Modular continuous wavelet processing of biosignals: extracting heart rate and oxygen saturation from a video signal

Paul S. Addison ✉

Medtronic Respiratory & Monitoring Solutions, Technopole Centre, Edinburgh EH26 0PJ, UK
✉ *E-mail: paul.addison@medtronic.com*

A novel method of extracting heart rate and oxygen saturation from a video-based biosignal is described. The method comprises a novel modular continuous wavelet transform approach which includes: performing the transform, undertaking running wavelet archetyping to enhance the pulse information, extraction of the pulse ridge time–frequency information [and thus a heart rate (HR_{vid}) signal], creation of a wavelet ratio surface, projection of the pulse ridge onto the ratio surface to determine the ratio of ratios from which a saturation trending signal is derived, and calibrating this signal to provide an absolute saturation signal ($S_{vid}O_2$). The method is illustrated through its application to a video photoplethysmogram acquired during a porcine model of acute desaturation. The modular continuous wavelet transform-based approach is advocated by the author as a powerful methodology to deal with noisy, non-stationary biosignals in general.

1. Introduction: Remote acquisition of physiological signals from video image streams is attracting much attention in the bioengineering space. In particular, the video photoplethysmogram may be used to derive the two main vital signs displayed on a pulse oximeter device: heart rate (HR) and oxygen saturation (SpO_2). HR has been tackled by many groups using a variety of signal processing approaches [1–8]. The more challenging task is the determination of oxygen saturation. Kong *et al.* [8] demonstrated a two-camera system for the determination of oxygen saturation based on visible light where two narrow-band filters (at 660 and 520 nm) were mounted on the cameras. They found good agreement with a traditional finger sensor. More recently, Tarassenko *et al.* [9] employed autoregressive modelling and pole cancellation of the red, green, blue (RGB) signals from a digital video camera to determine HR and respiratory rate from patients undergoing a dialysis session. They also demonstrated how desaturation events may be tracked in obstructive sleep apnea patients. The same group have also shown that it is possible to monitor HR, respiratory rate and changes in oxygen saturation in the neonatal intensive care unit [10]. The determination of oxygen saturation using standard video equipment would negate the requirement for specialist equipment and enable the rapid dissemination of such technologies. In the work reported here, a method for extracting HR and saturation trending from a standard video signal (HR_{vid} and $S_{vid}O_2$) is described which comprises a number of distinct algorithm modules based on the continuous wavelet transform (CWT) [11, 12].

2. Method: The flow diagram in Fig. 1 describes the method. Two input signals are used: the *R* and *G* video signals (from RGB). This was found to be the best performing of the three possible two-signal combinations. (Note that others have used the *R* and *B* channels [9, 10]). A CWT, $T(a, b)$, of each signal is first computed

$$T(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\psi^*\left(\frac{t-b}{a}\right) \mathrm{d}t \qquad (1)$$

where $\psi^*(t)$ is the complex conjugate of the wavelet function $\psi(t)$, a is the dilation or scale parameter of the wavelet, b is the location parameter of the wavelet and $x(t)$ is the signal under investigation [11]. A tunable Morlet-based continuous wavelet was employed

