



EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:

12.06.2019 Bulletin 2019/24

(21) Application number: **15808020.0**

(22) Date of filing: **08.12.2015**

(51) Int Cl.:

A61B 5/0295 ^(2006.01) **A61B 5/00** ^(2006.01)
A61B 5/026 ^(2006.01) **G06T 7/00** ^(2017.01)
A61B 5/08 ^(2006.01) **G06K 9/46** ^(2006.01)
A61B 5/0205 ^(2006.01) **A61B 5/024** ^(2006.01)
A61B 5/11 ^(2006.01) **A61B 5/113** ^(2006.01)

(86) International application number:
PCT/GB2015/053754

(87) International publication number:
WO 2016/092290 (16.06.2016 Gazette 2016/24)

(54) **METHOD AND APPARATUS FOR PHYSIOLOGICAL MONITORING**

VERFAHREN UND VORRICHTUNG ZUR PHYSIOLOGISCHEN ÜBERWACHUNG

PROCÉDÉ ET APPAREIL DE SURVEILLANCE PHYSIOLOGIQUE

(84) Designated Contracting States:

**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB
GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO
PL PT RO RS SE SI SK SM TR**

(30) Priority: **08.12.2014 GB 201421785**

(43) Date of publication of application:
18.10.2017 Bulletin 2017/42

(73) Proprietor: **Oxehealth Limited**
The Oxford Science Park
Oxford
OX4 4GA (GB)

(72) Inventor: **GUAZZI, Alessandro**
Oxford
Oxfordshire OX3 7DQ (GB)

(74) Representative: **J A Kemp**
14 South Square
Gray's Inn
London WC1R 5JJ (GB)

(56) References cited:

WO-A1-2011/080189 WO-A1-2014/009111
WO-A1-2014/155750 US-A1- 2010 016 693
US-A1- 2014 073 954

- **TARASSENKO L ET AL: "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models", PHYSIOLOGICAL MEASUREMENT, INSTITUTE OF PHYSICS PUBLISHING, BRISTOL, GB, vol. 35, no. 5, 28 March 2014 (2014-03-28) , pages 807-831, XP020262371, ISSN: 0967-3334, DOI: 10.1088/0967-3334/35/5/807 [retrieved on 2014-03-28]**
- **YUNYOUNG NAM ET AL: "Respiratory Rate Estimation from the Built-in Cameras of Smartphones and Tablets", ANNALS OF BIOMEDICAL ENGINEERING., vol. 42, no. 4, 23 November 2013 (2013-11-23), pages 885-898, XP055247279, GB ISSN: 0090-6964, DOI: 10.1007/s10439-013-0944-x**
- **MARTIN-MARTINEZ DIEGO ET AL: "Stochastic Modeling of the PPG Signal: A Synthesis-by-Analysis Approach With Applications", IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, IEEE SERVICE CENTER, PISCATAWAY, NJ, USA, vol. 60, no. 9, 1 September 2013 (2013-09-01), pages 2432-2441, XP011524036, ISSN: 0018-9294, DOI: 10.1109/TBME.2013.2257770 [retrieved on 2013-08-16]**
- **D MARTIN-MARTINEZ ET AL: "Cardiovascular signal reconstruction based on shape modelling and non-stationary temporal modelling", 2006 14TH EUROPEAN SIGNAL PROCESSING CONFERENCE, 1 January 2012 (2012-01-01), pages 1826-1830, XP055247266, ISSN: 2219-5491**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Description

[0001] The present invention relates to the remote monitoring of a human or animal subject to detect periodic physiological signals such as the heart rate or breathing rate.

[0002] The detection of periodic physiological signals, such as heart rate or breathing (respiration) rate, of a subject is a vital part of patient monitoring. Traditionally contact sensors have been used so that the heart rate can be detected by means of an ECG monitor or PPG finger probe and the breathing rate can either be derived from the ECG or PPG signals or measured directly by impedance pneumography. A PPG finger probe also gives the blood oxygen saturation (SpO₂). However contact-based sensors are not ideal for monitoring patients who move or are mobile, or for long-term monitoring. More recently, therefore, proposals have been made for non-contact monitoring. One such proposal is to derive measures of the heart rate and breathing rate from a photoplethysmographic image (PPGi). As discussed in the applicant's own published patent application WO-A2-2013/027027, a video image of the patient's skin can be analysed to detect the variation in transmittance or reflectance of light at certain wavelengths as the volume of blood in the skin capillaries varies with the cardiac cycle. Although invisible to normal sight, the skin can effectively be regarded as pulsing more red and less red with each heart beat. This change can be detected in a standard colour video image of the skin taken with a normal video camera such as a webcam. WO-A2-2013/027027 explains how the heart rate and breathing rate can be derived from such a colour video signal.

[0003] Figure 1 of the accompanying drawings schematically illustrates the PPGi analysis process. In step 100 a standard video signal with 3 colour channels (red, green, blue) is received, for example from a webcam of the type found in a tablet computer or smartphone, and in step 101 a region of interest positioned over bare skin in the image is defined and for each of the three colour channels, at each time point, the mean of the intensities over that region of interest is obtained. This gives a time-series of values for each colour, each value being the mean intensity for that colour over the whole region of interest. In step 102 the green channel is selected, as this has the most cardiac information. In step 103 the signal is divided into short overlapping time windows. Each time window is 15 seconds long and the window slides by one second at a time. Ultimately a heart rate or breathing rate estimate will be outputted for each window, thus one estimate will be outputted every second. In step 104 the time series in each window is detrended (by removing the best fit straight line from the data) and then filtered, e.g. using a zero-phase finite impulse response (FIR) filter of order 30, with cut-offs between 0.7 Hz and 2 Hz, corresponding to heart rates of 42 beats per minute and 120 beats per minute respectively - the usual range of adult resting heart rates.

[0004] The aim is then to spectrally analyse the signal to find a periodic variation corresponding to the heart rate (or breathing rate). In this prior art this is done in steps 105 to 107 by fitting autoregressive models to the signal. Thus in step 105 several autoregressive models of different orders are fitted to the signal (for example orders 8 to 20). As explained in WO-A2-2013/027027 autoregressive modelling can be regarded as filtering with an all-pole infinite impulse response filter (IIR) with a white noise input in which the number of poles is determined by the model order. In spectral analysis it is the properties of these poles which are of interest. The poles can be viewed as points lying in the complex plane in an area conscribed by the unit circle, having a phase angle that is related to a pure frequency component of the signal and a magnitude that is related to the strength of that component. The analysis of the signals in step 105 generates, for each model order, several poles representing different frequencies present in the signal. In order to find which pole represents the heart rate (or breathing rate) poles representing frequencies outside the allowed range for heart rate (or breathing rate) can be ignored and then the pole with the largest magnitude and/or lowest frequency, can be selected. The phase angle of the selected pole corresponds to a particular frequency - assumed to correspond to the heart rate estimate from that model. The estimates from each of the different model orders are taken and averaged in step 107, and in step 108 the heart rate estimate based on that average is output.

[0005] It will be appreciated that because the frequency estimate from each model order is based on the selected pole, the selection of that pole is of critical importance. However it can be difficult, in practice, to select the correct poles from the fitted autoregressive models. Selection of the wrong pole will result in an erroneous frequency estimate. Further, although it is anticipated that remote monitoring will be better than contact-based sensors in the case of moving subjects, movement of the subject in the video image can create movement artefacts which make signal processing difficult. It is very well known from video encoding technologies how to track parts of a video image from one frame to the next, and such tracking can be used to keep the region of interest on the skin of the subject. Nevertheless it would be useful to improve the motion compensation in order to improve the signal before spectral analysis. PPGi analysis can also be sensitive to lighting changes and it would also be useful to improve the robustness of the method to such changes, e.g. flicker in ceiling lights.

[0006] The paper "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models" by L. Tarassenko et al. PHYSIOLOGICAL MEASUREMENT, INSTITUTE OF PHYSICS PUBLISHING, BRISTOL, GB, vol. 35, no. 5, 28 March 2014, pp 807-831 discloses a method of obtaining physiological signals from a video image of a subject on which the precharacterizing portion of claim 1 is based.

[0007] One aspect of the present invention provides a method of detecting the frequency of a periodic physiological

signal of a subject from a video image of the subject, as defined by claim 1.

[0008] Thus with the present invention the pole selection from the fitted autoregressive model or models is conducted by a time domain analysis of the frequency that each pole represents with the signal to which the autoregressive model was fitted. Thus the spectral analysis using the autoregressive model is performed in the frequency domain, but the selection of the resulting pole is performed by reference back to the time domain. The spectral component (pole) which most closely represents the desired periodic physiological signal will have the highest cross-correlation in the time domain with the original signal.

[0009] The measure of similarity may be the cross-correlation. Any of the standard cross-correlation techniques well-known from signal analysis can be used.

