INVITED PAPER Special Section on Solid-State Circuit Design — Architecture, Circuit, Device and Design Methodology

Recent Progress of Biomedical Processor SoC for Wearable Healthcare Application: A Review

Masahiko YOSHIMOTO†, Fellow and Shintaro IZUMI††, Member

SUMMARY This paper surveys advances in biomedical processor SoC technology for healthcare application and reviews state-of-the-art architecture and circuits used in SoC integration. Particularly, this paper categorizes and describes techniques for improving power efficiency in communication, computation, and sensing. Additionally, it surveys accuracy enhancement techniques for bio-signal measurement and recognition. Finally, we have discussed the potential new directions for development as well as research.

key words: healthcare, low-power, SoC, wearable

1. Introduction

As the worldwide trend of aging population has increased in recent years, an increase in lifestyle diseases has occurred such as heart disease, hypertension, and diabetes [1]. Diabetes mellitus, a lifestyle-related disease, is said to increase the risk of developing dementia. For this reason, these diseases not only threaten health and longevity; they stand to raise national healthcare costs considerably. Early detection of cardiac and lifestyle diseases is necessary, but improving lifestyle habits is also important for disease prevention and for healthy life span extension. Wearable health care devices capable of routinely monitoring biological signals are expected to play an increasingly important role in meeting that social objective.

The heart beat is a biological signal used for cardiac disease detection, heart rate variability analysis [2], and exercise intensity estimation [3]. For measurements, an electrocardiograph (ECG) and a photoplethysmograph (PPG) are used. Furthermore, increased physical activity is particularly effective for lifestyle disease prevention. Triaxial acceleration data and other data are used to monitor physical activity in daily life. As presented in Fig. 1, improvement of a healthy lifestyle entails measurement of one’s own lifestyle, accumulation and analysis of data, and feedback about one’s own lifestyle based on those results. It is important to keep this cycle [4]. To make full use of the cycle effects, measured data must be obtained continuously over a long period. Required specifications for digital health system devices (Fig. 2) are a long operating time, high noise tolerance, wearable size, and reasonable cost. To fulfill these requirements, many researchers worldwide have undertaken VLSI and SoC development for digital health applications [5]–[27].

This paper describes technical trends of wearable processor SoC to meet the requirements above for healthcare application. Section 2 and Sect. 3 respectively present technical overviews and brief introductions of biosignals. Section 4 describes low-power design techniques. Section 5 explains accuracy improvement methods. These are followed by an explanation of future research directions in Sect. 6 and a summary in Sect. 7.

2. Technology Overview and Design Issues

Figure 3 portrays a block diagram of a typical
biomedical processor SoC consisting of a sensor front-end, an analog-to-digital converter (ADC), a general purpose processor core, memory, a DSP accelerator, RF circuits, and a power management circuit. Some of those components might not be integrated on the chip. Generally, a high-performance analog front-end circuit must be used to prevent the Signal-to-Noise ratio (SNR) degradation. Unfortunately, an analog circuit has large circuit area and high power consumption. Therefore, using high-performance amplifiers and high-quality analog filters is difficult. In accordance with Moore’s Law, the power of digital components decreases as process technology advances. In contrast, the power consumption of analog circuit does not decrease concomitantly. Therefore, weight is shifting to digital signal processing aimed at reducing the performance requirements of analog parts and minimizing the power consumption of the whole system. For further low-power implementation, hybrid architectures that use hardware accelerators to support software processing in microcomputers are the mainstream architectures in digital signal processing. Technology nodes with 180 nm to 130 nm processes have often been used: they provide sufficiently small leakage current. Table 1 summarizes several SoCs developed earlier for digital health applications [13]–[27].

The most important design issue is the operating time. Low power characteristics are indispensable for long daily continuous measurements under a finite battery lifetime. Even small batteries that realize wearable size must provide resources for long-term operation. The ultimate target of low power design is to guarantee continued operation in an autonomous energy supply environment by environmental power generation with an energy harvester using a heat source [10], [11], [28] and vibration [29]. The energy consumption of the entire sensor system node must be less than the environmental power generation energy. Figure 4 shows the annual trend of the power consumption characteristics of the developed SoCs summarized in Table 1, together with the power energy obtained by various environmental power generation. Power reduction in developed SoCs is progressing to a level where operation by some environmental power generation is possible, but in order to put it into practical use, the sensor node itself must be able to deal with the fluctuation of the generated power in environmental power generation [16].