as it provides the ability to alter the shape of the time–frequency footprints (Heisenberg boxes) of the wavelet in the transform domain [13]. Running wavelet archetypes (RWAs) are then generated from the two transforms. The RWA enables a time–frequency ensemble averaging which takes place in the transform domain and has the advantage over temporal methods of not requiring a knowledge of signal fiducial points [14]. The archetype transform, $T_{RWA}(a,b)$, is generated using a weighted
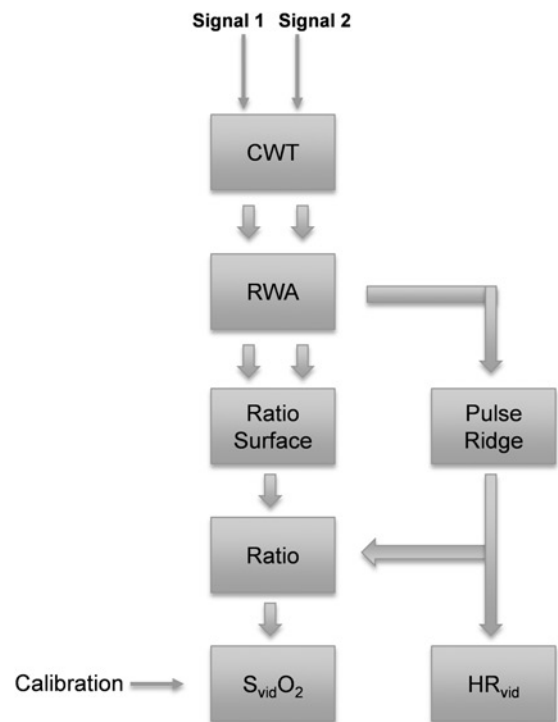

Fig. 1 *Flow diagram of method for determining oxygen saturation and HR from video signals*
Two CWTs are computed, one for each input signal, and two RWAs are computed. These are then combined to form a single ratio surface
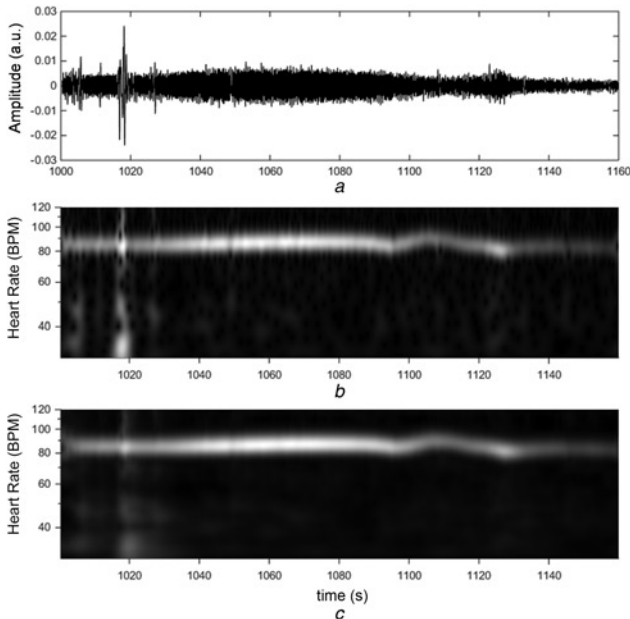
Fig. 2 *R signal with corresponding scalogram and associated RWA*
a R signal
b Transform modulus plot
c Corresponding RWA scalogram

averaging scheme as follows

$$T_{\mathrm{RWA}}(a, b) = wT(a, b) + (1 - w)T_{\mathrm{RWA}}(a, b - P(a)) \qquad (2)$$

where w is a predefined weight, $T_{\mathrm{RWA}}(a, b - P(a))$ is the previous archetype value separated from the current value by a period $P(a)$ and $T(a, b)$ is the currently computed wavelet transform. In the method, each time a value of $T(a, b)$ is computed, it is used with the previous archetype transform value to form a new value of the archetype transform. As the wavelet transform already separates out the signal information into natural temporal scales according to the periodicity of the wavelet function, the characteristic period of the wavelet may be used at each scale for $P(a)$. Thus, it can be seen that there is no requirement for the determination of fiducial points in the method, as the wavelet information is naturally 'rolled up' at each scale using $P(a)$ based on the wavelet function itself. The RWA method, fully described in [14], produces a more coherent manifestation of the pulse ridge in the transform domain.

