

Revealing Physiological Changes Through Colour

DPhil Thesis

Alessandro Guazzi

Kellogg College

Submitted in Partial Fulfilment of the Requirements
for the Degree Doctor of Philosophy

Supervisor:

Prof. Lionel Tarassenko



Institute of Biomedical Engineering
Department of Engineering Science
University of Oxford

Michaelmas Term, 14th January 2016

Declaration

I declare that this thesis is entirely my own work and, except where stated, describes my own research.

A. R. Guazzi
Kellogg College

Abstract

Revealing Physiological Changes Through Colour

This dissertation extends remote vital sign monitoring by examining whether it is possible to track oxygen saturation with visible-light video cameras by filming an exposed skin region (the face). Two further questions are also investigated as part of the presentation of this thesis: the influence of skin melanin on the ability of video cameras to track oxygen saturation, and the ability of video cameras to assess the absolute value of oxygen saturation at any given point in time.

A study involving human volunteers was designed to answer these research questions in a controlled environment, and 46 subjects were recruited. The subjects were divided into six different skin phenotype groups (Fitzpatrick Scale I–VI). For a period of up to one hour, the volunteers' oxygen saturation was varied between 80% and 100% and videos of their faces were recorded in that time. The volunteers' reference oxygen saturation was recorded using two pulse oximeters placed on the index fingers of each hand. Another study was carried out in a neonatal high-dependency unit to test the algorithms for remote vital sign monitoring in a challenging clinical setting.

A theoretical basis for the derivation of oxygen saturation monitoring algorithms is developed throughout the dissertation, and two methods (baseline and RoR) were built upon this. The baseline method has parallels with oximetry and is based on tracking overall changes in the colour of the skin while minimising the effect of lighting, while the RoR method mimics pulse oximetry in form and assesses the relative change in amplitude as measured by each colour channel while oxygen saturation changes. Both methods are found to work well for subjects of Fitzpatrick phenotypes I–IV (average coefficient of determination r^2 of 0.79 and 0.71 for baseline and RoR respectively across all Type I–IV subjects), variably for Type V subjects (average r^2 values of 0.64 and 0.48), and not work for Type VI subjects (average r^2 values of 0.19 and 0.01). Although the RoR method yields noisier results than the baseline method, it was found that the timecourse of the latter varied with respect to the oxygen saturation.

A simple linear model used to derive oxygen saturation directly from the video camera data was thus devised using the RoR method only, and applied to Type I–V subjects. The subjects were divided into randomly generated but equally split training sets and test sets and the model parameters learned from the training set were evaluated in the test set. This was repeated over 10 iterations and the overall mean absolute error between the estimated oxygen saturation using the video camera data and the reference oxygen saturation using the pulse oximeters was, for Type I–V subjects respectively: 3.0%, 2.3%, 2.5%, 2.5%, and 4.6%.

Acknowledgements

First and foremost I would like to thank my supervisor, Prof Lionel Tarassenko, for the confidence he gave me in pursuing this DPhil, as well as the care and support that allowed me to see it through. Of equal import are the members of the “Camera Group” – Mauricio Villarroel, David Clifton, Joao Jorge, Jonathan Daly, Hamish Tomlinson, and Sitthichok Chaichulee. These are the individuals who provided me with boundless intellectual support throughout and who helped setting up the innumerable camera studies I have been part of. I would also like to thank Dr Matthew Frise for his invaluable role and upbeat demeanour over the course of the Hypoxia study – data collection would not have worked so well without him and it would have been a great deal less enjoyable. I am also very grateful for the expertise and work that Dr Kenny McCormick, Dr Gabrielle Green, and Sara Davis have brought and continue to bring to the NICU studies and the analysis of the neonatal data.

I have had the privilege of undertaking this journey with a group of incredible people who have become close friends over the course of these years. I would like to thank Sultan, Selcuk, Simao, Tasos, Aleksandra, Luca, Anjali, Chris, Paul, Marie, Maxim, Monica, Andrei, Lise, Lisa, Cooper, Zoe, Kevin, Raphael, Maddy, Giacomo, Charles, Serena, Bea, Elena, Misra, Christina, Rupert, and many more for making these so much more than just four years of research. In addition to the company, I owe a huge debt of gratitude to Dr Tasos Papstylianou for the endless discussions on the topics that are presented here. Equally, I would like to thank my family for their continued and unwavering support and warmth throughout these years. For what it’s worth, this dissertation is dedicated to the newest addition to the family, that she may be less jealous of the time I spent with SOPHIA.

Contents

List of Figures	iv
List of Tables	xii
Commonly Used Abbreviations	xiii
1 Introduction and Motivation	1
1.1 Remote Vital Sign Monitoring	1
1.2 Clinical Motivation	2
1.3 Overview of the Thesis	3
1.4 Publications	4
2 Background	5
2.1 Overview	5
2.2 Blood Oxygen	6
2.2.1 The Role of Oxygen	6
2.2.2 Oxygen Transport	7
2.2.3 Oxygen Saturation	10
2.3 Skin and Light	13
2.3.1 Anatomy of the Skin	13
2.3.2 Light-Skin Interactions	17
2.4 Measures of Blood Oxygen	26
2.4.1 Physical Measures of Blood Oxygen	27
2.4.2 Optical Measures of Blood Oxygen	29
2.5 Remote Photoplethysmography	42
2.5.1 Origin of the Remote Photoplethysmogram	42
2.5.2 Video and Colour	44
2.5.3 From Video to rPPG	48
2.6 Oxygen Saturation with Cameras	52
2.6.1 Regression-Based Methods	53
2.6.2 Pulse Oximetry Based Methods	54
3 Experimental Methods	57
3.1 Overview	57
3.2 Neonatal Trial	58

3.2.1	Standard Protocol	58
3.3	Hypoxia Trial	62
3.3.1	Study Design	62
3.3.2	Setup and Protocol	64
3.3.3	Reference Equipment	68
3.3.4	Camera Set-up	71
3.3.5	Equipment Synchronisation	74
3.3.6	Population Statistics	74
3.4	Summary	77
4	Theoretical Framework and Data Preprocessing	78
4.1	Introduction	78
4.2	Theoretical Model	78
4.3	Model Assumptions and Implications on Analysis	81
4.3.1	Lighting and Sensor	81
4.3.2	Skin Detection	85
4.3.3	Motion	91
4.3.4	Summary of Image Processing Steps	96
5	Oxygen Saturation and Baseline Skin Colour	99
5.1	Introduction	99
5.2	Methods of Analysis	100
5.2.1	Extracting the Baseline Signal	100
5.2.2	Revising the Theoretical Model	102
5.2.3	Continuous Tracking of Oxygen Saturation via the Baseline Signal	103
5.3	Testing the Methodology: Results and Analysis	106
5.3.1	Light Intensity and Gaussianity	107
5.3.2	Multiple Linear Regression	110
5.3.3	The Effect of Skin Colour	114
5.3.4	Variable Interpretation	118
5.4	Conclusion	119
6	Oxygen Saturation and Changes in AC	121
6.1	Introduction	121
6.2	Theory	122
6.2.1	Theoretical Model	122
6.2.2	Amplitude Estimation	123
6.2.3	Spatially Increasing the SNR	125
6.3	Initial Results	129
6.3.1	Heart Rate	129
6.3.2	Oxygen Saturation	132
6.3.3	Weighted Averaging	137
6.4	ROI Selection	139
6.4.1	Selection Criteria	139
6.4.2	Selection Criteria Analysis	144

6.5	Regression Models	150
6.6	Chapter Conclusion	152
7	Thesis Results and Discussion	153
7.1	Introduction	153
7.2	Hypoxia Results and Analysis	154
7.2.1	Results Summary	154
7.2.2	Effect of ROI Selection on Baseline Estimates	158
7.2.3	Differences Between Baseline and RoR Estimates	161
7.2.4	Causes of Error in Estimates	164
7.3	Towards Absolute SpO ₂ Estimation	169
7.4	Conclusion	173
8	Conclusions and Further Work	174
8.1	Introduction	174
8.2	Thesis Review	174
8.2.1	On the Origin of the Colour Changes	174
8.2.2	Limitations	177
8.3	Further Work	183
8.3.1	Improving Estimates: Uncertainty and Data Fusion	183
8.3.2	Clinical Validation	185
8.3.3	Alternative Methodologies	189
8.4	Concluding Remarks	190
A	Mathematical Notes	191
A.1	Viola-Jones and Adaboost	191
A.2	Movement Thresholding	192
A.3	Bayesian Change Point Detection	194
A.4	Rodrigues Rotation Method	196
A.5	The CHROM Algorithm for Heart Rate	197
	Bibliography	199

List of Figures

2.1	Figure showing the types of blood vessels present in the circulatory system: the arteries bringing the oxygenated blood to the circulatory system; the microcirculation at the level of the arterioles, capillaries, and venules; and the veins carrying deoxygenated blood back to the heart and lungs. (Image created by Kelvin Song and distributed under CC BY-SA 3.0, https://commons.wikimedia.org/wiki/File:Blood_vessels-en.svg , retrieved on 12/08/2015)	8
2.2	Figure illustrating the oxyhaemoglobin dissociation curve. A normal dissociation curve (continuous line) may be shifted to the left (dotted line) or to the right (dashed line) as a result of environmental factors. An increase in pH or a decrease in temperature will lead to a shift to the left, whereas the inverse will lead to a shift to the right. Foetal haemoglobin is also shifted to the left so as to enable the exchange of oxygen from the mother's blood. The partial pressure of oxygen (PO_2) is expressed in millimetres of mercury ($1\text{mmHg}=133.32\text{Pa}$)	11
2.3	Schematic of the anatomy of human skin. (Source: SEER Training Modules. U. S. National Institutes of Health, National Cancer Institute. Accessed 16th of June 2015, http://training.seer.cancer.gov/) . . .	14
2.4	Figure showing the electromagnetic spectrum (log scale, above), and the visible spectrum (linear scale, below). Source: Image by Philip Ronan, CC BY-SA 3.0, accessed at https://commons.wikimedia.org/wiki/File:EM_spectrum.svg on the 28th of November 2015.	17
2.5	Figure evidencing a highlight due to the specular reflection of the incident light off the top of the author's head.	19
2.6	Curves showing the observed brightness of skin for a light source at three different angles of incidence ($\theta_i = 0^\circ, \theta_i = 30^\circ, \theta_i = 60^\circ$). These curves characterise the <i>bidirectional reflectance distribution function</i> of the skin - a perfect Lambertian surface would have no dependence on the angle of incidence as well as the angle of observation. Image taken from [118]. . .	21
2.7	Extinction coefficient of eumelanin and phaeomelanin at different wavelengths. Source: S. L. Jacques, "Optical absorption of melanin", Oregon Medical Laser Center, www.omlc.ogi.edu/spectra/melanin/ , retrieved June 2015. .	24

2.8	Extinction coefficients of deoxygenated (Hb) and oxygenated (HbO) haemoglobin are shown in the ultraviolet, visible, and near infrared wavelengths, together with the absorption spectra of bilirubin and β -carotene. Sources: S. Prahl, Oregon Medical Laser Center, www.omlc.ogi.edu/spectra/ , retrieved June 2015.	25
2.9	Base-10 log of the absorption coefficients of water and fat are shown in the visible and near infrared wavelengths. Sources: [224], [161], [150]. . .	26
2.10	Schematic of the Clark electrode. (Image by Alex Yartsev, copied with permission from http://ow.ly/WkMNR Chapter 2.0.5, on the 28th of November 2015.)	28
2.11	Graphical representation of the Beer-Lambert law. If a material is constituted of an attenuating species (dashed circles), the radiant flux transmitted across the material is a logarithmic function of the concentration of the species in the material.	31
2.12	The log absorption spectrum of oxygenated and deoxygenated haemoglobin. Sources: S. Prahl, Oregon Medical Laser Center, www.omlc.ogi.edu/spectra/ , retrieved June 2015.	34
2.13	The absorption of light in skin is due to pulsatile (AC) and non-pulsatile (DC) components, shown for the case in which the signal is reflected. The overall signal, composed of the AC and DC components, is known as the photoplethysmogram (PPG) – the signal traced by the absorption of differing volumes of light.	35
2.14	The path length and transmission and reflectance properties of blood are affected by the change in erythrocyte orientation as the pressure pulse passes through the vessel.	37
2.15	Pulse oximetry may be used in transmission or reflectance mode.	38
2.16	The arrival of a pressure pulse during systole is thought to cause an increase in volume in the blood vessels leading to more light being absorbed by the tissue.	43
2.17	The colour change observed by a camera as blood moves in and out of the skin was amplified in [238], providing further evidence for the blood-volume change hypothesis. Image from [238].	43
2.18	Figure showing different filter architectures. (a) The Bayer filter pattern array. (b) The dichroic prism for a 3-CCD system. (c) Stacked photodiodes with colour-specific sensitivities. (Source: Colin Burnett, CC BY-SA 3.0)	46
2.19	Figure illustrating some of the key components in the image acquisition and correction pipeline.	48
2.20	Schema illustrating the procedure common to all prior art by which the video is reduced to a set of one-dimensional signals and hence to a frequency.	48
3.1	Monitoring set-up used in the Neonatal study showing the video camera placed above the modified incubator.	58
3.2	Variation in the output frequency of the ECG, PPG, and IPG waveforms.	60

3.3	Sensor response for the blue, green, and red channels of the JAI video camera as a function of incident wavelength. Source: personal correspondence with JAI.	61
3.4	The custom-built study chamber in the Department of Physiology, Anatomy and Genetics, prior to the covering of internal surfaces with reflective material.	64
3.5	Reflectance properties of Tyvek papers, amongst which 1443R [194]. Tyvek materials reflect light uniformly thereby appearing white.	65
3.6	Spectrum of the light emitted by the 576-LED Mosaic Lighting Panel. The emitted spectrum was recorded using a Ocean Optics USB4000 spectrometer.	66
3.7	Figure showing the interior of the chamber prior to the beginning of the Hypoxia sub-study. The subject, with the reference equipment already strapped on, is visible seated facing the cameras.	67
3.8	Positioning of ECG leads (left) and thoracic and abdominal respiratory rate elastic belts (right) for the main Stowood device used to record the reference data. (Source: Stowood Scientific Ltd, Oxford, UK).	69
3.9	Positioning of the five cameras used during the hypoxia and hypercapnia sessions. The JAI camera was below the set of Point Grey Cameras in the vertical axis.	71
3.10	HR distribution as recorded by the two pulse oximeters, placed on the left and right index fingers of the subjects.	75
3.11	SpO ₂ distribution as recorded by the two pulse oximeters, placed on the left and right index fingers of the subjects.	76
3.12	BR distribution as recorded by the two chest belts and the nasal cannula, respectively placed on the thorax, abdomen, and under the nostrils of the volunteer.	77
4.1	Figure showing spectrum of the LED panel as sampled directly (blue) and reflected off Tyvek (red).	82
4.2	Figure showing the grey-scale timeseries of light reflected off the background (left), and its spectrum (right).	83
4.3	Figure showing the change in noise, measured as the variance of the frame values, for a blacked-out camera kept in the Hypoxia chamber.	83
4.4	Figure showing the histograms for the read-out noise of last camera frame (at 6100s in Figure 4.3), one for each channel. From left to right: R,G,B.	84
4.5	Figure showing face detection using the modified Viola-Jones algorithm for the six different skin phenotypes. Detected face bounding box shown in red.	87
4.6	Schematic diagram showing the skin segmentation algorithm steps.	88
4.7	The posterior probability for belonging to the skin class is shown in 4.7a, with histograms of the normalised red and normalised green distributions on the sides. The resulting segmentation for that frame is instead shown in 4.7b. From left to right: the original image, the posterior probability, the segmentation via the naive Bayesian criterion overlaid on the original image. (Image reproduced with consent of the subject).	90

4.8	Figure 4.8a shows the average red, green, and blue signals using two different techniques for tracking intensity changes that share the same starting point. The dashed lines show the effect of keeping the region of interest fixed and the continuous lines show the effect of tracking the region of interest throughout time. Motion occurs in the first 5 seconds and the participant remains relatively still thereafter. Figure 4.8b shows the position of the fixed region of interest in the first frame, in green, and Figure 4.8c shows the region of interest as it has moved through tracking in the last frame in red compared to the original region of interest in green.	92
4.9	Figure showing the Euclidean distance in pixels moved by the centroid from its starting position throughout two sessions, evaluated once every second, for two subjects. The top subject (HP3-15) moved little throughout the session, whereas the lower plot shows the same for a subject (HP3-28) who moved a significant amount, as may be noted by the different vertical axes in the two plots.	93
4.10	Figure showing the segmentation of the Euclidean motion data for the subject undergoing significant movement in Figure 4.9 (shown in red) into regions of no motion ($SQI_L = 0$, grey background) and rapid motion ($SQI_L = 1$, white background).	94
4.11	Figure showing the video and image pre-processing steps applied prior to estimating the oxygen saturation. The KLT fine tracking method and the morphological manipulations, occurring only over stable segments and frames respectively, have been omitted for clarity. AOI = Area of Interest.	96
5.1	Figure showing in (5.1a) the variation in red baseline intensity for different regions of interest during the course of a session. Gaps appear in the signals if the ROI is no longer imaging a skin segment. The lower plot (5.1b) shows the average reference values of SpO_2 from the finger pulse oximeters.	101
5.2	Figure showing how the RGB values of ROIs within a skin-mask (each generated by taking the average of the 12×12 pixels that form the ROI), shown in 5.2a, form a point cloud in RGB-space, shown in 5.2b. The first principal component of uncentred PCA, passing through the origin, is traced in red, while the cloud's first principal component by mean-centred PCA is shown in yellow.	104
5.3	Schematic diagram of the hypothesised rotation occurring to a distribution of skin pixels in RGB space (shown as an ellipsoid) between a time m and a time n . Not to scale.	106
5.4	Figure showing Pearson's r and Spearman's ρ for each channel-pair. Note that the horizontal axis is ordered according to the FitzPatrick Skin Type, from Type I on the left through to Type VI on the right.	108
5.5	Figure showing the the fraction of frames rejecting the null hypothesis of the Anderson Darling test with a p-value < 0.01 .	109
5.6	Figure showing the correlation between each channel-pair's correlation coefficients and the oxygen saturation.	110

5.7	Bar graphs showing the coefficients of determination between oxygen saturation and the regression equations (5.7) and (5.8) with no movement SQI and with the application of SQI_L , which rejects rapid motion.	111
5.8	Figure showing the proportion of windows that are considered to have no motion for each subject according to SQI_L	112
5.9	Figure showing the regressed camera data $\hat{S}_\beta$ against the reference SpO_2 (red trace) for each of the 12 test subjects from Table 5.1. The black line and the blue points respectively show the results using no SQI and SQI_L	113
5.10	Figure showing the values of the standardised linear regression coefficients β_1^s and β_2^s , indexed according to skin colour type (1-6).	114
5.11	Coefficient of determination r^2 for different values of β_1 and β_2 . Only values above the median r^2 of each subject are shown for clarity.	115
5.12	Figure showing the oxygen saturation (increasing from dark to light) as a function of the rotation components for all of the subjects at once. Type VI subjects have been removed for clarity. This plot makes the lack of any relationship between the skin colour, the values of (θ, ϕ) for each subject, and the SpO_2 dependence on (θ, ϕ) apparent. Outliers for HP3-06 may be seen around co-ordinate values $[0.66, 0.285]$ and $[0.64, 0.3]$	117
5.13	Plots of the movement of θ and ϕ (in blue) are shown plotted against the reference SpO_2 (in red, at the bottom) for three subjects. From these plots, ϕ can be seen to be negatively correlated with SpO_2 , whereas θ only seems to be affected by the rapid oxygenations (shown at positions A, B, and C) and the subsequent deoxygenations.	118
5.14	Even when using only one variable $(\frac{\pi}{2} - \phi)$, there seems to be no relationship between the skin colour of the subject as measured by the FitzPatrick Skin Type Scale and the gradient of the regression of $\frac{\pi}{2} - \phi$ onto the reference SpO_2	119
6.1	Welch periodogram of a 60 second green average of the forehead of HP3-15.	126
6.2	Performance of PTD, SA and EBM algorithms on modelled data, comparing the signal-to-noise ratio (measured in decibels, lower plot) to the absolute error between the algorithms' estimates and the true amplitude (top plot) for a signal-averaged over increasing numbers of pixels.	128
6.3	Coefficient of determination r^2 (colour-coded) found for each subject (x-axis) using the different combinations of algorithms (y-axis) and colour channels.	133
6.4	Relationship between oxygen saturation and colour channel ratios.	135
6.5	Regressed RoR camera estimates for all subjects using $\log(\frac{B}{R})$. The continuous red line indicates the reference oxygen saturation, taken as the mean of the two pulse oximeters, the continuous black line shows the regressed estimates with no SQI but matching HR, and the blue points are the regressed estimates using SQI_L	136
6.6	The M-step function $\mathcal{V}(f)$, shown for $M=2$	138
6.7	Coefficient of determination r^2 for results computed using the weighted average as compared to the simple average.	139

6.8	It is the relative gradient of the intensities surrounding the region of interest and the relative trajectory of the body with respect to the region of interest that determine the phase of the periodic signal over time if it is purely due to motion. Here, for two bodies (one with a black top and another with a white top) that move in an oscillatory fashion in the vertical direction only, and which at time $t = 0$ are moving upwards as indicated by the arrow, the mean values of the intensities of the regions of interest (shown in grey at the centre of the body) for the two bodies will be exactly in antiphase, although their frequency will be identical.	141
6.9	The green SNR_{HR} map on the left indicates regions of high SNR_{HR} both around the forehead and at the sides of the throat. The signal-averaged green channel forehead pulse (blue) is morphologically distinct from the signal-averaged throat pulse (red) - see right-hand plots.	142
6.10	Each condition acts to remove some of the ROIs that had been considered as skin by the skin detection algorithm. From left to right, top row: L_s , L_d , L_v , L_m . From left to right, lower row: L_p , L_a , L_f , L	143
6.11	The coefficient of determination for each subject both using and not using the ROI selection criteria.	144
6.12	Average proportion of ROIs accepted by each selection criterion as a proportion of those selected by the skin/non-skin classification method.	145
6.13	Confusion matrix indicating what percentage of ROIs were selected by any two methods on average across all subjects. Bright yellow indicates the most overlap and dark blue indicates the least overlap.	147
6.14	Regressed RoR camera estimates for all subjects using $\log(\frac{B}{R})$ and the $\mathbf{L}^{++}$ criterion. The continuous red line indicates the reference oxygen saturation, taken as the mean of the two pulse oximeters, the continuous black line shows the regressed estimates with no SQI but matching HR, and the blue points are the regressed estimates using SQI_L . A marked improvement is visible, particularly for HP3-31, with respect to Figure 6.5.	149
6.15	Relationship between the RoR initial conditions (on air at the beginning of each Hypoxia session) and the linear regression parameter λ_1 , showing the line of best fit in red.	151
6.16	Relationship between the RoR initial conditions (on air at the beginning of each Hypoxia session) and the linear regression parameter λ_0 , showing the line of best fit in red.	151
7.1	Bland-Altman plot for the estimated oxygen saturations using the baseline method (7.1b) and RoR (7.1a) method with the L^{++} criterion. Each estimate was first regressed linearly onto the reference oxygen saturation. Results are reported for $\text{SQI}_L = 1$ and each colour represents a different subject, excluding Type VI subjects.	157
7.2	Bland-Altman plot used to compare the two reference SpO_2 values for the two pulse oximeters. Subjects of Type VI are excluded. Each colour represents a different subject.	158

7.3	A comparison of the baseline estimates (top) and RoR estimates (middle), each linearly regressed onto the reference SpO_2 (in red in each plot). Baseline and RoR estimates are shown in black (no SQI) and blue (SQI_L). The lower plot shows the movement during that period, calculated by estimating the displacement of the facial bounding box centroid from the beginning of the session.	162
7.4	The effect of movement, shown in the lower plot as the Euclidean distance moved by the facial bounding box centroid, on the video camera estimates: baseline (top plot) and RoR (middle plot). Both estimates have been linearly regressed onto the reference oxygen saturation (shown in red). The black continuous line represents the estimates when no movement SQI is used, while the blue dots show the results when SQI_L is applied.	166
7.5	Effect of movement, calculated as the total standard deviation of the Euclidean distance travelled by the facial bounding box from its starting position, on the r^2 coefficient of determination for the baseline method (left) and the RoR method (right).	167
7.6	Plots showing the time delay between the reference oxygen saturation (red) and the camera estimates (in black for no movement SQI and blue for SQI_L) for the baseline method (top) and the RoR method (middle). The estimated and reference heart rates are shown in the lower plot, in black and red respectively. The phenomenon of baseline underestimation during resaturation is also visible.	167
7.7	Plots showing the two reference oxygen saturation values (dashed lines), from the left (red dashed) and right (blue dashed) index fingers plotted against the mean of the two.	168
7.8	Figures showing the performance of the absolute estimates for SpO_2 using learnt linear parameters. Figure 7.8a shows the Bland-Altman plot for the absolute estimates and the reference SpO_2 , including all sessions for each of the 10 random iterations. Figure 7.8b shows the spread of the mean error for each session over the ten iterations, for each Fitzpatrick phenotype, as a box-and-whiskers plot. It should be noted that the reference pulse oximeters, although measuring values that are above 100% internally, have an artificial hard stop at 100%, giving rise to the characteristic straight diagonal line in Figure 7.8a.	170
7.9	Plots showing the effect of using the learned regression parameters (last iteration) on the RoR data for the original test subjects (Fitzpatrick Skin Types I–V), with no SQI applied. The black points show the regressed estimates and the red line is the average reference oxygen saturation.	172
8.1	Variation of the mean absolute error (MAE) from window to window as a function of the number of ROIs used to calculate the RoR estimate. Each colour represents one of the nine subjects used in the derivation of the RoR method in Chapter 6 (HP3-06, HP3-04, and HP3-42 were excluded).	184
8.2	One frame from the NICU trial video used. The Colour Checker chart is visible placed next to the neonate.	186

-
- 8.3 Plots for just under 1 hour of a clinical session in the NICU. From left to right: reference oxygen saturation from a pulse oximeter; baseline estimates using the polar angle $\frac{\pi}{2} - \phi$; RoR estimates using the ratio of blue to red normalised amplitudes; detrended background changes. The straight red lines indicate the lowest point of major reference desaturations. 188
- A.1 Figure showing the four Haar features used in the Viola-Jones detection framework proposed in [228]. 192
- A.2 Figure showing the segmentation of the motion data, shown in red for the subject undergoing significant movement in Figure 4.9, using the method proposed. The areas shaded in grey are areas classified as still. By way of comparison, the Bayesian online change-point detection algorithm proposed by [2] was also used on the same data, and the changepoints found are shown in black on the top plot, while the bottom plot shows the run lengths in red, and the posterior probability of the run length. . . 194

List of Tables

2.1	Main factors affecting arteriole muscle tone in the skin, both locally and in response to hormones.	9
2.2	The Fitzpatrick Phenotype Score [55]	24
3.1	Output frequencies from the Philips IntelliVue MX800 Patient Monitor .	60
3.2	Recording reference devices and output frequencies.	70
3.3	Features and specifications of cameras employed in the Hypoxia Study. (Sources: JAI, PtGrey)	72
3.4	Settings for cameras employed in the Hypoxia Study	73
3.5	Lenses used during hypoxia.	73
3.6	Lenses used during hypercapnia	73
3.7	Distribution of Fitzpatrick skin Types	74
5.1	Two subjects were randomly chosen from each of the six FitzPatrick Skin Types to test the methods proposed.	107
6.1	Model parameters for an HR+noise-based single-pixel model	127
6.2	Heart Rate MAE - CHROM [45] and HR_{gf} [202] algorithms as explained in the text.	131
6.3	Proportion of estimated HR with a difference greater than 4bpm with respect to the reference HR for the two algorithms considered.	131
6.4	A comparison of the individual performances of the selection criteria in terms of the correlation r^2 of the resulting signal with the reference oxygen saturation.	146
6.5	Comparing the performance of different selection criteria, from $\mathbf{L}$ (most ROI conditions) to $\mathbf{L}'''$ (least ROI conditions).	148
6.6	Performance of amplitude-based criteria	148
7.1	Results using the baseline method (developed in Chapter 5) and the RoR method with the $L++$ condition (developed in Chapter 6) on the remaining Hypoxia sessions, listed by Fitzpatrick type. The 12 test subjects are also present in italics.	156
7.2	Effect of the RoR selection criteria on the r^2 values obtained with the baseline estimates	160

Commonly Used Abbreviations

AOI		Area of Interest
BR		Breathing Rate
BSS		Blind Source Separation
DTFT		Discrete Time Fourier Transform
EBM		Energy Based Method
ECG		Electrocardiogram
EM		Expectation-Maximisation
FFT		Fast Fourier Transform
FPGA		Field Programmable Gate Array
HDU		High Dependency Unit
HSV		Hue, Saturation, Value
HR		Heart Rate
ICA		Independent Component Analysis
IPG		Inductance Pneumogram
KLT		Kanade-Lucas-Tomasi
MAP		Mean Arterial Pressure
MAE		Mean Absolute Error
NICU		Neonatal Intensity Care Unit
PCA		Principal Component Analysis
PERF		Perfusion
PO₂		Oxygen partial pressure

PaO₂		Oxygen partial pressure in arterial blood
PtcO₂		Oxygen partial pressure in cutaneous tissue
PvO₂		Oxygen partial pressure in mixed venous blood
PPG		Photoplethysmogram
PTD		Peak-Trough Difference
PTT		Pulse Transit Time
RGB		Red, Green, Blue
ROI		Region of Interest
RoR		Ratio of Ratios
SA		Signal Averaging
SNR		Signal to Noise Ratio
SO₂		Oxygen saturation
SaO₂		Arterial oxygen saturation
SpO₂		Peripheral oxygen saturation
StO₂		Tissue oxygen saturation
SvO₂		Venous oxygen saturation
VLS		Visible Light Spectroscopy

Chapter 1

Introduction and Motivation

1.1 Remote Vital Sign Monitoring

The last two decades have seen an exponential rise in the development of digital camera sensors, greatly driven by the increased consumer demand for ever-cheaper ever-better sensors in common electronic appliances such as personal computers, tablet computers, and, perhaps most importantly, mobile phones. This growth has catalysed the creation of a novel field of research in biomedical engineering: that of remote vital sign monitoring using cameras. The measurement of vital signs – heart rate (HR), breathing rate (BR), peripheral oxygen saturation (SpO_2), temperature (T), and blood pressure (BP) – is fundamental to current clinical practice and one of the most basic tools in the arsenal of a physician wishing to assess the health of a patient. The shift from contact to non-contact methods of measuring these vital signs presents numerous advantages in terms of compliance, as well as in managing skin irritation and infection.

Of course, a number of other methods for monitoring vital signs remotely exist already, such as through the use of radar [26, 131] or thermography [1]. However, the ubiquity, high sophistication and low price of video cameras have made them an attractive choice for remote monitoring, particularly as camera-based technologies could be immediately

integrated into telecare systems and furthermore have the potential of turning any consumer-level device into a medical device. Of the five standard vital signs, commercial visible-light cameras are thought to be able to estimate or approximate four – HR, BR, SpO₂, and BP – although prior to this work only the first two had been validated in a clinical setting [202], and the accurate assessment of blood pressure remains an unsolved problem. Finally, video camera-based methods have the potential for linking vital signs to a visual assessment of the patient’s movement and facial expressions [203], thereby also automating the softer measures clinicians use for health assessment.

This thesis in Healthcare Innovation thus sits within the context of future monitoring systems, not only capable of assessing the standard vital signs, but also provide other clinically relevant information. The primary objective of the thesis is to determine whether it is possible to measure absolute values of, or changes in, oxygen saturation using visible light video cameras. Secondary objectives include the determination of which factors influence the accuracy of the measurements made, and in particular the effect of skin pigmentation.

1.2 Clinical Motivation

As early as 1919, one of the founding fathers of physiology, John Scott Haldane, noted that “[the lack of oxygen] *means not a mere slowing down of life, but progressive and perhaps irreparable damage*” [64]. It comes to no surprise then, that oxygen is measured in multiple settings. Without oxygen, human brain cells can survive only for about one minute in normal conditions [122], and other organs are similarly compromised in its absence, therefore it comes to no surprise that blood-oxygen measurements are crucial in settings in which the patient is in a critical condition, such as in the ICU and in operating theatres [149]. Oxygen measurements further give an indication of the ability of the ventilatory system to cope with the body’s requirements, and are

therefore used in titration for patients requiring a continuous supply of oxygen [149], and to monitor patients undergoing anaesthesia with agents that have varying effects on the respiratory drive and blood pressure [41]. Oxygen monitoring is also crucial in paediatric and neonatal care, in which maintaining the correct oxygen saturation may make the difference between an increased risk of mortality on the one hand and an increased risk of disability on the other [171].

Continuous oxygen monitoring presents a number of advantages over intermittent measurements, and has been made possible thanks to the development of pulse oximetry, which allows the non-invasive monitoring of both pulse rate (that is equivalent to the heart rate) and oxygen saturation with reduced risk to the patient. The method however remains a contact-based one, making it unsuitable in scenarios involving burns, patients in isolation units, and patients susceptible to infection by cross-contamination [73]. Contact methods of vital sign monitoring, including pulse oximetry, have in fact been found to heighten the risk of hospital-acquired infections [135], with visible and non-visible blood even contaminating a large proportion of the monitoring equipment present in an operating theatre [155]. In addition to this, as with any device that is attached to the body, the technology has a number of risks associated with compliance and loss [151]. The development of a non-contact method of oxygen saturation monitoring that could further be integrated with existing non-contact methods for heart rate and breathing rate measurement offers the alluring prospect of a technology that could provide vital sign monitoring with no probes attached to the patient.

1.3 Overview of the Thesis

The thesis here presented is that blood oxygen levels can be tracked in ambient, broad-spectrum illumination using video cameras. The background necessary to understand this proposition is laid out in Chapter 2, in which the interactions of light with skin

are outlined in order to explain the current methods of measuring blood oxygen. The evidence gathered for the thesis is drawn from two studies, described in Chapter 3. The theoretical basis for the thesis is then developed in Chapter 4, and the influence of the approach on the data acquisition and preprocessing is discussed therein. The limitations of the theory mean that only one (the Hypoxia study) of the two sets of video data that have been acquired can be used so as to draw the strongest conclusions with regards to the object of the thesis. Finally, two separate methods for recording oxygen saturation under specific conditions are outlined in Chapters 5 and 6, and developed using a subset of the Hypoxia study videos. The final versions of these methods are then tested in Chapter 7, and the results of a simple method to measure oxygen levels directly are also presented. Finally, a general discussion of the nature of what is being measured by the video cameras and of the limitations of the methods proposed leads to the thesis' conclusions in Chapter 8, followed by indications of some of the further work necessary to advance the field.

1.4 Publications

The research culminating in this thesis has resulted in one first-author publication [63], which sets out a first attempt of some of the ideas presented in Chapters 4 and 6. Other second-author publications [202, 227] pre-date this, building the clinical case for the investigation here presented, and describing the studies that have been pursued as part of the general research objectives of remote vital sign monitoring with cameras.

Chapter 2

Background

2.1 Overview

The objective of this chapter is to establish the scientific background upon which the research described in this thesis has been built. The first two sections aim to give an understanding of the biological and physical processes relevant to the measurements made, whereas the last three outline the engineering methods which have served as a starting point for the work. Section [2.2](#) examines the importance and significance of oxygen saturation by describing the action, transport, and delivery of oxygen, and Section [2.3](#) explores the anatomy of the skin and the way visible light interacts with it, grounding later assumptions on the nature of the measurements to be made. Section [2.4](#) introduces the ways oxygen has been and is currently measured, clearly outlining the history of the development of the different methods in order to understand the context in which they were developed, and to examine the similarities and advantages of the various methods. As pulse oximetry is ultimately the methodology most related to the content of the thesis, it is discussed in some detail within the chapter. This leads naturally to Section [2.5](#) which outlines the background to remotely obtaining a photoplethysmographic signal using video cameras. Finally, Section [2.6](#) relates the prior work on obtaining estimates

of oxygen saturation with video cameras, both visible and infrared, and examines the different methodologies proposed in some detail.

2.2 Blood Oxygen

2.2.1 The Role of Oxygen

In order to sustain life and reverse the thermodynamic tendency to increasing levels of entropy, organisms need to generate energy continuously [160]. Cells accomplish this by using energy-carrying molecules such as adenosine triphosphate (ATP) that can exchange energy in different parts of the cell. Oxidative phosphorylation is the preferred aerobic process by which cells such as those in the human body generate ATP, a process which necessitates oxygen [211]. As a result, a continuous supply of oxygen needs to be maintained in order for the cell to survive, and in complex organisms in which simple oxygen diffusion is not sufficient this function is provided by the cardiovascular system.

The strong requirement for oxygen means that the cardiovascular system, which will be examined in greater detail in the following sections, needs to be tightly regulated. A lack of blood supply to parts of the body leads to cell damage and eventually cell death, and may lead to the death of the organism itself if key organs and systems such as the heart, the brain, or the central nervous system are affected [31]. In a clinical context, it is therefore essential to be able to assess the functioning of the cardiovascular system and its oxygenation as it is a direct indicator of a patient's health. Abnormal oxygenation levels are broadly divided into two categories: abnormally high levels of oxygen in the circulatory system, or *hyperoxia*, and abnormally low levels of oxygen in the circulatory system, termed *hypoxia*.

2.2.2 Oxygen Transport

Air is drawn into the lungs as a result of the mechanical action of the intercostal muscles and diaphragm. These cause an increase in the volume of the thoracic cavity, leading to a pressure gradient forming between the atmosphere and the lungs. Oxygen, which normally accounts for around 21% of air and at sea level has a partial pressure of 21.3 kilopascals, passes into the alveoli, specialised lung compartments with thin walls to allow the diffusion of gases from the lungs to the blood stream. As a result of differing concentration gradients, oxygen diffuses from the alveoli to the blood while the waste products of aerobic respiration, carbon dioxide and water, diffuse from the blood to the alveoli to then be expelled. It is because of this gas exchange that the partial pressure of oxygen in the alveoli is slightly less than the atmospheric partial pressure as it becomes a function of the respiratory rate of the cells. A constant pressure gradient is maintained by the movement of blood, pumped by the heart,

that carries newly oxygenated blood to the tissues through the arterial system and carries oxygen-depleted blood to the lungs via the venous system.

The blood itself is composed of cells suspended in a fluid that is mostly composed of water: the plasma. Nutrients and waste products such as carbon dioxide dissolve into the plasma – the exchange of carbon dioxide is principally plasma-based [6] – through the process of diffusion, that is affected by temperature and the relative partial pressures of the species. However, although oxygen is also dissolved in the blood, the oxygen requirement is such that this process accounts for an insignificant proportion of the total oxygen transported [124]. In order to cater for the high-energy requirements of the mammalian body, specialised cells, known as erythrocytes or red blood cells, are responsible for the transport of oxygen in the blood. Erythrocytes have a characteristic double-concave disk shape and each carries between 200 and 300 million haemoglobin molecules – making up to 96% of their dry weight — which give them their particular red colour and name [66, 207].

It is in fact these molecules, with a quaternary structure such that it leaves iron ions exposed, to which most of the oxygen transported by the circulatory system binds [183]. Haemoglobin has a much higher oxygen affinity than the plasma, with up to four oxygen molecules binding to each haemoglobin molecule. Furthermore, the structure is such that the affinity of the haemoglobin increases every time an oxygen molecule binds to one of the oxygen-active sites, with implications for oxygen saturation that shall be explored in Section 2.2.3.

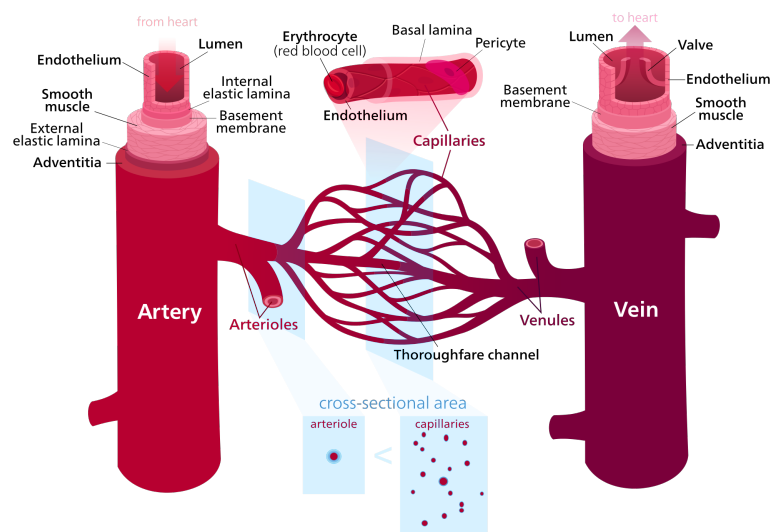


Figure 2.1: Figure showing the types of blood vessels present in the circulatory system: the arteries bringing the oxygenated blood to the circulatory system; the microcirculation at the level of the arterioles, capillaries, and venules; and the veins carrying deoxygenated blood back to the heart and lungs. (Image created by Kelvin Song and distributed under CC BY-SA 3.0, https://commons.wikimedia.org/wiki/File:Blood_vessels-en.svg, retrieved on 12/08/2015)

From the lungs, the blood passes into the left half of the heart and is thence pumped into the systemic circulatory system (shown in Figure 2.1) through the aorta, which branches into the major arteries to supply oxygen to the body. The blood is supplied to two main categories of organs and tissues within the body: those that ensure the homeostasis of the body and that are characterised by a much larger supply of blood than the demand of the tissue, such as the skin, and those that are essential to the function of

the body and that have a high demand for nutrients, such as the brain [211]. The role of the arteries is to carry oxygenated blood to all tissues and to aid the pumping of the circulatory system by sustaining a high blood pressure throughout. To this effect, they have an anatomy that is usually characterised by thick walls, and muscular or elastic tissue depending on which category of tissue they are perfusing. From the arteries, the blood is passed into the arterioles, smaller vessels that branch from the arteries and into the capillaries, thereby allowing microcirculation. Due to the drop in the velocity of blood as it passes from the larger arteries into the smaller-sized capillaries, arterioles have thick muscular walls to withstand and regulate the pressure difference. Furthermore, the arterioles are innervated by the autonomic nervous system, that regulates the tone of the muscular walls, and also respond locally to changes in the environment, including changes in hormone levels, pH, CO₂, stretch, and oxygen. An increase in muscular tone is termed vasoconstriction, whereas a decrease in tone is termed vasodilation. Table 2.1 shows the main factors affecting vasodilation and vasoconstriction in the skin, in which the effect of the hormones is particularly important as it enables thermoregulation, the main function of the blood flow to the skin [211].

Table 2.1: Main factors affecting arteriole muscle tone in the skin, both locally and in response to hormones.

	Vasodilation	Vasoconstriction
Local	pH, Carbon Dioxide, Potassium ions, Hydrogen ions	Oxygen, Stretch
Hormones	Beta-2 receptor activators	Alpha receptor activators

Microcirculation occurs in the network formed of arterioles, capillaries and venules. As discussed above, the arterioles feed the oxygenated blood from the arteries into the capillaries. It is in the capillaries, which have no smooth external muscles and are not innervated, that gas exchange occurs by a process of diffusion through the thin vessel walls into the tissue. This exchange is described by Fick's first law of diffusion:

$$\frac{\Delta N}{\Delta t} = KA \frac{\Delta P}{\Delta x} \quad (2.1)$$

where $\frac{\Delta N}{\Delta t}$ is the number of gas molecules exchanged per unit time, K is the diffusion coefficient, A is the surface area available for diffusion, ΔP is the pressure difference of the gas across the capillary wall, and Δx is the thickness of the capillary wall [211]. The gas exchange is further facilitated by the already mentioned affinity profile of oxygenated haemoglobin, which will favour the release of more oxygen as each oxygen molecule is released from the binding sites. The deoxygenated blood is then collected into the venules, characterised by thinner layers of smooth muscle compared to the arterioles. Multiple venules then feed into veins, that conduct deoxygenated blood back to the heart and thence into the lungs for it to be reoxygenated. Similarly to the venules, and as opposed to arteries and arterioles, veins typically have thinner smooth muscle around the vessel wall, and as a result do not usually expand and contract in response to the pressure wave generated by the contraction of the heart. Veins do however have valves in them to prevent the return of blood due to the low pressure in the venous tract. Despite the low pressure, the majority of the blood volume in the body is located in the venous compartment [92].

2.2.3 Oxygen Saturation

As has been examined in the previous section, the affinity of the haemoglobin molecule to oxygen molecules plays a very important role in oxygen transport: it is in fact this affinity that defines what is known as oxygen saturation. Oxygen saturation is a measure of the percentage of haemoglobin binding sites in the bloodstream occupied by oxygen molecules, ignoring the small amount of oxygen dissolved in the plasma [132]. As a result of the change in affinity of haemoglobin to oxygen, oxygen saturation may be examined as a function of the partial pressure of oxygen in the blood (the pressure exerted by the

oxygen within the blood, also known as the oxygen tension, and often denoted as PO_2) in what is known as the oxy-haemoglobin dissociation curve, as illustrated in Figure 2.2.

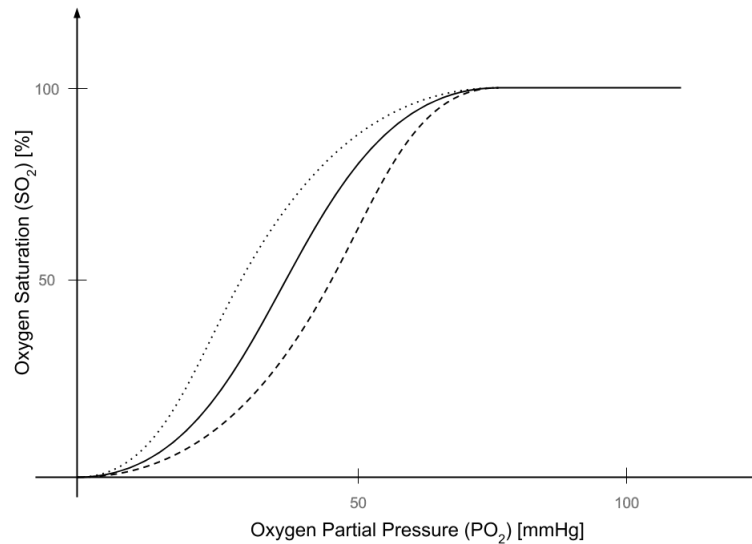


Figure 2.2: Figure illustrating the oxyhaemoglobin dissociation curve. A normal dissociation curve (continuous line) may be shifted to the left (dotted line) or to the right (dashed line) as a result of environmental factors. An increase in pH or a decrease in temperature will lead to a shift to the left, whereas the inverse will lead to a shift to the right. Foetal haemoglobin is also shifted to the left so as to enable the exchange of oxygen from the mother's blood. The partial pressure of oxygen (PO_2) is expressed in millimetres of mercury (1mmHg=133.32Pa)

The sigmoid shape of the dissociation curve describes haemoglobin affinity for oxygen: the flat part of the curve means that the haemoglobin molecule will take up oxygen even when most of the oxygen has been transferred from its gas phase whereas the steep part of the curve means the oxygen is rapidly taken up from the alveoli and released into the tissues at the capillary level. The curve is however affected by a number of factors, which cause a displacement left or right, such as the Bohr effect that leads to the curve shifting to the right in the presence of a low pH. This means that oxygen is released more easily at the tissue level but requires a higher partial pressure to bind at the lungs. Similarly, the curve for foetal haemoglobin is shifted to the left, allowing oxygen to be transferred from the mother's blood to the foetus [132].

At the clinical level, a difference must be noted between *functional oxygen saturation* and *fractional oxygen saturation*, that are simply two ways of reporting the oxygen saturation [46], and are often confused [209]. The former is defined as the percentage of oxyhaemoglobin with respect to the total number of functional haemoglobin:

$$SO_2 = \frac{HbO_2}{Hb + HbO_2} \quad (2.2)$$

It is this value that is commonly estimated via pulse oximetry and CO-oximetry (see Section 2.4), and that will be used interchangeably with oxygen saturation throughout this document. In turn, functional oxygen saturation may be expressed as arterial oxygen saturation (SaO_2), the functional oxygen saturation of the arterial blood, or as the peripheral oxygen saturation (SpO_2), the oxygen saturation of the arterial blood as estimated by a pulse oximeter (see Section 2.4.2). Typical values of the functional oxygen saturation in healthy humans are between 94%-98%.

It should be noted, however, that the haemoglobin molecule is not just bound to oxygen or free: it may be bound to other molecules. This nuance is better described by the fractional oxygen saturation, which is instead defined as the percentage of oxyhaemoglobin with respect to all types of haemoglobin:

$$FO_2 = \frac{HbO_2}{Hb_{all}} \quad (2.3)$$

This fractional oxygen saturation may only be measured by CO-oximeters operating at multiple wavelengths or by some modern gas analysers, and incorporates other types of bound haemoglobin (known as dyshaemoglobins), such as carboxyhaemoglobin ($HbCO$), methaemoglobin ($HbMet$), and sulphaemoglobin (HbS). Dyshaemoglobins are usually formed as a result of exposure to drugs and/or smoke, but also as genetic variations and naturally within the body [11].

2.3 Skin and Light

The optics of skin are key to the understanding of how the colour of skin arises and how changes in colour may be used to infer physiological parameters. Fortunately, this research area is of interest to a wide number of fields outside of biomedical engineering – from cosmetics to computer graphics – giving rise to an extensive body of literature that covers both experimental data on the physiology of the skin and theoretical and experimental approaches describing the spectral properties of the constituent elements of the tissue [9]. The colour of skin arises mainly from the distribution of two main pigments, melanin and haemoglobin [21], which absorb light at characteristic frequencies. However, it is also influenced by the scattering and reflective properties of these, and other, components. This section will explore the structure of the skin from the point of view of its histology, and then examine the influence of the biological parameters on the way light interacts with it.

2.3.1 Anatomy of the Skin

The skin is the largest organ in the human body, covering a surface area of around 2 m² in adults and contributing up to 6% of the total body mass [208]. It is the principal constitutive part of the integumentary system, whose main functions are to maintain homeostasis with regard to temperature, waste, and water; to allow sensation; to store and synthesise particular molecules; and to provide the first layer of protection against external agents, whether living (such as pathogens) or environmental (such as UV radiation). Although it carries out a critical role in controlling the amount of water that is lost to the outside environment, the skin is impermeable to water so as to ensure that the balance of nutrients within the body is not upset [208].

The social and evolutionary factors that have influenced the evolution of the human skin have also been recently researched [77]. The human skin has in fact been found to

have an important role in communicating the state of health and mood of an individual [192], thereby being directly related to the attractiveness of the individual. This is thought to have also influenced the development of the human eye, which is particularly sensitive to slight changes in colour that might indicate alterations in mood or health [27].

Structurally, the skin may be described as an inhomogeneous organ (see Figure 2.3). This may be subdivided into three main super-layers: the epidermis, papillary dermis, and the hypodermis, each of which is constituted of several sub-layers with more specialised functions [208]. The thickness and composition of each of these layers will further vary depending on the function of the skin at the location, which will also influence whether the skin is glabrous or hairy. Compared to hairy skin, glabrous skin is usually characterised by a thicker epidermis, the presence of sweat glands, and a higher receptor density [61].

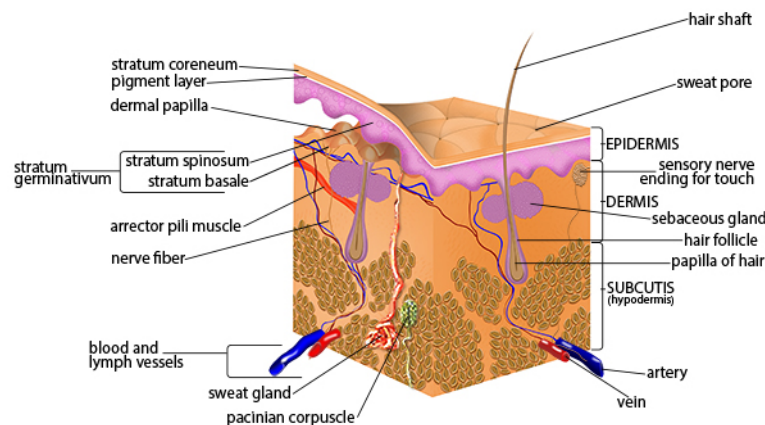


Figure 2.3: Schematic of the anatomy of human skin. (Source: SEER Training Modules. U. S. National Institutes of Health, National Cancer Institute. Accessed 16th of June 2015, <http://training.seer.cancer.gov/>)

The epidermis is the outermost super-layer of the skin and its main role is to protect internal organs from pathogens, water, and radiation, as well as to regulate the water in the organism. It is principally composed of keratinocytes, differentiated cells specific to

the epidermis that form a protective layer through a process known as cornification or keratinisation. No vessels reach the epidermis, and as a result cell oxygenation occurs by a process of diffusion with the environment and neighbouring cells [195]. The epidermis is structurally divided into four layers (present in all skin), with one additional layer that is only present in the regions where skin is thicker such as in the palm. The uppermost layer is the *stratum corneum*, constituted of a 0.01-0.02mm thick layer of dead cells within a lipid matrix [21, 201]. Underneath the stratum corneum lie three 0.027-0.15mm thick layers [76, 245] that are principally characterised by the stage of cornification of the keratinocytes: the *stratum basale*, the deepest, in which the keratinocytes are formed, the *stratum spinosum*, in which cornification commences, and finally the *stratum granulosum*, in which the keratinocytes lose their nuclei and release lipids into the extracellular space. A further layer, the *stratum lucidum* lies between the *stratum granulosum* and the *stratum corneum* in the skin of the palms and the soles [35].

The *stratum basale* is also constituted of melanocytes, which produce melanin. Melanin is a pigment that is found in the iris of the eye and hair as well as in the skin, and it is melanin that is the principal determinant of human skin colour, accounting for the majority of visible light absorption in the epidermis, although it is the production rate of melanin that varies amongst humans rather than the concentration of melanocytes [38]. The principal role of melanin in the skin is to prevent ultraviolet (UV) radiation from causing cell damage in the layers beneath the skin. The melanin is transferred from the melanocytes to the keratinocytes within specialised vesicles named melanosomes [208]. Two main typologies of melanin exist: pheomelanin (yellow/red pigmentation), and eumelanin (brown/black pigmentation), and the proportions and concentrations of the two types of pigment vary between individuals; however eumelanin usually accounts for most of the skin colour [206].

The dermis is separated from the epidermis by a basement membrane, and is composed of two main sublayers: the papillary region and the reticular region [208]. The division

of the dermis into the papillary and reticular dermis is based on the size of the blood vessels that perfuse the structure (finer vessels are present in the papillary dermis and serve to nourish the epidermis whereas larger vessels are present in the reticular dermis), and the total structure itself is 0.6-3mm thick [9]. The volume fraction of blood in the tissue varies within the 0.2-7% range and is found to increase almost linearly with depth [221] – the majority of the remaining volume is occupied by a fibrous extracellular matrix, itself mostly composed of collagen fibres [208]. Blood however contributes the main cutaneous pigments in this layer, as discussed in Section 2.2.2: haemoglobin, bilirubin and β -carotene, that contribute an orange/yellow colour to the skin. The pigments are found prevalently in the blood-stream but β -carotene may also be found in the epidermis and stratum corneum [104]. The vasculature of the dermis is a mix of arterioles, venules, and capillaries. As such, the oxygen saturation of the blood in the dermis is difficult to assess and few authors make an attempt at a “typical” value, preferring instead to state the oxygen saturation of blood in arteries and veins. A number of studies have however shown a strong correlation between the partial pressure of oxygen from arterial samples and that from capillary samples, as measured by blood gas analysis [67, 68, 133, 242], whereas other authors have found a strong correlation between the overall cutaneous oxygen saturation and the venous oxygen saturation [75, 134], explained as a result of the overall preponderance of venous blood volume throughout the body. Partly linked to the blood volume, but also a function of the interstitial tissue fluid volume, the water content of the skin can vary considerably, both in time and across different skin layers. However it has been estimated to increase mostly linearly between the *stratum corneum* (estimated to have an average water content of around 5%) and the reticular dermis (where the content is instead around 70%) [125].

The layer directly underneath the skin, the hypodermis, composed mainly of fat tissue, is not usually considered to be part of the skin and its size varies considerably depending on the location of the tissue. In regions such as the ears or the eyelids, it is

entirely absent [9].

2.3.2 Light-Skin Interactions

Visible light constitutes the part of the electromagnetic spectrum that is visible to the human eye, characterised by photons with wavelengths ranging from around 390nm to 700nm. As a result of the wave-particle duality of photons, a conundrum to physicists since Isaac Newton and Christian Huygens, the interactions of light may be described in a number of interchangeable models, that are valid approximations in particular conditions. The particle nature of light is key to geometric optics, whereby light is treated as non-interacting rays with a given energy and a geometric behaviour, as well as to radiative transfer, which instead looks at the interaction of light with matter and incorporates descriptions of emission, absorption, and scattering. The wave nature of light is instead useful to describe wave-specific behaviours that cannot be dealt with by particle models and that are instead the remit of wave optics: interference, diffraction, and polarisation.