Another important issue related to wearable sensors is high-accuracy design. Techniques can improve the signal level by improving noise tolerance characteristics [79]. A major noise that deteriorates the SNR of ECG or other

![Fig. 3 Block diagram of typical biomedical processor SoC.](image)

![Fig. 4 Technology trend.](image)

<table>
<thead>
<tr>
<th>Year</th>
<th>Technology</th>
<th>Power consumption</th>
<th>Supply voltage</th>
<th>Chip area (mm²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>180 nm</td>
<td>770 µW</td>
<td>0.77 V</td>
<td>2.3 x 2.3</td>
</tr>
<tr>
<td>2011</td>
<td>130 nm</td>
<td>500 µW</td>
<td>0.7 V</td>
<td>3.0 x 3.0</td>
</tr>
<tr>
<td>2012</td>
<td>90 nm</td>
<td>300 µW</td>
<td>0.7 V</td>
<td>2.5 x 2.5</td>
</tr>
<tr>
<td>2013</td>
<td>65 nm</td>
<td>150 µW</td>
<td>0.7 V</td>
<td>2.0 x 2.0</td>
</tr>
</tbody>
</table>

Table 1 Performance comparison of biomedical processing SoC

**Biologics**

<table>
<thead>
<tr>
<th>ECG</th>
<th>BCG</th>
<th>EMG</th>
<th>SMG</th>
<th>STR</th>
<th>ECG (3ch)</th>
<th>ECG (4ch)</th>
<th>ECG (5ch)</th>
<th>ECG (6ch)</th>
<th>ECG (7ch)</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

**Hardware components**

<table>
<thead>
<tr>
<th>ADC</th>
<th>10-bit</th>
<th>10-bit</th>
<th>12-bit</th>
<th>12-bit</th>
<th>8-bit</th>
<th>8-bit</th>
<th>12-bit</th>
<th>12-bit</th>
<th>8-bit</th>
<th>8-bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MCU</th>
<th>1-bit</th>
<th>1-bit</th>
<th>1-bit</th>
<th>1-bit</th>
<th>1-bit</th>
<th>1-bit</th>
<th>1-bit</th>
<th>1-bit</th>
<th>1-bit</th>
<th>1-bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RF</th>
<th>433.92 MHz</th>
<th>15.65 MHz</th>
<th>Inductive and 20-40 MHz</th>
<th>3.5 GHz</th>
<th>8-bit</th>
<th>8-bit</th>
<th>8-bit</th>
<th>8-bit</th>
<th>8-bit</th>
<th>8-bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Accelerometer (dedicated hardware) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Memory | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
signals is motion artifacts caused by skin stretching, pressure, and deformation. Those are about ten times as strong as the ECG signal, severely affecting the system operation. The other technique is a high-accuracy heartbeat and heart-beat variation estimation technique using different modalities [25], [26].

Wearable size is also an important issue. Inter-electrode distance minimization and battery miniaturization for biological information data measurement are effective for realizing a wearable sensor size. However, the former reduces the signal-to-noise ratio and degrades the measurement accuracy. ECG measures the potential difference on the body surface attributable to the electrical activity of the heart. The potential difference is decreased when the electrode distance is reduced. Thus, the SNR of ECG is affected by the electrode distance. The latter also presents a difficulty: the operation time can not be lengthened. Technology developments that resolve tradeoffs among miniaturization, accuracy, and operation time are proceeding in several lines of research.

3. Biosignal

Signals used to ascertain health parameters include Electrocardiogram (ECG), photoplethysmogram (PPG), Acceleration, and Electrodermal activity (EDA) including galvanic skin response (GSR) and bioimpedance (BIO-Z). ECG is used for heart rate analysis and cardiac activity measurement. EDA represents the stress level and skin activity.

3.1 Electrocardiogram (ECG)

An electrocardiograph [30]–[32] measures the potential difference on the body surface attributable to the electrical activity of the heart. The electrocardiogram records temporal changes in the action potential of the heart: generally P, Q, R, S, and T waves (Fig. 5). The P wave results from excitation of the atrium muscle, the Q, R, S waves respectively reflect the excitement of the ventricular muscles. The T wave shows their recovery disappearance. The number of R wave appearances during a given time is designated as the average heart rate (AHR). The time from the R wave to the next R wave is called the RR interval (RRI). The reciprocal of the RR interval is the instantaneous heart rate (IHR). Heart rate variability represents a spectrum analysis of IHR and reflects the heart rate fluctuation [5].

3.2 Photoplethysmogram (PPG)

A PPG sensor is commonly used in recent wearable devices to detect cardiovascular information including the heartbeat. The PPG sensor irradiates green light to the body surface and measures the amount of light absorption by hemoglobin related to the volume change of blood vessels [33] (Fig. 6). However, pulse wave signal measurement consumes large amounts of power because PPG requires one or more LEDs and a photodiode to measure the reflected light. Because the battery capacity of the wearable device is strictly limited, the power consumption should be reduced to realize an easy-to-use daily life monitoring system.