[0010] The synthetic signal is preferably sinusoidal.

[0011] The method may further comprise the step of defining a similarity threshold and inhibiting the outputting step if the similarity of the synthetic signal with the maximum similarity with the selected ratio signal is below the threshold. This avoids outputting poor estimates of the frequency of the physiological signal.

[0012] The signal derived from the video image may be a single colour channel of video source data, and the colour channel may be visible or infra-red (IR). The signal derived from the video image can be a ratio of two colour channels of video source data, the channels being in the visible or IR region.

[0013] Alternatively the signal derived from the video image can be a sequence of co-ordinates obtained by tracking movement in the image, e.g. a sequence of co-ordinates of a physical feature or region of high contrast being tracked.

[0014] The signal derived from the video image is derived by the following steps: receiving multiple colour channels of video source data representing said video image, each channel comprising a time series of intensity data for that colour channel; for each different pairing of the colour channels calculating the ratio of the intensity at each time point in a first of the pairing to the intensity at the same time point in the second of the pairing to produce multiple ratio signals; performing source separation on the ratio signals and selecting the output component which is most periodic. The autoregressive models are fitted not to the colour channel signals themselves but to the most periodic component in the ratio signals resulting from dividing the signal in one colour channel by the signal in another. Preferably the ratios are the red to green ratio, the red to blue ratio, and the green to blue ratio (or their reciprocals). Each of the colour channels can be normalised by dividing by its mean value before the ratio calculation. Because changes in lighting of the image or movement of the subject tend to affect all channels at once, taking the ratios of the colour channels eliminates to some degree the effect of such lighting changes or movement.

[0015] It is possible to generate a synthetic time-domain signal for each spectral component from each model and calculate its cross-correlation with the signal to which the autoregressive model was fitted, or alternatively only poles in a selected frequency range (e.g. the allowed range for the physiological signal of interest) can be selected, or only the dominant pole (i.e. pole with the greatest magnitude) from each model is used. Choosing only the dominant pole of each model speeds-up the signal processing.

[0016] In an alternative embodiment the signal derived from the video image can be derived by the following steps: receiving multiple sequences of co-ordinates of one or more physical features being tracked; performing source separation on the sequences and selecting the output component which is most periodic.

[0017] The source separation step mentioned above may be by Principal Component Analysis. The selection of the most periodic output component is preferably by selecting the component having the greatest peakiness of frequency spectrum. In the method the three ratio signals resulting from dividing one of the colour channel signals by another are subject to principal component analysis and whichever component is the most periodic is selected as the signal for autoregressive analysis for that window. The selection as "most periodic" can be achieved by selecting the signal with the peakiest spectrum within the physiological frequency range of interest, i.e. which has a peak with the highest power as a proportion of the power at all frequencies. The selection of the most periodic of the signals results in a further improvement in movement compensation. Other methods of choosing the most periodic signal, such as analysis of the autocorrelation of each signal, can be used.

[0018] Autoregressive models of order 8-20, more preferably 7 to 11, can be fitted to the signal derived from the video image.

[0019] Preferably a synthetic signal is generated corresponding to the frequency of only the dominant spectral component for each order model.

[0020] Each colour channel is preferably normalised by dividing by its mean before the step of calculating the ratio.

[0021] The signal derived from the video image is preferably temporally windowed, e.g. into overlapping windows, for example from 4-30 seconds long with an overlap from 0.5 to 10 seconds.

[0022] The video source data may be a time series of intensity data for each colour channel for a region of interest defined in the video image. The region of interest is preferably on the skin of the subject and the video source data is a photoplethysmographic image. The region of interest may include a periodically moving part of subject.

[0023] The periodic physiological signal may be the heart rate or respiration rate of the subject.

[0024] The invention is embodied as a computer program comprising program code means for executing the method

on a computer system according to claim 14. According to a non-claimed example it can be incorporated into a dedicated patient monitor.

[0025] Thus another aspect of the invention provides apparatus according to claim 15 for detecting the frequency of a periodic physiological signal of a subject from a video image of the subject comprising: an input for receiving one or more colour channels of video source data representing said video image, each channel comprising a time series of intensity data for that colour channel; a processor for processing video source data; the processor being configured to execute the steps of the method above; the apparatus further comprising an output to output as the physiological signal frequency the frequency of the synthetic signal with the maximum cross-correlation with the signal derived from the video image. As above the colour channel may be visible or IR.

[0026] The invention will be further described by way of example with reference to the accompanying drawings in which:-

Figure 1 is a flowchart explaining a prior art method of physiological signal analysis;

Figure 2 is a flowchart explaining a method of physiological signal analysis according to an embodiment of the present invention;

Figures 3A, B and C schematically illustrate apparatus and a video image in one embodiment of the invention; and Figures 4A, B, C, D, E, F and G illustrate data from steps in the method for an example in which heart rate is estimated.

[0027] Figure 3A schematically illustrates in side view a subject 1 facing a device 3 which is integrally provided with a video camera such as a webcam 4 and a display screen 5. The device 3 can be a tablet or notebook computer, a mobile telephone (smart phone) or a dedicated patient monitor or other electronic equipment provided with a video camera 4. As will be explained below, device 3 is loaded with a software application which obtains the red, green and blue video output from the video camera 4 and analyses it to obtain a measurement of one or more of the subject's periodic vital signs such as the heart rate or breathing rate. These measurements are stored and displayed to the patient (on demand) and may also be sent by a network or Internet connection (not shown) to a remote server (not shown). The despatch and storage of signals and/or measurements on the remote server would allow automatic analysis and alerting, as well as allowing remote clinicians to review the subject's condition and, if necessary, advise the patient either via a messaging application within the software application itself or by other communication channels such as conventional messaging or telephones.

[0028] Figure 3C illustrates an example of the monitoring system intended for monitoring patients in hospitals. As illustrated, the monitor includes a video camera 4a mounted above the patient 1a for example on the ceiling or a frame 17, where it can view the patient 1a while the patient is on a bed 2. The output from the video camera 4a is supplied to a processing unit 3a via a conventional wired or wireless connection.

[0029] Figure 3B illustrates schematically the image obtained by the video camera 4. In general the image will contain an image 1' of the subject 1 and a background 20. One or more regions of interest (ROIs) 22 are defined in the image 1'. For the detection of a PPGi signal the region of interest 22 needs to be on an area of skin, e.g. the face of the subject. A breathing rate signal can also be obtained by looking for periodic movement of the body of the subject, e.g. the chest, or movement of something in contact with the subject, e.g. pillow or bedclothes, in which case the region of interest 22a can be defined on the chest. Conventional recognition and tracking algorithms can be used to find and track a suitable region 22, 22a in the image 1'. The region of interest 22, 22a can also be set initially by an operator or clinician, and then maintained automatically by tracking.

[0030] Figure 2 sets out the process for analysing the signals from the video camera 4 to detect the frequency of a periodic physiological signal in the video image 1'. This may be a PPGi signal taken from a region of interest on the skin or a periodic movement signal obtained from, for example, the region of interest 22a defined on the chest. The movement signal can be the periodic variation in the coordinates of a tracked feature in the image. In step 20 the video signal consisting of red, green and blue channels is received. The signal will consist of one spatial frame of data for each time point (the frame rate varies with the application but can be 12 or 24 frames per second). In step 22 a representative intensity for the region of interest in each frame for each of the three colour channels is obtained. This may, for example, be the mean intensity of each colour taken over the whole region of interest (i.e. the mean of all the red pixel values in the region of interest, the mean of all the green pixel values in the region of interest and the mean of all the blue pixel values in the region of interest).

$$\overline{R^{ROI}} = \frac{\sum_{ROI} Red \text{ pixel intensities}}{\text{number of red pixels in ROI}}$$

$$\overline{G^{ROI}} = \frac{\sum_{ROI} Green \text{ pixel intensities}}{\text{number of green pixels in ROI}}$$

$$\overline{B^{ROI}} = \frac{\sum_{ROI} \text{Blue pixel intensities}}{\text{number of blue pixels in ROI}}$$

[0031] However alternatively the mode of the distribution of intensities for each of the three colour channels within the region of interest can be used, or another representative intensity for each channel in the region of interest.