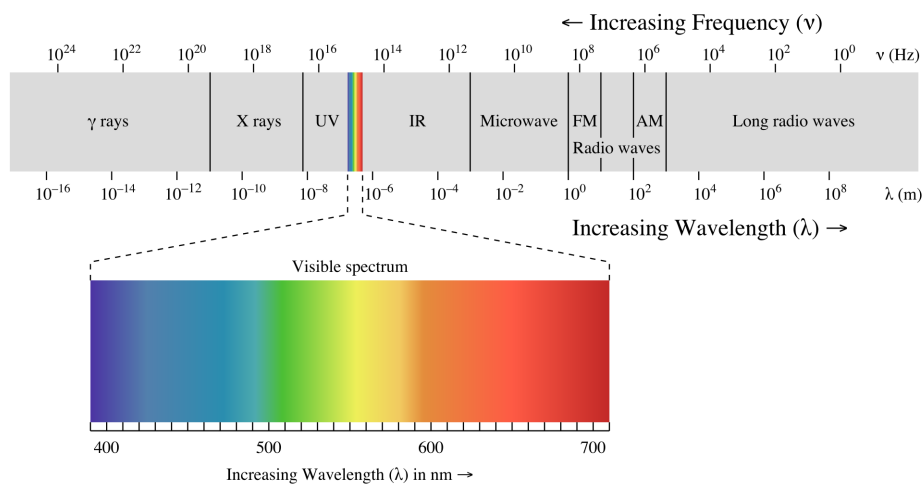


Figure 2.4: Figure showing the electromagnetic spectrum (log scale, above), and the visible spectrum (linear scale, below). Source: Image by Philip Ronan, CC BY-SA 3.0, accessed at https://commons.wikimedia.org/wiki/File:EM_spectrum.svg on the 28th of November 2015.

An understanding of the dual nature of light is important to the content of this thesis as it allows an understanding, for example, of what is meant by the absorption spectrum of a given pigment. However the concepts used in radiative transfer (also known as radiation transport) will generally be sufficient to explain most of the interactions between skin and light, and most of the work carried out in modelling the non-quantum interactions of light with skin has tended to use this approach [9], often based on models previously formulated for atmospheric physics and astronomy.

Geometric Optics

Geometric optics allows an immediate understanding of some of the phenomena that occur as light reaches an interface between two macroscopic media, such as that between the air and the *stratum corneum* when light first meets the most superficial layer of the skin, and may be expressed entirely by two laws: those of reflection and refraction ¹.

The law of reflection,

$$\theta_r = \theta_i \tag{2.4}$$

whereby the angle of reflection θ_r of a ray reflecting off the surface of a material is equal to the angle of incidence θ_i of the ray with respect to the surface normal may be observed in many real-life scenarios under the name of *specular reflections*, giving rise to what are known as *highlights* or *gloss* on the skin surface (as clearly visible at the top of the head of the author in Figure 2.5). These occur when a surface is smooth, and within the context of the skin will be affected by the hydration and tightness of the skin [9]. Importantly, highlights do not actually contain any information relative to the colour of the body imaged, as they are simply a reflection of the incident light.

¹These two laws may in turn be derived from the even more succinct principle that took Fermat's name, which states that light travels through the path by which it can reach its destination in the least amount of time. Unfortunately, however elegant, their derivation is outside even the most tangential scope of this thesis and the interested reader is referred to a Physics textbook such as [101].



Figure 2.5: Figure evidencing a highlight due to the specular reflection of the incident light off the top of the author's head.

Reflection also occurs at other interfaces within the skin, although it is particularly important at the interface between the dermis and the hypodermis, as the presence of the fat in the hypodermis presents a large smooth surface that reflects almost all the light back to upper layers. However, light is not always only reflected at an interface. The second law of geometric optics, the law of refraction, describes how the direction of the ray of incident light changes as a result of a property of the medium in which it is travelling, known as the complex index of refraction N , which is given for a particular wavelength of light λ and is expressed as

$$N(\lambda) = \eta(\lambda) + jk(\lambda) \quad (2.5)$$

where the real term $\eta(\lambda)$ is the refractive index, that indicates how much slower light will travel in the material than in a vacuum, whereas the imaginary term $k(\lambda)$ represents the extinction coefficient and describes how easily light can penetrate the material. The refractive index is key to the law of refraction, whereby

$$\eta_t \sin(\theta_t) = \eta_r \sin(\theta_i) \quad (2.6)$$

which states the relation between the angle of a transmitted ray θ_t as the ray passes from one medium to another in terms of the angle of incidence θ_i and the refractive indices of the two media. The laws of reflection and refraction may be combined into the Fresnel equations [101] to describe what fraction of light is reflected and what fraction of light is transmitted from one medium to another. For the interface between air and the *stratum corneum*, using these equations it has been calculated that around 4-7% of visible light incident at an angle close to the normal is reflected at that interface [59]. However, it should be noted that geometric optics will not hold accurately as other phenomena such as scattering will occur due to the topology of the stratum corneum, characterised by folds, pores, wrinkles, etc [98, 136]. Indeed, geometric optics may only be held to be valid when the wavelength of the incident light is significantly smaller than all other length scales in the system, the limit at which the electromagnetic treatment of wave optics through Maxwell's equations may be simplified to ray optics [107]. As a result, the skin does not always behave as a perfect specular reflector, and is often better approximated by a Lambertian reflector, a surface that is equally bright regardless of the angle between the observer and the light source, with a strong angular dependence as shown in Figure 2.6 [118].

Scattering

To understand the interactions that do occur at such a rough, inhomogeneous surface, the concept of scattering must be introduced. Scattering describes the behaviour of light when it is deflected by a body and may be broadly described as being either *elastic* or *inelastic*, whereby the energy (and thus the wavelength) of the photon is conserved or not, and as *forward* or *backward*, describing the direction of travel as a result of the interaction. The amount of scattering as a beam of light passes through a given volume is given by the scattering coefficient, which is defined for a radiant flux Φ_e at a given wavelength λ over a path length z as:

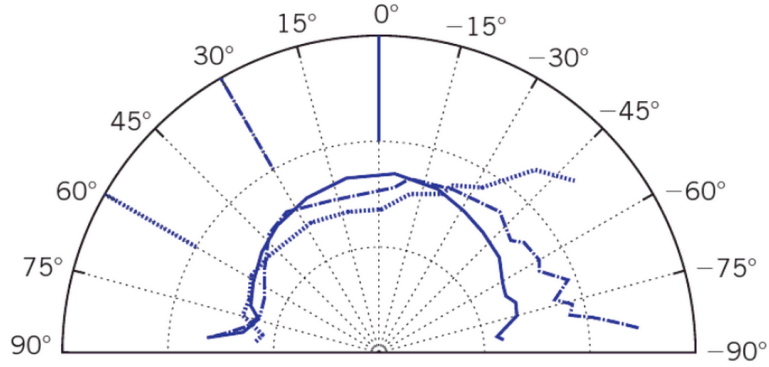


Figure 2.6: Curves showing the observed brightness of skin for a light source at three different angles of incidence ($\theta_i = 0^\circ, \theta_i = 30^\circ, \theta_i = 60^\circ$). These curves characterise the *bidirectional reflectance distribution function* of the skin - a perfect Lambertian surface would have no dependence on the angle of incidence as well as the angle of observation.

Image taken from [118].

$$\mu_s(\lambda) = -\frac{1}{\Phi_{e,\lambda}} \frac{d\Phi_{e,\lambda}}{dz} \quad (2.7)$$

The modes of scattering are usually grouped into three main regimes, determined by the size of the scattering body and the wavelength of the incident light. The dimensionless size α of a spherical scattering body of radius r with respect to the wavelength of the light λ may be written as:

$$\alpha = \frac{2\pi r}{\lambda} \quad (2.8)$$

When $\alpha \gg 1$, the scattering regime simplifies to the geometric optics case, as discussed above. On the other hand the regime $\alpha \approx 1$ describes Mie scattering, in which the wavelength is approximately equal to the size of the scattering body. This kind of visible light scattering occurs prevalently in the epidermis, and it has been found that it is characterised by forward scattering [23], indicative of relatively large particles although this is difficult to assert as Mie scattering assumes a spherical body. It should be noted however that the *stratum corneum* is also characterised by strong forward scattering,

and that this is mostly thought to be due to the alignment of collagen fibrils [170].

In addition to this, forward scattering in all regions of the epidermis has also been found to be wavelength dependent [79]. The dermis is instead characterised by both Mie scattering and Rayleigh scattering [79]. Rayleigh scattering occurs when the wavelength is much greater than the particle, i.e. $\alpha \ll 1$ and is characterised by a strong wavelength dependence, such that it is proportional to the fourth power of the frequency². This in turn means that light with shorter wavelengths (violet-blue in the visible light spectrum) is scattered more strongly than light in the red part of the spectrum, resulting in multiple scattering events before it is scattered into another layer or absorbed. It is for this reason that red and infrared wavelengths penetrate the dermis to a much greater extent than shorter wavelengths [21].

Absorption

The characteristic colour of skin is however determined by another optical property – that of absorption. Absorption may occur as a result of the presence of a pigment, a material that absorbs at particular wavelengths and therefore changes the colour of the light that is reflected or transmitted from it if it is absorbant in the visible part of the spectrum. The amount of absorption a beam of light undergoes as it passes through a medium is measured by the absorption coefficient, which, similarly to the scattering coefficient, is defined for a radiant flux Φ_e at a given wavelength λ over a path length z as:

$$\mu_a(\lambda) = -\frac{1}{\Phi_{e,\lambda}} \frac{d\Phi_{e,\lambda}}{dz} \quad (2.9)$$

Another common way of expressing this is instead in terms of the molar extinction

²It should be noted that this distinction between Mie and Rayleigh as different forms of scattering is used here as it is common in the biomedical literature but also technically incorrect. What is known as Mie scattering is in fact the solution of Maxwell's equations for a spherical scattering body (valid for all sizes), and Rayleigh scattering is nothing but the Rayleigh limit of Mie scattering for bodies with a diameter much smaller than the scattering body.

coefficient ϵ , which treats the absorption property as intrinsic to the pigment itself rather than considering a volume of medium through which the light passes. The molar extinction coefficient can be described in terms of the absorption coefficient μ_a , and the molar concentration c as:

$$\epsilon = \frac{1}{\ln 10} \frac{\mu_a}{c} \quad (2.10)$$

where the molar concentration is defined as the number density of the species divided by Avogadro's constant.

As noted in Section 2.3.1, there are several different pigments in the skin, each with a particular signature that contributes to the overall colour. The absorption spectra of each of these pigments have been studied extensively [79], and are briefly reported here.

Epidermis The main pigment in the epidermis is melanin. The absorption spectra in the visible part of the electromagnetic spectrum of the two main types of melanin, phaeomelanin and eumelanin, are shown in Figure 2.7. As may be observed from the figure, melanin is heavily absorbant in the blue part of the spectrum, while remaining relatively transparent to the longer wavelengths. This effect compounds the effect of Rayleigh scattering and leads to blue light being absorbed very quickly in the skin, particularly in individuals with a higher concentration of melanosomes, which can vary from a volume fraction of 1% in light-skinned caucasians to 43% in darkly pigmented Africans [78].

Although a number of objective ways, both *in vitro* and *in vivo*, of assessing the skin colour based on the concentration of melanosomes exist [121], a score known as the Fitzpatrick Phenotype Score is used as the standard in clinical practice to assess the colour of a subject [55]. The score should however be interpreted with caution as it was developed for a predominantly caucasian population and it is mostly concerned with the skin damage risks associated with skin colour rather than the colour itself. As a result,

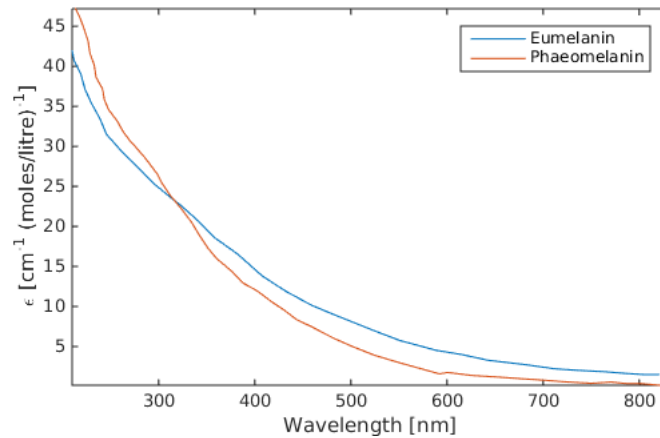


Figure 2.7: Extinction coefficient of eumelanin and phaeomelanin at different wavelengths. Source: S. L. Jacques, “Optical absorption of melanin”, Oregon Medical Laser Center, www.omlc.ogi.edu/spectra/melanin/, retrieved June 2015.

it is extremely unreliable as a gold standard [54, 162]. The Fitzpatrick scale, presented in Table 2.2 divides the population into six skin categories, from lightest (I) to darkest (VI) based on answers to phenotypic questions, exposure to sunlight, and penchant for burning.

Table 2.2: The Fitzpatrick Phenotype Score [55]

Skin Type	Appearance	Tanning
I	Very white or freckled	Always burns
II	White	Usually burns
III	White to olive	Sometimes burns
IV	Brown	Rarely burns
V	Dark brown	Very rarely burns
VI	Black	Never burns

Blood The absorption spectra of the main pigments in the blood – bilirubin, β -carotene, and oxygenated and deoxygenated haemoglobins may be found in Figure 2.8. The marked differences between the two haemoglobin species, and particularly the increased absorption of deoxygenated haemoglobin in the red part of the spectrum, are what give venous blood its characteristically darker colour with respect to oxygenated blood.

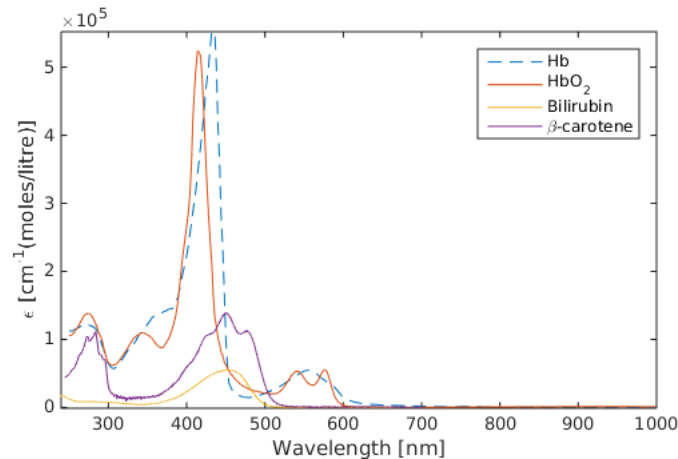


Figure 2.8: Extinction coefficients of deoxygenated (Hb) and oxygenated (HbO₂) haemoglobin are shown in the ultraviolet, visible, and near infrared wavelengths, together with the absorption spectra of bilirubin and β -carotene. Sources: S. Prahl, Oregon Medical Laser Center, www.omlc.ogi.edu/spectra/, retrieved June 2015.

The isobestic points of the two haemoglobin species, which are the points at which the extinction coefficients of the two species are equal, are also of note for this particular optical property. Also of note is the high absorption of bilirubin and β -carotene in the green part of the spectrum.

Other Pigments The absorption spectra of the two other species that are most preponderant in the skin – fat and water – are also shown in Figure 2.9. The very low absorption coefficient of water in and around the visible part of the spectrum gives rise to what is known as the “optical window”, a spectral region from around 650nm to around 1200nm in which tissue is relatively transparent to light.

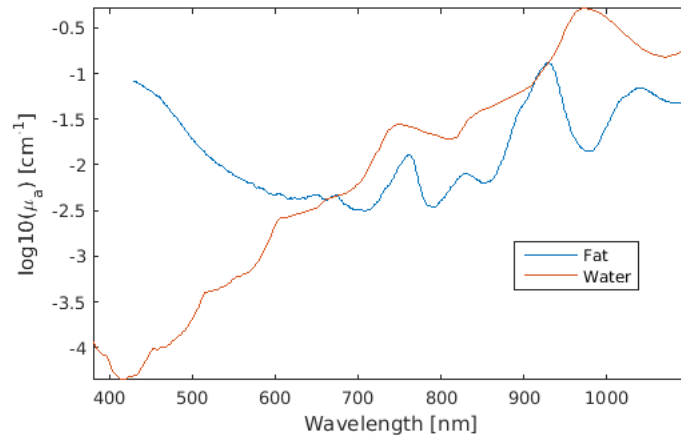


Figure 2.9: Base-10 log of the absorption coefficients of water and fat are shown in the visible and near infrared wavelengths. Sources: [224], [161], [150].

2.4 Measures of Blood Oxygen

Cyanosis, a bluish discolouration of the skin due to low oxygen saturation, made the link between the function of the lungs and the optical properties of the skin apparent to physicians as far back as the 18th century [130, 176]. However, it was only in the early 20th century that the role of oxygen as a cause of cyanosis was confirmed [111, 191]. Thereafter, cyanosis was routinely used as a clinical measure to evaluate anoxaemia (the lack of oxygen in the arterial blood) until the 1950s [33], in spite of the fact that the first studies [112, 191] had clearly demonstrated that only a severe anoxaemia ($80\%<$) would be visible to the naked eye, and it had been found that, even at 75% oxygen saturation, over one in ten experienced clinicians could not tell whether a patient was cyanotic [33]. Over the course of the 20th century, several different invasive and non-invasive methods, both *in vitro* and *in vivo*, of measuring the oxygen content of blood were developed to improve upon the accuracy of measurement. A small selection of the most relevant amongst these will be briefly reviewed in the following sections, leading up to the development of pulse oximetry, which, being the basis for much of the work carried out in this thesis, will be examined in some detail.

2.4.1 Physical Measures of Blood Oxygen

Blood gases have been extracted and measured since Boyle and Hooke devised a pump especially for that purpose in 1670 [22]. However, it was only with Lavoisier's discovery of oxygen itself in the late 18th century that the development of instruments to measure oxygen in blood could begin in earnest. Research carried out in the 19th century continued to employ blood-pumps to measure the volume of blood oxygen present in a sample, but the method was fraught with problems and required large quantities of blood to be drawn from the subject. With the turn of the century, improvements to the techniques allowed for smaller volumes and increased accuracy [176]. The true turning-point occurred with the accidental advent of polarography³ in 1925 and its later development in the form of the Clark electrode in 1956, which enabled PO_2 to be measured directly in blood [182]. This development led to the current gold standards for oxygen measurement: an updated version of the Clark electrode is in fact used to test blood samples *in vitro*, and the development of intravascular electrodes based on the same principles allowed for continuous monitoring of blood oxygen *in vivo* with catheters [177].

The principle behind the Clark electrode is that of other amperometric electrodes. The Clark electrode is constituted of a silver anode and a platinum cathode separated by a chlorine-rich electrolyte. When placed in a blood sample, oxygen diffuses across a membrane into the electrolyte. A voltage potential is applied to the electrodes and the current generated as the oxygen is reduced at the cathode and the silver anode oxidises is measured for different potentials. The current reaches a stable value at the potential at which the diffusion of oxygen limits the rate of reaction, and the plateau current value is correlated with the PO_2 in the electrolyte. Thus the Clark electrode measures the equivalent partial pressure of the oxygen dissolved in the blood, which constitutes one of

³Also known as voltammetry, polarography is an electrochemical method whereby the current passing through a sample in which two electrodes have been placed is measured as the electric potential between them is varied. Different chemical elements and compounds may be identified by the shape of the curve formed by the relationship between the current and the voltage.

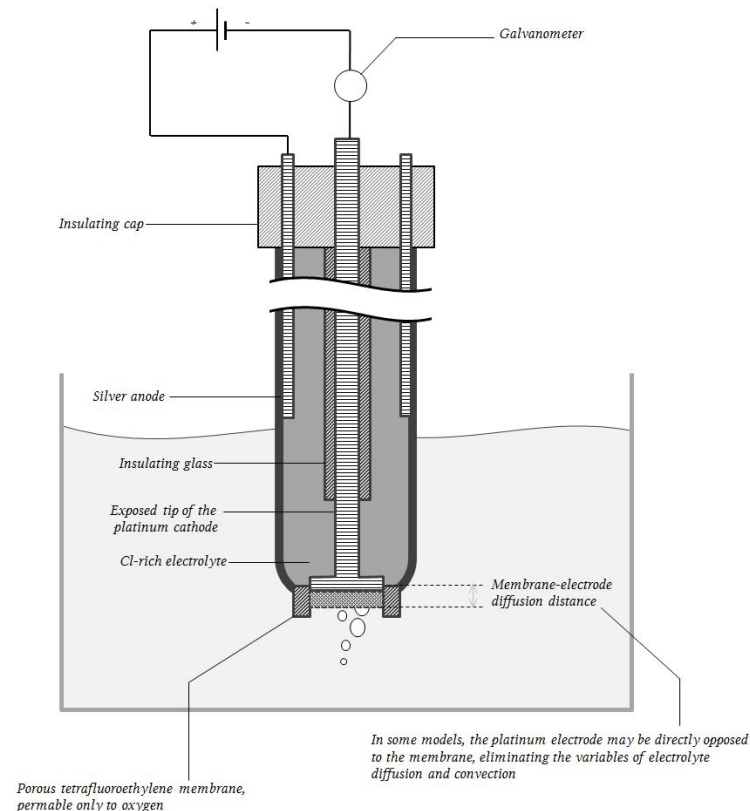


Figure 2.10: Schematic of the Clark electrode. (Image by Alex Yartsev, copied with permission from <http://ow.ly/WkMNR> Chapter 2.0.5, on the 28th of November 2015.)

the main advantages of polarographic techniques over the spectroscopic techniques that will be discussed in the following sections.

However, the gold standard suffers from a clear limitation and draw-back: the electrode needs to be in contact with the blood for the method to work. This has important implications for its clinical adoption and use, as the invasive nature of taking blood samples or using arterial lines means that the method exposes patients to increased risks of infection [173].

Transcutaneous measurements of PO_2 , in which a heated Clark electrode is placed on the surface of the skin, was an attempt to circumvent the invasive nature of PO_2 measurement by estimating the PO_2 of arterial blood (PaO_2) through the PO_2 of the skin (PtcO_2) [214]. The technique was first devised in 1851 by von Gerlach [230], but

subsequently benefited from the development of the Clark electrode.

In 1972 two groups reported that the $PtcO_2$ in neonates closely approximated their PaO_2 [47, 48, 70]. The relationship between the two however did not hold in adults or infants [214], as it was found that although the correlation between the two held in haemodynamically stable patients, the values of $PtcO_2$ were affected by ageing [212]. Furthermore, $PtcO_2$ was found to be a poor measure of PaO_2 during haemorrhagic shock in animal experiments, correlating instead with cardiac output, mean arterial pressure, and mixed venous PO_2 (PvO_2) [152, 166, 213]. Although the authors concluded that this meant $PtcO_2$ was a more useful measure than PaO_2 to infer tissue oxygenation, the variable results in clinical studies meant that the measure was regarded as unreliable [214]. As a result, the technique was largely supplanted by the contemporary development of pulse oximetry in clinical practice [178]. The inferences from this brief window of innovation are however extremely useful, as they are very much paralleled by the difficulty methods for measuring superficial tissue oxygenation have had in establishing a theoretical basis.

2.4.2 Optical Measures of Blood Oxygen

An altogether different method of measuring blood oxygen is that based upon the spectral characteristics of blood, and particularly haemoglobin, in its oxygenated and deoxygenated states, as considered in Section 2.3.2. The apparatus, based on the spectroscope developed by Bunsen and Kirchhoff in 1860 and rooted in Newton's study of the "spectrum" of colour produced by light passing through a prism ⁴ [91], was made possible by the development of photoelectric cells and the discovery of haemoglobin in the second half of the 19th century [181]. The technique, now known as oximetry, was developed by von Vierordt in the late 19th century but only revived by Nicolai over half

⁴It is curious to note that while Hooke was developing the aforementioned blood-pump, he was also actively hampering the development of the optical techniques that would one day supersede his invention, by sustaining that the spectrum of colours was created by the prism itself and delaying Newton's publication by over 10 years [181].

a century later [137]. Oximetry knew a period of rapid development as a result of the studies on aviation physiology during World War II, and the field was advanced by the likes of Kramer [97] and Matthes [120] in Germany and Squire and Millikan in the UK and US [128, 190]. However, due to the unreliability of the electronic equipment available, and despite the advances made by Wood in the post-war period, oximetry never made it to clinical practice, and was confined to the realms of *in vitro* cuvette analysis. It was only with the development of pulse oximetry in the 1970s that clinicians began to think in terms of oxygen saturation (SO_2) rather than solely in terms of oxygen tension (PO_2) [178].

Oximetry

In his seminal work published in 1932, Nicolai employed a mercury vapour light and green-blue filters to determine changes in oxygenation as the circulation to a hand was reduced [137]. His associate, Kramer instead used an incandescent light and a red filter to measure saturation in a cuvette [97], and later showed that oxygen saturation could be measured continuously. Crucially, Kramer showed that the Beer-Lambert law, whereby the intensity of transmitted light is a logarithmic function of the density of the absorbing molecules [13], held both for haemoglobin solutions and approximately for whole blood [97]. The law, first reported by Bouguer and Lambert in the 18th century and later rediscovered by Beer in the 19th, is based on the property of transmittance T of a material, defined as:

$$T = \frac{\Phi^t}{\Phi^i} \quad (2.11)$$

where Φ^t is the radiant flux transmitted across the material and Φ^i is the radiant flux incident on the material. This property may also be defined in terms of the optical depth τ of the material and the absorbance A of the material by:

$$T = e^{-\tau} = 10^{-A} \quad (2.12)$$

The Beer-Lambert law (graphically shown in Figure 2.11) simply states that for N different attenuating species present in the material, the optical depth may be represented as:

$$\tau = \sum_{j=1}^N \tau_j = \sum_{j=1}^N \sigma_j \int_0^l n_j(z) dz \quad (2.13)$$

where σ_j is the attenuation cross-section of the attenuating species j in the material, n_j is the density of the attenuating species j in the material, and l is the path length of the light through the material. Similarly, the material's absorbance may be expressed as:

$$A = \sum_{j=1}^N A_j = \sum_{j=1}^N \epsilon_j \int_0^l c_j(z) dz \quad (2.14)$$

where ϵ_j is the molar attenuation coefficient of the attenuating species j in the material and c_j is the concentration of the attenuating species j in the material.

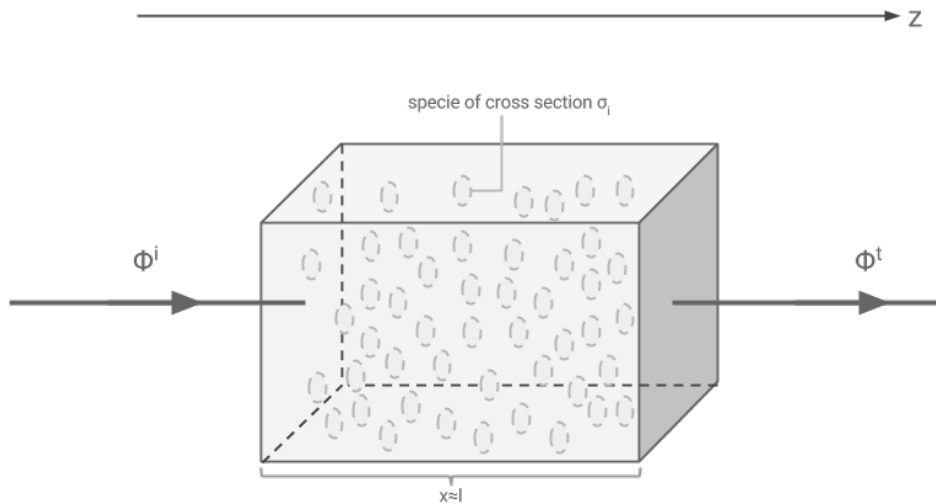


Figure 2.11: Graphical representation of the Beer-Lambert law. If a material is constituted of an attenuating species (dashed circles), the radiant flux transmitted across the material is a logarithmic function of the concentration of the species in the material.

Importantly, this law means that, for a constant path length l , and for known (different) values of $\epsilon_i(\lambda)$, the relative concentrations of the different species may be inferred if the intensity of the incident light at wavelength λ is known and the intensity of the transmitted light is measured. It was this principle that underpinned all advances in oximetry, culminating in the development of the multiple-wavelength CO-oximeter in the 1960s. This oximeter allowed the measurement of many of the pigments present in the blood and is now the *de-facto* gold standard for *in vitro* measurements of oxygen saturation in arterial blood gas (ABG) analysis [132].

Although the same principle is also the basis for *in vivo* methods in direct contact with the blood such as fiberoptic catheter oximeters [181], the Beer Lambert law only simplifies to analytical solutions for monochromatic light in a homogenous, non-scattering medium [132]. These conditions are only true for haemoglobin solutions, as whole blood is a non-homogeneous medium with non-linear scattering properties [159]. The Beer-Lambert law usually fails when applied to transcutaneous blood oxygen measures, as the path length is not truly related to the distance between the light source and the photodetector. It is also difficult to apply to whole blood unless all other pigment species have been accounted for. As a result, most CO-oximeters now employ a large range of different wavelengths in the visible part of the spectrum, and require haemoglobin solutions rather than whole blood as an input. [132].

Transcutaneous measurements also had the drawback that the light was passing through both the blood and the tissue surrounding the blood. Furthermore, it was impossible to distinguish the venous blood from the arterial blood and the measurements made were influenced by both. Millikan's ear oximeter [128], which used a red filter to track the oxygen-dependent component and a green filter to track the oxygen-independent component, was in fact only a relative measure of oxygen saturation (it could track changes but would need to be calibrated first to obtain absolute values), and it was also discovered that the green filter and sensor were actually responding to infrared light due

to the opacity of the ear lobe to the green light [181]. Absolute measurements were only made possible when Wood modified the Millikan oximeter [237] so as to mechanically compress the tissue to obtain a zero-value that could be subtracted from subsequent readings. The Wood oximeter however never became accepted in clinical practice due to the variability and inaccuracies in the instrumentation used [181].

Recently, a *tissue* oximetry – more akin to the PtcO_2 measurements with the Clarke Electrode than to pulse oximetry – has had a brief resurgence thanks to a renewed interest in measuring the oxygen saturation of local regions or organs, for example to determine the viability of skin flaps in surgery [89], or to estimate changes in the oxygen saturation of the brain [174]. A large body of research has been dedicated to the study of near-infrared spectroscopy (NIRS) to evaluate oxygenation in large organs which can be characterised by broad ranges of oxygen saturation (ranges as wide as 48-88% have been reported [100, 129]) [49, 56, 233]. Indeed, the choice of near-infrared wavelengths is ideal to probe tissues which are several centimetres below the skin due to the relative transparency of tissue to infrared wavelengths, and it has been shown to be responsive both to ischaemia and to hypoxaemia, as well as predictive of the development of organ dysfunction during shock resuscitation [32].

Other studies have instead been carried out using visible light spectroscopy (VLS) to evaluate local estimates of the microvascular oxygen saturation (StO_2 ⁵) by matching the reflected spectrum to the known spectra of the constituent elements via least-squares fitting [15, 239]. These also indicated that VLS is sensitive to both regional and global ischemia and hypoxemia, with the results from the VLS oximeter strongly correlated with those from a standard pulse oximeter [15]. Furthermore, a high correlation was found between the StO_2 measured using a VLS oximeter placed on the buccal mucosa and central venous oxygen saturation [239]. A strong correlation between StO_2 and SaO_2 was also reported by [134], who measured cutaneous oxygen saturation using red and

⁵ StO_2 is still a poorly defined concept, including anything from the tissue oxygenation of the brain, muscle, or skin. Here we take StO_2 as meaning the cutaneous oxygen saturation.

near-infrared light, but found that StO_2 was more closely related to the mixed venous blood oxygen saturation SvO_2 .

Pulse Oximetry

The development of pulse oximetry marked a breakthrough for patient monitoring, as it allows accurate and continuous measurements of arterial oxygen saturation. This was achieved through the serendipitous realisation that it is principally the arterial blood that has a pulsatile component and that this knowledge may be used to isolate the arterial blood from the surrounding tissue and from the venous blood without the mechanical application of pressure [178]. The physiological reason behind this is that arteries have elastic walls that expand more during systole than during diastole as a result of the volume change whereas the veins do not behave in a contractile manner. Another contributing factor to the rapid popularisation of the methodology was the development and mass-production of LEDs, that approximated the monochromatic requirements of spectrophotometry [132].

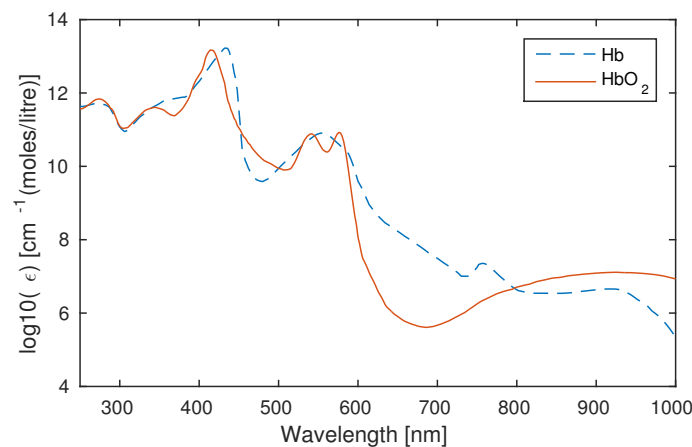


Figure 2.12: The log absorption spectrum of oxygenated and deoxygenated haemoglobin. Sources: S. Prahl, Oregon Medical Laser Center, www.omlc.ogi.edu/spectra/, retrieved June 2015.

The method works as follows. Two lights at two different wavelengths are shone through, or reflected back from, the tissue. One of the wavelengths is in the red part of

the spectrum and one in the infrared part of the spectrum, such that in the former the oxygenated blood is more absorbant and in the latter the deoxygenated haemoglobin is more absorbant (see Figure 2.12). Each light gives rise to a *photoplethysmogram* (PPG), a time-varying signal caused by changes in the volume of blood imaged: as the blood in the tissue increases during systole, a higher proportion of the light is absorbed. For each signal, the cardiac-synchronous pulsatile component, usually named AC because of the electrical engineering roots of the technology, is normalised with respect to the baseline, or DC, value (see Figure 2.13). The ratio of these two normalised signals is related to the arterial oxygen saturation via empirically determined look-up libraries [132].

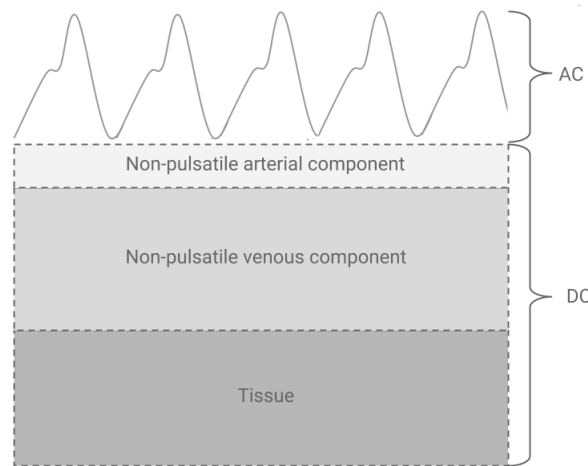


Figure 2.13: The absorption of light in skin is due to pulsatile (AC) and non-pulsatile (DC) components, shown for the case in which the signal is reflected. The overall signal, composed of the AC and DC components, is known as the photoplethysmogram (PPG) – the signal traced by the absorption of differing volumes of light.

The technique, first invented by Takuo Aoyagi in 1972 [180], had an initial theoretical grounding in the Beer-Lambert Law, and although it was soon demonstrated that the equations did not apply [186], both for the reasons discussed in Section 2.4.2, and for others associated with the pulsatility itself, it is useful to review the original theory as it provides a starting framework upon which several other authors have based their own

work and to pin-point simplifying assumptions and discuss their validity.

From Equation 2.14, for a monochromatic light passing through an non-scattering, homogeneous medium composed only of two species (in this case oxygenated haemoglobin and deoxygenated haemoglobin) with molar attenuation coefficients $\epsilon_o(\lambda)$ and $\epsilon_r(\lambda)$, we may express the concentrations as $c_o l(\lambda)$ and $c_r l(\lambda)$ (homogeneous and non-scattering assumption), where $l(\lambda)$ is the path length of the light through the medium. The relationship between the incident light intensity $I_i(\lambda)$ and the transmitted light intensity $I_t(\lambda)$ may then be written as:

$$I_t(\lambda) = I_i(\lambda) 10^{-l(\lambda)(\epsilon_o(\lambda)c_o + \epsilon_r(\lambda)c_r)} \quad (2.15)$$

The next assumption made is that of photoplethysmography (PPG): the pulsatile changes are due to light being absorbed as the result of blood volume changes that cause a change in path length, such that it may be expressed as:

$$l(\lambda, t) = l_{DC}(\lambda) + l_{AC}(\lambda, t) \quad (2.16)$$

which in turn means that the transmitted signal may now be expressed in terms of a DC component and an AC component, such that:

$$I_t(\lambda) = I_i(\lambda) 10^{-(l_{DC}(\lambda) + l_{AC}(\lambda, t))(\epsilon_o(\lambda)c_o + \epsilon_r(\lambda)c_r)} \quad (2.17)$$

From this equation, it is then easy to show that the so-called “ratio of ratios”:

$$R = \frac{\log_{10}\left(\frac{I_{AC}(\lambda_1)}{I_{DC}(\lambda_1)}\right)}{\log_{10}\left(\frac{I_{AC}(\lambda_2)}{I_{DC}(\lambda_2)}\right)} \quad (2.18)$$

is related to the fractional definition of oxygen saturation, namely:

$$SpO_2 = \frac{c_o}{c_o + c_r} \quad (2.19)$$

via the equation:

$$SpO_2 = \frac{\epsilon_r(\lambda_1) - \epsilon_r(\lambda_2)R}{\epsilon_r(\lambda_1) - \epsilon_o(\lambda_1) + [\epsilon_o(\lambda_2) - \epsilon_r(\lambda_2)]R} \times 100\% \quad (2.20)$$

The concept of photoplethysmography and its relation to the path length is key to pulse oximetry, as it allows the division of the signal into the AC and DC components. However, it is not a complete description of reality, as studies have found that a pulsatile signal will also appear when blood is pumped through a rigid vessel, and that this was as a result of the erythrocyte axis changing as a result of the pressure wave [139], as shown in Figure 2.14. In addition to this, the shape and size of the photoplethysmograph is also affected by changes in the total concentration of red blood cells, rouleau formations (red blood cell stacking), and local blood velocity [117], all of which will have an effect on the amount of light scattered and absorbed by the blood.

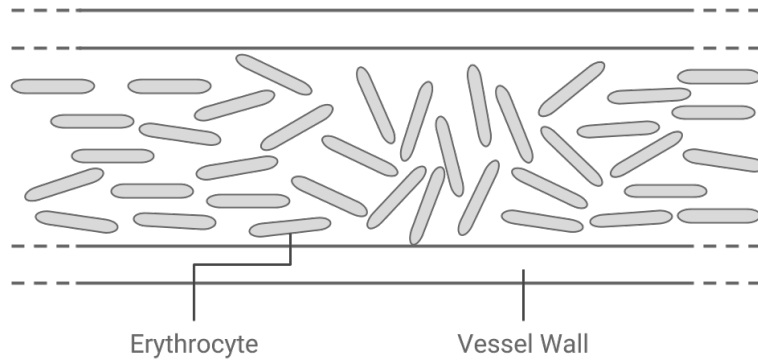


Figure 2.14: The path length and transmission and reflectance properties of blood are affected by the change in erythrocyte orientation as the pressure pulse passes through the vessel.

The invalidity of many of the assumptions made have led to the use of look-up tables in pulse oximeters, with values that are derived from studies conducted on sample populations of human volunteers (usually non-smokers with haemoglobin type A as this is the most prevalent in the population) rather than calculating the oxygen saturation directly from the ratio of ratios [132]. Several different calibration methods have been proposed, accounting for differences either by making the ratio of ratios R the argument

of a polynomial function [52]:

$$SpO_2 = (k_1 + k_2R + k_3R^2) \times 100\% \quad (2.21)$$

or by replacing the physiological variables that are part of Equation 2.20 with empirically determined constants [126]

$$SpO_2 = \frac{k_1 + k_2R}{k_3 + k_4R} \times 100\% \quad (2.22)$$

These improvements have allowed pulse oximetry to be routinely used in the clinic, with an accuracy of between 2-3% in measuring arterial oxygen saturation over 80% (it is unethical to test lower values) with respect to the CO-oximeter [65]. The technique provides a continuous sampling of oxygen saturation, allowing clinicians to spot acute hypoxia rapidly.

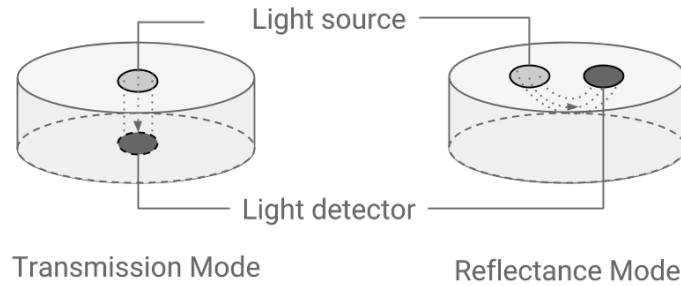


Figure 2.15: Pulse oximetry may be used in transmission or reflectance mode.

Pulse oximeters function both in transmission and in reflectance mode. In transmission mode the photodetectors are placed at a known distance from the light-sources either side of the tissue. In reflectance mode, the photodetectors are instead placed beside each other, next to the light sources, and the light is scattered back (see Figure 2.15). Transmission pulse oximetry sites are chosen such that there is an abundant vasculature

but not too much tissue to absorb the totality of the signal or cause burns due to the high light intensity [24]. Ear lobes, fingers, or toes, all of which benefit from their role in heat transfer as they tend to have a blood flow that is greater than that needed for tissue metabolism [145], are therefore usually chosen, and fingers are preferred wherever possible [132, 138]. Although reflectance mode pulse oximetry normally has a lower signal-to-noise ratio than transmission pulse oximetry [127, 200], the versatility of the reflectance pulse oximeter has allowed pulse oximetry to be employed not only in the periphery of the body but in other locations with superficial vasculature such as the forearm, chest and forehead. Furthermore, the dependence of transmission pulse oximetry sites on regulation by the autonomic system has led to reflection pulse oximetry being held as more accurate in subjects with low peripheral perfusion, for example as a result of lower cardiac output or lower environmental temperatures [3, 127, 172].

The differences in the pulses obtained from different sites has been studied extensively [243], and simultaneous multi-site measurements are thought to give increased information as to the state of the cardiovascular system, such as blood pressure and vessel elasticity.

Limitations of Pulse Oximetry

Although in the clinic it is often assumed that SpO_2 , the peripheral oxygen saturation as measured by pulse oximetry, is equivalent to the arterial oxygen saturation SaO_2 , the relationship is not always clear nor valid. Clinical studies have shown that the accuracy of pulse oximetry varies significantly in the 80% to 100% operating range, and although manufacturers claim an accuracy of 2% (where the accuracy is taken as the standard deviation of the difference between the SpO_2 and the SaO_2 [12]), the accuracy for single measurements was found to worsen to between 3% and 4% [12, 145]. A study on critically ill patients in a general intensive care unit (ICU) has further shown only a moderate correlation ($r^2 = 0.37$) between spontaneous SpO_2 and SaO_2 changes [154], leading the authors to conclude that SpO_2 is not a reliable measure of arterial oxygen saturation for

this patient population.

The reason behind many of these discrepancies in measurement is to be found in the empirical nature of the calibration, which is carried out on healthy volunteers and only for oxygen saturation values above 80% [132, 145]. These problems are exacerbated by the presence of other haemoglobin types, such as foetal haemoglobin, or by the binding of haemoglobin to other molecules, as discussed in Section 2.2.2. In recent years, a number of alternative techniques have emerged in order to obtain calibration-free measurements of oxygen saturation, relying on the morphology of the photoplethysmogram [164], on increasingly accurate physiological constants [144], or on the availability of multiple wavelengths [143], however these have yet to be tested extensively [145].

Foetal haemoglobin can affect pulse oximeter readings due to the differences in the oxygen-haemoglobin dissociation curves between foetal blood and infant blood and this increases the error expected of pulse oximetry in neonates up to between 3-4% [145]. As a result, authors have recommended a much higher and narrower safety range for pulse oximeters in neonates (95-97%) [185]. Three key studies have instead highlighted the importance of accurate pulse oximetry in preterm infants: COT (Canadian Oxygen Trial), BOOST (Benefits of Oxygen Saturation Targeting) II, and SUPPORT (Surfactant, Positive Pressure, and Oxygenation Randomised Trial) [171]. These studies examined the effect of maintaining the SpO_2 either at low (85%–89%) or high (91%–95%) values in a large population (4911) of preterm newborns, finding an increased mortality rate in the patients in the former group but also finding an increase in the incidence of retinopathy in the latter group. The authors of the meta-analysis concluded that preterm infants of a gestational age of under 28 weeks should be maintained at a target SpO_2 of 90 to 95%, clearly indicating the need for accurate pulse oximeters.

Dysfunctional haemoglobins (dyshaemoglobins), also have an effect on measurements with pulse oximeters as they too are pigments in the wavelengths selected for pulse oximetry. The two main dyshaemoglobins (MetHb and COHb) have a negligible effect

on measurements in the low concentrations that are present in healthy volunteers, but immediately begin to have an effect in patients with higher levels, such as patients with CO poisoning. It has been estimated that every percentage change in the concentration of COHb will change the pulse oximeter reading by an equivalent percentage point [57]. As with CO-oximeters, some manufacturers have attempted to get round this problem by increasing the number of wavelengths used in pulse oximeters [145].

In addition to the role of the species of haemoglobin present in the blood, other effects also impinge on the accuracy of SpO₂ readings. These include venous pulsations, in which pulsatile veins and venules contribute to the signal that is measured with the pulse oximeter and decrease the estimate [193, 199]; the effects of local hyperthermia, hypothermia, and vascular tone, on the recorded SpO₂, thought to arise as a result of changes in the reflectance properties of the skin which in turn affect the calibration of the pulse oximeter [145]; and the effect of skin melanin concentration, as to which a number of conflicting studies exist [132]. Finally, hypoxia itself has an effect on the volume of blood available to be sampled, with acute hypoxia reducing the blood flow to the skin and chronic hypoxia having a vasodilatory effect [20].

Technical limitations also contribute to inaccuracies in pulse oximeters. Mechanical artefacts caused by patient movement or tremor decrease the quality of the signal, and manufacturers have applied a number of signal processing techniques to assess the signal quality in these instances [156]. Light will also interfere with the functioning of the pulse oximeter and needs to be eliminated. This is particularly troublesome in pulsatile lighting such as that from fluorescent lamps as aliasing will lead to an artefactual periodic trace appearing [132]. Finally, malpositioning leads to the penumbra effect, whereby the SpO₂ readings are erroneously low, due to the reduced amplitude of the AC component that may lead to inaccurate calculations of R [10].

Despite these limitations, pulse oximetry has become ubiquitous in and out of the clinic, mainly thanks to the recognised importance of being able to tell whether the oxygen

saturation of a subject is changing, regardless of the accuracy [145] of the measurements. The safety of the device has also contributed to its popularity, although it still retains rare complications associated with compliance, burns, and pressure necroses [24, 132].

2.5 Remote Photoplethysmography

2.5.1 Origin of the Remote Photoplethysmogram

Key to the determination of heart rate within the remote vital sign monitoring framework is the concept of remote or imaging photoplethysmography (rPPG or iPPG respectively). The concept, first introduced by Blazek for infrared light in 2000 [18] and by Verkruyse for ambient visible light in 2008 [225], relies on the same phenomenon as that which gives rise to reflectance pulse oximetry. Ambient or directed light is absorbed by tissue much in the same way as it would be in the case of the pulse oximeter and the video camera is capable of sensing a change in the blood pressure pulse during systole and diastole as a result of a change in volume, and in the absorptive or reflective properties of the perfused tissue.

The exact nature of the origin of the photoplethysmogram as detected by the camera is still under debate. In addition to the explanations based on change of volume as in the conventional pulse oximetry literature, other theories have been put forward. The original authors of rPPG for near-infrared postulated that the photoplethysmogram would typically be caused by a change in volume in the venous beds, leading to a higher absorption by the blood [18]. This was generally taken to be the case in all subsequent experiments using visible light cameras, although most authors assumed that the vasculature imaged with visible light would be the capillary bed [225] (see Figure 2.17).

More recently, part of the photoplethysmogram seen by visible light cameras has been shown to originate from changes in reflectance from the skin as a result of global

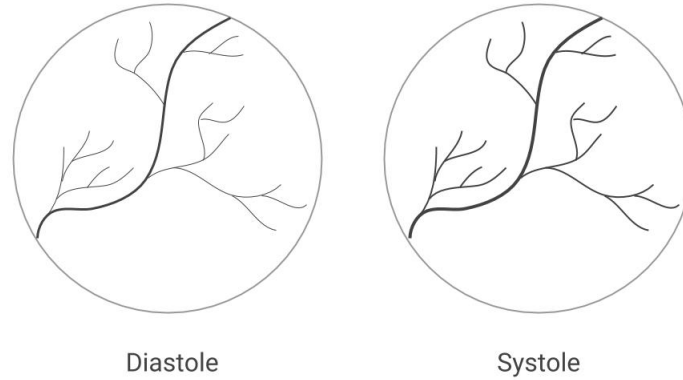


Figure 2.16: The arrival of a pressure pulse during systole is thought to cause an increase in volume in the blood vessels leading to more light being absorbed by the tissue.



Figure 2.17: The colour change observed by a camera as blood moves in and out of the skin was amplified in [238], providing further evidence for the blood-volume change hypothesis. Image from [238].

musculoskeletal motion due to the subtle movement of the head as the blood is ejected into the carotid artery [8], and this would indicate that part of the remote photoplethysmogram would therefore have no blood-absorption component at all. Finally, [87] have recently published results that would indicate that the blood-induced component is due to an elastic-mechanical compression of the upper dermal layers by deeper arteries as they expand. This last theory is strongly disruptive with respect to previous models, whether assumed tacitly [157, 225] or explicitly [18], and indicates that PPG signals from different regions of interest (ROIs) may be in antiphase with one another (this will be explored in greater detail in Section 6.4).

2.5.2 Video and Colour

In order to understand how a video camera may record a photoplethysmographic signal, an understanding of the functioning of video cameras and of the origin of the colour signals produced by said cameras is required.

Imaging Sensors

The core component in a digital video camera system is the imaging sensor, which registers the intensity of light incident on it and outputs a current or a charge that is proportional to that intensity, which is then converted into a digital signal before it can be processed. Two competing types of imaging array technologies currently lie behind the imaging sensors in use in most video cameras: charged coupled device arrays (CCD) and complementary metal oxide semiconductor arrays (CMOS). Both image sensor types are constituted of a grid of light-sensitive elements that create an electrical signal proportional to the intensity of light incident on them. CCD sensors rely on the build-up of negative charge in a potential well which is then read out at the end of the exposure time [204]. CMOS sensors instead measure the difference in charge between a reading prior to the exposure period and the reading following that period, as charge drains from a capacitor to a photodiode in proportion to the intensity of incident light upon the CMOS sensor [16, 204].

The difference in read-out of these two imaging arrays leads directly to the different properties of the video cameras that employ them. The additional electronics required for each imaging element within CMOS sensors (which typically require a capacitor and a transistor), reduce the surface area available for the photo-sensor itself with respect to CCDs. On the other hand, due to the over-accumulation of charge, CCD elements may overflow onto neighbouring pixels in an effect known as blooming, which does not occur in CMOS sensors. Furthermore, the per-pixel processing that can be achieved with CMOS sensors means that the noise levels may be decreased, the acquisition speed

increased, and the pixels addressed and buffered individually. As a result, although CCDs had typically been considered the technology of choice for scientific applications [114], they have been largely displaced by the cheaper and lower-power CMOS sensors in the consumer market and the performance of the two is virtually indistinguishable in more specialised domains [25].

Colour

For obvious reasons, digital video cameras in many ways mimic the physiology of the human eye in terms of how colour is recorded. Most cameras rely on a system whereby each photo-sensitive element is overlaid with a colour filter that only allows light of certain wavelengths to pass. The filters are usually chosen such that they correspond to the three primary colours – red, green, and blue (RGB) – that parallel the response of the different types of cone cells present in the retina (types L, M, and S respectively). However, other filter types exist, including the RGBE filter (red, green, blue, emerald), and the CYGM filter (cyan, yellow, green, magenta). The colours are arranged into particular patterns in order to sample the entire image as accurately as possible with respect to the human vision. In RGB cameras the green filter tends to be present more often than others as it tends to emulate the action of the M and L cones combined. The most common of these patterns or mosaics is the Bayer filter shown in Figure 2.18a, whereby the filters are arranged in a 2×2 square pattern with two green filters and one red and blue filter (Figure 2.18) [110].

In order to derive the actual colour at any given pixel, neighbouring pixels are interpolated so as to obtain individual pixel values in terms of RGB. As a result of the interpolation, some spatial colour information is lost when acquiring an image with a Bayer-type filter. Other methods of acquiring colour at a higher resolution exist, such as by stacking colour sensors with particular colour sensitivities on top of one another [62] as shown in Figure 2.18c, or by splitting the light with a dichroic prism prior to

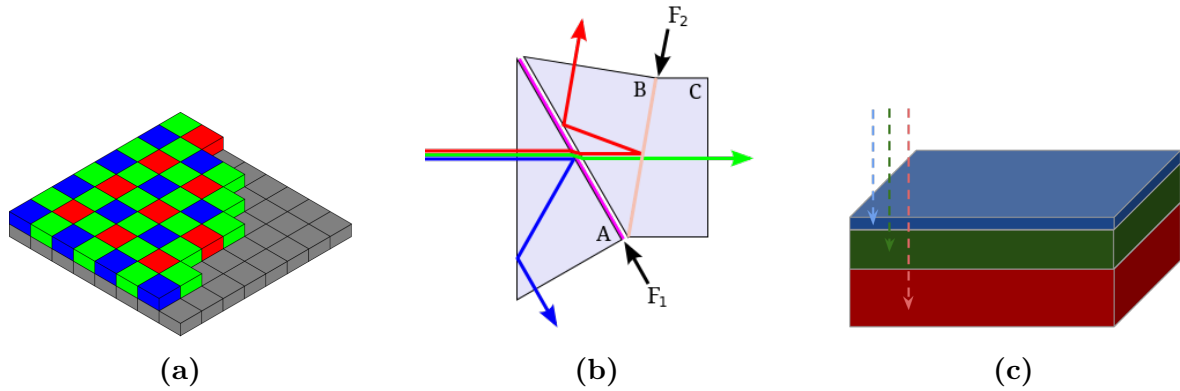


Figure 2.18: Figure showing different filter architectures. (a) The Bayer filter pattern array. (b) The dichroic prism for a 3-CCD system. (c) Stacked photodiodes with colour-specific sensitivities. (Source: Colin Burnett, CC BY-SA 3.0)

detecting it on three or more separate image arrays as shown in Figure 2.18b.

Once acquired, the RGB values represent a particular colour within the device's RGB colour space, a three dimensional volume characterised by a subset of the colours (the gamut) that may be expressed within the RGB additive colour model. A number of different colour spaces exist within the RGB colour space family, each having a different subset of colours that it can express. Of these, the most common is sRGB, used in most consumer-grade digital video cameras and monitors. In addition to the RGB colour space family, other colour space families also exist, such as the luminance-chromaticity colour models (that include HSV and HSL colour spaces (hue, saturation, and value or lightness), where the properties of colour are given by the brightness of the colour (the luminance/value/lightness property) and its chromaticity (that is independent of brightness and is given by the hue and saturation); and the luminance-chrominance models (that include the YCbCr, the YUV, and the $L^*a^*b^*$ colour models to name but three popular ones), where again the colour is given by the dual properties of brightness and chrominance, but where chrominance, unlike chromaticity, is scaled if the colour is attenuated. Each of these colour spaces represent colour in different ways (and often have different definitions for the same quantities) but transforms exist to convert some of the colour spaces into others, although this may be meaningless when attempted between

poorly defined colourspaces or between colourspaces with different gamuts [74].

The role of colour spaces in digital cameras is to convert the colour from the device's RGB colour space to one or more standards in order for it to be displayed correctly on a screen (as with the conversion from the device RGB to sRGB), or to facilitate the compression of colour without any significant loss in image quality, as occurs in many consumer grade cameras when the picture is transformed from sRGB to YCbCr for it to be stored in a lossy (JPEG) format. However, the different representations of colour have also been used to facilitate certain calculations with colour in computer vision, such as the identification of skin [28], much in the same way that cylindrical or spherical co-ordinates may facilitate the solution of a problem with respect to cartesian co-ordinates.

Finally, in addition to the individual, per-pixel changes colour identification, cameras must also simulate another property of the human visual system: that of adapting to the colour of the light source(s). This is termed white balance, as the human brain will attempt to render the colour as if it were illuminated by a white light source regardless of the illumination incident on the colour.

Image Acquisition and Correction

It should be clear by this point that the imaging sensor is only a small component, however crucial, within a digital camera system, and that several other systems come into play to ensure that the images acquired are of a high quality to a viewer regardless of the scene. Figure 2.19 illustrates some of these components (a good number are omitted for clarity, but the reader is directed to [102, 232] for further details). Not only do these components change the true nature of the colour acquired at the sensor (as would be output by the RAW data), but they also define the core acquisition properties of the video camera: the bit depth (the number of bits available to describe the intensity of the light of a given colour), the frame rate (the speed at which the frames are acquired), and the image resolution (the total number of pixels available).

