolation method to form a new length, L_2, where $L_2 = \alpha_2 * L$;~~

~~applying a detrending algorithm based on smoothness priors' formulation for the upsampled pixel intensity value in the time series for each frequency band wherein a smoothness factor, λ, of the detrending algorithm is set based on the following expression, $\lambda = \alpha_3 * L_2$; and~~

~~applying Independent Component Analysis (ICA) to extract PPG signal from the detrended pixel intensity value in the time series for each frequency band, wherein values of $\alpha_1, \alpha_2, \alpha_3$ & L are calibrated based on the imaging sensor~~.

21~~2~~. The method according to claim 19~~21~~ wherein the step of applying Independent Component Analysis (ICA) to extract PPG signal from the detrended pixel intensity value in the time series for each frequency band comprises:

grouping the detrended pixel intensity value in the time series for each frequency band into a set of Red PPG signal and a set of NIR PPG signals wherein the Red PPG signal is derived from the frequency bands of Red, Green and Blue and the NIR PPG signal is derived from the frequency bands of NIR, Green and Blue;

applying a band pass filter with cutoff frequency of between 0.16 and 2 Hz to the Red PPG signal; and

applying a band pass filter with cutoff frequency of between 0.16 and 2 Hz to the NIR PPG signal.

223. The method according to claim 212 wherein the step of estimating vitals of the person from the extracted PPG signals comprises:

computing Ratio (R) of time-domain range for NIR PPG signal and Red PPG signal wherein the Ratio (R) can be expressed with the following expression,

$$Ratio\ (R) = \frac{Amplitude\ range\ of\ Red\ PPG\ signal \Big/ Arithmetic\ mean\ of\ Red\ PPG\ signal}{Amplitude\ range\ of\ NIR\ PPG\ signal \Big/ Arithmetic\ mean\ of\ NIR\ PPG\ signal}$$

where amplitude range of Red PPG signal relates to difference between minimum and maximum values of the amplitude of the Red PPG signal in time domain, arithmetic mean of Red PPG signal relates to arithmetic mean of the Red PPG signal in between minimum and maximum values of the amplitude of the Red PPG signal in time domain, amplitude range of NIR PPG signal relates to difference between minimum and maximum values of the amplitude of the NIR PPG signal in time domain and arithmetic mean of NIR PPG signal relates to arithmetic mean of the NIR PPG signal in between minimum and maximum values of the amplitude of the NIR PPG signal in time domain;

estimating the Blood Oxygen Saturation with the following expression,

$$Blood\ Oxygen\ Saturation\ (SPO_2) = \beta_1 + \beta_2 * R$$

where β_1 refers to off-set parameter and β_2 refers to slope parameter of the linear regression curve which maps the Ratio (R) to the blood oxygen saturation (SPO_2).

234. The method according to claim 212 or 223 wherein the step of estimating vitals of the person from the extracted PPG signals comprises:

detecting Systolic Peak (P_s), Diastolic peak (P_d), and global minima (P_m) from the time-domain NIR PPG signal;

estimating the Systolic and Diastolic Blood Pressure with the following expressions,

$$\text{Systolic Blood Pressure (SBP)} = e^{(\gamma_1 * \log(HR) + \gamma_2 * \log(P_s - P_m) + \gamma_3)}$$

$$\text{Diastolic Blood Pressure (DBP)} = e^{(\gamma_4 * \log(HR) + \gamma_5 * \log(P_d - P_m) + \gamma_6)}$$

where $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6$ are parameters of the regression curve which maps the Heart Rate and NIR PPG signal amplitude features to the systolic and diastolic blood pressure, $\gamma_1 \; and \; \gamma_4$ are frequency-slope parameters for systolic and diastolic blood pressures respectively, $\gamma_2 \; and \; \gamma_5$ are amplitude-slope parameters for systolic and diastolic blood pressures respectively, $\gamma_3 \; and \; \gamma_6$ are off-set parameters for systolic and diastolic blood pressures respectively.

245. The method according to claim 234 wherein the step of detecting Systolic Peak (P_s), Diastolic peak (P_d), and global minima (P_m) from the time-domain NIR PPG signal comprises:

applying a peak detection algorithm on the NIR PPG signal to obtain a second order differential;

identifying a global maxima in the NIR PPG signal as the Systolic Peak (P_s), a peak in the NIR PPG signal corresponding to a minima following the Systolic Peak (P_s) in the second order derivative as Diastolic peak (P_d), a maxima immediately after the systolic peak in the second order derivative as the Diatric notch, a minima of the entire NIR PPG signal as the global minima (P_m).

256. The method according to any one of claims 19-2420-25 wherein the step of estimating vitals of the person from the extracted PPG signals comprises:

applying fast fourier transform to the NIR PPG signal to transform the NIR PPG signal from time-domain to frequency domain;

computing a plurality of local maxima of a power spectral density (PSD) of the NIR PPG signal in frequency domain;

in response to a local maxima between the frequency range of 0.8 and 2 Hz, estimating a Heart Rate (HR) with the following expression,

$$Heart\ Rate\ (HR) = 60 * f_{hr}$$

where f_{hr} refers to the local maxima that is between the frequency range of 0.8 and 2 Hz; and

in response to a local maxima between the frequency range of 0.1 and 0.5 Hz, estimating a Respiration Rate (RR) with the following expression,

$$Respiration\ Rate\ (RR) = 60 * f_{rr}$$

where f_{rr} refers to the local maxima that is between the frequency range of 0.1 and 0.5 Hz.

26~~7~~. The method according to claim 21~~2~~ wherein the step of estimating vitals of the person from the extracted PPG signals comprises applying a machine learning model based on a linear regression model.