國立交通大學

電子工程學系 電子研究所碩士班

碩士論文

整合於可攜式腦心監護系統之高能源效率
4 通道獨立成份分析處理器

A Power-Efficient VLSI Design of a 4-Channel Independent Component
Analysis Processor for Portable Brain-Heart Monitoring Systems

研 究 生： 傅致中

指導教授： 方偉騏 教授

中 華 民 國 九 十 九 年 十 月
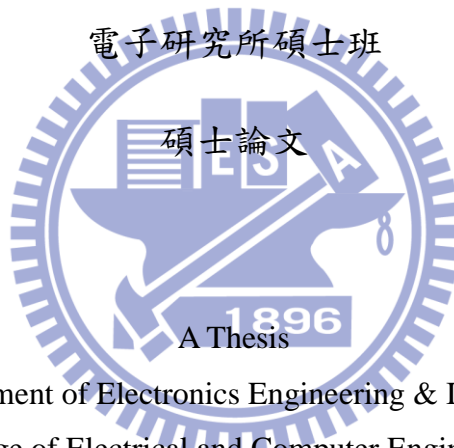
整合於可攜式腦心監護系統之高能源效率

4 通道獨立成份分析處理器

A Power-Efficient VLSI Design of a 4-Channel Independent Component
Analysis Processor for Portable Brain-Heart Monitoring Systems

研究生： 傅致中　　Student : 　Chih-Chung Fu

指導教授： 方偉騏　　Advisor : 　Wai-Chi Fang

國立交通大學　電子工程學系

電子研究所碩士班

碩士論文

A Thesis

Submitted to Department of Electronics Engineering & Institute of Electronics

College of Electrical and Computer Engineering

National Chiao Tung University

In partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Electronics Engineering

October 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年十月

整合於可攜式腦心監護系統之高能源效率

4 通道獨立成份分析處理器

學生：傅致中　　指導教授：方偉騏 教授

國立交通大學電子工程學系 電子研究所碩士班

中文摘要

近年來快速增加的老年人口比例已然成為各國必需面臨的重要問題，整合型健康照護系統已經成為電子領域發展的重點。本論文由三個應用情境包括緊急醫療需求(如救護車上之緊急量測)、長期觀測與照護(老年退化性疾病)與腦認知科學的研究為出發點，提出一針對腦電訊號(EEG)、心電訊號(EKG)處理分析與擴散光學腦部影像重建(DOT)所構成之整合型系統之雛型設計，進以推動針對此三類系統的可攜性整合型醫療儀器之發展。

由於生理電信號中最微弱的腦電信號通常與肌電信號(EMG)中的眼動信號與眨眼信號一起混合並量測，以獨立成份分析為方法的人工雜訊濾除技術已發展許久。但由於獨立成份分析的運算複雜度過高，腦波的應用通常受到離線運算的嚴重限制。本論文針對所提出整合型系統中的腦電信號處理所使用的四通道即時獨立成份分析器之設計與實作以一完整章節加以詳述。由於可攜式儀器的基本需求即為低功率與低成本，多種設計技巧與最佳化規格分析如三重循環記憶體的配置、鏡像非線性查表單元的設計與 ICA 訓練、成份萃取間的管線排程皆被用來降低功率消耗與硬體成本。此 ICA 硬體設計已由聯電 90 奈米製程下線並測試。晶片的核心面積為 0.58 平方毫米。量測數據顯示若使用 80 Hz 的取樣頻率，並使用 0.5 MHz 的工作頻率與 0.6 V 的核心電壓時，可達到最低 0.312 毫瓦的功率消耗。

此獨立成份分析模組也與一心率變異率分析器、一擴散光學影像重建模組一

同整合於一實驗性腦心監護系統之中。由前端訊號擷取模組所得到的生醫訊號被傳送至相應的即時運算引擎進行分析處理，處理完的結果與原始訊號皆由一無損失性生醫信號壓縮模組進行資料壓縮，再以一商業藍牙模組傳至臨近的生醫資訊工作站進行 3D 顯像與遠端觀察與診斷。此生醫訊號壓縮模組的平均壓縮率可達 2.5，此壓縮率可被視為在無線傳輸上面的功率節省。系統中的資料流順序主要由一固定優先權資料選擇器與一三級向後資料流控制機制所影響，而這樣的設計也能提高各模組的輸出緩衝記憶體使用率，如此一來可以造成較少的傳輸緩衝記憶體使用。獨立成份分析與心率變異率分析引擎皆以真實生理訊號驗證，並顯示優良的分析結果。而腦影像重建引擎則以一模型來顯示其分析與真實情況的一致性。

關鍵字：資訊最大化、獨立成份分析、腦波訊號處理、心率變異率分析、擴散式光學影像重建、整合型生醫系統、藍牙傳輸、可攜式系統、數位信號處理

A Power-Efficient VLSI Design of a 4-Channel Independent Component Analysis Processor for Portable Brain-Heart Monitoring Systems

Student: Chih-Chung Fu   **Advisor: Dr. Wai-Chi Fang**

Department of Electronics Engineering and Institute of Electronics

National Chiao Tung University

Abstract

Since the twenty-first century, the fast increment of an aged population has become a worldwide problem. Therefore, integrated health-care systems have become an important topic for electrical engineers. In this thesis, focusing on three application scenarios including emergency medical care (e.g. EEG, EKG measurements on ambulance or DOT for fast cerebral hemorrhage check), long-term observation and monitoring (for patients suffer from chronic ailment) and researches on brain and cognitive science, we propose a preliminary design of an integrated health-care system comprising electroencephalogram (EEG) and electrocardiogram (EKG) signal analysis and processing together with diffuse optical tomography for brain imaging. The significance of this system is to enable the practical development of such portable health-care devices for brain heart monitoring.

Since the EEG is the feeblest one of all physiological electrical signals usually contaminated by ocular artifacts (e.g. eye-blink artifact and eye-movement artifact), the artifact removal techniques using independent component analysis (ICA) has been developed for a long time. Because of the compelling computation complexity of ICA

algorithm directly inherits from the possible dependency in each channel, applications that analyze EEG signals are usually heavily restricted by the off-line ICA computation. One complete chapter is used to describe the design and implementation of the 4-channel ICA processor employed in the proposed integrated system as the EEG processing element. Since the two basic requirements for portable instruments are low-power and low-cost, various design techniques and optimized specification analyses like three-bank circular memory allocation, an mirrored non-linear lookup unit and the operation pipelining between the ICA training and component extraction are all adopted to reduce the power consumption and hardware cost. The designed ICA processor is fabricated using UMC 90 nm CMOS technology, and the core area of the chip is 0.58 mm^2. Performance measurements done by Agilent 93000 SoC Tester have shown that when using 80 Hz sample rate, 0.5 MHz operation frequency and 0.6 V core power, the lowest power consumption of 0.312 mW is achieved under the worst cast of 512 training iterations.

