process it, is dependent on the clock rate. The clock controls the duty cycle. Duty cycle is defined as the fraction of time the output is high compared to the total time. When designing such subsystems we have to examine the duty cycle.

8.3.4.1 Clock generator and timer circuit. A 555 timer can be used to generate a *n*-minute timer, but isn't accurate enough for this kind of application. For more precise timing we usually use a signal derived from a crystal-controlled oscillator. This clock is stable but is too high in frequency to drive a processor interrupt input directly. Therefore, it is divided with an external counter device to an appropriate frequency for the interrupt input. Usually such a system contains counter devices such as the Intel 8253 or 8254, which can be programmed with instructions to divide an input frequency by any desired number.

The big advantage of using these devices is that you can load a count into them, and start them and stop them with instructions in a software program. Sometimes addition of a wait state may be needed along with this device to compensate for the delay due to the decoders and buffers on board.

We usually reset the circuit using simple resistors and capacitors, which are held low during power-on. This maintains the logic at a known state, while the crystal oscillator and the power supplies stabilize.

The timer circuit could control the following units on the subsystem:

1. Set the baud rate of the UART communication network.
2. Generate interrupts for controlling the display circuit, as these are usually multiplexed to avoid use of high current.
3. Audio frequency generator, for alarms.
4. Clock frequency for the notch filter, used to suppress the power line noise.
5. Synchronous circuit operation for pattern generator.

8.3.4.2 Watchdog timer circuit. This is a kind of fail-safe timer circuit, which turns the oximeter off if the microprocessor fails.

A counter controls the input to a D flip-flop, which is tied to a shutdown signal in the power supply. The counter is reset using a control signal from the microprocessor and a latch. Using some current-limiting protection, this signal is ac-coupled to the reset input of the counter. Therefore if the counter is not reset before the counter output goes high, the flip-flop gets set and the power supply is turned off.

8.3.4.3 UART. Within a MBS, data are transferred in parallel, because that is the fastest way to do it. Data are sent either *synchronously or asynchronously*. A UART (Universal Asynchronous Receiver Transmitter), is a device which can be programmed to do asynchronous communication.

8.3.5 Pattern generator

This is a multipurpose section which is primarily used to generate timing patterns used for synchronous detection gating, LED control, and for synchronizing the power supply. The heart of this system is the EPROM (Erasable Programmable ROM). Here preprogrammed bit patterns are stored and are cycled out through the counter, and are tapped off using certain address lines. Using the address the bit pattern is sent out and latched using an octal latch. There may be an additional

latch used to deglitch the system, in which the last byte is held until the counter increments itself to the next address and the next pattern is obtained. Various patterns within the EPROM are used to select the sampling speeds of the LEDs or the synchronous detector pulse, the calibration patterns and diagnostic timing.

8.4 ANALOG PROCESSING SYSTEM (NELLCOR®)

8.4.1 Analog signal flow

Signals obtained are usually weak and may have electromagnetic interference. These signals must be filtered and then amplified. Usually a 50 or 60 Hz low-pass (for example may be a 2nd order Butterworth) filter is used. The signal is then ac coupled to stages of amplifiers and depending on the kind of response, variable gain circuits can be designed. The aim here is to maximize the signal before it enters the detector circuit, where the IR and red signal are separated, so the signal-to-noise ratio (SNR) is also kept as high as possible.

8.4.2 Coding resistor, temperature sensor, and prefiltering

Before examining the analog signal flow path, it is necessary to mention how the MBS decides what compensation to use for those LEDs which do not have their peak wavelength at the desired value. As LEDs are manufactured in bulk and tested in a random fashion, the probes may not always have the LEDs with the desired wavelength. Therefore the MBS generates some compensation, in order to solve this problem. Probes must be calibrated. Pulse oximeter systems have a coding resistor in every probe connector. A current is provided to the probe which allows the MBS to determine the resistance of the coding resistor by measuring the voltage drop across it. Thus the particular combination of LED wavelengths can be determined. Following this the MBS can then make necessary adjustments to determine the oxygen saturation.

As the wavelengths of the LEDs depend on the temperatures, for accurate measurements the effects of the temperatures must also be known, for adequate compensation (Cheung *et al* 1989). A temperature sensor may be employed, whose signal along with that of the coding resistor is used to select the calibration curves which are to be employed for compensation.

Despite efforts to minimize ambient light interference via covers over the probes and sometimes red optical filters, interfering light does reach the photodiode. Light from the sun and the incandescent lamp are continuous. The fluorescent light source emits ac light. This may overload the signal produced by the photodiode in response to the light received.

8.4.3 Preamplifier

The photodiode generates a current proportional to the light incident upon it. The signal from the photodiode is received by a preamplifier. Figure 8.4 shows that the preamplifier consists of a differential current-to-voltage amplifier and a single ended output amplifier. A gain determination resistor converts the current flowing through it into voltage. But along with the current-to-voltage conversion, external interference is also amplified, making the true signal difficult to extract

from the resulting output. The differential amplifier produces positive and negative versions of the output. This dual signal is then passed via a single ended amplifier with unity gain, which results in a signal with twice the magnitude of that of the input. Due to the opposite signs of the outputs of the differential amplifiers, the external interference is canceled out. As the noise factor increases by a marginal factor the signal-to-noise ratio improves. The mixed signal is then fed into two sample-and-hold (S/H) circuits whose timings are controlled such that each circuit samples the signal input to the demodulator during the portion of the signal corresponding to the wavelength to which it responds.

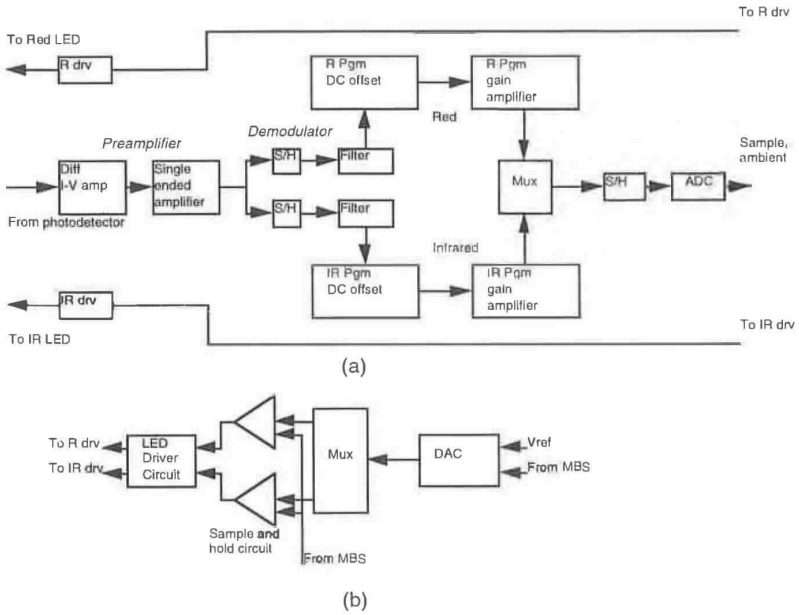

Figure 8.4 The analog signal flow path along with the signal demodulator and modulator circuit (from Cheung *et al* 1989).

8.4.4 Demodulator and filtering

This section splits the IR and the red signals from the mixed signal from the photodiode. The mixed signal is demultiplexed synchronously and steered depending on the type of signals present. The inputs to this circuit are the photodiode output and the timing or control signal from the MBS. The microprocessor along with the information stored in the EPROM calculates the time period each signal component is present in the photodiode output. Switching at the right time results in the two components getting separated. In order to eliminate the high-frequency switching noise, low-pass filters are provided. To optimize cost, size and accuracy, switched capacitor filters are used. These filters

cause the two signals (red and infrared) to be identical in gain and phase frequency response. In order to filter out the noise generated by this switched capacitor a second filter follows in the cascade to filter out the switching frequency noise. This stage is a high roll-off stage, allowing the first stage to be the dominating one, resulting in higher accuracy. Then using programmable DC offset eliminators and programmable gain amplifiers. the two signals are multiplexed along with other analog signals prior to being fed into an ADC. Offset amplifiers offset the signals by a small positive level. This ensures that the offsets caused by the chain of amplifiers do not allow the signal to be negative as this is the input to the ADC, and the ADC only accepts inputs from 0 to 10 V. Also sometimes the gain of the red or the IR channel may be greater than the other, and therefore the offset must be compensated accordingly.

8.4.5 DC offset elimination

To exploit the entire dynamic range of the ADC the two signals (red and IR) have to be processed further. Before discussing how this processing is done let us examine why this is done.

We know that the mixed signal consists of a pulsatile and a nonpulsatile component. The nonpulsatile component approximates the intensity of the light received at the photodiode when only the absorptive nonpulsatile component is present at the site (finger, earlobe, etc). This component is relatively constant over short periods, but due to probe position variation and physiological changes this component may vary significantly over large intervals. But as we analyze these signals in small interval windows, this is not a major problem. Figure 8.5 shows that this nonpulsatile component may be S_LOW, with the difference between S_HIGH and S_LOW being the varying pulsatile component, due to the arterial pulsations at the site. This pulsatile component is very small compared to the nonpulsatile component. Therefore great care must be taken when determining and eventually analyzing these values, as we desire the pulsatile component.

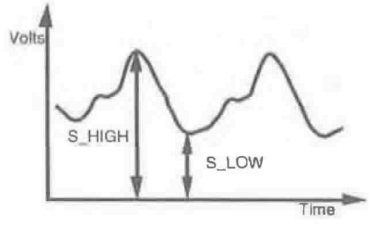

Figure 8.5 The nature of the signal transmission received by the photodiode circuit.

Amplifying and converting to digital form the substantial nonpulsatile component will use up most of the resolution of the ADC. Therefore in order to exploit the entire dynamic range we must eliminate this component, digitize it and later add it back to the pulsatile component.

For example, consider a ADC having an input range of 0 to 10 V. From figure 8.5 the AC may be 1% of the DC, let S_HIGH = 5.05 V and S_LOW = 5 V. For a 12-bit ADC, the resolution of this device is almost 2^{12}. This means that

the total signal is discretized into 4096 levels. Therefore from the above value of the pulsatile component (S_HIGH – S_LOW), we see that only 20 levels are utilized. Therefore if the nonpulsatile component is removed we can use all the 4096 levels, improving the resolution of the ADC.

Cheung *et al* (1989) discuss this concept of nonpulsatile component elimination and addition. The photodiode output contains both the nonpulsatile and the pulsatile component. The programmable subtractors (offset amplifier) remove a substantial offset portion of the nonpulsatile component of each signal and the programmable gain amplifiers increase the gain of the remaining signal for conversion by the ADC. A digital reconstruction of the original signal is then produced by the MBS, which through the use of digital feedback information removes the gain and adds the offset voltages back to the signal.

Feedback from the MBS to the analog and the digital sections of the board is required for maintaining the values for the offset subtraction voltage, gain, and driver currents at levels appropriate for the ADC. Therefore for proper operation, the MBS must continuously analyze and respond to the offset subtraction voltage, gain, and driver currents.

Figure 8.6 shows that thresholds L1 and L2 are slightly below and above the maximum positive and negative excursions L3 and L4 allowable for the ADC input and are established and monitored by the MBS at the ADC. When the signal at the input of the ADC or at the output of the ADC exceeds either of the thresholds L1 or L2, the LED driver currents are readjusted to increase or decrease the intensity of light impinging upon the photodiode. In this manner the ADC is protected from overdrives and the margins between L3, L1, and L2, L4 helps ensure this even for rapidly varying signals. An operable voltage margin for the ADC exists outside the threshold, allowing the ADC to continue operating while the appropriate feedback does the required adjustments.

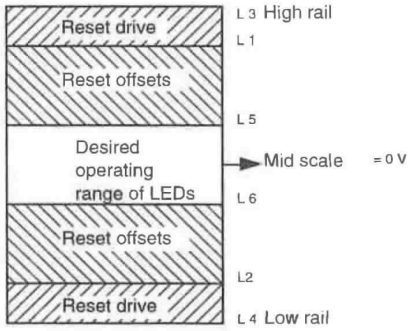

Figure 8.6 When the signal exceeds thresholds, the LED driver currents are readjusted to prevent overdriving the ADC (from Cheung *et al* 1989).

When the signal for the ADC exceeds the desired operating voltage threshold, L5 and L6, the MBS responds by signaling the programmable subtractor to increase or decrease the offset voltage being subtracted.

The instructions for the MBS program that controls this construction and reconstruction are stored in the erasable, programmable, read-only memory (EPROM).

8.4.6 Timing diagram (Nellcor®)

Figure 8.7 shows that the Nellcor pulse oximeter system uses a four state clock, or has a duty-cycle of 1/4, as compared to a Ohmeda system, where the duty-cycle is 1/3.

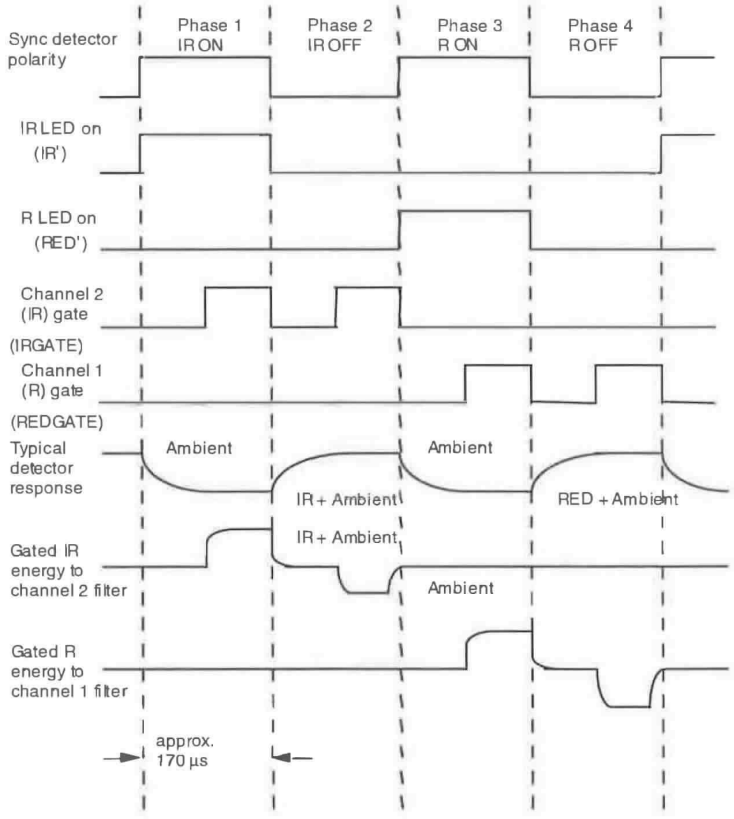

Figure 8.7 Timing diagram (reprinted with permission from Nellcor, Inc. ©Nellcor, Inc. 1989). Note that the typical detector response is inverted.

In the first quarter the IR LED is on and in the third quarter the R LED is on. In the second and the fourth quarters these LEDs are off. It is during this period that the ambient light measurements are done. The gate pulses are the sampling pulses applied to the input signal to separate out the R and the IR components from the input signal. The sample pulse during the OFF period of the respective LED is used to sample the ambient. The gradual rise or fall is due to the transients, which are smoothed out using low-pass filers. The ambient component is larger in the fourth quarter, compared to the value in the second.

Using suitable values for the gain in the programmable DC offset amplifiers we can eliminate this ambient component. The AC plus the DC components of the R and IR signals are digitized and sent to the MBS.

8.4.7 LED driver circuit

The need to drive both LEDs at different intensities requires analog switches that are used for gating the separate drive voltages. The factor that influences the amount of drive voltage necessary is the signal level from the photodiode and this value is set by the sample-and-hold section. The necessary control signals come from the pattern generator. The main purpose of this drive circuit is to convert this drive voltage to drive current.

Figure 8.8 shows the drive voltages VIR and VR and gating signals IR' and R'. These signals are used to multiplex the two signals back into one, and are fed into a voltage-to-current (V–I) converter such that the output of this V–I converter is a bipolar current signal, that is used to light up only one LED at a time. As the two LEDs are tied in a back to back configuration, this bipolar current drive ensures that only one LED is on at a time.

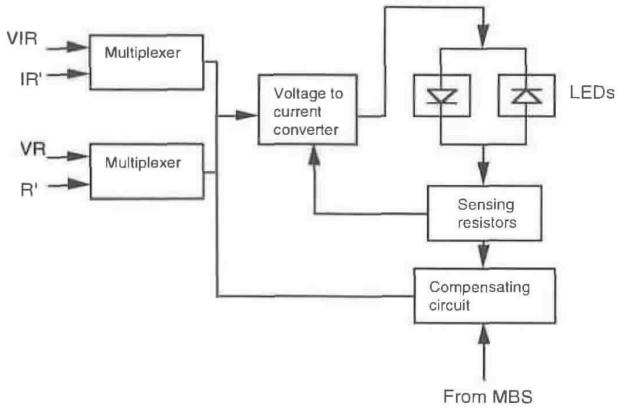

Figure 8.8 LED driver circuit with the sensor resistors to monitor and control the amount of current into the LED (adapted from Nellcor N-200® (Nellcor 1989)).

The drive current requires a control to convert the specified voltage to the proportional drive current. Within the voltage to current converter is an error amplifier that compares the voltage from the current sensing resistors with the specified voltage. There are two bridge networks with current boosters and drive and steering transistors which steer current around this conversion network. The drive output is connected to a pair of parallel back-to-back IR/R LEDs. The current through the LEDs is determined by a sensing resistor and fed back to the error amplifier to maintain a constant current proportional to the desired output voltage and to be independent of the other voltages present across the bridge circuit. Maximum LED current at 25% duty cycle is approximately 120 mA. The back-to-back configuration is such that when one LED is forward biased the

other is not. Chapter 5 describes the LED driver circuit used in the Nellcor® system.

8.4.8 Analog processing system (Ohmeda®)

The main block diagram indicated how different signals on board a pulse oximeter system flow, and showed the signal transfer from one major block (for e.g. ECG, probe, MBS, power supply, etc) to the other. This section will elaborate on the analog signal flow from the photodiode output until the analog signal is ready to drive the LEDs to make another measurement. See figure 8.9.