A trans-impedance amplifier (TIA, Fig. 7) is commonly used in recent PPG sensors [26], [34], [35]. The TIA converts the minute current $I_{in}$ generated by a photodiode into a voltage with transimpedance $R_f$. The amplified output voltage of this circuit $V_o$ is expressed as shown below.

$$V_o = R_f \cdot I_{in}$$

Actually, $C_f$ in Fig. 7, which is indispensable for negative feedback stability, is calculated as shown below [36].

$$C_f = \frac{C_{PD} + C_{in}}{2\pi R_f (GBW)}$$

Here, $C_{PD}$ represents the parasitic capacitance of the photodiode. Also, $C_{in}$ and $GBW$ respectively denote the input capacitance and the gain band width of an operational amplifier. $R_f$ is a resistance determining the trans-impedance.
gain. Pulse oximetry (SpO 2) sensor consists of similar measurement circuits. Cuff-less blood pressure measurement using PPG is a hot topic.

3.3 Acceleration

A triaxial accelerometer is used to monitor physical activity and thereby elucidate lifestyle habits (Fig. 8). For example, a physical activity classification algorithm [37], [38] uses two indices calculated using triaxial acceleration. Values of metabolic equivalents (METs) [39] have been used widely as indicators to quantify physical activity intensity. The METs value is the amount of oxygen consumed at rest. The amount of oxygen uptake at rest must be measured to obtain an accurate METs value. However, the method of gathering exhaled gases is too stressful to continue measurements over long periods. Therefore, a method of estimating METs values using a triaxial acceleration has been developed as a less burdensome measurement method.

3.4 Electrodermal Activity (EDA)

Electrodermal activity (EDA) [40], which reflects the electrical properties of human skin, is affected by perspiration. The signal is useful for quantifying autonomic nerve responses as a parameter of sweat gland function. Two EDA measurement approaches are used: exosomatic and endosomatic. Exosomatic techniques examine the apparent resistance change on the skin by application of a weak current between the electrodes. By contrast, endosomatic techniques measure the potential between one electrode pair on the skin.

A galvanic skin response (GSR) is a generic name for a skin potential response (SPR), a skin resistance change (SRC), and a skin conductance change (SCC). Actually, EDA is a more general term than GSR. For wearable sensors, GSR is often used.

The bio-impedance (Bio-Z) is the electrical impedance of body tissues. The Bio-Z can be used to measure the amounts of body fluids.

4. Low-Power Design

4.1 Communication Power Reduction

An RF circuit used for wireless communication is an important wearable sensor component. Bluetooth Low Energy (BLE), Zig-Bee, and Body Area Network (BAN) are candidate wireless communication standards for wearable biomedical sensors [10]. Many power reduction methods have been proposed for them. Nevertheless, it is difficult to reduce the transmission and receiving energy drastically for one-bit information. In contrast, the transmitted data size reduction is more effective for system-level power reduction. On-node processing is a modern approach to reduce the node data using data compression or feature extraction.

4.1.1 Compressed Sensing

The most straightforward way to reduce the transmitted data amount is data compression. Many data compression algorithms have been proposed. Compressed sensing [41] is especially attracting attention as a compression technique used for biosensing.

Compressed sensing assumes that the measured signal is sparse with low-rank. This algorithm recovers the sparse vector \( x \) from measured vector \( y \). The known measured vector \( y \) can be expressed as

\[ y = \Phi x, \]

where \( x \in \mathbb{R}^n \) is the original sparse signal and \( \Phi \in \mathbb{R}^{m \times n} \) is a random measurement matrix. Then, compressed signal \( y \) has smaller order \( m \) than the original signal order \( n \). Therefore, the transmitted data amount can be reduced. The process of reconstruction from \( y \) to \( x \) can be implemented in a base station (Fig. 9).

Dedicated hardware and sensing circuits have been proposed [42]–[46]. According to this theory, one reported PPG-sensing SoC [27] reduces the sampling rate. For example, 10 times compressed ratio can be achieved for ECG measurement with 2-bpm error as shown in Ref. [27].

4.1.2 Spectral Analysis

When the application layer only requires frequency features, one can reduce the data transmission amount using frequency analysis of the node.

Spectral analysis is used widely for digital signal processing for time series data such as vital signal and speech analysis. Discrete Fourier transform (DFT) is used widely for spectral analysis because the fast Fourier transform (FFT) algorithm obtains an accurate spectrum with few calculations. Moreover, it can be implemented easily using dedicated hardware because most calculations are simply multiplication and accumulation.

However, the DFT frequency resolution depends on
Compressed sensing for transmission data compression.

The sampling rate and the number of input data. More input points must be used to obtain higher frequency resolution. In some actual applications, the data length cannot be increased. Frequency characteristics must be extracted instantaneously from measured data, for instance, for biological signal measurements that require real-time data analysis. The short data length also reduces the processor and memory power consumption.