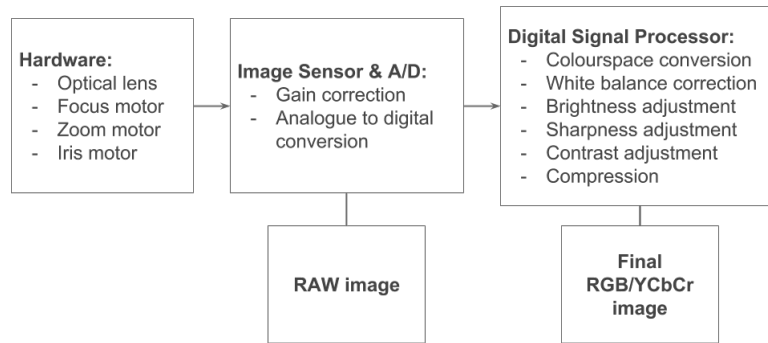


Figure 2.19: Figure illustrating some of the key components in the image acquisition and correction pipeline.

2.5.3 From Video to rPPG

As mentioned in the introductory section of this chapter, the video camera functions as a remote pulse oximeter operating in broad-band visible light, sensing changes in the perfusion of the tissue in order to convert a video stream into a signal with a pulsatility synchronous to the beating of the heart. Despite the great variation in methods used to determine which part of the skin and body to use as a source of the signal, and an even wider plethora of methods employed to filter the signal, the basic method to convert the video into a one-dimensional signal in time has remained virtually unchanged since it was first developed [18, 45, 71, 99, 105, 157, 158, 202, 225, 227, 236].

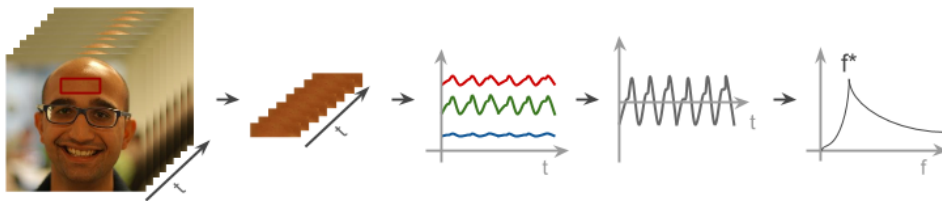


Figure 2.20: Schema illustrating the procedure common to all prior art by which the video is reduced to a set of one-dimensional signals and hence to a frequency.

The method, shown in Figure 2.20 is as follows. A region of interest (ROI) is first identified and tracked such that it should remain relatively constant with respect to the skin imaged, and the spatial mean of each colour channel of that region of interest is

computed throughout the video segment of interest. This yields the set of one-dimensional signals for each channel. Each of the signals can then be processed and analysed spectrally in order to find the frequency of the pulsatility.

The identification and tracking of the region of interest has so far been treated fairly simplistically in the literature, perhaps in recognition of the fact that this is a separate topic within the field of computer vision. Early research was conducted on motionless subjects facing the camera directly and the region of interest was either selected manually [225, 227] or by detecting the face via a trained Viola-Jones cascade detector [30, 157]. Their inherent simplicity has meant that these two methods have remained popular in most subsequent research, particularly when introducing novel concepts with little or no bearing to motion [29]. More advanced algorithms have also been introduced to select specific regions of interest within the face based on the position of detectable landmark features such as the edges of the eyes, nose and mouth [105]. These however require the subject to remain facing the camera and thus do not cope well with motion. Other authors have instead used the whole face as a single region of interest, using skin detection algorithms to create a mask within the region of the face [7, 19], whereas [223] propose the division of the entire frame into automatically generated triangular sub-regions that can then be analysed in order to differentiate skin from non-skin and optimally combined.

As noted by [202], the size of the region of interest is of considerable importance as it determines the level of noise present in the signal. The sensitivity of a video camera – its signal to noise ratio (SNR) – is too low for the heart-rate pulsatility to be apparent when taking the value of a single pixel or even the mean of a small subset of skin pixels, however by averaging over a larger area the noise will be decreased [99], assuming that the noise of each pixel is independent. In addition to this, smaller regions of interest run into the problem that breathing rate and motion components will dominate the signal [202]. This occurs because highly contrasting features have a larger effect on smaller regions of interest, as will be considered in detail in Chapter 6.

Further to the size of the region of interest, other factors come into play when determining its location and the number of regions chosen. Regions of interest have always been centred on regions of exposed skin, with authors obtaining signals from the forehead [196], cheeks [198], hands [73, 167, 168], fingertips [167, 244], and even soles of the feet [165]. Often, in order not to be restricted to a large single region of interest as in [94], authors have instead averaged multiple smaller regions of interest, taken from the face [99, 223], or from the hands [86, 87, 205]. The particular location of the region of interest is usually dictated by two main requirements: to obtain a pulsatile signal with a high SNR [99], or to examine the relative perfusion of tissue [86, 167, 202], in which case all areas of exposed skin are usually segmented into smaller sub-regions [167, 202, 205].

In all cases, signal quality remains extremely important, not only to determine the heart rate accurately, but also to determine the power of the signal [71, 99, 202], or its morphology [123]. From the raw RGB camera data, the green channel has been found to perform best ⁶[216], and this has been explained by the fact that oxyhaemoglobin is highly absorbant at the frequencies to which the green sensor is most responsive [157, 225]. This was in accordance with early research carried out by [39, 40, 222] in the last century and more recent work carried out using a spectrometer in the 380nm–980nm range that found that the SNR was greater than 1.5 between 480nm–610nm and between 800nm–925nm [37].

Nonetheless, the red and blue channels also contain some heart rate information. Although some authors only use the green channel for simplicity or speed [202], different methods have been devised to extract the maximal amount of heart rate information from a combination of the channels available, by attempting to find a weighting vector $\mathbf{w}$ that converts the multiple-channel data $\mathbf{C}$ into the optimal, single-channel “blood pulse volume” output signal $\mathbf{s}$.

⁶It should be noted here that such sweeping generalisations are often found in the literature. Presumably the authors assume that the reader is aware of the fact that both RGB camera sensors and lighting conditions are highly variable, and that what works in one setting has very little to do with another, as these are seldom specified.

$$\mathbf{s} = \mathbf{w}\mathbf{C} \quad (2.23)$$

Amongst these, independent component analysis (ICA) is the most commonly used. First proposed for this application by [157], it has found its way with minor variations in a number of subsequent publications [50, 51, 105, 123, 227], and a comparison of the performance of different ICA algorithms was investigated by [116]. Another blind source separation (BSS) technique was proposed by [105], who investigated the use of principal component analysis (PCA) as a method to choose between the components generated by the ICA algorithm, rather than selecting the signal with the highest signal-to-noise ratio in the frequency domain as had been done previously. A deterministic linear combination of the colour channels was instead proposed by [103, 108], in which the weighting coefficients were found for the data available. This approach was furthered by [45] whose chromaticity-based method was developed to reduce the effect of specular reflections. The same group subsequently noted that the characteristic signature of a person's pulse at rest in normalised RGB space could be used to determine yet another weighting for the colour channels [44]. Finally, alternative colour spaces were also investigated: [165, 216] noted an improvement in transforming the RGB channels into HSV and then either performing ICA on the three channels or simply taking the hue, while [19] identified the u^* channel in the standard $L^*u^*v^*$ space (a colour space of the luminance-chromaticity family) as a more appropriate one.

A further approach within the framework set out by Equation 2.23 instead aims to obtain a global improvement using multiple regions of interest and exploiting similarities in colour or pulsatility. Thus, [231] employed a spatial pruning method whereby they rejected regions of interest that had dissimilar colours or that had an angular movement in RGB space that was not smooth, whereas [99] used a weighted average of different regions of interest, with weightings proportional to the signal-to-noise ratio of each ROI. Finally, a complete departure from the linear-transform model was instead investigated by [234],

who investigated the use of Laplacian Eigenmaps in order to reduce the dimensionality of the problem.

In addition to the transformations that aim to maximise the cardiac information available, standard signal processing techniques have uniformly been applied to identify the cardiac frequency itself. Although the details of the implementations varied, these can generally be characterised by a two-stage process consisting of a pre-processing step, in which a filter is applied in order to minimise the signal of unwanted frequencies, followed by the transformation of the signal into the frequency domain in order to estimate the frequency of interest. For the pre-processing step, a band-pass filter defined by the physiological limits of the human heart rate (around 40-120 beats per minute for adults and around 90-180 beats per minute for neonates) has been typically employed to remove lower-frequency variations due to breathing or movement and higher frequency variations due to noise.

The estimation of the frequency of the heart-rate synchronous signal has been the object of several different spectral approaches, generally using short overlapping windows to account for the fact that the signal is non-stationary. A non-parametric approach was favoured by [44, 45, 157, 225], who employed Fourier-based techniques, and by [19, 71], who instead applied a wavelet-based analysis. Parametric spectral analysis was instead introduced to this field by [202, 227] through autoregressive models.

2.6 Oxygen Saturation with Cameras

As we have seen in Section 2.4.2, the development of pulse oximetry took advantage of the properties of the photoplethysmogram, isolating the arterial oxygen saturation by tracking the relative change in the amplitudes of the normalised cardiac-synchronous components at two or more different wavelengths. This thesis puts forward a method based on similar assumptions, building upon the recent developments mentioned above

in acquiring remote PPG signals in order to estimate changes in oxygen saturation using broadband RGB cameras. Prior to this work, however, a number of attempts at estimating SpO_2 were made, both using near infrared and visible light. This section aims to review the prior art, evaluating the strengths and weaknesses of the different approaches, by distinguishing between methods that used an approach informed by pulse-oximetry and those that instead attempted to regress the RGB values onto a modelled reflectance spectrum of the skin.

2.6.1 Regression-Based Methods

These methods attempt to find the concentration and distribution of different chromophores from an image and observe how these evolve over time. These methods are therefore not only concerned with the estimation of the oxygen saturation, but also with the total relative concentration of haemoglobin and melanin.

The first application of these approaches was however aesthetic rather than medical. At the turn of the millennium, [217] reported that if an RGB colour image of human skin was reduced to two components via PCA and then further separated by ICA, the resulting components had meaningful values, representing the strength (or concentration) of haemoglobin and melanin, that could then be used to make an image of a person's face look more tanned or more perfused. The medical value of the discovery was not lost on the author, who a year later published another paper which attempted to match the two components to the spectral absorbance of melanin and haemoglobin respectively [218]. The work was later used by [82, 83] to map chromophores (including oxyhaemoglobin) with a non-contact multispectral imaging device, but it was not until 2001 that RGB cameras were employed to estimate the concentration and distribution of oxyhaemoglobin [219]. This was accomplished by matching the transformed RGB values to a look-up table that would convert the three main components into melanin, oxyhaemoglobin, and deoxyhaemoglobin. A similar idea was used by [140, 141, 142], who first converted RGB

values into the CIE XYZ colour space in order for them to be device independent, and then estimated the relative concentrations of melanin, oxygenated haemoglobin and deoxygenated haemoglobin by multiple regression analysis onto models of skin reflectance created by a Monte Carlo simulation of light in tissue. A more simplistic approach was instead suggested by [80] who inferred the mixing matrix between concentration and RGB colour directly from physiological values for the absorption coefficients of each chromophore, an approach that was followed by [93], but this time using a 17-LED multispectral approach.

The limitations of this family of methods is inherent in its core formulation: they are only as accurate as the model from which they infer the parameters. Even then, there is an inherent inaccuracy associated with the fact that a set of continuous spectra will be ill-defined by only three points in colour. Finally, the work is only applicable in a lab setting as the model cannot even account for shadows. This notwithstanding, the greatest issue with these approaches is that they have never really left the field of aesthetics: despite multiple authors working over a cumulative period of over fifteen years, no concrete evidence that these methods work in practice has ever been produced. The majority of the tests carried out revolve around a purely qualitative metric – the behaviour of the chromophore concentrations before, during, and after the occlusion of a limb. When one of the tests was made semi-quantitative by comparing the RGB-estimated spectra to a reference spectrum, these were found to be utterly different in the red part of the spectrum and yielded vastly different results for almost all of the concentrations [140].

2.6.2 Pulse Oximetry Based Methods

Pulse oximetry-based methods are rooted in the idea that the same concepts can be applied to non-contact versions of a pulse oximeter (i.e. a sensor with a narrow-band filter) and hence extended to an RGB camera (i.e. broad-band filters). The former may

be seen as an extension of the pulse oximeter, and was first realised *in vivo* by [72, 73] by using LEDs emitting at 760nm and 880nm in conjunction with a CMOS sensor, and *in vitro* by [236]. However, the short distance required between the imaging apparatus and the finger from which a PPG signal was detected made this impractical as a non-contact method. The use of an RGB camera directly may instead be traced to [175] in the contact domain, who used the camera of a mobile phone together with its flash light to obtain a red and a blue PPG. As with pulse oximetry, the logarithm of the ratio of two amplitudes of the pulsatile cardiac-synchronous components of the two signals was then taken to be proportional to oxygen saturation.

A non-contact architecture using ambient light and narrow-band filters ($660\pm 5\text{nm}$ and $520\pm 5\text{nm}$) was proposed by [94], who, by using the entire face as a region of interest, gained enough of a signal-to-noise ratio to enable remote monitoring of oxygen saturation. The first contribution in this domain by the research group in which this thesis was undertaken was to demonstrate that this could also be done without a filter in a non-contact manner using a 3-CCD camera and the ratio of the red and blue channels [202]. This was later validated by further work on a separate cohort of patients [227], and separately by an author not attached to the group [7].

In addition to the experimental work described, a basic theoretical framework for pulse oximetry based methods has also been proposed by [36] using a Monte Carlo simulation of light passing through tissue. This work suggests that the following specifications are required in order for a camera with filters at 633nm and 940nm to detect a change in oxygen saturation of 1%: a minimum dynamic range of 77dB and a minimum frame rate of 40 frames per second.

As with regression-based methods, however, the main issue with the methods proposed so far is the lack of experimental data. While [72] reported a single value, and [73] worked on synthetic data, the results from [236] were taken from a phantom. Of the remaining publications, [7, 94, 175] each had their volunteers hold their breath for an extended

period of time, but in two out of three cases this only led to a maximum drop of 5% in SpO_2 , a change that is not sufficient to determine any relation between the estimated values and the SpO_2 itself ⁷. Even for the two publications from the research group within which this thesis was undertaken, which reported repeated, significant changes in oxygen saturation as a result of clinical conditions, the data was acquired for a period of tens minutes and only for single patients [202, 227]. In addition to this, the data analysis in the three publications that reported decreases of SpO_2 significant enough to constitute the basis for any kind of reasonable analysis (the contact-based publication of [175], and the group's two publications) was at best semi-quantitative, and the visual plots clearly showed the estimates diverging significantly from the reference values over time.

The shortcomings of the published research point to the need for a structured study that can demonstrate that oxygen saturation may be tracked continuously over an extended period of time using an RGB camera, whether by using narrow-band filters or simply using the available RGB sensors directly. The objective of this thesis is to contribute to the creation of this body of evidence and pinpoint the factors that will affect the measurements.

⁷Interestingly, [175] reported a decrease of over 20% in the reference SpO_2 for the single subject analysed, who was instructed to “hold [his] breath until [he] felt discomfort”!

Chapter 3

Experimental Methods

3.1 Overview

Two key studies were set up to provide the data necessary for the development of the algorithms described in this thesis. These studies, a Neonatal Trial and a Hypoxia Study, were very different in their objectives. The subjects in the Neonatal Trial were preterm infants and the objective of the trial was to establish whether the algorithms for heart rate and breathing rate that had previously been developed for adult humans could be adapted with equal success to the very different physiology of neonates. The Hypoxia Study only involved healthy adult volunteers and had the main objective of determining whether video cameras could be employed to determine changes in oxygen saturation.

With regards to oxygen saturation data, the two studies provided different, but complementary, datasets: in the Neonatal Trial, oxygen desaturations occurred as a result of an apnoea or a true change in the physiology of the infant. In the Hypoxia Study oxygen desaturations were induced, and were representative only of a change in the oxygen partial pressure in the environment. The former study yielded clinically-relevant data on neonates that could be used to derive algorithms specific to the neonatal physiology. The Hypoxia Study was instead conducted in carefully controlled conditions, and provides

the core data from which the basis of the use of video cameras to determine oxygen saturation can be investigated. However, as the root cause of the desaturations is not physiological, the results will need to be interpreted carefully.

3.2 Neonatal Trial

3.2.1 Standard Protocol

The Neonatal Trial received Medical Research Ethics Committee approval (REC reference number: 13/SC/0597) and was carried out at the John Radcliffe Hospital, where 30 pre-term infants were monitored in the High Dependency Care Unit (HDU) of the Neonatal Ward. Each infant was double monitored using a digital video camera and the standard reference monitor (Philips IntelliVue MX800 Patient Monitor).



Figure 3.1: Monitoring set-up used in the Neonatal study showing the video camera placed above the modified incubator.

The set-up used for the recordings is shown in Figure 3.1. The video camera is clearly visible above the incubator containing a mannequin of a baby. The set-up was purpose-built and comprised a modified Giraffe Omnibed (GE Healthcare, Little Chalfont UK) incubator with a 3cm diameter hole that was cut into the perspex canopy

(but otherwise identical to those in current use in the Neonatal HDU and ICU at the John Radcliffe Hospital), and a mobile unit with the video camera (attached to the unit via an extensible arm), two workstations (one for the video camera and one for the reference equipment), a monitor and a mouse and keyboard mounted onto a roll stand. The modifications to the canopy allowed the video camera to record video data without any interference due to reflections from the perspex glass itself. The effects of these modifications were evaluated over a period of two weeks to determine whether the introduction of the hole caused any changes to the humidity and temperature of the incubator environment [227]. All parts of the mobile unit were chosen so as to diminish the risks of infection and any parts which came into contact with the incubator were routinely sterilised. As a further precaution, a new transparent glass surface was attached to the camera lens every day.

The temperature and humidity are maintained constant within a neonatal incubator, however the light that was incident on the infant could not be controlled. Depending on the position of the incubator, the season, and the time of day during which the session was recorded, the light incident on the infant's skin would have been a mixture of sunlight and artificial fluorescent lighting (T12 80W tubes with starterless ballast, T8 58W tubes with electronic ballast, and T5 36W tubes with electronic ballast¹). To mitigate for these potential changes in lighting conditions throughout a recording session, a sterilised colour checker chart (BST13, Danes Picta, Prague, CZE) was placed next to the infant for the duration of the session, so as to have a multi-colour reference when environmental lighting changed. The colour checker was replaced every day to minimise the risk of infection.

The Philips reference monitor allowed the measurement of three vital signs (heart rate, breathing rate, and peripheral arterial oxygenation) by using an electrocardiogram (ECG) and an impedance pneumogram (IPG) recorded using the two ECG leads on the

¹The starterless ballasts have a frequency of 50Hz and the electronic ballasts have a nominal frequency upwards of 20kHz

chest. Finally, the pulse oximeter (OEM III Module – Masimo Corp, Irvine, US) was usually attached to the infant’s foot, and occasionally on the wrist in order to measure the photoplethysmogram (PPG) and derived peripheral oxygen saturation (SpO_2). The ECG, PPG and IPG waveforms were output at 500Hz, 125Hz, and 62.5Hz respectively. The variations in output frequency around the nominal sampling frequency are shown in Figure 3.2. The derived heart rate, breathing rate, and peripheral arterial oxygenation measures were nominally read at a frequency of 1Hz. The output frequency of these was found to have a median value of 0.9766Hz (mean: 0.9761Hz over 4 hours) (see Table 3.1)

Waveforms	Nominal Frequency (Hz)	Median Frequency (Hz)
ECG	500	499.9766
PPG	125	125.0016
ICG	62.5	62.4999

Derived	Nominal Frequency (Hz)	Median Frequency (Hz)
HR	1	0.9766
BR	1	0.9766
SpO_2	1	0.9766
PERF	1	0.9766

Table 3.1: Output frequencies from the Philips IntelliVue MX800 Patient Monitor

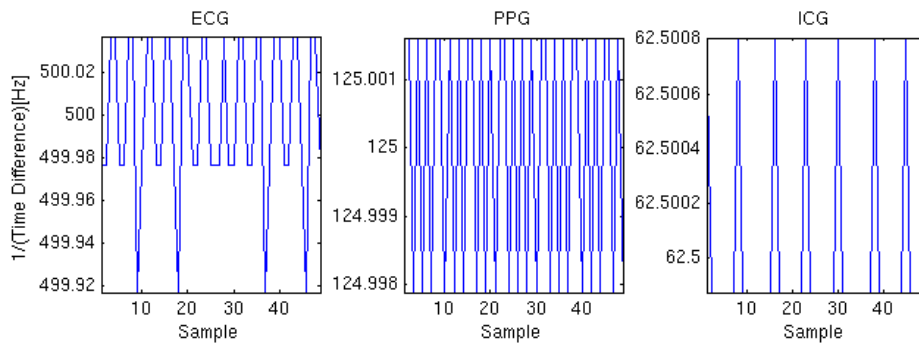


Figure 3.2: Variation in the output frequency of the ECG, PPG, and ICG waveforms.

The monitor also recorded a measure of perfusion (PERF) as estimated by the pulse oximeter and also nominally output at 1Hz. Intermittent, mean-arterial pressure (MAP) measurements were made by the nursing staff with a cuff, from which, diastolic and

systolic blood pressures were estimated. Additionally, non-numeric outputs from the patient monitor were also recorded, such as warnings, when the vital signs changed to alarming levels, occurrences of anomalies such as bradycardia, tachycardia, and apnoea whenever they were detected by the monitor. All data from the monitor was acquired using the ixTrendExpress 2.0 software (ixellence, GmbH) installed on a computer running Windows 7 and was exported to comma-separated-value (csv) files.

Other than the reference data acquired by the Philips IntelliVue monitor, notes were taken throughout each session by the clinical staff, recording both clinical procedures and clinically-relevant events. The notes were then converted into an electronic format and stored in the OpenClinica database, where they could be accessed as extensible markup language (xml) files. Finally, the sound level within the incubator was also recorded using a battery-powered microphone placed next to the infant and which sampled the noise at 1Hz, storing it onto an SD card.

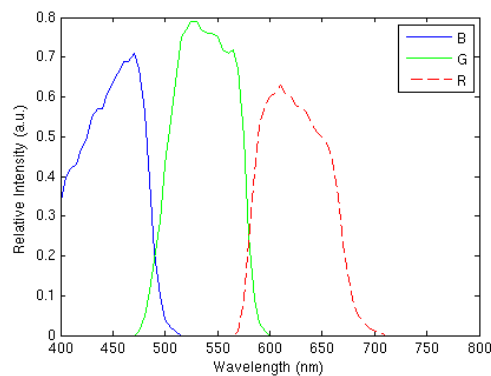


Figure 3.3: Sensor response for the blue, green, and red channels of the JAI video camera as a function of incident wavelength. Source: personal correspondence with JAI.

The video data was obtained using a 3-CCD (charged-couple device) progressive scan camera, the JAI AT-200CL. The video camera's 3-CCD configuration allows the red, green, and blue channels to be recorded almost independently of each other, with minimal overlap between the channels' response curves (see Figure 3.3). The lack of a Bayer-type filter to separate the three channels provides an improved spatial resolution. Optical

separation of the light is achieved by means of a prism onto three separate CCD arrays. Each CCD array in the JAI AT-200CL then employs Sony ICX-424AL sensors to detect the light intensity incident upon it. In the Neonatal Trial, the videos were acquired at 20 frames per second (fps), with 8-bit channel resolution and a spatial resolution of 1628x1236. The images were acquired, processed, and stored using a Xilinx Spartan Field Programmable Gate Array (FPGA) board, which was installed on a DELL PowerEdge T420, TPM machine using the Red Hat Fedora 20 operating system.

The video camera was linked to the FPGA via the Camera Link standard and the FPGA acquired the raw data from the image sensors, performed bit and channel swapping, added padding to bytes and provided direct access to memory on the hard disks used to store the images.

3.3 Hypoxia Trial

The Hypoxia (control) study was conducted in the Department of Physiology, Anatomy and Genetics of the University of Oxford, which has the facilities for, and extensive experience of, inducing oxygen desaturation for the purpose of research on low-oxygen (e.g. high altitude) conditions [69]. The study gained approval from the Medical Sciences Research Ethics Committee (CUREC number: MSD-IDREC-C3-2014-003) for the monitoring of up to 60 subjects divided into 6 equal groups of differing skin phenotypes as distinguished by the FitzPatrick Scale.

3.3.1 Study Design

The protocol for each session of the study was comprised of two separate sub-studies: the hypoxia sub-study, with the objective of tracking the oxygen saturation of the volunteer, over the course of which the oxygen saturation of the subject would be lowered to 80%; and the hypercapnia sub-study, with the objective of changing the breathing rate of

the volunteer by stimulating an increased respiratory response as a result of a gradual increase in the amount of the carbon dioxide inhaled by the volunteer. Each component of the session lasted approximately an hour, and were separated by a short break during which the subject was allowed to drink water.

Volunteer Consent and Confidentiality

The study gained University Ethics approval for the video recordings of the study to be used only within the research group and by authorised individuals from the University of Oxford. The videos collected are held in an anonymised and secure format. Volunteers were recruited through the distribution of approved advertising material in appropriate locations (gyms, colleges, etc.). Volunteers were screened by phone or email prior to the study date to ensure that they were eligible for the study. They were sent an approved Participant Information Sheet that contained all study details. Furthermore, volunteers had access to the study clinician for any questions prior to the study.

Eligibility and Exclusion Criteria

Volunteers were eligible for the study if they met the following criteria:

- Willing and able to give informed consent for participation in the study
- Male
- Aged between 18 and 40
- In good health

Volunteers were excluded from the study on the basis of the following criteria:

- Contraindication to exposure to hypoxia or hypercapnia, as judged by the research clinician, such as epilepsy or significant cardiac or respiratory disease
- Claustrophobia precluding spell in hypoxia chamber

Volunteers were free to withdraw at any time and were compensated for their time regardless of whether they withdrew or not. There were two levels of withdrawal of consent: no further contact and no further use.

3.3.2 Setup and Protocol



Figure 3.4: The custom-built study chamber in the Department of Physiology, Anatomy and Genetics, prior to the covering of internal surfaces with reflective material.

Chamber Set-up

Each session involved one subject who was placed in a purpose-built study chamber (Figure 3.4) in which the relative concentrations of oxygen, carbon dioxide, and nitrogen could be modified so as to induce hypoxia and hypercapnia [69]. All windows in the study chamber were covered so as to minimise the amount of stray incident lighting from any source other than that in the chamber itself. The interior of the chamber was furthermore lined with sheets of Tyvek 1443R to ensure a maximal and equal reflection of all wavelengths from the source (see Figure 3.5) so as to create an environment of diffuse and spectrally uniform lighting.

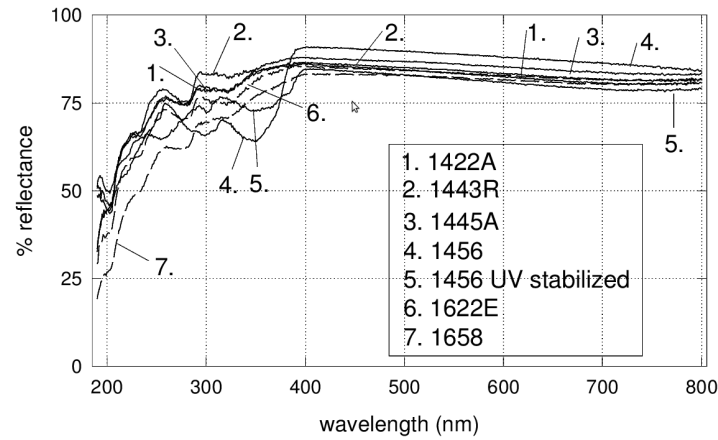


Figure 3.5: Reflectance properties of Tyvek papers, amongst which 1443R [194]. Tyvek materials reflect light uniformly thereby appearing white.

Lighting

The lighting selected to illuminate the subject was a 576-LED Mosaic 30cm lighting panel (Limelite – Bowens International Ltd., Colchester, UK). Two such LED panels were employed throughout the trial. The lights were directed at a reflective set-up that diffusely illuminated the subject's face. The output spectrum of the LEDs (measured prior to the start of the experiments with an Ocean Optics USB4000 spectrometer) is shown in figure 3.6, and was characterised by continuous spectrum with a sharp peak in the blue part of the spectrum (centred at 410nm) and a wider crest in the green and red part of the spectrum.

Hypoxia Sub-Study

During the hypoxia sub-study, the subjects were seated upon a reclining chair (see Figure 3.7, with their heads kept upright and at a constant angle to the video camera by a white V-shaped pillow placed around their shoulders. They were asked to face the video cameras and remain as motionless as possible while the nitrogen in the chamber was gradually increased so as to bring their peripheral oxygen saturation down to 80%. This

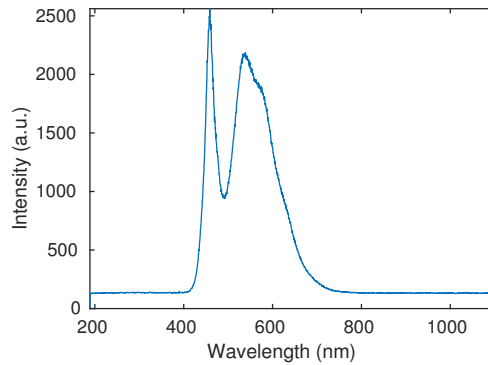


Figure 3.6: Spectrum of the light emitted by the 576-LED Mosaic Lighting Panel. The emitted spectrum was recorded using a Ocean Optics USB4000 spectrometer.

was done by lowering the oxygen saturation by approximately 5% at a time in steps of 7 minutes. The chamber carbon dioxide was not clamped during this procedure, but allowed to vary. Once the lowest oxygen saturation level (80%) had been reached and maintained for a period of up to 7 minutes, oxygen was delivered to the subjects via a nasal canula, usually at a rate of 1 litre per second, leading to a rapid resaturation. As soon as the oxygen saturation reached a plateau at around 95%, oxygen delivery was halted, in turn engendering a very fast decrease in oxygen saturation as the chamber atmosphere was still oxygen-depleted. Once the subject's oxygen saturation reached approximately 85%, the rapid resaturation-desaturation procedure was repeated. Finally, the subject was resaturated by delivering oxygen nasally at a slower rate (usually 0.5 litres per second).

Two sets of reference equipment were used throughout the procedure, the safety reference equipment that was used to monitor the condition of the subject, and the data-collection reference equipment that was instead collected as the standard reference. Of the safety equipment, the probe used to monitor the peripheral oxygen saturation was a GE TruSignal (TS-E4-H) ear-clip sensor or a Datex-Ohmeda OXY-F-UN finger probe if the signal on the former was too noisy. Both probes were connected to a Datex-Ohmeda 3900 Tru-Trak pulse oximeter via a Datex-Ohmeda OXY OL-3 cable. If at any point



Figure 3.7: Figure showing the interior of the chamber prior to the beginning of the Hypoxia sub-study. The subject, with the reference equipment already strapped on, is visible seated facing the cameras.

of the procedure the volunteer showed any sign of discomfort or dropped to an oxygen saturation level below 80%, the volunteer would either be fully or partially resaturated at the discretion of the clinician involved in the study.

Hypercapnia

The subjects were also seated in the reclining chair for the hypercapnia section of the protocol, however more movement was allowed. To this end, no prop was given to keep the head upright, allowing for more natural body positions. Once again, the volunteers were asked to remain as still as possible while facing the video cameras. After 3 minutes of time during which the carbon dioxide was kept at a normal level, the carbon dioxide in the chamber was increased by 12 Torr at a time in steps of 8 minutes. This was continued until the volunteer showed signs of discomfort or until the chamber carbon dioxide pressure was the same as the expired carbon dioxide pressure, a very potent breathing stimulus [188]. Once the maximal carbon dioxide pressure was reached, air would be allowed to enter the chamber. The hypercapnia section terminated 3 minutes after the beginning of the chamber air exchange.

3.3.3 Reference Equipment

The quantities measured by the data-collection reference equipment, the output sampling rates, and the devices used to measure each quantity are summarised in Table 3.2. In addition to the data from the equipment specifically devoted to recording the data-collection references, some additional data was extracted from the safety reference equipment specific to the chamber the volunteers were in, and this is reported as the chamber's equipment.

Two pulse oximeters were used in each session to record oxygen saturation and heart rate continuously, as well as a PPG waveform (a processed combination of the infrared and red channel PPGs), sampled at 32Hz and 256Hz respectively. Each pulse oximeter was attached to an index finger of the subject. The sensor used was the Masimo Corp M-LNCS DBI soft adult finger sensor, and each sensor was in turn connected to a Visi BlackShadow vital sign monitoring device (Stowood Scientific Ltd, Oxford, UK) via an M-LNC-1 Patient Cable (Masimo Corp, Irvine, US). The sensor is reported to have an averaging time of 2–4 seconds and an accuracy of $\pm 2\%$ in stationary adults in the SpO₂ range between 70% and 100%.

One of the two Visi Blackshadow devices employed was also used to record a 3-lead ECG, outputting a signal sampled at 256Hz; and to record the three signals used to estimate breathing rate. Two inductance pneumography (IPG) signals were derived from belts placed on and below the thorax, that measured chest and abdomen excursions respectively (XactTrace Thorax and Abdomen snap sensors attached to the XactTrace belt, Embla Systems, Kanata, US). A medium belt was used for all subjects. The third respiratory trace, that was only present in the hypercapnia part of the session, was the air pressure trace obtained via a disposable adult nasal pressure cannula (Embla Systems, Kanata, US) placed underneath the nostrils of the volunteer. These three signals were each recorded at a sampling rate of 64Hz.

As mentioned previously, the chamber was also equipped to monitor the volunteer

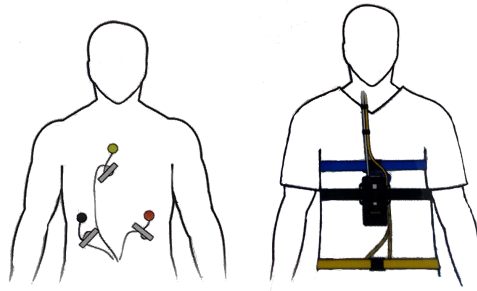


Figure 3.8: Positioning of ECG leads (left) and thoracic and abdominal respiratory rate elastic belts (right) for the main Stowood device used to record the reference data. (Source: Stowood Scientific Ltd, Oxford, UK).

and the chamber environment over the course of each study to ensure the safety of the participant. Three additional ECG snap-leads (ECG cable 0740-000, Kontron AG, Augsburg, DE), connected to a Micromon N-xt (Charter-Kontron Ltd, Milton Keynes, UK) ECG device were used to monitor the volunteer's heart rate, and a gas sampling line (73319, Datex Ohmeda - GE Healthcare, Little Chalfont, UK) that fed into a capnometer (Normocap Oxy, Datex-Engstrom - GE Healthcare, Little Chalfont, UK) was placed underneath the volunteer's nostrils, thereby giving both a breathing rate and inspired and expired (end-tidal) gas concentrations (PIO_2 , $PETO_2$, and $PICO_2$, $PETCO_2$ for oxygen and carbon dioxide inspired and end-tidal partial pressures respectively). To monitor oxygen saturation, the afore-mentioned GE earclip pulse oximeter was attached to the left ear-lobe. However, as this was predominantly a safety feature that was critical to the hypoxia part of the session, if the signal was not reliable enough to carry out the experiment safely, the aforementioned Datex Ohmeda fingerclip was used instead, attached to the right middle finger of the volunteer.

The monitoring data was also recorded using the chamber's bespoke software platform. The platform recorded inspired and end-tidal partial pressures for both oxygen and carbon dioxide, as well as the oxygen saturation from the chamber ("safety") pulse oximeter, on a per-breath basis rather than on at a constant sampling rate. As the breaths were

determined by the software itself and derived from the pressure trace originating from the nasal prongs that were also used for oxygen delivery, no data was available from the monitoring equipment during oxygen delivery in the hypoxia part of the session.

Measuring Device	Measured Quantity	Output Frequency
Chamber Software ²	SpO ₂	N/A
	PIO ₂	N/A
	PICO ₂	N/A
	PETO ₂	N/A
	PETCO ₂	N/A
Visi BlackShadow I	HR	32
	SpO ₂	32
	ECG	256
	PPG	256
	IPG - Thorax	64
	IPG - Abdomen	64
	Air pressure trace	64
	PTT (1)	5
	PTT (2)	5
	R-R	5
	Battery	32
Visi BlackShadow II	PPG	256
	HR	32
	SpO ₂	32
	Battery	32
Humidity/Temperature SD500 Dataloggers	Humidity	0.2
	Temperature	0.2

Table 3.2: Recording reference devices and output frequencies.

The humidity and temperature of the chamber were also recorded throughout the study, using two ExTech Instruments SD500 Humidity/Temperature Dataloggers (FLIR

²In addition to the quantities listed, the device also reported a running average for the CO₂ and O₂ end-tidal partial pressures, a running average for the oxygen saturation, and the mode value for the CO₂ and O₂ per breath.

Commercial Systems) that saved the data onto SD cards. These were placed close to the volunteer and close to one of the main workstations respectively.

3.3.4 Camera Set-up

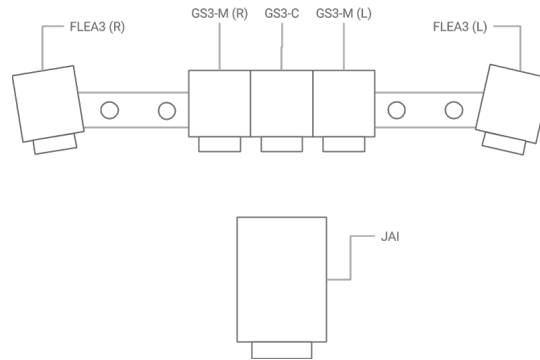


Figure 3.9: Positioning of the five cameras used during the hypoxia and hypercapnia sessions. The JAI camera was below the set of Point Grey Cameras in the vertical axis.

The recordings made in the Hypoxia and Hypercapnia sessions were made with six different cameras, so as to address multiple research questions at the same time. The set-up of the video cameras may be seen in Figure 3.9 and the features of each camera are listed in Table 3.3. A 3-CCD camera (AT-240CL – JAI, Yokohama, Japan) and a colour single-CCD Grasshopper3 camera (GS3-C – Point Grey Research Inc., Richmond, Canada) with a Bayer filter were used, and positioned centrally with respect to the volunteer. Two monochrome Grasshopper3 single-CCD cameras (GS3-M – Point Grey Research Inc., Richmond, Canada) were positioned on a purpose-built frame either side of the GS3-C, each with a narrow (10nm half-width bandpass filter): one centred at 520nm and the other centred at 660nm (Edmund Optics Ltd., York, UK). The choice of the filters was determined by three factors: maximising the quantum efficiency of the video cameras within the visible spectrum; maximising the difference between the absorption of the oxygenated and deoxygenated haemoglobin across the two filters (see Figure 2.12); and minimising the gradient change across the spectral bands. Finally,

two monochrome high-speed single-CCD cameras (FLEA3 – Point Grey Research Inc., Richmond, Canada) were positioned at the far ends of the frame, with the objective of using them as a stereo pair.

	JAI	GS3-C	GS3-M	FLEA3
Imaging Sensor	Sony ICX274AL	Sony ICX674	Sony ICX674	Semi VITA1300
Sensor Class	1/1.8"-CCD	2/3"-CCD	2/3"-CCD	1/3"-CMOS
Pixel Size (μm)	4.4 $\times$ 4.4	4.54 $\times$ 4.54	4.54 $\times$ 4.54	4.8 $\times$ 4.8
A/D Converter	14-bit	14-bit	14-bit	10-bit
Pixel Resolution	1628 $\times$ 1236	1920 $\times$ 1440	1920 $\times$ 1440	1280 $\times$ 1024
Max Frame Rate (Hz)	20	26	26	150
Video output	RGB	RGB	Mono	Mono
Lens-mount	C-Mount	C-Mount	C-Mount	C-Mount
Interface	CameraLink	USB 3.0	USB 3.0	USB 3.0

Table 3.3: Features and specifications of cameras employed in the Hypoxia Study. (Sources: JAI, PtGrey)

The video cameras were connected to three separate workstations to meet the storage requirements of the study: the JAI camera was connected to a PowerEdge T420 server (Dell Inc., Round Rock, US) via CameraLink; the two GS3-M and the GS3-C cameras were connected to a Precision T7610 (Dell Inc., Round Rock, US) workstation via USB3; and the two FLEA3 cameras were connected to another Precision T7610 workstation, also via USB3. On each workstation, a purpose-built software developed by the research group (Pebase), automatically set and saved the recording parameters via the video cameras' APIs. As a result of the large storage requirements, the videos from each session were copied onto external encrypted hard disks configured in RAID-1 format and thence transferred to the Institute of Biomedical Engineering's research storage cluster, where each video was compressed via lossless ZIP-based compression.

The settings at which each camera was run are reported in Table 3.4. These were dependent on a trade-off between the maximal resolution, bit-depth, and speed (frames per second) of each camera, the storage requirements, and the speed of the recording software. Minimising frame loss during the recording and saving process was determined

	JAI	GS3-C	GS3-M	FLEA3
Frame Size	1628 × 1236	1800 × 1200	1800 × 1200	1200 × 800
Bit Depth	12	16	16	16
Frame Rate	16	26	26	60
Exposure Time	N/A	OFF	OFF	OFF
Shutter Time(ms)	N/A	20	20	15

Table 3.4: Settings for cameras employed in the Hypoxia Study

to be the most important optimisation factor and several different configurations were attempted for each camera to ensure that each configuration would not lead to any lost frames while maximising all other parameters.

	JAI	GS3-C	GS3-M	FLEA3
Lens name	HF12.5SA-1	HF25HA-1B	SV-2514H	SV-1614H
Lens manufacturer	Fujinon	Fujinon	VS Technology	VS Technology
Focal length (mm)	12.5	25	25	16
Iris range	F1.4:F22	F1.2:F22	F1.4:16	F1.4:16

Table 3.5: Lenses used during hypoxia.

	JAI	GS3-C	GS3-M	FLEA3
Lens name	GMHR3D26018C	HF25HA-1B	SV-2514H	SV-0614H
Lens manufacturer	Goyo Optical	Fujinon	VS Technology	VS Technology
Focal length (mm)	6	25	25	6
Iris range	F1.8:F16	F1.2:F22	F1.4:16	F1.4:16

Table 3.6: Lenses used during hypercapnia

The lenses used in the hypoxia sessions and the hypercapnia sessions are detailed in Tables 3.5 and 3.6 respectively. Different lenses were chosen for every camera except the GS3-M and the GS3-C so as to ensure that during hypoxia the face would occupy the majority of the frame whereas the whole of the upper body would be visible during hypercapnia. This was done so as to maximise the amount of skin visible in the hypoxia session, while ensuring that movement from the abdomen and chest was visible during the hypercapnia session.

3.3.5 Equipment Synchronisation

All the equipment used in the study was synchronised with respect to the study laptop's time. The laptop used was a Latitude E5540 (Dell Inc., Round Rock, US) running the Windows 7 operating system. It was connected to the three recording workstations in a local area network (LAN) through an SD2008 8-port network switch (Linksys International Inc.). Immediately prior to running the video cameras, each workstation time would synchronise with the laptop time, and the time-difference between each machine and the study laptop was also recorded at the end of each session. The temperature and humidity dataloggers were synchronised to the laptop time manually at the beginning of the day, while the Visi Blackshadow devices were synchronised to the laptop time at the beginning of each hypoxia and hypercapnia session through VisiDownload, the instrument's bespoke bluetooth set-up software (Stowood Scientific Ltd, Oxford, UK).

3.3.6 Population Statistics

Skin Type Distribution

Of the 46 volunteers recruited for the Hypoxia Study, Table 3.7 shows that more than half belonged to three Fitzpatrick skin types only (types II-IV), each composed of 9 subjects, while only 5 Fitzpatrick Type VI subjects were recruited. Imbalances in skin groups will need to be taken into account when investigating the effect of skin colour on the ability to track oxygen saturation accurately.

Table 3.7: Distribution of Fitzpatrick skin Types

Fitzpatrick Type	I	II	III	IV	V	VI
Number of Subjects	7	9	9	9	7	5

Heart Rate and SpO₂ Distribution

The distribution of heart rates across the Hypoxia sub-study sessions was calculated using the two data-collection reference pulse oximeters, each placed on an index finger of the subjects. As may be seen in Figure 3.10, the distributions are in an overall agreement and show values typical of an adult population, with a quasi-normal distribution that spans heart rates from around 45 to 110 beats per minute (bpm). Although the volunteers recruited were generally young adults and were known to not have any medical conditions, the heart rates are perhaps a little higher than what would be expected of young, healthy adults at rest, and this could be as a result of two factors: the excitatory influence of low oxygen conditions on the heart rate, and some degree of stress as a result of undergoing the study.

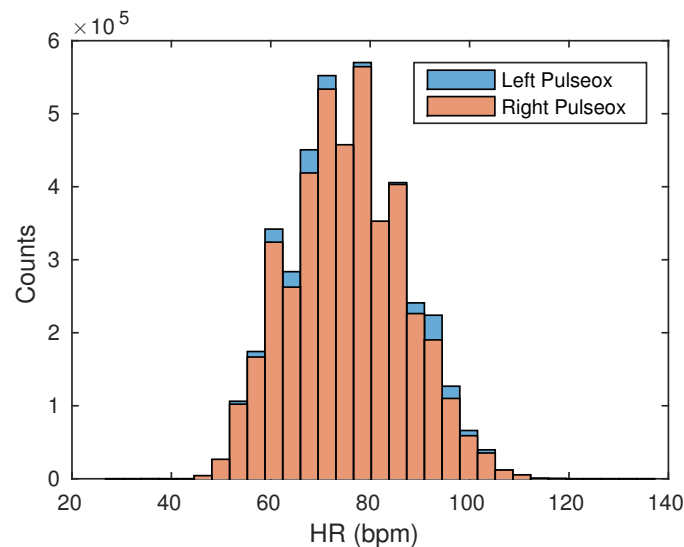


Figure 3.10: HR distribution as recorded by the two pulse oximeters, placed on the left and right index fingers of the subjects.

The data-collection reference pulse oximeters were similarly used to review the distribution of oxygen saturation values across all subjects recruited for the hypoxia sub-study. As a result of the format of the study, whereby the subjects spend 7 minutes on air prior to spending 7 minutes at 95%, more data has been collected for in the 92.5%-100% interval than any other equally sized interval. Similarly the distribution

shows that the timing of the stepped oxygen descent were fairly consistent, as the 78.5%-92.5% interval is approximately flat.

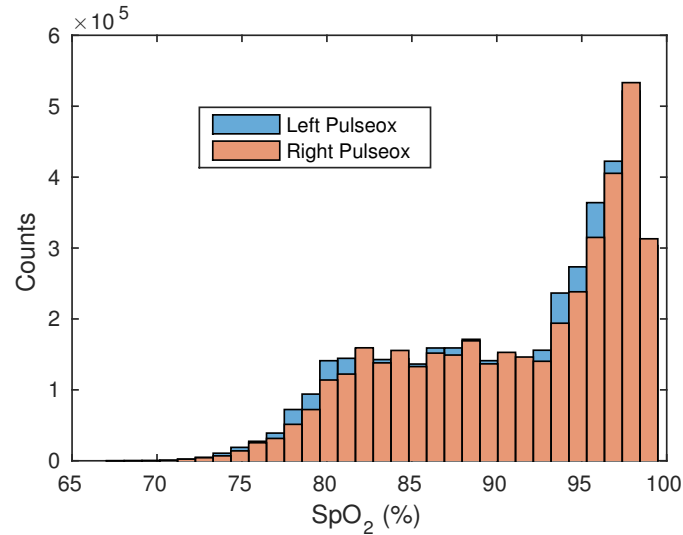


Figure 3.11: SpO₂ distribution as recorded by the two pulse oximeters, placed on the left and right index fingers of the subjects.

Respiratory Rate Distribution

The distribution of respiratory rates across the Hypoxia sessions was calculated using three separate sources of data: the airflow, the thorax impedance pneumograph, and the abdominal impedance pneumograph. The signals were processed by identifying the frequency corresponding to the peak spectral power when using a short-time Fourier transform (STFT) with a running 20 second window shifted by 10 seconds. As a result, the distribution of respiratory rates is binned into discrete bins separated by 1.82bpm. Unlike the HR histogram, which was more smoothly distributed, a large majority of BR values are all within the 16.8-18.8bpm range, although the breathing rate spans rates from 6.6bpm to 30bpm.

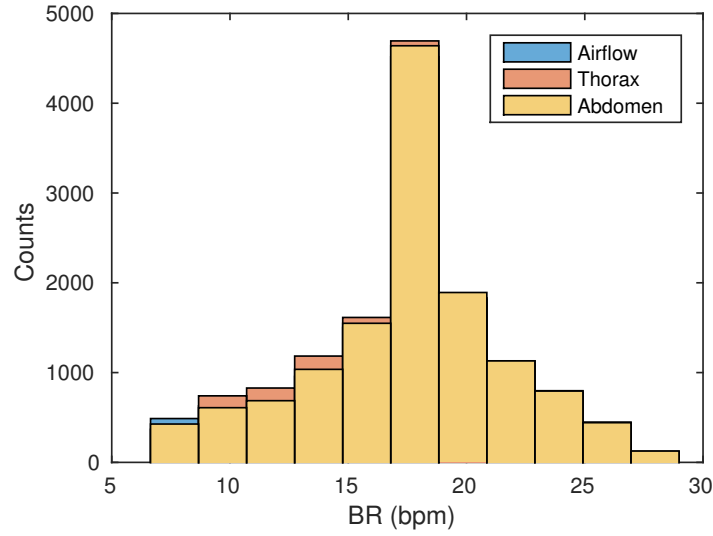


Figure 3.12: BR distribution as recorded by the two chest belts and the nasal cannula, respectively placed on the thorax, abdomen, and under the nostrils of the volunteer.

3.4 Summary

Video data was acquired in two different settings: clinical conditions (NICU trial) and controlled conditions (Hypoxia trial). In the former, pre-term infants were monitored using a single RGB camera (8-bit JAI) under variable visible illumination, while in the latter young, male, healthy volunteers were filmed while undergoing a procedure to make them hypoxic (Hypoxia sub-study) and then hypercapnic (Hypercapnia sub-study). Recordings were made with five different cameras to study various aspects of non-contact vital sign monitoring with cameras. Throughout the course of this thesis, we shall consider only the data acquired using the 12-bit JAI camera used in the Hypoxia sub-study, although the NICU data will be considered in Chapter 8.

Chapter 4

Theoretical Framework and Data Preprocessing

4.1 Introduction

The objective of this chapter is to establish a basic theoretical framework to be used in the development of methods to monitor oxygen saturation changes using changes in colour as measured by a digital video camera positioned some distance from the subject. This is set out and discussed in Section 4.2. In Section 4.3, the implications of the assumptions that are inherent to the model are discussed within the context of the Hypoxia study, and applied to the image processing methods developed. These are then summarised in Section 4.3.4 in order to construct a coherent picture of the image-analysis pipeline employed. The output of this pipeline will be used in future chapters to establish the relationship between oxygen saturation and colour itself.

4.2 Theoretical Model

In order to understand how the relationship between the colour measured by the video camera and oxygen saturation may be found, a general expression for the origin of the

time-varying colour signal will be derived, firstly for the simplest case of a flat surface imaged under ambient lighting, that is illumination of equal intensity incident on the surface from all directions equally and scattered in all direction equally [34].

Considering a single sensor within a sensor array, positioned at x and characterised by its colour-dependence c , the response of the sensor ρ_c^x may be described as the integral over the electromagnetic spectrum of the product of the light incident on the sensor and the sensor response $R_c(\lambda)$ [115]. In the case of uniform ambient lighting, we may further express the light incident on the sensor as the geometrical fraction ξ of the product of light incident on the surface $L(\lambda)$ and the surface reflectance function of the area $d\vec{A}$ imaged by the sensor at x , $S^x(\lambda)$, such that:

$$\rho_c^x = \xi \int L(\lambda) S^x(\lambda) R_c(\lambda) d\lambda \quad (4.1)$$

For light incident on an area $d\vec{A}$ of skin, we may separate the surface reflectance function into two components: a function $B^x(\lambda)$ describing the interaction of the light with blood and a function expressing the interaction of light with the remaining tissue $M^x(\lambda)$

$$\rho_c^x = \xi \int L(\lambda) M^x(\lambda) B^x(\lambda) R_c(\lambda) d\lambda \quad (4.2)$$

While $M^x(\lambda)$ will be influenced by the structure of the skin and the concentration of scatterers and absorbers such as melanin and collagen, $B^x(\lambda)$ will be dependent on the volume of blood present at any given time and the spectral characteristics of blood. In line with the Beer-Lambert law introduced in Chapter 2, the two functions are combined multiplicatively. Furthermore, the pulsatile nature of blood flow may be incorporated into this model by considering that the light will interact with a static, or baseline, body of blood that is present at all times and whose volume will depend on the overall perfusion of the skin, and a time-varying (AC) component that is due to the change

in blood pressure between systole and diastole. The reflectance functions of these two components must however be treated differently in terms of wavelength-dependence as we assume that the time-varying component is due to changes in the volume of arterial blood whereas the baseline component depends on mixed venous and arterial blood [4, 87]. These considerations yield the expression:

$$\rho_c^x = \xi \int L(\lambda) M^x(\lambda) \left(v_b^x(t) \hat{B}_b^x(\lambda) + v_{AC}^x(t) \hat{B}_{AC}^x(\lambda) \right) R_c(\lambda) d\lambda \quad (4.3)$$

where the factors $v_b^x(t)$ and $v_{AC}^x(t)$ describe the volume dependence of the baseline and pulsatile blood components and the normalised functions $\hat{B}_b^x(\lambda)$ and $\hat{B}_{AC}^x(\lambda)$ describe the dependence on chromophores such as oxygen saturation. If all other chromophores are stable over time, we may then decompose ρ_c^x into two oxygen-saturation dependent components, $\rho_{c,b}^x$ and $\rho_{c,AC}^x$

$$\rho_c^x = \rho_{c,b}^x + \rho_{c,AC}^x \quad (4.4)$$

with

$$\begin{aligned} \rho_{c,b}^x &= \xi v_b^x(t) \int L(\lambda) M^x(\lambda) \hat{B}_b^x(SO_2) R_c(\lambda) d\lambda \\ \rho_{c,AC}^x &= \xi v_{AC}^x(t) \int L(\lambda) M^x(\lambda) \hat{B}_{AC}^x(SO_2) R_c(\lambda) d\lambda \end{aligned} \quad (4.5)$$

The time-dependence of the cardiac-synchronous pulsatile component $v_{AC}^x(t)$ is characterised by the heart rate frequency f_{HR} and the perfusion of the tissue. The time dependence of the baseline component instead only depends on the perfusion $v_b^x(t)$. However $\rho_{c,b}^x$ and $\rho_{c,AC}^x$ both also depend on oxygen saturation SO_2 . Through the use of a suitable low-pass filter, the AC variations of the signal may be eliminated and the two components separated.

4.3 Model Assumptions and Implications on Analysis

The simplistic treatment and assumptions behind the model described by Equation 4.3 limit its potential for application in the real-world. The assumptions implicit in the model derived in the previous section will firstly be reviewed and translated into requirements for data analysis, and the relationships found will then be tested on the data acquired in the Hypoxia study.

4.3.1 Lighting and Sensor

The first assumption in the model, one that is already present in Equation 4.1, is that of the time-independence of the illumination term $L(\lambda)$ and of the sensor $R_c(\lambda)$. Inherent in the model is also the assumption of ambient lighting. To a first degree of approximation, these assumptions are valid in the Hypoxia study since the light source is limited to a single type of LED, with the light reflecting off surfaces covered in the same, spectrally flat Tyvek material, and because only one camera (JAI) is used for each analysis. Throughout this thesis, the results that will be reported are those for the JAI camera. This camera was chosen over others for oxygen saturation investigations as the properties of the camera provide less overlap and cross-talk between different colour channels, and a higher pixel resolution. There is also the possibility of extending the work of the Hypoxia study to the NICU dataset which was recorded with the same camera. The spectra of the light as sampled directly using a spectrometer and reflected off the Tyvek material are shown in Figure 4.1 and indicate that the material does not affect the shape of the spectrum.