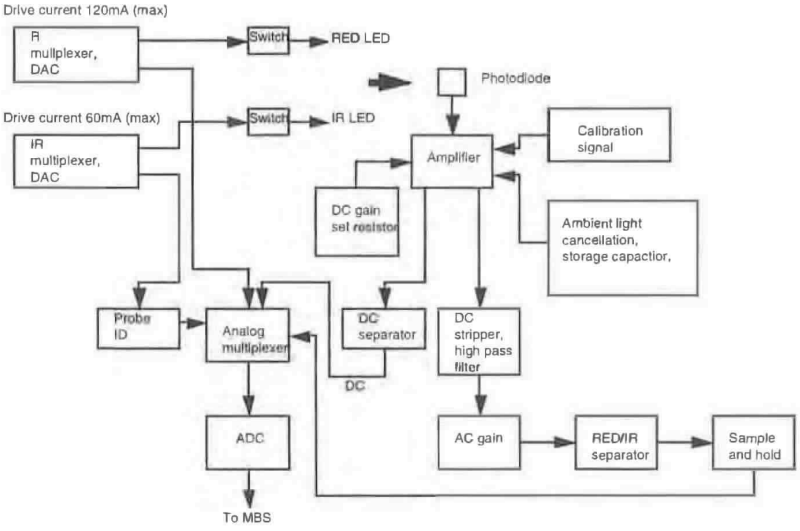

Figure 8.9 Functional block diagram of the pulse oximeter system showing all the main blocks involved in analog signal processing (adapted from Ohmeda 3740® (Ohmeda 1988)).

8.4.8.1 LED drive and monitor. The probe consists of the LEDs and the photodiode. The currents through the LEDs are controlled by a pair of multiplexers and switches and digital-to-analog converters (DACs). The maximum drive current is 120 mA through the R LED and 60 mA through the IR LED. The multiplexer and the switch turn the R and the IR LED drive on and off. The timing signal from the MBS controls the switches. The duty cycle of this timing signal is approximately 1/3 (Note that the duty cycle in devices from Nellcor® is 1/4). Therefore the subsequent hardware and analog and digital signal processing is different. The notable differences are in the multiplexers and sample and hold circuit. In the Ohmeda version, as the duty cycle is 1/3, first the R, then the IR, and finally the ambient component (measured when both the R and IR LEDs are off), are separated. In the Nellcor version as the duty cycle is 1/4 (see figure 8.7), the ambient component is measured twice.

The LED drive currents are monitored by switches and capacitors when both the R and IR LEDs are on individually and when both of them are off.

8.4.8.2 Calibration test signal. The signal received by the photodiode contains information on the AC and DC components of the pulsatile arterial blood flow measured by both the R and IR LEDs and also the ambient light component which is measured when both the R and IR LEDs are off. The calibration signal is a test signal injected into the signal path. The calibration signal is used to emulate the photodiode amplifier output which represents a known oxygen concentration and pulse rate of 150 to 210 beats per minute. The MBS checks the calibration of the oximeter by setting a test signal. This selects the calibration signal to be passed through the switch of the multiplexer in place of the photodiode amplifier output.

8.4.8.3 Ambient light cancellation. Ambient light cancellation is done to remove the effects of ambient light from the photodiode signal. A capacitor and a switch of a multiplexer are used to first charge up this capacitor to a voltage difference between the input signal and ground, when the input signal contains only the ambient component (R and IR LEDs are off). After this phase this voltage is subtracted from the input signal, now containing the R and IR components.

8.4.8.4 DC gain set resistor. The DC gain of the input signal is set under the control of the MBS. One resistor from a resistor bank is selected and along with another fixed resistor is used to set the gain of the amplifier.

8.4.8.5 DC separator. This block separates the DC components of the R and IR signals. This section consists of multiplexers and low-pass filters. The switches, controlled by the MBS, allow the red or the infrared component of the signal to pass through the low-pass filter. A set of amplifiers amplifies this DC before it is sent into the analog-to-digital (ADC) converter for conversion before being fed into the MBS. As a result of this stage we obtain the DC components of the R and the IR signals.

8.4.8.6 Low-pass filtering and DC stripping. A switched capacitor low-pass filter is used in this section. Since the DC components have been separated and measured previously, it is not necessary to filter during the ambient time. The R and IR components are low-pass filtered during the R and IR. time.

DC stripping is used to separate the pulsatile component from the signal. The low-passed signal is sent via a high-pass switching filter, and depending on the R and IR LED times, the pulsatile or the AC components of the R and IR signals are generated. This stage yields the AC components of the R and the IR signals.

8.4.8.7 Red/infrared separator. Multiplexers separate the red and infrared pulsatile signals into two independent channels. Low-pass filters are also used to smooth the separated signals. To compensate for the gain differences between the red and infrared signal paths, the gain of the infrared amplifier is adjustable by potentiometer.

8.4.8.8 Sample and hold circuits. Sample and hold circuits sample the red and infrared pulsatile signals simultaneously so that they can be measured by the ADC. An additional sampling signal controls the timing of the sampling of the pulsatile components at a rate synchronous to the power line frequency. This sampling frequency helps to suppress interference generated from sources connected to the line power.

8.4.8.9 Probe identification. This is the voltage generated by passing a known amount of current through the probe coding resistor to identify the wavelengths associated with the probe. This signal is digitized, compared to a lookup table in the MBS's memory, and the associated wavelength values are used for further processing.

An analog multiplexer is used to choose one of the many inputs and feed it to the ADC. The MBS for the Ohmeda system is similar to the one used in Nellcor, but uses Zilog's Z-8002. Motion artifact elimination using the R wave (ECG synchronization), as seen in Nellcor N-200 is not present in this system.

8.4.8.10 Timing diagram. In the Ohmeda Biox 3700® oximeter the LED on-off cycle is repeated at a rate of 480 Hz (figure 8.10). This cycling allows the oximeter to know which LED is on at any instant of time (Pologe 1987). The duty cycle in this system is 1/3. The red LED is on for the first 1/3 of the cycle, the infrared LED is on for the second 1/3 and both LEDs are off for the third 1/3, allowing for the ambient light measurements. This kind of measurement of ambient light is necessary so that it can be subtracted from the levels obtained when the LEDs are on.

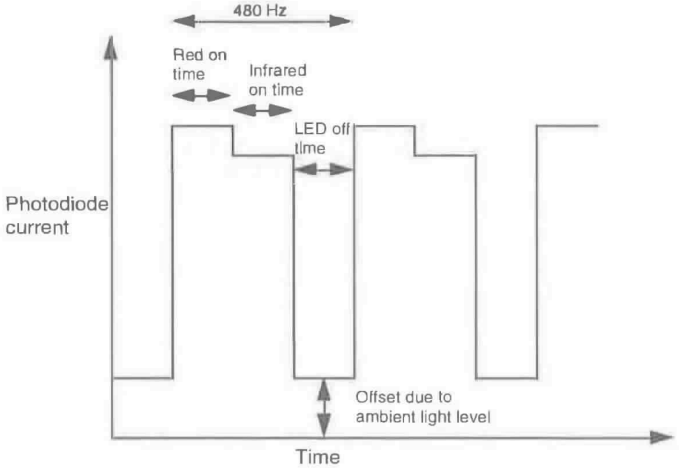

Figure 8.10 Output of the photodiode of the pulse oximeter system (adapted from Ohmeda 3700® (Ohmeda 1988)).

8.5 ECG SECTION

Pulse oximeters use the ECG to eliminate disturbances caused by motion artifacts and ambient light. There is a time delay between the electrical and the mechanical activity of the heart. When an ECG QRS electrical complex is detected, a mechanical pulse will be detected at the sensor after a transit delay of about 100 ms. This delay depends on factors such as the heart rate, the compliance of the

arteries, and the distance of the probe from the heart. The pulse oximeter computes this delay and stores it, and an average delay is generated after a few pulses. This average delay is used to establish a time window, during which the pulse is expected at the probe site. So if a pulse is received within this time window, it is treated as real and is processed. Any pulse arriving outside this window is simply rejected. Note that the time averaging and the time window are constantly updated to account for the patient's physiological changes.

8.5.1 Active filters

Figure 8.11 shows that the ECG signal from the patient has to pass through a series of filters before it is used for processing. Usually these filter stages provide gain, as the signal level received is very small. The most commonly used filters are as follows.

1. A low-pass filter with a corner frequency of 40 Hz.
2. A switched capacitor notch filter at the power line frequency. The capacitor switching frequency is determined by the timer pulse, which is in turn set by the microprocessor. The microprocessor along with its associated circuitry determines the power line frequency, and accordingly sets the notch frequency.
3. A second 40 Hz low-pass filter may be used to filter out the transients generated by the capacitor switching.
4. A high-pass filter, with a corner frequency of 0.5 Hz, is used for the lower end of the range. This filter has a substantial gain and has a long time constant. The reset condition discharges this capacitor. The most common situations desiring a reset are the lead fall off condition, muscle contraction under the electrode, or a sudden shift in the baseline of the ECG, due to the already high combined gain due to the front end section and the filters preceding this stage.

When pulse oximeters are used in electromagnetic environments (MRI 1992), such as magnetic resonance imaging (MRI), special care has to be taken, as EMI interferences are quite disturbing for pulse oximeters. Probes and connectors are shielded, using faraday cages, and additional EMI elimination filters are incorporated in the pulse oximeter (see chapter 11).

8.5.2 Offset amplifiers

Analog-to-digital converters have a specified input dynamic range for obtaining the maximum digitized output. Usually these are in the positive range, from 0 to 5 or 0 to 10 V. Therefore an amplifier that can offset the analog signals to a value beyond 0 V and convert its peak value to 5 to 10 V is needed. For example if there is a signal from –0.7 V to +6.7 V and an ADC with dynamic range of 0 to 5 V, the offset amplifier will convert this range to 0 to 5 V. Then we can make use of the entire resolution of the ADC.

8.5.3 Detached lead indicator

ECG signals are sensed by the electrodes placed on the body and the signals are transferred from the site to the pulse oximeter via leads. If the electrode falls off

from the surface of the body, the pulse oximeter's front end display must indicate this. The indicator system consists of a voltage comparator, absolute value circuit and a latching flip flop. This stage examines the ECG signal at the input to the switched capacitor notch filter (60 Hz), after it has passed through the low-pass stage preceding it. There is a biasing resistor network that drives the ECG signal to either ±15 V, if one or more ECG leads are detached from the patient's body. If the signal rails to −15 V, it is converted to positive voltage by a level shifter amplifier. Using this signal, the data are latched in a flip flop and the processor is notified that a lead has fallen off. After the processor recognizes this, it resets this latching flip flop so it is now ready to sense any other fall off.

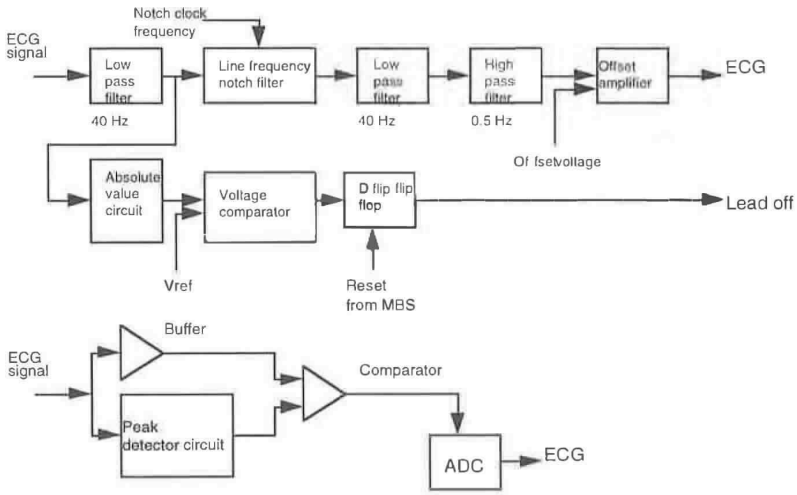

Figure 8.11 ECG signal processing section along with the lead fall off indicator and peak detector unit (adapted from Nellcor N-3000® (Nellcor 1991)).

8.5.4 Power line frequency sensing

Electric devices usually have the main power ac signal transformed to a root mean square value for power line analysis. A voltage comparator is used to generate a signal that interrupts the microprocessor at the frequency of the ac power line. This signal is then used to set the notch filter at the line frequency to eliminate the line frequencies. Most devices have provision for a 50 Hz or 60 Hz line.

8.5.5 ECG output

This section is used to generate pulses to synchronize the processor with the R-wave arrival. There is a peak follower circuit that stores the peak R-wave pulse in a slowly decaying fashion. This is employed to ensure that the capacitor doesn't discharge before the next R wave arrives. An adjustable threshold is provided for sensing each R-wave peak. This parameter is set by the processor, which in turn

is influenced by many external parameters. A voltage comparator produces a high output whenever the ECG input exceeds the adjustable threshold determined by the previous R-wave peak.

8.6 SIGNAL CONVERSION

The signal conversion unit consists of an ADC or DAC. Signals have dc offsets subtracted and even amplified before processing. This enables us to extract signals having low modulation and riding on a high DC, and this helps improve the response time of the system.

8.6.1 Analog-to-digital conversion technique

Analog-to-digital conversion on the processor board is accomplished by using a sample-and-hold circuit, which holds a voltage until it is sampled by a routine written in the memory of the processor. Both software and hardware play a important role in the conversion.

Figure 8.12 shows that first the processor writes to an analog multiplexer to select one of its several analog inputs that desire digital conversion. These signals could be the demultiplexed and filtered IR or the red photodiode channel signal, the filtered ECG waveform, or filtered voltage from the coding resistor. The selected signal is first latched and the analog circuitry is notified of the amplitude level. This helps to set the gain of the programmable amplifiers, so that the voltages at the input of the ADC do not exceed the maximum range. This ensures that the entire range of the ADC is used. A sample is generated by the processor to trigger the sample-and-hold circuit. This causes the sample-and-hold chip to hold the current channel for conversion. The processor begins executing the appropriate software for conversion. Usually the conversion routine adopted is the successive approximation routine (SAR). In this system, the SAR performs a binary check, by setting up a voltage using the DAC, which is compared to the currently held voltage signal via a comparator. The result of this comparison is polled by the processor. This process continues till the least significant bit is converted. Usually a 12-bit conversion is done in approximately 100 μs.

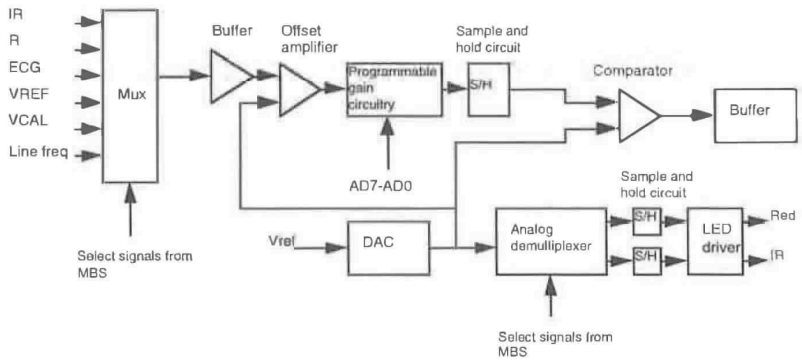

Figure 8.12 Generic analog-to-digital conversion circuit.

8.6.2 Digital-to-analog conversion

This is used to aid in digitizing the voltage at the ADC input, using the SAR. The DAC also has the following important application. The DACs have analog sample-and-hold circuits which are made using the analog demultiplexers and a series of variable gain amplifiers. Usually a DAC is used to update and store signals like the IR/red LED brightness control, or the speaker volume control. The analog signals are routed using the microprocessor. The processor puts out the analog voltage to the analog demultiplexer using the DAC. The processor selects which output will be written, using the address lines.

8.6.3 Sample-and-hold circuit

The analog conversion circuitry contains an n-bit DAC, a 1-to-8 bus-compatible analog multiplexer, switches for selecting the full scale analog output voltage range, and the analog sample-and-hold network. The DAC puts out an address of the task to be sampled and this information is decoded by the analog multiplexer and the desired sample-and-hold circuit is selected.

The sample-and-hold circuit is made up of storage capacitors and unity gain amplifiers. These amplifiers are usually the FET high-input-impedance devices. These circuits are protected via zener diodes that are needed to eliminate the short lived voltage transients. As we are driving high capacitive loads we need resistances to minimize these transients.

8.7 TIMING AND CONTROL

The time required to access a memory or an external device is as important as controling the various instruction executions within a microprocessor subsystem. The microprocessor adopts two techniques for the timing control. These are the polled processor I/O signal and the processor interrupts.

8.7.1 Polling and interrupt

In the polling technique the microprocessor has a signal that constantly polls or scans the various input waiting for a response. As soon as a valid signal is received at the polled input, the microprocessor starts the required task execution.

In the interrupt technique, the various chips and the inputs on the system are tied to the microprocessor via dedicated input lines. These lines are asserted high when a device requests help from the microprocessor. The microprocessor interrupts its current functionality and starts executing the interrupt routine. In the case of important activities these interrupt lines could be masked. Inputs may be provided with priority interrupt levels. When two or more interrupts are initiated at the same time, the higher priority interrupt performs first. Nested processing is done, in which within one interrupt execution another interrupt request can be answered.

For example, while the oxygen saturation is being measured along with the ECG signal, if there is a lead fall off situation, in which the device loses the ECG signal, a number of processes have to be initiated. First of all, the program in progress, calculating the oxygen saturation using the calibration tables has to be

interrupted, as the R wave detected is no longer valid, and therefore the software used to eliminate motion artifacts will have to be terminated. An interrupt to the display/audio driver will start a routine to display the lead fall off information and generate some alarms. After this problem has been fixed, another interrupt would trigger the operation to resume. During interrupt routine execution the MBS stalls for a while until the process generating the interrupt has been serviced.

Also, if the physician wants to refer to the pulse rate of the patient recorded a few minutes back, the interrupt raised will cause the current routine to branch, retrieve the data from the memory and continue with the recording. Usually while user-triggered interrupts are generated, the main routine continues with the measurements and has this raised interrupt serviced in parallel.

8.8 POWER SUPPLY

The power supplies present on most boards are switched mode power supplies (SMPS). A SMPS-based power supply is either in the flyback or the flyforward converter mode. Figure 8.13 shows the power supply present on the Nellcor N-200®, which contains switching power supplies in flyback converter configuration. These power supplies are capable of generating low voltages at high currents. The supply is capable of providing 2 A at +5 V and 100 mA at ±18 V.

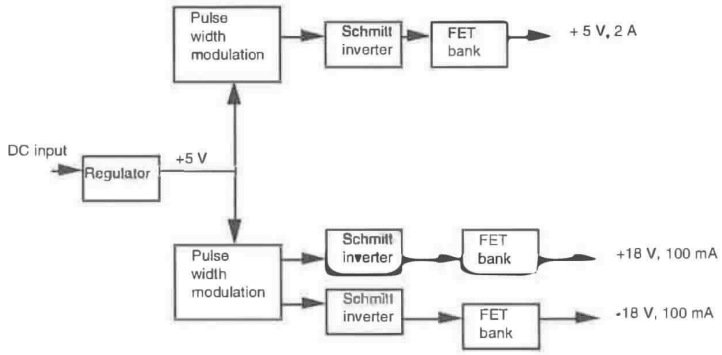

Figure 8.13 Basic power supply block diagram (adapted from Nellcor N-200® (Nellcor 1989)).