[0032] Once the representative intensity for each colour channel for each frame has been obtained, a time series of these intensities is assembled for a series of frames in a time window of, for example, 15 seconds. The length of the time window can be different, for example from 8 seconds to 1 minute. Each window overlaps its neighbour by a small time period, for example 1 second, though different overlaps, for example from 0.5 seconds to 5 seconds are usable.

$$\overline{R_{t1}^{ROI}} \dots \dots \overline{R_{twindow}^{ROI}} \quad \overline{G_{t1}^{ROI}} \dots \dots \overline{G_{twindow}^{ROI}} \quad \overline{B_{t1}^{ROI}} \dots \dots \overline{B_{twindow}^{ROI}}$$

[0033] In step 26 the signal values for each channel in each window are normalised by dividing each value by the mean of the representative intensities for that colour channel over that time window.

$$n_R(t) = \frac{\overline{R_{t1}^{ROI}}}{\mu_{window}^{Red}} \dots \dots \frac{\overline{R_{twindow}^{ROI}}}{\mu_{window}^{Red}}; \quad n_G(t) = \frac{\overline{G_{t1}^{ROI}}}{\mu_{window}^{green}} \dots \dots \frac{\overline{G_{twindow}^{ROI}}}{\mu_{window}^{green}}; \quad n_B(t) = \frac{\overline{B_{t1}^{ROI}}}{\mu_{window}^{blue}} \dots \dots \frac{\overline{B_{twindow}^{ROI}}}{\mu_{window}^{blue}};$$

[0034] In step 28 the ratio of the normalised channels to each other is obtained. That is to say for each time point in the sequence, the normalised value for the red channel is divided by the normalised value for the green channel, the normalised value for the red channel is divided by the normalised value for the blue channel and the normalised value for the green channel is divided by the normalised value for the blue channel. This generates three different ratio signals consisting of a ratio value for each of the frames (each time point) in the window.

$$R1(t) = \frac{n_R(t)}{n_G(t)}, \quad R2(t) = \frac{n_R(t)}{n_B(t)}, \quad R3(t) = \frac{n_G(t)}{n_B(t)}$$

[0035] For each sequence of ratios Principal Component Analysis (PCA) is carried out in step 30 and the output components are detrended and filtered (e.g. by a bandpass digital filter whose pass band is the range of physiological signal frequencies) in step 32.

[0036] Then in step 34 whichever of the detrended and filtered components is most periodic is judged. This is achieved in this embodiment by selecting the signal with the peakiest spectrum, i.e. that which has the peak with the highest power as a proportion of the total power at all physiologically-possible frequencies. For example, the criterion can be to be maximise the value of:-

$$\max \text{peak}(F)^2 / \text{area}(F)$$

[0037] Where F is the signal in the frequency domain, $\max \text{peak}(F)$ is the peak power, and $\text{area}(F)$ is the area underneath the power spectral density curve in the frequency range where the physiological signal may exist. This can easily be calculated from a Fast Fourier Transform of the detrended and filtered PCA output.

[0038] As a result whichever of the three components is regarded as most periodic is then the subject of autoregressive modelling in step 36 by fitting autoregressive models of multiple orders to the sequence of values for that window. In this embodiment AR models of order 7-11 are fitted to each sequence. However the order and number of model orders can be varied for different applications.

[0039] In the prior art method of Figure 1 the dominant pole from each order model would be taken and the frequencies to which they correspond are averaged to provide an estimate of the periodic physiological signal. However in accordance with the invention, in this embodiment one or more of the poles from each of the models of different order are taken and in step 38 the frequency which they represent (i.e. their phase angle) is calculated. In step 40 a pure sinusoid of that frequency is generated, this sinusoid being of the same time length as the detrended and filtered ratio signal and with an amplitude proportional to the standard deviation of the detrended and filtered signal from step 34.

[0040] Then in step 42 the detrended and filtered component from step 34 is cross-correlated with the synthesised signal to find the coefficient of correlation c , this being repeated for all possible non-identical phase differences. In this

repetition the synthesised signal may be moved relative to the ratio signal by one sample step each time, or a larger step can be used. Whichever phase difference gives the highest cross-correlation c is taken as the best fit for that model order. This procedure is repeated for poles from each of the different model orders. In step 44 the pole with the largest cross-correlation coefficient c is retained.

[0041] Steps 40, 42 and 44 can be conducted for only the dominant pole for each model order (i.e. the pole with the largest magnitude, possibly within an allowed frequency range for the physiological signal of interest), or it can be repeated for all poles within the allowed frequency range, or all poles from the model. The fewer poles processed, the quicker the processing.

[0042] In step 46 the frequency of the synthetic signal with the highest cross-correlation with the original ratio signal is selected as the estimated frequency of the physiological signal and is outputted. The output is preferably presented as, for example, a heart rate in beats per minute or a respiration rate in breaths per minute (by calculating the frequency in Hz times 60).

[0043] Steps 42 and/or 44 may be supplemented by a check that the cross-correlation c is above a predetermined threshold. If the cross-correlation is not above the predetermined threshold then the frequency may not be selected. Alternatively step 46 can be supplemented by the step of checking the cross-correlation coefficient c of the selected frequency against the predetermined threshold and if it is less than the threshold then the estimate is not output. This avoids outputting an estimate based on a poor level of periodic information in the input signal.

[0044] Figure 4 shows data at different steps for an estimation of heart rate from a webcam. Three channels of video data are captured using a webcam (in this case operating at 12 frames per second), by averaging (mean) the red, green and blue components of the pixels in a region of interest on the subject's forehead, the three channels being shown in Figure 4A. The channels are normalised by dividing each value by the mean for that channel for the whole time window (8 seconds in the illustrated example), resulting in the plots shown in Figure 4B, and ratios are then taken of each pair of channels as shown in Figure 4C. Principal Component Analysis is performed (resulting in the components shown in Figure 4D) and the components are detrended and filtered resulting in Figure 4E. For each principal component shown in Fig. 4E, the component with greatest "peakiness" in the frequency spectrum is chosen as the most periodic component. In this example the amplitude spectrum is calculated using the Fast Fourier Transform and the peak of the spectrum is found (marked with an arrow in Figure 4F). The "peakiness" of the spectrum (which is used as a measure of periodicity) is calculated as the squared height of this peak divided by the area underneath the amplitude curve. In this example, the first principal component has the highest "peakiness" and hence the greatest periodicity. This component with greatest "peakiness" in the frequency spectrum is chosen as the most periodic component and autoregressive models of different orders are fitted to this component. Each of the pole frequencies from the autoregressive models is tested by calculating the maximum cross-correlation of a sinusoid at that frequency with the detrended, filtered component (Figure 4G). Pole frequencies from all autoregressive model orders are tested, provided that the frequency is in the expected physiological range of heart rates. In this example, the highest cross-correlation was obtained for a frequency of 84 beats per minute, shown as a circle (which is from a pole of an autoregressive model with order 9). An estimated heart rate of 84 beats per minute is therefore outputted. This matches a simultaneous reference measurement of 84 beats per minute which was obtained from a finger pulse oximeter.

Claims

1. A method being executed on a computer system of detecting the frequency of a periodic physiological signal of a subject (1) from a video image of the subject (1), comprising the execution by a processor of the steps of: fitting (36) autoregressive models of one or more orders to a signal derived from the video image to detect spectral components in the signal; **characterized by:**

for each spectral component generating (40) a synthetic signal of the same frequency and calculating (42) its similarity in the time domain with the signal derived from the video image; outputting (46) as the physiological signal frequency the frequency of the synthetic signal with the maximum similarity with the signal derived from the video image.

2. A method according to claim 1 wherein the measure of similarity is cross-correlation.

3. A method according to claim 1 or 2 wherein the synthetic signal is sinusoidal.

4. A method according to claims 1, 2 or 3 further comprising the step of defining a similarity threshold and inhibiting the outputting step (46) if the similarity of the synthetic signal with the maximum similarity with the signal derived from the video image is below the threshold.

5. A method according to any one of the preceding claims wherein the signal derived from the video image is one of: a single colour channel of video source data; a sequence of co-ordinates obtained by tracking physical movement in the image; a ratio of two colour channels of video source data; or a sequence of co-ordinates of a physical feature being tracked.

6. A method according to claims 1, 2, 3 or 4 wherein the signal derived from the video image is derived by the following steps:

receiving (20) multiple colour channels of video source data representing said video image, each channel comprising a time series of intensity data for that colour channel;
for each different pairing of the colour channels calculating (28) the ratio of the intensity at each time point in a first of the pairing to the intensity at the same time point in the second of the pairing to produce multiple ratio signals;
performing (30) source separation on the ratio signals and selecting (34) the output component which is most periodic.

7. A method according to claims 1, 2, 3 or 4 wherein the signal derived from the video image is derived by the following steps:

receiving multiple sequences of co-ordinates of one or more physical features being tracked;
performing source separation (30) on the sequences and selecting (34) the output component which is most periodic.

8. A method according to claims 6 or 7 wherein the selection (34) of the most periodic output component is by selecting the component having the greatest peakiness of the frequency spectrum.

9. A method according to any one of the preceding claims wherein a synthetic signal is generated (40) corresponding to the frequency of only the dominant spectral component for each order model.

10. A method according to claim 6 wherein the colour channels are red, green and blue and the pairings are red and green, red and blue, green and blue.