Once the RWA is computed from the CWT, a ridge extraction algorithm is employed to extract the instantaneous frequency of the pulse ridge from the RWA transform. The ridge algorithm detects local maxima of the pulse band with respect to wavelet characteristic frequency. The ridge locus provides the instantaneous video-based HR signal, HR_{vid}. In addition, a wavelet ratio surface is computed from the RWAs by dividing the modulus of the two RWA transforms. (This is similar to the method described in [15], but using the RWA transform moduli instead of the original CWT moduli). The ratio surface provides a time–frequency decomposition of all ratios which may be used to compute a value of oxygen saturation. The HR_{vid} ridge locus is projected onto the ratio surface to determine the optimal ratio through time. This ratio is then used to determine the oxygen saturation through calibration with a reference pulse oximeter device (Nellcor N600x, Medtronic, Boulder, CO).

3. Monitoring during an acute hypoxic event: The method described above was applied to a standard video sequence from a commercially available standard RGB video camera

(Panasonic HDC-TM10 High Definition Video Camera). Video footage was acquired during a porcine model undergoing a rapid desaturation event where the arterial oxygen saturation dropped from 100% to around 40% for 2 min before being brought back to 100%. A Nellcor pulse oximeter (Medtronic, Boulder, CO) was also attached and provided a reference HR and oxygen saturation. The R and G signals were extracted from the video image, which was zoomed in on visible skin around the face region. The AC components from each signal were extracted by high-pass filtering at 0.25 Hz to remove respiratory and other lower frequency signal components, and then normalised by their DC component. The normalised AC signals were then wavelet transformed using a Morlet wavelet with characteristic frequency $\omega_0 = 15$ [13]. The R signal is shown in Fig. 2*a* with the corresponding wavelet modulus plot in Fig. 2*b*. A distinct pulse band can be seen across the transform modulus plot at around 85 beats per minute. The corresponding processed RWA scalogram is shown in Fig. 2*c*, where the smoothing of the band relative to the original scalogram in Fig. 2*b* is evident.

The wavelet ratio surface derived by dividing the R RWA transform modulus by the G RWA transform modulus is shown in Fig. 3. There is an obvious flat 'valley' in the surface coincident where the two dominant pulse bands divide each other out. Noise manifests on the surface as distinct spikes in the off-pulse band regions. An optimal ratio is extracted through time by simply projecting the pulse rate onto the ratio surface. The pulse rate may be derived by extracting the wavelet ridge frequency (i.e. the video-based HR signal) over time. This is shown in Fig. 4*a*. Apart from localised edge effects, the video and pulse oximeter HRs stay within 2 BPS of each other during the period of investigation (see also Fig. 5*c*). The projection of the pulse rate onto the ratio surface is shown in Fig. 4*b* and the subsequent extracted ratio is shown in Fig. 4*c*. This ratio is then calibrated using the existing oximeter reading. The calibrated video saturation signal is presented in Fig. 5*d* along with the original oximeter-based saturation (SpO$_2$) signal. The original R and G signals and the extracted HRs are also shown in Figs. 5*a–c*, respectively. Note that the S$_{\mathrm{vid}}$O$_2$ appears to fall more rapidly than the SpO$_2$. This is a manifestation of the difference in filtering characteristics of the two signals: one is wavelet based, applied to a video signal, and the other uses traditional pulse oximeter algorithmic components and is applied to a photoplethysmogram acquired at the finger.
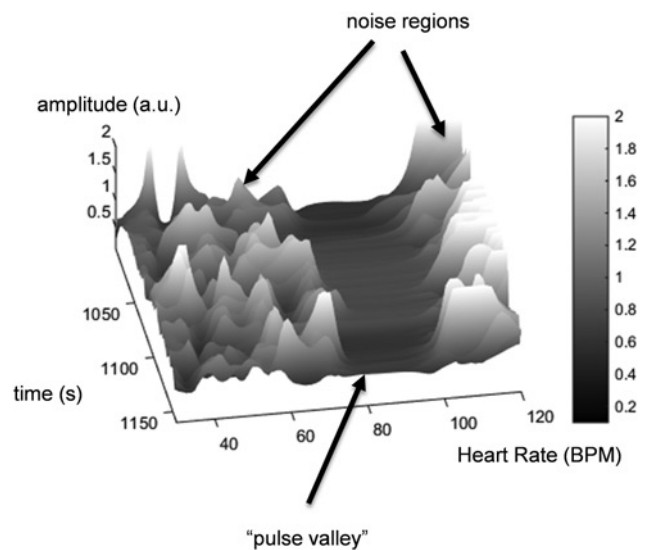