The essence of a SMPS supply is the pulse width modulator (PWM). In figure 8.13 the two PWMs control the 5 V and the ±18 V supply. The PWM senses the dc voltages at their inputs and controls the pulse width at the gates of switching FETs. A voltage regulator provides reference voltages for the two PWMs.

Field effect transistors are characterized by the rise and fall times of their drain currents. As the gates of the FETs are slightly capacitive, there is a need to minimize the rise and fall times of the drain current. Schmitt inverters are present to provide low impedance active current drive to these capacitive gates.

8.8.1 Recharging

Battery charger circuits are necessary to charge up a battery in case of a power line failure. In such an application when the main system is on line a battery charging circuit charges up a battery making use of the ac line voltage. In case of emergencies, because of a line failure, this system is set into the battery operated mode. However there is only a limited usage time available. Moreover the system becomes more bulky.

Figure 8.14 shows that ac power is taken from one of the secondaries of the transformers. It is rectified using a diode bridge arrangement (full wave rectifier) and filtered using a capacitor, to provide a positive voltage to the voltage regulator. Current limiting action is present via the use of a current-sensing resistor and a set of current-limiting transistors. Potentiometers are provided to trim the battery charging voltage. In order to avoid back discharge from the battery when the ac power is removed, a diode is present. Keeping in mind the efficiency of a power system, the expected voltage is 85% of the voltage provided by the battery charging circuit.

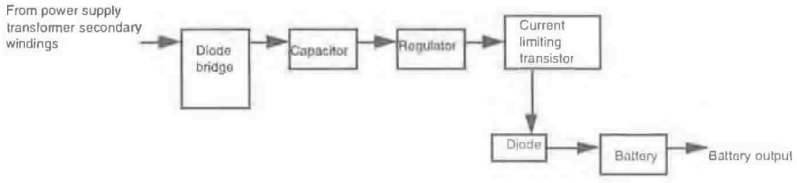

Figure 8.14 Battery charger block diagram (adapted from Nellcor N-200® (Nellcor 1989)).

8.9 ALARMS

When using pulse oximeters in critical applications, alarms are essential to give an indication to the physician that something is wrong. These alarms have to be in both audio and visual form. Comparators, power amplifiers, drivers and speakers constitute the audio alarm section. LCD bar graphs and blinkers are used for the visual section. Certain guidelines have been formulated by standardizing agencies such as the American Society for Testing and Materials (ASTM) regarding the color of the indicator, frequency of the indicator and the tone, audio level etc. Also the signals that need to be treated as emergency signal are classified (pulse rate, detached lead, etc).

8.10 STORAGE

Data concerning oxygen saturation and pulse rate can be collected and stored in memory. This may be used in the future to train the pulse oximeter system, using neural networks and artificial intelligence to generate control signals.

Memories are selected using address lines and data lines are used to load and unload data from them. In order to make the most efficient use of this storage

mechanism, some care has to be taken while designing the memory system. When no power is applied to the memory system there is danger of losing data.

8.11 FRONT END DISPLAY

This section includes the display terminal on the front end of the pulse oximeter. Liquid crystal displays (LCD) or LED displays are used depending on the clarity, resolution, power consumption, and even aesthetics. Push buttons in the form of feather touch buttons or conventional switches are provided. Interface points, alarm indicators, and other important features are also displayed.

8.11.1 Front end driver circuit

Figure 8.15 shows that the driver circuit consists of two major driving techniques, one for the digit and bargraph and the other for lightbar and other front panel indicators.

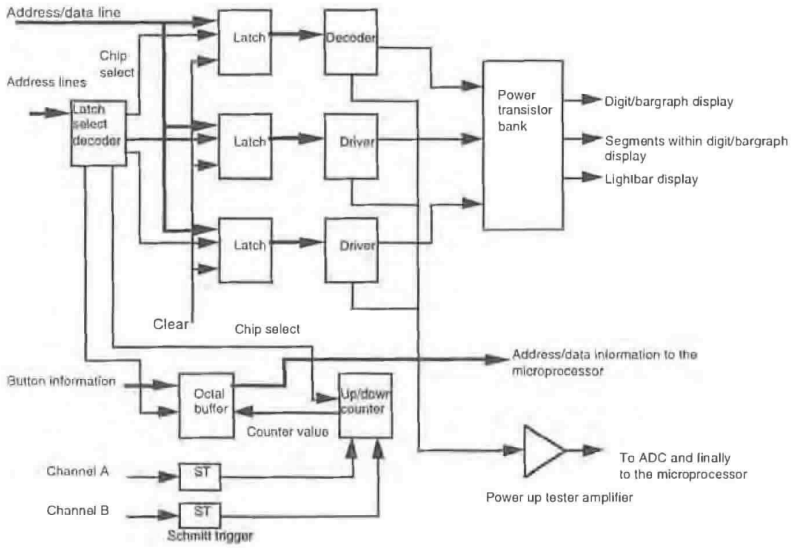

Figure 8.15 Generic display driver circuit.

In order to clear the front panel display, a reset circuit consisting of capacitors and resistors is used. During reset, these components generate a small duration reset pulse that is sent to all the latches on the driver board. This reset pulse clears all the front panel display elements.

Latches and decoders are used to generate the signals required to display information on the display elements. A set of power transistors are used to

generate the drive current required to turn on the display elements. A chip select decoder is used to select the latch–decoder combination depending on the type of display needed. In the circuit layout, signals are required for the digit/bargraph display, to select particular segments within the digit/bargraph display and signals for light bar display. The latches generate the information to be displayed via address information that comes from the microprocessor. After the microprocessor-based system has decided what is to be displayed, address information is sent to Character Generator ROMs (CG-ROMs) or Dynamic Display RAMs (DD-RAMs) which generate the digit/display information. In these devices, bit information pertaining to a particular character is stored at a specific address location. Depending on the address at the input, the required character is generated.

8.11.2 Front panel control

The chip select decoder is used to select the octal buffer, which reads in inputs from the buttons on the front panel and the up/down counter which reads in the control knob rotation information which is relayed through it.

The control knob consists of a two-channel optical chopper, with the two channels mechanically 90 degrees out of phase with each other, and a dual channel optical slot detector. There are two Schmitt triggers, one per channel, to eliminate any transients present and to clean up the signal. The two channels are used to send control signals to the up/down counter. Depending on the direction in which the knob is turned, either the up or the down mode is selected. Channel A provides the clocking pulses for the counter and channel B provides the direction control, whether it is up or down. The processor reads the counter output to determine a change in the up/down mode of the counter. It then adds the count to the accumulated count. The processor then resets the counter.

8.11.3 Power up display tests

When we power up the system for the first time the system runs a few initialization tests. The software tests run are discussed in detail in chapter 9. The primary concern is to ensure that all the display elements are operational. We therefore have a power up tester amplifier which senses the power return line from the driver ICs by monitoring a voltage developed across a resistor. The driver ICs are used to boost the drive current into the segments of the digital displays. This is amplified and given to the ADC. The processor uses this signal during start up to check whether the display is faulty.

8.12 SPEAKERS

The speaker is an inductive load needing a positive and a negative signal. Figure 8.16 shows that currents to these two inputs are controlled by two different paths. Depending on the address/data information the demultiplexer generates many signals like the VRED, VIR and the volume control signal. A sample-and-hold circuit is used to hold this signal. This signal is then passed via a series of power transistors to boost the current flowing into the speaker.

A timer and counter chip generates a count using certain address/data information and temporarily saves it into a buffer. This tone signal is used to

control a FET switch which alternately connects or disconnects the speakers negative input to ground. The frequency of the tone signal (determined by the timer/counter chip) determines the pitch of the sound produced. A capacitor is present to smoothen the sound. A diode is also present to suppress any transients from the inductive load.

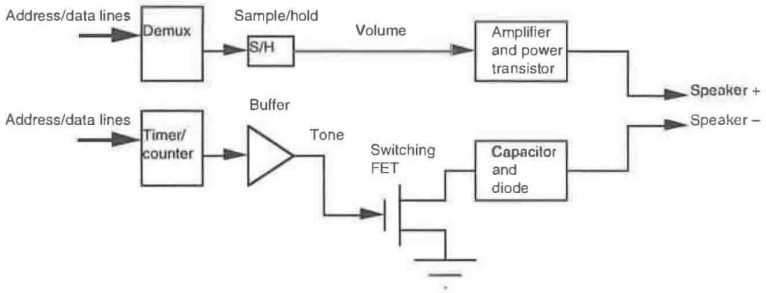

Figure 8.16 Speaker driver block diagram (adapted from Nellcor N-200® (Nellcor 1989)).

REFERENCES

Cheung P W, Gauglitz K F, Hunsaker S W, Prosser S J, Wagner D O and Smith R E 1989 Apparatus for the automatic calibration of signals employed in oximetry *US patent 5,259,381*
Corenman J E, Stone R T, Boross A, Briggs D A and Goodmann D E 1990 Method and apparatus for detecting optical pulses *US patent 4,934,372*
MRI 1992 *Service Manual, model 3500 MR-compatible oximeter* (Bay Shore, NY: MRI)
Nellcor 1989 *Service Manual, N-200 Pulse Oximeter* (Pleasanton, CA: Nellcor)
Nellcor 1991 *Service Manual, N-3000 Pulse Oximeter* (Pleasanton, CA: Nellcor)
New W Jr 1987 Pulse oximeter monitor *US patent 4,653,498*
Nielsen L L 1983 Multi-wavelength incremental absorbance oximeter *US patent 4,167,331*
Ohmeda 1988 *Service Manual, Model 3740 Pulse Oximeter* (Louisville, CO: Ohmeda)
Pologe J A 1987 Pulse oximetry: Technical aspects of machine design *Int. Anesth. Clinics* **25** (3) 137–53
Protocol 1991 *Service Manual* (Beaverton, OR: Protocol)
Sobusiak A C and Wiczynski G 1995 Specificity of SIF co-operating with optoelectronic sensor in pulse oximeter system *Proc. SPIE* **2634**
Wilber S A 1985 Blood constituent measuring device *US patent 4,407,290*
Yoshiya I, Shimada Y and Tanaka K 1980 Spectrophotometric monitoring of arterial oxygen saturation in the fingertip *Med. Biol. Eng. Comput.* **18** 27

INSTRUCTIONAL OBJECTIVES

8.1 Sketch the block diagram of the microprocessor subsystem, and highlight at least three main features that you think are vital for optimum operation of this system.
8.2 Explain the signal flow in the pulse oximeter from the photodiode to the front-end display.
8.3 Explain the kind of circuit protection associated with a patient module.
8.4 Explain how communication is established between the various chips on the microprocessor based system.
8.5 Describe the timing control involved in the microprocessor-based system.

8.6 Explain the operation of the synchronous detector and the demultiplexer in the pulse oximeter system.
8.7 Mention the need for active amplifiers and low-pass filters.
8.8 Explain the analog-to-digital conversion action involved in the pulse oximeter.
8.9 Explain the function of the pattern generator.
8.10 It is decided to improve the resolution of the ADC. List the steps you will take to improve the existing system. Explain how this will affect the system operation.
8.11 Explain the motivation for subtraction of the DC-level in the photodiode signal before the ADC.
8.12 Describe the components of an input module of a pulse oximeter.

CHAPTER 9

SIGNAL PROCESSING ALGORITHMS

Surekha Palreddy

Pulse oximeters measure and display the oxygen saturation of hemoglobin in arterial blood, volume of individual blood pulsations supplying the tissue, and the heart rate. These devices shine light through the tissue that is perfused with blood such as a finger, an ear, the nose or the scalp, and photoelectrically sense the transmittance of the light in the tissue. The amount of light that is transmitted is recorded as an electric signal. The signal is then processed using several signal processing algorithms to estimate the arterial oxygen saturation reliably in the presence of motion and other artifacts. Signal-processing algorithms implemented both in hardware and software play a major role in transforming the signals that are collected by the sensors and extracting useful information. In this chapter, the signal-processing to calculate S_aO_2 is discussed and ECG synchronization algorithms that enhance the reliability of S_aO_2 estimation and improve the signal-to-noise ratio are discussed. Commercial pulse oximeters use various algorithms for ECG synchronization. Some of these algorithms are discussed with reference to commercially available pulse oximeters such as from Nellcor® and Criticare®.

9.1 SOURCES OF ERRORS

The three general sources of errors dealt with by signal-processing algorithms are the *motion artifact*, *reduced saturation levels* (<80%) and *low perfusion levels* (Goodman and Corenman 1990). The motion artifact is a major problem that is usually due to the patient's muscle movement proximate to the oximeter probe inducing spurious pulses that are similar to arterial pulses. The spurious pulses when processed can produce erroneous results. This problem is particularly significant in active infants, and patients that do not remain still during monitoring. The quantity of motion required to disturb the signal is very small. Shivering and slight flexing of the fingers can make the signal erroneous.

Another significant problem occurs in circumstances where the patient's blood circulation is poor and the pulse strength is very weak. For example, poor circulation occurs in cases of insufficient blood pressure or reduced body temperature. In such conditions, it is difficult to separate the true pulsatile component from artifact pulses because of the low signal-to-noise ratio. Several time-domain and frequency-domain signal-processing algorithms are proposed to

enhance the performance of pulse oximeters with improved rejection of noise, spurious pulses, motion artifact, and other undesirable aperiodic waveforms.

This chapter describes the algorithms required to estimate the arterial oxygen saturation based on the Beer–Lambert law.

9.2 BEER–LAMBERT LAW

Pulse oximetry measures the effect of arterial blood in tissue on the intensity of the transmitted light (Cheung *et al* 1989). The volume of blood in the tissue is a function of the arterial pulse, with a greater volume present at systole and a smaller volume present at diastole. Because blood absorbs most of the light passing through the tissue, the intensity of the light emerging from the tissue is inversely proportional to the volume of the blood present in the tissue. The emergent light intensity varies with the arterial pulse and can be used to indicate a patient's pulse rate. In addition, the absorbance coefficient of oxyhemoglobin is different from that of deoxygenated hemoglobin for most wavelengths of light. Differences in the amount of light absorbed by the blood at two different wavelengths can be used to indicate the hemoglobin oxygen saturation, which equals

$$\%S_{\mathrm{a}}\mathrm{O}_2 = [\mathrm{HbO}_2]/([\mathrm{Hb}] + [\mathrm{HbO}_2]) \times 100\%. \tag{9.1}$$

The Beer–Lambert law governs the absorbance of light by homogeneous absorbing media. The incident light with an intensity I_0 impinges upon the absorptive medium of characteristic absorbance factor A that indicates the attenuating effect and a transmittance factor T that is the reciprocal of the absorbance factor ($1/A$). The intensity of the emerging light I_1 is less than the incident light I_0 and is expressed as the product TI_0. The emergent light intensity I_n transmitted through a medium divided into n identical components, each of unit thickness and the same transmittance factor T is equal to T^nI_0. I_n can be written in a more convenient base by equating T^n to $\mathrm{e}^{-\alpha n}$, where α is the absorbance of medium per unit length and is frequently referred to as the relative extinction coefficient. The relative extinction coefficient α is related to the extinction coefficient ε (discussed in chapter 4) as $\alpha = \varepsilon C$, where C is the concentration of the absorptive material. The expression for the intensity of the light I_n emerging from a medium can be given by the following general equation called the Beer–Lambert law.

$$I_n = I_0\,\mathrm{e}^{-\alpha d} \tag{9.2}$$

where I_n is the emergent light intensity, I_0 is the incident light intensity, α is the absorbance coefficient of the medium per unit length, d is the thickness of the medium in unit lengths, and the exponential nature of the relationship has arbitrarily been expressed in terms of base e. Equation (9.2) is commonly referred to as the Beer–Lambert law of exponential light decay through a homogeneous absorbing medium (figure 9.1).

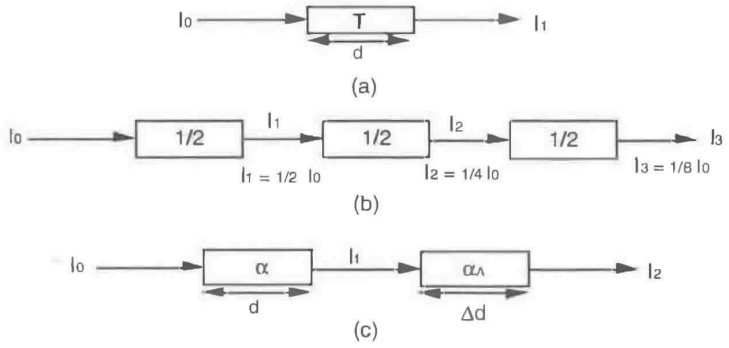

Figure 9.1. A block diagram illustrating the transmittance of light through a block model of the components of a finger. (a) Incident light having an intensity of I_0 impinges upon an absorptive medium with a characteristic transmittance factor T. (b) The effect of a medium divided into n identical components of unit thickness and same transmittance factor T on incident light intensity I_0. (c) For a finger model, the baseline component of the unchanging absorptive elements and the pulsating component of the changing absorptive portion are represented (Cheung *et al* 1989).

9.2.1 Estimation of oxygen saturation using the Beer–Lambert law

The absorbance coefficients of oxygenated and deoxygenated hemoglobin are different at most wavelengths, except at the *isosbestic* wavelength. If a finger is exposed to incident light and the emergent light intensity is measured, the difference between the two is the amount of light absorbed, which contains information relating to the oxygenated hemoglobin content of the blood in the finger. The volume of blood contained in the finger varies with the arterial pulse. The thickness of the finger also varies slightly with each pulse, changing the path length for the light that is transmitted through the finger. Also, the precise intensity of the incident light applied to the finger is not easily determined. Hence, it is desirable to eliminate the effects of intensity of the incident light and the thickness of the path length in estimating oxygen saturation. The Beer–Lambert law needs to be modified to eliminate the input light intensity and length of the path as variables.

9.2.1.1 Eliminating the input light intensity as a variable. The intensity of light transmitted through a finger is a function of the absorbance coefficient of both fixed components, such as bone, tissue, skin, and hair, as well as variable components, such as the volume of blood in the tissue. The intensity of light transmitted through the tissue, when expressed as a function of time is often said to include a baseline component, which varies slowly with time and represents the effect of the fixed components on the light, as well as a periodic pulsatile component, which varies more rapidly with time and represents the effect that changing tissue blood volume has on the light (Cheung *et al* 1989). The baseline component modeling the unchanging absorptive elements has a thickness d and an absorbance α. The pulsatile component representing the changing absorptive portion of the finger has a thickness of Δd and the relative absorbance of α_A representing the arterial blood absorbance (figure 9.1(c)).