Together with an HRV and fNIR-DOT processor, the designed ICA processor is integrated in an experimental brain heart monitoring system. EEG, EKG and near-infrared signals acquired from the analog front-end IC are processed in real-time or bypassed according to user configurations. Processed data and raw data are compressed by a lossless biomedical data compressor and sent to a remote science station by a commercial Bluetooth module for further analysis, 3-D visualization and remote diagnosis. The biomedical signal compressor achieves an average compression ratio (CR) of 2.5 which is translated into power saving during wireless transmission. The data flow in the system is mainly controlled by a prioritized data selector and a three-stage backward handshaking mechanism, and the design can increase the utilization of the output buffers inside each processor so that the data transmission

buffer can be reduced. The ICA and HRV processor are verified by real EEG and EKG signals while the DOT processor is verified by an experimental model.
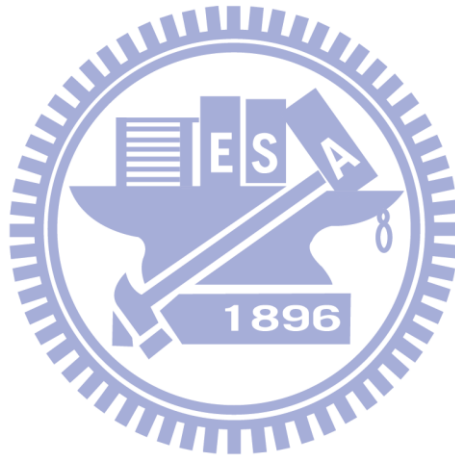
誌 謝

　　本論文的完成要感謝許多背後支持的力量，首先是兩年以來我的指導老師方偉騏教授在研究上的關心與指導。謝謝方老師特別從 NASA 回到母校交大來從事教職。方老師提供了理想的研究環境與豐富的資源，使得研究能夠很快的進入狀況。除了課業上的支持外，老師也讓我學習到諸多寶貴的經驗如與人相處之道與團體合作的技巧。另外，要感謝我的口試委員：林進燈教務長與范倫達教授，他們所提出的建議與指教使得本論文更為完整與充實。而林老師與范老師實驗室中的學長姊們也給予了非常寶貴的經驗。

　　研究的過程中，陳秋國博士提供了最直接的支持。無論是生活上的問題或是學業上的困難，秋國學長都給與非常實在的建議與指導。學長對學弟們的關心使的實驗室的運作順暢與和諧。感謝實驗室中一起幫忙完成 ICA 設計的同學：秋國、Ericson、少彥、適揚以及兩位宗翰。有了你們的幫助，晶片的設計、整合、測試才能夠順利完成。另外，實驗室中的學弟們：康適、世揚、偉業、佳慶、敬儒、奕仲、瑞傑，在學業的最後階段你們真的提供了非常大的助力。如果之後有我幫的上忙的地方，我也一定不遺餘力地提供協助。
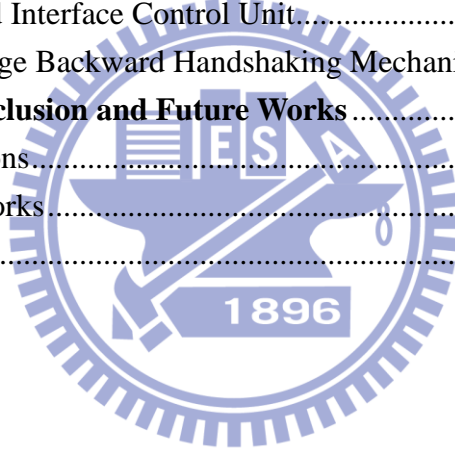
　　最後，我的爸爸、媽媽、姊姊、姊夫所給予的支持與關心使我能在研究時沒有後顧之憂，並且渡過所有難關。特別感謝我的女友軒念，少了妳的陪伴我可能無法如此順利的達成這艱難的任務。

謹以本論文獻給我的家人及所有關心我的師長、同學與朋友們。

Contents

List of Figures

x

List of Tables

Chapter 1   Introduction

Since the twenty-first century, the fast increment of an aged population is emerging as a preeminent worldwide phenomenon. Most of the elderly suffer from chronic ailments and illnesses related to central nervous system (CNS) in their later life. To ease the problems caused by insufficient nursing personnel, many health-care systems focusing on biomedical signal processing and monitoring have been developed. Traditional EEG measuring equipments require the patients to be confined to a small area due to their large size, bringing tremendous inconvenience to them. Therefore, integrated portable health-care systems have become an increasingly important topic.

1.1 Three Common Human Health Indicators

Recent studies have shown that combined analysis of EEG together with heart rate variability or brain fNIR can aid in better diagnosis and treatment. For example EEG and HRV data were jointly analyzed for the automatic detection of seizures in newborns [1] and sleep apnea in hospital patients [2], while the advantage of combined analysis of EEG and fNIR data for cognitive rehabilitation and post traumatic stress syndrome was presented in [3]. Despite these studies indicating the need for joint monitoring of brain fNIR, EEG and HRV, an integrated brain-heart monitoring SoC solution has not been developed.
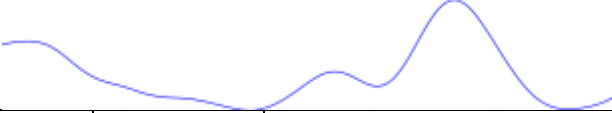
1.1.1 Electroencephalogram
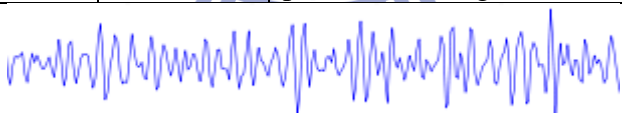
Human cerebral cortex has a large amount of neurons, and the activities of these neurons have some degree of regularity, so pairs of electrodes on the scalp can be recorded from the cerebral cortex on the next generation of potential changes. The potential changes are composed of the **Electrical Rhythms** and **Transient Discharge**. These changes of waveform are called brain potential signal.