11. A method according to any one of the preceding claims wherein the signal derived from the video image is temporally windowed

12. A method according to any one of the preceding claims wherein the video source data is either: a time series of intensity data for each colour channel for a region of interest (22, 22a), optionally wherein the region of interest is on the skin of the subject and the video source data is a photoplethysmographic image; or a sequence of co-ordinates obtained by tracking physical movement of the subject (1) in a region of interest in the image.

13. A method according to any one of the preceding claims wherein the periodic physiological signal is the heart rate or respiration rate of the subject (1).

14. A computer program comprising program code means for executing on a computer system the method of any one of the preceding claims.

15. Apparatus (3) for detecting the frequency of a periodic physiological signal of a subject (1) from a video image of the subject (1) comprising:

an input for receiving one or more colour channels of video source data representing said video image, each channel comprising a time series of intensity data for that colour channel;
a processor for processing video source data;
the processor being configured to execute the steps of any one of claims 1 to 13; the apparatus further comprising an output to output as the physiological signal frequency the frequency of the synthetic signal with the maximum cross-correlation with the signal derived from the video image.

Patentansprüche

1. Verfahren, das auf einem Computersystem ausgeführt wird, eines Erfassens der Frequenz eines periodischen physiologischen Signals eines Subjekts (1) aus einem Videobild des Subjekts (1), umfassend die Ausführung durch einen Prozessor der folgenden Schritte:
Anpassen (36) autoregressiver Modelle einer oder mehrerer Ordnungen an ein Signal, das aus dem Videobild abgeleitet ist, um Spektralkomponenten im Signal zu erfassen; **gekennzeichnet durch**:
Erzeugen (40) eines synthetischen Signals derselben Frequenz und Berechnen (42) seiner Ähnlichkeit im Zeitbereich mit dem aus dem Videobild abgeleiteten Signal für jede Spektralkomponente;
Ausgeben (46) als die physiologische Signalfrequenz der Frequenz des synthetischen Signals mit der maximalen Ähnlichkeit mit dem aus dem Videobild abgeleiteten Signal.
2. Verfahren nach Anspruch 1, wobei das Ähnlichkeitsmaß Kreuzkorrelation ist.
3. Verfahren nach Anspruch 1 oder 2, wobei das synthetische Signal sinusförmig ist.
4. Verfahren nach Ansprüchen 1, 2 oder 3, ferner umfassend den Schritt eines Definierens eines Ähnlichkeitsschwellenwerts und Verhinderns des Ausgangsschritts (46), wenn die Ähnlichkeit des synthetischen Signals mit der maximalen Ähnlichkeit mit dem aus dem Videobild abgeleiteten Signal unter dem Schwellenwert liegt.
5. Verfahren nach einem der vorhergehenden Ansprüche, wobei das aus dem Videobild abgeleitete Signal eines aus Folgendem ist: einem einzelnen Farbkanal von Videoquelldaten; einer Sequenz von Koordinaten, die durch Verfolgen physischer Bewegung im Bild erhalten wird; einem Verhältnis von zwei Farbkanälen von Videoquelldaten; oder einer Sequenz von Koordinaten eines physischen Merkmals, das verfolgt wird.
6. Verfahren nach Ansprüchen 1, 2, 3 oder 4, wobei das aus dem Videobild abgeleitete Signal durch die folgenden Schritte abgeleitet wird:
Empfangen (20) mehrerer Farbkanäle von Videoquelldaten, die das Videobild darstellen, wobei jeder Kanal eine Zeitreihe von Intensitätsdaten für diesen Farbkanal umfasst;
Berechnen (28) des Verhältnisses der Intensität zu jedem Zeitpunkt in einem ersten der Paarung zur Intensität zum selben Zeitpunkt im zweiten der Paarung für jede unterschiedliche Paarung, um mehrere Verhältnissignale zu produzieren;
Durchführen (30) von Quellentrennung an den Verhältnissignalen und Auswählen (34) der Ausgangskomponente, die am meisten periodisch ist.
7. Verfahren nach Ansprüchen 1, 2, 3 oder 4, wobei das aus dem Videobild abgeleitete Signal durch die folgenden Schritte abgeleitet wird:
Empfangen von mehreren Sequenzen von Koordinaten eines oder mehrerer physischer Merkmale, die verfolgt werden;
Durchführen von Quellentrennung (30) an den Sequenzen und Auswählen (34) der Ausgangskomponente, die am meisten periodisch ist.
8. Verfahren nach Ansprüchen 6 oder 7, wobei die Auswahl (34) der am meisten periodischen Ausgangskomponente durch Auswählen der Komponente mit der größten Peakigkeit (peakiness) des Frequenzspektrums erfolgt.
9. Verfahren nach einem der vorhergehenden Ansprüche, wobei ein synthetisches Signal entsprechend der Frequenz von nur der dominanten Spektralkomponente für jedes Ordnungsmodell erzeugt wird (40).
10. Verfahren nach Anspruch 6, wobei die Farbkanäle rot, grün und blau sind und die Paarungen rot und grün, rot und blau, grün und blau sind.
11. Verfahren nach einem der vorhergehenden Ansprüche, wobei das aus dem Videobild abgeleitete Signal temporär gefenstert ist.
12. Verfahren nach einem der vorhergehenden Ansprüche, wobei die Videoquelldaten entweder eine Zeitreihe von

Intensitätsdaten für jeden Farbkanal für einen Bereich von Interesse (22, 22a), optional wobei der Bereich von Interesse auf der Haut des Subjekts ist und die Videoquelldaten ein photoplethysmographisches Bild sind; oder eine Sequenz von Koordinaten, die durch Verfolgen physischer Bewegung des Subjekts (1) in einem Bereich von Interesse im Bild erhalten wird, sind.

13. Verfahren nach einem der vorhergehenden Ansprüche, wobei das periodische physiologische Signal die Herzfrequenz oder Atemfrequenz des Subjekts (1) ist.

14. Computerprogramm, umfassend Programmcodemittel zum Ausführen des Verfahrens nach einem der vorhergehenden Ansprüche auf einem Computersystem.

15. Vorrichtung (3) zum Erfassen der Frequenz eines periodischen physiologischen Signals eines Subjekts (1) aus einem Videobild des Subjekts (1), umfassend:

eine Eingabe zum Empfangen eines oder mehrerer Farbkanäle von Videoquelldaten, die das Videobild darstellen, wobei jeder Kanal eine Zeitreihe von Intensitätsdaten für diesen Farbkanal umfasst; einen Prozessor zum Verarbeiten von Videoquelldaten; wobei der Prozessor konfiguriert ist, die Schritte nach einem der Ansprüche 1 bis 13 auszuführen; wobei die Vorrichtung ferner einen Ausgang umfasst, um die Frequenz des synthetischen Signals mit der maximalen Kreuzkorrelation mit dem aus dem Videobild abgeleiteten Signal als die physiologische Signalfrequenz auszugeben.

Revendications

1. Procédé étant exécuté sur un système informatique consistant à détecter la fréquence d'un signal physiologique périodique d'un sujet (1) à partir d'une image vidéo du sujet (1), comprenant l'exécution par un processeur des étapes de :

adaptation (36) de modèles autorégressifs d'au moins un ordre à un signal issu de l'image vidéo pour détecter des composantes spectrales dans le signal ; **caractérisé par** :

pour chaque composante spectrale la génération (40) d'un signal de synthèse de la même fréquence et le calcul (42) de sa similitude dans le domaine temporel avec le signal issu de l'image vidéo ;

l'émission (46) en tant que fréquence de signal physiologique de la fréquence du signal de synthèse avec la similitude maximale avec le signal issu de l'image vidéo.

2. Procédé selon la revendication 1, dans lequel la mesure de la similitude est une corrélation croisée.

3. Procédé selon la revendication 1 ou 2, dans lequel le signal de synthèse est sinusoïdal.

4. Procédé selon les revendications 1, 2 ou 3, comprenant en outre l'étape de définition d'un seuil de similitude et l'inhibition de l'étape d'émission (46) si la similitude du signal de synthèse avec la similitude maximale avec le signal issu de l'image vidéo se situe sous le seuil.

5. Procédé selon l'une quelconque des revendications précédentes, dans lequel le signal issu de l'image vidéo est : un canal de couleur distincte de données de source vidéo ; une séquence de coordonnées obtenues par le suivi d'un mouvement physique dans l'image ; un rapport de deux canaux de couleur de données de source vidéo ou une séquence de coordonnées d'une caractéristique physique étant suivie.

6. Procédé selon les revendications 1, 2, 3 ou 4, dans lequel le signal issu de l'image vidéo est obtenu par les étapes suivantes :

réception (20) de multiples canaux de couleur de données de source vidéo représentant ladite image vidéo, chaque canal comprenant une série temporelle de données d'intensité pour ce canal de couleur ;

pour chaque appariement différent des canaux de couleur, le calcul (28) du rapport de l'intensité à chaque instant dans un premier élément de l'appariement à l'intensité au même instant dans le deuxième élément de l'appariement pour produire de multiples signaux de rapport ;

la réalisation (30) d'une séparation de sources sur les signaux de rapport et la sélection (34) de la composante

de sortie qui est la plus régulière.