The light emerging from the baseline component can be written as a function of the incident light intensity I_0 as follows

$$I_1 = I_0 \, e^{-\alpha d}. \tag{9.3}$$

Likewise, the intensity of light I_2 emerging from the pulsatile component is a function of its incident light intensity I_1 and can be written as follows

$$I_2 = I_1 e^{-\alpha_A \Delta d}. \tag{9.4}$$

Substituting the expression of I_1 in the expression for I_2, the light emerging from the finger as a function of the incident light intensity I_0 is as follows

$$I_2 = I_0 e^{-[\alpha d + \alpha_A \Delta d]}. \tag{9.5}$$

The effect of light produced by the arterial blood volume is given by the relationship between I_2 and I_1. Defining the change in transmittance produced by the arterial component as $T_{\Delta A}$, we have

$$T_{\Delta A} = I_2 / I_1. \tag{9.6}$$

Substituting the expressions for I_1 and I_2 in the above equation yields the following:

$$T_{\Delta A} = (I_0 e^{-[\alpha d + \alpha_A \Delta d]}) / (I_0 e^{-\alpha d}). \tag{9.7}$$

The term I_0 in the numerator and the denominator can be canceled by eliminating the input light intensity as a variable in the equation. Therefore, the change in arterial transmittance can be expressed as

$$T_{\Delta A} = e^{-\alpha_A \Delta d}. \tag{9.8}$$

A device employing this principle in operation is effectively self-calibrating, and is independent of the incident light intensity I_0.

9.2.1.2 Eliminating the thickness of the path as a variable. The changing thickness of the finger, Δd, produced by the changing arterial blood volume remains a variable in equation (9.8). To further simplify the equation, the logarithmic transformation is performed on the terms in equation (9.8) yielding the following

$$\ln T_{\Delta A} = \ln (e^{-\alpha_A \Delta d}) = -\alpha_A \Delta d. \tag{9.9}$$

The variable Δd can be eliminated by measuring arterial transmittance at two different wavelengths. The two measurements at two wavelengths provide two equations with two unknowns. The particular wavelengths selected are determined in part by consideration of a more complete expression of the arterial absorbance α_A

$$\alpha_A = (\alpha_{OA})(S_aO_2) - (\alpha_{DA})(1 - S_aO_2) \tag{9.10}$$

where α_{OA} is the oxygenated arterial absorbance, α_{DA} is the deoxygenated arterial absorbance, and S_aO_2 is the oxygen saturation of arterial Hb. α_{OA} and α_{DA} are substantially unequal at all light wavelengths in the red and near infrared wavelength regions except for the isosbestic wavelength of 805 nm. With an S_aO_2 of approximately 90%, the arterial absorbance α_A is 90% attributable to the oxygenated arterial absorbance α_{OA}, and 10% attributable to the deoxygenated arterial absorbance α_{DA}. At the isosbestic wavelength, the relative contribution of these two coefficients to the arterial absorbance α_A is of minimal significance in that both α_{OA} and α_{DA} are equal (figure 4.2).

Wavelengths selected are in a range away from the approximate isosbestic wavelength that is sufficient to allow the two signals to be easily distinguished. It is generally preferred that the two wavelengths selected fall within the red and infrared regions of the electromagnetic spectrum. The ratio of the transmittance produced by the arterial blood component at red and infrared wavelengths follows from equation (9.9).

$$\frac{\ln T_{\Delta AR}}{\ln T_{\Delta AIR}} = \frac{-\alpha_A(\lambda_R)\Delta d}{-\alpha_A(\lambda_{IR})\Delta d} \tag{9.11}$$

where $T_{\Delta AR}$ equals the change in arterial transmittance of light at the red wavelength λ_R and $T_{\Delta AIR}$ is the change in arterial transmittance at the infrared wavelength λ_{IR}. If the two sources are positioned at approximately the same location on the finger, the length of the light path through the finger is approximately the same for light emitted by each LED. Thus, the change in the light path resulting from arterial blood flow Δd is approximately the same for both the red and infrared wavelength sources. For this reason, the Δd term in the numerator and the denominator of the right side of equation (9.11) cancel, producing

$$\frac{\ln T_{\Delta AR}}{\ln T_{\Delta AIR}} = \frac{\alpha_A(\lambda_R)}{\alpha_A(\lambda_{IR})}. \tag{9.12}$$

Equation (9.12) is independent of the incident light intensity I_0 and the change in finger thickness Δd, attributable to arterial blood flow. Because of the complexity of the physiological process, the ratio indicated in equation (9.12) does not directly provide an accurate measurement of oxygen saturation. The correlation between the ratio of equation (9.12) and actual arterial blood gas measurement is therefore relied upon to produce an indication of the oxygen saturation. Thus, if the ratio of the arterial absorbance at the red and infrared wavelengths can be determined, the oxygen saturation of the arterial blood flow can be extracted from independently derived, empirical calibration curves in a manner dependent on I_0 and Δd. For simplicity, a measured ratio R_{OS} is defined from equation (9.12) as

$$\text{Ratio} = R_{OS} = \frac{\alpha_A(\lambda_R)}{\alpha_A(\lambda_{IR})}. \tag{9.13}$$

9.3 RATIO OF RATIOS

The Ratio of Ratios (R_{OS}) is a variable used in calculating the oxygen saturation level. It is typically calculated by taking the natural logarithm of the ratio of the peak value of the red signal divided by the valley measurement of the red signal. The ratio is then divided by the natural logarithm of the ratio of the peak value of the infrared signal divided by the valley measurement of the infrared signal (Cheung *et al* 1989).

9.3.1 Peak and valley method

A photodiode placed on the side of a finger opposite the red and infrared LEDs receives light at both wavelengths transmitted through the finger. The received red wavelength light intensity varies with each pulse and has high and low values R_H and R_L, respectively. R_L occurs during systole when arterial blood volume is at its greatest, while R_H occurs during diastole when the arterial blood volume is lowest (figure 9.2). Considering the exponential light decay through homogeneous media, it is observed that

$$R_L = I_0 e^{-[\alpha(\lambda_R)d+\alpha_A(\lambda_R)\Delta d]}. \tag{9.14}$$

Similarly,

$$R_H = I_0 e^{-\alpha(\lambda_R)d}. \tag{9.15}$$

Taking the ratio of equations (9.14) and (9.15) and simplifying, we have

$$\frac{R_L}{R_H} = e^{-\alpha_A(\lambda_R)\Delta d}. \tag{9.16}$$

Taking the logarithm of both sides of equation (9.16) yields

$$\ln\left(\frac{R_L}{R_H}\right) = -\alpha_A(\lambda_R)\Delta d. \tag{9.17}$$

Similar expressions can be produced for the infrared signal.

$$\ln\left(\frac{IR_L}{IR_H}\right) = -\alpha_A(\lambda_{IR})\Delta d. \tag{9.18}$$

The ratiometric combination of equations (9.17) and (9.18) yields

$$\frac{\ln\left(\frac{R_L}{R_H}\right)}{\ln\left(\frac{IR_L}{IR_H}\right)} = \frac{-\alpha_A(\lambda_R)\Delta d}{-\alpha_A(\lambda_{IR})\Delta d}. \tag{9.19}$$

Because the Δd terms in the numerator and denominator of the right side of the equation (9.19) cancel, as do the negative signs before each term, equation (9.19) when combined with equation (9.13) yields

$$\text{Ratio} = R_{OS} = \frac{\alpha_A(\lambda_R)}{\alpha_A(\lambda_{IR})} = \frac{\ln\left(\frac{R_L}{R_H}\right)}{\ln\left(\frac{IR_L}{IR_H}\right)}. \tag{9.20}$$

Thus, by measuring the minimum and the maximum emergent light intensities of both the red and infrared wavelengths (R_L, R_H, IR_L, IR_H), a value for the term R_{OS} can be computed. Empirically derived calibration curves are then used to determine the oxygen saturation based on R_{OS}.

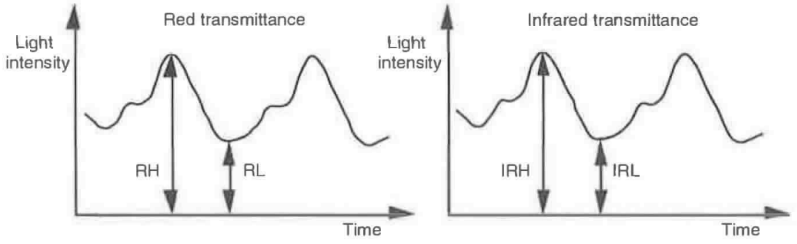

Figure 9.2. A graphical plot of transmitted light intensity converted into voltage. High (H) and low (L) signals are shown as a function of time of the transmittance of red (R) and infrared (IR) light through the finger.

9.3.2 Derivative method: noise reduction software

Yorkey (1996) derives the Ratio of Ratios by calculating using the separated AC and DC components of the measured signal. This mathematical derivation of the ratio of ratios is performed using the Beer–Lambert equation.

$$I_1 = I_0 e^{-\alpha L} \tag{9.21}$$

where I_1 is the emerging light intensity, I_0 is the incident light intensity, α is the relative extinction coefficient of the material and L is the path length. In this method, the Ratio of Ratios is determined using the derivatives. Assuming the change in path length is the same for both wavelengths during the same time interval between samples, the instantaneous change in path length (dL/dt) must also be the same for both wavelengths.

We can extend the general case of taking the derivative of e^u to our case

$$\frac{de^u}{dt} = e^u \frac{du}{dt} \tag{9.22}$$

$$\frac{dI_1}{dt} = I_0 e^{-\alpha L}\left(-\alpha \frac{dL}{dt}\right) \tag{9.23}$$

Therefore,

$$\frac{(dI_1/dt)}{I_1} = -\alpha \frac{dL}{dt}. \tag{9.24}$$

Here, I_1 is equal to the combined AC and DC component of the waveform and dI_1/dt is equal to the derivative of the AC component of the waveform. Using two wavelengths we have

$$\text{R of R} = \frac{(dI_R/dt)/I_R}{(dI_{IR}/dt)/I_{IR}} = \frac{-\alpha(\lambda_R)}{-\alpha(\lambda_{IR})}. \tag{9.25}$$

Instead of using the previous method of calculating the Ratio of Ratios based on the natural logarithm of the peak and valley values of the red and infrared signals, the value of the R of R can be calculated based on the derivative value of the AC component of the waveform.

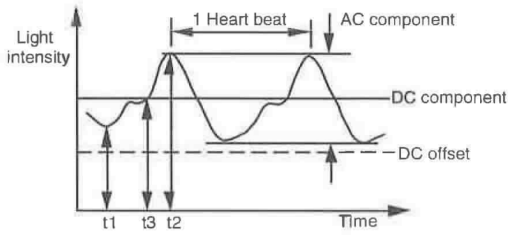

Figure 9.3. A waveform of the transmitted light intensity through a finger showing the AC component, the DC component and the DC offset.

Note in discrete time

$$\frac{dI_R(t)}{dt} \approx I_R(t_2) - I_R(t_1). \tag{9.26}$$

If we choose t_2 and t_1 to be the maximum and minimum of the waveform, we can refer to this difference as the AC value, and the denominator above evaluated at some point in time t_3 in between t_2 and t_1 as the DC value. So,

$$\frac{\dfrac{dI_R(t)/dt}{I_R}}{\dfrac{dI_{IR}(t)/dt}{I_{IR}}} \approx \frac{I_R(t_2) - I_R(t_1)}{I_R(t_3)} \Big/ \frac{I_{IR}(t_2) - I_{IR}(t_1)}{I_{IR}(t_3)} = \frac{\dfrac{AC_R}{DC_R}}{\dfrac{AC_{IR}}{DC_{IR}}} = R. \tag{9.27}$$

Potratz (1994) implemented another improved method for noise reduction called the derivative method of calculating the Ratio of Ratios. To calculate the Ratio of Ratios based on the derivative formula, a large number of sampled points along the waveform are used instead of merely the peak and valley measurements. A series of sample points from the digitized AC and AC + DC values for the infrared and red signals are used to form each data point. A digital FIR filtering step essentially averages these samples to give a data point. A large number of data points are determined in each period. The period is determined after the fact by noting where the peak and valley occur (figure 9.3).

From the AC signal, a derivative is then calculated for each pair of data points and used to determine the ratio of the derivatives for R and IR. A plot of these ratios over a period will ideally result in a straight line. Noise from the motion artifact and other sources will vary some values. But by doing the linear regression, a best line through a period can be determined, and used to calculate the Ratio of Ratios.

A problem with other systems was DC drift. Therefore, a linear extrapolation was performed between two consecutive negative peaks of the waveform. This adjusts the negative peak of the waveform as if the shift due to the system noise did not occur. A similar correction can be calculated using the derivative form of the waveform. In performing the correction of the DC component of the waveform, it is assumed that the drift caused by noise in the system is much slower than the waveform pulses and the drift is linear. The linear change on top of the waveform can be described by the function

$$g(t) = f(t) + mt + b \tag{9.28}$$

where m is equal to the slope of the waveform and b is equal to a constant.

The linear change added to the waveform does not affect the instantaneous DC component of the waveform. However, the derivative of the linear change will have an offset due to the slope of the interfering signal:

$$\mathrm{d}(f(t) + mt + b)/\mathrm{d}t = \mathrm{d}f(t)/\mathrm{d}t + m. \tag{9.29}$$

if we assume that the offset is constant over the period of time interval, then the Ratio of Ratios may be calculated by subtracting the offsets and dividing:

$$\text{R of R} = \frac{Y}{X} = \frac{(y - m_y)}{(x - m_x)} \tag{9.30}$$

where y and x are the original values and m_x and m_y are the offsets.

Since the Ratio of Ratios is constant over this short time interval the above formula can be written as

$$\frac{(y - m_y)}{(x - m_x)} = R. \tag{9.31}$$

Therefore,

$$y = Rx - Rm_x + m_y. \tag{9.32}$$

Since it was assumed that m_1, m_2, and R are constant over the time interval, we have an equation in the form of $y = mx + b$ where m is the Ratio of Ratios. Thus,

we do a large number of calculations of the Ratio of Ratios for each period, and then do the best fit calculation to the line $y = Rx + b$ to fit the optimum value of R for that period, taking into account the constant b which is caused by DC drift.

To determine the Ratio of Ratios exclusive of the DC offset we do a linear regression. It is preferred to take points along the curve having a large differential component, for example, from peak to valley. This will cause the mx term to dominate the constant b:

$$R = \frac{n\sum x_j y_j - \sum x_j \sum y_j}{n\sum x_j^2 - (\sum x_j)^2} \tag{9.33}$$

where n = # of samples, j = sample #, $x = I_R \, dI_{IR} / dt$, $y = I_{IR} \, dI_R / dt$.

Prior sampling methods typically calculate the Ratio of Ratios by sampling the combined AC and DC components of the waveform at the peak and valley measurements of the waveform. Sampling a large number of points on the waveform, using the derivative and performing a linear regression increases the accuracy of the Ratio of Ratios, since noise is averaged out. The derivative form eliminates the need to calculate the logarithm. Furthermore doing a linear regression over the sample points not only eliminates the noise caused by patient movement of the oximeter, it also decreases waveform noise caused by other sources.

9.4 GENERAL PROCESSING STEPS OF OXIMETRY SIGNALS

The determination of the Ratio of Ratios (R_{OS}) requires an accurate measure of both the baseline and pulsatile signal components (Frick *et al* 1989). The baseline component approximates the intensity of light received at the detector when only the fixed nonpulsatile absorptive component is present in the finger. This component of the signal is relatively constant over short intervals and does not vary with nonpulsatile physiological changes, such as movement of the probe. Over a relatively long time, this baseline component may vary significantly. The magnitude of the baseline component at a given point in time is approximately equal to the level identified as R_H (figure 9.2). However, for convenience, the baseline component may be thought of as the level indicated by R_L, with the pulsatile component varying between the values of R_H and R_L over a given pulse. Typically, the pulsatile component may be relatively small in comparison to the baseline component and is shown out of proportion in figure 9.3. Because the pulsatile components are smaller, greater care must be exercised with respect to the measurement of these components. If the entire signal, including the baseline and the pulsatile components, were amplified and converted to a digital format for use by microcomputer, a great deal of the accuracy of the conversion would be wasted because a substantial portion of the resolution would be used to measure the baseline component (Cheung *et al* 1989).

In this process, a substantial portion of the baseline component termed offset voltage V_{OS} is subtracted off the input signal V_1. The remaining pulsatile component is amplified and digitized using an ADC. A digital reconstruction is then produced by reversing the process, wherein the digitally provided information allows the gain to be removed and the offset voltage added back.

This step is necessary because the entire signal, including the baseline and pulsatile components is used in the oxygen saturation measurement process.

Feedback from the microcomputer is required to maintain the values for driver currents I_0, V_{os} and gain A at levels appropriate to produce optimal ADC resolution (figure 9.4). Threshold levels L1 and L2 slightly below and above the maximum positive and negative excursions L3 and L4 allowable for the ADC input are established and monitored by the microcomputer (figure 9.5). When the magnitude of the input to and output from the ADC exceeds either of the thresholds L1 or L2, the drive currents I_D are adjusted to increase or decrease the intensity of light impinging on the detector. This way, the ADC is not overdriven and the margin between L1 and L3 and between L2 and L4 helps assure this even for rapidly varying signals. An operable voltage margin for the ADC exists outside of the thresholds, allowing the ADC to continue operating while the appropriate feedback adjustments to A and V_{os} are made. When the output from the ADC exceeds the positive and negative thresholds L5 or L6, the microcomputer responds by signaling the programmable subtractor to increase or decrease the voltage V_{os} being subtracted. This is accomplished based on the level of the signal received from the ADC. Gain control is also established by the microcomputer in response to the output of the ADC (Cheung *et al* 1989).

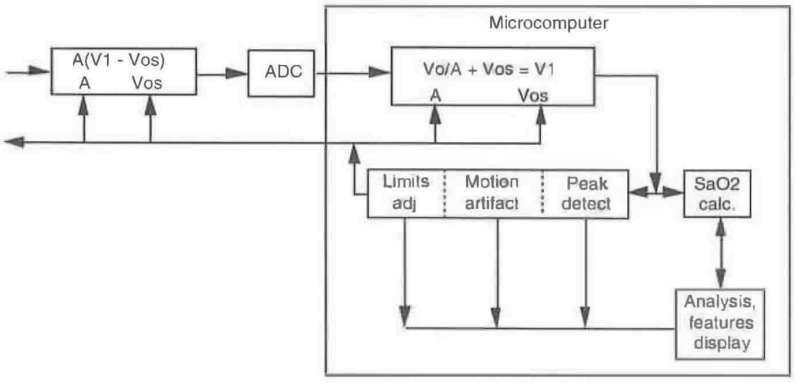

Figure 9.4. A functional block diagram of the microcomputer feedback illustrating the basic operation of the feedback control system. The DC value of the signal is subtracted before digitizing the waveform to increase the dynamic range of conversion. The removed DC value is later added to the digitized values for further signal processing (Cheung *et al* 1989).