An autoregressive (AR) model can realize accurate and low-power spectral analysis with short data length. The AR model is a linear predictive modeling technique that assumes that the current value of a signal can be described by a finite linear aggregate of the preceding values. As an alternative to the Fourier transform, the AR model is useful for spectral analysis [47]. In this method, time series data are input and a specific linear system, the AR model, is output. Subsequently, the frequency spectrum is obtained by comparing the AR model with the variance of white noise. Using this method, the frequency resolution can be ascertained arbitrarily. This parametric method can yield higher resolutions than nonparametric methods such as DFT in cases with short data length.

However, the AR-model-based spectral analysis requires a larger number of calculations than FFT does, especially with the higher model order. An efficient hardware design of the AR model estimation has been studied [48]. Figure 10 shows the signal processing flow of the hardware-implemented fast Burg’s method [49]. The calculation process is mainly divided into two stages. One is the stage of calculating autocorrelations (AC stage). The calculation of this stage increases in proportion to the input data length. Another is a stage to estimate AR model (AR stage). Figure 11 shows dedicated hardware for Fast Burg’s method. Here, the AR order is a model order, which determines the complexity of expressible models. A higher AR order increases the computational amounts. As presented in Fig. 12, the implemented hardware can achieve almost identical FoM [51] compared to state-of-the-art FFT hardware [50]–[52].

$$\text{FoM} = \frac{\text{Normalized Energy}}{\text{Execution}}$$
Fig. 12  FoM comparison with earlier works of FFT hardware implementation [50]–[52]. AR order is set to 5. Input data length is set from 128 points to 2048 points.

By transmitting only the AR coefficients, it is possible to reduce the amount of data to be transmitted as compared with the case of transmitting the spectrogram calculated by FFT. Assuming application to heart rate variability analysis, if the data length is 256 and AR order is 64, the data amount is reduced to almost half. This is because the reduction rate is almost AR order/128. In that case, the accuracy deterioration is 1% or less.

4.1.3 Feature Extraction

Even when an application uses a feature in the time-domain, the transmitted data amount can be reduced by on-node feature extraction. For example, in an arrhythmia detection application from ECG, the wave form and heartbeat interval variation have important features. One earlier report [19] proposed an on-node PQRST wave extraction method. The extracted PQRST components are useful for determining Q-T distance and T wave height for arrhythmia detection.

Cuff-less blood pressure monitoring methods using ECG and PPG waveforms have also been proposed [53]–[56]. ECG measures the potential difference on the body surface attributable to electrical activity of the heart. By contrast, a PPG sensor irradiates green or red light to the body surface and measures the amount of light absorption by hemoglobin related to the blood vessel volume change. Therefore, several factors affect the PPG waveform and the peak delay between ECG and PPG: blood vessel hardness from the heart to the wrist, the arm position with respect to the heart, and the blood pressure. Moreover, one can infer blood pressure and blood vessel hardness using features extracted from ECG and PPG. Rather than transmitting raw data, the amount of transmitted data can be greatly reduced to one data transmission per heartbeat, by executing on-node

Table 2  Frequency components of biological signals

<table>
<thead>
<tr>
<th>Signal</th>
<th>Frequency Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrocardiogram (ECG)</td>
<td>0.1 – 150 Hz</td>
</tr>
<tr>
<td>Electroencephalogram (EEG)</td>
<td>0.5 - 60 Hz</td>
</tr>
<tr>
<td>Photoplethysmogram (PPG)</td>
<td>1 - 20 Hz</td>
</tr>
<tr>
<td>Electromyogram (EMG)</td>
<td>Few-kHz</td>
</tr>
</tbody>
</table>

Fig. 13  Conceptual timing diagram of normally-off computing with nonvolatile memory.

4.1.4 Machine Learning

By processing not only feature extraction but also recognition on node, the amount of transmission data can be minimized. Machine learning approaches using Support Vector Machine (SVM) [19] and neural networks [57] are usually employed.

In recent years, Deep Neural Networks (DNNs) have received particular attention as signal processing approaches for ECG and other bio-signals. Several reports of the relevant literature [58]–[63] describe that DNN can improve bio-signal detection and classification accuracy. Atrial fibrillation detection [64]–[66] and biometric identification using ECG [67] are hot issues in this research area.

4.2 Computation Power Reduction

4.2.1 Normally-Off Computing and Nonvolatile Memory

Power reduction techniques for signal processing circuits are described. Table 2 presents the frequency components of several biological signals. Because the frequency range of biological signals is low, both standby power reduction and sleep time maximization are effective for system level power reduction. Therefore, normally off computing technology was introduced using nonvolatile memory [68]–[71]. This technology is a control that turns off the power when the circuit is not used. In Fig. 13, a blue line represents a power consumption transition in the case of volatile memory, and a red line represents the power consumption transition in the case of a nonvolatile memory. In nonvolatile memory, because the power supply is gate-controlled in the standby mode, the leakage current can be greatly reduced. Izumi et al. used FeRAM as nonvolatile memory.