Electroencephalogram (EEG) is a non-invasive tool for recording of electrical activity along the scalp produced by the firing of neurons within the brain. EEG measurement of different locations, frequency ranges, amplitudes, waveforms and periodicities can be used to distinguish different generation of EEG. The EEG provides important information about the health of the central nervous system (CNS), particularly in the newborn [4]. In medical application of neurology, it is common to use EEG to diagnose such as epileptic, coma, encephalopathy and brain death.

A typical voltage range of EEG signal is about 10 μVolt to 100 μVolt, and the frequency domain is less than 100 Hz. In addition, there are five major bands of continuous rhythmic sinusoidal EEG activity. They are recognized as δ (delta, below 4Hz), θ (theta, 4-8Hz), α (alpha, 8-12Hz), β (beta, 12-30Hz) and γ (gamma, above 30Hz) waves, and their characteristics are listed in Table 1.1. (activities below or above these range is likely to be taken artifactual noise, under standard clinical recording techniques)

Table 1.1   Classification of the continuous rhythmic sinusoidal EEG activities

Type	Frequency Range (Hz)	Common Amplitude Range (V)	Description
Delta (δ)	0 ~ 4	-	Delta is often associated with the very young and certain encephalopathies and underlying lesions. It is seen in stage 3 and 4 sleep.
Theta (θ)	4 ~ 7	Below 20μ	Theta is associated with drowsiness, childhood, adolescence and young adulthood. This EEG frequency can sometimes be produced by hyperventilation. Theta waves can be seen during hypnagogic states such as trances, hypnosis, deep day dreams, lucid dreaming and light sleep and the preconscious state just upon waking, and just before falling asleep.

Alpha (α)	8 ~ 12	20μ ~ 80μ	Alpha is characteristic of a relaxed, alert state of consciousness. For alpha rhythms to arise, usually the eyes need to be closed. Alpha attenuates with drowsiness and open eyes, and typically come from the occipital (visual) cortex. An alpha-like normal variant called mu is sometimes seen over the motor cortex (central scalp) and attenuates with movement, or rather with the intention to move.
Beta (β)	12 ~ 30	Below 20μ	Beta rhythms with low amplitude or multiple and varying frequencies is often associated with active, busy or anxious thinking and active concentration. Rhythmic beta with a dominant set of frequencies is associated with various pathologies and drug effects, especially benzodiazepines.
Gamma (γ)	30 ~ 100	-	Gamma rhythms may be involved in higher mental activity, including perception, problem solving, fear, and consciousness.

In clinical experiments, EEG signal is displayed based on the location of the electrode that affects the amplitude, phase and frequency. EEG measurements can be divided into monopolar derivation and bipolar derivation. The monopolar derivation uses a probe electrode and a reference electrode fixed on the scalp surface, and it measures the relative value of the probe electrode and reference electrode. The bipolar derivation is induced with two probe electrodes and a reference electrode fixed on the scalp. The potential difference between the two probe electrodes detects EEG signal reflects and acquires relatively small EEG amplitude.

EEG system is no longer limited to the interception and analysis of the signals. Today, there are many proposed identification systems and human brain wave techniques for medical diagnosis and treatments. For example, Fuzzy C-means (FCM) algorithm can be used to identify epileptic seizures and cerebral palsy [5]. However EEG signal is very sensitive, and very often may be contaminated by various disturbances like ocular artifact, EMG and electrical noise from nearby instruments [6], and they largely restricts the precision of the identifications and analysis.

Fortunately, this problem can be alleviated by algorithms the independent component analysis (ICA) algorithm [7], which separates artifacts and noise from the measured EEG signals. Wavelet [8] and Spatially-Constrained [9] techniques can be used to identify the artifacts and eliminate them. As a result, we can derive clean EEG signals after the noise channel is eliminated and remixed. However, the computation complexity is so intense that real-time ICA analysis is not feasible for pc-based bio-science station. Therefore, in recent years, the researches on hardware implementation of ICA engines are blooming.

1.1.2 Near-Infrared Spectrogram on Human Tissue

Since DOT (Diffuse Optical Tomography) technology is a kind of non-invasive and real-time radiography, it has been widely used to detect tumors in the breast and imaging the brain in recent years. Many researches are involved in DOT technology and have made rapid progress and development. DOT can be used to detect oxygenated hemoglobin (HbO) and deoxygenated hemoglobin (Hb) concentration and volume with bi-wavelength Near-Infrared. Therefore, in clinical application, the main uses of DOT are monitoring blood flow, blood volume, oxygen saturation, tumors within the brain, and detecting breast cancer [10]. By measuring different characteristics of the diffused near-infrared, DOT can be generally divided into three

main categories: the Continuous Wave (CW), Frequency Domain and Time Domain.

Table 1.2   Characteristics of the three main types of diffuse optical measurements

Type	Advantages	Disadvantages
Time Domain (TD)	1. Spatial resolution 2. Penetration depth 3. Most accurate separation of absorption and scattering coefficients	1. High sampling rate 2. Instrument size and weight 3. Stabilization and cooling 4. Cost
	Example Uses: Imaging cerebral oxygenation and breast imaging	
Frequency Domain (FD)	1. Relatively low sampling rate 2. Relatively accurate separation of absorption and scattering coefficients	1. Penetration depth 2. Instrument size and weight 3. Cost
	Example Uses: Cerebral and muscle oximetry, breast imaging	
Continuous Wave (CW)	1. Low sampling rate 2. Instrument size, weight and simplicity 3. Low cost	1. Penetration depth 2. Difficult to separate absorption and scattering coefficients
	Example Uses: Finger pulse oximeter, functional brain experiments, cerebral hemorrhage	

Table 1.2 shows the characteristics of different DOT systems. The CW system provides advantages such as low cost, high portability, low power consumption and computation overhead, although lack of depth information [11]. The volume of the CW-DOT system can be miniaturized which is the biggest advantage than the other algorithms. Therefore, there exists the possibility of implementing hardware architecture for CW systems. However, little literature has been published on hardware architecture of CW-DOT signal processing. Most of CW-DOT systems post-process the signal by computer such as [12] and [13]. This will demolish the feature of portability, and make it difficult to miniaturize the system.

1.1.3 Electrocardiogram

Electrocardiography (EKG) is an interpretation of the electrical activity of the heart over time captured and externally recorded by skin electrodes [14]. It is a noninvasive recording produced by an electrocardiographic device. The EKG is an essential tool for health professionals in making a diagnosis of abnormal heart

rhythms when one is suspected.