Although no interference was expected as LED lighting was used, the videos were examined to determine whether there was any lighting-dependent interference due to lights external to the room or the LEDs. The frequency of the interference would be visible as an aliased component on a periodogram of the light reflected off a motionless surface in the background, particularly if taken over a long time period. Figure 4.2 shows

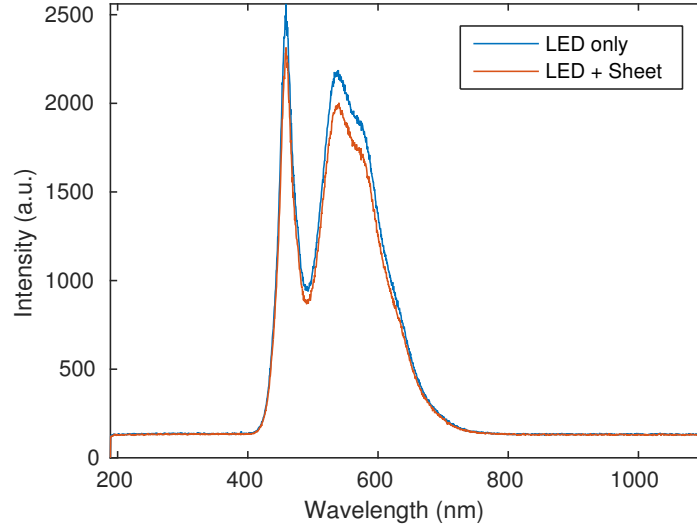


Figure 4.1: Figure showing spectrum of the LED panel as sampled directly (blue) and reflected off Tyvek (red).

both the timeseries and the spectral components of the mean total intensity (R+G+B) resulting from a 12×12 background region. This allows the hypothesis of lighting flicker to be tested over a 40-second window. The result shown indicates that this effect is either not present or not significant, as the g-statistic ¹ for the maximal peak has a p-value of 0.98 and the frequency at which it occurs does not yield an integer N when substituted into:

$$f_{alias} = |f - Nf_s| \quad (4.6)$$

where f_{alias} is the expected aliased frequency, f is the true flickering frequency (expected at 50Hz or 100Hz, the fundamental or second harmonic of mains frequency in the UK), f_s is the 16Hz sampling frequency of the camera sensor and N is an unknown integer.

The sensor noise itself will have several different sources, the most important of which are: time-dependent drift noise, that may be expected as a result of changes in

¹Fisher's g-statistic is the ratio of the maximal value of the periodogram to the sum of all values within the periodogram [53], with the p-value calculated using [235]

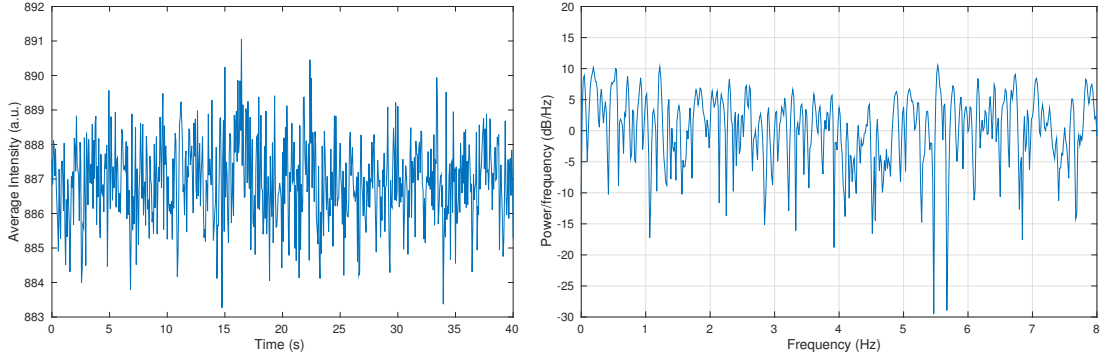


Figure 4.2: Figure showing the grey-scale timeseries of light reflected off the background (left), and its spectrum (right).

the temperature of the CCD; spatial noise that may be due to slight differences in the manufacturing of each pixel; shot noise that is instead due to the random nature of the arrival of photons, and is hence governed by a Poisson distribution; and finally the read-out noise as the charge is converted into voltage changes and then digitised, which is expected to be dependent on the sampling frequency. Throughout the study, the 3-CCD JAI camera acquired data at the same sampling frequency; however a time-dependence may be introduced as the camera heats up during operation. The importance of this effect is shown in the figure below, showing the frame variance over the course of just over 100 minutes, during which time the video camera was kept running while covered and in the darkened room.

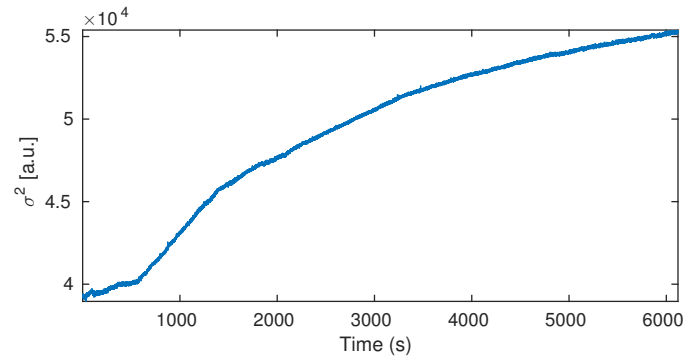


Figure 4.3: Figure showing the change in noise, measured as the variance of the frame values, for a blacked-out camera kept in the Hypoxia chamber.

Similarly, the bias of the read-out was also examined in order to establish whether the noise was truly Gaussian, and was calculated for one in every 16 frames. For each frame evaluated, the average of the frames acquired one minute before and one minute after was subtracted so as to remove the pixel bias, and the Gaussianity of the frame was then established by applying the one-sample Kolmogorov-Smirnov test² and plotting the p-values of the test against time. These were found to not vary over time and to be of the lowest value computable, thereby estimating the distribution of pixels as Gaussian. The distribution for one of the frames is shown in Figure 4.4.

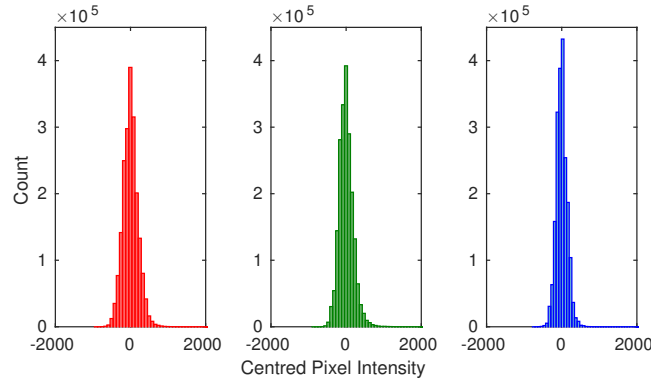


Figure 4.4: Figure showing the histograms for the read-out noise of last camera frame (at 6100s in Figure 4.3), one for each channel. From left to right: R,G,B.

As the shot noise and the read-out noise can be modelled as white noise, this noise can be reduced within frames by taking a spatial average over several pixels rather than simply considering each pixel individually. For this reason and also to decrease the computational cost of further operations, the first step in the analysis was to reduce the videos into 12×12 non-overlapping pixel regions, yielding an improvement of a factor of 12 in the signal-to-noise ratio (SNR) in the ideal-case scenario. As a result of this initial step, a 12×12 region of interest (ROI) is the smallest possible for all subsequent analysis. In applying this step, the sums of the 12×12 regions were saved to disk rather

²The one-sample Kolmogorov-Smirnov test is a non-parametric hypothesis test used to determine whether sample data comes from a population with a normal distribution. The Kolmogorov-Smirnov statistic measures the distance between the empirical cumulative distribution function and the expected cumulative distribution function [119].

than the averages so as to ensure that if larger regions of interest were constructed from sums of the 12×12 regions, roundoff errors would not be accumulated.

4.3.2 Skin Detection

In Equation 4.2, the model expands the surface reflectance function under the assumption that the recorded video camera data is a measurement of the reflectance from a region of skin, a necessary pre-condition in order to obtain oxygen saturation. To ensure that the pixels considered in the analysis are indeed from the skin, two features of the skin under investigation may be employed: the heart rate dependence of the skin colour and the colour of the skin itself. The former is strongly affected by the signal-to-noise ratio of the cardiac-synchronous component in the video camera signal which depends both on the size of the ROI and the concentration of melanin present. As a result, its usefulness as a criterion is limited by having to use more pixels (and therefore potentially a subset of non-skin pixels) in order to estimate it and its dependence on skin colour. Furthermore, non-skin pixels may also exhibit heart-rate synchronous pulsations due to the movement of the body and head and connected regions [8].

A large body of literature exists around the field of skin detection and segmentation using colour; however as the detection of skin is not the primary objective of the thesis, an approach that makes use of known priors is used to facilitate its implementation for the specific case of the data acquired in the Hypoxia study.

Face Detection

Firstly, all participants faced the camera and this means that a simple face detection algorithm may be employed to ensure that fewer false-positives are included in the skin detection method. The Viola-Jones method of object detection [106, 228] is a fast and well established approach to face detection provided that the face is not occluded and has a frontal view that is not at an angle. As the Hypoxia dataset is constrained in

a way that meets these requirements, a specialised training dataset was not necessary in order to perform the face detection and a state-of-the-art frontal face detector was chosen instead³, the details of which may be found in Appendix A.1.

To approximately match the size of the faces used to train the Matlab Viola-Jones implementation, the images acquired by the JAI video camera were resized by a factor of 0.1. The number of candidate faces output by the Viola-Jones algorithm was also reduced to one by selecting the largest candidate in the frame. The results of the face detection algorithm are shown in Figure 4.5 for six of the participants, partitioned according to skin type. The performance of the algorithm was also tested on the annotated CMU+MIT testing dataset of frontal images [197] in order to evaluate the Matlab implementation on a public database that would enable its comparison with the original algorithm devised [228] and then tested [229] by Viola and Jones. On the CMU+MIT dataset, the Matlab implementation used throughout this thesis had a sensitivity of 0.77, a specificity of 0.92, and an accuracy of 0.84 with a total of 9 false positives - a performance surpassing that reported in [229].

Skin Segmentation

Once the bounding box of the region segmented by the face detector has been obtained, it may be used to create a model specific to the colour of the participant's skin to enhance the accuracy of skin detection (Figure 4.6). This is the inverse of the usual technique of employing skin detection to aid in face recognition, and is based on the realisation that skin pixels are highly separable from non-skin pixels in a number of colour spaces [84, 226]. In particular, the rg chromaticity space (where $r = \frac{R}{R+G+B}$ and $g = \frac{G}{R+G+B}$, form a chromaticity space as they are independent of light intensity, as per Section 2.5.2) was used as this has been used widely and is known to separate skin independently of illumination [241]. Noting that the inner part of the bounding box selected by the face

³CascadeObjectDetector of the Matlab Computer Vision System Toolbox, Version 6.1 (R2014b)



Figure 4.5: Figure showing face detection using the modified Viola-Jones algorithm for the six different skin phenotypes. Detected face bounding box shown in red.

detector will contain a majority of skin pixels but will also contain some non-skin pixels corresponding to hair, eye, and background regions, the contents of the bounding box may be classified as being skin or non-skin (step E.1 in Figure 4.6). A probability distribution function (pdf) can be constructed for both by taking the pixel values within several facial bounding boxes for selected faces. The modelling of the distributions by a Gaussian pdf, rather than the direct use of the pixel values such as by using histogram-based comparisons, allows a generalisation for pixel values that were not present in the images [226].

A bimodal multivariate Gaussian mixture model in rg -space is fitted to the data acquired from within the facial bounding box and across different frames (step F.1 in Figure 4.6), allowing the main skin distribution to be separated from the background distribution. The mixture model is initialised by estimating the peaks in the histograms of r and g respectively and selecting the two peaks that are furthest apart as the mean locations of the two distributions. In order to diminish the noise in estimating the peaks, the histogram is first smoothed with a five-point median filter. The covariance matrix

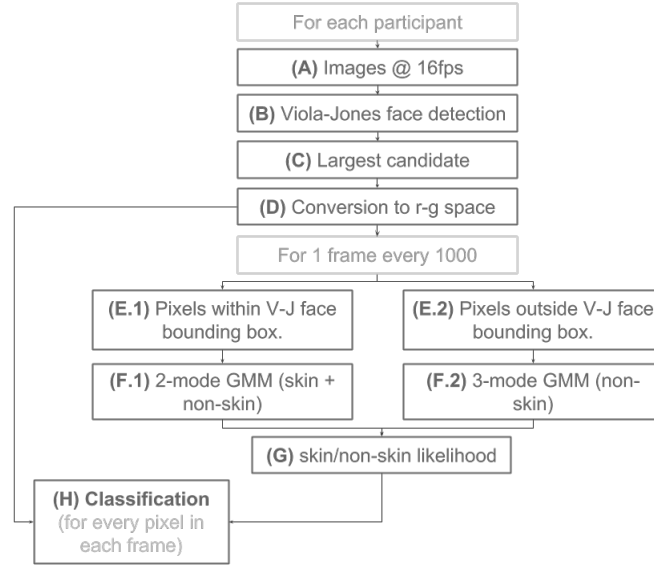


Figure 4.6: Schematic diagram showing the skin segmentation algorithm steps.

is initialised as a diagonal matrix with the terms corresponding to the assumed skin distribution as having two-thirds of the variance of the r and g distributions and the terms corresponding to the non-skin distribution as having one-third of the variance of the r and g distributions, reflecting a prior belief about the preponderance of skin pixels.

The Gaussian mixture model likelihood is then optimised via the iterative Expectation - Maximisation (EM) algorithm. Finally, the Gaussian in the mixture model describing the largest number of pixels is taken to be the likelihood function for the skin class $p(\mathbf{c}|\text{skin})$ and is described by:

$$p(\mathbf{c}|\text{skin}) = \frac{1}{(2\pi)^{\frac{1}{2}}|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{c}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{c}-\boldsymbol{\mu})} \quad (4.7)$$

where $\mathbf{c}$ is the vector for the position of the colour in rg -space, $\boldsymbol{\mu}$ is the mean estimate of the skin chromaticity across the dataset, and Σ is the diagonal covariance matrix estimated over the same dataset.

Given that the purpose of the model is solely to express the colour of the skin of a single individual across the entire session accurately rather than construct a model

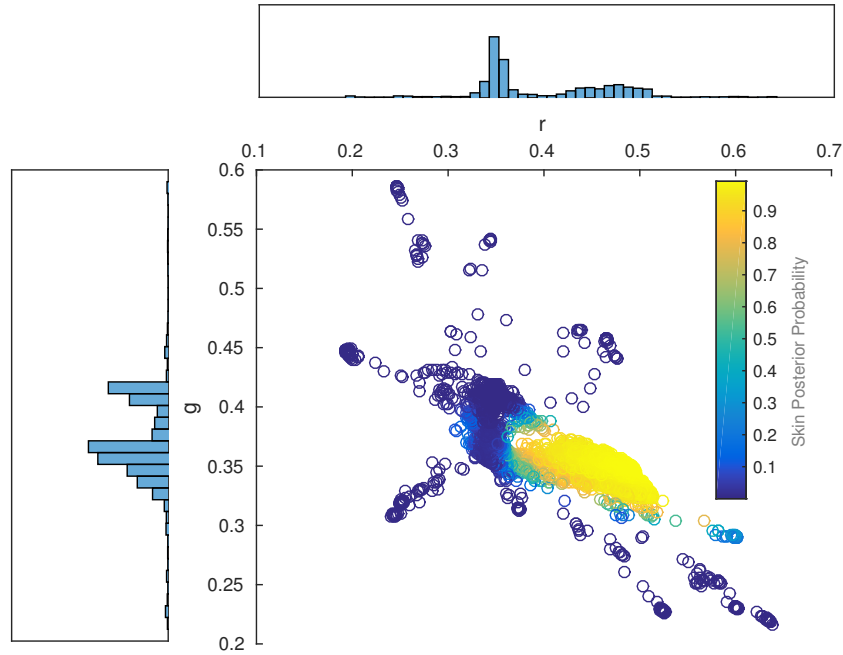
for multiple types of skin, a single Gaussian is sufficient [84]. The background region is modelled using the remaining Gaussian found within the facial bounding box and a further trimodal Gaussian Mixture Model is fitted to all pixels outside the facial bounding box (steps E.2 and F.2 in Figure 4.6), yielding the likelihood for the non-skin class $p(\mathbf{c}|\neg\text{skin})$. The trimodal Gaussian mixture model is initialised using the k-means++ algorithm [5] with a k-value of 3. When the two mixture models are combined (step G), a naive Bayesian classifier can be built [81] such that any new pixel of chromaticity c will be classified as skin if:

$$p(\mathbf{c}|\text{skin})p(\text{skin}) \geq p(\mathbf{c}|\neg\text{skin})p(\neg\text{skin}) \quad (4.8)$$

where the priors for skin and non-skin ($\neg\text{skin}$) are taken from the frequency of occurrence of each (step H in Figure 4.6). The results of the skin classification step may be seen in Figure 4.7, which gives an example of the distribution of points in rg -space for one frame of a participant, and the classification of the points based on the posterior distribution found using the dataset for that participant (built using one frame in every thousand for that session). As may be seen, the misclassification error is still relatively high, in that many non-skin features in the background are still classified as skin.

In order to eliminate unwanted background pixels, the index positions of all pixels that are classified as skin – known as the skin mask, is first morphologically dilated⁴ in order to fill in any non-skin areas which may split the face region into multiple segments. The resulting mask is then labelled into separate regions, under the assumption that the face will constitute the largest region compared to any background labelled as skin, and all regions except the largest are removed. Finally, the resulting mask is multiplied with the original skin mask in order to remove the effects of the dilation.

⁴Morphological dilation [220] is a rule for binary images whereby the immediate neighbours of any pixel that is classified as foreground also become foreground, enlarging the boundaries of foreground pixels.



(a)



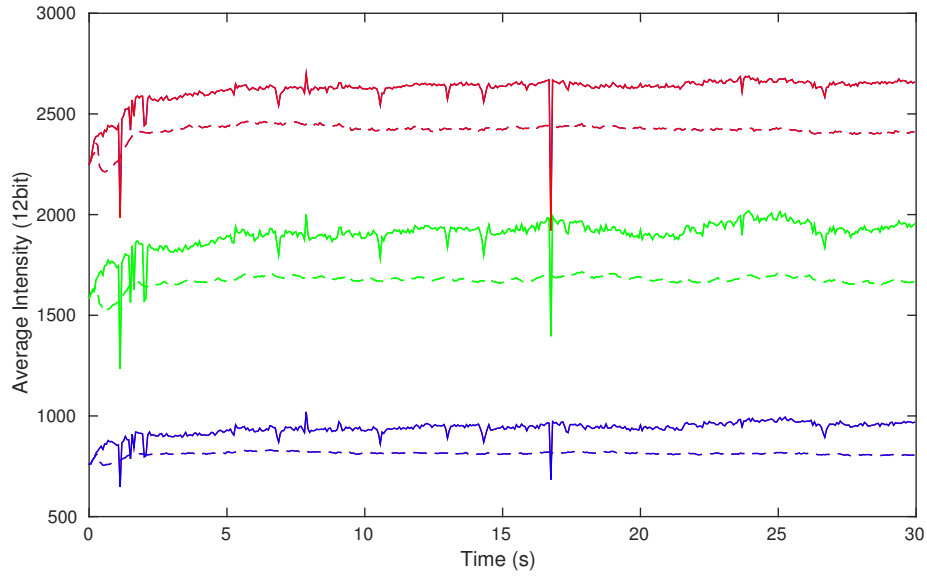
(b)

Figure 4.7: The posterior probability for belonging to the skin class is shown in 4.7a, with histograms of the normalised red and normalised green distributions on the sides. The resulting segmentation for that frame is instead shown in 4.7b. From left to right: the original image, the posterior probability, the segmentation via the naive Bayesian criterion overlaid on the original image. (Image reproduced with consent of the subject).

4.3.3 Motion

A key aspect of real-life monitoring that is not incorporated by the model outlined in the theory is motion. Motion will affect the video camera signals in two main ways: if the location of the region of interest is kept constant and the subject moves, a different part of the skin will be visualised, with a different angle to the camera and illuminating light, and possibly a different average path for the light passing through the tissue; on the other hand, if the region of interest moves with the subject, the same effects will occur as the angle between the camera and the illuminating light also changes. The top graph in Figure 4.8 demonstrates this. The continuous line shows the mean red, green, and blue (RGB) values for a single 12×12 pixel region of interest as tracked using the relative distance from eye corner feature points during and after motion, whereas the dashed line shows the mean RGB values for the original fixed ROI. The two regions of interest coincide for the first frame but then differ as the participant moves, and as a result their intensities also differ. The features that were used to track the motion (the corners of the eyes) and shift the region of interest accordingly were selected manually, and do not deform independently of the subject's rotation. Some degree of instability is seen in the spikes that occur (at around 17 seconds) even when the subject is relatively stable. These do not occur for the fixed region and are simply due to the relative sensitivity of the feature points selected to small movements.

Motion may be tracked simply by determining the position of the face within the frame using the Viola-Jones algorithm. Once the bounding box of the face has been obtained, the Cartesian co-ordinates of the centroid of the bounding box can be determined, and the Euclidean distance between the centroid and any fixed point will represent the distance travelled by the centroid and hence motion. Figure 4.9 shows the Euclidean distance between the location of the centroid and its original location over the course of two sessions, evaluated once every 16 frames (once a second) for the duration of the sessions, the top figure showing a subject who moved relatively little and the lower one showing a



(a)



(b)



(c)

Figure 4.8: Figure 4.8a shows the average red, green, and blue signals using two different techniques for tracking intensity changes that share the same starting point. The dashed lines show the effect of keeping the region of interest fixed and the continuous lines show the effect of tracking the region of interest throughout time. Motion occurs in the first 5 seconds and the participant remains relatively still thereafter. Figure 4.8b shows the position of the fixed region of interest in the first frame, in green, and Figure 4.8c shows the region of interest as it has moved through tracking in the last frame in red compared to the original region of interest in green.

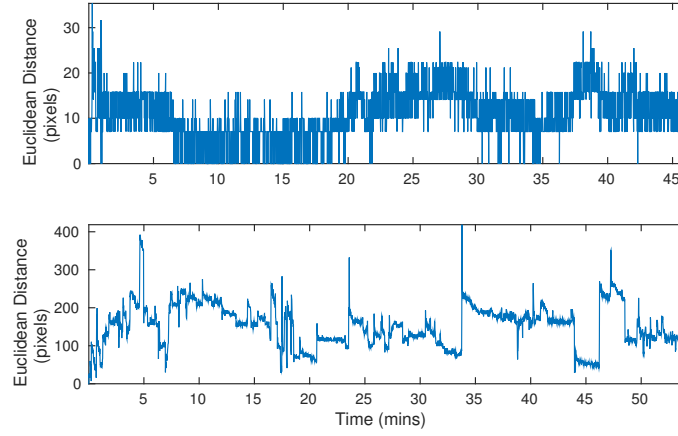


Figure 4.9: Figure showing the Euclidean distance in pixels moved by the centroid from its starting position throughout two sessions, evaluated once every second, for two subjects. The top subject (HP3-15) moved little throughout the session, whereas the lower plot shows the same for a subject (HP3-28) who moved a significant amount, as may be noted by the different vertical axes in the two plots.

subject who moved significantly throughout his session.

Motion Detection

Although the head and body have multiple degrees of freedom, whereas the method described above can only represent translations accurately, it is both fast and useful in distinguishing periods of rapid, impulsive motion from periods in which the participants are relatively still. There is always a degree of movement, since breathing causes the chest and head to move, and, as discussed in Chapter 2.5, some heart-rate related movement is also expected.

The variance of the signal may therefore be used to assess the level of motion occurring at any given point, and a number of methods are available that segment a timeseries based on changes to its mean or variance, such as Online Bayesian Change-Point Detection [2] or Gaussian Process Change Point Models [169]. However, the goal here is to find periods of minimum variance per participant, as each participant had different levels of motion and was characterised by different stable periods. For each participant, a threshold

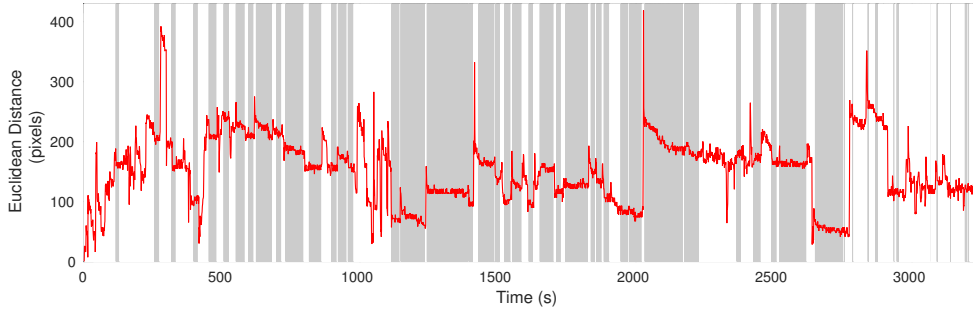


Figure 4.10: Figure showing the segmentation of the Euclidean motion data for the subject undergoing significant movement in Figure 4.9 (shown in red) into regions of no motion ($SQI_L = 0$, grey background) and rapid motion ($SQI_L = 1$, white background).

was chosen such that, if the standard deviation of a sliding window passed through the Euclidean movement signal was larger than that threshold, the window was rejected (the method used to select the threshold is described fully in Appendix A.2). This resulted in the generation of a signal quality index (SQI) for movement based on centroid location – SQI_L , essentially classifying each session into windows that were either characterised by no motion ($SQI_L = 0$), or by rapid motion ($SQI_L = 1$), as shown in Figure 4.10.

Movement Tracking

Once the time-periods during which the subject is relatively still have been identified, the motion of the head may be tracked throughout time. This is useful as it can yield an independent assessment of breathing and an understanding of what small movements are occurring that may lead to changes in reflectance. An established method of feature tracking is the Kanade-Lucas-Tomasi (KLT) feature tracking algorithm [109, 184, 210], which estimates the displacement d of a feature by minimising the residue error ϵ when aligning a template window of dimensions W in an image I at time t to an input image J at time $t + \tau$ with a weighting function w , where $\mathbf{x} = (x, y)^T$ is the position vector describing the pixel co-ordinates:

$$\epsilon = \int_W [I(\mathbf{x} - \mathbf{d}, t) - J(\mathbf{x}, t + \tau)]^2 w(\mathbf{x}) d\mathbf{x} \quad (4.9)$$

As well as providing a formalised method to track the features, the KLT algorithm may also be used to derive the optimal features to track based on the vertical and horizontal intensity gradients within the feature windows [184, 210]. For small displacements, the intensity function may be approximated by limiting its Taylor expansion to the linear term by:

$$I(\mathbf{x} - \mathbf{d}) = I(\mathbf{x}) - \mathbf{g} \cdot \mathbf{d} \quad (4.10)$$

where time has been omitted for simplicity and $\mathbf{g}$ is the derivative of I . By differentiating this equation with respect to the displacement term and rearranging, a system of two scalar equations may be obtained, such that:

$$\left(\int_W \mathbf{g} \mathbf{g}^T w dA \right) \mathbf{d} = \int_W (I - J) \mathbf{g} w dA \quad (4.11)$$

The left part of the equation yields a 2×2 coefficient matrix, with eigenvalues that represent the intensity profile of the window. The KLT algorithm selects windows for which the smallest of its two eigenvalues λ_1, λ_2 are larger than a given threshold λ_t .

The distance travelled by each feature from its initialisation may then be used to calculate the breathing rate and the displacement from the beginning, yielding additional information for the analysis if a higher resolution understanding of the displacement is necessary. One draw-back of the KLT algorithm is that it does not cope with occluded regions; however this is mitigated by the fact that the regions can be re-initialised in each steady period as determined by the previous motion segmentation analysis.

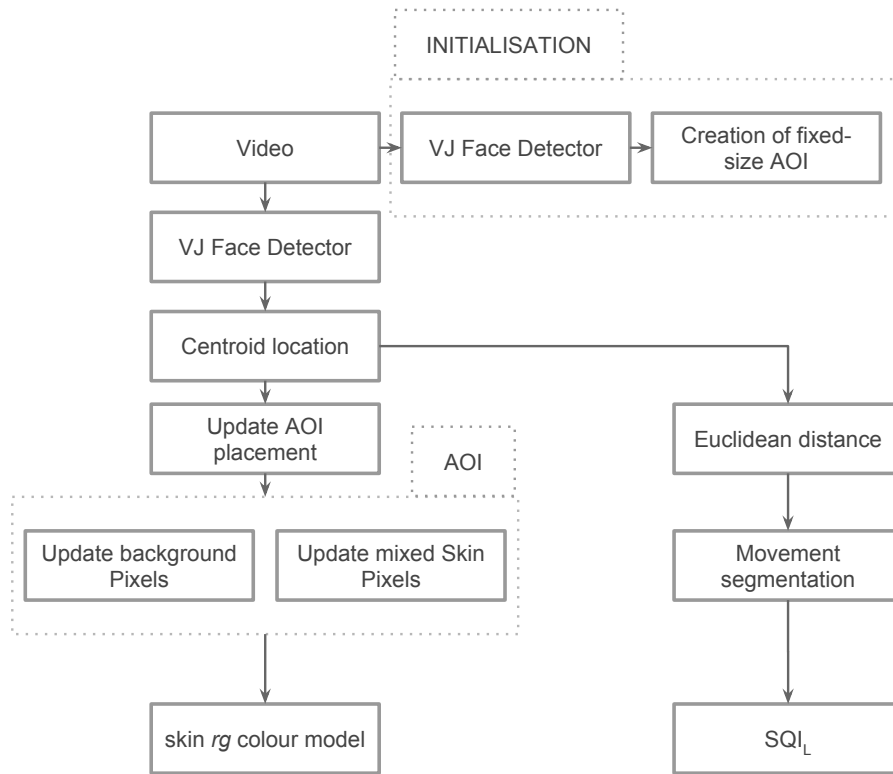


Figure 4.11: Figure showing the video and image pre-processing steps applied prior to estimating the oxygen saturation. The KLT fine tracking method and the morphological manipulations, occurring only over stable segments and frames respectively, have been omitted for clarity. AOI = Area of Interest.

4.3.4 Summary of Image Processing Steps

The model introduced in Section 4.3 outlined a set of conditions that need to be met in order for the theory to be applicable. These have been translated into processing steps that are mostly concerned with limiting the data to be analysed in order to reduce the amount of noise and the number of irrelevant pixels as much as possible prior to attempting to estimate the oxygen saturation with the methods that will be described in the following chapters. The video pre-processing pipeline, containing all the steps mentioned above, is summarised in Figure 4.11.

For each video, a Viola-Jones detector trained on frontal faces is used to detect the

position of the face. The first time the face detector is applied, the area of the bounding box around the face is expanded into the *area of interest* (AOI) and subdivided into a grid of 12×12 pixel regions of interest (ROI). The AOI will contain both all the skin pixels of the face and some background. The centroid of the face bounding box will then be tracked throughout time, and if it moves more than a certain threshold the AOI is re-initialised about the new centroid. The value of the sums of the 12×12 RGB pixels constituting each element in the larger region of interest are stored per frame throughout time to decrease computational costs and reduce rounding errors.

Once the face detector algorithm has been applied to the video data, a colour model for the subject is created by taking the elements of the AOI contained within the original region detected by the face detector and the pixels in the surrounding regions as the mixed skin and background pixels respectively. The model is created using one frame every thousand across the video (approximately equivalent to taking a frame every minute of recording). The mean element values (in RGB) are converted into *rg* values. A bimodal Gaussian model is then fitted to the skin pixels and a trimodal Gaussian model is fitted to the background pixels. The two models are then combined so as to obtain a single-mode multivariate Gaussian model for the skin pixels and a four-modal multivariate Gaussian model for the non-skin pixels. When processing the frames in later steps, each region of interest is classified either as skin or as background based on the maximum likelihood estimator for that region of interest. Finally, the morphological operations are applied to the skin mask in each frame, and all regions far from the largest continuous region of skin are rejected as non-skin.

The videos are also segmented into regions of stationarity and regions of motion as described in Section 4.3.3, and a signal quality index, SQI_L , is derived to distinguish periods of motion from periods during which the subject is still. When processing the video in segments made up of multiple frames, a set of features are initialised within the AOI using the KLT algorithm to track the motion to a greater accuracy and obtain the

respiratory trace. All the features that are not within one ROI of the skin segmented region are rejected to avoid tracking features in the background. The maximal feature motion expected by the KLT algorithm is determined by setting it to the maximum centroid motion recorded for that segment during the motion segmentation step. These features are then tracked for the duration of the stationary segment and the horizontal and vertical displacements are recorded.

Chapter 5

Oxygen Saturation and Baseline Skin Colour

5.1 Introduction

As explained in Chapter 2, it has been known since at least the early 19th century that the overall colour of the skin (equivalent to what is described as the baseline signal throughout this thesis) will change with oxygen saturation, and that even slight colour changes due to oxygenation changes may be measured using visible light. The previous chapter has also shown that this signal is however very sensitive to other factors, such as motion (see Figure 4.8), lighting, and skin tone. As a result, in previous research the baseline component has typically only been used in conjunction with the AC component, drawing from the field of pulse oximetry [63, 175, 202]. The aim of this chapter is to determine whether, with all other factors maintained as constant as possible, baseline changes in colour channels as a result of changes in oxygen saturation can be measured, above the noise level with the equipment used in the Hypoxia study. Other questions relate to the nature of the relationship for that equipment, and whether the baseline component can be used as an estimator for oxygen saturation on its own.

5.2 Methods of Analysis

5.2.1 Extracting the Baseline Signal

In order to extract the baseline component within each {R,G,B} channel, we will first make a number of assumptions about the model and its application to the data collected in the Hypoxia Study. As was stated in the derivation of Equation 4.5, the time-dependence of the AC component arises through changes in the optical path length (whether due to blood volume, capillary compression, or erythrocyte axis rotation) and changes in oxygen saturation. Whereas the former are expected to be synchronous with the heart rate, changes in oxygen saturation generally occur over longer periods of time. Similarly, the breathing rate is expected to cause changes in both the amplitude of the AC component and in the reflectance of the skin, yielding a time-varying periodic signal with varying amplitude of lower frequency. It is possible to apply a low pass filter with a window length w_L for which $B_{AC}(SO_2) = \text{constant}$ (Equation 4.5) to remove the pulsatile components. One such smoothing mechanism is the simple running-average filter,

$$M(x) = \frac{1}{2N+1} \sum_{k=-N}^N x[n+k] \quad (5.1)$$

characterised in this case by a window length of $2N+1$ time-points and an overlap of δ time-points.

If such a filter were applied to the skin ROIs of a static subject (the reader may imagine the use of a stereotactic frame), the output signal should only be affected by physiological changes to the blood and skin, such as a decrease in oxygen saturation and perfusion levels. Deriving the oxygen saturation would then only be a question of finding the optimal mapping from the colour signal to the reference oxygen saturation signal. Furthermore, once found, this mapping would be applicable to any and all of the 12×12 ROIs of the skin, as all are expected to be identical within the model proposed in Section 4.2.

However, the skin mask in Figure 4.7b shows that this is not the case. As the face is not a flat surface, the intensity of the light incident on each ROI will differ slightly from region to region. This then becomes an even greater problem when the subject moves even slightly, as the intensity measured at each ROI will be dependent on the angle of the skin with respect to the light and the camera. As a result of this, the light intensity measured from a given ROI during a static period prior to motion and the intensity measured from the same ROI during a static period following the motion may differ significantly without there having been any physiological change. These discontinuities, as well as the noise caused by the intensity changes during motion, make it impossible to track oxygen saturation from the baseline signal using one ROI alone unless the subject is extremely still, while the spatial variations in light intensity are not covered by the theoretical model, which expects constant, uniform illumination across all regions of interest.

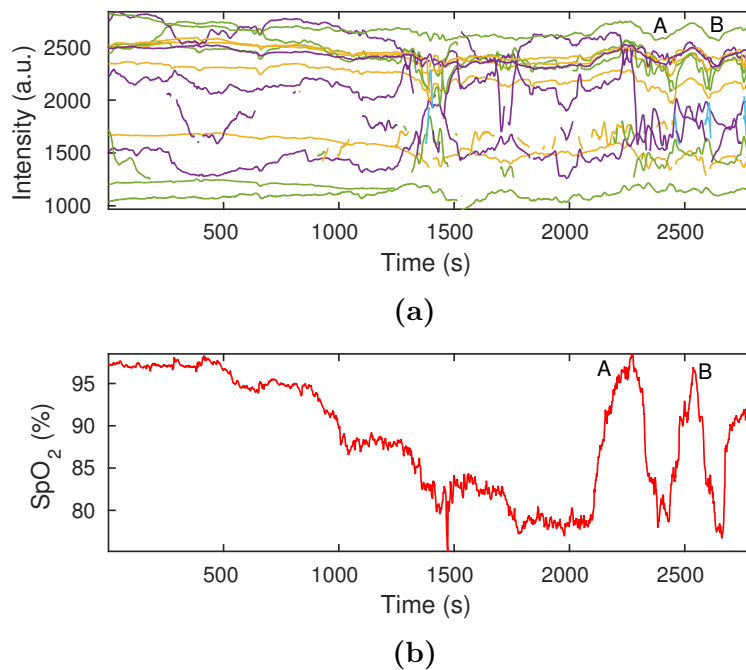


Figure 5.1: Figure showing in (5.1a) the variation in red baseline intensity for different regions of interest during the course of a session. Gaps appear in the signals if the ROI is no longer imaging a skin segment. The lower plot (5.1b) shows the average reference values of SpO₂ from the finger pulse oximeters.

This key issue in the analysis of oxygen saturation using a baseline signal is demonstrated particularly well by Figure 5.1, which shows in 5.1a the red baseline traces for one in every ten 12×12 randomly chosen skin ROI for a single subject compared to the reference oxygen saturation plot for the same subject (in 5.1b). The baseline traces in this case were generated using a running average with a window length $2N + 1 = 15$ seconds and $\delta = 14$ seconds. As a result of every timeseries corresponding to the average light intensity of that region only, the overall change in oxygen saturation may be inferred from overall trends in the data (as may be seen at points A and B, where the fast resaturations and desaturations are visible in the top signals), but, due to the unpredictability of motion and lighting change, there is no single combination of the timeseries that will yield a suitable mapping to the oxygen saturation.

5.2.2 Revising the Theoretical Model

The spatial variations in intensity of light incident on the skin are not covered by the original theoretical model, and to reflect these Equation 4.1 may be rewritten as:

$$\rho_c^x = \xi \int I^x(\lambda) \hat{L}(\lambda) S^x(\lambda) R_c(\lambda) d\lambda \quad (5.2)$$

where $I^x(\lambda)$ now represents the light intensity at position x and $\hat{L}(\lambda)$ is the normalised illumination incident on the surface. A simplification may then be made by assuming that an increase in light intensity will result in a linear increase in the intensity of each {R,G,B} channel, as also proposed by [231], such that we may write:

$$\rho_c^x = i_c^x \xi \int \hat{L}(\lambda) S^x(\lambda) R_c(\lambda) d\lambda \quad (5.3)$$

where i_c^x is the linear coefficient term for the intensity. This assumption will be tested by taking the RGB values of all ROIs within a given frame and looking at the Spearman [189] and Pearson [153] coefficients for the correlations between each pair of channels to

determine statistically whether there is a linear relationship between them that would indicate a collinearity caused by the incident light intensity (results reported in Section 5.3.1).

5.2.3 Continuous Tracking of Oxygen Saturation via the Baseline Signal

This expression of the effect of light intensity has an important consequence. If a subject's skin ROI has an $\{R,G,B\}$ value¹ $[R_i, G_i, B_i]$, according to this model other skin ROIs should have $\{R,G,B\}$ values that are multiples of this, with a multiplicative factor that depends on the intensity of the light incident on them. These points will therefore form a line in RGB-space (a Cartesian 3-dimensional space where we use the convention $R = x, G = y, B = z$) that expresses the effect of intensity on that particular colour, passing through the origin to describe the expected colour when no light is present. However, not all skin ROIs will be the same, as they will have different histological compositions and reflectance properties, so some spread is to be expected about this line. In this way, rather than considering three 1-D timeseries (one for each channel), to describe the time-course of an ROI, we may now consider all of the average skin ROIs within a frame as forming a cloud of points, and observe the movement of this cloud throughout time (Figure 5.2).

To help track the movement of the point cloud, we further assume each pixel value ρ_c^x originates from a Gaussian distribution $\rho_c \sim \mathcal{N}(\mu_{\rho_c}, \sigma_{\rho_c}^2)$. This assumption means that the average of the point cloud is representative of the sample distribution. The validity of this assumption will be evaluated by testing the hypothesis that for each ROI the $[R_c^x, G_c^x, B_c^x]$ triplet was generated by a normal distribution in each channel against the null hypothesis that it was not, using the Anderson-Darling test [146] (see Section 5.3.1).

¹From Chapter 4.3.1, each ROI has a single RGB value that is the average of the individual 12×12 pixels that make up the ROI itself.

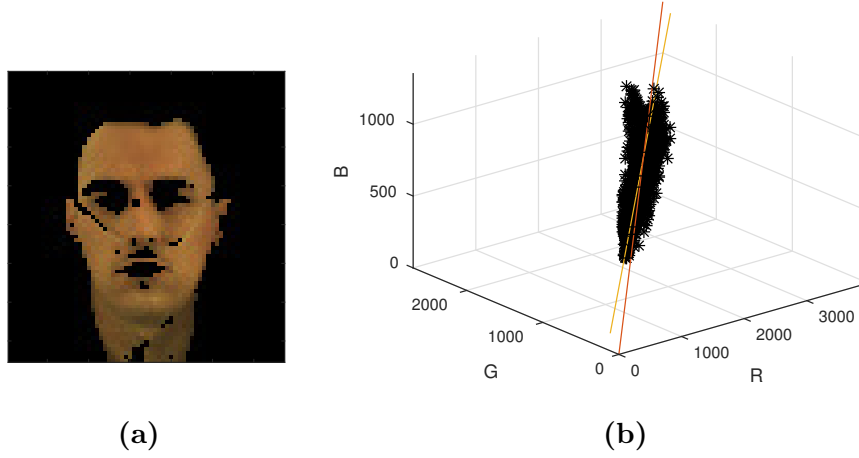


Figure 5.2: Figure showing how the RGB values of ROIs within a skin-mask (each generated by taking the average of the 12×12 pixels that form the ROI), shown in 5.2a, form a point cloud in RGB-space, shown in 5.2b. The first principal component of uncentred PCA, passing through the origin, is traced in red, while the cloud’s first principal component by mean-centred PCA is shown in yellow.

Expressing the skin as a 3-dimensional Gaussian distribution of points distributed about a line corresponding to an “axis of intensity” means that the impact of varying intensity levels (and therefore also movement) can be minimised by considering the movement of this axis, as we assume that the projection of the points in RGB-space onto this line will describe the intensity of the light itself, while the movement of the line will describe changes in physiology, including the oxygen saturation. The line thus has an equation:

$$\mathbf{l} = k\hat{\mathbf{v}} \quad (5.4)$$

where, using an approach successfully used in [231] to improve on heart rate estimation, the directional unit vector $\hat{\mathbf{v}}$ may be found by taking the first principal component $\hat{\mathbf{e}}_0$ via an uncentred Principal Component Analysis (PCA), such that $\hat{\mathbf{v}} = \hat{\mathbf{e}}_0$. Uncentred PCA highlights the importance of the origin to the distribution, rather than measuring the inclination of the cloud about its own centre of gravity [90]. By taking a line passing

through the origin as a model for the light intensity distribution of the skin, the model incorporates the assumption that at zero light intensity the measurement will be at the origin, ignoring baseline noise. Furthermore, the inclusion of the origin reduces the variability that could be expected of centred PCA due to outliers. As the vector moves about the origin, its movement may be described by a single transformation: rotation, such that a line at time m and a line at a subsequent time n may be related by the rotation matrix $\mathbf{R}_{mn}$ acting on the directional vector $\mathbf{v}_m$.

The rotation of the lines in time is calculated by finding the 3×3 rotation matrix $\mathbf{R}_{mn}$ between the first directional unit vector $\hat{\mathbf{v}}_m$ and the directional unit vectors $\hat{\mathbf{v}}_n$ at subsequent time intervals n :

$$\hat{\mathbf{v}}_n = \mathbf{R}_{mn} \hat{\mathbf{v}}_m \quad (5.5)$$

where the rotation matrix may be evaluated by the Rodrigues rotation method [42, 60] (see Appendix A.4 for further details). Equivalently, the rotational component may be tracked by observing the spherical-coordinate representation of the directional unit vector of each distribution in RGB-space throughout time in terms of the elevation angle ϕ and azimuthal angle θ . We define the elevation angle ϕ as the angle between the plane spanned by the R-G axes and the B-axis and the azimuthal angle θ as the rotation about B starting at the R-axis. By equivalence with the transformations between spherical and Cartesian co-ordinate systems, the vector defined by (θ, ϕ) may be expressed as:

$$\begin{aligned} \theta &= \arctan \frac{G}{R} \\ \phi &= \arctan \frac{B}{\sqrt{R^2 + G^2}} \end{aligned} \quad (5.6)$$

The measurement of (θ, ϕ) (shown diagrammatically in Figure 5.3) allows an evaluation of changes in the baseline values of the skin that, according to the hypothesis put forward, should minimise the effect of changes in illumination and therefore allow a continuous

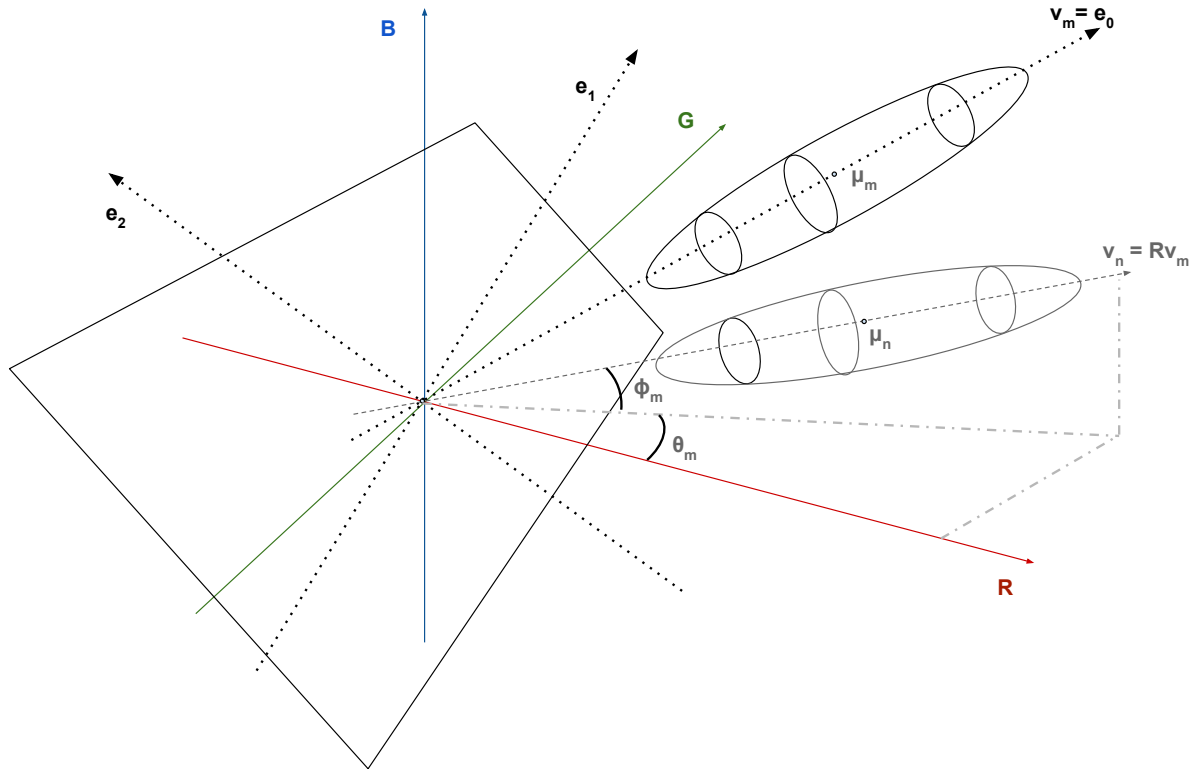


Figure 5.3: Schematic diagram of the hypothesised rotation occurring to a distribution of skin pixels in RGB space (shown as an ellipsoid) between a time m and a time n . Not to scale.

measurement of oxygen saturation. The results of this analysis are presented in Section 5.3.

5.3 Testing the Methodology: Results and Analysis

The theoretically-derived methods of analysis proposed in Section 5.2.3 were tested on a subset (12 subjects) of the Hypoxia dataset (prefix: HP3, dropped where necessary for clarity). The subjects were chosen randomly subject to the condition that two subjects would be available for each Fitzpatrick Skin type (I–VI) so as to be able to infer the effect of different skin colours to the data. The correspondance between the subject identification and Fitzpatrick skin type for the subjects analysed is summarised in Table

5.1. Note that, for the sake of clarity in the presentation of the results, the subjects are always ordered by Fitzpatrick skin type. The pre-processing of the data (the selection of the skin mask, the reduction of the ROIs to the averages of each 12×12 region, and the calculation of SQI_L) was carried out as described in Section 4.3 and a moving average filter with a 15-second window length and 14-second overlap was used to obtain the baseline signal for each ROI.

Table 5.1: Two subjects were randomly chosen from each of the six FitzPatrick Skin Types to test the methods proposed.

	FitzPatrick Skin Type					
	I	II	III	IV	V	VI
Subject ID (HP3)	20, 31	3, 25	14, 15	29, 7	24, 6	4, 42

The reference timeseries was obtained by using the average oxygen saturation from the two pulse oximeters placed on the index fingers of each hand. As the reference equipment and the camera were not perfectly synchronised, a coarse estimate of the heart rate of the subject was estimated using the camera data by taking the peak frequency component in the range $[0.7\text{Hz}, 2.2\text{Hz}]$ of the green signal for 15-second windows each with an overlap of 14 seconds (i.e. the same as the baseline signal windows). This signal was cross-correlated with the heart rate timeseries from each pulse oximeter in order to find the approximate time-delay between the two. Each of the corrected reference oxygen saturation timeseries was then resampled at the timestamps of the camera data. Only at this point were the two reference SpO_2 timeseries averaged.

5.3.1 Light Intensity and Gaussianity

In order to ascertain the linearity of the colour channels with respect to light intensity across the face, the inter-channel correlations were found throughout the video for the pixels labelled as skin. In order to distinguish between a linear correlation and another monotonic correlation, both Spearman's coefficient ρ and Pearson's coefficient r are

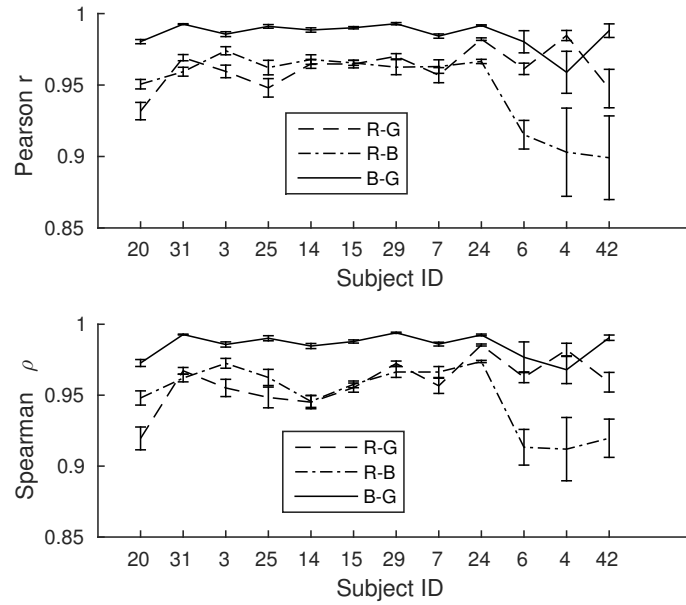


Figure 5.4: Figure showing Pearson's r and Spearman's ρ for each channel-pair. Note that the horizontal axis is ordered according to the FitzPatrick Skin Type, from Type I on the left through to Type VI on the right.

reported in Figure 5.4, using the average and standard deviation of the correlations over each session for each subject.

As may be observed, Pearson's correlation coefficient r between colour channels is consistently both positive and close to 1.0, similar to Spearman's coefficient ρ , indicating that with a high degree of probability a linear relationship exists with light intensity, as expected. The correlation between the blue and red channels in particular however is affected by skin colour as the last three subjects on the right of the plot are both characterised by the highest skin melanin content and the lowest R-B correlation. The normality of the individual colour-channel distributions was also tested per window using the Anderson-Darling test. The results are reported in Figure 5.5 as the fraction of windows for which the Anderson-Darling normality test rejected the null hypothesis with a p-value smaller than 0.01.

A further observation was made by examining whether the oxygen saturation itself

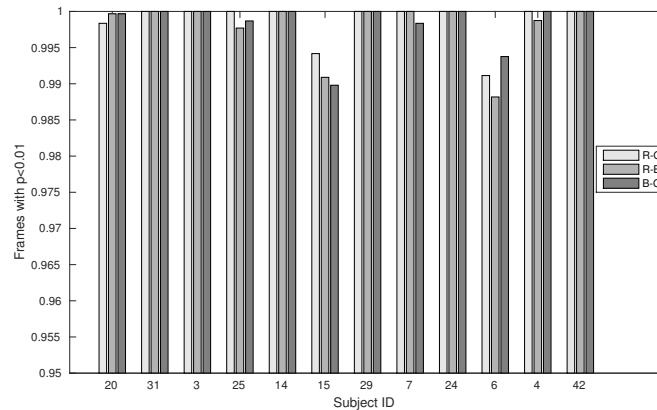


Figure 5.5: Figure showing the the fraction of frames rejecting the null hypothesis of the Anderson Darling test with a p-value < 0.01 .

may have an effect on the degree of correlation between colour channels, and the results are reported in Figure 5.6. For each subject, the correlation between the three colour channel pairs was recorded and the correlation of that with the oxygen saturation trace was reported. The results show a surprising relationship, as they show that R-based pairs are negatively correlated with oxygen saturation, indicating that the coefficient of correlation between the two colour channels increases with decreasing SpO_2 . Such a relation could be caused by a tightening of the skin-mask cloud in RGB space (as shown in Figure 5.2) with decreasing SpO_2 , or by an increasingly non-linear relationship between the colour channels with increasing SpO_2 . On the other hand, the high variability in this phenomenon from subject to subject could mean the effect is simply artefactual.

These results demonstrate that the assumption of Gaussianity, and to some extent that of linearity for the colour-intensity relationship, are valid. As these assumptions are necessary for the method of tracking the colour distribution by its rotation, we may now investigate the relationship between these measurements and the oxygen saturation.

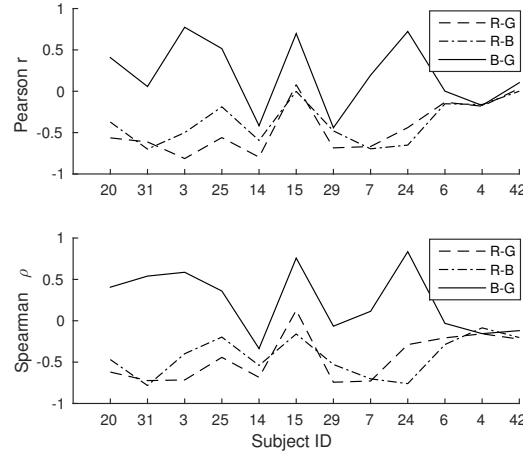


Figure 5.6: Figure showing the correlation between each channel-pair's correlation coefficients and the oxygen saturation.

5.3.2 Multiple Linear Regression

To examine the relationship between the baseline values of the RGB channels and oxygen saturation, the rotational data, tracked using the azimuthal angle θ and the elevation angle ϕ were related to the oxygen saturation. This relationship was investigated per subject using the linear model

$$\hat{S}_\beta = \beta_0 + \beta_1\theta + \beta_2\phi \quad (5.7)$$

where $\hat{S}_\beta$ is the estimator of the reference oxygen saturation S with coefficients that minimise the error between the actual and estimated values. The strength of the correlation between the estimator and the reference oxygen saturation signal is reported for each subject in Figure 5.7 using the coefficient of determination r^2 . Each histogram also shows the effect of the signal quality index SQI_L , which rejects rapid movement as described in Chapter 4.3.3. As a strong negative correlation was found between the elevation and the oxygen saturation, the following linear model was also investigated, using the polar angle $(\frac{\pi}{2} - \phi)$:

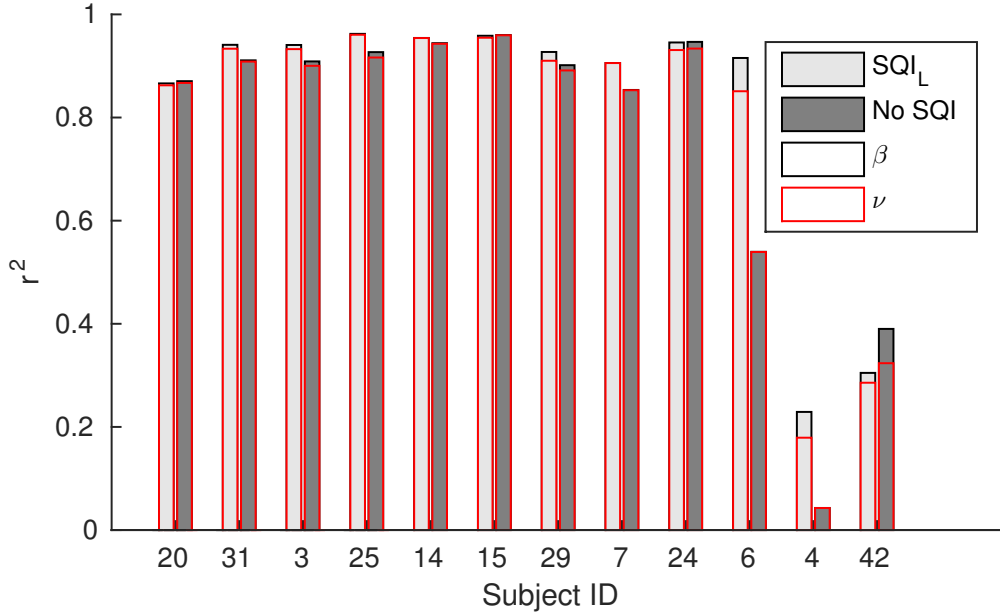


Figure 5.7: Bar graphs showing the coefficients of determination between oxygen saturation and the regression equations (5.7) and (5.8) with no movement SQI and with the application of SQI_L, which rejects rapid motion.

$$\hat{S}_\nu = \nu_0 + \nu_1\left(\frac{\pi}{2} - \phi\right) \quad (5.8)$$

The proportion of data when using the signal quality index (SQI) is shown in Figure 5.8. As may be observed, in each session around half of the windows are rejected due to motion by SQI_L, but the use of the SQI does not significantly improve the r^2 value, as shown in Figure 5.7, with the exception of subjects HP3-06 and HP3-04.