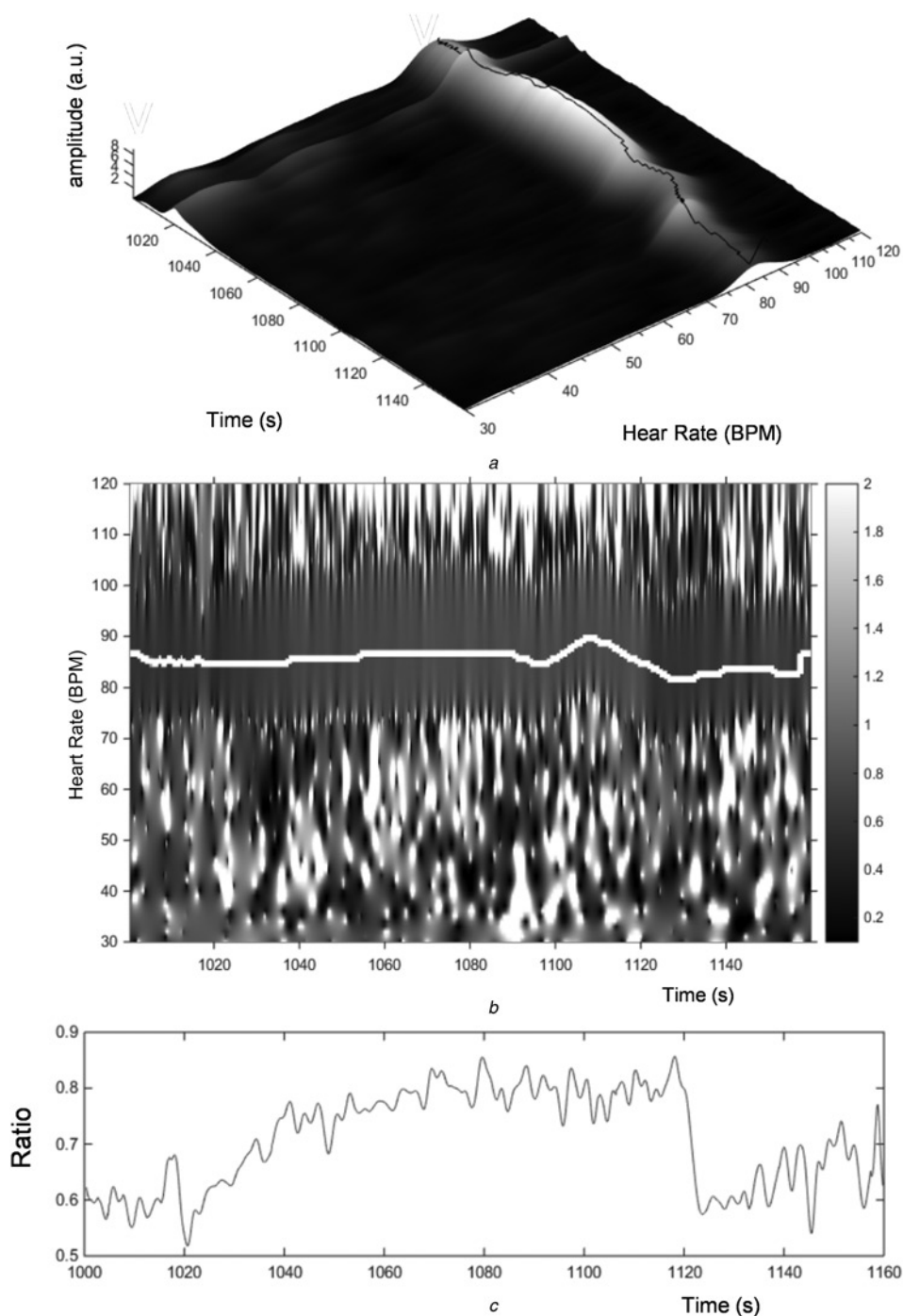**Fig. 3** *Wavelet ratio surface constructed from two RWAs*