The EKG works mostly by detecting and amplifying the tiny electrical changes on the skin that are caused when the heart muscle depolarizes during each heart beat. Usually more than 2 electrodes are used and they can be combined into a number of pairs. The output from each pair is known as a "lead". Different types of EKG measurements can be referred to by the number of leads that are recorded, for example 3-lead, 5-lead or 12-lead EKGs. A 12-lead EKG is one in which 12 different electrical signals are recorded at approximately the same time and will often be used as a one-off recording of an EKG, typically printed out as a paper copy. 3- and 5-lead ECGs tend to be monitored continuously and viewed only on the screen of an appropriate monitoring device, for example during an operation or whilst being transported in an ambulance.

A typical EKG waveform shown in Figure 1.1 is composed of P peak, QRS peaks and T peak. How these peaks in the EKG are originated is explained in Table 1.3.
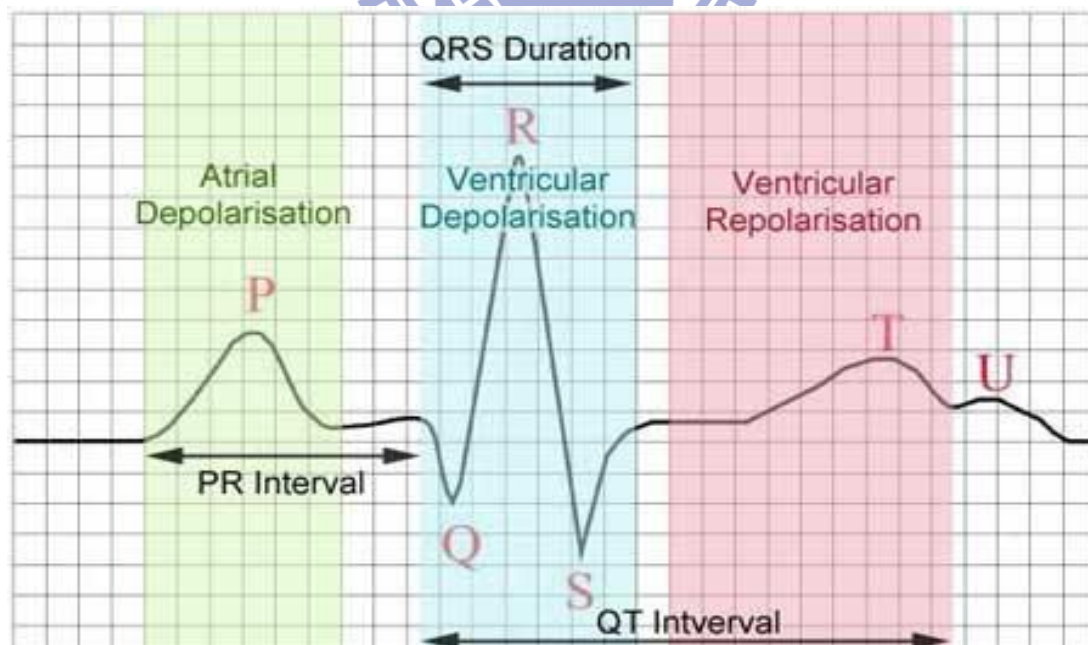
Figure 1.1   A typical EKG waveform

Table 1.3   Different peaks in a typical EKG waveform

Peak	Origination and Description
P	Systole (depolarization) of the atrium.
QRS	Systole (depolarization) of the ventricle. The amplitudes of QRS peaks are usually larger than P and T peaks, because the muscle of the ventricle is stronger.
T	Repolarization of the ventricle.

Intervals between each peak can indicate the health of heart. The most common used three kinds of interval are listed in Table 1.4 with their usages and descriptions.

Table 1.4   Different types of peak interval that can be used to evaluate the health of the heart

Interval	Description
RR interval	Two adjacent R peaks can represent for the heart rate. The normal heart rate is between 50 bpm to 100 bpm (beat per minute).
PR interval	It is usually 120 to 200 ms long. The PR interval reflects the time the electrical impulse takes to travel from the sinus node through the AV node and entering the ventricles. The PR interval is therefore a good estimate of AV node function. ● A long PR interval (of over 200 ms) may indicate a first degree heart block. Prolongation can be associated with hyperkalemia or acute rheumatic fever. ● A short PR interval may indicate a pre-excitation syndrome via an accessory pathway that leads to early activation of the ventricles, such as seen in Wolff-Parkinson-White syndrome. ● A variable PR interval may indicate other types of heart block.
QT interval	The QT interval generally represents electrical depolarization and repolarization of the left and right ventricles. A prolonged QT interval is a risk factor for ventricular tachyarrhythmias and sudden death.

Heart rate (HR) is a non-stationary value; it can vary as the body's need to absorb oxygen and excrete carbon dioxide changes, such as during exercise or sleep. The measurement of heart rate is used by medical professionals to assist in the diagnosis and tracking of medical conditions. It is also used by individuals, such as athletes, who are interested in monitoring their heart rate to gain maximum efficiency from their training.

Heart rate variation (HRV) may contain indicators of current disease, or warnings about impending cardiac diseases [14]; it has proved to be a valuable tool to investigate the sympathetic and parasympathetic function of the ANS, especially in diabetic and postinfarction patients [14]. Sympathetic activity is associated with the low frequency range (0.04–0.15 Hz) while parasympathetic activity is associated with the higher frequency range (0.15–0.4 Hz) of modulation frequencies of the HR. This difference in frequency ranges allows HRV analysis to separate sympathetic and parasympathetic contributions evidently [14].

On the other hand, time-frequency parameters calculated using wavelet transform and extracted from the nocturnal heart period analysis appeared as powerful tools for obstructive sleep apnoea syndrome diagnosis. Time-frequency domain analysis of the nocturnal HRV using wavelet decomposition could represent an efficient marker of obstructive sleep apnoea syndrome [15].

1.2 The Need for an Integrated Health-Care Solution

In recent years, many portable bio-signal acquisition systems are proposed in academic research, and in the business community, plenty of tiny bio-status recorder systems have already been sold in the health-care market for years. A major imperfection of such systems is that an integrated multiple bio-signals recording device, for example simultaneously recording EEG, ECG and fNIR, is not proposed lately. Data of multiple kinds of bio-signals recorded in synchronized timeline is much more useful than single kind recorded bio-signal. For example, modern people suffer more from the pressure living their life aberrantly, and many of them go to the hospital for evaluation of their sleeping quality, because they experience insomnia. In fact, according to one epidemiological study [16], about one-third of the adult population exhibit at least one symptom of insomnia. In the sleeping quality

evaluation process, the patients are required to be monitored by EEG, ECG, EMG, fNIR, respiration, posture and sound. When the observing target need to go to the restroom, he has to bring the EEG measuring headgear together with him, so it is very inconvenient, and at the same time the recording is interrupted although the information at the period is not really necessary. However, in such situation, an integrated portable system with the ability to wirelessly transmit data to the science station will bring the following advantages:

- Much more comfortable for the observing target

- Short wiring for feeble physiological electrical signals

- Decrease the chance of inaccurate measurement caused by the discomfort

- Decrease the chance of sensor fall off

- Lower the cost

- Extend the applicable range of the system

Such systems allow the observed target to move freely around an area with health-care science station service, while not losing the quality of the measurement. Therefore, the integrated portable health-care device has become an inevitable trend. In next section, three major target application scenarios will be point out, and in chapter 4, a complete architecture for portable brain-heart monitoring system will be proposed.