7. Procédé selon les revendications 1, 2, 3 ou 4, dans lequel le signal issu de l'image vidéo est obtenu par les étapes suivantes :

la réception de multiples séquences de coordonnées d'au moins une caractéristique physique étant suivies ;
la réalisation d'une séparation de sources (30) sur les séquences et la sélection (34) de la composante de sortie qui est la plus régulière.

8. Procédé selon les revendications 6 ou 7, dans lequel la sélection (34) de la composante de sortie la plus régulière consiste à sélectionner la composante présentant le caractère le plus culminant du spectre de fréquences.

9. Procédé selon l'une quelconque des revendications précédentes, dans lequel un signal de synthèse est généré (40) correspondant à la fréquence de seulement la composante spectrale dominante pour chaque modèle d'ordre.

10. Procédé selon la revendication 6, dans lequel les canaux de couleur sont rouge, vert et bleu et les appariements sont rouge et vert, rouge et bleu, vert et bleu.

11. Procédé selon l'une quelconque des revendications précédentes, dans lequel le signal issu de l'image vidéo est temporairement en fenêtre.

12. Procédé selon l'une quelconque des revendications précédentes, dans lequel les données de source vidéo sont soit : une série temporelle de données d'intensité pour chaque canal de couleur pour une région d'intérêt (22, 22a), la région d'intérêt étant éventuellement sur la peau du sujet et les données de source vidéo étant une image photopléthysmographique ; soit une séquence de coordonnées obtenue par le suivi d'un mouvement physique du sujet (1) dans une région d'intérêt dans l'image.

13. Procédé selon l'une quelconque des revendications précédentes, dans lequel le signal physiologique périodique est le rythme cardiaque ou la fréquence respiratoire du sujet (1).

14. Programme informatique comprenant des moyens de code de programme destiné à exécuter sur un système informatique le procédé selon l'une quelconque des revendications précédentes.

15. Appareil (3) destiné à détecter la fréquence d'un signal physiologique périodique d'un sujet (1) à partir d'une image vidéo du sujet (1) comprenant :

une entrée permettant de recevoir au moins un canal de couleur de données de source vidéo représentant ladite image vidéo, chaque canal comprenant une série temporelle de données d'intensité pour ce canal de couleur ;

un processeur permettant de traiter des données de source vidéo ;

le processeur étant conçu pour exécuter les étapes selon l'une quelconque des revendications 1 à 13 ; l'appareil comprenant en outre une sortie pour émettre en tant que fréquence de signal physiologique la fréquence du signal de synthèse avec la corrélation croisée maximale avec le signal issu de l'image vidéo.