A program of instructions executed by the Central Processing Unit of the microcomputer defines the manner in which the microcomputer provides servosensor control as well as produces measurements for display. The first segment of the software is the interrupt level routine.

9.4.1 Start up software.

The interrupt level routine employs a number of subroutines controlling various portions of the oximeter. At the start up, calibration of the oximeter is

performed. After calibration, period zero subroutine is executed which includes five states, zero through four (figure 9.6).

Period zero subroutine is responsible for normal sampling

State 0: Initialize parameters
State 1: Set drive current
State 2: Set offsets
State 3: Set gains
State 4: Normal data acquisition state.

Probe set-up operations are performed during the states zero to three of this subroutine. During these states probe parameters including the amplifier gain A and offset voltage V_{os} are initialized, provided that a finger is present in the probe. State 4 of the interrupt period zero subroutine is the normal data acquisition state. The signals produced in response to light at each wavelength are then compared with the desired operating ranges to determine whether modifications of the driver currents and voltage offsets are required. Finally state 4 of the period zero subroutine updates the displays of the oximeter. Sequential processing returns to state 0 whenever the conditions required for a particular state are violated (Cheung *et al* 1989).

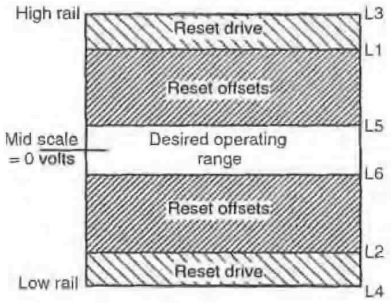

Figure 9.5 A graphical representation of the possible ranges of digitized signal, showing the desired response of the I/O circuit and microcomputer at each of the various possible ranges (Cheung *et al* 1989).

9.5 TRANSIENT CONDITIONS

The relative oxygen content of a patient's arterial pulses and the average background absorbance remain about the same from pulse to pulse. Therefore, the red and infrared light that is transmitted through the pulsatile flow produces a regularly modulated waveform with periodic pulses of comparable shape and amplitude and a steady state background transmittance. This regularity in shape helps in accurate determination of the oxygen saturation of the blood based on the maximum and minimum transmittance of the red and infrared light.

Changes in a patient's local blood volume at the probe site due to motion artifact or ventilatory artifact affect the absorbance of light. These localized

changes often introduce artificial pulses into the blood flow causing the periodic pulses ride on a background intensity component of transmittance that varies as blood volume changes. This background intensity component variation, which is not necessarily related to changes in saturation, affects the pulse to pulse uniformity of shape, amplitude and expected ratio of the maximum to minimum transmittance, and can affect the reliability and accuracy of oxygen saturation determination (Stone and Briggs 1992).

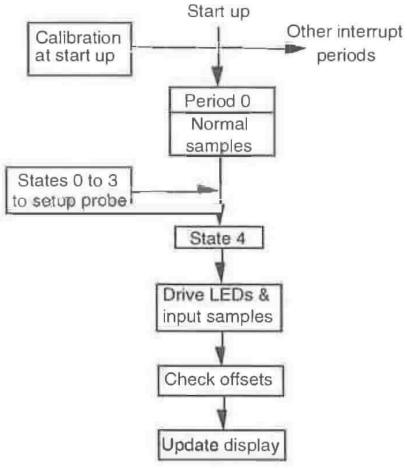

Figure 9.6. Flow chart of a portion of an interrupt level software routine included in the microcomputer (Cheung *et al* 1989).

In addition, there are times when the patient's background level of oxygen saturation undergoes transient changes, for example, when the patient loses or requires oxygen exchange in the lungs while under gaseous anesthesia. The transient waveform distorts the pulse shape, amplitude, and the expected ratio of the pulses, which in turn affects the reliability and accuracy of the oxygen saturation determination.

With changes in the background intensity absorbance component due to artifacts from changes in blood volume or transient saturation changes, the determined saturation value is not accurate and it would not become accurate again until the average absorbance level stabilizes.

The saturation calculations based upon transient signals provide an overestimation or underestimation of the actual saturation value, depending upon the trend. The transmittance of red light increases as oxygen saturation increases resulting in a signal value having a smaller pulse, and the transmittance of the infrared light decreases as saturation increases resulting in the infrared pulsatile amplitude increasing. For these wavelengths, the transmittance changes with saturation are linear in the range of clinical interest, i.e., oxygen saturation between 50% and 100%. The accuracy of the estimation is of particular concern during rapid desaturation. In such a case, the determined saturation based on the

detected signals indicates a greater drop than the actual value. This underestimation of oxygen saturation may actuate low limit saturation alarms that can result in inappropriate clinical decisions.

The pulsatile amplitude is usually quite small, typically less than 5% of the overall intensity change and any small change in overall or background transmittance, such as slight changes in average blood saturation, can have a relatively large effect on the difference in maximum and minimum intensity of the light levels. Because the change in transmittance with changing oxygen saturation is opposite in direction for the red and infrared, this can result in overestimation of the pulsatile ratio during periods when saturation is decreasing, and underestimation during periods when saturation is increasing. It is therefore essential to compensate for the effects of transient conditions and localized blood volume changes on the actual signal, thereby providing a more accurate estimation of the actual oxygen saturation value.

This can be achieved by using a determined rate of change from pulse to pulse, using interpolation techniques and by using the low frequency characteristics of the detected signal values.

The transient error is corrected by linear interpolation where the determined maxima and minima for a first and second optical pulses are obtained, the second pulse following the first. The respective rates of change in the transmittance due to the transient are determined from the maximum transmittance point of the first detected pulse to the second detected pulse (Stone and Briggs 1992). The determined rates of change are then used to compensate any distortion in the detected transmittance of the first detected pulse introduced by the transient in accordance with the following algorithm

$$V_{\max}(n)^{*} = V_{\max}(n) + [V_{\max}(n) - V_{\max}(n+1)] \times \frac{[t_{\max}(n) - t_{\min}(n)]}{[t_{\max}(n+1) - t_{\max}(n)]} \quad (9.34)$$

where $t_{\max}(n)$ is the time of occurrence of the detected maximum transmittance at the n maximum, $t_{\min}(n)$ is the time of occurrence of the detected minimum transmittance of the wavelength at the n minimum, $V_{\max}(n)$ is the detected optical signal maximum value at the maximum transmittance of the wavelength at the n maximum $V_{\max}(n)^{*}$ is the corrected value, for n being the first optical pulse, and $n + 1$ being the second optical pulse of that wavelength.

By application of the foregoing linear interpolation routine, the detected maximum transmittance value at $t_{\max}(n)$ can be corrected, using the values $t_{\max}(n+1)$, detected at the next coming pulse, to correspond to the transmittance value that would be detected as if the pulse were at steady state conditions. The corrected maximum value and the detected (uncorrected) minimum value thus provide an adjusted optical pulse maximum and minimum that correspond more closely to the actual oxygen saturation in the patient's blood at that time, not withstanding the transient condition. Thus, using the adjusted pulse values in place of the detected pulse values in the modulation ratio for calculating oxygen saturation provides a more accurate measure of oxygen saturation than would otherwise be obtained during transient operation.

Similarly, the respective rates of change in the transmittance are determined from the minimum transmittance point of the first detected pulse to the minimum of the second detected pulse. The determined rates of change are then used to compensate for any distortion in the detected minimum transmittance of the

second detected pulse introduced by the transient in accordance with the following algorithm

$$V_{\min}(n)^{*} = V_{\min}(n-1) + [V_{\min}(n) - V_{\min}(n-1)] \times \frac{[t_{\max}(n) - t_{\min}(n-1)]}{[t_{\min}(n) - t_{\min}(n-1)]} \quad (9.35)$$

where $t_{\max}(n)$ is the time of occurrence of the detected maximum transmittance at the n maximum; $t_{\min}(n)$ is the time of occurrence of the detected minimum transmittance of the wavelength at the n minimum; $V_{\min}(n)$ is the detected optical signal minimum value at the minimum transmittance of the wavelength at the n minimum; $V_{\min}(n)^{*}$ is the corrected value, for n being the second optical pulse, and $n-1$ being the first optical pulse of that wavelength.

By application of the foregoing linear interpolation routine, the detected minimum transmittance value at $t = n$ can be compensated using the detected values at the preceding pulse $t = n - 1$, to correspond to the transmittance value that would be detected as if the pulse were detected at steady state conditions. The compensated minimum value and the detected (uncompensated) maximum value thus provide an adjusted optical pulse maximum and minimum that correspond more closely to the actual oxygen saturation in the patient's blood at that time, notwithstanding the transient condition. Thus, using the adjusted pulse values in place of the detected pulse values in the modulation ratio for calculating oxygen saturation provides a more accurate measure of oxygen saturation than would otherwise be obtained during transient operation.

As is apparent from the algorithms, during steady state conditions the compensated value is equal to the detected value. Therefore, the linear interpolation routine may be applied to the detected signal at all times, rather than only when transient conditions are detected. Also, the algorithm may be applied to compensate the detected minimum or maximum transmittance values by appropriate adjustment of the algorithm terms. The amount of oxygen saturation can then be determined from this adjusted optical pulse signal by determining the relative maxima and minima as compensated for the respective wavelengths and using that information in determining the modulation ratios of the known Lambert–Beer equation.

The Nellcor® N-200 oximeter is designed to determine the oxygen saturation in one of the two modes. In the unintegrated mode the oxygen saturation determination is made on the basis of optical pulses in accordance with conventional pulse detection techniques. In the ECG synchronization mode the determination is based on enhanced periodic data obtained by processing the detected optical signal and the ECG waveform of the patient.

The calculation of saturation is based on detecting maximum and minimum transmittance of two or more wavelengths whether the determination is made pulse by pulse (the unintegrated mode) or based on an averaged pulse that is updated with the occurrence of additional pulses to reflect the patient's actual condition (the ECG synchronized mode).

Interrupt programs control the collection and digitization of incoming optical signal data. As particular events occur, various software flags are raised which transfer operation to various routines that are called from a main loop processing routine.

The detected optical signal waveform is sampled at a rate of 57 samples per second. When the digitized red and infrared signals for a given portion of

detected optical signals are obtained, they are stored in a buffer called DATBUF and a software flag indicating the presence of data is set. This set flag calls a routine called MUNCH, which processes each new digitized optical signal waveform sample to identify pairs of maximum and minimum amplitudes corresponding to a pulse. The MUNCH routine first queries whether or not there is ECG synchronization, then the MUNCH routine obtains the enhanced composite pulse data in the ECG synchronization mode. Otherwise, MUNCH obtains the red and infrared optical signal sample stored in DATBUF, in the unintegrated mode. The determined maximum and minimum pairs are then sent to a processing routine for processing the pairs. Preferably, conventional techniques are used for evaluating whether a detected pulse pair is acceptable for processing as an arterial pulse and performing the saturation calculation, whether the pulse pair is obtained from the DATBUF or from the enhanced composite pulse data.

The MUNCH routine takes the first incoming pulse data and determines the maximum and minimum transmittance for each of the red and infrared detected optical signals, and then takes the second incoming pulse data, and determines the relative maximum and minimum transmittance. The routine for processing the pairs applies the aforementioned algorithm to the first and second pulse data of each wavelength. Then the oxygen saturation can be determined using the corrected minimum and detected maximum transmittance for the second pulses of the red and infrared optical signals. Some of the examples demonstrate the above application.

Example 1

Figure 9.7(*a*) shows the representative plethysmographic waveforms in a steady state condition for the red and infrared detected signals. $V_{max}R(1)$ equals 1.01 V, and $V_{min}R(1)$ equals 1.00 V, for $n = 1$, 2 and 3 pulses. $V_{min}R(n)$ is the detected optical signal minimum value at the minimum transmittance at the n pulse minimum. The modulation ratio for the maxima and minima red signal is:

$$\frac{V_{max}R(n)}{V_{min}R(n)} = \frac{1.01v}{1.00v} = 1.01.$$

For the infrared wavelength, $V_{max}IR(n)$ equals 1.01 V and $V_{min}IR(n)$ equals 1.00 V and the determined modulation ratio is 1.01.

Using these determined modulation ratios in the formula for calculating the ratio R provides:

$$R = \frac{\ln[V_{max}R(n)/V_{min}R(n)]}{\ln[V_{max}IR(n)/V_{min}IR(n)]} = \frac{0.01}{0.01} = 1.00.$$

A calculated $R = 1$ corresponds to an actual saturation value of about 81% when incorporated into the saturation equation. A saturation of 81% corresponds to a healthy patient experiencing a degree of hypoxia for which some corrective action would be taken.

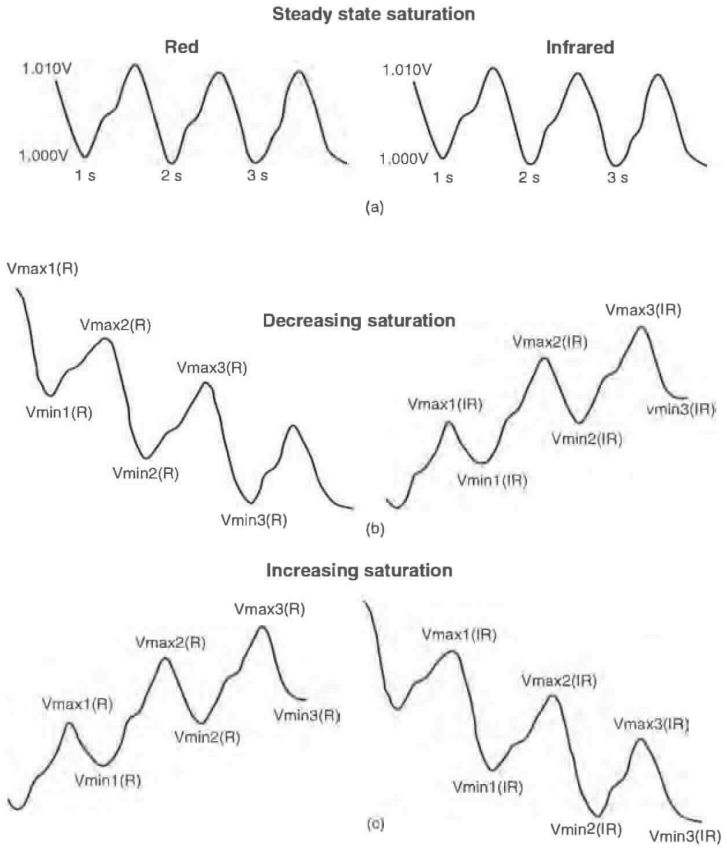

Figure 9.7. Graphical representation of detected optical signals during the steady state and transient conditions (Stone and Briggs 1992).

Example 2

Figure 9.7(*b*) shows the representative plethysmographic waveforms for a patient during desaturation or decreasing saturation transient conditions for the red and infrared detected signals having optical pulses $n = 1$, 2, and 3. However, in this transient example, it is known at $n = 1$, that the actual saturation of the patient is very close to that during the steady state conditions in example 1. In this transient example, the detected values are as follows for both the red and infrared signals:

$t_{max}(1) = 1.0$ s	$V_{max}R(1) = 1.012$ V	$V_{max}IR(1) = 1.008$ V
$t_{min}(1) = 1.2$ s	$V_{min}R(1) = 1.000$ V	$V_{min}IR(1) = 1.000$ V
$t_{max}(2) = 2.0$ s	$V_{max}R(2) = 1.002$ V	$V_{max}IR(2) = 1.018$ V
$t_{min}(2) = 2.2$ s	$V_{min}R(2) = 0.990$ V	$V_{min}IR(2) = 1.010$ V
$t_{max}(3) = 3.0$ s	$V_{max}R(3) = 0.992$ V	$V_{max}IR(3) = 1.028$ V
$t_{min}(3) = 3.2$ s	$V_{min}R(3) = 0.980$ V	$V_{min}IR(3) = 1.020$ V

Calculating the oxygen saturation ratio R at $n = 1$, using the detected optical signal provides the following

$$\begin{aligned} R &= \frac{\ln[V_{max}R(1)/V_{min}R(1)]}{\ln[V_{max}IR(1)/V_{min}IR(1)]} \\ &= \ln[1.012/1.000]/\ln[1.008/1.000] \\ &= \ln[1.012]/\ln[1.008] \\ &= 0.012/0.008 \quad = 1.5. \end{aligned}$$

The calculated saturation ratio of 1.5 based on the detected transmittance corresponds to a calculated oxygen saturation of about 65 for the patient, which corresponds to severe hypoxia in an otherwise healthy patient. This contrasts with the known saturation of about 81% and demonstrates the magnitude of the underestimation of the oxygen saturation (overestimation of desaturation) due to the distortion in transmittance of the red and infrared light caused by transient conditions.

Applying the correction algorithm to correct the distorted maximum transmittance point of the detected red signal during the transient condition:

$$\begin{aligned} V_{max}R(1)^* &= V_{max}R(1) - [V_{max}R(1) - V_{max}R(2)] \times \frac{[t_{max}(1) - t_{min}(1)]}{[t_{max}(2) - t_{max}(1)]} \\ &= 1.012 - [1.012 - 1.002] \times [1.0 - 1.2]/[1.0 - 2.0] \\ &= 1.010. \end{aligned}$$

and correspondingly for the maximum transmittance of the detected infrared signal

$$\begin{aligned} V_{max}IR(1)^* &= 1.008 - [1.008 - 1.018] \times [1.0 - 1.2]/[1.0 - 2.0] \\ &= 1.010 \end{aligned}$$

Thus, by replacing $V_{max}R(n)$ with $V_{max}R(n)^*$ and replacing $V_{max}IR(n)$ with $V_{max}IR(n)^*$ in the calculations for determining the oxygen saturation ratio R, we have

$$\begin{aligned} R &= \frac{\ln[V_{max}R(1)^*/V_{min}R(1)]}{\ln[V_{max}IR(1)^*/V_{min}IR(1)]} \\ &= \ln[1.010/1.00]/\ln[1.010/1.00] \\ &= 0.01/0.01 \\ &= 1.0. \end{aligned}$$

Thus, basing the saturation calculations on the corrected maximum transmittance values and the detected minimum transmittance values, the corrected R value corresponds to the same R for the steady state conditions and the actual oxygen saturation of the patient.