Figure 14 shows the average consumption current
when various mixed memories are used for data logging. By making FeRAM operate with intermittent operation (normally off), average current less than that of SRAM is realized in the heart rate and electrocardiographic region. Because the SRAM cannot turn off the power supply, standby leakage current exists so that the current continues to flow even in regions with low access frequency.

To realize both nonvolatility and high-speed access characteristics, a 6T-4C memory cell coupling an SRAM cell and ferroelectric capacitor was developed. Furthermore, nonvolatile technology was introduced into all the flip-flops embedded in Nonvolatile MCU [72, 73] as shown in Fig. 15. Thereby, the transition overhead is minimized so that the standby period can be maximized. Results show that the average operating current was reduced to 6.14 μA in the ECG processing circuit section [68] (see Fig. 16).

4.2.2 Hardware Accelerator

An efficient architecture to minimize power consumption is a combination of dedicated HW + general-purpose processors. Cortex M0 and RISC core are often used as general-purpose processors. They carry a dedicated H/W circuit for biometric signal processing with a high computation load. Functions integrated as an accelerator in electrocardiogram wave analysis and acceleration sensing are presented in Table 3, which includes autocorrelation, frequency analysis (FFT, AR model), noise removal, compression processing, filtering, feature extraction, and SVM.

Recent deep learning and neural network development opens the possibility of extracting more detailed information from data acquired using a wearable sensor. Because performance and accuracy are extremely important, deep learning, sensor fusion, and heavy calculation processing are executed off-platform on the cloud. Conventional signal processing and feature extraction processing are executed by the dedicated circuits of the node.

4.3 Sensing Power Reduction

4.3.1 ECG Sampling Rate Reduction

Nishikawa et al. proposed a sampling rate reduction method for heart rate variability analysis for low-power wearable devices [74]. With the introduction of interpolation processing and autocorrelation operation, the sampling rate is reduced to 32 Hz while maintaining a heartbeat interval error of 1 ms or less. Better results are obtained for PPG waveforms distributed at frequencies lower than ECG.

4.3.2 PPG Sampling Rate Reduction

Reduction of the sampling rate is particularly effective at reducing the power used by the PPG sensor. The power
Table 3 Functions integrated as accelerator

<table>
<thead>
<tr>
<th>Biosignal</th>
<th>Processing</th>
<th>Algorithm</th>
<th>Computation</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrocardiogram</td>
<td>Instantaneous heart rate/Heart beat</td>
<td>Peak detection</td>
<td>max/min. search</td>
<td>[81]</td>
</tr>
<tr>
<td></td>
<td>extraction</td>
<td>Auto correlation</td>
<td>multiply/accumulate or L1-norm.</td>
<td>[79]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Template matching</td>
<td>multiply/accumulate or L1-norm.</td>
<td>[68]</td>
</tr>
<tr>
<td></td>
<td>Arrhythmia detection</td>
<td>Waveform features</td>
<td>zero-cross detection, multiply/accumulate, max/min. search</td>
<td>[57-59]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neural network</td>
<td>multiply/accumulate</td>
<td></td>
</tr>
<tr>
<td>Heart-rate variability</td>
<td></td>
<td>FFT or Autoregressive model</td>
<td>multiply/accumulate</td>
<td>[2]</td>
</tr>
<tr>
<td>Noise reduction</td>
<td>Noise reduction</td>
<td>Wavelet transform</td>
<td>Add/Sub, Shift</td>
<td>[82-92]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Moving average</td>
<td>Add/Sub, Shift</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adaptive filtering</td>
<td>multiply/accumulate</td>
<td></td>
</tr>
<tr>
<td>Pulse wave velocity</td>
<td></td>
<td>Phase difference and waveform features</td>
<td>zero-cross detection, multiply/accumulate, max/min. search</td>
<td>[53-56]</td>
</tr>
<tr>
<td>Data Compression</td>
<td></td>
<td>Compressed sensing, adaptive sampling</td>
<td>multiply/accumulate</td>
<td>[41-46]</td>
</tr>
<tr>
<td>Acceleration</td>
<td>Gravity component elimination / Noise</td>
<td>2nd-order IIR filter</td>
<td>multiply/accumulate</td>
<td>[37, 38, 78, 100]</td>
</tr>
<tr>
<td></td>
<td>reduction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Physical activity</td>
<td>Multiple regression analysis</td>
<td>polynomial</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Behavior classification</td>
<td>SVM</td>
<td>polynomial</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neural network</td>
<td>multiply/accumulate</td>
<td></td>
</tr>
</tbody>
</table>

consumption of LEDs and PDs are dominant in PPG. Watanabe et al. proposed a low rate sampling technique combining Correlated Double Sampling technology with the linear interpolation process and autocorrelation operation described above [75]. A complete PPG sensor equipped with MCU, LED, PD, and other technologies was implemented. It achieved MAE error of 5 ms, which is sufficiently low for heart rate variability analysis with a 150 μs LED turn-on time and a 16 Hz sampling rate. Results show that the LED power was reduced by 85.9% compared with the conventional device. The power consumption of the whole PPG sensor was reduced to 22 μA.