Fig. 4 *Ratio through time from ratio surface*
a Ridge found from the RWA pulse band maxima
b Pulse band projected onto the ratio surface shown in Fig. 3
c Extracted surface heights (=ratio)

4. Conclusion: A method has been detailed for the computation of HR and oxygen saturation from video signals comprising a series of code modules built around manipulations of the CWT. The method allows for the main vital signs currently measured by a pulse oximeter to be measured using a video system. However, the use of such CWT-based modular approach is advocated by the author as a powerful methodology to deal with noisy, non-stationary biosignals in general. In addition, other modules could be added to further enhance the method. These include reassignment or synchrosqueezing to enhance the definition of the pulse band [16–19], cross-wavelet transforms to measure correlation and phase coupling between the signals as a measure of stability and/

or quality [20–23] and secondary wavelet feature decoupling to extract additional (e.g. respiratory) information from the signals [24]. In addition, alternative wavelet-based modules for determining the optimal ratio-of-ratios may be employed; such as one based on the three-dimensional Lissajous method [25].

The signals used in this study came from an opportunistic capture of video during a porcine model of desaturation involving an anesthetised animal where ambient light and motion noise was not an issue. However, significant noise on the signal before and after the hypoxic episode was observed when movement of the investigators in the room caused major fluctuations in ambient light levels. This caused poor-quality signal prior to the desaturation
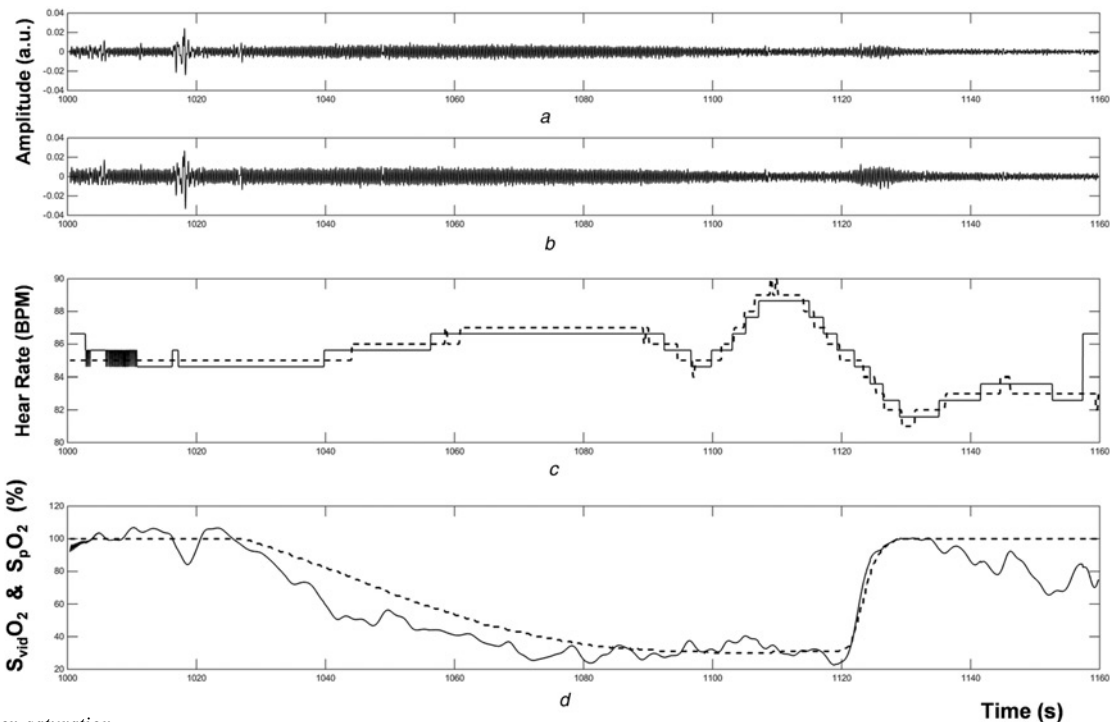
Fig. 5 *Oxygen saturation*
a Video R signal
b Video G signal
c Dashed line: pulse oximeter HR_p. Full line: video HR_{vid}
d Dashed line: pulse oximeter SpO_2. Full line: video $S_{vid}O_2$