Example 3

Figure 9.7(*c*) shows the representative plethysmographic waveforms for a patient during desaturation or decreasing saturation transient conditions for the red and infrared detected signals having optical pulses $n = 1$, 2 and 3. However, in this transient example, it is known that at $n = 2$, the actual saturation of the patient is very close to that during the steady state conditions in example 1. In this transient example, the detected values are as follows for both the red and infrared signals:

$t_{max}(1) = 1.0$ s	$V_{max}R(1) = 1.022$ V	$V_{max}IR(1) = 1.002$ V
$t_{min}(1) = 1.2$ s	$V_{min}R(1) = 1.008$ V	$V_{min}IR(1) = 0.992$ V
$t_{max}(2) = 2.0$ s	$V_{max}R(2) = 1.012$ V	$V_{max}IR(2) = 1.012$ V
$t_{min}(2) = 2.2$ s	$V_{min}R(2) = 0.998$ V	$V_{min}IR(2) = 1.002$ V
$t_{max}(3) = 3.0$ s	$V_{max}R(3) = 1.002$ V	$V_{max}IR(3) = 1.022$ V
$t_{min}(3) = 3.2$ s	$V_{min}R(3) = 0.988$ V	$V_{min}IR(3) = 1.012$ V

Calculating the oxygen saturation ratio R at $n = 2$, using the detected optical signal provides the following

$$R = \frac{\ln[V_{max}R(2)/V_{min}R(2)]}{\ln[V_{max}IR(2)/V_{min}IR(2)]}$$
$$= \ln[1.012/0.998]/\ln[1.012/1.002]$$
$$= 0.01393/0.0099 \quad = 1.4.$$

Thus, the calculated saturation ratio of 1.4 based on the detected transmittance corresponds to a calculated oxygen saturation of about 51% for the patient, which corresponds to severe hypoxia in an otherwise healthy patient. This contrasts with the known saturation of about 81% and demonstrates the magnitude of the underestimation of the oxygen saturation (overestimation of desaturation) due to the distortion in transmittance of the red and infrared light caused by transient conditions.

Applying the correction algorithm to correct the distorted minimum transmittance point of the detected red signal during the transient condition, we find the following:

$$V_{min}R(2)^* = V_{min}R(2) - [V_{min}R(2) - V_{min}R(1)] \times \frac{[t_{max}(2) - t_{min}(1)]}{[t_{min}(2) - t_{max}(1)]}$$
$$= 1.008 - [0.998 - 1.008] \times [2.0 - 1.2]/[2.2 - 1.2]$$
$$= 1.0$$

and correspondingly for the minimum transmittance of the detected infrared optical signal we have:

$$V_{\min}\mathrm{IR}(2)^{*} = 0.992 - [1.002 - 0.992] \times 0.8$$
$$= 1.0.$$

Thus, by replacing $V_{\min}\mathrm{R}(n)$ with $V_{\min}\mathrm{R}(n)^{*}$ and replacing $V_{\min}\mathrm{IR}(n)$ with $V_{\min}\mathrm{IR}(n)^{*}$ in the calculations for determining oxygen saturation ratio R we have:

$$R = \frac{\ln[V_{\max}\mathrm{R}(2) / V_{\min}\mathrm{R}(2)^{*}]}{\ln[V_{\max}\mathrm{IR}(2) / V_{\min}\mathrm{IR}(2)^{*}]}$$
$$= \ln[1.012 / 1.0] / \ln[1.012 / 1.0]$$
$$= 1.0.$$

Thus, basing the saturation calculations on the corrected minimum transmittance values and the detected maximum transmittance values, the corrected R value corresponds to the same R for the steady state conditions and the actual oxygen saturation of the patient.

9.6 ECG SYNCHRONIZATION ALGORITHMS

Electrical heart activity occurs simultaneously with the heartbeat and can be monitored externally and characterized by the electrocardiogram waveform. The ECG waveform comprises a complex waveform having several components that correspond to electrical heart activity of which the QRS component relates to ventricular heart contraction. The R wave portion of the QRS component is typically the steepest wave therein having the largest amplitude and slope, and may be used for indicating the onset of cardiac activity. The arterial blood pulse flows mechanically and its appearance in any part of the body typically follows the R wave of the electrical heart activity by a determinable period of time. This fact is utilized in commercially available pulse oximeters to enhance their performance. Another advantage of recording ECG is that it provides a redundancy in calculating the heart rate from both the ECG signal and the optical signal to continuously monitor the patient even if one of the signals is lost (figure 9.8).

With ECG synchronization, the pulse oximeter uses the electrocardiographic (ECG) QRS complex as a timing indicator that the optical pulse will soon appear at the probe site. The R portion of the ECG signal is detected and the time delay by which an arterial pulse follows the R wave is determined to establish a time window an arterial pulse is to be expected. By using the QRS complex to time the oximeter's analysis of the optical pulse signal, ECG processing synchronizes the analysis of oxygen saturation and pulse rate data. The established time window provides the oximeter with a parameter enabling the oximeter to analyze the blood flow only when it is likely to have a pulse present for analysis. This method of signal processing passes those components of the signal that are coupled to the ECG (i.e., the peripheral pulse), while attenuating those components that are random with respect to the ECG (e.g., motion artifact or other noise in the signal).

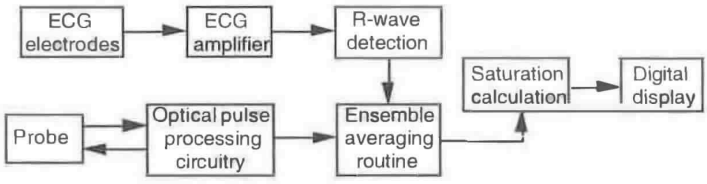

Figure 9.8. Block diagram illustrating the ECG processing components, its subcomponents and their relationship in an oximeter.

9.6.1 Nellcor® system

C-LOCK ECG synchronization enhances the signal-processing capabilities of Nellcor® systems such as the N-200 pulse oximeter and the N-1000 multifunction monitor. This improves the quality of the optical signal in certain clinical settings in which the performance of a conventional pulse oximeter may deteriorate, e.g. when a patient is moving or has poor peripheral pulses. Consequently, C-LOCK signal processing extends the range of clinical situations in which pulse oximetry may be used. Patient movement and poor peripheral pulses present similar problems for a conventional pulse oximeter: performance may deteriorate because the oximeter is unable to distinguish between the true optical pulse signal and background noise. C-LOCK ECG synchronization improves signal quality in these difficult signal-detection settings (Goodman and Corenman 1990).

The digital optical signal is processed by the microprocessor of the Nellcor N-1000 Pulse Oximeter in order to identify individual optical pulses and to compute the oxygen saturation from the ratio of maximum and minimum pulse levels as seen by the red wavelength compared to the pulse seen by the infrared wavelength.

Noninvasive pulse oximeters process optical signals which are prone to motion artifacts caused by the muscle movement proximate to the probe site. The spurious pulses induced in the optical signals may cause the pulse oximeter to process the artifact waveform and provide erroneous data. This problem is particularly significant with infants, fetuses, or patients that do not remain still during monitoring. Another problem exists in circumstances where the patient is in poor condition and the pulse strength is very weak. In continuously processing the optical data, it can be difficult to separate the true pulsatile component from the artifact pulses and noise because of low signal to noise ratio. Inability to reliably detect the pulsatile component in the pulsatile signal may result in a lack of the information needed to calculate oxygen blood saturation.

By incorporating the patient's heart activity into the pulse oximeter, problems due to motion artifact and low signal-to-noise ratio can be solved. Processing of the signals that occur during a period of time when the optical pulses are expected to be found, increases the likelihood that the oximeter will process only optical waveforms that contain the pulsatile component of arterial blood, and will not process spurious signals.

The software incorporated into the microprocessor for processing the ECG signals and displaying the calculated ECG pulse rate receives the digitized version of diagnostic ECG signal (DECG) and filtered ECG signals (FECG). The microprocessor calculates the amplitude of the ECG waveform and controls the AGC (automatic gain control) amplifier, so that DECG and FECG will fall within the voltage range limits of the electronic circuitry used to process these signals.

The microprocessor regularly searches a status input latch at a rate of 57 cycles per second. The output of detected R wave (DRW) sets the latch to a logical 1 when the R wave is detected. Depending on the status, the microprocessor selects the next operation and resets the DRW latch to 0. At this first level, the microprocessor counts the time interval beginning from the detection of an R wave pulse until the occurrence of the next logical 1 at the status input latch. Based on this time interval, the pulse oximeter displays the pulse rate. After averaging several time intervals and establishing a regular ECG pulse rate, the microprocessor will change to the second level of processing.

After the detection of an R wave pulse, the microprocessor separately analyzes the digital optical signal and correlates the period of time by which an optical pulse follows the detected R wave pulse to establish the time window during which the optical pulse is likely to occur. During this second level, the pulse oximeter just calculates and displays the time period or pulse rate between DRW pulses.

The third level of processing starts after a time window has been established. On detecting an R wave pulse, the microprocessor activates the time window so that only optical signals detected within the time window following the occurrence of an R wave pulse will be evaluated for acceptance or rejection and for use in calculating and displaying vital measurements such as oxygen saturation, pulse flow, and pulse rate. The evaluation of a detected pulse is made in conjunction with a preselected confidence factor that is associated with the quality of the optical signals. The higher the optical signal quality, the better the correlation between the recorded pulse history and the detected pulse, and the higher the confidence level. The confidence level may be set automatically by the microprocessor, or it may be adjusted by the operator. The microprocessor will reject any detected pulses occurring outside the time window. A typical time window for an adult male using a fingertip oximeter probe may be about 50 ms ±10 ms after the occurrence of an R wave. The oximeter will also reject any additional pulses detected after an optical pulse is detected within the same time window, even though the time window has not expired.

However, if the optical pulse is not found within an opened time window, the microprocessor will continue to search for optical pulses using the degraded criteria during the time window period for about three successive detected R wave (DRW) pulses, after which it continues to search with degraded criteria. After a specific interval, e.g. 10 s, without detecting an optical pulse, the microprocessor will revert to independent or nonintegrated processing of the optical and ECG signals, returning the pulse oximeter to startup conditions. Therefore, if the oximeter cannot establish or maintain a reliable correlation between the R wave and the optical pulse, the waveforms will be processed independently. The display will indicate whether the pulse oximeter is operating in integrated or nonintegrated mode. After attaining the third level of processing, losing either the ECG or optical pulse signals will activate an alarm and return the program to the startup condition.

9.6.1.1 R-wave determination routine. The R-wave determination routine begins with electric signals received from the ECG leads and calculating the R–R period RRPER between the last detected R wave and the present R wave (figure 9.9). The average period HISTORY from the previous R waves and the present R wave is calculated and the determined RRPER is compared to the average period HISTORY (Goodman and Corenman 1990). If RRPER does not correspond to HISTORY, the R wave ECG flag is reset and the routine is exited to await another R wave. If RRPER does correspond to HISTORY, a timer is activated to measure the interval from the occurrence of the R wave to the occurrence of the optical pulse. Output HR (ECG heart rate) is calculated based on successive R waves. The system determines whether a series of R–R periods have been synchronized (ECG synchronization). If not synchronized, then the system checks for alarms by comparing output HR to a preselected heart rate and generates an alarm if the output HR is too low. If the ECG is synchronized but the optical pulse to optical pulse is not synchronized, the output HR is sent to the display and then checked for alarms. If the optical signal is synchronized, then the system just checks for alarms. Only if the ECG is synchronized, the optical pulse is not synchronized, and the R wave looks like a valid R wave by comparison with HISTORY, then HISTORY is updated using the new R wave. After updating HISTORY, the system itself is updated (TIME OUT) to maintain synchronization. If TIME OUT is not updated for a period of five seconds, then ECG synchronization is lost and the routine must begin building a new history.

9.6.1.2 The systems routine. The system routine for processing digital optical pulse information for optical pulses to send to LEVEL 3 is flow charted (figure 9.10). The system begins by continuously evaluating the data from the detected digital optical signal (Goodman and Corenman 1990). The data are first evaluated for compatibility with signal processing. If the data are over or undervalued electronically, i.e., beyond the voltage range of the circuitry, then the system exits the routine, and the LED intensities are adjusted to correct the electrical values accordingly. When the data are compatible, they are next evaluated for a maximum signal. A relative maximum is determined and saved. The next value is compared to the saved value, and if it is a new maximum, it is saved instead. When the value found is not a new maximum, then a MAX FLAG is set. Thereafter, the system evaluates the following data received, by passing the maximum value section, to find the maximum slope, again by successive comparisons. When the largest slope value is found, it is saved and the SLOPE FLAG is set. Thereafter, the following data are evaluated, by passing the maximum and slope calculations, to find the minimum value corresponding to the end of the pulse. When the smallest minimum is found, it is saved and the slope value that was saved is compared with a pre-established minimum threshold to determine whether it is large enough to be a possible optical pulse. If it is not large enough, then the pulse is rejected, the flags are reset, and the routine begins processing the next possible pulse. If the slope is large enough, then the pulse parameters, maximum, minimum, and slope, are saved in memory for use by LEVEL 3 processing in evaluating the possible pulse. Then, the time delay from the R wave to the possible pulse is calculated. Thereafter, the DATA FLAG is set indicating to LEVEL 3 that there is a possible pulse to be evaluated, the MAX and SLOPE FLAGs are reset, and the routine begins again to process the following data, looking for new maximum values corresponding to possible pulses.

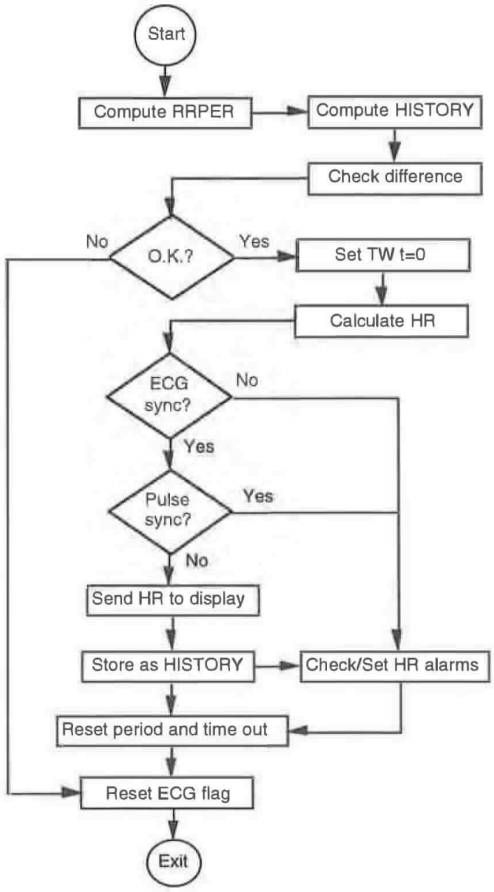

Figure 9.9. The R wave determination routine calculates RRPER, compares it with the average period HISTORY. If RRPER corresponds to HISTORY, the interval between the occurrence of R wave and occurrence of pulse is measured. The algorithm checks for ECG synchronization, alarms and displays heart rate (HR) (Goodman and Corenman 1990).

9.6.1.3 LEVEL 3 software. Figure 9.11 shows LEVEL 3 of software for computing the saturation measurements (Goodman and Corenman 1990). The system starts by acquiring a potential optical pulse after a DATA FLAG has been set and inquiring whether there is ECG synchronization i.e., a regular ECG period has been established. If a DATA FLAG has not been set, then the system exits the routine. If there has not been ECG synchronization, then the microprocessor processes the optical pulse signals independent of the ECG.

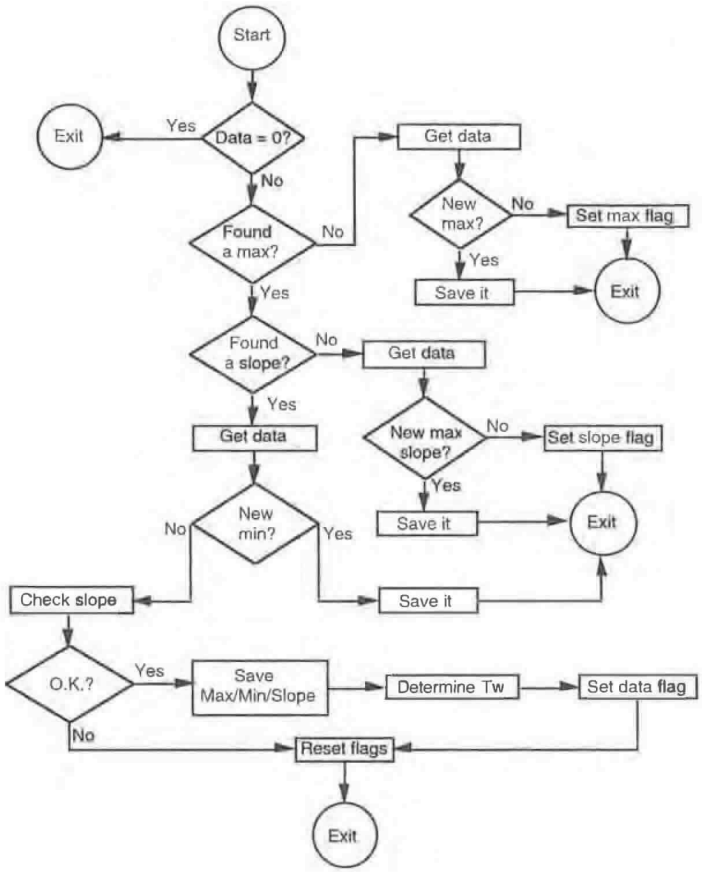

Figure 9.10. The system routine measures the maximum and minimum values in the data presented and calculates the largest slope. The slope value is compared with the normal expected values to determine whether it is a possible optical pulse (Goodman and Corenman 1990).

If there is ECG synchronization, but no R wave has occurred, then the system exits and the pulse is not processed. If there is ECG synchronization and a R wave has occurred, then the microprocessor processes the pulse. The LED intensity is evaluated to see if adjustment is necessary. The reset system gain, based on minimum LED intensity required for adequate signal strength, is checked to see if adjustment is required to the optical pulse historic period, amplitude and ratio. The system then inquires whether the ECG apparatus is operating between an R wave and the following optical pulses for the previous four pulses is computed to give the TIME WINDOW (TW). Then the pulse waveform is analyzed to see if it is a dicrotic notch rather than a real optical pulse. The downward slope of a dicrotic notch or other artifact can be

misinterpreted as an optical pulse, but typically the pulse amplitude is less than half the amplitude of an actual pulse. If the pulse is determined to be a notch or artifact, then the system exits and the next pulse presented will be processed. If not determined to be a notch, then it is analyzed to determine if it is a pulse.

Assuming the ECG is synchronized, then the system determines if two criteria are met. The first is whether the time delay falls within the above-computed TIME WINDOW. If it does not, then the microprocessor rejects the pulse. The second criterion tested is whether or not the ratio is within acceptable limits. Only if the pulse satisfies both criteria is the pulse accepted and a saturation calculation made.

If the ECG is not synchronized then the pulse must pass any two of three criteria regarding (1) pulse period, (2) amplitude, and (3) ratio, to be accepted, e.g., pulse and period, period and amplitude, pulse and amplitude, or all three. If the pulse is accepted, then the oxygenation saturation is calculated.