Fig. 1

Prior art

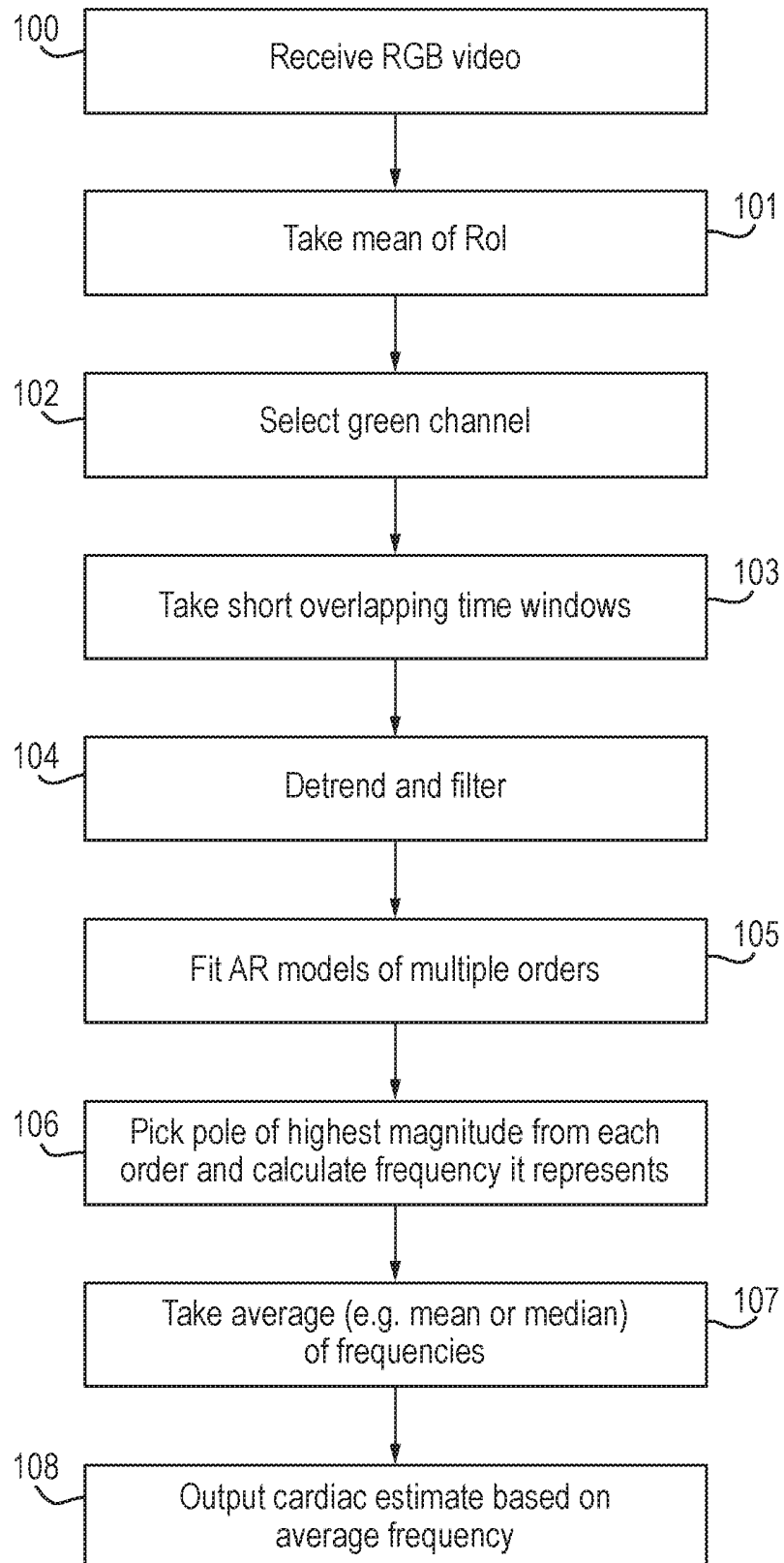


Fig. 2

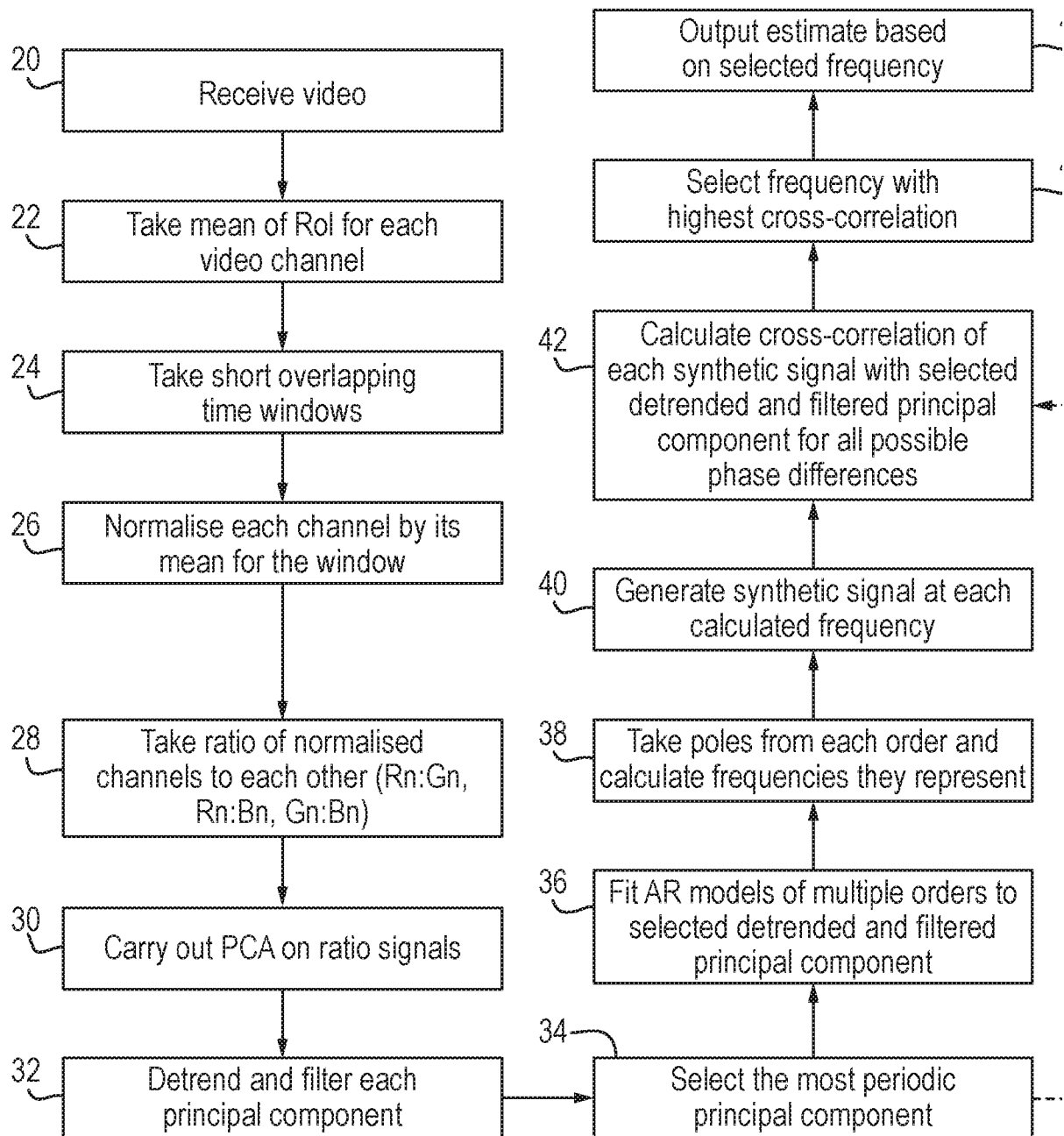


Fig. 3A

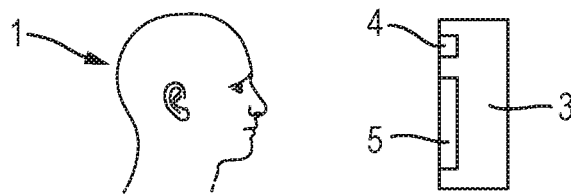


Fig. 3B

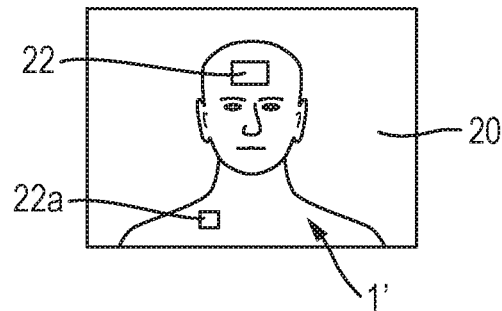


Fig. 3C

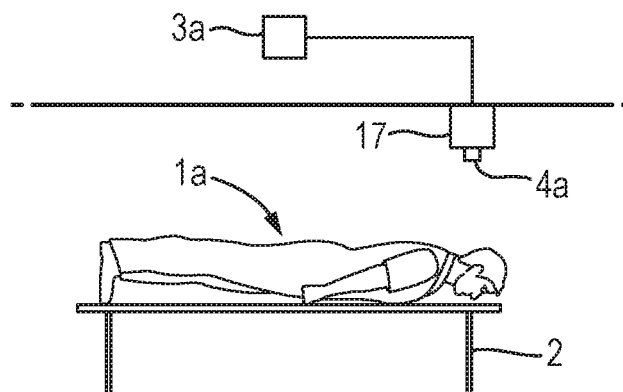


Fig. 4A

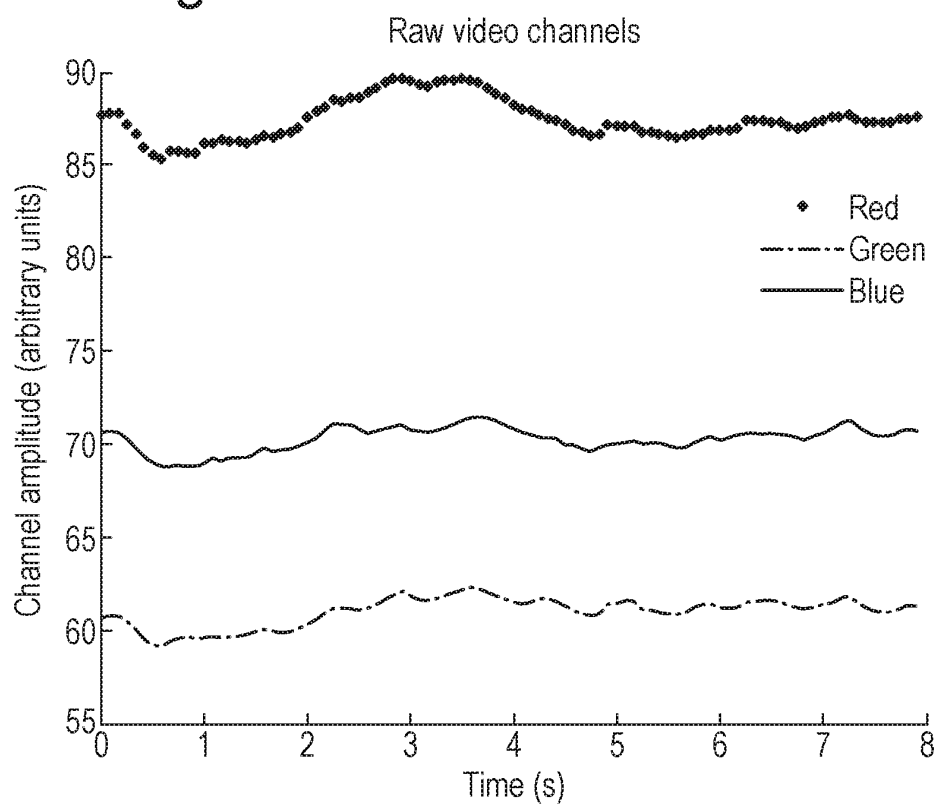


Fig. 4B

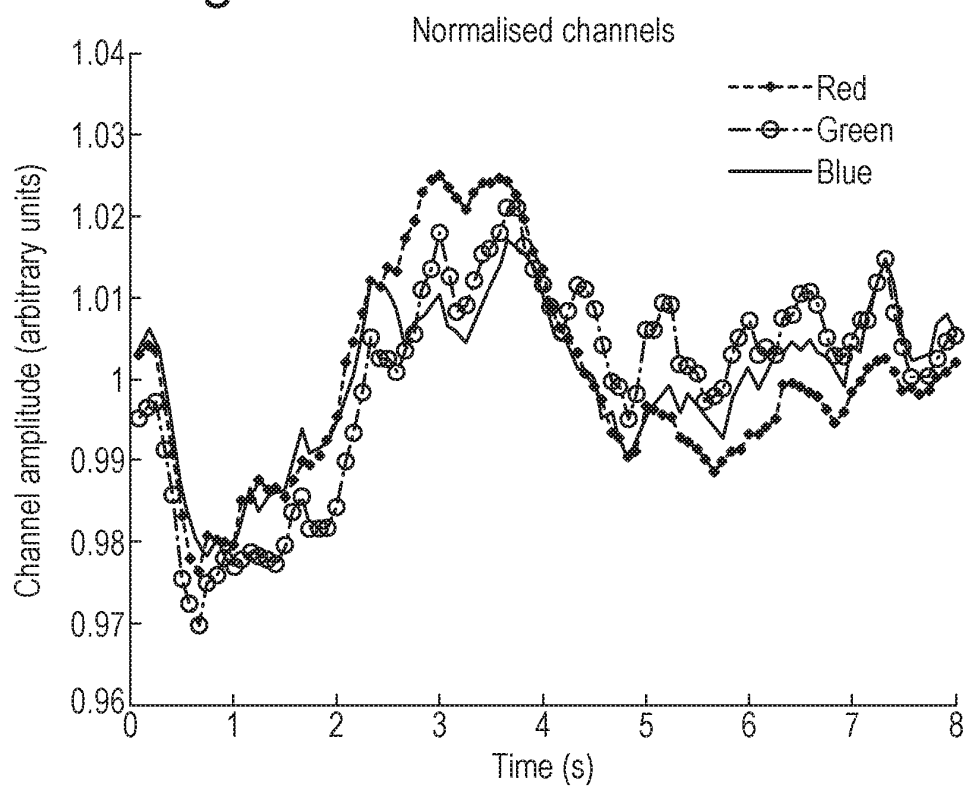


Fig. 4C

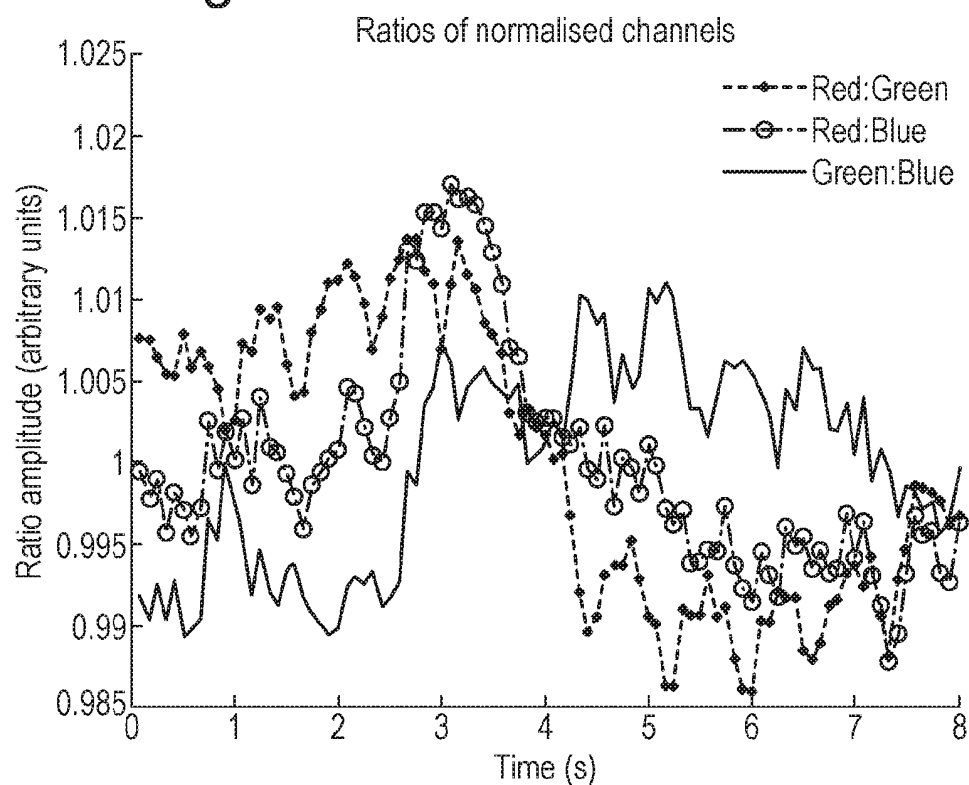


Fig. 4D

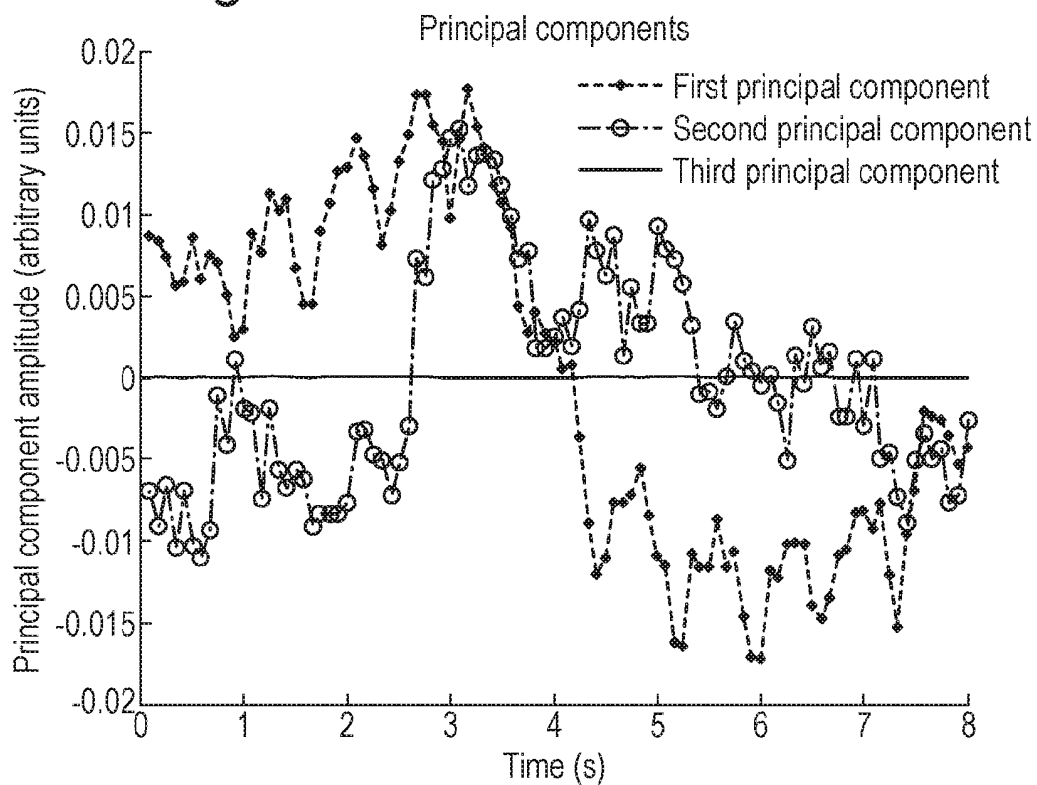


Fig. 4E

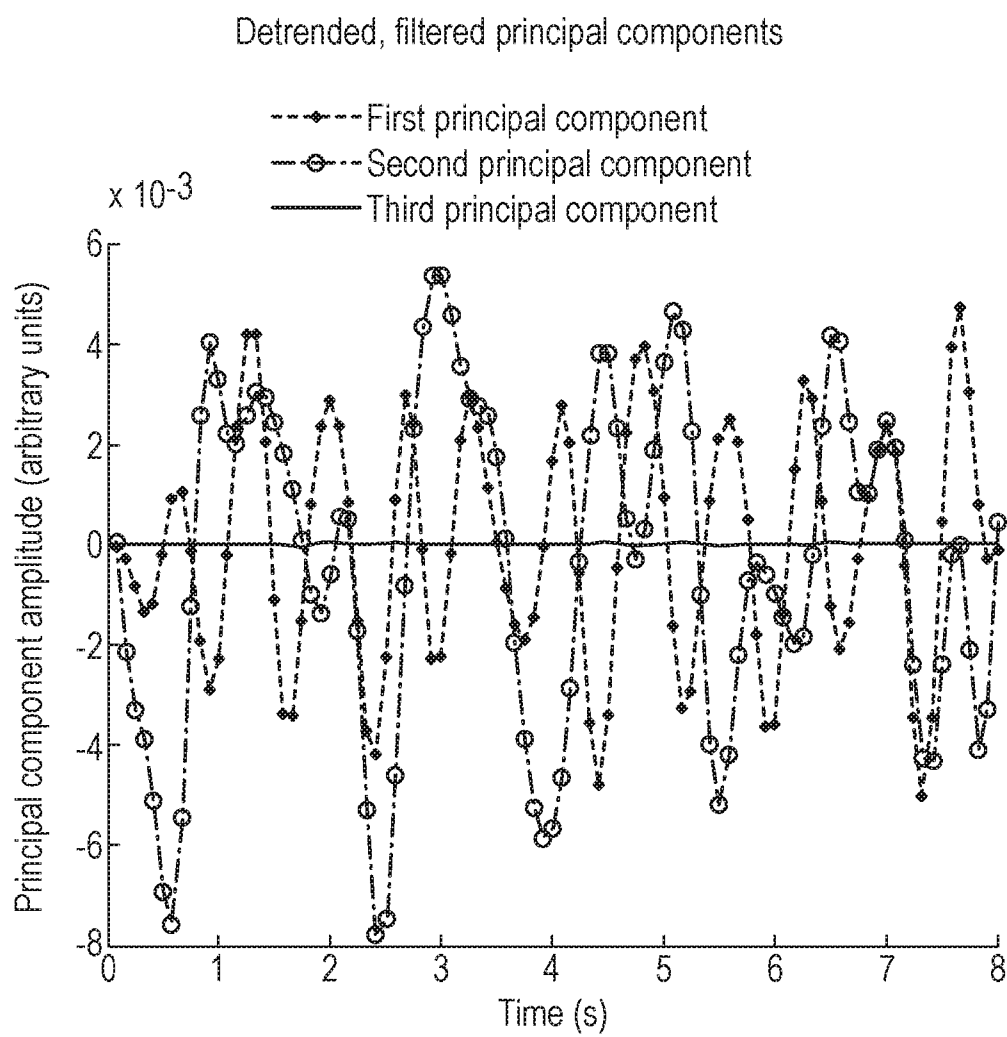


Fig. 4F

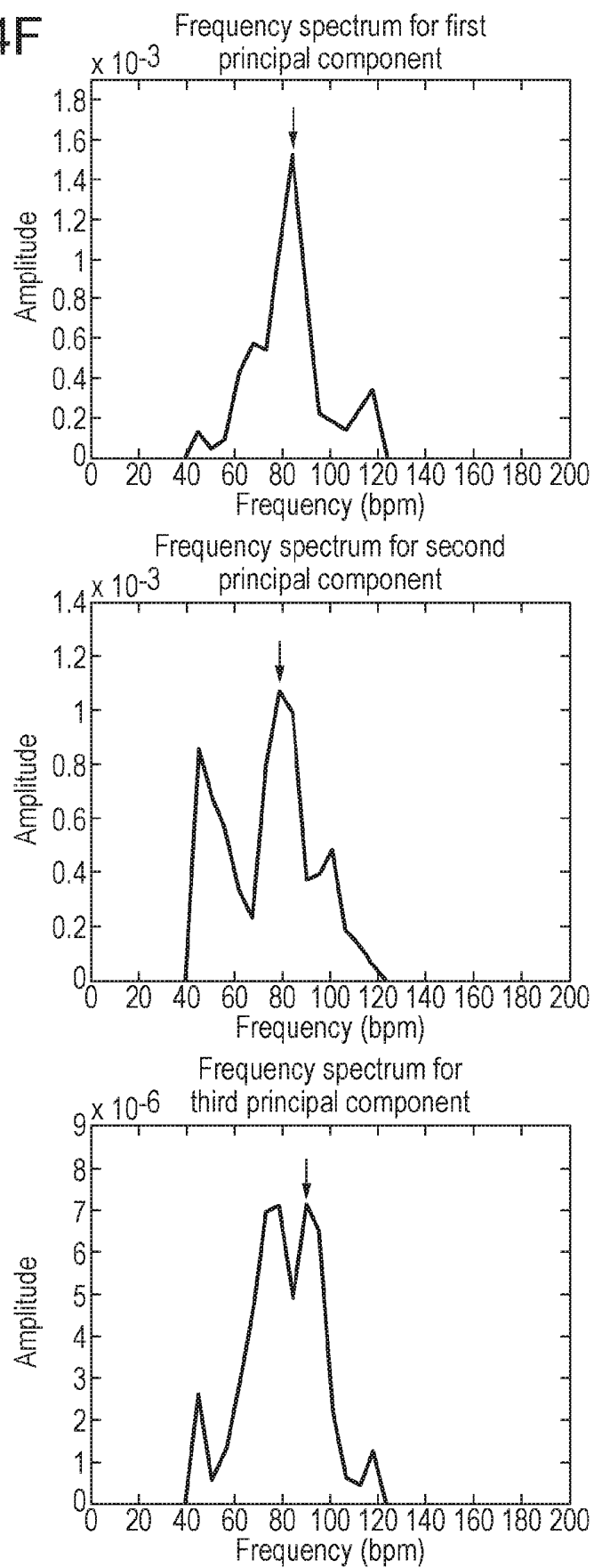
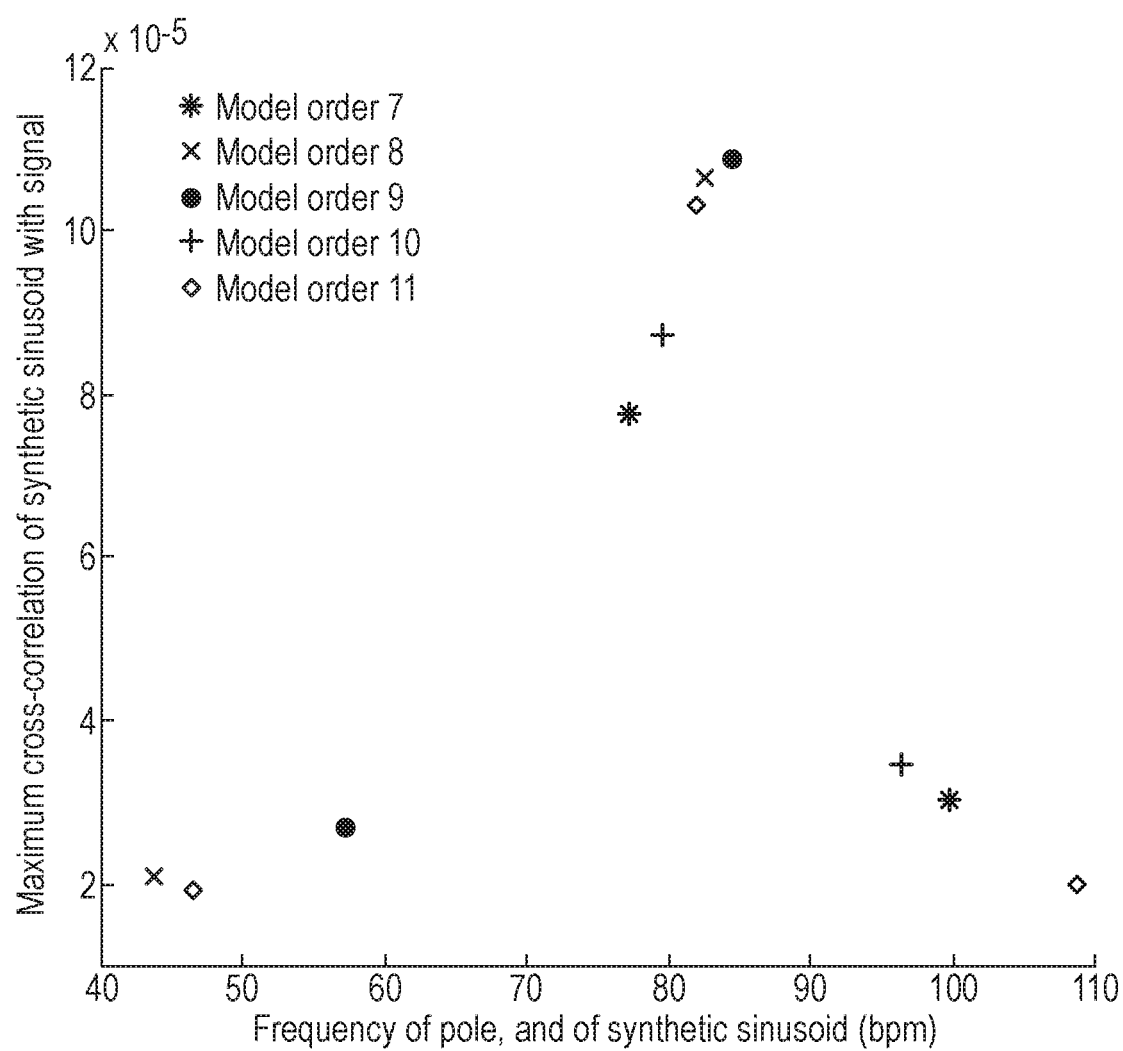