1.3 Application Scenarios

Before presentation of the detail design of the proposed system, three major target application scenarios are shown below. The system targets the usage in emergency condition, long-term medical observation and monitoring and potential researches on brain function and cognitive science, and the implemented chip can be integrated in portable devices in the following application scenarios.

1.3.1 Emergency Use

Current brain imaging technologies used in hospitals are not useful in emergency conditions. For example, we cannot equip a CT on an ambulance, but cerebral hemorrhage is a common case for the car accident victims. When we can use portable CW DOT devices on the ambulance to check if the patient suffers from cerebral hemorrhage, the medical personnel on the ambulance can inform the medical team in the hospital if cerebral hemorrhage happened on the patient.

Another common case happens when an infant falls over with his head knocking on the ground. Sometimes brain injury is hard to find out without the help of medical equipments, but people frequently ignore the possible dangerous behind this not to mention a baby that can barely talk. With the help of a cheap and portable DOT device, we can avoid many tragedies like this.

1.3.2 Long-term Observation and Monitoring

The traditional electroencephalogram acquisition systems and DOT systems are very large in size and also very heavy. Between the patient and the acquisition instruments, there are many connecting wires (one wire for each channel). Figure 1.2 in the next page shows the traditional electroencephalogram and DOT acquisition instruments.

| (a) | (b) |

<div align="center">(c)   (d)</div>

Figure 1.2   Traditional EEG measuring equipment (b) Traditional EEG System enlargement at head part (c) Frequency Domain DOT System (d) Frequency domain DOT system enlargement at head part

Such systems bring tremendous inconvenience to the patients especially in long-term monitoring cases, for example the patients suffer from seizure or the measuring object in the research of degenerative brain diseases. For these patients, wearing the wired headgear means the restriction of free movement. Furthermore, these equipments are always located in facilities like hospitals and health-care centers, so the patients and observing targets are required to stay in a restricted area. Figure 1.3 shows an advanced setup for our brain monitoring system.
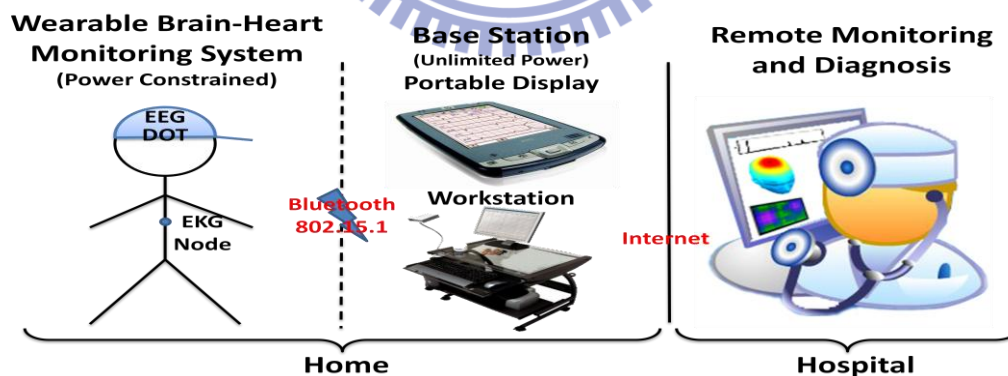
Figure 1.3   Application scenario for the proposed wearable brain monitoring system

The system is interfaced with an analog front-end chip and a commercial Bluetooth module. NIR, EEG and ECG signal are acquired, processed and transmitted via the Bluetooth wireless link. Biomedical data received at the base station will be

decompressed, displayed in real-time on the screen of the science station, and finally stored into non-volatile storage media for further off-line processing, analysis and diagnosis. In addition, the data can be sent from the base station to a remote workstation for online monitoring and diagnosis by doctors in hospitals.

1.3.3 Research on Brain and Cognitive Science

The past researches have clearly shown that electroencephalogram (EEG) contains important information about human cognitive process. Human brain cognitive science has become a very important and challenging research direction since the twenty-first century, and the related topic includes brain computer interface [17], artificial intelligence, electronic prosthesis and even artificial neural tissues. Not only has the electroencephalogram shown its potential capabilities for human brain cognitive research, the information from functional near-infrared also provide cortical hemodynamic response and shows which area of the brain is currently active. The flexible nature of DOT, which uses a wearable imaging cap (Figure 1.2 (d)), makes it well-suited to human brain studies in enriched environments and for a wide range of behavioral paradigms and activations [18], including visual [19], during motor tasks [20], somatosensory system [21], auditory [22][23], and language [24][25] Although the instruments like Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) together with multiple channel EEG signal can provide significantly valuable brain activity information, but their high cost and huge size result in the low availability for academic research.

1.4 Importance of this Work

A power and area efficient 4-channel ICA processor involved with various design and optimization techniques is presented in this thesis. Low-power and low-cost make it possible to be used in portable devices, which are usually constrained by the limited

power consumption. Prolonged battery life also makes it possible to be integrated in portable long-term observation and monitoring systems.

An experimentally integrated system which comprises a novel functional near-infrared (fNIR) diffuse optical tomography system for taking brain image, an independent component analysis (ICA) processor for artifact removal from electroencephalogram (EEG) signal, and a heart rate variability (HRV) analysis processor for electrocardiogram (ECG) signal is implemented. The significance of this SoC is to enable practical developments of portable real-time brain-heart monitoring systems.

1.5 Organization of the Thesis

The organization goes as follows. In chapter 2, a hardware and power efficient 4-channel ICA processor is presented from the theory of ICA to the tape-out summary of the designed processor. Various design and optimization techniques including optimized data windowing, 3-bank circular memory allocation, an optimized mirrored non-linear lookup unit and operation pipelining between the ICA training and component extraction are all presented in section 2.2. Performance analysis of the 4-channel ICA processor using both certain super-gaussian random pattern and real EEG signal with or without eye-blink artifacts is demonstrated and compared with off-line result analyzed by EEGLab in section 2.3 with an example of eye-blink artifact removal. In section 2.4, the tape-out summary of the designed ICA processor using UMC 90nm CMOS technology is presented. In addition to functional verification and power consumption analysis, an FPGA-based Testbed built to provide an experiment and demonstration platform is shown in section 2.4.3.