The accuracy of the linear regression expressed in Equation 5.7 is also shown in Figure 5.9, where $\hat{S}_\beta$ is shown plotted against the true oxygen saturation for each of the subjects. The red line in each plot is the average oxygen saturation whereas the continuous black line is the regressed camera estimate of the oxygen saturation $\hat{S}_\beta$ with no motion SQI. The regressed camera estimate using SQI_L is overlaid with blue points.

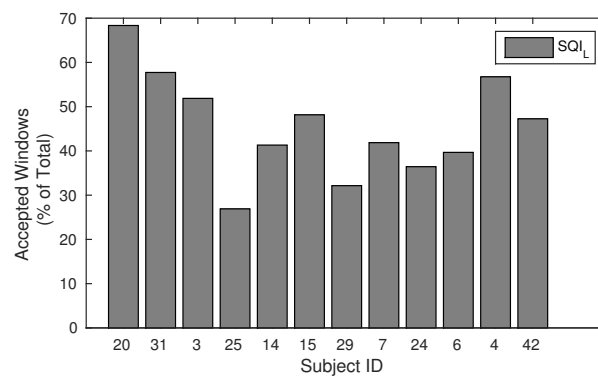


Figure 5.8: Figure showing the proportion of windows that are considered to have no motion for each subject according to SQL_L.

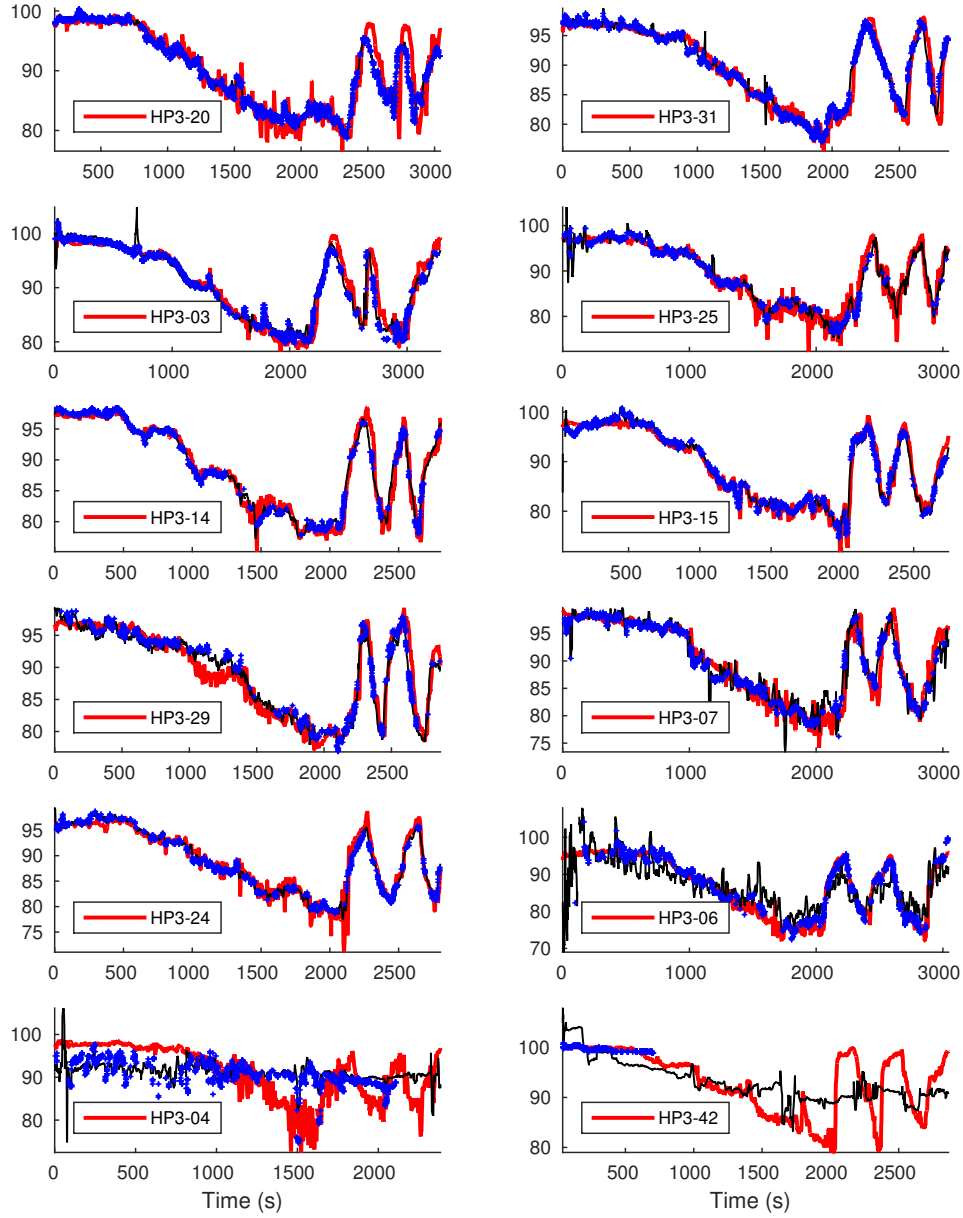


Figure 5.9: Figure showing the regressed camera data $\hat{S}_\beta$ against the reference SpO_2 (red trace) for each of the 12 test subjects from Table 5.1. The black line and the blue points respectively show the results using no SQI and SQI_L .

5.3.3 The Effect of Skin Colour

Given the high correlation ($r^2 > 0.85$) for skin types I through to V, the relationship between the relative importance of the two coefficients of regression and the skin colour was investigated. The results, summarised in Figure 5.10, are inconclusive, as they do not show any clear skin colour-related effect. Standardised regression coefficients (β_1^s and β_2^s) are used in order to compare the relative effect of the coefficients without the influence of the scale of the variables. To obtain the standardised regression coefficients, the regression was carried out on mean-centred and unit variance data, and the motion SQI_L was used for optimal fit.

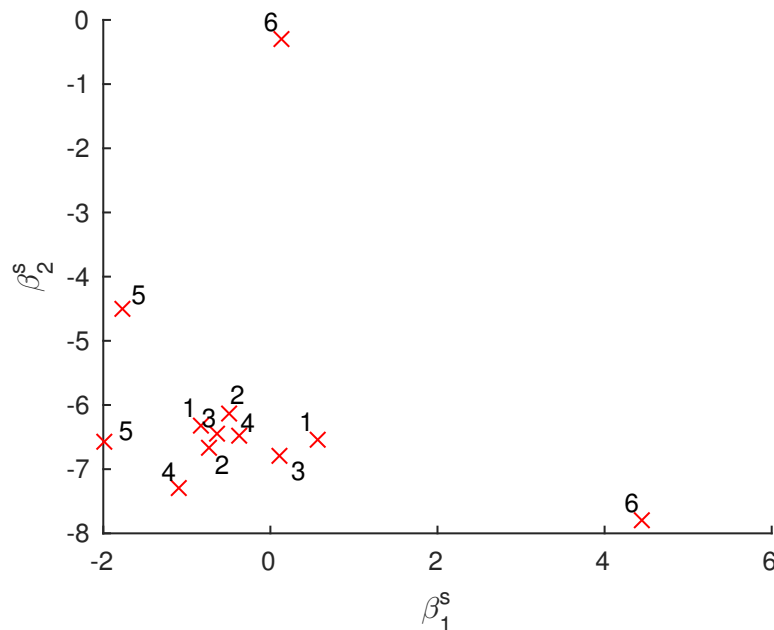


Figure 5.10: Figure showing the values of the standardised linear regression coefficients β_1^s and β_2^s , indexed according to skin colour type (1-6).

The lack of correlation between the standardised regression coefficients and the skin colour may be due to another unknown variable affecting the relationship, however Figure 5.10 also reveals that the elevation angle is the better predictor for oxygen saturation. This may indicate a low sensitivity of the coefficients to the regression, i.e.

because one coefficient dominates, different coefficient combinations may yield similar performances. In order to test this, a sensitivity analysis was carried out on the coefficient of determination as a function of the values of β_1 and β_2 , each varied between -20 and 20 in steps of 1. The results, shown in Figure 5.11, suggest that the regression coefficients are in fact relatively insensitive to these variations, and mostly cluster around the $\beta_1 = 0$ axis, indicating that the elevation component is far more important than the azimuthal component.

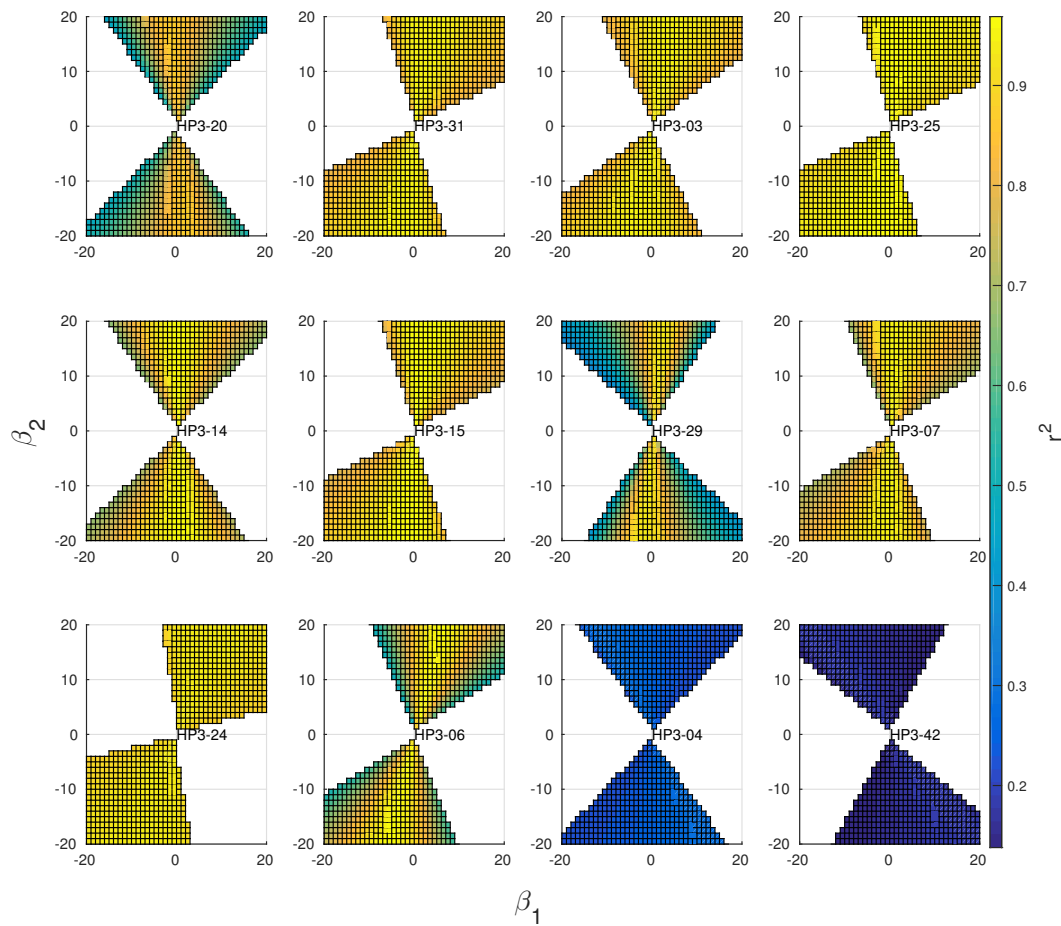


Figure 5.11: Coefficient of determination r^2 for different values of β_1 and β_2 . Only values above the median r^2 of each subject are shown for clarity.

Yet another possibility is that the FitzPatrick scale is unable to capture properly the

differences in melanin between subjects, as it is both subjective and based on very coarse phenotypic features, or that it did not capture other light-interacting properties of the skin such that the trend would be hidden. The colour signature of the subjects, defined by its (θ, ϕ) co-ordinates should however capture this information objectively. Figure 5.12 shows all of the (θ, ϕ) co-ordinates for all subjects plotted onto a single graph with the reference oxygen saturation overlaid in colour (Type VI subjects were removed for clarity). The main direction of the variation in oxygen saturation is shown as an arrow on three of the subjects, and the lack of any relationship between the direction of SpO_2 variation and the relative starting positions of the subjects in (θ, ϕ) is made clear. A quick comparison with Figure 5.10 also reveals that there is no apparent relationship between the relative weighting of the variables and their values in (θ, ϕ) space.

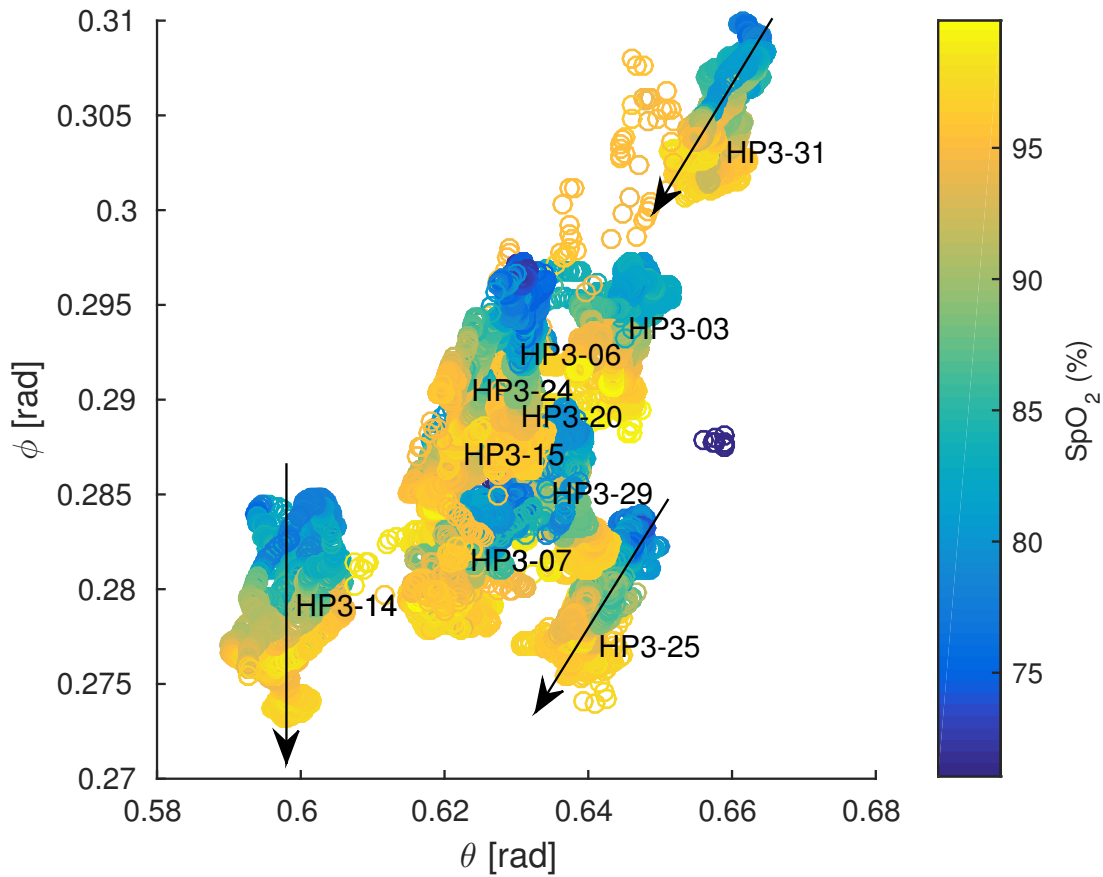


Figure 5.12: Figure showing the oxygen saturation (increasing from dark to light) as a function of the rotation components for all of the subjects at once. Type VI subjects have been removed for clarity. This plot makes the lack of any relationship between the skin colour, the values of (θ, ϕ) for each subject, and the SpO_2 dependence on (θ, ϕ) apparent. Outliers for HP3-06 may be seen around co-ordinate values $[0.66, 0.285]$ and $[0.64, 0.3]$.

5.3.4 Variable Interpretation

Figures 5.7, 5.10, and 5.12 suggest that the polar, or elevation, rotation ϕ is the single most important contributing variable. This is in line with the general principle noted since at least the 18th century that skin will tend towards a bluer colour as a result of hypoxaemia as the rotation essentially describes a proportional decrease of red and green with respect to blue.

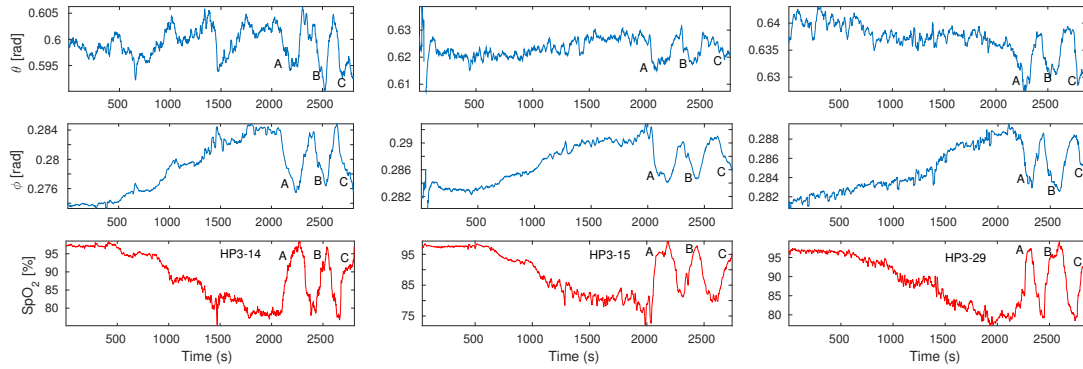


Figure 5.13: Plots of the movement of θ and ϕ (in blue) are shown plotted against the reference SpO_2 (in red, at the bottom) for three subjects. From these plots, ϕ can be seen to be negatively correlated with SpO_2 , whereas θ only seems to be affected by the rapid oxygenations (shown at positions A, B, and C) and the subsequent deoxygenations.

The linear models described by the set of equations 5.7 consider the azimuthal rotation θ a further independent factor related to the oxygen saturation. However a compelling hypothesis arising from the results is that, in the same way that the polar component ϕ can be said to be related to oxygen saturation as a colour change towards blue, the azimuthal component θ may be related to perfusion (the volume of blood present in the skin imaged) and vessel tone by the relative colour change between red and green.

Figure 5.13 shows representative plots of θ , ϕ and the reference SpO_2 for three of the subjects, and these results are in agreement with the hypothesis in two main ways. Firstly, whereas ϕ is consistently negatively correlated with SpO_2 , θ is not as consistent. What is consistent between the subjects is that during the reoxygenation stage (positions A, B,

C in each plot) there is a marked decrease in θ , i.e. the skin-mask cloud in RGB-space rotates towards the red axis and away from the green. Reoxygenation may lead to hyperoxia – an excess of oxygen available to tissue that has been reported to be a strong stimulus for vasoconstriction [240]. Because this occurs suddenly, there is less time for the homeostatic balancing that might be occurring during the slower deoxygenations, thereby leading to a more marked change in vessel tone.

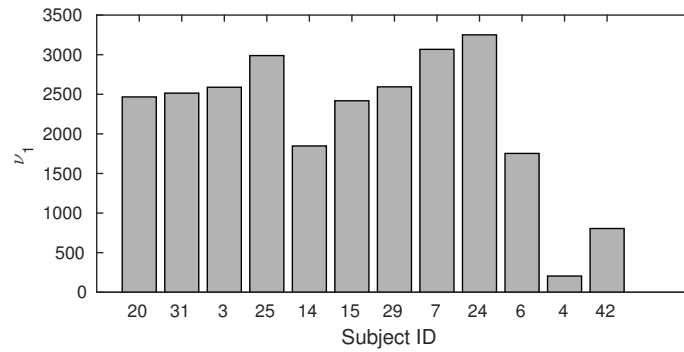


Figure 5.14: Even when using only one variable ($\frac{\pi}{2} - \phi$), there seems to be no relationship between the skin colour of the subject as measured by the FitzPatrick Skin Type Scale and the gradient of the regression of $\frac{\pi}{2} - \phi$ onto the reference SpO_2 .

Both the sensitivity analysis and the initial comparison of coefficients of determination in Figure 5.7 lead to the conclusion that the trend in oxygen saturation for skin types I-V may be well described by the variable ϕ alone. However Figure 5.14 shows that the coefficient ν varies greatly across subjects with no clear relation to skin colour (Figure 5.10).

5.4 Conclusion

This chapter has introduced a novel method for the detection of changes in oxygen saturation using the baseline component of the skin signal that has been derived from first principles. The method was evaluated on a subset of the videos of the Hypoxia dataset. The hypothesis that cameras could be used to track changes in oxygen saturation

was thus confirmed ($r^2 > 0.85$) on subjects with skin types I through to V. However, the method was unable to detect changes in oxygen saturation for subjects of skin type VI. Of the dependent variables examined, it was found that the most relevant was the elevation angle ϕ , which is negatively correlated with oxygen saturation.

Chapter 6

Oxygen Saturation and Changes in AC

6.1 Introduction

Changes in the AC component of heart-synchronous timeseries of transmitted or reflected light have been used to calculate oxygen saturation since the advent of pulse oximetry. As stated in the background chapter, the crucial advantage of the AC component in classical pulse oximetry is that it yields results that can be related to oxygen saturation through a look-up table without the need for constant calibration as had been the case for oximetry-based devices.

The objective of this chapter is to investigate the relationship between the AC signal and the oxygen saturation for the equipment used in the Hypoxia study. The theoretical basis for the use of the AC signal in the detection of oxygen saturation is first outlined in Section [6.2](#). Methods to detect the photoplethysmogram and recover the signal amplitude reliably using standard signal processing techniques are described here and tested in Section [6.3](#).

These initial results inform a refinement of the methods used to select the regions

of interest (ROIs), and these are investigated in Section 6.4. Finally, in Section 6.5 a set of linear models are used to investigate the relationship between oxygen saturation, changes in AC, and skin colour.

6.2 Theory

6.2.1 Theoretical Model

The variations in the amplitude of the AC component may be clearly related to changes in oxygen saturation by using Equation 4.5, the theory introduced in Chapter 4 and the introduction of the intensity term in Chapter 5 such that:

$$\begin{aligned}\rho_{c,b}^x &= i_c^x \xi v_b^x(t) \int \hat{L}(\lambda) M^x(\lambda) \hat{B}_b^x(SO_2) R_c(\lambda) d\lambda \\ \rho_{c,AC}^x &= i_c^x \xi v_{AC}^x(t) \int \hat{L}(\lambda) M^x(\lambda) \hat{B}_{AC}^x(SO_2) R_c(\lambda) d\lambda\end{aligned}\tag{6.1}$$

Here the treatment is different to that of the baseline signal: the inter-ROI variations in intensity i_c^x may simply be removed by taking the ratio of the AC of one channel to the baseline of the same channel for any given ROI. It is also immediately apparent that according to the model proposed the ratio of ratios of two colour channels over a single ROI will also remove the volume-related terms v_b^x and v_{AC}^x , as, for two channels c_1 and c_2 we have:

$$\frac{\rho_{c_1,AC}^x / \rho_{c_1,b}^x}{\rho_{c_2,AC}^x / \rho_{c_2,b}^x} = \frac{\frac{\int \hat{L}(\lambda) M^x(\lambda) \hat{B}_{AC}^x(SO_2) R_{c_1}(\lambda) d\lambda}{\int \hat{L}(\lambda) M^x(\lambda) \hat{B}_b^x(SO_2) R_{c_1}(\lambda) d\lambda}}{\frac{\int \hat{L}(\lambda) M^x(\lambda) \hat{B}_{AC}^x(SO_2) R_{c_2}(\lambda) d\lambda}{\int \hat{L}(\lambda) M^x(\lambda) \hat{B}_b^x(SO_2) R_{c_2}(\lambda) d\lambda}}\tag{6.2}$$

Extrapolating for multiple regions, for conditions where $M^x(\lambda) = M(\lambda)$, $\hat{B}_b^x(SO_2) = \hat{B}_b(SO_2)$ and $\hat{B}_{AC}^x(SO_2) = \hat{B}_{AC}(SO_2)$, this relationship holds also for averages of $\rho_{c,b}^x$ and $\rho_{c,AC}^x$ over multiple x . These conditions essentially represent cases for which the various skin elements sampled are anatomically similar and for which the oxygen saturation of the baseline and AC are the same. As we will see in the following sections, this extrapolation

will allow the actual implementation of the AC-based oxygen saturation tracking.

The rationale for using a ratio of ratios (RoR) to track oxygen saturation thus emerges from the application of few assumptions to a model that is different to that used in conventional pulse oximetry, as it does not rely on the Beer-Lambert law and the manipulation of a single exponential. However, contrary to the latter, it does not provide any information as to the nature of the relationship between the ratio of colour channels and the oxygen saturation. This will have to be determined empirically and will only be valid for the particular lighting conditions given by the normalised light distribution $\hat{L}(\lambda)$.

6.2.2 Amplitude Estimation

A number of strategies may be employed in order to estimate the amplitude of a periodic signal, a field well studied within conventional pulse oximetry. Conceptually, the simplest of these is to establish the values of consecutive peaks and troughs and evaluate the peak-trough difference (PTD). In order to produce a robust estimate, as the peak positions and heights are prone to noise, the median PTD (PTD_{med}) or the average PTD (PTD_{μ}), over a series of peaks may be found. For this method to be applicable, a high signal-to-noise ratio is required: the signal cannot however be filtered as the characteristic sawtooth shape of a photoplethysmographic signal is due to its higher harmonics: even an ideal filter with a flat pass-band would still risk removing the information associated with such harmonics.

Another time-domain method that further reduces the noise present in a signal by averaging is known as signal averaging (SA), whereby a representative waveform is constructed by using a reference beat to align and average together multiple sections of a noisy timeseries. By taking a number of samples slightly greater than that which corresponds to half the period of the dominant frequency either side of the waveform peak (a characteristic feature of the PPG waveform, robust to noise), the signal-to-noise

ratio is increased over time for any periodic signal having noise that has a mean of zero.

The amplitude of the signal may also be determined in the frequency domain by examining the power of the spectrum of the Discrete-Time Fourier Transform (DTFT) at the frequency of interest f_{HR} . As the discrete frequencies in the DTFT are separated by $\frac{f_s}{N}$, where f_s is the sampling frequency and N is the length of the input time series, an accurate estimate of the amplitude of a periodic signal will only occur if there is no spectral leakage. This occurs when $\delta = 0$ for

$$\frac{f_{HR}}{f_s} = \frac{l + \delta}{N} \quad (6.3)$$

where l and δ are the integer and fractional parts of the full cycle of f_{HR} within the window [14]. In most applications, however, $\delta \neq 0$, and a number of different methods may be employed to minimise the amplitude estimation error. One way of accomplishing this is to zero-pad the signal so as to minimise δ , while [148] and [14] offer a mathematical solution (the energy-based method or EBM) to the problem, relating the amplitude $\hat{A}$ to the DTFT of a windowed signal $X_w(\lambda) = \text{DTFT}\{x_w(n)\}$ for an H -order cosine window

$$w(n) = \sum_{h=0}^{H-1} (-1)^h a_h \cos\left(2\pi h \frac{n}{N}\right), \quad n = 0, 1, \dots, N-1 \quad (6.4)$$

by the equation

$$\hat{A} = \frac{2}{N} \sqrt{\frac{\sum_{k=-H}^H |X_w(l+k)|^2}{\text{NNPG}}} \quad (6.5)$$

where NNPG, the normalised noise power gain, is defined as

$$\text{NNPG} = \sum_{m=0}^{N-1} \frac{w^2(m)}{N} \quad (6.6)$$

As with the signal averaging method, the DTFT-based methods are computationally efficient and robust to uncertainties [14], but they are limited to cases for which the

signal has a constant frequency. The main drawback of DTFT estimates, however, is that they are unable to assess the true waveform peak amplitude for waveforms composed of multiple sinusoids. While tracking the amplitude of the fundamental may lead to an accurate estimate of oxygen saturation changes, this will not equate to the peak-trough difference for anything other than a perfect sinusoid.

6.2.3 Spatially Increasing the SNR

The methods described above all require a relatively high SNR to accurately estimate the amplitude of the waveform peak, and each uses the periodicity of the signal to improve on the signal-to-noise ratio. Another way of increasing the SNR is instead to average spatially (i.e. over the pixels) rather than, or as well as, temporally. The relationship between the size of the ROI and SNR_{HR} has been established in previous literature devoted to the development of heart rate monitoring using video cameras and ambient light (see Chapter 2.5.3). As discussed in Chapter 4, it may be explained by a simple model with an underlying periodic signal corresponding to the heart rate, noise, and the quantisation effect of the camera. The increase in SNR_{HR} obtained by averaging over a larger area of skin results from the decrease in the noise associated with quantisation and other sources (see Chapter 4.3 for further details on the possible origin of the noise). This assumption has already been used in the spatial down-sampling of the videos, where the mean values of 12×12 ROIs are taken for each video rather than the original frames. However, a model of this process is useful as it allows the investigation of the effects of the ROI size on the results. Crucially, the performance of the three amplitude estimation methods can be evaluated on a known dataset before being used to investigate oxygen saturation, and furthermore the model allows an exploration of the relationship between the SNR and the amplitude estimate errors.

The model parameters for the heart rate signal were generated based on those found for the green channel of a Type III subject (HP3-15) resulting from averaging over

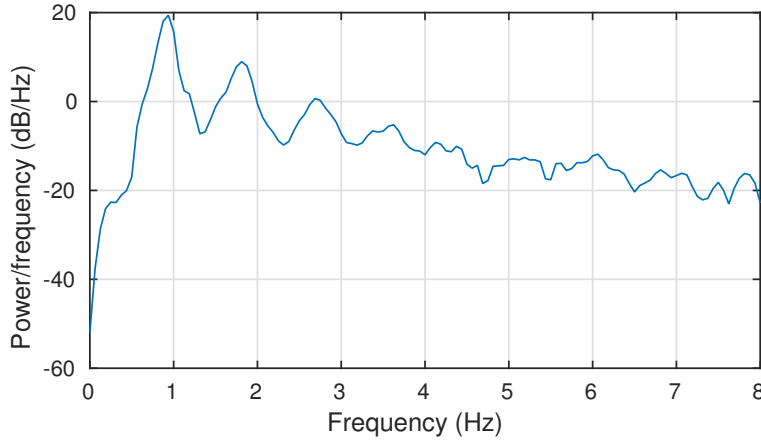


Figure 6.1: Welch periodogram of a 60 second green average of the forehead of HP3-15.

a large (160×70 pixels) forehead region during a manually selected 1-minute period in which the subject was still and there was no oxygen saturation change. The green signal was detrended and the Welch periodogram of the signal was computed, giving the result shown in Figure 6.1. The periodogram was generated using six windowed (Blackman-Harris) segments with half overlap. Three peaks, corresponding to the first three harmonics, are clearly visible, and the amplitude of the harmonics as determined by the EBM was used in the generation of the coefficients of the individual sinusoidal components ($\alpha_1, \alpha_2, \alpha_3$) of the model heart rate signal x_{HR} with a fundamental angular frequency ω_{HR} :

$$x_{HR} = \alpha_1 \sin(\omega_{HR}t) + \alpha_2 \sin(2\omega_{HR}t) + \alpha_3 \sin(2\omega_{HR}t) \quad (6.7)$$

The amplitude of the Gaussian white noise applied to the model was instead calculated using the average power (σ_n^2) of the detrended signals from each pixel within the forehead ROI. The quantisation introduced by the camera's analogue-to-digital converter (ADC) was modelled as a mid-tread uniform quantiser, such that each model pixel is now defined as

Table 6.1: Model parameters for an HR+noise-based single-pixel model

Parameter [unit]	Value
Sampling frequency (f_s) [Hz]	16
HR frequency (f_{HR}) [Hz]	1.1
Window length [s]	60
Fundamental amplitude (a) [a.u.]	7.0
2 nd harmonic amplitude (b) [a.u.]	1.4
3 rd harmonic amplitude (c) [a.u.]	0.4
Noise amplitude (σ_n) [a.u.]	32
Number of harmonics (M)	3

$$p_{HR} = \lceil x_{HR} + \epsilon \rceil, \quad \epsilon \sim \mathcal{N}(0, \sigma_n) \quad (6.8)$$

The effect of the ROI size was then related to the signal-to-noise ratio of the heart rate component of the signal, calculated in decibels (SNRdB). It is essential to relate the ROI size to the signal-to-noise ratio as the former depends on a number of factors, including distance from the camera, camera lens, skin colour and skin perfusion. The SNRdB parameter, on the other hand, although not a perfect estimator (as it is also expected to vary independently of noise due to oxygen saturation and skin perfusion) can at least be related from one real-world dataset to the next.

With the properties of the model thus defined (the model parameters used may be found in Table 6.1), the performance of each of the methods to calculate the total peak amplitude a_{HR} described in the previous section may be compared (Figure 6.2). The figure shows the results for PTD, SA and EBM₁ (the amplitude of the fundamental frequency as determined by the EBM algorithm) for one instance of the model. It is indicative of a number of important features of the estimators and the effect of noise on the data.

Firstly, the relationship between the number of pixels used to and the SNRdB is logarithmic, demonstrating that there is indeed a linear relationship between SNR and ROI size for this model. Secondly, the predicted error of the EBM method, due to

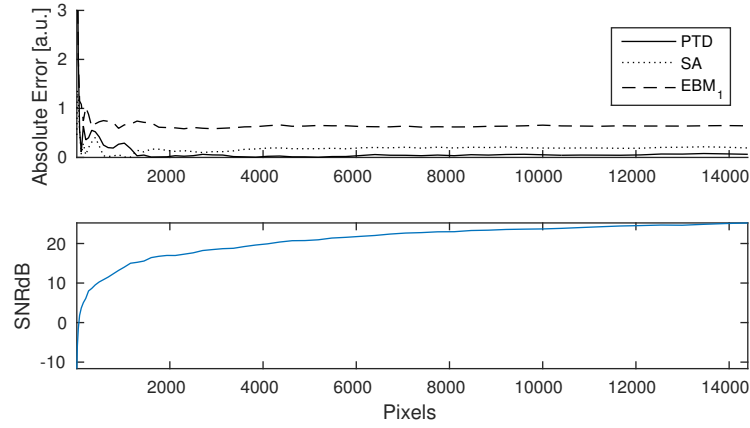


Figure 6.2: Performance of PTD, SA and EBM algorithms on modelled data, comparing the signal-to-noise ratio (measured in decibels, lower plot) to the absolute error between the algorithms' estimates and the true amplitude (top plot) for a signal-averaged over increasing numbers of pixels.

the fact that there are multiple harmonics present, is clearly apparent, as it tends towards a constant error with respect to the true amplitude. Thirdly, there is a high variability in the output of all amplitude estimation algorithms at low signal-to-noise ratios ($\text{SNRdB} < 0\text{dB}$), with the estimates stabilising at around 15dB for this particular case. This indicates that, as hypothesised, averaging over a large number of ROIs improves the accuracy of the amplitude estimate. In addition, the variability in estimates may explain the result found in [88], where it is noted that the ratio of colour amplitudes would change for different ROI sizes, for a study in which the experiments were conducted with a contact-based method, pressing the camera onto the finger. Most interestingly however, the absolute errors (AE) for PTD and SA do not tend to zero, and the values of all three estimators continue to vary even at high SNRs. This in turn indicates that some variability is expected from one AC measurement to another that will not be explained by SpO_2 nor any other factor other than noise.

6.3 Initial Results

The theory and the initial methods for oxygen saturation estimation proposed in the previous sections were tested on the data from the same 12 Hypoxia study subjects as had been used in the tests for the baseline signal. The following research questions were investigated:

- whether the AC signal contained any information as to the oxygen saturation of the subjects
- which individual channel or ratio of channels had the best performance with respect to SpO_2
- whether the three methods (PTD, SA, and EBM) proposed were equally robust in estimating oxygen saturation
- whether the method of evaluating the baseline signal had an impact on results.

The results were obtained by taking 15 sliding windows across each session with a 14 second overlap. All ROIs classified as skin according to the method described in Chapter 4.3.2 (see Figures 4.7 and 5.2 for examples of such skin masks) were used as signal sources.

6.3.1 Heart Rate

Unlike the baseline method proposed, the AC methods rely on an accurate estimation of the heart rate to work properly, as the latter is used in every amplitude estimate proposed except for the simple peak-detection method. The error in the heart rate estimate with respect to the pulse oximeter is limited by two main factors: the parameters used in its estimation by the video-based system, and the parameters used by the pulse oximeter, including any further post-processing. In particular, the camera-based system uses a

15-second window and a Fast Fourier Transform (FFT), yielding a resolution of 0.067Hz, equivalent to 4 beats per minute (bpm).

Two methods for heart rate estimation were tested to determine which would provide the best estimate for the Hypoxia dataset (adult subjects, limited movement, variability in skin colour): (a) the estimate corresponding to the frequency of every green ROI (HR_{gf}), band-passed filtered using a zero-phase bandpass FIR filter with a pass-band between 0.7 and 2.1Hz (42–132 beats per minute¹) – the standard method for heart rate extraction in use throughout the heart rate literature and found to perform adequately in conjunction with auto-regressive models in [202] – and (b) the method proposed in [45], known as CHROM, which was tested by the authors on several different skin colour types. This method, described in detail in Appendix A.5, uses a linear combination of all three channels which has been empirically shown to give improved heart rate estimates. The same bandpass filter as described above was used within the CHROM algorithm. The performance of these two methods is expressed in Table 6.2 as the mean absolute error between the estimates from each method and the left and right pulse oximeters (MAE_L and MAE_R respectively).

To take the maximal resolution of the FFT into account, the results are also shown in Table 6.3, expressed as the percentage of windows for which the estimated heart rate using each method is more than 4bpm away from the reference heart rates. As may be observed, these represent a clear minority of estimates in many subjects (particularly HP3-24, HP3-04, and HP3-42) and are ignored in all subsequent analysis. The results presented in these two tables highlight the CHROM algorithm put forward in [45] as being significantly more effective than a simple filtering operation on the green channel in determining the heart rate in Type V (HP3-06 and HP3-24) and VI (HP3-04 and HP3-42) subjects. Throughout the rest of this thesis unless otherwise stated, the heart rate will therefore be calculated using the CHROM algorithm to maximise performance.

¹These correspond to the range of the heart rates within the Hypoxia dataset (see Chapter 3)

Table 6.2: Heart Rate MAE - CHROM [45] and HR_{gf} [202] algorithms as explained in the text.

	CHROM		HR _{gf}	
	MAE _L	MAE _R	MAE _L	MAE _R
HP3-20	1.83	1.86	1.80	1.84
HP3-31	1.39	1.36	1.47	1.44
HP3-03	1.65	1.56	1.78	1.64
HP3-25	2.68	2.70	2.87	2.88
HP3-14	2.11	2.14	2.09	2.13
HP3-15	1.66	1.66	1.64	1.65
HP3-29	1.20	1.23	1.22	1.25
HP3-07	1.96	1.90	2.46	2.36
HP3-24	1.11	1.07	1.10	1.07
HP3-06	1.60	1.63	2.43	2.42
HP3-04	3.82	3.78	9.55	9.48
HP3-42	4.70	4.65	6.51	6.44

Table 6.3: Proportion of estimated HR with a difference greater than 4bpm with respect to the reference HR for the two algorithms considered.

	CHROM		HR _{gf}	
	MAE _L	MAE _R	MAE _L	MAE _R
HP3-20	0.12	0.12	0.12	0.12
HP3-31	0.04	0.04	0.05	0.05
HP3-03	0.09	0.08	0.10	0.09
HP3-25	0.25	0.25	0.26	0.27
HP3-14	0.16	0.16	0.16	0.16
HP3-15	0.09	0.10	0.09	0.09
HP3-29	0.02	0.02	0.02	0.03
HP3-07	0.13	0.12	0.15	0.15
HP3-24	0.01	0.01	0.01	0.01
HP3-06	0.06	0.06	0.11	0.11
HP3-04	0.29	0.28	0.49	0.48
HP3-42	0.22	0.22	0.37	0.37

6.3.2 Oxygen Saturation

The three different methods described in Section 6.2.2 were tested on the 12 test subjects from Table 5.1. The heart rate used in the EBM method, as well as the heart rate signal used for peak detection in the PTD method and peak alignment in the SA method was calculated using the CHROM method operating on all skin ROIs. Two different methods of normalising the AC signals were used: using the average DC (DC_{μ}), whereby each window is normalised using the average of all samples within the window, and a low-pass filter normalisation (DC_{lpf}), whereby the AC signal is normalised by a low-pass filtered version of the overall signal (the low-pass filter used was a 48th order FIR filter with a cut-off frequency of 0.67Hz). The results, reported as the coefficient of determination r^2 between the mean reference oxygen saturation of the two pulse oximeters and the video camera signal, are reported in Figure 6.3. For each method the correlation was estimated on single channels (R,G,B) and ratios of the channels (R/G, R/B, G/B, G/R, B/R, B/G).

Figure 6.3 is useful in comparing the different methods and the relationships between the three colour channels and oxygen saturation. Subjects are grouped in pairs of the same Fitzpatrick skin colour type going from Skin Type I to Skin Type VI from bottom to top. It is immediately obvious that the signal averaging (SA) method consistently outperforms all other methods. Regardless of the choice of method or channel combination subjects HP3-31, HP3-06, HP3-04, and HP3-42, yielded very poor results. This may be due to increased levels of motion or the effect of melanin in light absorption, as the last three subjects have the highest Fitzpatrick Skin Types (V and VI). The type of baseline normalisation used (using a low-pass filtered signal or the average over the window) was explored for the SA method and the low-pass filtered version (SA_{lpf}) was found to perform best.

Of the colour channels, it is the amplitude of the red channel that is most affected by oxygen saturation: the ratios with the red normalised AC signal in the numerator or

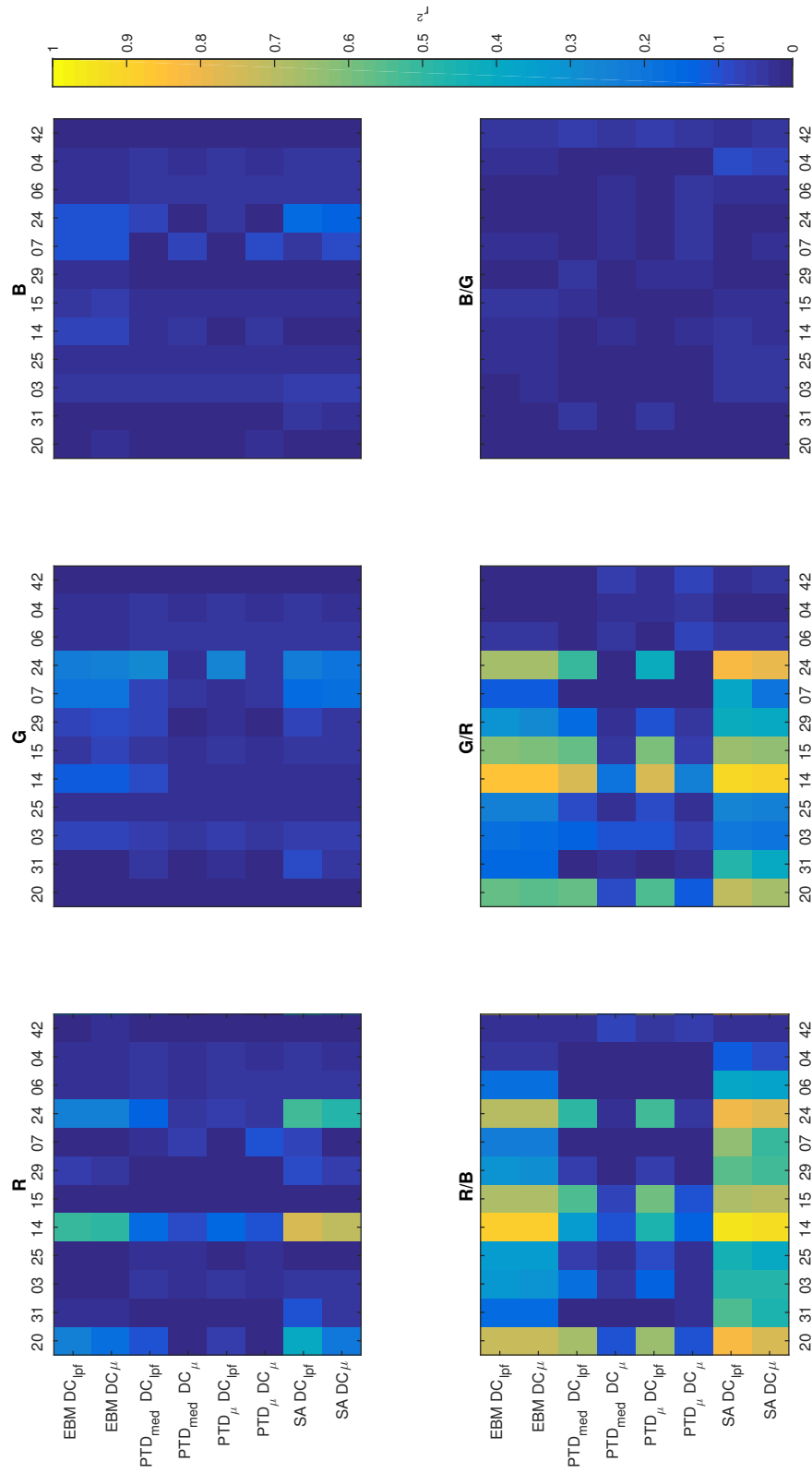


Figure 6.3: Coefficient of determination r^2 (colour-coded) found for each subject (x-axis) using the different combinations of algorithms (y-axis) and colour channels.

denominator have the highest correlation with the oxygen saturation. Of the different colour channel ratio combinations, the highest correlation is found for R/B. For HP3-24, SA_{lpf} yields a higher value of r^2 when applied to the R/G ratio. When investigated, it was found that, for cases in which the amplitude of the normalised red channel was in the numerator, there was a negative correlation with oxygen saturation, while a positive correlation was found for cases for which the amplitude of the normalised red channel was in the denominator. This dual relationship could indicate a model of the sort:

$$S \propto \log\left(\frac{C}{R}\right) \quad (6.9)$$

where S is the oxygen saturation, R is the red channel and C is another channel that is independent of oxygen saturation. This logarithmic model, which would explain the correlations found in the data, was tested using both green and blue as C , and the results are shown in Figure 6.4. The results indicate that a model of type $S \propto -\frac{R}{B}$ is equivalent in performance to a logarithmic model; however the latter also has the advantage that the normalised amplitude of the red will be most sensitive to noise when it is smallest (i.e. at high saturations) and the logarithm should attenuate the magnification of the noise. Throughout the remainder of this chapter, the logarithmic form will therefore be used for the definition of the ratio of ratios (RoR).

The ratio of ratio method based on normalised AC amplitudes gives much lower performance values of r^2 than the baseline method in estimating oxygen saturation (compare Figure 6.4 with Figure 5.7), and the reasons for this may be seen in Figure 6.5. Here, the logarithms of the ratio of ratios of blue to red AC normalised amplitudes are regressed for each subject onto the reference oxygen saturation using the linear regression from Equation 6.9. Compared to the equivalent plot for the baseline method in Chapter 5, Figure 6.5 shows estimates that are much less stable and oscillate significantly about the mean. Although the application of the motion SQI from Chapter 4 (SQI_L) does remove some of the larger changes, the estimates still remain very noisy.

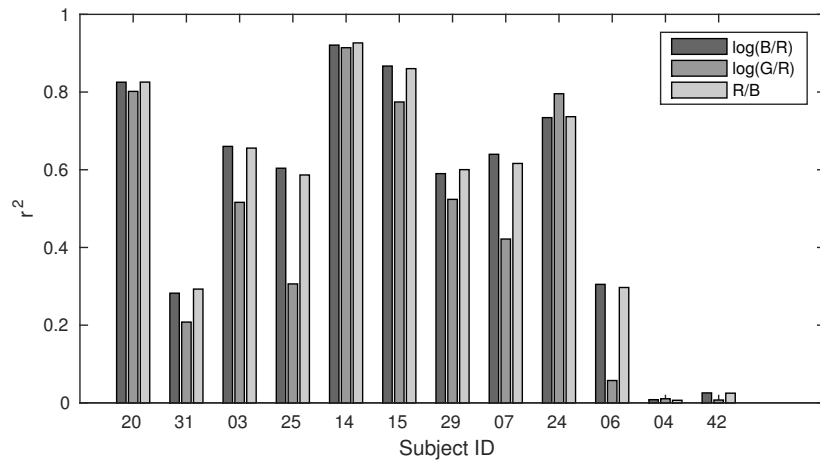


Figure 6.4: Relationship between oxygen saturation and colour channel ratios.

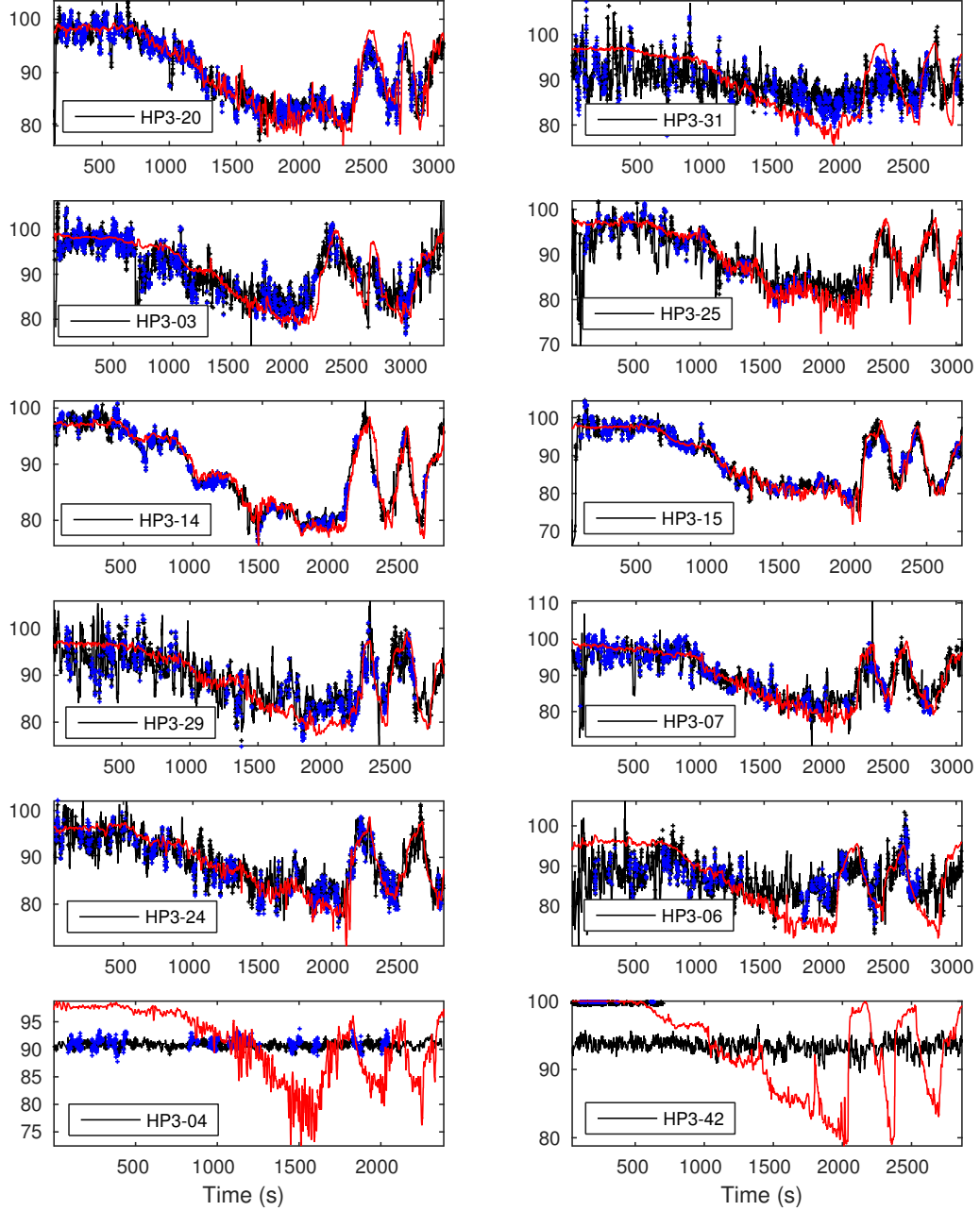


Figure 6.5: Regressed RoR camera estimates for all subjects using $\log(\frac{B}{R})$. The continuous red line indicates the reference oxygen saturation, taken as the mean of the two pulse oximeters, the continuous black line shows the regressed estimates with no SQI but matching HR, and the blue points are the regressed estimates using SQI_L .

6.3.3 Weighted Averaging

The ROIs are selected solely on the basis of the skin detection algorithm, which, working solely on the basis of colour, may include non-skin pixels such as hair or background that may be distorting the signal obtained by simple averaging. This hypothesis may be tested by introducing a weighting to the average that would emphasise ROIs having a clear periodic component with a frequency corresponding to the heart rate. This may be achieved by evaluating the heart rate SNR (SNR_{HR}) of each ROI and giving it a weighting corresponding to its strength relative to the total. The signal-to-noise ratio (SNR, or SNRdB for the decibel-based SNR) specific to f_{HR} , SNR_{HR} , can then be calculated using the SNR function SNR [44, 63, 157] on each channel of each ROI, where $\text{SNRdB} = 10 \log(\text{SNR})$ and SNR is defined as:

$$\text{SNR} = \text{SNR}\{x(t), \hat{f}\} = \frac{\int_a^b \mathcal{V}(\hat{f}) |\mathcal{F}\{x(t)\}|^2 df}{\int_a^b [1 - \mathcal{V}(\hat{f})] |\mathcal{F}\{x(t)\}|^2 df} \quad (6.10)$$

for an integral over frequencies f from $f = a$ to $f = b$, the Fourier transform of a detrended and appropriately filtered timeseries $x(t)$, $\mathcal{F}\{x(t)\}$, and an M-step function (shown in Figure 6.6) $\mathcal{V}(f)$ defined by the convolution:

$$\mathcal{V}(f) = \left[\sum_{m=1}^M \delta(f - m\hat{f}) \right] * \Pi(\pm f_h) \quad (6.11)$$

centred on the fundamental frequency of interest $\hat{f}$ (in this case f_{HR}) and the M harmonics of interest, with δ as the Dirac delta function and Π as the rect function of half-width f_h . The weighting w_i attributed to each ROI $i = 1, \dots, Q$ is then given by:

$$w_i = \frac{\text{SNR}_{HR}}{\sum_{i=1}^Q \text{SNR}_{HR}} \quad (6.12)$$

To evaluate the effect of the introduction of the weighted averaging, the coefficients of determination for the R-B RoR using the SA method with DC_{lpf} while normalising each

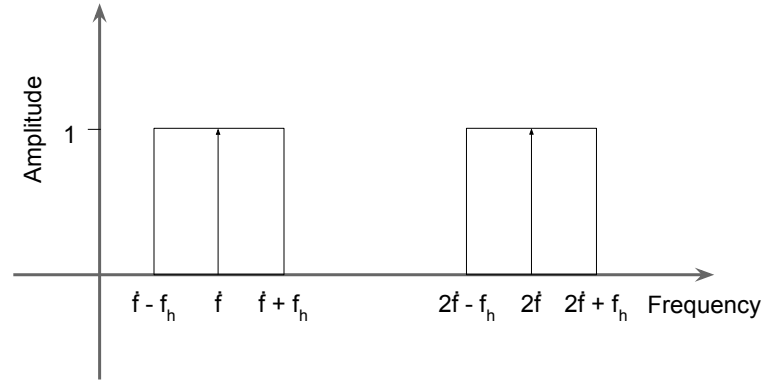


Figure 6.6: The M-step function $\mathcal{V}(f)$, shown for $M=2$.

ROI prior to averaging were found for weightings that used the SNR_{HR} computed using the green channel. The results are shown in Figure 6.7 and indicate that the use of the weighted average does bring about some improvement but not consistently so, as shown by the results for subjects HP3-25, HP3-29, and HP3-07. Because of the extremely low correlation between RoR and oxygen saturation, Type VI subjects will be ignored for now.