Fig. 4G

Cross-correlation for each pole frequency



REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO 2013027027 A2 [0002] [0004]

Non-patent literature cited in the description

- Non-contact video-based vital sign monitoring using ambient light and auto-regressive models. **L. TARASSENKO et al.** PHYSIOLOGICAL MEASUREMENT. INSTITUTE OF PHYSICS PUBLISHING, 28 March 2014, vol. 35, 807-831 [0006]

专利名称(译)	用于生理监测的方法和设备		
公开(公告)号	EP3229676B1	公开(公告)日	2019-06-12
申请号	EP2015808020	申请日	2015-12-08
[标]申请(专利权)人(译)	oxehealth		
申请(专利权)人(译)	oxehealth有限		
当前申请(专利权)人(译)	oxehealth有限		
[标]发明人	GUAZZI ALESSANDRO		
发明人	GUAZZI, ALESSANDRO		
IPC分类号	A61B5/0295 A61B5/00 A61B5/026 G06T7/00 A61B5/08 G06K9/46 A61B5/0205 A61B5/024 A61B5/11 A61B5/113		
代理机构(译)	J A KEMP		
优先权	2014021785 2014-12-08 GB		
其他公开文献	EP3229676A1		
外部链接	Espacenet		

摘要(译)

自回归建模用于识别周期性生理信号，例如受试者图像中的心率或呼吸率。通过将每个信号除以其平均值来对视频信号的颜色通道进行加窗和归一化。找到归一化信道相互之间的比率，并对比率信号进行主成分分析。选择最周期的主成分，并将一个或多个不同指令的自回归模型拟合到所选成分。采用不同阶的拟合自回归模型的极点，产生对应于每个极点频率的纯正弦曲线，并找到它们与原始分量的互相关。选择与具有最大互相关的正弦曲线相对应的任何极点作为原始视频信号中的周期性生理信息的频率的最佳估计。该方法可以用在患者监视器中或支持网络摄像头的设备中，例如平板电脑或智能电话。

$$\overline{R_{ROI}} = \frac{\sum_{ROI} \text{Red pixel intensities}}{\text{number of red pixels in ROI}}$$