4.3.3 Adaptive Sampling of Acceleration

Tsukahara et al. realized low-power consumption by adaptively changing the sampling frequency of the acceleration sensor according to behavior classification: 32 Hz, 16 Hz, 8 Hz, 4 Hz [76], [77]. In a simulation result resembling that of an actual use, about 56% power reduction effect was confirmed. Additionally, the error accompanying the frequency change was confirmed to be as small as 0.14 METs in the root mean square error (RMSE) with respect to the case in which the sampling frequency was fixed in physical activity estimation accuracy.

Nakanishi et al. proposed a highly accurate technique of behavior classification considering gravitational acceleration and a control method that directly changes the sampling frequency to a specific frequency according to behavior classification results [78]. The proposed method reduced the acceleration sensor active rate by about 23%. Results show that the behavior classification system operated at about 9.8 μA current consumption.

5. Accuracy Enhancement

5.1 Noise Reduction and Noise Tolerance

Figure 17 presents the inter-electrode distance dependence of the SNR [79]. The SNR improves as the distance between the electrodes increases. To reduce the sensor size, a shorter inter-electrode distance is better. A tradeoff exists between wearable size realization and SNR enhancement. Because ECG measures the potential difference on the body surface, the SNR is affected by the electrode distance. Therefore, it is apparent that a design with high noise tolerance not only improves the reliability of measured data: it also contributes greatly to sensor size reduction.

A widely used approach for IHR extraction from ECG is R-wave extraction using threshold determination. The Pan–Tompkins (PT) algorithm [82], which is commonly used for beat detection, uses band-pass filtering, differentiation, squaring, and moving window integration. Periodically, the threshold is adjusted automatically using QRS morphology and the heart rate. The SQRS [83] algorithms, which have been published in PhysioNet, can detect QRS based on the ECG slope. The SQRS uses band pass
filtering for noise reduction, which uses only the integer coefficient. The Discrete Wavelet Transform (DWT) [81], [84], [85] uses a wavelet transform with quadratic spline wavelet (QSW). The threshold is calculated using the root mean square value of the wavelet transform. This algorithm has been used in robust ECG monitoring LSIs [81], [86], [87]. The QSW requires few calculations and low hardware costs because it can be implemented merely by using adders and shift operators. The Quad Level Vector (QLV) algorithm [88] is used with dedicated hardware for ECG monitoring LSI [80], [89]. The QLV is generated using DWT and the adaptive threshold. Then the threshold is ascertained from the maximum mean deviation (MD) of prior heartbeats. The Continuous Wavelet Transform (CWT) algorithm [90]–[92] uses a Mexican hat wavelet in the frequency interval of 15–18 Hz. The R-peak can be extracted using the adaptive threshold, which is calculated using the modulus maxima of the CWT. This algorithm was also implemented in an earlier study [80].

The ECG waveform used for wearable sensors is often contaminated by noise from various sources such as baseline drift, hum noise, muscle noise, and motion artifacts, as portrayed in right portion of Fig. 18 [79]. Left portion of Fig. 18 presents frequency characteristics of PT, SQRS, and DWT with a 128 Hz sampling rate. A baseline drift and a hum noise can be removed easily using digital filters. However, unfortunately, muscle noise and motion artifacts are difficult to remove because they have a similar frequency range to ECG. Therefore, false positive and false negative detections occur frequently in the conventional threshold approach.

Short-term autocorrelation (STAC) has been proposed for IHR extraction to remove muscle noise and motion artifacts, as portrayed in right portion of Fig. 18 [79]. Autocorrelation [93], [94] and template matching [95] use similarity of QRS complex waveforms. They have no threshold calculation process. Figure 20 presents the performance of hardware implementation of the heart rate extraction algorithm [23] with MIT-BIH arrhythmia [96] and noise stress [97] database. The STAC technique has higher noise tolerance and minimum power overhead compared to previous results found for hardware implemented heartbeat detectors.

5.2 Recognition Accuracy

Heart rate can be measured using ECG, but also by PPG and BIO-Z. Combining these methods can considerably improve the reliability and robustness of measurement data. More accurate and reliable health assessment can be provided over a wider range. In such cases, electronic systems must be able to synchronize and measure multiple modalities accurately. Konijnenburg et al. proposed a battery-driven multi-parameter recording platform with built-in simultaneous support of ECG, BIO-Z, GSR, and PPG [25]. Because the data are collected on a single chip, very accurate synchronization between the data streams is possible. Correlation techniques between the data streams can be implemented. Such techniques support the study of blood pressure estimation by combining ECG and PPG measurements by pulse arrival time analysis. Furthermore, by combining different sensing modalities such as ECG, PPG, and BIO-Z, reliable estimates of hemodynamic parameters and heart rate variability are obtained.