One reason for the uneven performance of weighted averaging may be the over-emphasis given to false signals: because of the greater weighting given to them, ROIs that are not due to colour change may be skewing the results. Such ROIs, as discussed in Chapter 2.5.3, are characterised by two different phenomena: movements deriving from heart pulsations and random movements that happen to be heart rate-synchronous. The former category can be further subdivided into local heart rate movements, due to local skin movements such as those reported by [238] when observing pulse-related movements on a wrist, which lead to local changes in superficial light reflectance; and global movements due to heart rate, that are instead the result of pressure changes along the entire vasculature leading to slight inelastic movements of the head, as studied in some detail by [8]. Within the proposed method, these signals appear as a result of reflectance changes caused by changes in the angle of reflection or caused by intensity changes as features of different intensities move in and out of an ROI (see Figure 6.8). Movements independent of heart

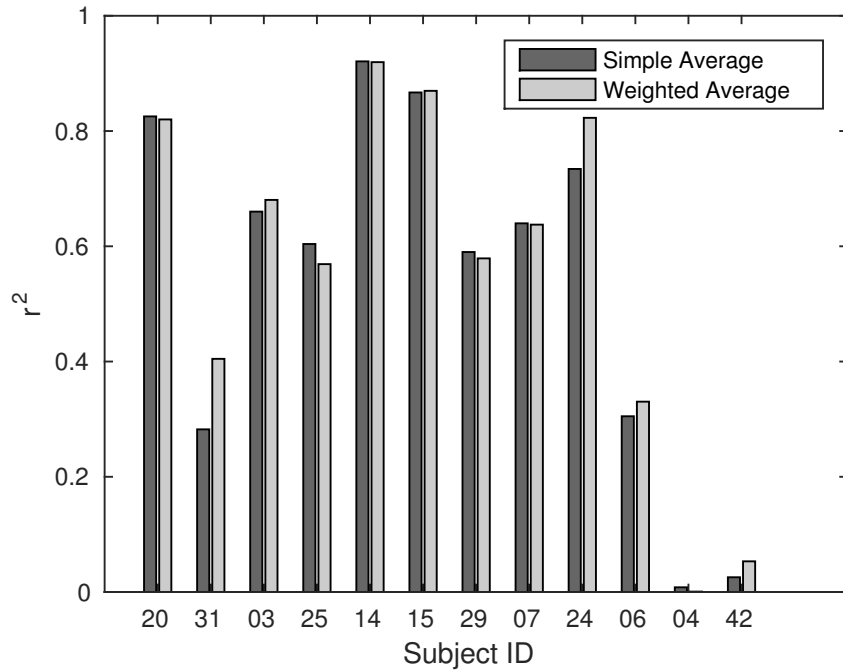


Figure 6.7: Coefficient of determination r^2 for results computed using the weighted average as compared to the simple average.

rate may instead be caused by (a) random cardiac-synchronous body movements, or (b) blinking, or lighting aliasing, occurring at a frequency within the expected heart rate range.

6.4 ROI Selection

6.4.1 Selection Criteria

A careful preselection of ROIs is therefore crucial to obtain an accurate AC-based estimate of oxygen saturation, more so than it was for estimates based on changes in baseline because of the nature of what is analysed. In order to distinguish between those ROIs showing a cardiac-synchronous periodicity due to colour and those that are instead due to some type of movement, prior domain knowledge of the form of the expected signal and the occurrence of the false signals may be used.

Firstly, if a signal is clearly detectable, its morphology may be used. A PPG signal will have a number of properties in common with other PPG signals: it should be approximately in phase with any other PPG signal in the vicinity, with a possible time delay introduced by the pulse transit time (PTT) between one area of the face and another. The pulse transit time itself is correlated with vessel tone and as such is the object of active research [43]; however to the author's knowledge no published work has investigated the pulse transit time between two regions on the face. An upper bound of $\approx 350\text{ms}$ can be placed with a reasonable degree of confidence on the maximal time difference between two areas of skin on the face based on the average pulse transit time between heart and finger [187]. In practice, the actual maximal time difference is likely to be significantly lower as the relative distance travelled by the pulse will be smaller (see: [43]). Pulsatile components due to global movements may instead be completely out of phase with the PPG depending on the intensity differential across the region of interest and the direction of movement, as shown schematically in Figure 6.8. However, as the motion is that of a non-rigid body, further phase distortions may be present depending on the time lag between the change in pressure and the movement itself.

Another means of distinguishing between a PPG signal and cardiac-synchronous motion is the morphology of the signal itself. This is demonstrated clearly in Figure 6.9, showing a signal from the forehead in blue with a characteristic long-rise time and short fall time (similar to other reflectance PPG waveforms shown in the literature) compared with a signal from the side of the throat in red that is instead due to motion and that has a very short rise time and an extended fall time. Interestingly, the latter is much more like the PPG waveform in contact pulse oximetry, even exhibiting a characteristic dicrotic notch.

The amplitude of the PPG signal will also be consistent across all skin regions, with signals with very high amplitudes likely to be caused by motion. This is a similar argument to that put forth by [44] for selecting pixels to be included in the estimation of

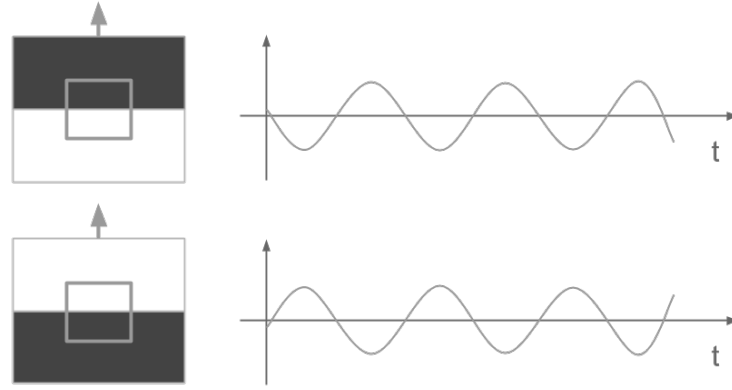


Figure 6.8: It is the relative gradient of the intensities surrounding the region of interest and the relative trajectory of the body with respect to the region of interest that determine the phase of the periodic signal over time if it is purely due to motion. Here, for two bodies (one with a black top and another with a white top) that move in an oscillatory fashion in the vertical direction only, and which at time $t = 0$ are moving upwards as indicated by the arrow, the mean values of the intensities of the regions of interest (shown in grey at the centre of the body) for the two bodies will be exactly in antiphase, although their frequency will be identical.

heart rate based on closeness in RGB space, leading to the rejection of ROIs that have very high AC variances compared to others.

Thus the phase and morphology of the signal may also be employed to select ROIs. Based on prior literature on the subject (see Chapter 2), the green AC signal is expected to have a greater amplitude than the red or blue signals, and the heart rate frequency should be the single most dominant frequency present in the green detrended signal. Finally, the baseline of the signal may be used to improve the skin segmentation by noting that skin pixels in ambient light will have a greater red component than other channels [96, 226], thus requiring the red baseline to be always greater than the green or blue.

The set of conditions an ROI i has to meet in order to be included into the weighted averaging of the signals may therefore be represented as:

$$L^i = L_s^i \cap L_d^i \cap L_v^i \cap L_m^i \cap L_p^i \cap L_a^i \cap L_f^i \quad (6.13)$$

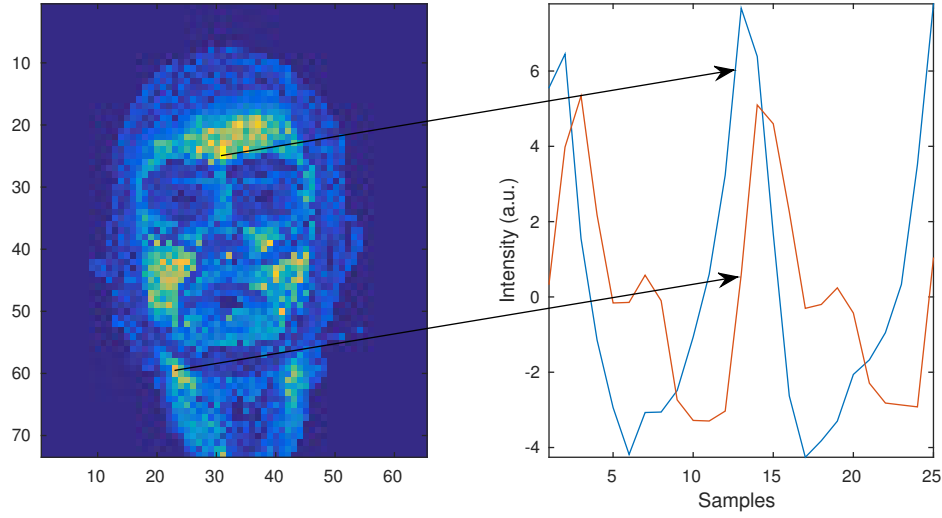


Figure 6.9: The green SNR_{HR} map on the left indicates regions of high SNR_{HR} both around the forehead and at the sides of the throat. The signal-averaged green channel forehead pulse (blue) is morphologically distinct from the signal-averaged throat pulse (red) - see right-hand plots.

Where

- L^i is the cumulative binary selection condition
- L_s^i is the skin/non-skin classification criterion discussed in Chapter 4.3.2
- L_d^i is the requirement for the red component of the baseline to be greater than the blue or green components
- L_v^i characterises those ROIs having an AC variance (i.e. after the removal of the baseline signal) within two standard deviations of the mean AC variance of all ROIs
- L_m^i is the requirement that, for each ROI, the signal-averaged waveform must have a greater rise time than fall time
- L_p^i is the requirement that, for each ROI, the signal-averaged waveform must have a delay of less than half a period with respect to the signal-averaged waveform of

all ROIs averaged together.

- L_a^i is the condition that for each ROI the amplitude of the signal-averaged green channel should be greater than both the amplitude of the signal-averaged red signal and the amplitude of the signal-averaged blue signal
- L_f^i is the condition that, for each ROI, the heart rate frequency should be the dominant frequency of the filtered green signal.

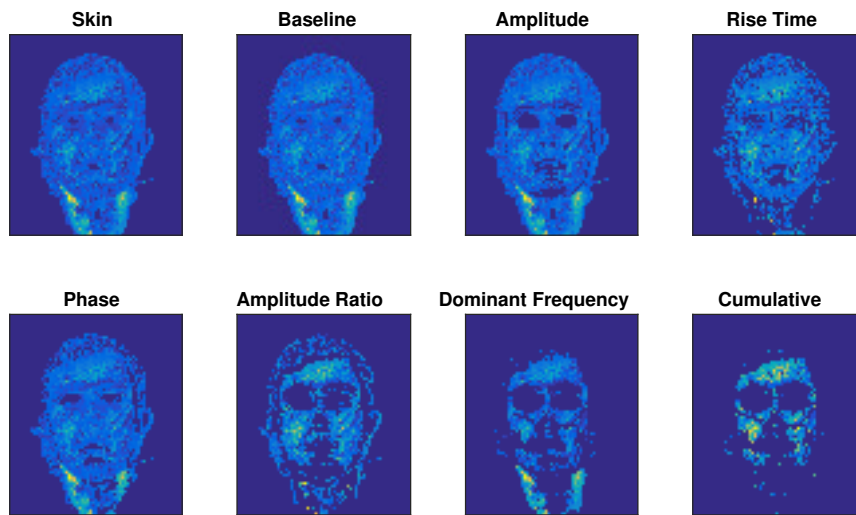


Figure 6.10: Each condition acts to remove some of the ROIs that had been considered as skin by the skin detection algorithm. From left to right, top row: L_s , L_d , L_v , L_m . From left to right, lower row: L_p , L_a , L_f , L .

The individual and cumulative effects of the conditions on the ROIs available over the course of one window are shown in Figure 6.10. As may be observed, while the skin detector has classified the hair as skin, this is removed by the condition of having the heart rate as the dominant frequency. While amplitude alone serves to eliminate the eye pixels (characterised by high-amplitude changes because of blinking), the amplitude ratio and morphology conditions both help remove the throat pixels associated with movement.

The application of the set of conditions given in Equation 6.13 to the 12 test subjects gives mixed results, with improvements in performance for many subjects but also decrease in performance for some subjects (e.g. HP3-07) as may be observed in Figure 6.11.

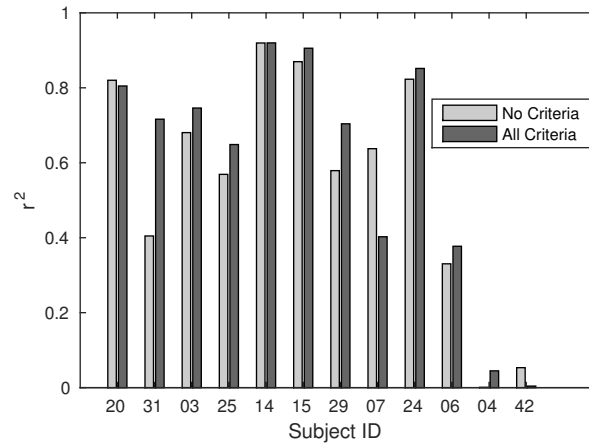


Figure 6.11: The coefficient of determination for each subject both using and not using the ROI selection criteria.

6.4.2 Selection Criteria Analysis

In order to evaluate the impact of each selection criterion on the results, the number of ROIs chosen by each criterion was investigated, reported as the average number of ROIs accepted by each method as a proportion of those selected by the skin detection criterion over the course of each session (Figure 6.12). In addition to determining which criteria were the most selective, the overlap between different methods was also investigated. This was done by examining the average number of ROIs coinciding for any two methods as a fraction of the number of ROIs selected by the most selective method (so as to avoid redundancy with Figure 6.12). The results are shown as a confusion matrix in Figure 6.13. Finally, the effect of every individual criterion and different combinations of criteria was evaluated by examining its effect on the coefficient of determination between the RoR and the reference oxygen saturation (Tables 6.4, 6.5, and 6.6).

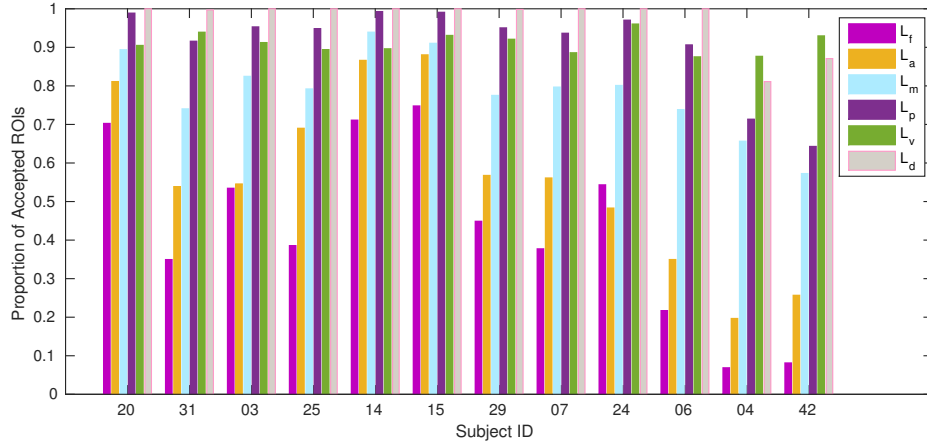


Figure 6.12: Average proportion of ROIs accepted by each selection criterion as a proportion of those selected by the skin/non-skin classification method.

The requirement for every ROI to generate an AC signal that is clearly related to heart rate is by far the most stringent condition imposed, shortly followed by the requirement for the green channel AC signal to have the largest amplitude, and that of the rise time being greater than the fall time. Furthermore, the amplitude and the dominant signal conditions have the least overlap, with over 25% of the ROIs selected by the dominant signal condition not corresponding to any selected by the amplitude condition, as may be observed from Figure 6.13. This indicates that these three conditions are the ones which have the most impact on the results. A comparative analysis was therefore run to determine whether reducing the stringency of the conditions imposed on the ROI selection would affect the correlation with the reference oxygen saturation.

Three different overall selection criteria were constructed, reducing the stringency of the selection criteria imposed:

$$L'^i = L_s^i \cap L_d^i \cap L_v^i \cap L_p^i \cap L_m^i \cap L_a^i \quad (6.14)$$

that removes the dominant frequency selection criterion, and:

Table 6.4: A comparison of the individual performances of the selection criteria in terms of the correlation r^2 of the resulting signal with the reference oxygen saturation.

	$\mathbf{L}_s$	$\mathbf{L}_d$	$\mathbf{L}_v$	$\mathbf{L}_p$	$\mathbf{L}_m$	$\mathbf{L}_a$	$\mathbf{L}_f$
HP3-20	0.82	0.82	0.82	0.82	0.81	0.83	0.78
HP3-31	0.40	0.40	0.49	0.39	0.53	0.76	0.32
HP3-03	0.68	0.68	0.71	0.67	0.63	0.77	0.61
HP3-25	0.57	0.57	0.62	0.57	0.58	0.68	0.55
HP3-14	0.92	0.92	0.92	0.92	0.92	0.92	0.91
HP3-15	0.87	0.87	0.88	0.87	0.86	0.91	0.88
HP3-29	0.58	0.58	0.65	0.54	0.45	0.72	0.49
HP3-07	0.64	0.64	0.68	0.65	0.59	0.72	0.50
HP3-24	0.82	0.82	0.84	0.82	0.84	0.86	0.86
HP3-06	0.33	0.33	0.40	0.33	0.30	0.38	0.26
HP3-04	0.00	0.00	0.02	0.00	0.00	0.03	0.00
HP3-42	0.05	0.06	0.05	0.02	0.03	0.01	0.00

$$L''^i = L_s^i \cap L_d^i \cap L_v^i \cap L_p^i \cap L_m^i \quad (6.15)$$

which removes both the dominant frequency selection criterion and the amplitude selection criterion, and:

$$L'''^i = L_s^i \cap L_d^i \cap L_v^i \cap L_p^i \quad (6.16)$$

that further removes the morphology criterion (rise time). The performances of these three methods against the initial overall criterion L are shown in Table 6.5. The results demonstrate that whereas L_f adds contributes little information to the estimation of oxygen saturation, L_a is often able to reduce the effect of false ROIs, giving L' a performance gain with respect to L'' and L''' .

Table 6.5 would also seem to indicate that removing L_f and L_m improve performance, whereas, in agreement with Table 6.4, removing L_a decreases performance. Furthermore, the results using multiple ROI selection criteria are still inferior to simply using L_a . A new combination of criteria was thus tested to emphasise the role of L_a , whereby:

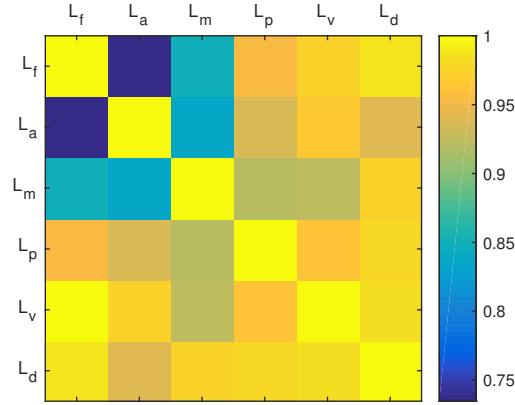


Figure 6.13: Confusion matrix indicating what percentage of ROIs were selected by any two methods on average across all subjects. Bright yellow indicates the most overlap and dark blue indicates the least overlap.

$$L^{i+} = L_s^i \cap L_d^i \cap L_v^i \cap L_p^i \cap L_a^i \quad (6.17)$$

$$L^{i++} = L_s^i \cap L_d^i \cap L_v^i \cap L_a^i \quad (6.18)$$

$$L^{i+++} = L_s^i \cap L_d^i \cap L_a^i \quad (6.19)$$

The results using these selection criteria are shown in Table 6.6, showing a modest improvement for $\mathbf{L}^{++}$ compared to all other selection criteria. The final results, using the $\mathbf{L}^{++}$ criterion on all 12 subjects is instead shown in Figure 6.14.

In conclusion, the single most important criterion to determine that the signal is derived from a colour change is the condition for the green AC amplitude to be greater than the blue and red AC amplitudes, with the L_d and L_v conditions occasionally yielding a very marginal improvement. This is likely to be due to the fact that, as shown in Figures 6.12 and 6.13, the L_a condition is both fairly stringent and has a good overlap with all other criteria.

Table 6.5: Comparing the performance of different selection criteria, from $\mathbf{L}$ (most ROI conditions) to $\mathbf{L}'''$ (least ROI conditions).

	$\mathbf{L}$	$\mathbf{L}'$	$\mathbf{L}''$	$\mathbf{L}'''$
HP3-20	0.80	0.82	0.81	0.82
HP3-31	0.72	0.74	0.59	0.48
HP3-03	0.75	0.75	0.67	0.70
HP3-25	0.65	0.68	0.62	0.62
HP3-14	0.92	0.92	0.92	0.92
HP3-15	0.91	0.91	0.88	0.89
HP3-29	0.70	0.73	0.53	0.62
HP3-07	0.40	0.69	0.66	0.70
HP3-24	0.85	0.85	0.85	0.84
HP3-06	0.38	0.39	0.40	0.40
HP3-04	0.04	0.05	0.01	0.01
HP3-42	0.00	0.00	0.02	0.05

Table 6.6: Performance of amplitude-based criteria

	$\mathbf{L}^+$	$\mathbf{L}^{++}$	$\mathbf{L}^{+++}$
HP3-20	0.83	0.83	0.83
HP3-31	0.77	0.78	0.77
HP3-03	0.76	0.78	0.77
HP3-25	0.69	0.70	0.68
HP3-14	0.92	0.92	0.92
HP3-15	0.91	0.91	0.91
HP3-07	0.68	0.73	0.72
HP3-29	0.73	0.73	0.72
HP3-24	0.86	0.86	0.86
HP3-06	0.40	0.40	0.38
HP3-04	0.02	0.01	0.02
HP3-42	0.01	0.01	0.01

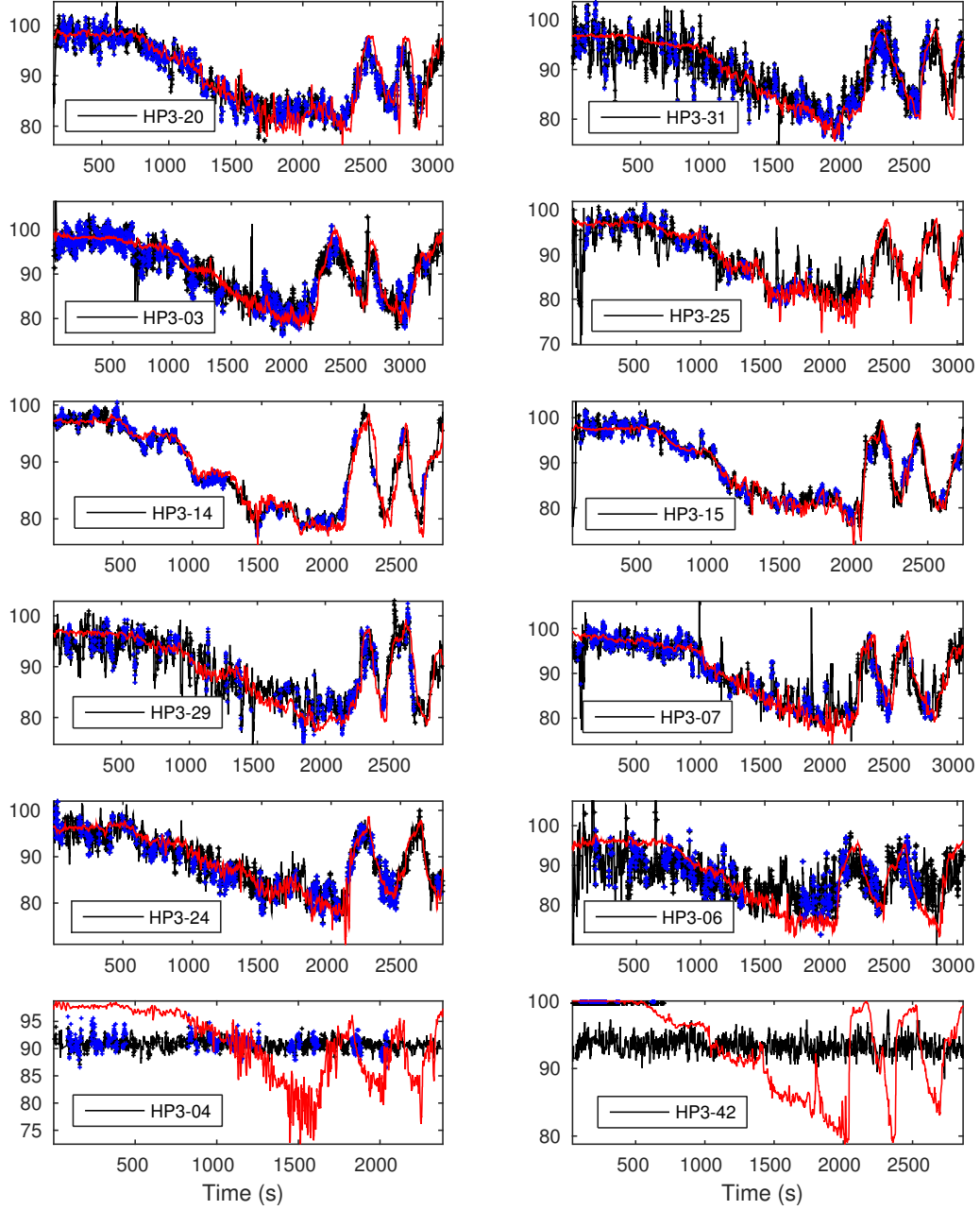


Figure 6.14: Regressed RoR camera estimates for all subjects using $\log(\frac{B}{R})$ and the $\mathbf{L}^{++}$ criterion. The continuous red line indicates the reference oxygen saturation, taken as the mean of the two pulse oximeters, the continuous black line shows the regressed estimates with no SQI but matching HR, and the blue points are the regressed estimates using SQI_L . A marked improvement is visible, particularly for HP3-31, with respect to Figure 6.5.

6.5 Regression Models

If the linear model used to regress the RoR estimates onto the reference oxygen saturation is considered, whereby:

$$S = \lambda_0 + \lambda_1 \log(R) \quad (6.20)$$

where R is the ratio of the blue normalised AC to the red normalised AC, calculated using the L^{++} criterion, the dependence of λ_1 on the skin's initial conditions (at the start of each hypoxia session – when the subject was breathing air) was explored for all subjects excluding HP3-06, HP3-04, HP3-42. As may be observed in Figure 6.15, although no relationship is particularly visible between the Fitzpatrick scale and λ_1 , a weak linear relationship was found between λ_1 and the RoR for red and blue and for green and blue at the beginning of the session, calculated by taking the average ratio of normalised amplitudes during the time when the subject was breathing air.

This relationship would indicate the possibility of quantifying to some degree of accuracy the changes in oxygen saturation so as to move from a method that tracks relative changes to one that tracks absolute changes in oxygen saturation under the conditions of the Hypoxia dataset. The relationship between λ_0 and the starting conditions was also evaluated in the same manner and a stronger relationship was found both for the red and blue log-ratio and the green and blue log-ratio (Figure 6.16).

These results show that the linear relationship between the camera-derived estimates (using the RoR method) and oxygen saturation may be expressed as a function of the starting conditions and further equations, as explored in the next Chapter. There we will examine the implications of such a relationship and attempt to derive the values of oxygen saturation based on values derived for these equations.

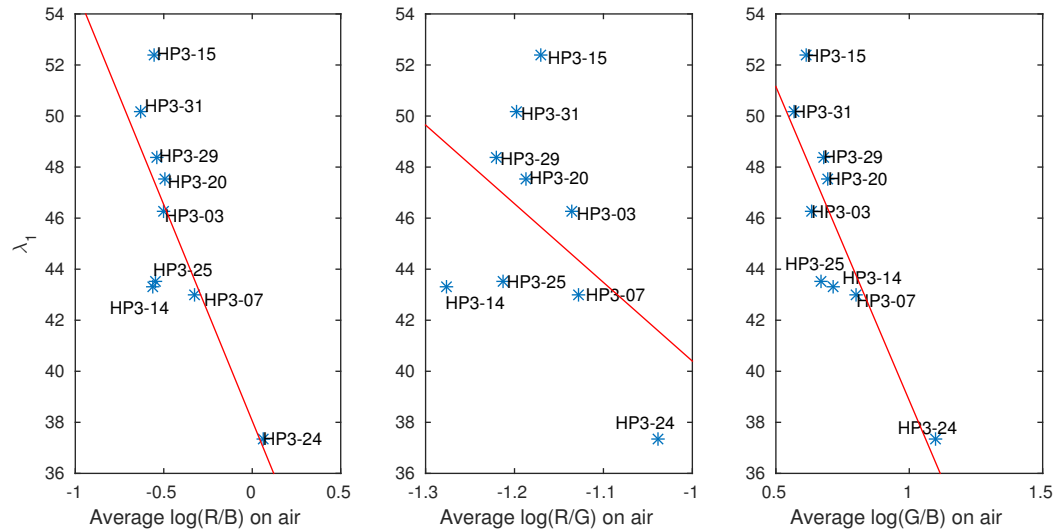


Figure 6.15: Relationship between the RoR initial conditions (on air at the beginning of each Hypoxia session) and the linear regression parameter λ_1 , showing the line of best fit in red.

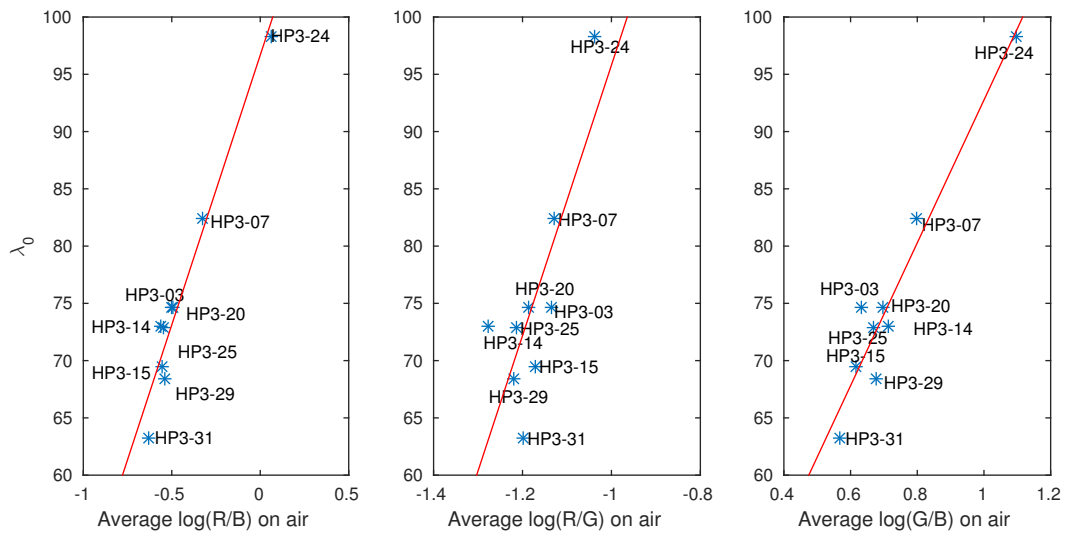


Figure 6.16: Relationship between the RoR initial conditions (on air at the beginning of each Hypoxia session) and the linear regression parameter λ_0 , showing the line of best fit in red.

6.6 Chapter Conclusion

In this chapter, the ratio of ratios (RoR) method, was derived from first principles and applied to the data from the 12 test subjects of the Hypoxia study. It was found that the ratio of the normalised blue amplitude to the normalised red amplitude gave the best results within these conditions, although estimates were again unusable for Fitzpatrick Type VI subjects. The method is more noisy than the baseline method derived in Chapter 5, and a number of selection criteria were put forward and tested to improve the signal-to-noise ratio. A link was found between the starting conditions of the subject (i.e. the starting colour) and the parameters used for the linear regression onto the oxygen saturation, hinting at the possibility of an absolute measurement of oxygen saturation using video cameras.

Chapter 7

Thesis Results and Discussion

7.1 Introduction

This chapter evaluates further the baseline-derived method and the ratio-of-ratios (RoR) method that have been developed in Chapters 5 and 6 respectively for tracking changes in oxygen saturation. The remaining 34 Hypoxia sessions are used to evaluate the methods in Section 7.2, using the coefficient of determination (r^2) and the mean absolute error between the reference oxygen saturation and the linearly regressed baseline and RoR methods as the key metric (Section 7.2.1). The effect of applying the ROI selection methods used in the RoR estimates are investigated in Section 7.2.2, while the differences between the baseline and RoR methods are analysed in Section 7.2.3. Section 7.2.4 investigates the root causes of poor performance on particular subjects.

Finally, an initial attempt at an absolute estimation of oxygen saturation from the RoR data available is made in Section 7.3. The linear fit between RoR and oxygen saturation is parameterised and modelled as dependent on the starting values of RoR on air. To demonstrate that an absolute derivation of oxygen saturation is possible given similar ambient conditions, the model parameters are estimated using half of the subjects and tested on the other half. The linear fit is only sought for Fitzpatrick Types I-V.

7.2 Hypoxia Results and Analysis

7.2.1 Results Summary

The performance of the baseline method and RoR method with the $L++$ condition on all the participants (12 test subjects and the remaining 34 subjects) is summarised in Table 7.1. The effect of the movement SQI described in Chapter 4.3.3, is also investigated, with the right-hand side of the table reporting the percentage of windows that are accepted when using SQI_L and the corresponding MAE and r^2 values for both the baseline (Chapter 5) and RoR (Chapter 6) methods. As already reported in chapters 5 and 6, both methods fail on Type VI subjects, with no correlation ($r^2 \leq 0.06$) found between the RoR method and the reference oxygen saturation, and a low correlation ($r^2 \leq 0.51$) found between the baseline method and the reference oxygen saturation.

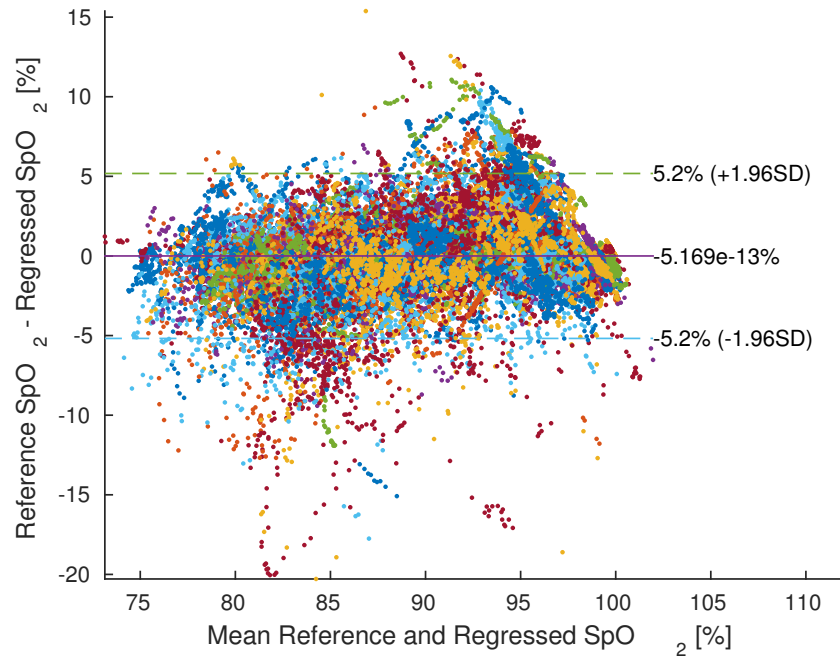
The overall performance of the methods, on all subjects excluding Type VIs, is also analysed using Bland-Altman plots, as shown in Figure 7.1. As in previous chapters, the Bland-Altman plot is generated by comparing the reference oxygen saturation with a linear regression of the method's estimates onto the reference estimates themselves, and should therefore represent the ideal linear mapping between the two. A consequence of this approach is that the average difference between the two is by definition zero, so the distribution of points about this mean is the most relevant metric, and shows the variance of the method's results. As may be seen from the plots, the baseline method yields results with low variance, whereas the RoR method varies significantly around the value, giving normal standard deviations (1.96σ) of 5.2% and 5.9% respectively. A diagonal cut-off is present at the right-hand end of both plots. This is because the reference oxygen saturation has an upper limit at 100%, whereas camera-derived estimates have no upper bound.

As a means of comparison, the two reference values (for the pulse oximeters placed on the subjects' left – Ref_L – and right – Ref_R – index fingers respectively) are also plotted

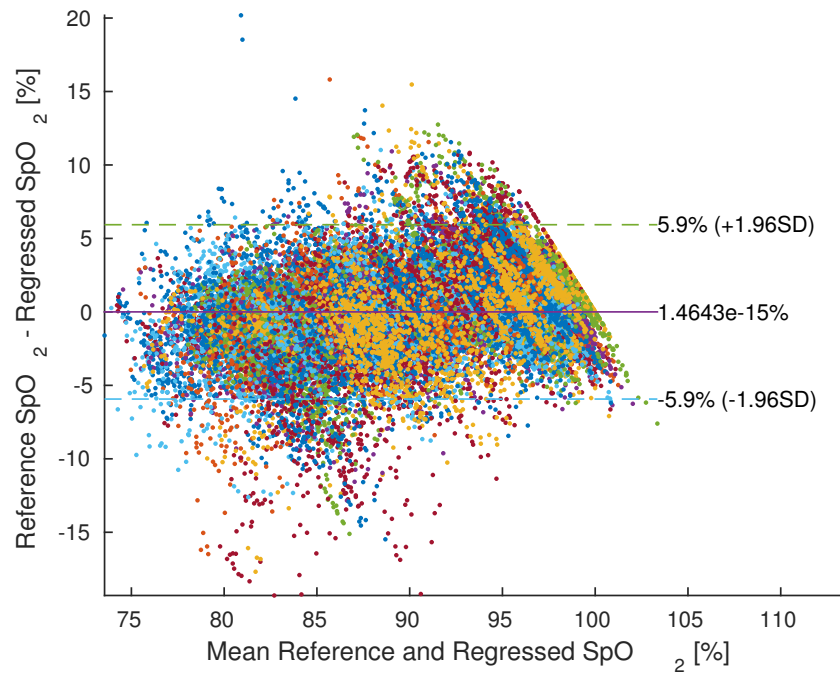
on a Bland-Altman plot as shown in Figure 7.2. Unlike Figure 7.1, the plot shows a small bias (of 0.45%) between the two reference SpO₂ values; however, the spread is much smaller, with a standard normal deviation of 2.35%. This is not surprising, as both pulse oximeters are from the same manufacturer, they were both initialised in the same way, and the measurements made with them were from similar physiological sites. The difference between these two pulse oximeters is within the ranges described in the literature [12, 145], where a mean difference of up to 4% has been reported.

Subject (HP3-)	Fitzpatrick Skin Type	NO SQI				SQI _L				
		Baseline r^2	Baseline MAE (%)	RoR r^2	RoR MAE (%)	Valid Windows (%)	Baseline r^2	Baseline MAE (%)	RoR r^2	RoR MAE (%)
22	I	0.83	2.17	0.68	3.07	49	0.82	2.23	0.65	3.18
23	I	0.85	1.31	0.70	1.88	40	0.86	0.99	0.76	1.52
40	I	0.64	1.58	0.42	2.14	56	0.77	1.21	0.51	1.95
46	I	0.90	1.25	0.76	1.96	64	0.90	1.21	0.78	1.93
47	I	0.93	0.63	0.76	1.07	36	0.93	0.65	0.73	1.18
<i>20</i>	<i>I</i>	<i>0.89</i>	<i>1.48</i>	<i>0.79</i>	<i>2.21</i>	<i>74</i>	<i>0.88</i>	<i>1.51</i>	<i>0.83</i>	<i>2.13</i>
<i>31</i>	<i>I</i>	<i>0.93</i>	<i>1.16</i>	<i>0.77</i>	<i>2.32</i>	<i>63</i>	<i>0.95</i>	<i>1.04</i>	<i>0.78</i>	<i>2.16</i>
08	II	0.91	1.36	0.81	2.06	51	0.93	1.18	0.85	1.85
09	II	0.95	1.23	0.86	2.09	44	0.95	1.24	0.87	2.12
10	II	0.87	1.95	0.88	1.90	52	0.91	1.63	0.91	1.73
12	II	0.76	2.61	0.73	2.74	53	0.83	2.15	0.81	2.25
21	II	0.49	3.60	0.52	3.53	42	0.87	1.55	0.56	3.15
32	II	0.88	1.59	0.66	2.72	54	0.89	1.50	0.73	2.35
34	II	0.71	2.25	0.67	2.29	49	0.79	1.77	0.75	2.03
<i>03</i>	<i>II</i>	<i>0.89</i>	<i>1.58</i>	<i>0.72</i>	<i>2.39</i>	<i>49</i>	<i>0.91</i>	<i>1.41</i>	<i>0.78</i>	<i>2.14</i>
<i>25</i>	<i>II</i>	<i>0.91</i>	<i>1.50</i>	<i>0.75</i>	<i>2.46</i>	<i>31</i>	<i>0.96</i>	<i>1.05</i>	<i>0.70</i>	<i>2.23</i>
05	III	0.89	1.39	0.81	2.00	48	0.90	1.45	0.84	1.97
11	III	0.27	3.95	0.38	3.90	56	0.32	3.69	0.37	3.89
16	III	0.79	1.48	0.68	2.10	42	0.91	0.93	0.78	1.67
19	III	0.69	2.34	0.75	2.10	60	0.73	2.26	0.83	1.84
28	III	0.18	4.87	0.14	4.86	41	0.40	3.70	0.55	2.84
30	III	0.94	1.24	0.91	1.55	48	0.93	1.32	0.92	1.48
<i>14</i>	<i>III</i>	<i>0.94</i>	<i>1.17</i>	<i>0.91</i>	<i>1.45</i>	<i>45</i>	<i>0.95</i>	<i>1.03</i>	<i>0.92</i>	<i>1.57</i>
<i>15</i>	<i>III</i>	<i>0.96</i>	<i>1.07</i>	<i>0.92</i>	<i>1.52</i>	<i>53</i>	<i>0.95</i>	<i>1.20</i>	<i>0.91</i>	<i>1.41</i>
13	IV	0.94	1.17	0.94	1.28	53	0.95	1.21	0.94	1.29
26	IV	0.49	3.75	0.74	2.64	63	0.58	3.50	0.77	2.57
27	IV	0.56	3.01	0.46	3.20	72	0.61	2.63	0.56	2.70
33	IV	0.87	1.75	0.84	2.09	74	0.89	1.62	0.84	2.07
36	IV	0.86	1.53	0.80	1.90	53	0.90	1.32	0.84	1.73
37	IV	0.90	1.61	0.92	1.27	59	0.90	1.55	0.92	1.25
38	IV	0.72	2.25	0.57	2.90	44	0.84	1.66	0.67	2.46
39	IV	0.83	1.78	0.58	3.35	48	0.90	1.23	0.67	3.02
<i>29</i>	<i>IV</i>	<i>0.87</i>	<i>1.69</i>	<i>0.72</i>	<i>2.60</i>	<i>38</i>	<i>0.90</i>	<i>1.68</i>	<i>0.73</i>	<i>2.53</i>
<i>07</i>	<i>IV</i>	<i>0.84</i>	<i>2.06</i>	<i>0.74</i>	<i>2.41</i>	<i>44</i>	<i>0.88</i>	<i>1.60</i>	<i>0.73</i>	<i>2.44</i>
41	V	0.78	2.25	0.71	2.46	51	0.75	2.18	0.73	2.22
43	V	0.35	3.86	0.19	4.31	58	0.54	3.19	0.40	3.63
44	V	0.73	2.27	0.53	3.06	67	0.69	2.33	0.51	2.91
48	V	0.03	2.19	0.21	1.91	11	0.01	2.56	0.53	1.57
50	V	0.82	1.50	0.53	2.61	52	0.84	1.51	0.58	2.47
<i>24</i>	<i>V</i>	<i>0.93</i>	<i>1.17</i>	<i>0.84</i>	<i>1.84</i>	<i>41</i>	<i>0.92</i>	<i>1.17</i>	<i>0.86</i>	<i>1.81</i>
<i>06</i>	<i>V</i>	<i>0.82</i>	<i>2.43</i>	<i>0.38</i>	<i>4.77</i>	<i>49</i>	<i>0.86</i>	<i>2.29</i>	<i>0.40</i>	<i>4.26</i>
35	VI	0.00	5.10	0.03	5.05	20	0.21	2.61	0.05	3.45
45	VI	0.14	4.67	0.01	4.99	36	0.49	2.95	0.03	4.74
49	VI	0.37	3.73	0.01	4.87	54	0.42	3.14	0.00	4.31
<i>04</i>	<i>VI</i>	<i>0.03</i>	<i>4.89</i>	<i>0.00</i>	<i>4.98</i>	<i>49</i>	<i>0.13</i>	<i>4.01</i>	<i>0.01</i>	<i>4.38</i>
<i>42</i>	<i>VI</i>	<i>0.42</i>	<i>3.52</i>	<i>0.01</i>	<i>5.20</i>	<i>45</i>	<i>0.87</i>	<i>1.53</i>	<i>0.01</i>	<i>4.74</i>

Table 7.1: Results using the baseline method (developed in Chapter 5) and the RoR method with the $L++$ condition (developed in Chapter 6) on the remaining Hypoxia sessions, listed by Fitzpatrick type. The 12 test subjects are also present in italics.



(a)



(b)

Figure 7.1: Bland-Altman plot for the estimated oxygen saturations using the baseline method (7.1b) and RoR (7.1a) method with the $L + +\text{criterion}$. Each estimate was first regressed linearly onto the reference oxygen saturation. Results are reported for $\text{SQI}_L = 1$ and each colour represents a different subject, excluding Type VI subjects.

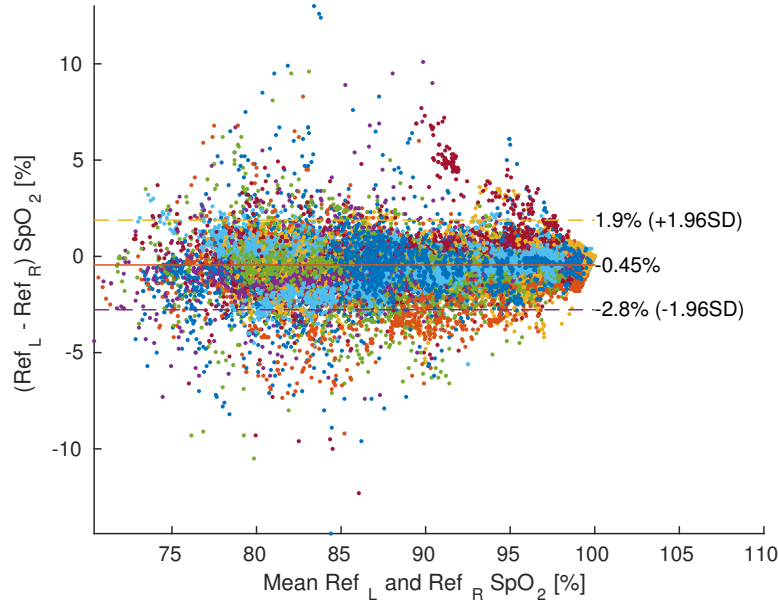


Figure 7.2: Bland-Altman plot used to compare the two reference SpO_2 values for the two pulse oximeters. Subjects of Type VI are excluded. Each colour represents a different subject.

7.2.2 Effect of ROI Selection on Baseline Estimates

In Chapter 6, the need for methods of ROI selection arose as a result of the increased level of accuracy required to measure minute changes in signal amplitude, leading to a consistent improvement in the performance of the RoR method. Table 7.2 shows the effect of applying the most important selection criteria of Chapter 6 to the baseline estimates derived from the method of Chapter 5. It might be expected that the use of these conditions should improve the baseline estimates but this is not the case.

Although the addition of the basic L_d condition (that takes advantage of the fact that skin pixels are expected to be more red than other pixels in the image) performs as well as the original algorithm of Chapter 6 (and better in the case of HP3-26), the addition of further criteria does not lead to any further improvements. This is different to the case of the RoR algorithm, in which the addition of the $L++$ condition improved the r^2 values. As the baseline method is effectively measuring an overall average, both in space

and time, outlier elimination on the average ROI-cloud in RGB-space may not be as important as in the calculation of the RoR estimates if there are many ROIs from which to draw. If there are only a few available ROIs, the centroid of such a sparse ROI-cloud will move more as this will lead to less accurate estimates by the baseline method.

Table 7.2: Effect of the RoR selection criteria on the r^2 values obtained with the baseline estimates

Subject (HP3-)	Fitzpatrick Skin Type		L_d	L_a	L^{+++}	L^{++}
22	I	0.83	0.83	0.80	0.80	0.79
23	I	0.85	0.85	0.83	0.83	0.82
40	I	0.64	0.64	0.66	0.66	0.65
46	I	0.90	0.90	0.90	0.90	0.90
47	I	0.93	0.93	0.93	0.93	0.93
20	I	<i>0.89</i>	<i>0.89</i>	<i>0.89</i>	<i>0.89</i>	<i>0.89</i>
31	I	0.93	0.93	<i>0.91</i>	<i>0.91</i>	<i>0.90</i>
08	II	0.91	0.91	0.91	0.91	0.91
09	II	0.95	0.95	0.95	0.95	0.95
10	II	0.87	0.87	0.85	0.85	0.84
12	II	0.76	0.76	0.42	0.42	0.38
21	II	0.49	0.49	0.55	0.55	0.54
32	II	0.88	0.88	0.88	0.88	0.88
34	II	0.71	0.71	0.71	0.71	0.71
03	II	0.89	0.89	<i>0.86</i>	<i>0.86</i>	<i>0.86</i>
25	II	0.91	0.91	<i>0.88</i>	<i>0.88</i>	<i>0.87</i>
05	III	0.89	0.89	0.88	0.88	0.87
11	III	0.27	0.11	0.23	0.10	0.11
16	III	0.79	0.80	0.77	0.77	0.77
19	III	0.69	0.71	0.50	0.50	0.50
28	III	0.18	0.18	0.07	0.07	0.06
30	III	0.94	0.94	0.92	0.92	0.92
14	III	0.94	0.94	<i>0.93</i>	<i>0.93</i>	<i>0.94</i>
15	III	0.96	0.96	<i>0.95</i>	<i>0.95</i>	<i>0.95</i>
13	IV	0.94	0.94	0.93	0.93	0.93
26	IV	0.49	0.89	0.63	0.75	0.75
27	IV	0.56	0.58	0.42	0.44	0.44
33	IV	0.87	0.87	0.85	0.85	0.85
36	IV	0.86	0.86	0.83	0.83	0.82
37	IV	0.90	0.90	0.86	0.86	0.85
38	IV	0.72	0.81	0.45	0.39	0.40
39	IV	0.83	0.83	0.76	0.76	0.75
29	IV	0.87	0.87	<i>0.77</i>	<i>0.77</i>	<i>0.76</i>
07	IV	0.84	0.84	<i>0.69</i>	<i>0.69</i>	<i>0.69</i>
41	V	0.78	0.78	0.67	0.67	0.66
43	V	0.35	0.35	0.43	0.43	0.42
44	V	0.73	0.73	0.62	0.62	0.62
48	V	0.03	0.26	0.00	0.04	0.03
50	V	0.82	0.82	0.64	0.64	0.63
24	V	0.93	0.93	<i>0.91</i>	<i>0.91</i>	<i>0.91</i>
06	V	0.82	0.82	<i>0.52</i>	<i>0.53</i>	<i>0.51</i>
35	VI	0.00	0.00	0.01	0.01	0.00
45	VI	0.14	0.14	0.46	0.46	0.36
49	VI	0.37	0.17	0.05	0.08	0.05
04	VI	0.03	<i>0.01</i>	<i>0.01</i>	<i>0.01</i>	<i>0.01</i>
42	VI	0.42	<i>0.34</i>	<i>0.15</i>	<i>0.09</i>	<i>0.10</i>

7.2.3 Differences Between Baseline and RoR Estimates

Both Table 7.1 and the Bland-Altman plots in Figure 7.1 show a marked difference between the results from the baseline method and from the RoR method. As already mentioned, the main reason for this difference is the high variability of the RoR estimates compared to the relative stability of the baseline estimates. As shown in the preceding Section, the large number of regions of interest available to the baseline method will smooth the estimate, whereas the variance of the RoR method depends on both the number and the quality of the ROIs it selects, such that a careful balance must exist between allowing as many ROIs as possible and restricting them to those which met a set of conditions. In addition to this, the baseline method is essentially a temporal- and spatial-average over all available pixels over a 15-second window, and as a result of this any erroneous changes will be smoothed. Although an averaging step does occur in the derivation of the RoR estimate, this is both dependent on the accuracy of peak selection and on the interference caused by changes in breathing rate, changes in light intensity, or movement, so that slight changes in these can have a large impact on the estimation of the normalised wave amplitude.

The differences in variance between the two methods can therefore be ascribed to the stability of the methods and the way the ROIs are used rather than to the type of oxygen saturation sampled by each method (mixed, venous, or arterial oxygen saturation). After all, for very stable subjects such as HP3-37, the baseline MAE is in fact larger than the RoR MAE. This in turn leads to the question as to whether, if the variance of the two methods were the same, the accuracy (or bias) would differ. A good example of a case in which the accuracy does differ is Figure 7.3, showing the linearly regressed traces of the baseline and RoR estimates plotted on top of the reference oxygen saturation for Subject HP3-13. As may be observed, the baseline estimates appear to follow the fast, double resaturation and desaturation stage poorly, such that estimates that the elevation angle ϕ for an oxygen saturation of 95% after a rapid resaturation is different from the

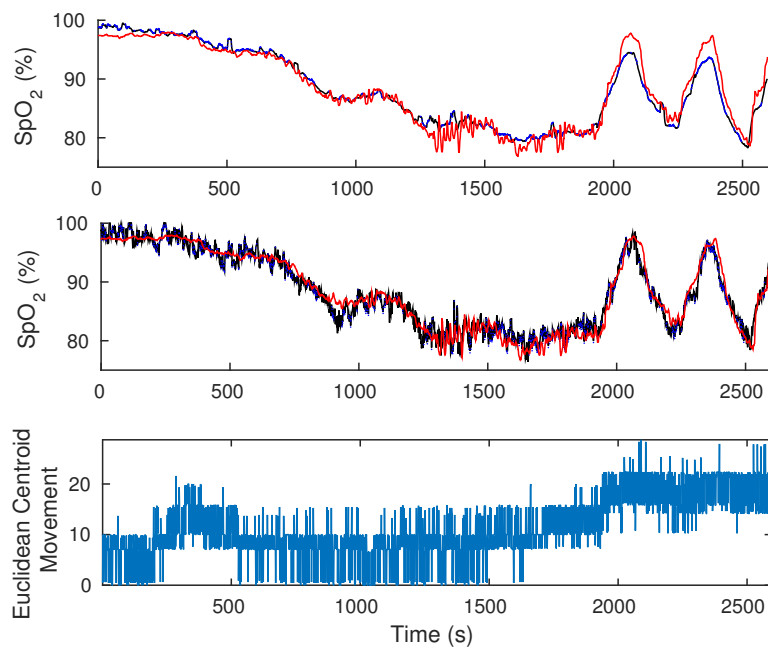


Figure 7.3: A comparison of the baseline estimates (top) and RoR estimates (middle), each linearly regressed onto the reference SpO_2 (in red in each plot). Baseline and RoR estimates are shown in black (no SQI) and blue (SQI_L). The lower plot shows the movement during that period, calculated by estimating the displacement of the facial bounding box centroid from the beginning of the session.

elevation angle ϕ for an oxygen saturation of 95% at the beginning of the session when the subject was breathing air. On the other hand the RoR method continues to track the oxygen saturation with a similar accuracy throughout the session. Although this phenomenon does not occur in every case, it indicates that the accuracy of the baseline estimate changes over time – i.e. a time-dependent bias is introduced.

This may be due to a number of reasons. One possibility is that the baseline method is strongly affected by movement of the subject, such that the baseline-derived estimates need to be re-initialised after movement. Although this may be true for larger movements, as indeed may be seen by the large step-changes in estimates that are rejected by SQL_L in Figure 5.9, this is unlikely to be the cause in this case as there is very little movement according to the lower plot in Figure 7.3, and that which does occur does not lead to a step change in the baseline estimates. A second possibility is that the baseline method used to evaluate the oxygen saturation, which uses a linear regression with only one variable (ϕ), is not able to capture the full extent of the blood-oxygen change. In Chapter 5, the hypothesis that the change in the azimuthal angle θ between red and green could be describing changes in perfusion was formulated. If the hypothesis were correct, there might be an effect due to the change in blood volume on the relationship between the polar angle ϕ and the oxygen saturation. A third possibility is that the measurement is in fact accurate, but that the quantity measured by the baseline estimate is different to that measured by the pulse oximeter. As the baseline estimates are more akin to an oximetry measurement (i.e. measuring oxygen saturation in both venous blood and arterial blood) than pulse oximetry, this could mean that the venous blood is reperfusing differently to the arterial blood, clamping the baseline-measured oxygen saturation at a lower level.

Although the last two hypotheses are more likely than the first, they unfortunately cannot be tested with the data currently available. Regardless of this, the analysis of the differences between the baseline and RoR estimates underlines one important distinction

between the two: although the former are more precise they are also less accurate than the latter. By extension, the RoR estimates may be said to be a better estimator of the peripheral oxygen saturation than the baseline estimates.

7.2.4 Causes of Error in Estimates

Although a baseline estimate having a more pronounced bias than an RoR estimate may yield a worse MAE after regression, the coefficient of determination itself is likely to be less affected. Nevertheless, there are many cases in Table 7.1 in which both the baseline method and the RoR method have high values of MAE and low values of r^2 . In this section we attempt to classify these errors and speculate on what causes them.