phase and also subsequent to resaturation. In addition, the method as currently set up requires the input of calibration information from an existing oximeter. In practice, this system may be used as a trend and/or event monitor. Alternatively, the absolute saturation value may be calculated in a system that employs intermittent calibration during bedside observation by the caregiver using a pulse oximeter device. In the latter mode, the $S_{vid}O_2$ approximation of SpO_2 may be calibrated during spot checks and calculated continuously between these manual checks using the method described here.

In the preliminary investigation described in this Letter, close agreement between the video-derived parameters and pulse oximeter references was observed. A fundamental problem in attempting to compare the ability of video-derived HR and SpO_2 to track oximeter-derived values is that both 'devices' may produce a different time sequence of output values due to internal filtering characteristics and/or internal reporting delays. Thus, matching the two signals may be difficult during rapid changes in the parameter (e.g. as observed in Fig. 5*d*). The study described is, in effect, a tougher challenge than that for a 'standard' oximeter hypoxia study [26] where the measured SpO_2 is altered in discrete steps (e.g. from around 100, to 90, to 80, to 70% then back up again), and where at each step or 'plateau' the values are measured only after they become stable. These are then compared with either another device or a reference comprising several blood draws for arterial oxygen saturation values. In contrast, in the study described here, the values were compared dynamically, which is much more challenging and does not lend itself to rigorous comparison. However, this is a consequence of the nature of the study where the data was collected opportunistically and there was no fine control over the SpO_2 changes.

Future work will include opportunistically adding video capture to other ongoing in-house human hypoxia studies. However, in addition, more specific trials are planned to consider motion and lighting noise in a rigorously controlled environment with reference to the ISO 8060-2-61:2011 (2011) [26] pulse oximeter standard for the evaluation and qualification of pulse oximeter devices.

5 References

[1] Li M.H., Yadollahi A., Taati B.: 'A non-contact vision-based system for respiratory rate estimation'. 36th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC), 2014, pp. 2119–2122

[2] Poh M.Z., McDuff D.J., Picard R.W.: 'Advancements in noncontact, multiparameter physiological measurements using a webcam', *IEEE Trans. Biomed. Eng.*, 2011, **58**, (1), pp. 7–11

[3] Scalise L., Bernacchia N., Ercoli I., ET AL.: 'Heart rate measurement in neonatal patients using a webcamera'. IEEE Int. Symp. on Medical Measurements and Applications Proceedings (MeMeA), 2012, May 2012, pp. 1–4

[4] Kwon S., Kim H., Park K.S.: 'Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone'. Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC), 2012 August, 2012, pp. 2174–2177

[5] Sun Y., Hu S., Azorin-Peris V., ET AL.: 'Noncontact imaging photoplethysmography to effectively access pulse rate variability', *J. Biomed. Opt.*, 2013, **18**, (6), pp. 061205–061205

[6] Bousefsaf F., Maaoui C., Pruski A.: 'Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate', *Biomed. Signal Process. Control*, 2013, **8**, (6), pp. 568–574

[7] Cennini G., Arguel J., Akşit K., ET AL.: 'Heart rate monitoring via remote photoplethysmography with motion artifacts reduction', *Opt. Express*, 2010, **18**, (5), pp. 4867–4875