In chapter 3, an experimental brain-heart monitoring system with the ICA processor integrated is shown. Overall system architecture and specifications are

documented in section 3.1 and 3.2. Three bio-signal processing processors that perform ICA, DOT and HRV function are briefly described in section 3.3. From section 3.4 to 3.6, designs and behaviors of the other system peripherals and data flow control units are described in detail. Finally we conclude the current achievements and the future works in chapter 4.

Chapter 2   4-Channel Independent Component Analysis Processor

In this chapter, the design of a 4-channel independent component analysis (ICA) processor adopted in the brain-heart monitoring system proposed in Chapter 3 is shown. It is employed to perform artifact removal from 4-channel EEG signals. First, the independent component analysis algorithm and the reason why the system comprises an ICA processor are introduced and described in section 2.1. The design of a hardware and power efficient 4-channel independent component analysis processor is then shown in section 2.2. In section 2.3, the performance analysis using super-gaussian random pattern and real EEG pattern recorded by NeuroScan system is presented and described in detail to prove the validity of the designed processor, and comparison with other hardware ICA implementation are also done In section 2.3. Finally in section 2.4, the physical information, tape-out summary, chip testing and power measurement of the fabricated chip using UMC 90nm CMOS technology is presented.

2.1 Independent Component Analysis

In recent years, independent component analysis (ICA) had been applied to different signal processing applications, such as speech enhancement, telecommunication, feature extraction and artifact removal from signals. In this section, we will define and explain blind source separation (BSS) problem and introduce the principles and the algorithm of ICA.

2.1.1 Blind Source Separation

Blind source separation (BSS), also known as blind signal separation, is the separation of a set of signals from a set of mixed signals, without the aid of information (or with very little information) about the source signals or the mixing process. Blind signal separation relies on the assumption that the source signals do not

correlate with each other. For example, the signals may be mutually statistically independent or decorrelated. Blind signal separation thus separates a set of signals into a set of other signals, such that the regularity of each resulting signal is maximized, and the regularity between the signals is minimized (i.e. statistical independence is maximized). In Figure 2.1, the problem definition of the blind source separation (BSS) is shown. The original source signals are mixed in the path transmitted to the sensors. Processed by the blind source separation, the unknown source signals may be revealed again.
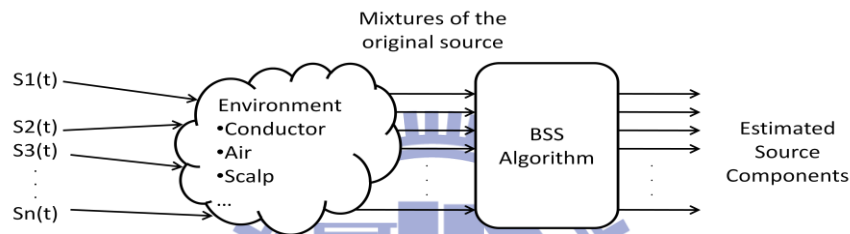
Figure 2.1   Problem definition of the blind source separation (BSS)

Blind source separation (BSS) problems are universal for signals acquired from natural sources for example light intensity, sound wave and electrical potential. The first problem of blind source separation (BSS) had been proposed to solve individual speech at a noisy cocktail party. It posed a serious problem that humans cannot understand their conversation when more than one person is speaking, and the concept using ICA to separate the two independent speeches is shown in Figure 2.2. There are also many similar problems, for example, EEG signals disturbed by artifacts, a speech in a noisy environment, and the signals include various sources that we want and do not want in the same location where sensors are placed.
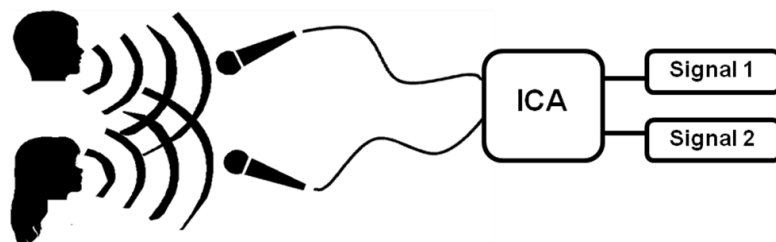
Figure 2.2   The concept of using ICA for speech separation in a cocktail party

2.1.2 Entropy and Mutual Information

Entropy and Mutual Information are the foundation of information maximization algorithms that includes the Infomax independent component analysis (Infomax ICA) algorithm chosen for the designed 4-channel ICA processor. Therefore, before introduction to Infomax ICA algorithm, the concept of Entropy and Mutual Information used are first clarified in this section.

The **entropy** is a function of random variable which tries to describe the "unpredictability" of a random variable with non-negative values, and the value of entropy becomes zero when the input random variable is "certain" when predicted. The entropy function of a random variable X is defined by:

$$H(X) = \sum_X P(X) \log \frac{1}{P(X)} \tag{2.1}$$

The X in equation 2.1 is a random variable and P(X) is the probability distribution function (PDF) of X, also known as the probability density function. The value of entropy is not only influenced by the value territory of the random variable, but also influenced by the probability distribution.

The **joint entropy** is defined as the entropy of a joint probability distribution of two or more random variables, or a multi-valued random variable. For two random variables X and Y, the joint entropy is defined by

$$H(X, Y) = \sum_{x,y} p(x, y) \log \frac{1}{p(x,y)} \tag{2.2}$$

The **conditional entropy** is a statistics that summarizes the randomness of Y given knowledge of X. It is defined by:

$$H(Y|X) = \sum_{x,y} p(y|x) \log \frac{1}{p(y|x)} \tag{2.3}$$

The conditional entropy H(Y|X) is H(Y) without H(X) and the conditional entropy H(X|Y) is H(X) without H(Y). If there exists correlation between H(Y) and H(X),

then $H(X, Y) \leq H(X) + H(Y)$, and if two random variables X and Y are statistically independent, the joint entropy H(X,Y) equals to the sum of the independent entropies , that is H(X)+H(Y).