After the system is turned on (POWER UP) after a TIME OUT alarm (a 10 s period with no valid optical pulse found) a series of consistent pulses must be found to generate an optical pulse history before the oxygenation saturation will be sent to the display. Thus, if there is no optical pulse synchronization, there will be no saturation display. All optical pulses, those accepted and those not accepted, excluding pulses rejected as artifacts, enter the calculation routine section. If the ECG is not synchronized then a pulse-to-pulse period and either an amplitude or a ratio must exist for the optical heart rate (OHR) calculation to be made. If either the ECG or the optical pulse is synchronized, then the HR calculation made will be displayed. If there is no synchronization, then the OHR is not displayed. The system is evaluating the status for pulse evaluation, i.e., whether signals should continue to be processed after a TIME WINDOW period has expired then TIME WINDOW is closed until opened by the detection of the next R wave. The blood oxygen saturation is calculated using the Ratio of Ratios.

9.6.2 Criticare® systems

The patient wears three standard ECG electrodes which provide the pulse oximeter with an ECG signal which if present is used to enhance the quality of the optical waveforms. The oximeter computes oxygen saturation from the enhanced waveform and displays it on a screen (Conlon *et al* 1990).

An ECG amplifier and an R-wave detection algorithm routine process the ECG signal provided by the electrodes and determine the timing for an ensemble averaging algorithm routine. An oxygen saturation value is calculated by a microcomputer in a calculation algorithm routine using the ensemble averaged waveform as input, and is then displayed digitally on a screen.

If an ECG signal is not present, the absence is detected by the R wave detection algorithm routine which causes the ensemble averaging routine to be bypassed and the unenhanced optical pulse to be input into the calculation algorithm routine. The microcomputer executes the software comprising the R wave detection, ensemble averaging, calculation, and display algorithm routines.

The three lead ECG signal is amplified by a differential amplifier. This amplifier amplifies the differential component of the signal, which is the desired ECG waveform, while rejecting a large portion of the common-mode voltage. The output of this amplifier is AC-coupled by a capacitor to an amplifier which provides further gain. The gain provided by the amplifier is adjustable and can be set to 1/2 or 2 by the microprocessor. The amplifier can also accept an additional

high level input which is intended to be connected to the output of an external ECG monitoring device, thus obviating the need for an additional set of ECG electrodes on the patient. The output of the amplifier is processed by a low-pass filter to remove the unwanted artifact such as 60 Hz and electrosurgery induced noise, and is converted to a serial, digital signal by an ADC. The digitized signal then passes through an optoisolator to a serial port which resides on the bus of the microcomputer. The optoisolator serves to isolate the patient ECG leads from the external power supply and is incorporated for reasons of patient safety.

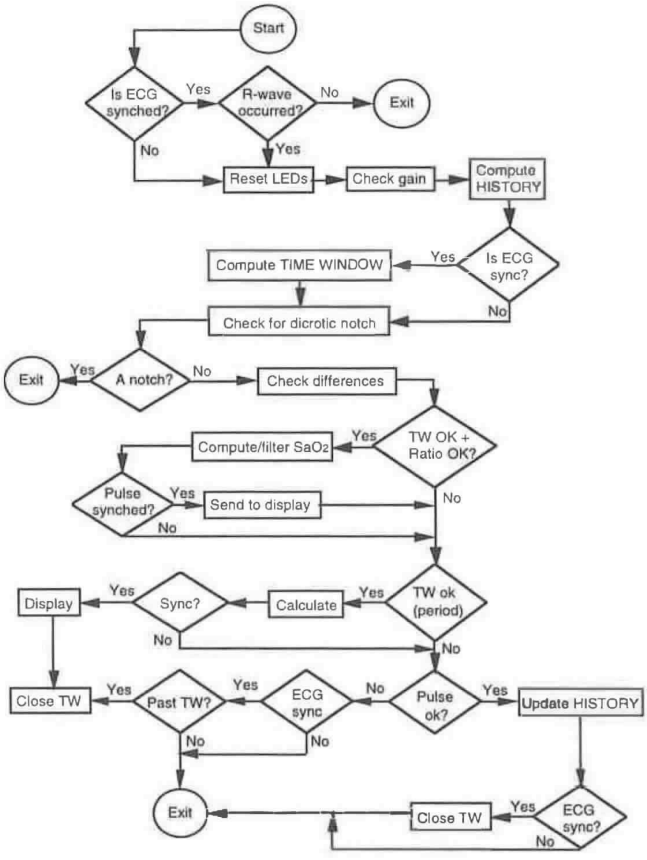

Figure 9.11. The LEVEL 3 software checks for ECG synchronization and processes the data appropriately to calculate the oxygen saturation (Goodman and Corenman 1990).

The oximeter is software driven and the operation of the software involves the process of removing motion artifact and enhancing waveform quality in low perfusion situations.

ECG synchronization is used to provide a reliable time frame upon which to base ensemble averaging, and a robust and accurate R wave detection algorithm is an integral part of the system. The R wave detection process involves three stages of processing: a low-pass digital filter, a peak excursion finding algorithm and a peak discrimination algorithm. The ECG input signal from the ADC is sampled at a rate of 240 Hz. The resulting digital waveform is low-pass filtered, with a corner frequency of 12 Hz, to remove artifact such as 60 Hz and muscle noise.

9.6.2.1 Peak excursion finding algorithm. The filtered ECG waveform then undergoes transformation by the *peak excursion finding algorithm*. The purpose of this transformation is to amplify those characteristics of the ECG waveform which are inherent in QRS complexes while inhibiting those which are not (Conlon *et al* 1990). This algorithm continually matches the ECG waveform to one of the two templates as shown in figure 9.12. The algorithm maintains a queue buffer of length N, which is searched in order to determine the parameters P1, P2, P3, and P4. The algorithm routine is called for $N = 8$, 12, 16, 20, and 24, and the individual excursion values are summed so as to give a total transformation value. More weight is placed on lower values of N in order to emphasize narrower spikes over wider ones. The newest sample is added to the buffer at each instant and the oldest sample is removed from the buffer. The maximum and the minimum values and their positions are searched in the buffer and depending on their relative positions, the matched template is chosen. The parameters P2 and P3 are assigned the appropriate maximum and minimum values accordingly. The parameters P1 and P4 are then found based on the template. For example, if the buffer matches template (a), the maximum value after P2 is assigned to P1, and the minimum before P3 is assigned to P4. Finally, the peak closed excursion on the interval N is computed as (P3 – P2 – (P4 – P1)) if the buffer matches template (a) or (P2 – P3 – (P1 – P4)) if the buffer matches template (b).

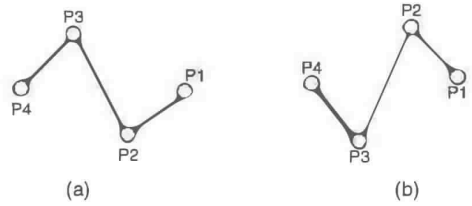

Figure 9.12. Two ECG waveform templates utilized in R-wave detection.

9.6.2.2 Peak discrimination algorithm. After transformation of the ECG waveform, the peak discrimination algorithm classifies the spikes found in the transformed waveform as either QRS complexes or artifact. The peak discrimination algorithm is a state machine with three states: peak, valley, and noise peak. The thresholds are set based upon the past history of the ECG waveform.

The algorithm enters the peak state if the algorithm is in the valley or noise state and exceeds a set threshold (threshold 2). The algorithm exits the peak state and enters the valley state when the waveform drops below one fourth of the maximum value attained in the peak state. The algorithm in valley state enters the

noise state whenever the waveform climbs above four times the minimum value attained during the valley state. The algorithm in noise state enters the valley state when the waveform drops half the distance between the maximum value attained during the noise state and the minimum during the previous valley state. The detection of QRS spike is signaled upon the transition into the peak state. The conditions for state changes are summarized in the table 9.1. The algorithm maintains an average of the last eight QRS peaks in order to set the threshold for detecting the next peak in the waveform. An average of the noise peak levels found between the last four QRS peaks, is also maintained to aid the rejection of artifact while accepting valid QRS spikes. The averages are updated whenever there is a transition between the peak and valley states.

Table 9.1 A summary of conditions for state changes.

Present state	Condition	Next state
Valley or Noise states	Exceeds a set threshold	Peak state
Peak state	< 1/4 max in peak state	Valley state
Valley state	> 4* min in valley state	Noise state
Noise state	< 1/2 (max in noise – min in previous valley)	Valley state

Additional rejection of artifact is gained by examining the length of time which has elapsed between a new peak and the last accepted peak (interval figure 9.15). If it is less than 5/8 of the previous R–R interval, the spike is assumed to be noise and is not counted as a QRS spike. If it is greater than 7/8 of the previous R–R interval, it is accepted unconditionally. If it is greater than 5/8, but less than 7/8 of the previous R–R interval, the spike is accepted on probation as long as it exceeds a second threshold (threshold 1) which is set based on the noise peaks encountered during the last four beats. It is counted as a valid QRS spike but the previous state information is also saved in order to undo acceptance of the spike if a better candidate is found. The probation interval is equal to 9/8 of the previous R–R interval minus the length of time which has elapsed since the last accepted peak. During this interval any spike which meets the threshold requirements overrides the acceptance of the spike in question.

When the algorithm is found to be in the peak state, the maximum value encountered in this state is noted. If the waveform is not a local maximum, the routine checks to see if the waveform has fallen to one fourth of the last local maximum. If it has, the routine determines whether the current peak is a noise peak or a QRS peak. If it was a noise peak, the average of the noise levels over the last four beats is calculated. If it was a QRS peak, the average of the last eight QRS peaks is updated using the local maximum. The threshold values needed to detect the next QRS peak are then determined. Threshold 1 is halfway between the current eight-beat peak average and the current four-beat noise average. Threshold 2 is one-half of the current eight beat peak average (figure 9.13).

Before exiting the peak discrimination algorithm, parameters reflecting the quality of the ECG waveform are tested. If the time elapsed since the spike was accepted exceeds four times the R–R interval and/or the baseline of the transformed signal exceeds one-half the peak value, the ECG waveform is assumed to be lost and the routine disengages the synchronization.

9.6.2.3 Ensemble averaging algorithm. The ensemble averaging algorithm makes use of the output of the R–wave peak discrimination algorithm to enhance that

part of the red and infrared plethysmographic waveforms which are correlated with the ECG, while diminishing all which is unrelated, to yield a signal with an improved signal to noise ratio (Conlon *et al* 1990).

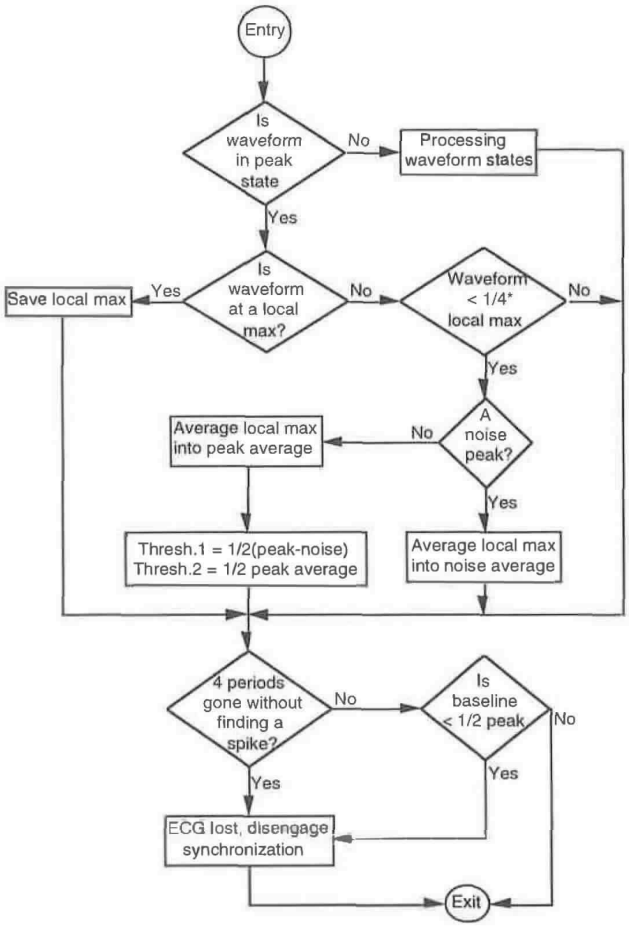

Figure 9.13. The R wave peak discrimination algorithm (Conlon *et al* 1990).

The algorithm relies on the assumption that instances of moderate to severe motion, and of low perfusion, can be detected as the plethysmographic waveforms are being sampled (figure 9.14). To do this, it was found to be advantageous to buffer these waveforms while they are being sampled, and to delay the actual averaging until the R peak is detected. The averaging weight of the current waveform cycle can then be adjusted, depending on whether the plethysmographic waveform just acquired is weak or exhibits the influence of

excessive motion artifact. An additional benefit of this buffering stage is that the oximeter is able to discard waveform pulses during which optical pulse processing circuitry has saturated and distorted the waveform. Yet another benefit of this buffering stage is that it allows a level of error tolerance in the R wave detection process whereby the peak_discrimination routine can accept certain marginal QRS spikes on probation while maintaining the flexibility to correct the error if a better candidate is subsequently detected (figure 9.15).

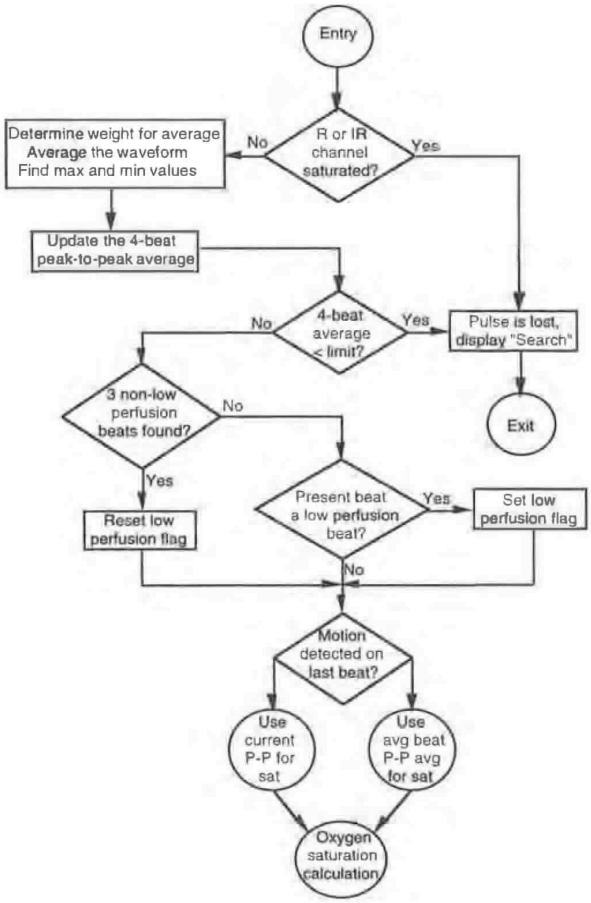

Figure 9.14. The ensemble averaging algorithm (Conlon *et al* 1990).

9.6.2.4 Motion determination algorithm. In order to give less weight to waveform pulses which are distorted by motion artifact, a criterion by which motion can be measured is established. This routine assumes that a plethysmograph unaffected by motion varies only slightly between one pulse and

the next. In addition, a change in the amplitude, not shape, of the pulse comprises the majority of the observed difference between one pulse and the next. A plethysmograph containing artifact, however, differs greatly from the previous signal. A point-by-point subtraction of the latest pulse from the one preceding it yields a signal with an average amplitude less than that of the signal. The value of the difference signal is more or less constant while the signal itself changes rapidly. The integration of the difference signal yields a good indication of the amount of motion present in the pulse. A noisy signal yields a large value on integration compared to a clean signal. This routine checks for the occurrence of an R-wave spike which would be detected by the R-wave detection algorithm. If a spike was detected, the routine saves the integrated value as an indication of the level of motion present in the pulse, and initializes the variables to prepare for the next pulse (Conlon *et al* 1990).

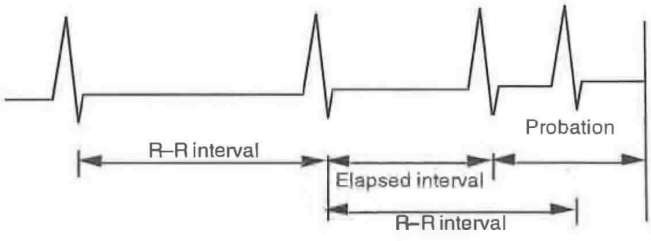

Figure 9.15. R wave artifact rejection timing subroutine (Conlon *et al* 1990).

9.6.2.5 Pulsatile waveform weight determination algorithm. It is generally known that ensemble averaging with a set of N waveforms increases signal-to-noise ratio by a factor of the $\sqrt{N}$ for uncorrelated, random noise. Thus, ensemble averaging will decrease the influence of the uncorrelated motion artifact and will enhance a low perfusion signal (which may be buried in noise) at the expense of response time. At the same time, a maximum limit on response time is set in order to ensure that the displayed saturation value is reasonably current (Conlon et al 1990).

The variable weight average is used in order to provide flexibility over a broad spectrum of pulsatile waveforms. It attempts to give a large weight to waves which are largely motion-free, while diminishing the weight given to those which have motion. Additionally, if a low perfusion situation is detected, less weight is given to all pulses until several strong pulses are found. Furthermore, the algorithm takes into account the pulse rate when determining the averaging weight. Since the averaging occurs each time a beat is detected, more averaging can be used on a patient with a fast pulse rate than one with a slow pulse rate while maintaining a constant response time. More averaging is needed in cases of motion artifact and low perfusion because the signal-to-noise ratio of these pulses is less than normal pulses (figure 9.16).

The weight determination algorithm uses two empirically determined thresholds to determine whether the motion is significant. One of these thresholds applies during the normal perfusion, while the other is used in cases of low perfusion. The algorithm decides which of the two thresholds to use by checking for the low perfusion state. If the low perfusion has not been detected, the

algorithm checks the motion against the high motion threshold. If significant motion is not found, the algorithm checks whether the heart rate is above 120 bpm. If it is, the beat is assigned an average weight of 1/8, otherwise an average weight of 1/4. If significant motion is found, the algorithm checks for the heart rate and if it is above 120 bpm, assigns an average weight of 1/16. Further, if the heart rate is below 60 bpm, the algorithm assigns an average weight of 1/8.