Skin Type

Skin melanin content was originally measured indirectly through the subjective Fitzpatrick Skin Type assessment and it became apparent during the development of the baseline and RoR methods that both algorithms failed altogether on subjects of Type VI. This finding was confirmed by the results shown for Type VI subjects in Table 7.1. Results are worse for the majority of Type V subjects, indicating that the boundary for algorithm failure occurs within the Type V class.

Movement

The impact of movement if no movement SQI were applied is shown in Figure 7.4 (subject HP3-28). Here the baseline and RoR regressed estimates, both overlaid onto the SpO₂ reference values, are shown together with the movement metric of choice: the Euclidean distance moved by the face mask centroid from its initial position over time. This is perhaps an extreme case, with large and frequent movement compounded by a very fast-changing oxygen saturation in the supposedly slow desaturation phase. To understand the influence of movement more generally, the overall impact of movement

on the measurements was assessed by examining whether there was any link between the total movement within a session, as estimated by the standard deviation of the Euclidian distance of the face from its original position, and the coefficient of determination. To exclude the impact of movement during the movement itself, the movement SQI (SQI_L) was used in this analysis, while only the data for subjects of Type I-V was analysed to remove the influence of skin colour. The results are shown for both the RoR method and the baseline method in Figure 7.5. These indicate that, although large levels of movement ($\sigma_{movement} > 30$) do have a negative impact on the estimates, these alone cannot explain all decreases in performance, particularly for sessions with similar levels of overall movement but very different performances ($\sigma_{movement} < 30$).

Measurement-Reference Differences

Figure 7.6 (subject HP3-22, Fitzpatrick Type I) helps shed some light on cases in which both the low level of movement and the skin colour should lead to a good overall performance but for which this does not happen in practice. The figure, showing the regressed plots for both the baseline and RoR (with the $L++$ condition) methods overlaid on the reference oxygen saturation values from the pulse oximeter, clearly shows a time delay between the estimated oxygen saturation and the reference one. A shift in time, particularly when combined with a signal that changes rapidly in time (as occurs during the fast resaturation-desaturations) will not only give rise to high MAE values, but may also lead to low r^2 values.

Such a time delay is clearly visible in other records (HP3-27, HP3-34, HP3-39), but not in all sessions (it is for example notably absent in Figure 7.3). As the heart rate is well synchronised, the delay cannot be ascribed to a global delay (for example due to a clock mismatch) between the video camera and the pulse oximeters. Similarly, as it occurs across different Fitzpatrick Skin Types, the delay cannot be said to be linked to melanin content and is unlikely to have been caused by a particular setting or environmental

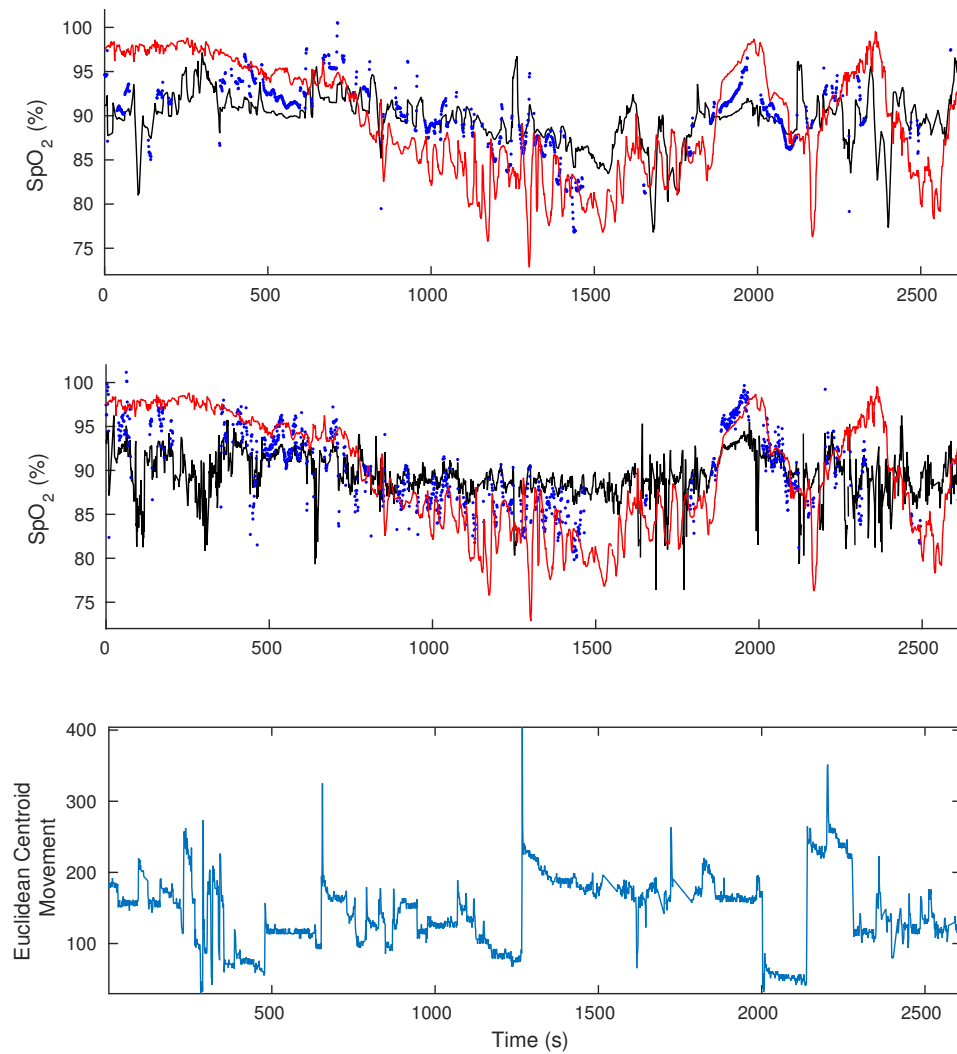


Figure 7.4: The effect of movement, shown in the lower plot as the Euclidean distance moved by the facial bounding box centroid, on the video camera estimates: baseline (top plot) and RoR (middle plot). Both estimates have been linearly regressed onto the reference oxygen saturation (shown in red). The black continuous line represents the estimates when no movement SQI is used, while the blue dots show the results when SQI_L is applied.

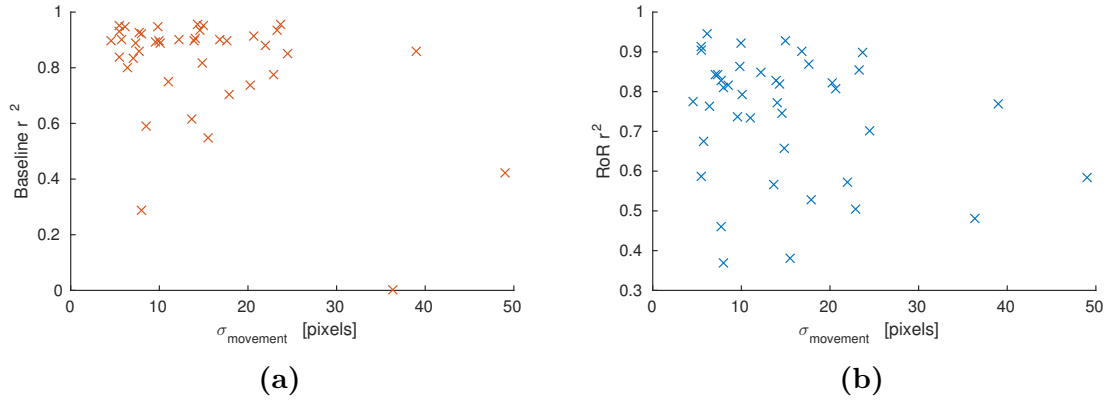


Figure 7.5: Effect of movement, calculated as the total standard deviation of the Euclidean distance travelled by the facial bounding box from its starting position, on the r^2 coefficient of determination for the baseline method (left) and the RoR method (right).

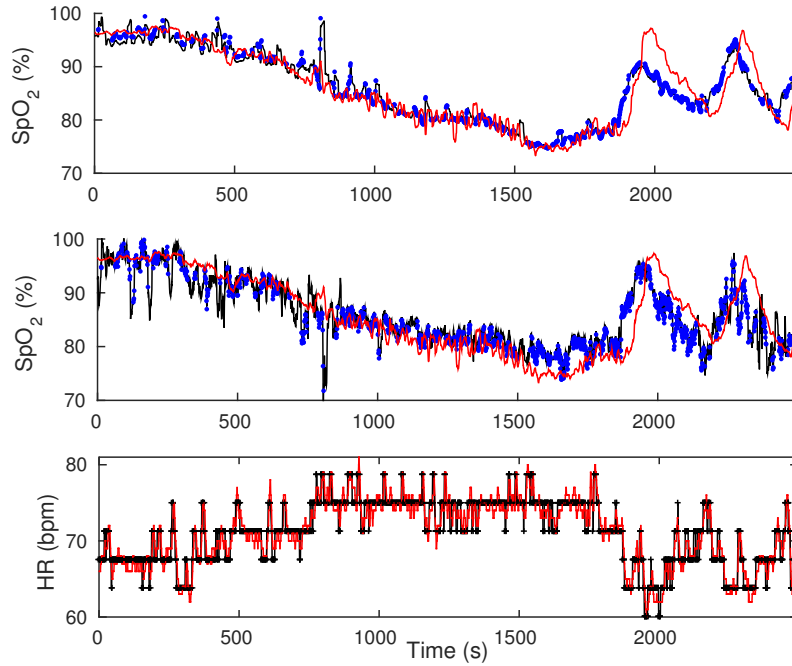


Figure 7.6: Plots showing the time delay between the reference oxygen saturation (red) and the camera estimates (in black for no movement SQI and blue for SQI_L) for the baseline method (top) and the RoR method (middle). The estimated and reference heart rates are shown in the lower plot, in black and red respectively. The phenomenon of baseline underestimation during resaturation is also visible.

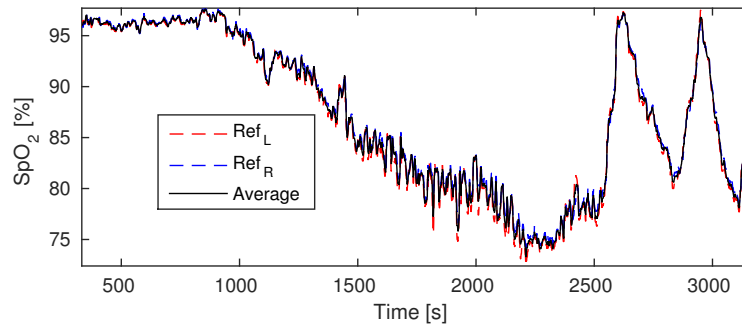


Figure 7.7: Plots showing the two reference oxygen saturation values (dashed lines), from the left (red dashed) and right (blue dashed) index fingers plotted against the mean of the two.

condition. Figure 7.7 further shows that the two reference values for SpO_2 are in close agreement with each other in the case of subject HP3-22.

In addition to this, the time delay between the reference values and the camera-derived estimates is of the order of 35 seconds. The reference pulse oximeter's reported averaging time is 2-4 seconds, whereas the camera estimates are derived using 15-second windows. Although further post-processing techniques may be employed by the pulse oximeter's manufacturer to estimate oxygen saturation in poor conditions, a delay of over 30 seconds is unlikely, particularly considering that in other subjects the two signals remained synchronous throughout the session. The case of the data for subject HP3-22 is further particularly instructive as:

- The subject was of Fitzpatrick Skin Type I and should therefore pose no problem to the pulse oximeter algorithms.
- The delay only becomes visible during the resaturation: if the pulse oximeter algorithm had a problem with the skin type (perhaps through over-saturation), the delay should have been constant from the start.
- The subject moved very little throughout the session: there was no further cause for the pulse oximeter to increase window length or otherwise delay the estimate.

Two possible explanations therefore remain. The first is that the two camera estimates are measuring some other quantity synchronous with oxygen saturation, such as vessel dilation, and that this quantity changes faster than the peripheral oxygen saturation. The other is simply that the blood perfusing the tissues around the face is oxygenated faster than the peripheral blood in the fingers. A combination of these two explanations seems most likely.

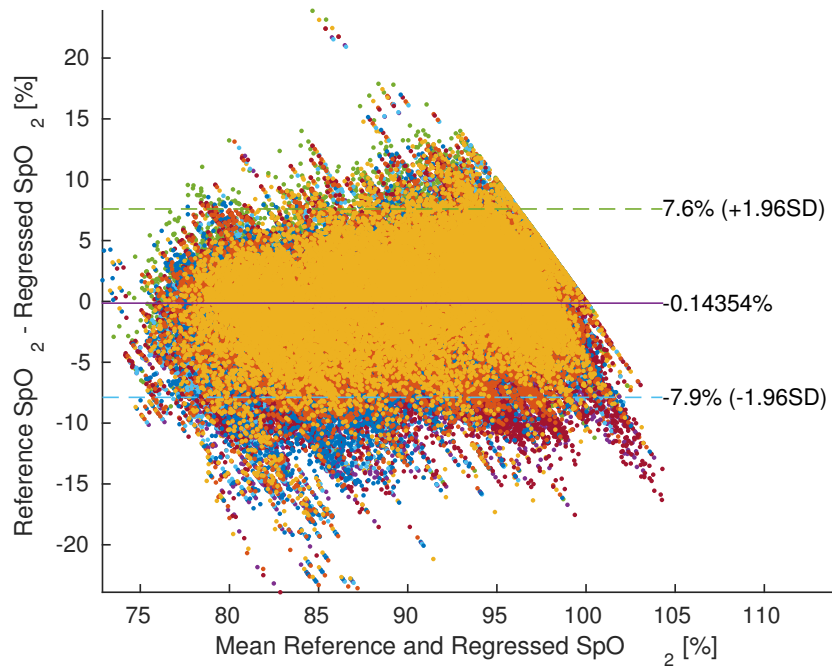
7.3 Towards Absolute SpO₂ Estimation

To determine whether an absolute value of SpO₂ could be obtained directly from camera data if a database with sufficient numbers of subjects were assembled, the sessions for all subjects of Fitzpatrick Type I-V (40 in total) were randomly split into two datasets of equal size, partitioned so as to ensure that every Fitzpatrick class was equally represented in the two datasets (training data and testing data). From Chapter 6.5, a linear relationship is expected between the average value of the ratio of ratios when the subject is breathing air at the start of the session (R_{air}) and each of the regression parameters λ_0 and λ_1 of Equation 6.20, of type:

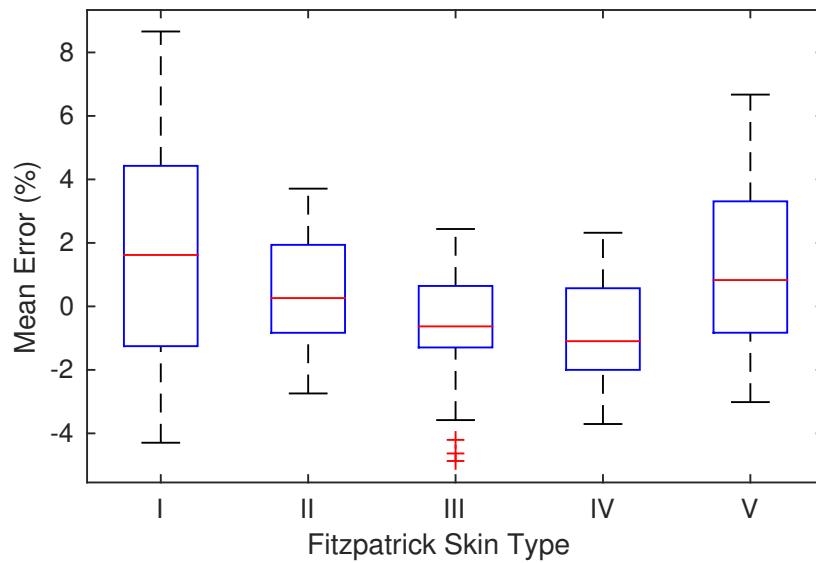
$$\begin{aligned}\lambda_0 &= k_0 + k_1 R_{air} \\ \lambda_1 &= k_2 + k_3 R_{air}\end{aligned}\tag{7.1}$$

This model was tested by obtaining the values of the meta-parameters k_0, k_1, k_2, k_3 using the training dataset (50% of the data) and evaluating the model on the test data (the remaining 50% of the data). Both when obtaining the meta-parameters, and when testing the models, the RoR estimates and the reference oxygen saturation were first cross-correlated values to find the delay between them. The delay was then corrected prior to finding the regression parameters or testing the goodness-of-fit.

This was repeated 10 times, randomly dividing the data for all subjects of Type I-V



(a)



(b)

Figure 7.8: Figures showing the performance of the absolute estimates for SpO₂ using learnt linear parameters. Figure 7.8a shows the Bland-Altman plot for the absolute estimates and the reference SpO₂, including all sessions for each of the 10 random iterations. Figure 7.8b shows the spread of the mean error for each session over the ten iterations, for each Fitzpatrick phenotype, as a box-and-whiskers plot. It should be noted that the reference pulse oximeters, although measuring values that are above 100% internally, have an artificial hard stop at 100%, giving rise to the characteristic straight diagonal line in Figure 7.8a.

into two training and test sets at each iteration as described above. The ability of the RoR-based model to estimate the absolute value of oxygen saturation is evaluated and is characterised by the Bland-Altman plot and the box plots shown in Figure 7.8. There is only a slight bias present in the Bland-Altman plot, indicating that the linear fit for λ_0 is fairly accurate. The box plots show that the extrema of the Fitzpatrick scale have the most spread in terms of error, and also have the greatest bias. In fact, perhaps unexpectedly, it is the Fitzpatrick Type I class that is the worst described by the linear model in terms of mean error, possibly because the effects of perfusion would be most visible in Type I subjects and these may offset the estimates.

To visualise the effect of the absolute SpO₂ estimation, the parameters learned in the last (10th) iteration were used to regress 10 (Fitzpatrick Skin Types I–V) of the original 12 subjects used in the derivation of the baseline and RoR methods in Chapters 5 and 6. This is shown in Figure 7.9. Although, as noted previously, the estimates suffer from a high variability, the fit is fairly accurate for all subjects except HP3-31 (Type I) and HP3-06 (Type V). However, while the estimates in the former are clearly wrong due to an over-estimation, it is not clear how poorly the fit works in HP3-06 due to the high variance of the results.

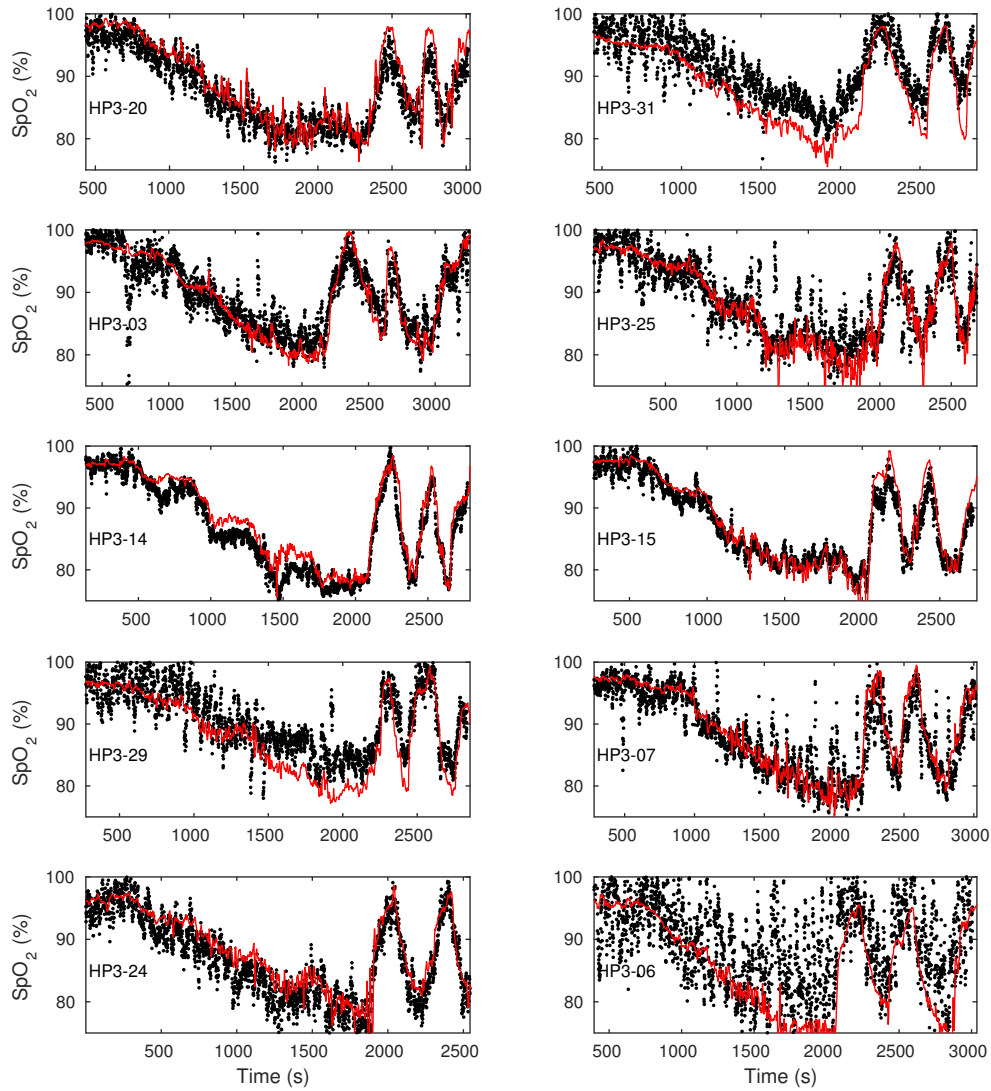


Figure 7.9: Plots showing the effect of using the learned regression parameters (last iteration) on the RoR data for the original test subjects (Fitzpatrick Skin Types I–V), with no SQI applied. The black points show the regressed estimates and the red line is the average reference oxygen saturation.

7.4 Conclusion

The methods developed in previous chapters have been tested on all subjects filmed in the course of the Hypoxia study, and the strengths and weaknesses of each method evaluated. A simple approach for obtaining the oxygen saturation values directly from the video data has also been attempted, showing promising results for future work in remote oxygen saturation monitoring using video cameras. The results and the features peculiar to the camera data will be further analysed in the following chapter, which will also discuss the main avenues for future work.

Chapter 8

Conclusions and Further Work

8.1 Introduction

In this Chapter, the results obtained and explanations given in Chapters 5–7 are reviewed in Section 8.2. This is done with the objective of revisiting one more time issue of which quantity is measured by the two methods proposed (Section 8.2.1), and secondly, discuss the factors which influence and limit these oxygen saturation estimates (Section 8.2.2). Some of the questions left unanswered after these discussions, as well as new avenues opened up for research within this field as a result of the work carried out in this thesis, are then addressed in Section 8.3.

8.2 Thesis Review

8.2.1 On the Origin of the Colour Changes

Chapter 7 demonstrates that, in well-controlled circumstances, oxygen saturation can be tracked with a video camera using the methods proposed in Chapters 5 and 6. However, although a large body of evidence has been accrued and presented to that effect, it is not clear whether the two methods are measuring tissue or skin oxygen saturation or indeed

another physiological change synchronous with it. A good example of one such change is that observed in Figure 7.6, showing the heart rate change with oxygen saturation. An algorithm for estimating oxygen saturation from the heart rate would achieve a reasonably high degree of accuracy in this particular case. Similarly, both the baseline and RoR methods may be measuring the effect of vasodilations and vasoconstrictions that accompany, or (as seen in Chapter 7) even precede, changes in peripheral oxygen saturation.

Nevertheless, a number of compelling arguments that suggest that the two methods are indeed sensing changes in oxygen saturation may be put forward. Firstly, each method is linked to existing methods for measuring oxygen saturation: the baseline method with oximetry and visual measurements of cyanosis, and the RoR method with pulse oximetry. Furthermore, the baseline method evaluates the relative proportion of “blueness” of the skin via the relationship captured by Equation 5.6, whereas the RoR method is measuring relative changes in amplitude, with the red channel driving the changes. In addition, the time-course of the RoR estimates, despite large oscillations about the “true” estimate, is close in shape to the reference oxygen saturation plots. If the RoR estimates were of some physiological variable other than blood oxygen saturation, they would be likely to diverge more often, much in the same way as the baseline estimates do not return to their original values after each resaturation cycle. The effect of hypoxia on vasodilation is in fact far from uniform: acute hypoxia is known to reduce blood flow to the skin, but at the same time chronic hypoxia may increase it [20]. Finally, the time delay between the camera-derived estimates and the oxygen saturation values from the pulse oximeters attached to the volunteers’ fingers, that may indicate that the video camera estimates are tracking some other quantity other than SpO_2 (such as vasodilation) as discussed in Chapter 7.2.4, is a phenomenon that has already been reported in the conventional pulse oximetry literature [113, 179, 215], for low-perfusion conditions. When comparing forehead (reflectance) or ear (transmittance) with finger (transmittance) pulse oximeters

in a variety of conditions, the latter were found to be delayed with respect to the former with differences of over 30s. According to [113], this is due to the high vasoactivity of the fingers compared to the forehead.

If it is oxygen saturation that is being measured, then, the question still remains as to which type of oxygen saturation it is. From the literature review conducted in Chapter 2, we know that pulse oximetry was devised specifically to obtain a better estimate of the arterial (peripheral) oxygen saturation, whereas oximetry samples mixed venous and arterial blood. Similarly, the baseline method's dependence on the overall colour of the skin is likely to indicate that both venous and arterial blood oxygen contribute to the camera-derived estimates, while the RoR method is measuring changes in a pulsatile compartment of the blood. However, unlike conventional pulse oximeters, the video cameras sense only visible light (the JAI camera has a near-infrared filter), as a result of which the blood saturation information comes from a depth that will not exceed 2-3mm [17], with the red light probing deeper layers of skin than the blue light. From Chapter ??, this would correspond to the upper layer of the dermis: the papillary dermis. The blood in this region mostly flows through capillaries as well as the supplying arterioles and exiting venules, with the pulsatile effect arising mostly from the arterioles themselves, or, as hypothesised by [87], from the compression by larger vessels below. Regardless of the mechanisms leading to the pulsatility, the blood oxygen saturation measured here is thus that of perfused skin tissue rather than that of the central arterial blood.

A parallel to these can also be found in visible light spectroscopy (VLS), discussed in Chapter 2.4.2. Benaron *et al.*, [15] examined the use of VLS in animal models, and found a strong correlation between SpO_2 and what they termed (cutaneous) microvascular tissue oxygenation (StO_2). If such a parallel were found to hold for our camera-derived estimates, then the measurements are more likely to be sensitive to mixed venous blood oxygen saturation (SvO_2) than arterial blood oxygen saturation (SaO_2) [134]. This parallel is informative and should be analysed in light of the differences between the two

methods employed, as shown in Figure 7.6, and of the theory espoused by [87].

The baseline method is clearly imaging an average of the changes in the blood present in a thin layer of skin tissue, and it is therefore very likely that it is in fact measuring cutaneous StO_2 . This would explain the different responses to reoxygenation with respect to the reference values, which are instead measuring peripheral arterial oxygen saturation. The RoR method, on the other hand, follows the form of the peripheral oxygen saturation changes accurately if not precisely. A reason may be found in the motivation for pulse oximetry: the RoR method is only dependent on the blood in pulsatile vessels (i.e. arterioles) and as such is closer to SpO_2 measurements than mixed venous-arterial oxygen saturation. This would, however, be in direct contradiction to the hypothesis put forward by [87]. They argue that the pulsatility that is seen when imaging an area of skin using visible light is caused by the compression of capillaries towards the upper compartments of the skin by larger vessels in the reticular dermis. If that were the case, since the oxygen saturation of the capillaries (unlike that of the arterioles) is determined by the mixed arterial and venous blood oxygenation, the RoR estimates will have a similar time-course as that of the baseline estimates. This is not what is observed (Figure 7.6) and therefore our experimental evidence casts some doubt on the hypothesis of [87].

8.2.2 Limitations

The secondary objective of the Hypoxia study was to determine which factors would limit the accuracy of the relationship found between the estimates obtained using a video camera and the reference oxygen saturation. The limitations are twofold. Firstly there are the limitations imposed by the experimental conditions, the most important amongst which are:

- The illumination spectrum
- The lighting arrangement

- The camera sensors used
- The small range of movement of volunteers during the study, compared to most real-time monitoring applications
- The type of hypoxia induced
- The limited range of physiological variation amongst the volunteers (who were all young, male, stationary, and healthy)

Secondly, the algorithms themselves introduced some limitations, empirically determined by testing them, as reported in Chapters 5, 6, and 7:

- The inability to track the oxygen saturation of all Type VI and some Type V subjects
- The inaccuracy of the baseline method
- The imprecision of the RoR method

The first set of limitations are extremely important as they indicate that, in the absence of a theoretical model or empirical data showing how the relationship between colour and oxygen saturation would change in different lighting conditions or for different sensors, many of the conclusions drawn from this thesis can only be applied to environments similar to that of the Hypoxia study (LED lamps and JAI camera). Furthermore, it is possible that the methodology for inducing hypoxia (whereby the volunteer breathed low-oxygen air, as opposed to the hypoxia being caused by a physiological change in the physiology of the volunteer) may have influenced both the hypoxic response of the volunteer and their skin response. Finally, the limits imposed on volunteer recruitment (male, 18-40 years of age), meant that the full spectrum of different skin types (for example the influence of wrinkles) and hypoxic responses was not studied, so that all inferences from this dataset are only applicable to a restricted segment of the population.

The theoretical derivation of the methods is, however, independent of these considerations, and relies solely on having two (for the RoR method) or three (for the baseline method) different channels having different responses to oxygen saturation changes. As such, although some of the ROI selection criteria may no longer be applicable, the same methodology could be applied to different sensor-lighting combinations.

The second set of limiting factors is perhaps more interesting, as it describes those situations within the chosen experimental conditions in which the algorithms were unable to work. First amongst these factors is skin colour. In order to test the influence of skin colour on the ability to determine oxygen saturation, volunteers were chosen amongst all Fitzpatrick skin Types (I-VI). The Fitzpatrick Skin Type is an assessment of risk to sunlight exposure and divides subjects into classes that should be related to the concentration of melanin in their skin. Although this method of categorising skin colours is highly subjective, it was found that it could at least classify the conditions under which the algorithms failed: both oxygen tracking methods failed on all Type VI subjects and some Type V subjects. This is likely to be due to the effect of increased melanin, and the consequent increase in the absorption of visible wavelengths, as discussed in Chapter 2.3.2. In particular, the strong absorption of higher frequency (blue) wavelengths for these darker-skinned subjects means that both the baseline and the RoR methods have a decreased signal-to-noise ratio. One way to quantify the effect of the absorptive properties of melanin on the resulting signal would be by covering a part of exposed skin on a Type I or Type II subject with an artificial absorber such as carbon black. This would allow a relation between the changes in signal-to-noise ratio to known quantities of the absorber to be explored, thereby giving a proxy assessment of the effect of pigmentation in a controlled setting.

Nonetheless, algorithm failure precludes their application to a large segment of the population, and indicates that more sensitive equipment is necessary to detect the weaker heart signal in dark-skinned individuals. Alternatively, unpigmented skin tissue (such as

that of the lips, the fingers, or the palms of the hands) may be imaged instead to bypass the effect of melanin.

The remaining limiting factors pertain to the properties of the techniques used for tracking oxygen saturation. In a surprising – and perhaps, as we shall see in Section 8.3, beneficial – duality, the baseline measure for oxygen saturation is very precise and varies little, whereas the RoR method for oxygen saturation is extremely noisy but follows the reference oxygen saturation more accurately. The reasons for these differences lie in the ways the two estimates are calculated: as already discussed in Section 8.2.1, the differences in accuracy are likely to be caused by differences in the type of blood oxygenation sampled. On the other hand, in Chapter 7.2.3, the differences in precision are ascribed to a combination of both the baseline estimates being more stable as a result of averaging, and the RoR estimates being less stable as a result of the variability of the amplitude calculation.

Both methods are equally affected by movement, as confirmed by the sudden movement visible at around 800s in Figure 7.6, which corrupts the estimates for both methods. The same plots, however, show there is no clear discontinuity between the periods prior to and following the movement, leading to the hypothesis that, for small movements such as those expected of volunteers sitting in a chair, these may be simply ignored, without having to compensate for them. This is also confirmed by the relative insensitivity to movement shown in Figure 7.5.

The example of Figure 7.6 also shows the difference in the effect of movement on both methods: whereas smaller movements are barely perceptible in the baseline signal (particularly in the first and last 500s of the plot), these are heavily accentuated in the output signal for the RoR method. Again, this is probably mostly a by-product of the method used in calculating the RoR: a sudden movement may not only modify a single time-averaged amplitude, but will be present in all four terms of the RoR equation (the AC and DC for both colour channels), greatly amplifying the error in the final result.

These limiting factors must be taken into account in the analysis of the first attempt at deriving an absolute measure of oxygen saturation. Having shown that SpO_2 can be *tracked* to varying degrees of accuracy, and that melanin can be an important limiting factor, we attempted to show that oxygen saturation could be estimated directly from the RoR estimates based on hyperparameters derived from a sub-set of the data (the training set). Only with the addition of such a step will the methods developed in this thesis become clinically useful. As there was no prior knowledge regarding the relationship between the video camera estimates and the oxygen saturation, a linear model was used (or to be precise, a linear-log model as the RoR is a log ratio of ratios) throughout the thesis to regress camera-derived estimates onto the reference oxygen saturation. In instances in which there was little or no noise in the data, as for HP3-13 in Figure 7.3, this simplistic model describes the oxygen saturation well (at least in the case of the RoR method).

To obtain an absolute estimate of oxygen saturation, however, the parameters of the linear model (λ_0 and λ_1 for the RoR model and ν_0 and ν_1 for the baseline model) need to be inferred from the starting conditions. As Chapters 5 and 6 have shown, there is no single set of constants that can be inferred from the data: each model is unique. By hypothesising that the parameters of the model are instead dependent on the starting conditions, the uniqueness of the model may be limited and an absolute estimate is recoverable only if the starting conditions are known. The hypothesis may otherwise be seen as simply incorporating the effect of the nature of the subject's skin: Equation 7.1 uses the starting values of the RoR (while the subject is breathing air) to gauge the effect of the particularities of the skin on the linear-log equation. Interestingly, although the most important factor within the skin should be melanin, the subjects in Figures 6.15 and 6.16 are not aligned by their Fitzpatrick class (something also found for the baseline model in Figure 5.12). This is probably an indicator of the fact that the Fitzpatrick scale is a poor assessment of the skin colour itself at the time of each session,

as it incorporates other factors such as eye and hair colour which may actually only be loosely related to the skin colour.

Since there was no prior, the parameters of the linear-log expression λ_0 and λ_1 were themselves modelled with the use of linear models, and it is the resulting four parameters (k_0 through to k_3), the hyperparameters, that were deemed to be fixed constants that could be learned from the data. The results shown in Chapter 7 are somewhat mixed. Half the data was used to obtain the hyperparameters and the other half was used to test it, and over 10 random iterations the mean bias from the Bland-Altman was comparable to that found between pulse oximeters. However, the Bland-Altman plot in Figure 7.8 also shows 1.96σ values around 7.5%, while for the individual fits of Chapter 6 these were 5.9% and for the inter-reference differences (between the two pulse oximeters) around 2.3%. Similarly, the boxplots show a very large range for the mean error, particularly in Fitzpatrick Skin Types I and V.

The results are indicative of the success in estimating λ_0 , as an error in estimating λ_0 would result in a significant bias, but also indicate that λ_1 is not being estimated correctly. In turn, this may be either because the simple linear model with only one variable (R_{air}) is insufficient to model λ_1 across different skin tones, or because λ_1 is highly sensitive and is affected by outliers.

Nevertheless, the results indicate that not only is it possible to track oxygen saturation, but to obtain an absolute value for it using video data alone. This constitutes the first step towards the adoption of remote monitoring of oxygen saturation, as, rather than needing multiple calibration values as would be the case for a tracking-only technology (in order to fit the linear-log line), only one calibration value is necessary (for example the oxygen saturation breathed at the beginning of the session and R_{air} in Equation 7.1). When combined with the camera-based vital sign monitoring techniques currently being developed for heart rate and breathing rate estimation, this has the potential to turn webcams and smart-phones into monitoring and diagnostic tools. In the next section of

this chapter, we shall examine the steps that need to be taken in order to reach this goal.

8.3 Further Work

Although there are many avenues that could be explored to investigate this field further, three main areas may be seen as crucial to any further meaningful development: (a) improving the current accuracy and precision of estimates within controlled experimental conditions; (b) testing the two approaches derived in a clinical scenario; (c) testing different light sources and sensors.

8.3.1 Improving Estimates: Uncertainty and Data Fusion

As well as showing the impact of the individual selection criteria, Figure 6.12 hints at another important relationship: the subjects that generate the best results for the RoR method have a higher overall proportion of accepted ROIs, independently of the ROI method of selection. The relationship between the error and the number of ROIs available to calculate the estimate was investigated for each subject used in the derivation of the baseline and RoR methods (excluding subjects HP3-06, HP3-04, and HP3-42). The absolute error between the linearly regressed R-B curves and the oxygen saturation was investigated as a function of the number of ROIs used to obtain the estimate, and is shown in Figure 8.1.

In accordance with the theoretical results of Chapter 6, Figure 8.1 demonstrates that there is a strong relationship between the spread in the mean absolute error and the number of ROIs selected. Since, as discussed in Chapter 6, the number of ROIs selected can instead be related to the total SNR, for example via an equation of type:

$$\text{SNR}_{tot} = \frac{1}{\sqrt{N}} \sum_{i=1}^N \text{SNR}_i, \quad (8.1)$$

where N is the number of ROI regions available, a relationship can be found between

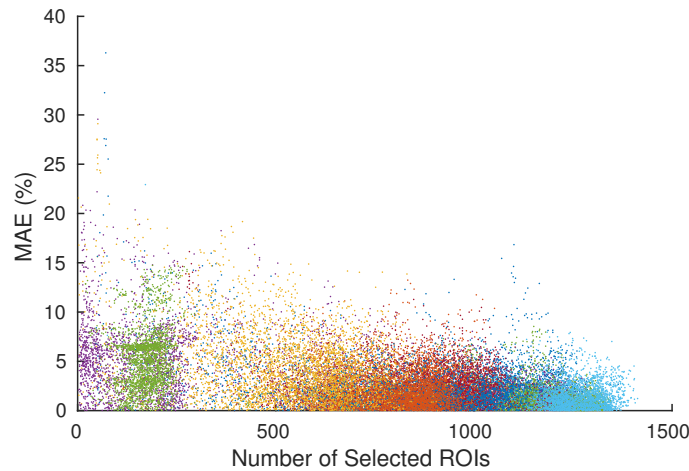


Figure 8.1: Variation of the mean absolute error (MAE) from window to window as a function of the number of ROIs used to calculate the RoR estimate. Each colour represents one of the nine subjects used in the derivation of the RoR method in Chapter 6 (HP3-06, HP3-04, and HP3-42 were excluded).

the SNR available within any given window and the uncertainty associated with the RoR estimate for that window. This is a particularly powerful tool, as the uncertainty may then be incorporated within a framework such as Gaussian Processes (GP) [163], or the Kalman Filter (KF) [85], allowing the selection of the optimal RoR estimate based on previous windows.

This could also be accomplished by taking advantage of the precision/accuracy split between the two types of video camera estimates. For example, within a KF-type approach, the baseline estimate could inform the prediction step of the Kalman Filter while the RoR estimate and its uncertainty could inform the update step.

As well as leading to an improved estimate of oxygen saturation through data fusion, the efficacy of each step in the derivation of the RoR algorithm can be assessed by the SNR measure as well. This would not only inform the theory behind the plot shown in Figure 8.1, but also allow an analysis of the sources of error that are added at each step, something which has not been covered in this dissertation, in which the steps were only evaluated by their final outcome on the SpO_2 results.

8.3.2 Clinical Validation

As mentioned in the Discussion, the way that hypoxia was induced in the Hypoxia study is in many ways different to hypoxia that occurs as the result of disease or apnoea. Replicating the results obtained within a clinical context is therefore essential for this technology to be adopted in clinical practice.

The clinical context presents a number of challenges, as noted in [202] in the first long-term heart rate and breathing rate remote monitoring work to be carried out in a hospital. In almost any clinical scenario, a lot more movement will be expected from the patient, who may move in and out of the scene. Furthermore, for long-term monitoring, both the lighting levels and the type of light are expected to vary over time – the time of day, clouds, and electrical lighting will all modify the type of illumination (if any) incident on the skin tissue. Last but not least, any movement by clinical staff or relatives between the patient and the light sources will cast a shadow on the subject that may cover different parts of the subject at different times.

All of these are first-order effects that will have a large impact on the measurements made by the video camera. Testing is required in order to evaluate whether they can be circumvented or constitute a limitation to the technology. Second-order effects instead include changes in the incident spectrum due to reflections off coloured surfaces (such as clothes or surrounding walls), phase delays between different ROIs when imaging large areas of skin (e.g. whole body imaging), and inter-ROI differences when imaging different parts of the skin (i.e. imaging the skin on the back of a subject may not be the same as imaging their face).

In order to illustrate the challenges that need to be overcome and suggest novel ways to do so, the algorithms developed within the Hypoxia study were applied, with some modifications, to a one hour section of a video from the NICU trial described in Chapter 3.2 (Figure 8.2 shows one frame from the video). Similarly to the Hypoxia study, the video data in the NICU trial was recorded using the JAI camera. However in the NICU

trial each channel only had an 8-bit range rather than the 12-bit used in the Hypoxia study so as to permit a higher sampling rate for the higher heart rates expected of infants. In order to compensate for this, the time-window was kept the same (15 seconds) despite neonatal heart rates being two to four times faster than those recorded in the course of the Hypoxia study. In addition, the regions of interest taken were enlarged from 12×12 to 48×48 , improving the SNR four-fold. The overlap between the lower range of neonatal heart rate frequencies (minima around 80bpm) and the upper range of neonatal breathing rate frequencies (maxima around 100bpm) meant that the amplitude ROI selection condition (L_a) could not be used, as the breathing rate could not be removed from the baseline without affecting the cardiac-synchronous AC signal and this therefore affected the calculation of the average amplitude in individual ROIs.



Figure 8.2: One frame from the NICU trial video used. The Colour Checker chart is visible placed next to the neonate.

The results for the one-hour session are shown in Figure 8.3, and illustrate the points made regarding the challenges concerning illumination in a standard clinical context. The left-hand plot, which shows the detrended light intensity on the cot in the vicinity of the neonate, is particularly instructive, as not only does it show intensity variations that are reflected in the outputs for both the baseline and the RoR methods (and occasionally

also seem to influence the reference), but also shows that the variations in intensity are far from uniform, with the individual composition of each lighting change varying significantly every time.

The results indicate clearly that the baseline method will be not be applicable in this setting unless it is heavily modified, for example by attempting to minimise the effect of the different lighting conditions on the skin. The RoR method is still able to track some oxygen saturation changes (as made clear by the last marked saturation – a particularly good example as it occurs independently of any background lighting change), however it is also strongly affected by lighting changes. Overall, if the coefficients of determination of the two methods (0.01 and 0.09 for baseline and RoR respectively) are compared with the average for the hypoxia data over a similar periods of time, it is clear that these methods will need a substantial amount of work before being applicable to the NICU dataset.

A number of methods can be used to improve the results under these conditions. Vital sign data fusion methods could help detect apnoeas during a lighting change or during movement. A model can further be constructed of the expected effect of different lighting on skin and used to compensate for lighting changes that are detected by background pixels. Because shadows will only cover part of the image, in such a scenario the skin pixels cannot be considered as a single block but should instead be divided according to the intensity/type of light incident on them.

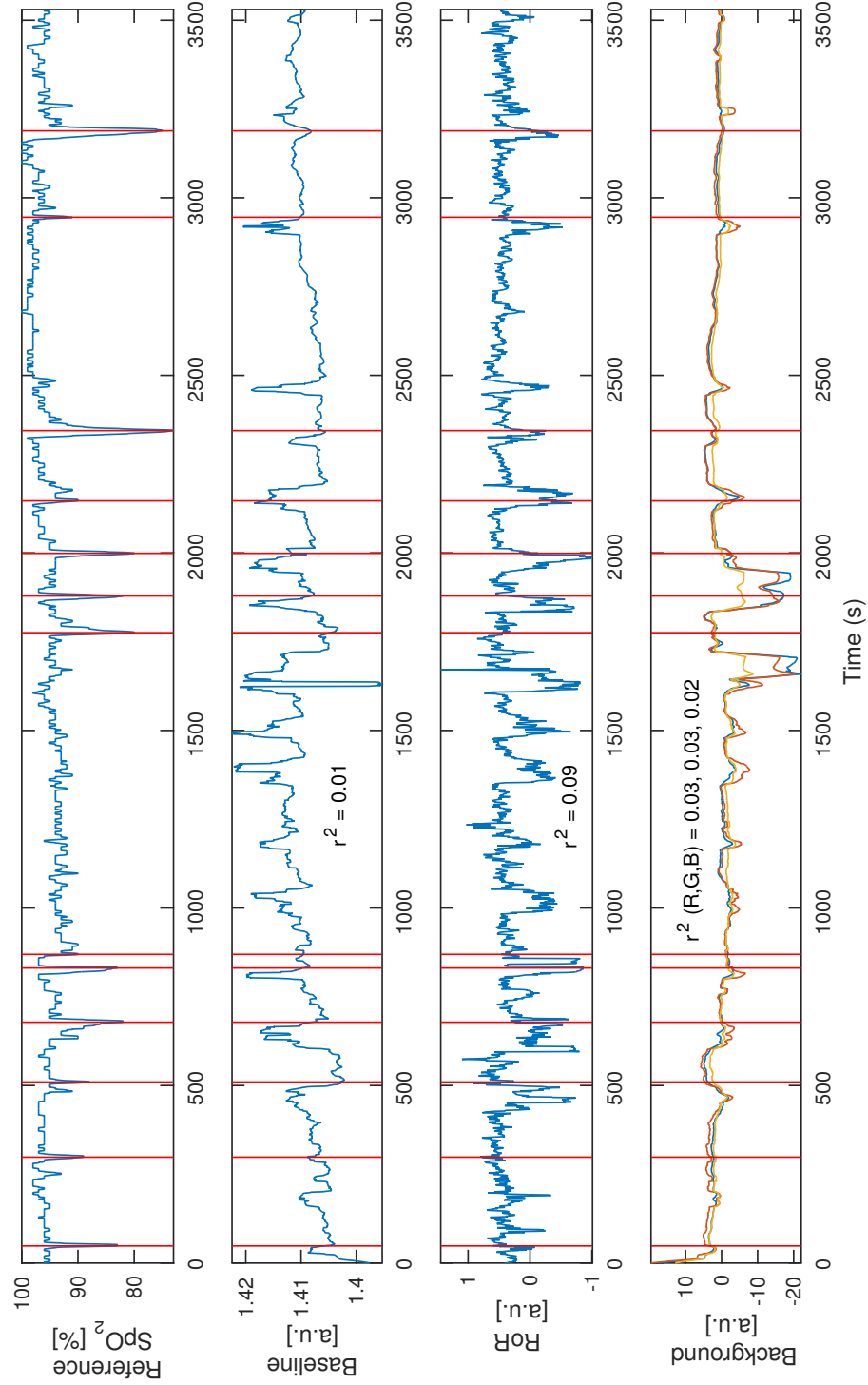


Figure 8.3: Plots for just under 1 hour of a clinical session in the NICU. From left to right: reference oxygen saturation from a pulse oximeter; baseline estimates using the polar angle $\frac{\pi}{2} - \phi$; RoR estimates using the ratio of blue to red normalised amplitudes; detrended background changes. The straight red lines indicate the lowest point of major reference desaturations.

8.3.3 Alternative Methodologies

An alternative to modifying or tuning the algorithms to different sensor-lighting configurations is to control the illumination itself. Other than the obvious advantage of being more generally applicable, this approach has a number of benefits. Background lighting, variable or not, may be measured by pulsing the controlled illumination in synchrony with the frame acquisition by the video camera, much in the same way as is done currently in pulse oximetry. The measurements made during the off period can then be used to normalise the readings during the on period. Such an arrangement would require a high frame rate – sufficiently high to preserve the waveform shape and allow background normalisation.

A further advantage of controlling the illumination is that the illuminating wavelengths can be selected so that they have useful properties to overcome the limitations discussed earlier in this chapter. The choice of wavelengths for which melanin has a low absorption coefficient would allow the monitoring of oxygen saturation in Type V and Type VI subjects. It is also essential that the illumination should not be visible to the subject. Although not directly relevant to the physiological measurements, the last condition will allow for monitoring at night, thereby enabling continuous 24-hour monitoring. These requirements are met by illumination within the 700-1000nm electromagnetic wavelength range: the near infrared (NIR). An approach similar to pulse oximetry may then be used to estimate the oxygen saturation from multiple wavelengths. The drawback of such a solution, however, is that the illumination needs to be controlled carefully as continuous retinal exposure to NIR light may be hazardous [95]. A device using such illumination technology will have to be regulated.

8.4 Concluding Remarks

Interest in remote vital sign monitoring using video cameras has been gaining momentum in the last decade, during which time the number of research groups actively working in this field has grown exponentially. Curiously, although the stated objective of the initial papers in this field was precisely to replicate the pulse oximeter in all of its functions, the focus of much of the research output has been the development of improved algorithms for heart rate estimation. The work presented in this thesis has attempted to lay the foundation for a renewed effort towards the development of a full remote pulse oximeter.

The potential of this technology is however much greater than merely replicating a device that is already low-cost, easy to use, and effective. The ubiquity of video cameras of course makes this an ideal solution for telehealth applications within a healthcare system that is beginning to make the home, rather than the hospital, the first line of care. However, it is the potential of video-based monitoring for going beyond pulse oximetry that is most exciting.

The key question that will continue to have to be addressed, especially in a clinical context, is that of movement. When a desaturation occurs, the subject is often likely to move at the same time, however all movement correction algorithms will hide an extremely important observation about the patient. How the subject is responding to the desaturation is a question that can, for the first time, be answered rigorously thanks to the use of cameras as pulse oximeters. If video cameras can be used to monitor vital signs in and out of the clinic, physiological variables such as pallor and other information such as movement and facial expressions (routinely used, perhaps even unconsciously, to inform the decisions of healthcare staff) would be rapidly incorporated into normal clinical care.

Appendix A

Mathematical Notes

A.1 Viola-Jones and Adaboost

The Matlab implementation of the Viola-Jones algorithm, (`CascadeObjectDetector` of the Matlab Computer Vision System Toolbox, Version 6.1 (R2014b)) uses Haar features, shown in Figure [A.1](#), rectangles comprised of two, three or four alternate smaller rectangles that are varied in size and moved around the image. At each different position the sum of the pixels in the light parts of the rectangle are subtracted from the features in the dark parts of the rectangle. The use of these simple features allows a rapid speed-up in computation thanks to the use of integral images, images defined such that any pixel at position (i, j) is characterised by a value V that is the sum of intensities I above it and to the left of it

$$V(i, j) = \sum_{\substack{i' \leq i \\ y' \leq y}} I(i', j') \quad (\text{A.1})$$

such that the sum of any rectangular region within the integral images only requires four corner values, allowing constant speeds of calculation regardless of the size of the rectangle used. The Haar features are trained on a dataset of 20×20 pixel images

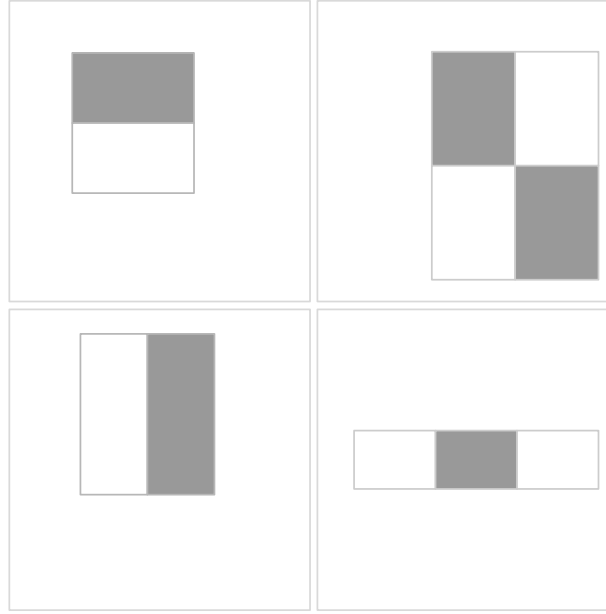


Figure A.1: Figure showing the four Haar features used in the Viola-Jones detection framework proposed in [228].

of frontal-facing faces, using a modified AdaBoost learning algorithm to reduce the number of features and train strong classifiers from a weighted linear sum of weak classifiers. These strong classifiers are then applied in successive, increasingly complex stages (cascade architecture) in which classifiers formed of increasing numbers of features are applied so as to reduce the number of false positives returned by the algorithm. In [106], different boosting algorithms were evaluated experimentally, and it was shown that the combination of Gentle AdaBoost and classification and regression trees (CART) performed best, and it is this combination that is used in the implementation used.

A.2 Movement Thresholding

The method is based on the concept that for the movement signal derived from the Euclidian distance moved by the centroid of the Viola-Jones face bounding box, in which the subject is expected to remain relatively still for at least some of the time, if a sliding standard deviation is calculated using variable-length windows the distribution of

standard deviations over the timeseries will tend to peak at low values for short windows, and only increase when the windows are so large that the majority of the windows include some motion. At that point the distribution of the standard deviations is expected to change significantly, as a step-change in the modal value for binned standard deviations is expected. The optimal value for the standard deviation threshold θ is thus chosen to be the mean value of the mode of the binned standard deviations prior to the step-change.

As there is a limit to the smallest time-window that may be used in estimating the heart rate (usually around 10-20 seconds), this may be used as the lower bound for the length of the running window. Similarly, an expectation of the longest time a person can remain without moving much may be used as a higher bound for the window length so as to decrease the computational time necessary to run the algorithm. The number of bins used to group the standard deviations across the timeseries is instead determined via the Freedman-Diaconis criterion [58], that states that the bin width for a dataseries of N points x with an interquartile range given by $\text{IQR}\{x\}$ should be:

$$h = 2 \frac{\text{IQR}\{x\}}{N^{1/3}} \quad (\text{A.2})$$

The mean value of the bin corresponding to the mode of the binned distribution is tracked for each iteration, in which the length of each sliding window is increased. The modal values are then plotted for each window-length and the jump in modal values is detected using a Bayesian single change-point detector based on [147]. The detector tests whether there is a point $m \in \mathbb{N}^+$ in the range $[1, W]$ where W is the number of windows of different lengths that have been used that maximises the probability of the points either side it being from two different distributions (for further details on the algorithm, please refer to Appendix A.3). The point prior to the changepoint is then chosen as having the optimal value for the standard deviation threshold θ for the window of choice. In Figure A.2, the method proposed is compared against the changepoints detected by the online Bayesian algorithm proposed by [2] for a window length of 15

seconds moved by 1 second at a time. As may be seen in Figure A.2, the classification works well visually, even detecting changepoints ignored by [2], however it fails to detect and segment slower motion, and this is particularly visible after 2000s, during which a slow change in the head position is classified as no motion.

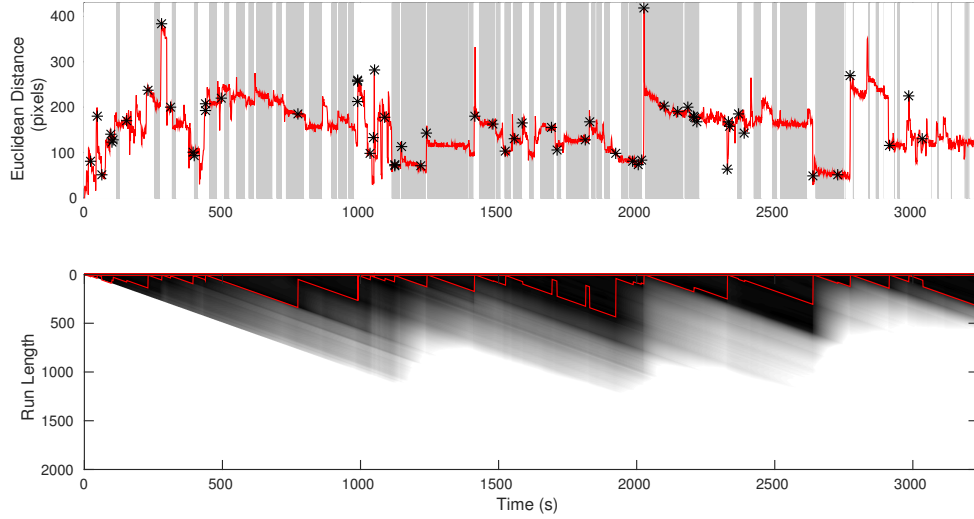


Figure A.2: Figure showing the segmentation of the motion data, shown in red for the subject undergoing significant movement in Figure 4.9, using the method proposed. The areas shaded in grey are areas classified as still. By way of comparison, the Bayesian online change-point detection algorithm proposed by [2] was also used on the same data, and the changepoints found are shown in black on the top plot, while the bottom plot shows the run lengths in red, and the posterior probability of the run length.