Simultaneous measurement of PPG and ECG is also suitable for measuring the pulse transit time, which is an important factor of correlation for cuffless blood pressure measurement [26]. Sharma et al. proposed a wearable
platform using SoC that simultaneously captures ECG and PPG. They developed a device that wirelessly transmits heartbeats to smartphones every 2 s. It operates continuously for more than 5 days with a 250 mAh battery.

Using a multimodal approach combining heart rate data and triaxial acceleration data, a highly accurate algorithm of behavior classification and SoC for accurate human behavior analysis has been developed [78], [98]–[101]. For energy consumption estimation, the physical activity classification algorithm uses three indicators: heart rate reserve (% HR\text{reserve}), synthetic acceleration (ACC\text{fil}), and ratio of filtered acceleration to unfiltered acceleration (RFU). The heart rate reserve is expressed as follows:

\[
% \text{HR}_{\text{reserve}} = \frac{\text{HR}_{\text{max}} - \text{HR}_{\text{rest}}}{\text{HR}_{\text{max}}} - \text{HR}_{\text{rest}}
\]

The heart rate during activity (HR\text{act}): the mean value of averaged heart rate in every 10 s.

The heart rate at rest (HR\text{rest}): the mean value of the averaged heart rate in a supine condition during 10 min.

The maximum heart rate (HR\text{max}): \text{HR}_{\text{max}} = 220 - \text{Age}.

The synthetic value of filtered triaxial acceleration is expressed as follows:

\[
\text{ACC}_{\text{fil}} = \sqrt{X_f^2 + Y_f^2 + Z_f^2}
\]

Here, \(X_f\), \(Y_f\), \(Z_f\) are 3D components of filtered triaxial acceleration.

RFU (Ratio of filtered and unfiltered synthetic acceleration) is expressed as follows:

\[
\text{RFU} = \frac{\sqrt{X_u^2 + Y_u^2 + Z_u^2}}{\sqrt{X_f^2 + Y_f^2 + Z_f^2}}
\]

Here, \(X_u\), \(Y_u\), \(Z_u\) are 3D components of unfiltered triaxial acceleration.

Using these three indices, a decision tree is constructed to classify physical activity into five classes: sedentary, household, moderate (excluding locomotive), locomotive, and vigorous. The algorithm has been implemented on the MCU embedded in the SoC. Evaluation results show that the average classification accuracy for 21 activities is 91% [98], [101].

A highly reconfigurable analog front-end (AFE) IC supporting multi-modal (bio)signal monitoring has been reported [102]. The reconfigurable AFE channel occupies an area of 1.1mm\(^2\) while supporting four acquisition modes, i.e. biopotential (ExG), bio-impedance (BioZ), galvanic skin response (GSR) and general purpose analog (GPA). It consumes 36\(\mu\)W maximum from a 1.2V supply.

6. Future Direction

Daily monitoring by biological information sensors will be incorporated into the medical system in near future. In other words, personal biological information is accumulated in a database of medical institutions via a network, which will be useful for disease treatment and health management. In that case, what is important is a security function and authentication technology [62], [103]–[107]. With biometric authentication, memorization and password input are unnecessary. Also, since it is not necessary to have any physical ID information, the burden on the user can be reduced. In addition, there is little risk that other people steal passwords. “Impersonation” is difficult. These facts indicate that biometrics authentication is a sufficiently safe and convenient approach for practical applications. Cardiovascular information related to cardiovascular apparatuses has attracted attention as biological information which is not used in conventional biometric authentication methods. For example, the electrocardiogram (ECG) shows the extent of electrical excitation of the heart and can clarify the individual difference of the QRS components representing the excitation of the ventricle. It is possible to generate the ID by extracting the features of the QRS component from a certain period of the ECG.

In order to make biological information monitoring of each individual everyday comfortable and to promote data dissemination, it is necessary to make noncontact monitoring technology more practical. Although a lot of wearable sensor devices have been developed for ECG and heart rate measurement, these systems require pasting of wet electrodes directly onto the skin, which presents shortcomings in usability. Therefore, a non-contact heart rate monitoring is required. To realize remote heart rate monitoring, the microwave Doppler sensors [108]–[113] have been proposed to detect the heart velocity. Imaging-based methods using the color change of the face, which indicates the pulse beat, has also been proposed [114]. Although these methods present severe noise contamination problems, the heart rate can be detected without direct skin contact. A noncontact measurement method using capacitively coupled electrodes has been also proposed [115]–[119]. The capacitively coupled ECG sensor has high usability because it can take measurements even if an insulator such as clothing is interposed between the electrode and the human body.