Mutual information is a quantity that measures a relationship between two random variables that are sampled simultaneously. The mutual information of two discrete random variables X and Y is defined by

$$I(X, Y) = \sum_x \sum_y p(x, y) \log \frac{p(x,y)}{p_1(x)p_2(y)} \qquad (2.4)$$
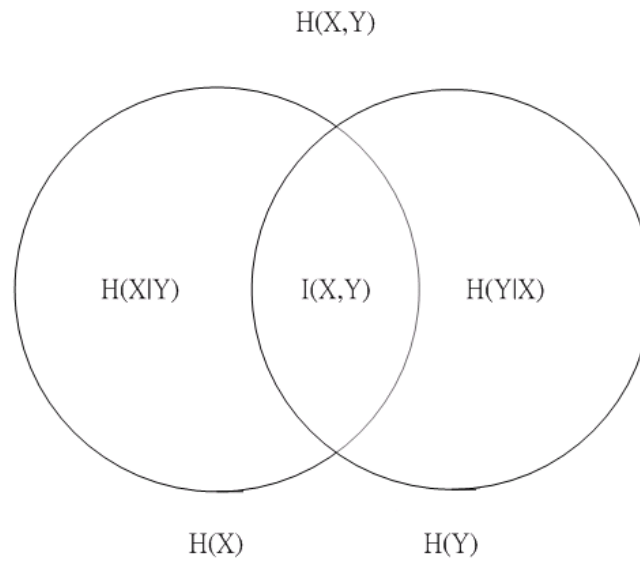
Figure 2.3   Entropy relationship presented by the concept of set

From the concept to express entropy using set in Figure 2.3, we define the entropy defined both in H(X) and H(Y) to be mutual information I(X,Y), with

$$H(X,Y) = H(X)+H(Y|X) = H(Y)+H(X|Y)$$
$$=H(X)+H(Y)-I(X,Y) \qquad (2.5)$$

Therefore the **mutual information** is then defined as

$$I(X,Y)=H(X)-H(X|Y)$$
$$=H(Y)-H(Y|X)$$
$$=H(X)+H(Y)-H(X,Y)$$
$$=H(X,Y)-H(X|Y)-H(Y|X) \qquad (2.6)$$

18

Minimization of the mutual information will lead to the maximization of the independence between the random variables, and if the mutual information between two random variables is zero, the two random variables are statistically independent. In the other hand, the smaller mutual information we derive, the more statistically independent two random variables we will have. With this conclusion, we will introduce the Infomax ICA algorithm using the concept in the next section.

2.1.3 Infomax ICA

Recently, blind source separation by Independent Component Analysis (ICA) has received attention because of its potential applications in signal processing such as in speech recognition systems, telecommunications and medical signal processing. The goal of ICA is to recover independent sources given only sensor observations that are unknown linear mixtures of the unobserved independent source signals. In contrast to correlation-based transformations such as Principal Component Analysis (PCA), ICA not only decorrelates the signals (2nd-order statistics) but also reduces higher-order statistical dependencies, attempting to make the signals as independent as possible.

There have been two different fields of research considering the analysis of independent components. On one hand, the study of separating mixed sources observed in an array of sensors has been a classical and difficult signal processing problem. The work on blind source separation by Jutten, Herault and Guerin (1988) where result in an adaptive algorithm using simple feedback architecture, and its learning rule was based on a neuromimetic approach that is able to separate simultaneously unknown independent sources. Furthermore, Comon (1994) introduced the concept of independent component analysis and proposed cost functions related to the minimization of mutual information between the sensors. On the other hand and in parallel to blind source separation studies unsupervised learning

rules based on information-theory have been proposed by Linsker (1992), Becker and Hinton (1992) and others. This idea is to maximize the mutual information between the inputs and outputs of a neural network. This approach is related to redundancy reduction which was suggested by Barlow (1961) as a coding strategy in neurons. Each neuron should encode features that are statistically independent from other neurons. This leads to the notion of factorial code that has been explored for the visual processing strategy by Attik (1992). Nadal and Parga (1994) showed that in the low-noise case, the maximum of the mutual information between the input and output of a neural processor implied that the output distribution was factorial. Roth and Baram (1996) and Bell and Sejnowski (1995) independently derived stochastic gradient learning rules for this maximization and applied them, respectively to forecasting and time series analysis, and the blind separation of sources. Bell and Sejnowski (1995) were the first explaining the blind source separation problem from an information-theoretic viewpoint and applying them to separation and deconvolution of sources.

Extensive simulations have been performed to demonstrate the power of the learning algorithm. However, instantaneous mixing and unmixing simulations are problems and the challenge lies in dealing with real world data. Makeig et al. (1996) applied the original Infomax algorithm to EEG and ERP data showing that the algorithm can extract EEG activations and isolate artifacts. Jung et al. (1997) show that the extended Infomax algorithm is able to linearly decompose EEG artifacts such as line noise, eye blinks, and cardiac noise into independent components with sub- and super-Gaussian distributions. McKeown et. al. (1997) have used the extended ICA algorithm to investigate task-related human brain activity in fMRI data. By determining the brain regions that contained significant amounts of specific

temporally independent components, they were able to specify the spatial distribution of transiently task-related brain activations.

As previously described, ICA algorithm not only decorrelates a signal (second-order statistical independence), but also reduce the dependency in higher order statistics. In other words, the goal of independent component analysis is to find a linear but not necessarily for the orthogonal coordinate system which can express multi-dimensional data. The independent multivariate random processing is involved in the various components of all the order of the statistics. The computation complexity of the statistics over second-order is significantly high, so using an adaptive learning unit which replaces the higher order calculation with adding a non-linear function g( ) after each component make u(t) approaches to s(t). The g( ) can be single-tone non-decreasing functions with values between 0 and 1, such as sigmoid and hyperbolic tangent function. Bell and Sejnowski (1995) [26] presented the Infomax ICA algorithm, which is suitable for separation of super-Gaussian sources.