A.3 Bayesian Change Point Detection

For a sequence x of W samples divided into two segments with different means (μ_1 and μ_2) and Gaussian noise ϵ that is independent of the mean, identically distributed and has standard deviation σ such that ϵ_i is drawn from $\mathcal{N}(0, \sigma^2)$:

$$x_i = \begin{cases} \mu_1 + \epsilon, & \text{if } i \leq m \\ \mu_2 + \epsilon, & \text{otherwise} \end{cases} \quad (\text{A.3})$$

As the noise is distributed normally, the probability density function for the noise samples is given by

$$P(\epsilon) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\epsilon-\mu)^2}{2\sigma^2}} \quad (\text{A.4})$$

Furthermore, the likelihood of the data given these parameters is given by the joint probability of each of the noise samples as each is independently occurring:

$$P(x|\mu_1\mu_2\sigma m) = \prod_{i=1}^W P(\epsilon_i) \quad (\text{A.5})$$

which can be expanded into

$$P(x|\mu_1\mu_2\sigma m) = (2\pi\sigma^2)^{-\frac{W}{2}} e^{-\frac{1}{2\sigma^2}(\sum_{i=1}^m (x_i-\mu_1)^2 + \sum_{i=m+1}^W (x_i-\mu_2)^2)} \quad (\text{A.6})$$

Bayes theorem holds that the posterior probability may be found from the likelihood, thus $P(\mu_1\mu_2\sigma m|x)$ may be inferred from $P(x|\mu_1\mu_2\sigma m)$ provided $P(x)$ and $P(\mu_1\mu_2\sigma m)$ are known. If all parameter values are equally likely, then we have $P(\mu_1), P(\mu_2), P(\sigma)$ all constant, with

$$P(\sigma) \propto \frac{1}{\sigma} \quad (\text{A.7})$$

as $P(\ln \sigma) = \text{constant}$. By Bayes we finally obtain:

$$P(\mu_1\mu_2\sigma m|x) \propto \frac{1}{\sigma} P(x|\mu_1\mu_2\sigma m) \quad (\text{A.8})$$

However μ_1, μ_2 , and σ are not of interest. These are “nuisance” parameters that

may be integrated away in order to leave an expression for $P(m|x)$ only ¹:

$$\begin{aligned}
 P(m|x) &= \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{1}{\sigma} P(x|\mu_1\mu_2\sigma m) d\mu_1 d\mu_2 d\sigma \\
 &\propto \frac{1}{\sqrt{m(N-m)}} \left[\sum_{i=1}^W x_i^2 - \frac{1}{m} \left(\sum_{i=1}^m x_i \right)^2 - \frac{1}{N-m} \frac{1}{m} \left(\sum_{i=m+1}^W x_i \right)^2 \right]^{\frac{N-2}{2}} \quad (\text{A.9})
 \end{aligned}$$

A.4 Rodrigues Rotation Method

The Rodrigues rotation method, developed independently by Euler and Rodrigues [42], allows the efficient calculation of the rotation between two vectors in three dimensions. It states that the rotation θ between a vector $\mathbf{a}$ and another vector $\mathbf{b}$ about a unit axis vector $\hat{\mathbf{k}}$ defining (i.e. orthogonal to) the rotation plane may be written as [60]:

$$\mathbf{b} = \mathbf{a} \cos \theta + (\hat{\mathbf{k}} \times \mathbf{a}) \sin \theta + \hat{\mathbf{k}}(\hat{\mathbf{k}} \cdot \mathbf{a})(1 - \cos \theta) \quad (\text{A.10})$$

where the unit axis vector is defined as:

$$\hat{\mathbf{k}} = \frac{\mathbf{a} \times \mathbf{b}}{|\mathbf{a} \times \mathbf{b}|} \quad (\text{A.11})$$

In matrix form, this may be expressed as the relation between $\mathbf{a}$ and $\mathbf{b}$ via the 3×3 rotation matrix $\mathbf{R}$:

$$\mathbf{b} = \mathbf{R}\mathbf{a} \quad (\text{A.12})$$

where $\mathbf{R}$ is given by:

$$\mathbf{R} = \mathbf{I} + \sin \theta \mathbf{K} + (1 - \cos \theta) \mathbf{K}^2 \quad (\text{A.13})$$

¹This is done by integrating over μ_1 and μ_2 using the integral identity for quadratic exponentials $\int_{-\infty}^\infty e^{-ax^2-bx-c} dx = \sqrt{\frac{\pi}{a}} \left(\frac{b^2}{4a} - c \right)$ and the gamma integral identity for σ later.

and $\mathbf{K}$ is the cross-product matrix for the unit axis vector $\hat{\mathbf{k}}$:

$$\mathbf{K} = \begin{bmatrix} 0 & -k_3 & k_2 \\ k_3 & 0 & -k_1 \\ -k_2 & k_1 & 0 \end{bmatrix} \quad (\text{A.14})$$

A.5 The CHROM Algorithm for Heart Rate

The CHROM algorithm² for heart rate monitoring using RGB video cameras in ambient lighting was first presented in [45], showing an improvement to heart rate detection in difficult conditions (movement, different lighting) with respect to previous methods while still retaining much of the computational efficiency of those methods. In addition to this, unlike other methods published by other research groups, the method was tested on a large sample of volunteers of different skin phenotypes, and can therefore be considered more reliable. The implementation of the CHROM algorithm within this thesis follows the original paper closely.

The method itself attempts to improve the SNR by taking a favourable combination of the three colour channels. It is based on the concept of chrominance, and on empirically derived constants with which mean-normalised colour channels (denoted $[R_n, G_n, B_n]$) can be corrected to account for non-white illuminants. The resulting chrominance vectors are:

$$\begin{aligned} X_s &= 3R_n - 2G_n \\ Y_s &= 1.5R_n + G_n - 1.5B_n \end{aligned} \quad (\text{A.15})$$

which are then each filtered using a zero-phase bandpass FIR filter with a pass-band between 42 and 132 beats per minute (0.7–2.1Hz) to yield X_f and Y_f , and then combined

²It should be noted that CHROM is not the method's baptismal name, but simply one it acquired later in the literature: in the paper in which it was introduced it was referred to as the $X_s \min \alpha Y_s$ algorithm

so as to give the final pulsatile signal

$$S = X_f - \alpha Y_f \quad (\text{A.16})$$

with:

$$\alpha = \frac{\sigma X_s}{\sigma Y_s} \quad (\text{A.17})$$

To obtain the heart rate, the FFT of this signal is then taken and the dominant frequency in the heart rate frequency range is selected.

The rationale for the last two equations is the following: as chrominance channels are thought to be equally affected by movement but differently affected by the pulsatile HR-synchronous signal, the ratio of the two chrominance channels should lead to a signal that is not affected by movement. The ratio of the chrominance channels can however be shown to be (by Taylor expansion), approximately equal to a difference between the chrominance channels, such that:

$$\frac{X_f}{Y_f} \approx X_f - Y_f \quad (\text{A.18})$$

and the addition of the α factor simply accounts for the minimisation of in-band disturbances that may have been produced by inaccuracies in the illumination correction.

Bibliography

- [1] A. K. Abbas, K. Heimann, K. Jergus, T. Orlikowsky, and S. Leonhardt. Neonatal non-contact respiratory monitoring based on real-time infrared thermography. *BioMedical Engineering OnLine*, 10(1):93, 2011.
- [2] R. P. Adams and D. J. C. Mackay. Bayesian Online Changepoint Detection. Technical report, Cambridge, UK, 2007.
- [3] G. S. Agashe, J. Coakley, and P. D. Mannheimer. Forehead Pulse Oximetry. *Anesthesiology*, 105(6):1111–1116, 2006.
- [4] J. Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, 28(3):R1–R39, 2007.
- [5] D. Arthur, D. Arthur, S. Vassilvitskii, and S. Vassilvitskii. k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, 8:1027–1035, Philadelphia, PA, USA, 2007. Society for Industrial and Applied Mathematics.
- [6] G. J. Arthurs and M. Sudhakar. Carbon dioxide transport. *Continuing Education in Anaesthesia, Critical Care & Pain*, 5(6):207–210, 2005.
- [7] U. Bal. Non-contact estimation of heart rate and oxygen saturation using ambient light. *Biomedical Optics Express*, 6(1):86, 2015.
- [8] G. Balakrishnan, F. Durand, and J. Guttag. Detecting Pulse from Head Motions in Video. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3430–3437, 2013.
- [9] G. Baranoski and A. Krishnaswamy. An introduction to light interaction with human skin. *Revista de Informática Teórica*, 11(1):33–62, 2004.
- [10] S. J. Barker, J. Hyatt, N. K. Shah, and Y. J. Kao. The effect of sensor malpositioning on pulse oximeter accuracy during hypoxemia. *Anesthesiology*, 79(2):248–54, 1993.
- [11] S. J. Barker and J. J. Badal. The measurement of dyshemoglobins and total hemoglobin by pulse oximetry. *Current Opinion in Anaesthesiology*, 21(6):805–810, 2008.

- [12] P. B. Batchelder and D. M. Raley. Maximizing the laboratory setting for testing devices and understanding statistical output in pulse oximetry. *Anesthesia and analgesia*, 105(6 Suppl):S85–94, 2007.
- [13] a. Beer. Versuch die absorptions-verhaltnisse des cordierites fur roth bestimmen. *Annalen der Physik und Chemie*, 84:37–52, 1851.
- [14] D. Belega and D. Petri. Accuracy analysis of the amplitude estimation of a sine-wave by the energy-based method using the direct and indirect procedures. *Measurement*, 45(9):2264–2274, 2012.
- [15] D. a. Benaron, I. H. Parachikov, D. Ph, S. Friedland, R. Soetikno, E. F. Fincher, C. P. Hsu, F. L. Clark, W.-f. Cheong, J. L. Duckworth, D. K. Stevenson, and M. D. Continuous, Noninvasive, and Localized Microvascular Tissue Oximetry Using Visible Light Spectroscopy. *Anesthesiology*, 100(6):1469–1475, 2004.
- [16] M. Bigas, E. Cabruja, J. Forest, and J. Salvi. Review of CMOS image sensors. *Microelectronics Journal*, 37(5):433–451, 2006.
- [17] T. Binzoni, a. Vogel, a. H. Gandjbakhche, and R. Marchesini. Detection limits of multi-spectral optical imaging under the skin surface. *Physics in medicine and biology*, 53(3):617–36, 2008.
- [18] V. Blazek, T. Wu, and D. Hoelscher. Near-Infrared CCD Imaging - Possibilities for Noninvasive and Contactless 2D Mapping of Dermal Venous Hemodynamics. In *BiOS 2000 The International Symposium on Biomedical Optics*, 3923:2–9, 2000.
- [19] F. Bousefsaf, C. Maaoui, and A. Pruski. Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate. *Biomedical Signal Processing and Control*, 8(6):568–574, 2013.
- [20] A. T. Boutin, A. Weidemann, Z. Fu, L. Mesropian, K. Gradin, C. Jamora, M. Wiesener, K.-U. Eckardt, C. J. Koch, L. G. Ellies, G. Haddad, V. H. Haase, M. C. Simon, L. Poellinger, F. L. Powell, and R. S. Johnson. Epidermal Sensing of Oxygen Is Essential for Systemic Hypoxic Response. *Cell*, 133(2):223–234, 2008.
- [21] P. Boutouyrie, M. Briet, C. Collin, S. Vermeersch, and B. Pannier. Assessment of pulse wave velocity. *Artery Research*, 3(1):3–8, 2009.
- [22] C. S. Breathnach. The development of blood gas analysis. *Medical History*, 16(1):51–62, 1972.
- [23] W. a. G. Bruls and J. C. Van Der Leun. Forward Scattering Properties of Human Epidermal Layers. *Photochemistry and Photobiology*, 40(2):231–242, 1984.
- [24] D. L. J. Bunker, R. Kumar, A. Martin, and S. P. Pegg. Thermal Injuries Caused by Medical Instruments. *Journal of Burn Care & Research*, 35(2):e132–e134, 2014.

- [25] B. Carlson. Comparison of modern CCD and CMOS image sensor technologies and systems for low resolution imaging, 2002.
- [26] C. G. Caro and J. a. Bloice. Contactless apnoea detector based on radar. *Lancet*, 2(7731):959–961, 1971.
- [27] M. a. Changizi, Q. Zhang, and S. Shimojo. Bare skin, blood and the evolution of primate colour vision. *Biology Letters*, 2(2):217–221, 2006.
- [28] J. M. Chaves-González, M. a. Vega-Rodríguez, J. a. Gómez-Pulido, and J. M. Sánchez-Pérez. Detecting skin in face recognition systems: A colour spaces study. *Digital Signal Processing: A Review Journal*, 20(3):806–823, 2010.
- [29] D.-Y. Chen, J.-J. Wang, K.-Y. Lin, H.-H. Chang, H.-K. Wu, Y.-S. Chen, and S.-Y. Lee. Image Sensor-based Heart Rate Evaluation from Face Reflectance Using Hilbert-Huang Transform, 2014.
- [30] A. Chung, X. Y. Wang, R. Amelard, C. Scharfenberger, J. Leong, J. Kulinski, A. Wong, and D. a. Clausi. High-resolution motion-compensated imaging photoplethysmography for remote heart rate monitoring. In *SPIE BiOS*, 9316:93160A, 2015.
- [31] J. P. C. Cobb, R. S. H. Otchkiss, I. E. K. Arl, and T. G. B. Uchman. Mechanisms of cell injury and death. *British Journal of Anaesthesia*, 77(1):3–10, 1996.
- [32] S. M. Cohn, A. B. Nathens, F. a. Moore, P. Rhee, J. C. Puyana, E. E. Moore, and G. J. Beilman. Tissue oxygen saturation predicts the development of organ dysfunction during traumatic shock resuscitation. *The Journal of trauma*, 62(1):44–54; discussion 54–55, 2007.
- [33] J. H. Comroe and S. Botelho. The unreliability of cyanosis in the recognition of arterial anoxemia. *The American Journal of the Medical Sciences*, 214(1):1–6, 1947.
- [34] R. L. Cook and K. E. Torrance. A Reflectance Model for Computer Graphics. *ACM Trans. Graph.*, 1(1):7–24, 1982.
- [35] P. Corcuff, C. Bertrand, and J. L. Leveque. Morphometry of human epidermis in vivo by real-time confocal microscopy. *Archives of dermatological research*, 285(8):475–81, 1993.
- [36] F. Corral, G. Paez, and M. Strojnik. A photoplethysmographic imaging system with supplementary capabilities. *Optica Applicata*, XLIV(2):191–204, 2014.
- [37] L. F. Corral Martinez, G. Paez, and M. Strojnik. Optimal wavelength selection for noncontact reflection photoplethysmography. In *22nd Congress of the International Commission for Optics: Light for the Development of the World*, 8011:801191–801191–7, 2011.

- [38] G.-E. Costin and V. J. Hearing. Human skin pigmentation: melanocytes modulate skin color in response to stress. *The FASEB Journal*, 21(4):976–994, 2007.
- [39] J. a. Crowe and D. Damianou. The wavelength dependence of the photoplethysmogram and its implication to pulse oximetry, 1992.
- [40] W. Cui, L. E. Ostrander, and B. Y. Lee. In vivo reflectance of blood and tissue as a function of light wavelength. *IEEE Transactions on Biomedical Engineering*, 37(6):632–639, 1990.
- [41] A. Dahan, D. J. F. Nieuwenhuijs, and E. Olofsen. Influence of propofol on the control of breathing. *Advances in experimental medicine and biology*, 523(1):81–92, 2003.
- [42] J. S. Dai. EulerRodrigues formula variations, quaternion conjugation and intrinsic connections. *Mechanism and Machine Theory*, 92:144–152, 2015.
- [43] J. Daly and L. Tarassenko. Method and apparatus for measuring and displaying a haemodynamic parameter, 2014.
- [44] G. de Haan and a. van Leest. Improved motion robustness of remote-PPG by using the blood volume pulse signature. *Physiological measurement*, 35(9):1913–1926, 2014.
- [45] G. De Haan and V. Jeanne. Robust pulse rate from chrominance-based rPPG. *IEEE Transactions on Biomedical Engineering*, 60(10):2878–2886, 2013.
- [46] J. P. de Kock and L. Tarassenko. Pulse oximetry: theoretical and experimental models. *Medical & biological engineering & computing*, 31(3):291–300, 1993.
- [47] P. Eberhard, K. Hammacher, and W. Mindt. Perkutane Messung des Sauerstoffpartialdruckes. In *Proc. Medizin-Technik*, Stuttgart:26, 1972.
- [48] P. Eberhard and W. Mindt. Skin sensors for continuous oxygen monitoring of newborns. *Biotelemetry*, 4(2):48–76, 1977.
- [49] a. E. El-Desoky, L. R. Jiao, R. Havlik, N. Habib, B. R. Davidson, and a. M. Seifalian. Measurement of hepatic tissue hypoxia using near infrared spectroscopy: comparison with hepatic vein oxygen partial pressure. *European surgical research. Europäische chirurgische Forschung. Recherches chirurgicales européennes*, 32(4):207–14, 2000.
- [50] J. R. Estepp, E. B. Blackford, and C. M. Meier. Recovering pulse rate during motion artifact with a multi-imager array for non-contact imaging photoplethysmography, 2014.
- [51] L. Feng, L.-M. Po, X. Xu, and Y. Li. Motion artifacts suppression for remote imaging photoplethysmography, 2014.
- [52] I. Fine and a. Weinreb. Multiple scattering effect in transmission pulse oximetry. *Medical and Biological Engineering and Computing*, 33(5):709–712, 1995.

- [53] R. a. Fisher. Tests of Significance in Harmonic Analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 125(796):54–59, 1929.
- [54] T. B. Fitzpatrick. The validity and practicality of sun-reactive skin types I through VI. *Archives of dermatology*, 124(6):869–871, 1988.
- [55] T. B. Fitzpatrick. Soleil et peau. *J Med Esthet*, 2(7):33–34, 1975.
- [56] P. M. Fortune, M. Wagstaff, and a. Petros. Cerebro-splanchnic oxygenation ratio (CSOR) using near infrared spectroscopy may be able to predict splanchnic ischaemia in neonates. *Intensive Care Medicine*, 27(8):1401–1407, 2001.
- [57] S. Fouzas, K. N. Priftis, and M. B. Anthracopoulos. Pulse Oximetry in Pediatric Practice. *Pediatrics*, 128(4):740–752, 2011.
- [58] D. Freedman and P. Diaconis. On the histogram as a density estimator: L2 theory. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57(4):453–476, 1981.
- [59] Z. Gajinovic, M. Matic, S. Prcic, and V. Duran. Optical properties of the human skin. *Serbian Journal of Dermatology and Venerology*, 2(4):131–136, 2010.
- [60] G. Gallego and A. Yezzi. A compact formula for the derivative of a 3-D rotation in exponential coordinates. *Journal of Mathematical Imaging and Vision*, 51(3):6, 2013.
- [61] T. Gambichler, R. Matip, G. Moussa, P. Altmeyer, and K. Hoffmann. In vivo data of epidermal thickness evaluated by optical coherence tomography: Effects of age, gender, skin type, and anatomic site. *Journal of Dermatological Science*, 44(3):145–152, 2006.
- [62] D. L. Gilblom, S. Keun, and P. Ventura. Operation and performance of a color image sensor with layered photodiodes. In *Proceedings of SPIE*, 5074(Figure 1):318–331, 2003.
- [63] A. R. Guazzi, M. Villarroel, M. C. Frise, P. a. Robbins, and L. Tarassenko. Non-contact Measurement of Oxygen Saturation With an RGB Camera. *Biomedical Optics Express*, 45(2008):3320–3338, 2015.
- [64] J. S. Haldane. A Lecture on the Symptoms, Causes, and Prevention of Anoxaemia (Insufficient Supply of Oxygen to the Tissues), and the Value of Oxygen in its Treatment. *British medical journal*, 2(3055):65–71, 1919.
- [65] C. D. Hanning, Alexander-Williams, and Jm. Pulse oximetry: a practical review. *Bmj*, 311(August):367–370, 1995.
- [66] D. Harmening. The red blood cell: Structure and function. In *Clinical hematology and fundamentals of hemostasis*, chapter 3, pages 54–70. F. A. Davis Co., Philadelphia, Pa, 4 edition, 1997.

- [67] K. Heidari, H. Hatamabadi, N. Ansarian, M. Alavi-Moghaddam, A. Amini, S. Safari, P. Darbandsar Mazandarani, and A. Vafaei. Correlation between capillary and arterial blood gas parameters in an ED. *The American journal of emergency medicine*, 31(2):326–9, 2013.
- [68] A. Honarmand and M. Safavi. Prediction of arterial blood gas values from arterialized earlobe blood gas values in patients treated with mechanical ventilation. *Indian journal of critical care medicine : peer-reviewed, official publication of Indian Society of Critical Care Medicine*, 12(3):96–101, 2008.
- [69] L. S. Howard, R. a. Barson, B. P. Howse, T. R. McGill, M. E. McIntyre, D. F. O'Connor, and P. a. Robbins. Chamber for controlling end-tidal gas tensions over sustained periods in humans. *Journal of applied physiology (Bethesda, Md. : 1985)*, 78(3):1088–1091, 1995.
- [70] a. Huch, R. Huch, and K. Meinzer. Eine schuelle. behizte Ptoberflachenelektrode zur kontinuierlichen uberwachung des POz beim menschen. Elektrodnaufbau und eigenschaften. *Stuttgart Proc Medizin-Technik*, page 26, 1972.
- [71] M. Huelsbusch and V. Blazek. <title>Contactless mapping of rhythmical phenomena in tissue perfusion using PPGI</title>. In *Medical Imaging 2002: Physiology and Function from Multidimensional Images*, pages 110–117, 2002.
- [72] K. Humphreys, T. Ward, and C. Markham. A CMOS Camera-Based Pulse Oximetry Imaging System. In *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 4:3494–7. Department of Electronic Engineering, National University of Ireland, Maynooth, Co. Kildare, Ireland., 2005.
- [73] K. Humphreys, T. Ward, and C. Markham. Noncontact simultaneous dual wavelength photoplethysmography: A further step toward noncontact pulse oximetry. *Review of Scientific Instruments*, 78(4):1–6, 2007.
- [74] N. a. Ibraheem, M. M. Hasan, R. Z. Khan, and P. K. Mishra. Understanding Color Models : A Review. *ARPJ Journal of Science and Technology*, 2(3):265–275, 2012.
- [75] M. Intaglietta, P. C. Johnson, and R. M. Winslow. Microvascular and tissue oxygen distribution. *Cardiovascular research*, 32(4):632–643, 1996.
- [76] N. G. Jablonski. The evolution of human skin and skin color. *Annual Review of Anthropology*, 33:585–623, 2004.
- [77] N. G. Jablonski and G. Chaplin. The evolution of human skin coloration. *Journal of Human Evolution*, 39(1):57–106, 2000.
- [78] S. L. Jacques. Origins of tissue optical properties in the UVA, visible, and NIR regions. *OSA TOPS on Advances in Optical Imaging and Photon Migration*, 2:364–371, 1996.

- [79] S. L. Jacques. Optical properties of biological tissues: a review. *Physics in Medicine and Biology*, 58(11):R37–R61, 2013.
- [80] D. Jakovels and J. Spigulis. RGB imaging device for mapping and monitoring of hemoglobin distribution in skin. *Lithuanian Journal of Physics*, 52(1):50–54, 2012.
- [81] M. J. Jones and J. M. Rehg. Statistical color models with application to skin detection. *International Journal of Computer Vision*, 46(1):81–96, 2002.
- [82] J. M. Kainerstorfer, M. Ehler, F. Amyot, M. Hassan, S. G. Demos, V. Chernomordik, C. K. Hitzenberger, A. H. Gandjbakhche, and J. D. Riley. Principal component model of multispectral data for near real-time skin chromophore mapping. *Journal of biomedical optics*, 15(4):046007, 2010.
- [83] J. M. Kainerstorfer, J. D. Riley, M. Ehler, L. Najafizadeh, F. Amyot, M. Hassan, R. Pursley, S. G. Demos, V. Chernomordik, M. Pircher, P. D. Smith, C. K. Hitzenberger, and A. H. Gandjbakhche. Quantitative principal component model for skin chromophore mapping using multi-spectral images and spatial priors. *Biomedical optics express*, 2(5):1040–58, 2011.
- [84] P. Kakumanu, S. Makrogiannis, and N. Bourbakis. A survey of skin-color modeling and detection methods. *Pattern Recognition*, 40(3):1106–1122, 2007.
- [85] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering-Transactions of the ASME*, 82(1):35–45, 1960.
- [86] A. a. Kamshilin, S. Miridonov, V. Teplov, R. Saarenheimo, and E. Nippolainen. Photoplethysmographic imaging of high spatial resolution., 2011.
- [87] A. a. Kamshilin, E. Nippolainen, I. S. Sidorov, P. V. Vasilev, N. P. Erofeev, N. P. Podolian, and R. V. Romashko. A new look at the essence of the imaging photoplethysmography. *Scientific reports*, 5:10494, 2015.
- [88] W. Karlen, J. Lim, J. M. Ansermino, G. Dumont, and C. Scheffer. Design challenges for camera oximetry on a mobile phone. *Proceedings of the 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012:2448–2451, 2012.
- [89] A. Keller. Noninvasive tissue oximetry for flap monitoring: an initial study. *Journal of Reconstructive Microsurgery*, 23(4):189–197, 2007.
- [90] C. Kim. *Principal Component Analysis*. Springer, 2nd edition, 2014.
- [91] G. Kirchhoff and R. Bunsen. Chemische Analyse durch Spectralbeobachtungen. *Annalen der Physik und Chemie*, 110(6):161–189, 1860.
- [92] R. E. Klabunde. *Cardiovascular physiology concepts*. Lippincott Williams & Wilkins, 2nd edition, 2004.

- [93] J. H. G. M. Klaessens, M. Nelisse, R. M. Verdaasdonk, and H. J. Noordmans. Multimodal tissue perfusion imaging using multi-spectral and thermographic imaging systems , applied on clinical data. In *Proceedings of Multimodal Biomedical Engineering VIII*, 8574:1–8, 2013.
- [94] L. Kong, Y. Zhao, L. Dong, Y. Jian, X. Jin, B. Li, Y. Feng, M. Liu, X. Liu, and H. Wu. Non-contact detection of oxygen saturation based on visible light imaging device using ambient light. *Optics Express*, 21(15):17464–17471, 2013.
- [95] N. Kourkoumelis and M. Tzaphlidou. Eye safety related to near infrared radiation exposure to biometric devices. *The Scientific World Journal*, 11:520–528, 2011.
- [96] J. Kovac, P. Peer, and F. Solina. Human skin color clustering for face detection. In *EUROCON 2003. Computer as a Tool. The IEEE Region 8*, 2:144–148 vol.2, 2003.
- [97] K. Kramer. Ein Verfahren zur fortlaufenden Messung des Sauerstoffgehaltes im stromenden Blut an uneröffneten Gefäßen. *Zeitschrift für Biologie*, 96(61-75):61–75, 1935.
- [98] A. Krishnaswamy and G. V. G. Baranoski. *A biophysically-based spectral model of light interaction with human skin*. Master thesis, University of Waterloo, 2004.
- [99] M. Kumar, A. Veeraraghavan, and A. Sabharwal. DistancePPG: Robust non-contact vital signs monitoring using a camera. *Biomedical Optics Express*, 6(5):1565, 2015.
- [100] C. D. Kurth, J. L. Steven, L. M. Montenegro, H. M. Watzman, J. W. Gaynor, T. L. Spray, and S. C. Nicolson. Cerebral oxygen saturation before congenital heart surgery. *The Annals of Thoracic Surgery*, 72(1):187–192, 2001.
- [101] T. F. Lectures, T. Principle, and L. Time. 26Optics : The Principle of Least Time. In *The Feynman Lectures on Physics, Volume I*, volume I, chapter 26, pages 1–14. 2015.
- [102] J. S. Lee, Y. Y. Jung, B. S. Kim, and S. J. Ko. An advanced video camera system with robust AF, AE, and AWB control. *IEEE Transactions on Consumer Electronics*, 47(3):694–699, 2001.
- [103] K. Z. Lee, P. C. Hung, and L. W. Tsai. Contact-free heart rate measurement using a camera, 2012.
- [104] R. Lee, M. M. Mathews-Roth, M. A. Pathak, and J. A. Parrish. The detection of carotenoid pigments in human skin. *Journal of Investigative Dermatology*, 64(3), 1975.
- [105] M. Lewandowska. Measuring pulse rate with a webcama non-contact method for evaluating cardiac activity. In *Computer Science ...*, pages 405–410, 2011.

- [106] R. Lienhart, A. Kuranov, and V. Pisarevsky. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. In B. Michaelis and G. Krell, editors, *Pattern Recognition SE - 39*, volume 2781 of *Lecture Notes in Computer Science*, pages 297–304. Springer Berlin Heidelberg, 2003.
- [107] E. Linfoot. *Principles of Optics*, volume 8. Cambridge University Press, 7 edition, 1961.
- [108] H. Liu, Y. Wang, and L. Wang. FPGA-based remote pulse rate detection using photoplethysmographic imaging, 2013.
- [109] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, pages 674–679, San Francisco, CA, USA, 1981. Morgan Kaufmann Publishers Inc.
- [110] R. Lukac and K. N. Plataniotis. Color filter arrays: design and performance analysis, 2005.
- [111] C. Lundsgaard. Studies on cyanosis : I. Primary causes of cyanosis. *The Journal of Experimental Medicine*, 30(3):259–269, 1919.
- [112] C. Lundsgaard, D. Van Slyke, and M. E. Abbott. Cyanosis. *Canadian Medical Association Journal*, 13(8):601–604, 1923.
- [113] D. B. MacLeod, L. I. Cortinez, J. C. Keifer, D. Cameron, D. R. Wright, W. D. White, E. W. Moretti, L. R. Radulescu, and J. Somma. The desaturation response time of finger pulse oximeters during mild hypothermia. *Anaesthesia*, 60(1):65–71, 2005.
- [114] P. Magnan. Detection of visible photons in CCD and CMOS: A comparative view. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 504(1-3):199–212, 2003.
- [115] L. T. Maloney and B. A. Wandell. Color constancy: a method for recovering surface spectral reflectance. *Journal of the Optical Society of America. A, Optics and image science*, 3(1):29–33, 1986.
- [116] K. Mannapperuma, B. D. Holton, P. J. Lesniewski, and J. C. Thomas. Performance limits of ICA-based heart rate identification techniques in imaging photoplethysmography. *Physiological Measurement*, 36(1):67–83, 2015.
- [117] P. D. Mannheim. The light-tissue interaction of pulse oximetry. *Anesthesia and Analgesia*, 105(6 Suppl):S10–7, 2007.
- [118] S. R. Marschner, S. H. Westin, E. P. Lafortune, K. E. Torrance, and D. P. Greenberg. Image-based BRDF measurement including human skin. In *10th Eurographics Workshop on Rendering*, pages 139–152, Aire-la-Ville, Switzerland, Switzerland, 1999. Eurographics Association.

- [119] F. J. Massey. The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951.
- [120] K. Matthes and F. Gross. Untersuchungen über die Absorption von rotem und ultraviolettem Licht durch kohlenoxydesattigtes und reduziertes Blut. *Archives of Experimental Pathology and Pharmacology*, 191:369–390, 1939.
- [121] P. J. Matts, P. J. Dykes, and R. Marks. The distribution of melanin in skin determined in vivo. *British Journal of Dermatology*, 156(4):620–628, 2007.
- [122] R. Maynard. Nunn’s applied respiratory physiology, 5th edition. *Occupational and environmental medicine*, 57(7):502, 2000.
- [123] D. McDuff, S. Gontarek, and R. W. Picard. Remote detection of photoplethysmographic systolic and diastolic peaks using a digital camera, 2014.
- [124] S. A. McLellan and T. S. Walsh. Oxygen delivery and haemoglobin. *Continuing Education in Anaesthesia, Critical Care & Pain*, 4(4):123–126, 2004.
- [125] I. V. Meglinski and S. J. Matcher. Quantitative assessment of skin layers absorption and skin reflectance spectra simulation in the visible and near-infrared spectral regions. *Physiological measurement*, 23(4):741–753, 2002.
- [126] Y. Mendelson and J. C. Kent. Variations in optical absorption spectra of adult and fetal hemoglobins and its effect on pulse oximetry. *IEEE transactions on bio-medical engineering*, 36(8):844–848, 1989.
- [127] Y. Mendelson and B. D. Ochs. Noninvasive pulse oximetry utilizing skin reflectance photoplethysmography. *IEEE Transactions on Biomedical Engineering*, 35(10):798–805, 1988.
- [128] G. A. Millikan. The oximeter, an instrument of measuring continuously the oxygen saturation of arterial blood in man. *Review of Scientific Instruments*, 13(1):432, 1942.
- [129] M. Misra, J. Stark, M. Dujovny, R. Widman, and J. I. Ausman. Transcranial cerebral oximetry in random normal subjects. *Neurological Research*, 20(2):137–141, 1998.
- [130] G. Morgagni. *De Sedibus et Causis Morborum Per Anatomen Indagatis Libri Quinque*. (Venice), 1761.
- [131] I. Mostafanezhad, O. Boric-Lubecke, V. Lubecke, and D. P. Mandic. Application of empirical mode decomposition in removing fidgiting interference in doppler radar life signs monitoring devices. *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009*, 2009:340–343, 2009.
- [132] J. Moyle. *Pulse Oximetry*. BMJ Books, 2nd edition, 2002.

- [133] R. Murphy, S. Thethy, S. Raby, J. Beckley, J. Terrace, C. Fiddler, M. Craig, and C. Robertson. Capillary blood gases in acute exacerbations of COPD. *Respiratory Medicine*, 100(4):682–686, 2006.
- [134] Y. Nagashima, Y. Yada, M. Hattori, and A. Sakai. Development of a new instrument to measure oxygen saturation and total hemoglobin volume in local skin by near-infrared spectroscopy and its clinical application. *International Journal of Biometeorology*, 44(1):11–19, 2000.
- [135] P. Nandy, A. D. Lucas, E. A. Gonzalez, and V. M. Hitchins. Efficacy of commercially available wipes for disinfection of pulse oximeter sensors. *American Journal of Infection Control*, 2015.
- [136] S. K. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection: Physical and geometrical perspectives, 1991.
- [137] L. Nicolai. Über Sichtbarmachung, Verlauf und chemische Kinetik der Oxyhemoglobinreduktion im lebenden Gewebe, besonders in der menschlichen Haut. *Archiv für die gesammte Physiologie des Menschen und der Thiere*, 229(1):372–389, 1932.
- [138] J. A. Nijboer and J. C. Dorlas. Comparison of plethysmograms taken from finger and pinna during anaesthesia. *British Journal of Anaesthesia*, 57(5):531–534, 1985.
- [139] J. A. Nijboer, J. C. Dorlas, and H. F. Mahieu. Photoelectric plethysmography—some fundamental aspects of the reflection and transmission method. *Clinical Physics and Physiological Measurement*, 2(3):205–215, 1981.
- [140] I. Nishidate, T. Maeda, K. Niizeki, and Y. Aizu. Estimation of Melanin and Hemoglobin Using Spectral Reflectance Images Reconstructed from a Digital RGB Image by the Wiener Estimation Method. *Sensors*, 13(6):7902, 2013.
- [141] I. Nishidate, K. Sasaoka, T. Yuasa, K. Niizeki, T. Maeda, and Y. Aizu. Visualizing of skin chromophore concentrations by use of RGB images. *Optics Letters*, 33(19):2263–2265, 2008.
- [142] I. Nishidate, N. Tanaka, T. Kawase, T. Maeda, T. Yuasa, Y. Aizu, T. Yuasa, and K. Niizeki. Noninvasive imaging of human skin hemodynamics using a digital red-green-blue camera. *Journal of Biomedical Optics*, 16(8):86012, 2011.
- [143] M. Nitzan and S. Engelberg. Three-wavelength technique for the measurement of oxygen saturation in arterial blood and in venous blood. *Journal of Biomedical Optics*, 14(2):24046, 2009.
- [144] M. Nitzan, S. Noach, E. Tobal, Y. Adar, Y. Miller, E. Shalom, and S. Engelberg. Calibration-free pulse oximetry based on two wavelengths in the infrared - a preliminary study. *Sensors*, 14(4):7420–7434, 2014.

- [145] M. Nitzan, A. Romem, and R. Koppel. Pulse oximetry: Fundamentals and technology update. *Medical Devices: Evidence and Research*, 7(1):231–239, 2014.
- [146] M. Noe. The Annals of Mathematical Statistics. *The Annals of Mathematical Statistics*, 1(1):1–2, 1930.
- [147] J. J. K. O Ruanaidh and W. J. Fitzgerald. *Numerical Bayesian Methods Applied to Signal Processing*. Springer-Verlag, New York, NY, USA, 1 edition, 1996.
- [148] C. Offelli and D. Petri. A frequency-domain procedure for accurate real-time signal parameter measurement. *Instrumentation and Measurement, IEEE Transactions on*, 39(2):363–368, 1990.
- [149] R. Ortega, C. J. Hansen, K. Elterman, and A. Woo. Pulse Oximetry. *New England Journal of Medicine*, 364(16):e33, 2011.
- [150] K. F. Palmer and D. Williams. Optical properties of water in the near infrared*. *Journal of the Optical Society of America*, 64(8):1107–1110, 1974.
- [151] G. Paré, M. Jaana, and C. Sicotte. Systematic review of home telemonitoring for chronic diseases: the evidence base. *Journal of the American Medical Informatics Association : JAMIA*, 14(3):269–277, 2007.
- [152] G. Parzinger. *Der transkutane PO₂ im hypo- und normovolamischen shock in beziehung gesetz zum kontinuierlich gemessenen art, und ven. PO₂m pH-werten und kardiozirkulatorischen parametern*. PhD thesis, Ludwig-Maximilians Universitat, Munich, 1977.
- [153] K. Pearson. Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58(347-352):240–242, 1895.
- [154] G. D. Perkins, D. F. McAuley, S. Giles, H. Routledge, and F. Gao. Do changes in pulse oximeter oxygen saturation predict equivalent changes in arterial oxygen saturation?, 2003.
- [155] S. M. Perry and W. P. Managhan. The prevalence of visible and/or occult blood on anesthesia and monitoring equipment. *Journal of the American Association of Nurse Anesthetists*, 69(1):44–48, 2001.
- [156] M. T. Petterson, V. L. Begnoche, and J. M. Graybeal. The effect of motion on pulse oximetry and its clinical significance. *Anesthesia and analgesia*, 105(6 Suppl):S78–84, 2007.
- [157] M.-Z. Poh, D. J. McDuff, and R. W. Picard. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 18(10):10762–10774, 2010.
- [158] M. Z. Poh, D. J. McDuff, and R. W. Picard. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE Transactions on Biomedical Engineering*, 58(1):7–11, 2011.

- [159] J. A. Pologe. Pulse oximetry: technical aspects of machine design. *International Anesthesiology Clinics*, 25(3):137–153, 1987.
- [160] R. Popa. *Between Necessity and Probability: Searching for the Definition and Origin of Life*. Springer-Verlag Berlin Heidelberg, 1 edition, 2004.
- [161] R. M. Pope and E. S. Fry. Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements. *Applied Optics*, 36(33):8710–8723, 1997.
- [162] F. H. Rampen, B. A. Fleuren, T. M. de Boo, and W. A. Lemmens. Unreliability of self-reported burning tendency and tanning ability. *Archives of Dermatology*, 124(6):885–888, 1988.
- [163] C. E. Rasmussen and C. K. I. Williams. *Gaussian processes for machine learning (adaptive computation and machine learning)*. The MIT Press, 2005.
- [164] K. A. Reddy, B. George, N. M. Mohan, and V. J. Kumar. A novel calibration-free method of measurement of oxygen saturation in arterial blood, 2009.
- [165] N. G. Roald. *Estimation of vital signs from ambient light non-contact photoplethysmography*. Masters thesis, Norwegian University of Science and Technology, 2013.
- [166] M. I. Rowe and G. Weinberg. Transcutaneous oxygen monitoring in shock and resuscitation. *Journal of Pediatric Surgery*, 14(6):773–778, 1979.
- [167] U. Rubins, R. Erts, and V. Nikiforovs. The blood perfusion mapping in the human skin by photoplethysmography imaging. In P. Bamidis and N. Pallikarakis, editors, *XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010*, 29:304–306. Springer Berlin Heidelberg, 2010.
- [168] U. Rubins, V. Upmalis, O. Rubenis, D. Jakovels, and J. Spigulis. Real-time photoplethysmography imaging system. In K. Dremstrup, S. Rees, and M. I. Jensen, editors, *15th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics (NBC 2011) SE - 46*, volume 34 of *IFMBE Proceedings*, pages 183–186. Springer Berlin Heidelberg, 2011.
- [169] Y. Saatci, R. D. Turner, and C. E. Rasmussen. Gaussian process change point models. In J. Fürnkranz and T. Joachims, editors, *Proceedings of the 27th International Conference on Machine Learning (ICML-10), June 21–24, 2010, Haifa, Israel*, pages 927–934. Omnipress, 2010.
- [170] I. S. Saidi, S. L. Jacques, and F. K. Tittel. Mie and Rayleigh modeling of visible-light scattering in neonatal skin. *Applied Optics*, 34(31):7410–7418, 1995.
- [171] O. D. Saugstad and D. Aune. Optimal oxygenation of extremely low birth weight infants: a meta-analysis and systematic review of the oxygen saturation target studies. *Neonatology*, 105(1):55–63, 2014.

- [172] L. Schallom, C. Sona, M. McSweeney, and J. Mazuski. Comparison of forehead and digit oximetry in surgical/trauma patients at risk for decreased peripheral perfusion. *Heart and Lung: Journal of Acute and Critical Care*, 36(3):188–194, 2007.
- [173] B. V. Scheer, A. Perel, and U. J. Pfeiffer. Clinical review: Complications and risk factors of peripheral arterial catheters used for haemodynamic monitoring in anaesthesia and intensive care medicine. *Critical Care*, 6(3):199–204, 2002.
- [174] T. W. L. Scheeren, P. Schober, and L. A. Schwarte. Monitoring tissue oxygenation by near infrared spectroscopy (NIRS): background and current applications. *Journal of Clinical Monitoring and Computing*, 26(4):279–287, 2012.
- [175] C. G. Scully, J. Lee, J. Meyer, A. M. Gorbach, D. Granquist-fraser, Y. Mendelson, and K. H. Chon. Physiological parameter monitoring from optical recordings with a mobile phone. *IEEE Transactions on Biomedical Engineering*, 59(2):2010–2013, 2012.
- [176] J. W. Severinghaus and P. B. Astrup. History of blood gas analysis. IV. Leland Clark’s oxygen electrode. *Journal of Clinical Monitoring*, 2(2):125–139, 1986.
- [177] J. W. Severinghaus and P. B. Astrup. History of blood gas analysis. V. Oxygen measurement. *Journal of Clinical Monitoring*, 2(3):174–189, 1986.
- [178] J. W. Severinghaus and Y. Honda. History of blood gas analysis. VII. Pulse oximetry. *Journal of Clinical Monitoring*, 3(2):135–138, 1987.
- [179] J. W. Severinghaus, K. H. Naifeh, and S. O. Koh. Errors in 14 pulse oximeters during profound hypoxia. *Journal of Clinical Monitoring*, 5(2):72–81, 1989.
- [180] J. W. Severinghaus. Takuo Aoyagi: Discovery of pulse oximetry. *Anesthesia and Analgesia*, 105(SUPPL. 6):S1–4, 2007.
- [181] J. W. Severinghaus and P. B. Astrup. History of blood gas analysis. VI. Oximetry. *Journal of Clinical Monitoring*, 2(4):270–288, 1986.
- [182] J. W. Severinghaus and a. F. Bradley. Electrodes for blood pO₂ and pCO₂ determination. *J Appl Physiol*, 13(3):515–520, 1958.
- [183] B. Shaanan. The iron-oxygen bond in human oxyhaemoglobin. *Nature*, 296(5858):683–684, 1982.
- [184] J. Shi and C. Tomasi. Good features to track, 1994.
- [185] S. Y. P. K. Shiao and C. N. Ou. Validation of oxygen saturation monitoring in neonates. *American Journal of Critical Care*, 16(2):168–178, 2007.
- [186] Y. Shimada, I. Yoshiya, N. Oka, and K. Hamaguri. Effects of multiple scattering and peripheral circulation on arterial oxygen saturation measured with a pulse-type oximeter. *Medical & biological engineering & computing*, 22(5):475–478, 1984.

- [187] R. P. Smith, J. Argod, J.-L. Pépin, and P. A. Lévy. Pulse transit time: an appraisal of potential clinical applications. *Thorax*, 54(5):452–457, 1999.
- [188] S. C. Sorensen. The chemical control of ventilation. *Acta Physiologica Scandinavica*, 361(SUPPL.):9–72, 1971.
- [189] C. Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101, 1904.
- [190] J. R. Squire. An instrument for measuring the quantity of blood and its degree of oxygenation in the web of the hand. *Clinical Science*, 4(1):331–339, 1940.
- [191] W. C. Stadie. The oxygen of the arterial and venous blood in pneumonia and its relation to cyanosis. *The Journal of Experimental Medicine*, 30(3):215–240, 1919.
- [192] I. D. Stephen, V. Coetzee, M. L. Smith, and D. I. Perrett. Skin blood perfusion and oxygenation colour affect perceived human health. *PLoS ONE*, 4(4):e5083, 2009.
- [193] K. G. Stewart and S. J. Rowbottom. Inaccuracy of pulse oximetry in patients with severe tricuspid regurgitation. *Anaesthesia*, 46(8):668–670, 1991.
- [194] S. Stoll. An investigation of the reflective properties of Tyvek papers and Tetratex PTFE Film. Technical report, Brookhaven National Laboratory, 1996.
- [195] M. Stücker, A. Struk, P. Altmeyer, M. Herde, H. Baumgärtl, and D. W. Lübbers. The cutaneous uptake of atmospheric oxygen contributes significantly to the oxygen supply of human dermis and epidermis. *The Journal of Physiology*, 538(Pt 3):985–994, 2002.
- [196] Y. Sun, C. Papin, V. Azorin-Peris, R. Kalawsky, S. Greenwald, and S. Hu. Use of ambient light in remote photoplethysmographic systems: comparison between a high-performance camera and a low-cost webcam. *Journal of Biomedical Optics*, 17(3):37005–37010, 2012.
- [197] K.-K. Sung, T. Poggio, H. A. Rowley, S. Baluja, and T. Kanade. CMU + MIT Frontal Face Dataset.
- [198] C. Takano and Y. Ohta. Heart rate measurement based on a time-lapse image. *Medical Engineering & Physics*, 29(8):853–857, 2007.
- [199] S. Takatani, C. Davies, N. Sakakibara, a. Zurick, E. Kraenzler, L. R. Golding, G. P. Noon, Y. Nose, M. E. DeBakey, S. Stoll, K. G. Stewart, S. J. Rowbottom, W. C. Stadie, S.-Y. P. K. Shiao, C.-N. Ou, J. W. Severinghaus, Y. Honda, P. B. Astrup, L. Schallom, C. Sona, M. McSweeney, J. Mazuski, O. D. Saugstad, D. Aune, G. D. Perkins, D. F. McAuley, S. Giles, H. Routledge, F. Gao, M. Nitzan, A. Romem, R. Koppel, J. a. Nijboer, J. C. Dorlas, J. B. Morgagni, Y. Mendelson, B. D. Ochs, C. Lundsgaard, D. Van Slyke, M. E. Abbott, L. S. Howard, R. a. Barson, B. P. Howse, T. R. McGill, M. E. McIntyre, D. F. O’Connor, P. a. Robbins, S. Fouzas, K. N.

- Priftis, M. B. Anthracopoulos, S.-J. Fearnley, J. M. Manners, J. H. J. Comroe, S. Botelho, S. J. Barker, J. Hyatt, N. K. Shah, Y. J. Kao, G. S. Agashe, J. Coakley, and P. D. Mannheim. De sedibus et causis morborum per antomen indagatis libri quinque. *Journal of clinical monitoring*, 105(3):87–88, 2007.
- [200] S. Takatani, C. Davies, N. Sakakibara, A. Zurick, E. Kraenzler, L. R. Golding, G. P. Noon, Y. Nose, and M. E. DeBakey. Experimental and clinical evaluation of a noninvasive reflectance pulse oximeter sensor. *Journal of Clinical Monitoring*, 8(4):257–266, 1992.
- [201] P. Talreja, G. B. Kasting, N. K. Kleene, W. L. Pickens, and T.-F. Wang. Visualization of the lipid barrier and measurement of lipid pathlength in human stratum corneum. *AAPS PharmSciTech*, 3(2), 2001.
- [202] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. a. Clifton, and C. Pugh. Non-contact video-based vital sign monitoring using ambient light and auto-regressive models. *Physiological Measurement*, 35(5):807–31, 2014.
- [203] H. E. Tasli, A. Gudi, and M. Den Uyl. Integrating remote PPG in facial expression analysis framework. In *Proceedings of the 16th International Conference on Multimodal Interaction*, pages 74–75, New York, NY, USA, 2014. ACM.
- [204] S. A. Taylor. CCD and CMOS imaging array technologies: technology review (Technical Report EPC-1998-106). Technical report, Xerox Research Centre Europe, 1998.
- [205] V. Teplov, E. Nippolainen, A. A. Makarenko, R. Giniatullin, and A. A. Kamshilin. Ambiguity of mapping the relative phase of blood pulsations. *Biomedical Optics Express*, 5(9):3123–3139, 2014.
- [206] A. J. Thody, E. M. Higgins, and K. Wakamatsu. Pheomelanin as well as eumelanin is present in human epidermis. *Journal of Investigative Medicine*, 97(2):340–344, 1991.
- [207] C. Thomas and A. B. Lumb. Physiology of haemoglobin. *Continuing Education in Anaesthesia, Critical Care & Pain*, 12(5):251–256, 2012.
- [208] D. J. Tobin. Biochemistry of human skin—our brain on the outside. *Chemical Society Reviews*, 35(1):52–67, 2006.
- [209] J. Toffaletti and W. G. Zijlstra. Misconceptions in reporting oxygen saturation. *Anesthesia and Analgesia*, 105(6 Suppl):S5–9, 2007.
- [210] C. Tomasi and T. Kanade. Detection and tracking of point features. Technical report, Carnegie Mellon University, 1991.
- [211] M. Townsley, R. McMillon, and A. Lee. *Regulation of tissue oxygenation*, volume 16. Morgan & Claypool Life Sciences, 1995.

- [212] K. K. Tremper and W. C. Shoemaker. Transcutaneous oxygen monitoring of critically ill adults, with and without low flow shock. *Critical Care Medicine*, 9(10):706–709, 1981.
- [213] K. K. Tremper, K. Waxman, and W. C. Shoemaker. Effects of hypoxia and shock on transcutaneous PO₂ values in dogs. *Critical Care Medicine*, 7(12):526–531, 1979.
- [214] K. K. Tremper. Transcutaneous PO₂ measurement. *Canadian Anaesthetists' Society Journal*, 31(6):664–677, 1984.
- [215] N. S. Trivedi, A. F. Ghouri, E. Lai, N. K. Shah, and S. J. Barker. Pulse oximeter performance during desaturation and resaturation: a comparison of seven models. *Journal of Clinical Anesthesia*, 9(3):184–188, 1997.
- [216] G. R. Tsouri and Z. Li. On the benefits of alternative color spaces for noncontact heart rate measurements using standard red-green-blue cameras. *Journal of Biomedical Optics*, 20(4):48002, 2015.
- [217] N. Tsumura, H. Haneishi, and Y. Miyake. Independent-component analysis of skin color image. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 16(9):2169–2176, 1999.
- [218] N. Tsumura, H. Haneishi, and Y. Miyake. Independent component analysis of spectral absorbance image in human skin. *Optical Review*, 7(6):479–482, 2000.
- [219] N. Tsumura, M. Kawabuchi, H. Haneishi, and Y. Miyake. Mapping pigmentation in human skin by multi- visible-spectral imaging by inverse optical scattering technique. *Journal of Imaging Science and Technology*, 45(5):444–450, 2001.
- [220] R. van den Boomgaard and R. van Balen. Methods for fast morphological image transforms using bitmapped binary images. *CVGIP: Graphical Models and Image Processing*, 54(3):252–258, 1992.
- [221] M. J. C. Van Gemert, S. L. Jacques, H. J. C. M. Sterenborg, and W. M. Star. Skin optics. *Biomedical Engineering, IEEE Transactions on*, 36(12):1146–1154, 1989.
- [222] E. J. Van Kampen and W. G. Zijlstra. Determination of hemoglobin and its derivatives. *Advances in Clinical Chemistry*, 8:141–187, 1965.
- [223] R. van Luijtelaar, W. Wang, S. Stuijk, and G. de Haan. Automatic ROI detection for camera-based pulse-rate measurement. In *Computer Vision-ACCV 2014 Workshops*, pages 360–374. Springer, 2014.
- [224] R. L. P. van Veen, H. J. C. M. Sterenborg, A. Pifferi, A. Torricelli, E. Chikoidze, and R. Cubeddu. Determination of visible near-IR absorption coefficients of mammalian fat using time- and spatially resolved diffuse reflectance and transmission spectroscopy. *Journal of Biomedical Optics*, 10(5):054004, 2005.

- [225] W. Verkruyse, L. O. Svaasand, and J. S. Nelson. Remote plethysmographic imaging using ambient light. *Optics Express*, 16(26):21434–21445, 2008.
- [226] V. Vezhnevets, V. Sazonov, and A. Andreeva. A survey on pixel-based skin color detection techniques. In *Proceedings of Graphicon 2003*, pages 85–92, 2003.
- [227] M. Villarroel, A. Guazzi, J. Jorge, S. Davis, P. Watkinson, G. Green, A. Shenvi, K. McCormick, and L. Tarassenko. Continuous non-contact vital sign monitoring in neonatal intensive care unit. *Healthcare Technology Letters*, 1(3):87–91(4), 2014.
- [228] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features, 2001.
- [229] P. Viola and M. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2004.
- [230] J. von Gerlach. Ueber das Hautathmen. *Archiv für Anatomie und Physiologie*, 431, 1851.
- [231] W. Wang, S. Stuijk, and G. de Haan. A novel algorithm for remote photoplethysmography: spatial subspace rotation. *Biomedical Engineering, IEEE Transactions on*, PP(99):1, 2015.
- [232] S. Watanabe. *Image sensors and signal processing for digital still cameras*. CRC Press, Inc., Boca Raton, FL, USA, 2006.
- [233] S. L. Watkin, S. A. Spencer, P. W. Dimmock, Y. A. Wickramasinghe, and P. Rolfe. A comparison of pulse oximetry and near infrared spectroscopy (NIRS) in the detection of hypoxaemia occurring with pauses in nasal airflow in neonates. *Journal of Clinical Monitoring and Computing*, 15(7-8):441–447, 1999.
- [234] L. Wei, Y. Tian, Y. Wang, T. Ebrahimi, and T. Huang. Automatic webcam-based human heart rate measurements using Laplacian eigenmap. In *Computer Vision ACCV 2012 SE - 22*, 7725:281–292. Springer, 2013.
- [235] S. Wichert, K. Fokianos, and K. Strimmer. Identifying periodically expressed transcripts in microarray time series data. *Bioinformatics*, 20(1):5–20, 2004.
- [236] F. P. Wieringa, F. Mastik, R. H. Boks, A. Visscher, A. J. J. C. Bogers, and A. F. W. der Steen. In vitro demonstration of an SpO₂-camera. In *Proc. Computers in Cardiology*, pages 749–751, 2007.
- [237] E. H. Wood and J. E. Geraci. Photoelectric determination of arterial oxygen saturation in man. *The Journal of Laboratory and Clinical Medicine*, 34(3):387–401, 1949.
- [238] H.-y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics (Proceedings SIGGRAPH 2012)*, 31(4), 2012.

- [239] A. Yadlapati, T. Grogan, D. Elashoff, and R. B. Kelly. Correlation of a novel noninvasive tissue oxygen saturation monitor to serum central venous oxygen saturation in pediatric patients with postoperative congenital cyanotic heart disease. *The journal of Extra-corporeal Technology*, 45(1):40–45, 2013.
- [240] F. Yamazaki, K. Takahara, R. Sone, and J. M. Johnson. Influence of hyperoxia on skin vasomotor control in normothermic and heat-stressed humans. *Journal of Applied Physiology*, 103(6):2026–2033, 2007.
- [241] M.-H. Yang, D. J. Kriegman, and N. Ahuja. Detecting faces in images: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1):34–58, 2002.
- [242] D. Yildizda, H. Yapiciolu, H. L. Yilmaz, and Y. Sertdemir. Correlation of simultaneously obtained capillary, venous, and arterial blood gases of patients in a paediatric intensive care unit. *Archives of Disease in Childhood*, 89(2):176–180, 2004.
- [243] G. H. Yont, E. A. Korhan, and L. Khorshid. Comparison of oxygen saturation values and measurement times by pulse oximetry in various parts of the body. *Applied Nursing Research*, 24(4):e39–43, 2011.
- [244] J. Zheng, S. Hu, V. Chouliaras, and R. Summers. Feasibility of Imaging Photoplethysmography. In *2008 International Conference on BioMedical Engineering and Informatics*, 2:72–75, 2008.
- [245] G. Zonios, J. Bykowski, and N. Kollias. Skin melanin, hemoglobin, and light scattering properties can be quantitatively assessed in vivo using diffuse reflectance spectroscopy. *Journal of Investigative Dermatology*, 117(6):1452–1457, 2001.