As mentioned in Sect. 2, a battery-less system is desirable for lightweight miniaturization and lower cost. As an energy source of energy harvesting for a wearable device, heat, vibration, light, electromagnetic waves and the like are conceivable. The energy harvester from ambient electromagnetic waves cannot generate sufficient energy yet with small antenna size applicable to wearable sensors. The expected power generation is less than 1 \(\mu\)W [120]. Thermal power generation uses Seebeck element to convert the temperature difference across the element into electric power. Body temperature can be used as a heat source in this application. Since the temperature difference of the Seebeck element is generated from the difference between the body temperature and the outside temperature, efficiency decreases when the element is covered with clothes or the outside air temperature is high. The power generation is on the order of 10 \(\mu\)W to 100 \(\mu\)W per square centimeter with the body temperature [120]. Although the photovoltaic energy harvester
can generate almost same energy from ambient light [120], it limits the position of wearable sensors. In vibration power generation, vibration is converted into electric power using a piezo element or an electret. When converting the vibration of the human body into electric power by these elements, the power which can be taken out on average is less than 10 μW [121]. However, since it is greatly influenced by the exercise intensity, it is necessary to devise measures such as combining with exercise intensity and an algorithm that adaptively processes according to the amount of power generation. Also, since the frequency of the vibration of the human body is low and it is not constant, it is difficult to increase the efficiency. A structure for upconverting the vibration frequency and a technique for widening the resonance frequency are studied. Although heat or vibration is considered to have the widest applicable range, at the present time, there is a problem with the amount of power generation and stable operation. Consequently development of power efficiency improvement technology suitable for the human body is awaited.

7. Conclusion

Digital health is one of the most promising fields of application in the Cyber-Physical System. Supported by the information network infrastructure, it will be incorporated into the social system and grow into an essential infrastructure of medical systems. And VLSI sensor devices promote their progress. Battery-less, low power consumption, high accuracy, wearable, low cost, security function are indispensable requirements, and it will be further evolving from now on. We expect this review to be useful for researchers, engineers, and policymakers working in the area of VLSI design for the healthcare application.

Acknowledgments

This paper is based on results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

References


[101] Private communication with author of literature [98].


Masahiko Yoshimoto joined the LSI Laboratory, Mitsubishi Electric Corp., Itami, Japan, in 1977. During 1978–1983 he was engaged in the design of NMOS and CMOS static RAM. From 1984, he was involved in the research and development of multimedia ULSI systems. He earned a Ph.D. degree in Electrical Engineering from Nagoya University, Nagoya, Japan in 1998. From 2000, he was a Professor of the Dept. of Electrical & Electronic System Engineering of Kanazawa University, Japan. From 2004, he was a professor of the Dept. of Computer and Systems Engineering in Kobe University, Japan. His current activities specifically emphasize research and development of ultra-low power sensor device for IoT application. He holds 70 registered patents. He served on the Program Committee of the IEEE International Solid State Circuit Conference during 1991–1993. Furthermore, he served as Guest Editor for Special Issues on Low-Power System LSI, IP and Related Technologies of IEICE Transactions in 2004. He was a chair of the IEEE Solid State Circuits Society (SSCS) Kansai Chapter during 2009–2010. He is also a chair of the IEICE Electronics Society Technical Committee on Integrated Circuits and Devices from 2011–2012. He received R&D100 awards from the R&D magazine for the development of the DSP and the development of the real-time MPEG2 video encoder chipset, respectively, in 1990 and 1996. He received the 21st TELECOM System Technology Award in 2006. He is a Fellow of IEICE.

Shintaro Izumi respectively received his B.Eng. and M.Eng. degrees in Computer Science and Systems Engineering from Kobe University, Hyogo, Japan, in 2007 and 2008. He received his Ph.D. degree in Engineering from Kobe University in 2011. He was a JSPS research fellow at Kobe University from 2009 to 2011, and an Assistant Professor in the Organization of Advanced Science and Technology at Kobe University from 2011 to 2018. Since 2018, he has been an Associate Professor in the Institute of Scientific and Industrial Research, Osaka University, Japan. His current research interests include biomedical signal processing, communication protocols, low-power VLSI design, and sensor networks. He has served as a Chair of IEEE Kansai Section Young Professionals Affinity Group, as a Technical Committee Member for IEEE Biomedical and Life Science Circuits and Systems, as a Student Activity Committee Member for IEEE Kansai Section, and as a Program Committee Member for IEEE Symposium on Low-Power and High-Speed Chips (COOL Chips). He was a recipient of 2010 IEEE SSCS Japan Chapter Young Researchers Award.