[8] Kong L., Zhao Y., Dong L., ET AL.: 'Non-contact detection of oxygen saturation based on visible light imaging device using ambient light', *Opt. Express*, 2013, **21**, (15), pp. 17464–17471

[9] Tarassenko L., Villarroel M., Guazzi A., ET AL.: 'Non-contact videobased vital sign monitoring using ambient light and auto-regressive models', *Physiol. Meas.*, 2014, **35**, (5), p. 807

[10] Villarroel M., Guazzi A., Jorge J., ET AL.: 'Continuous non-contact vital sign monitoring in neonatal intensive care unit', *Healthc. Technol. Lett.*, 2014, **1**, (3), pp. 87–91

[11] Addison P.S.: 'The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance' (CRC Press, 2002)

[12] Addison P.S., Walker J., Guido R.C.: 'Time–frequency analysis of biosignals', *IEEE Eng. Med. Biol. Mag.*, 2009, **28**, (5), pp. 14–29

[13] Addison P.S., Watson J.N., Feng T.: 'Low-oscillation complex wavelets', *J. Sound Vib.*, 2002, **254**, (4), pp. 733–762

[14] Addison P.S.: 'Running wavelet archetype: time–frequency ensemble averaging requiring no fiducial points', *Electron. Lett.*, 2015, **51**, (15), pp. 1153–1155

[15] Addison P.S., Watson J.N.: 'Oxygen saturation determined using a novel wavelet ratio surface', *Med. Eng. Phys.*, 2005, **27**, (3), pp. 245–248

[16] Kodera K., De Villedary C., Gendrin R.: 'A new method for the numerical analysis of non-stationary signals', *Phys. Earth Planet. Inter.*, 1976, **12**, (2), pp. 142–150

[17] Auger F., Flandrin P.: 'Improving the readability of time–frequency and time-scale representations by the reassignment method', *IEEE Trans. Signal Process.*, 1995, **43**, (5), pp. 1068–1089

[18] Daubechies I., Lu J., Wu H.T.: 'Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool', *Appl. Comput. Harmon. Anal.*, 2011, **30**, (2), pp. 243–261

[19] Addison P.S.: 'Identifying stable phase coupling associated with cerebral autoregulation using the synchrosqueezed cross-wavelet transform and low oscillation Morlet wavelets'. 37th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society, Milan, Italy, 25–29 August 2015

[20] Ge Z.: 'Significance tests for the wavelet cross spectrum and wavelet linear coherence', *Ann. Geophys.*, 2008, **26**, pp. 3819–3829

[21] Torrence C., Webster P.J.: 'Interdecadal changes in the ENSO-monsoon system', *J. Clim.*, 1999, **12**, (8), pp. 2679–2690

[22] Grinsted A., Moore J.C., Jevrejeva S.: 'Application of the cross wavelet transform and wavelet coherence to geophysical time series', *Nonlinear Process. Geophys.*, 2004, **11**, (5/6), pp. 561–566

[23] Addison P.S.: 'A review of wavelet transform time–frequency methods for NIRS-based analysis of cerebral autoregulation', *IEEE Rev. Biomed. Eng.*, 2015, **8**, pp. 78–85

[24] Addison P.S., Watson J.N.: 'Secondary transform decoupling of shifted nonstationary signal modulation components: application to photoplethysmography', *Int. J. Wavelets Multiresolution Inf. Process.*, 2004, **2**, pp. 43–57

[25] Addison P.S., Watson J.N.: 'A novel time–frequency-based 3D Lissajous figure method and its application to the determination of oxygen saturation from the photoplethysmogram', *Meas. Sci. Technol.*, 2004, **15**, (11), p. L15

[26] ISO 80601-2-61:2011: 'Medical electrical equipment – part 2–61: particular requirements for basic safety and essential performance of pulse oximeter equipment', *Int. Organ. Standard.*, 2011