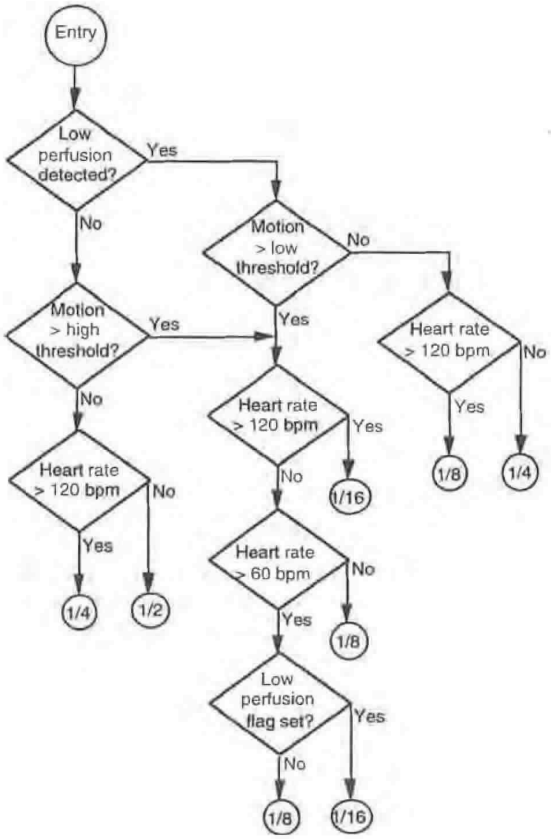

Figure 9.16. The weight determination algorithm (Conlon *et al* 1990).

If the heart rate is above 60 bpm and less than 120 bpm, the software has to differentiate between low perfusion with motion and motion alone. The algorithm checks the low perfusion flag and if set, assigns an average weight of 1/16, otherwise it assigns a weight of 1/8. The ensemble_averaging routine employs the

weight determination algorithm to find the average weight of the waveform and averages the buffered waveform with the composite averaged waveform stored in the microcomputer memory using a tail-weight average of the form (W × NEW) + (1 – W) × COMPOSITE, with W being the averaging weight. Because the averaged pulses are of varying duration, some pulses will overlay more points of the averaged waveform than others. Thus, the tail of the averaged waveform may not accurately reflect the most recent plethysmographic information. Hence the minimum and maximum of the averaged waveform were found only up to the minimum length of the last eight pulses. After determining the minimum and maximum values, the four beat average of peak-to-peak values are updated.

The algorithm then checks to ensure that the average has not fallen below the minimum low perfusion threshold. If it has, the pulse is considered lost. Then the algorithm checks if three non-low perfusion beats have been found. If so, it resets the low perfusion flag. If not, it checks if the current beat is a low perfusion beat, setting the perfusion flag appropriately. The algorithm then checks for motion in the last beat. If there is motion, it sends the four-beat, peak-to-peak average to the saturation_calculation algorithm routine. Otherwise, the last peak-to-peak value of the routine is sent to the saturation_calculation algorithm which calculates the oxygen saturation and displays it.

9.7 SPECTRAL METHODS OF ESTIMATING S_pO_2

Arterial oxyhemoglobin saturation (S_pO_2) values are currently computed using weighted moving average (WMA) techniques (Rusch *et al* 1994). These methods process the time domain signals and give a precision of no better than ± 2% (± one standard deviation). Researchers have explored other digital signal processing algorithms for improved estimation of S_pO_2. The fast Fourier transform (FFT) and discrete cosine transform (DCT) were identified as potentially superior algorithms (Rusch *et al* 1994) and useful to optimize the portability of pulse oximetry systems. Preliminary studies indicate that a 64-point FFT, with a 15 Hz sample rate, over a data collection period of 4.3 s was found to be the optimal combination for pulse oximetry applications, minimizing hardware expense, footprint, and power consumption. S_pO_2 values were calculated from a transform size of 64 points using

$$S_pO_2 = 110 - 25 \times R \tag{9.36}$$

where R is the ratio of the red and infrared normalized transmitted light intensity. The R value is

$$R = \frac{AC_R/DC_R}{AC_{IR}/DC_{IR}}. \tag{9.37}$$

The AC component is the signal variation at the cardiac frequency and the DC component is the average overall transmitted light intensity. The AC component is selected as the highest spectral line in the cardiac frequency band.

REFERENCES

Cheung P W, Gauglitz K, Mason L R, Prosser S J, Smith R E, Wagner D O and Hunsaker S W 1989 Feedback-controlled method and apparatus for processing signals used in oximetry *US patent 4,819,646*

Cheung P W, Gauglitz K, Mason L R, Prosser S J, Smith R E, Wagner D O and Hunsaker S W 1990 Method and apparatus for offsetting baseline portion of oximeter signal *US patent 4,892,101*

Conlon B, Devine J A and Dittmar J A 1990 ECG synchronized pulse oximeter *US patent 4,960,126*

Corenman J E, Stone R T, Boross A, Briggs D A and Goodman D E 1990 Method and apparatus for detecting optical signals *US patent 4,934,372*

Frick G, McCarthy R and Pawlowski M 1989 Waveform filter pulse detector and method for modulated signal *US patent 4,867,571*

Goodman D E and Corenman J E 1990 Method and apparatus for detecting optical signals *US patent 4,928,692*

Jaeb J P and Branstetter R L 1992 Composite signal implementation for acquiring oximetry signals *US patent 5,094,239*

Pologe J A 1987 Pulse oximetry: technical aspects of machine-design *Int. Anesthesiol. Clinics* **25** 137–53

Potratz R S 1994 Condensed oximeter system with noise reduction software *US patent 5,351,685*

Scharf J E and Rusch T L 1993 Optimization of portable pulse oximetry through fourier analysis *Proc. IEEE Twelfth Southern Biomedical Engineering Conf.* Tulane University pp 233–5

Smith R E 1989 Method and apparatus for processing signals used in oximetry *US patent 4,800,495*

Stone R T and Briggs D A 1992 Method and apparatus for calculating arterial oxygen saturation based plethysmographs including transients *US patent 5,078,136*

Yorkey T J 1996 Two 'rat rat' derivation *Personal communication* (Hayward, CA: Nellcor Inc)

INSTRUCTIONAL OBJECTIVES

9.1. Name the general sources of error that could be corrected with signal processing algorithms.
9.2. Explain the process of eliminating incident light intensity and thickness of the path as variables from Beer–Lambert law.
9.3. How is R_{OS} (Ratio of Ratios) estimated from the red and infrared optical signals?
9.4. Discuss the advantages of estimating R_{OS} using the derivative method over the peak and valley method. Explain how noise reduction is achieved using the derivative method.
9.5. Discuss the role of the construction–reconstruction process in improving the accuracy of S_aO_2 estimation.
9.6. Explain the function of the start-up interrupts.
9.7. Discuss the function of the five different states in the period zero subroutine.
9.8. Discuss the C-Lock ECG synchronization algorithm used in Nellcor®.
9.9. Explain the motion detection algorithm used in Criticare.
9.10. Name the advantages of using spectral methods in estimating oxygen saturation.
9.11. Explain the advantages of using ECG synchronization.

CHAPTER 10

CALIBRATION

Jeffrey S Schowalter

The calibration curves of *R* (Ratio of Ratios) values used to calculate oxygen saturation levels are critical to the accuracy of the entire pulse oximeter system. Without an accurate table of appropriate *R* values, the pulse oximeter has no way of determining oxygen saturation levels. As such, it is important to understand how the pulse oximeter calibration curve data are acquired. In addition, it is important to understand some of the past and present simulation techniques used to test the accuracy and functionality of pulse oximeters.

10.1 CALIBRATION METHODS

Chapter 4 states that Beer's law does not apply for a pulse oximetry system due to the scattering effects of blood. Therefore, pulse oximeter manufacturers are currently forced to use an empirical method of determining the percentage of arterial oxygen saturation for a given *R* ratio.

10.1.1 Traditional in vivo calibration

The traditional method of pulse oximeter calibration involves comparison of oximeter *R* value to the oxygen saturation ratio obtained from *in vivo* samples using human test subjects. In fact, this was the only method used to calibrate these devices up until 1993 (Moyle 1994). Although this method requires a variety of laboratory instrumentation and is typically done in a hospital setting, this data collection process is only required during the design and development of the device.

10.1.1.1 Procedure. In general, the calibration procedure is fairly straightforward. Test subjects are fitted with an indwelling arterial cannula, which is placed in the radial artery. A sample of blood is taken and analyzed with a CO-oximeter (see chapter 3) to determine the subject's levels of COHb and MetHb. In most cases, samples are taken over a broad population. Typically, data come from nonsmokers with background carboxyhemoglobin levels between 1% and 2%. Wukitsch *et al* (1988) mentions that subjects used for the Ohmeda Biox 3700 calibration had an average COHb level of 1.6% and a MetHb level of 0.4%.

Once a low level of COHb and MetHb are verified, the subject is also fitted with one or more pulse oximeter probes. The test begins by first ensuring that the subject is at the 100% oxygen saturation level. The subject breathes an oxygen/air mix so as to bring the arterial oxygen saturation level to 100% (as determined from arterial blood analyzed with the CO-oximeter). Oxygen saturation level is incrementally decreased by breathing gas mixtures of progressively less oxygen and more nitrogen. At each level where the pulse oximeter indicates a stable reading, an arterial blood sample is immediately taken and analyzed with the CO-oximeter. Corresponding readings are recorded and the data are then plotted with oxygen saturation percentage (as determined by the CO-oximeter) on the *y*-axis and *R* ratio (as determined by the pulse oximeter under test) on the *x*-axis yielding a traditional *R* curve as shown in figure 4.7. Typical values for the *R* ratio vary from 0.4 to 3.4 (Pologe 1989). A best fit calibration equation is then calculated from the data. If the pulse oximeter manufacturer has selected LEDs for their probes that have relatively narrow bands of center wavelength (as discussed in chapter 5), then only one curve is required. If they have a number of probes with differing red and infrared center wavelengths, then each probe with unique LED combination must be tested to obtain its unique curve characteristics. Some manufacturers have as many as 30 different probes.

10.1.1.2 Problems. One of the problems with this traditional method is the limited range of oxygen saturation that can be acquired. Ethical issues prevent intentional desaturation of healthy subjects below a certain point due to risk of hypoxic brain damage. As a result, saturation levels can only be reduced to around 60%. This leaves a large range of values on the curve that need to be calculated by extrapolation. This has the potential to induce errors and in fact, Severinghaus *et al* (1989) tested 14 pulse oximeter models and showed that most pulse oximeters performed poorly under relatively low levels of saturation (see chapter 11). Another problem of this calibration method is that it does not address the spacing and number of data points needed to build a curve. Moyle (1994) states that well spaced data points over the entire range from 100% down to 80% is more accurate than having many data points clustered between 95% and 100%.

There has been a great deal of debate over the years as to what the pulse oximeter is actually measuring and as such, a unique term has been created to specify an oxygen saturation reading as determined by a pulse oximeter. The problem is that the pulse oximeter uses two wavelengths to measure oxygen saturation. However, there are four common species of hemoglobin (Hb, HbO_2, COHb, and MetHb). Since there are routinely four light absorbing substances in a sample in a system which is assuming it is measuring only two substances, much discussion and misconception arise as to what the pulse oximeter is actually measuring (Pologe 1989). Equation (4.5) shows that functional S_aO_2 is the ratio of oxygenated hemoglobin to the sum of oxygenated and reduced hemoglobin. If a person were found that had no COHb or MetHb, this is what the pulse oximeter would measure. However, since some COHb and MetHb are typically present in everyone's blood, and these terms show up in the fractional S_aO_2 formula, it is easy to assume that the pulse oximeter is measuring fractional S_aO_2. However, this is not the case either. Moyle (1994) state that the conventional two-wavelength oximeter measures what should be defined as 'oxygen saturation as measured by a pulse oximeter', or S_pO_2.

Payne and Severinghaus (1986, p 47) state that the pulse oximeter reports

$$\frac{HbO_2 + COHb + MetHb}{HbO_2 + COHb + MetHb + Hb} \times 100\% \qquad (10.1)$$

and subtracting this quantity from 100% gives the percentage of reduced hemoglobin or Hb%. He suggests that to eliminate confusion pulse oximeters should display this value instead. If this were done, however, conventional thinking would have to change because readings would increase from zero as opposed to S_pO_2 readings which decrease currently from 100. The bottom line is that COHb and MetHb do have an effect on the accuracy of pulse oximeter readings (Reynolds *et al* 1993a,b) so they cannot be ignored as part of the calibration process.

10.1.1.3 Effects of COHb and MetHb. The effects of COHb and MetHb are typically handled in one of two ways. Some manufacturers subtract 2% for these factors so they are displaying fractional saturation (assuming a patient with nominal levels of COHb and MetHb) and others do not subtract this factor so they are displaying functional saturation (Ackerman and Weith 1995).

In a sense, the pulse oximeter will measure what it has been calibrated to measure based on the test subject profile. Tremper (Payne and Severinghaus 1986) states that Nellcor calibration data were originally based on five Olympic athletes in virtually perfect physical condition. These individuals probably had as low levels of COHb and MetHb as are found in humans. As such, anyone being tested with higher (normal) levels of COHb and MetHb yielded inaccurate readings. Today, by using a more representative subject to build the calibration curve, pulse oximeter manufacturers account for some of this during the calibration process. However, individuals with relatively high levels of COHb and MetHb will have inaccurate S_pO_2 readings.

10.1.1.4 Field calibration. Another issue of concern is field calibration. Using this technique, once the *R* curves are established, the transmitting wavelengths of the LEDs and corresponding *R* curve are provided via a coding resistor (see chapter 5) and as such only a two-point check to verify the correctly selected calibration curve is required. Typically this check will identify a problem due to a malfunctioning LED or photodiode or an incorrect coding resistor. However, other than this cursory check there is no type of field calibration done on the pulse oximeter. Cheung *et al* (1993) have proposed a system for compensating for the effects of temperature variations on the LEDs. Since the pulse oximeter photodiode cannot detect a shift in LED wavelength, the proposed system provides the capability for the temperature of the probe LEDs to be measured and thus an alternative calibration curve, as shown in figure 10.1, can be used for the new set of LED wavelengths. This system seems to be of limited usefulness however, since Reynolds *et al* (1991) have shown that the peak wavelength of a red LED will only shift by 5.5 nm and an infrared LED will shift by 7.8 nm with a temperature shift from 0 °C to 50 °C. Applying this information to a theoretical computer model based on Beer's law, causes negligible changes in accuracy of the pulse oximeter.

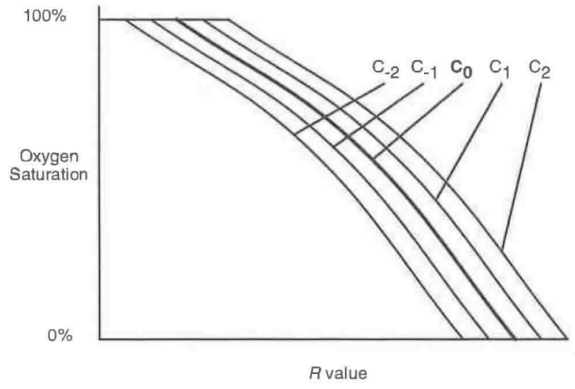

Figure 10.1 Temperature compensation curves as proposed by Cheung *et al* (1993). C values indicate different curves for use with different sensed temperatures.

10.1.2 In vitro calibration using blood

Figure 10.2 shows an *in vitro* test system that requires whole blood (Reynolds *et al* 1992). Blood is pumped through a cuvette acting as a model finger. The pulse oximeter probe is then attached to the model finger. Blood is pulsed within the system using a computer controlled peristaltic pump head capable of generating almost any shape of pulsatile waveform. Blood is oxygenated by passing through a membrane oxygenator using a gas mixture of O_2, N_2, and CO_2. The composition of the gas mixture passing through the membrane oxygenator is controlled with a gas mixing pump. A variety of model fingers were tried with the final model finger consisting of a cuvette made of two thin (0.5 mm) silicone rubber membranes and a rigid Plexiglas central section. When using whole blood, the model finger is covered with a diffuser made from translucent paper. Blood enters one end of the cuvette and flows in a thin (1 mm) layer through the cuvette over the fingertip end and back along the bottom side. Both inlet and outlet are tapered to prevent flow separation. The silicone rubber membrane is flexible enough such that pulsating blood produced volume changes in the tubes giving an AC/DC ratio in the physiological range. Readings from the pulse oximeter are recorded and a simultaneous sample taken from the sample port and analyzed by the CO-oximeter in a similar fashion to the procedure described in section 10.1.1. This system yields calibration values that are accurate to 50% and lower. Most pulse oximeters have no specified accuracy below 50%. One problem is that the system is sensitive to blood flow rate, due to changes in blood cell orientation with flow. This was verified using a hemoglobin solution instead of whole blood in the test system.

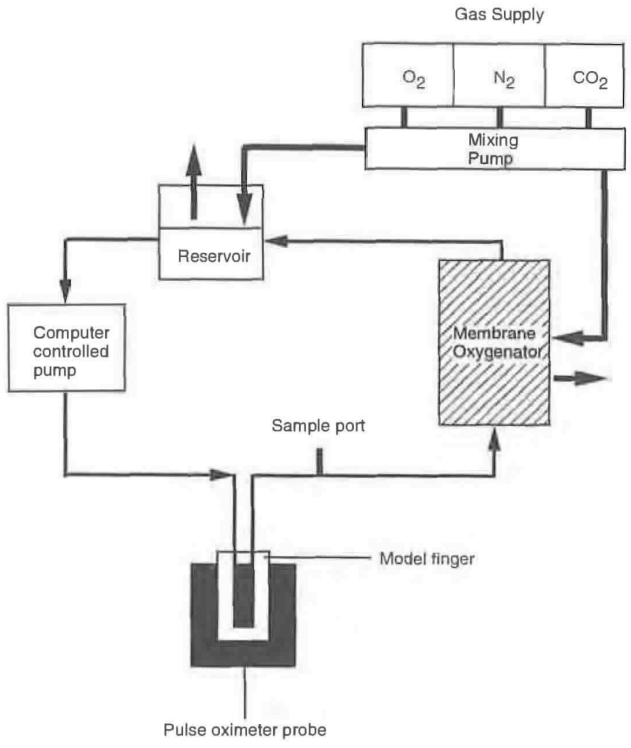

Figure 10.2 Block diagram of *in vitro* test system developed by Reynolds *et al* (1992).

10.2 TESTING SIMULATORS

Devices which check the functionality of pulse oximeters are becoming increasingly popular. Many of these devices use some type of artificial finger to verify that the pulse oximeter is functioning correctly. When pulse oximeters first came out, the only way technicians had to verify the functionality of the pulse oximeter was to use their own fingers. This, however, only indicates basic functionality at best with no way to control any parameters and with nothing with which to compare. Several devices have been developed that simulate the optical properties of the human finger and its pulsatile blood flow. In addition, optoelectronic systems, which simulate the human finger electronically, have also been developed. Finally, pulse oximeter manufacturers themselves have developed simple simulators that essentially simulate a probe's electron signals.