The Infomax method performs linear ICA based on a principle of maximum information preservation. However, it can also be seen as a maximum likelihood method, or as a method based on the minimization of mutual information between y = g(u)(g( ) is sigmod function) The goal of Infomax ICA is to find an unmixing weight *W* that can be used to estimate the independent component signals. The adaptive learning ICA algorithm is shown in Figure 2.4. The input X is the signal mixture, and *U* is calculated by

$$U = W*x \tag{2.7}$$

The output Y is expressed by

$$Y = g(u) \tag{2.8}$$
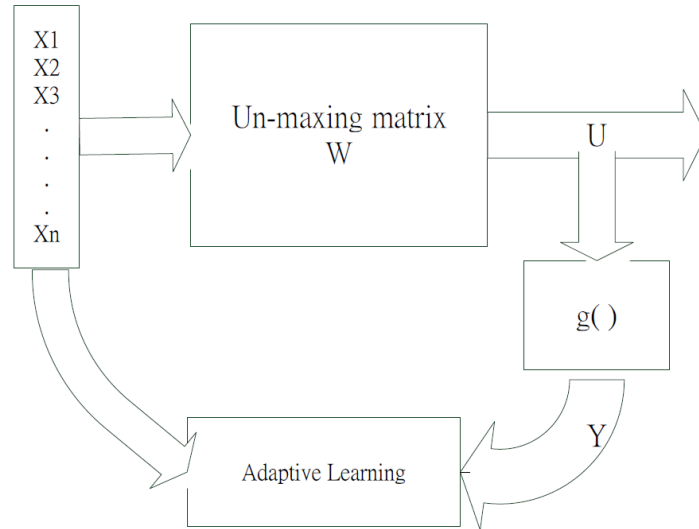
Figure 2.4   Adaptive learning independent component analysis

The information which output Y includes X is defined in equation 2.6.

$$I(Y,X)=H(Y)-H(Y|X)$$

Where H(Y) is the entropy of the output, while H(Y|X) is whatever information the output has which didn't come from the input. In the case, we have no noise, the mapping between X and Y is deterministic and H(Y|X) has its lowest possible value. We differentiate equation 2.6 as shown in equation 2.9

$$\frac{\partial}{\partial w}I(Y,X)=\frac{\partial}{\partial w}H(Y) \tag{2.9}$$

H(Y|X) do not rely on w, so the part $\frac{\partial}{\partial w}H(Y|X)$ is zero. The information transfer between the input X and output Y is maximized by maximizing the joint entropy of the output, H(Y). As discussed above, finding a function Y=f(X) that maximizes I(X,Y) is equivalent to maximizing H(Y).

The equation for the joint entropy of the output Y is the sum of the individual entropies minus the mutual information between them, and it's a way of information maximization that reduces statistical dependence. The equation is expressed as

$$H(Y)=H(y1)+H(y2)+.....+H(yn)-I(y1,y2,....yn) \tag{2.10}$$

22

From Equation 2.10, maximizing the output entropy H(Y) is equivalent to minimizing the mutual information of the extracted components y_i, and individual outputs will move towards the statistical independence. The transformation between y and u is a monotonic transform, and information maximization uses this concept to achieve the goal of ICA. The coefficient of adjustment from Infomax algorithm which was proposed was based on the conventional rules of stochastic gradient method. Gradient method will be introduced as follow.

First, consider an input variable, x, which passed through a transforming function, g(x), to produce an output variable, y. The probability density function (PDF) of the output p(y) can be expressed as a function of the PDF of the input p(x).

$$p(y) = \frac{p(x)}{\frac{\partial y}{\partial x}} \tag{2.11}$$

The entropy of the output, H(y), is given by

$$H(y) = E(\frac{1}{\ln p(y)}) = \int_{-\infty}^{\infty} p(y) \ln \frac{1}{\ln p(y)} dy \tag{2.12}$$

From equation 2.11 and 2.12, H(y) can be expressed as

$$H(y) = E(\ln \frac{\partial y}{\partial x}) - E[\ln p(x)] \tag{2.13}$$

In order to maximize the entropy of y by changing W, maximizing the first term which is the average log of how the input affects output need to be focused on. This can be done by considering the training set of x to approximate the density p(x), and deriving a stochastic gradient descent learning rule:

$$\Delta w \propto \frac{\partial H}{\partial w} = \frac{\partial}{\partial w}(\ln \frac{\partial y}{\partial x}) = (\frac{\partial y}{\partial x})^{-1} \frac{\partial}{\partial w}(\frac{\partial y}{\partial x}) \tag{2.14}$$

Y is the output of sigmoid function, $y = \frac{1}{1+e^{-u}}$, and $u = wx + w_0$ in which the input is multiplied by weight w and added to a bias w_0, and the calculation is

$$\frac{\partial y}{\partial x} = wy(1-y) \tag{2.15}$$

$$\frac{\partial}{\partial w}(\frac{\partial y}{\partial x}) = y(1-y)(1+wx(1-2y)) \tag{2.16}$$

Equation 2.15 and 2.16 are the learning rule, and equation 2.17 summarize it as

$$\Delta w \propto w^{-1} + x(1-2y) \tag{2.17}$$

To see the advantages of approach in artificial neural networks, the analysis of multi-input and multi-output can be described as follow. The input x is a vector in a network, and a weight matrix W and a monotonically transformed output vector $y = g(Wx + w_0)$, the multivariate probability density function of y can be written

$$p(y) = \frac{p(x)}{|J|} \tag{2.18}$$

where $|J|$ is the absolute value of the Jacobian of the transformation.

$$J = \det \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_n} \end{bmatrix} \tag{2.19}$$

For sigmoid function, $y = \frac{1}{1+e^{-u}}$, and $u = Wx + w_0$, and $\frac{\partial y}{\partial u} = y(1-y)$, so equation 2.18 can be written

$$J = (\det W) \prod_{i=1}^{n} y_i (1 - y_i) \tag{2.20}$$

From equation 2.12, the joint entropy of the output is

$$H(y) = E[\ln\frac{1}{p(y)}] = E[\ln|J|] - E[\ln p(x)] \tag{2.21}$$

Weights can be adjusted to maximize H(y), as before, and they only affect the $E[\ln|J|]$.

$$\Delta w \propto \frac{\partial H(y)}{\partial w} = \frac{\partial}{\partial w}(\ln|J|) = \frac{\partial}{\partial W}\ln|\det W| + \frac{\partial}{\partial W}\ln\prod_{i=1}^{n}|y_i(1-y_i)| \tag{2.22}$$

For the full weight matrix, the definition of inverse of a matrix and the adjoint matrix, adj W , is the transpose of the matrix of cofactors.

$$\frac{\partial}{\partial W}\ln\left|\det W\right|=[W^T]^{-1} \qquad (2.23)$$

The product splits up into a sum of log-terms, only one of which depends on particular w.

$$\frac{\partial}{\partial W}\ln\prod_{i=1}^{n}|\,y_i(1-y_i)\,|=\frac{\partial}{\partial W}\ln\prod_{i=1}^{n}|\frac{\partial y}{\partial x}|=\prod_{i=1}^{n}(\frac{\partial y}{\partial x})^{-1}\frac{\partial}{\partial W}(\frac{\partial y}{\partial x})=(1-2y)x^T \quad (2.24)$$

The resulting learning rules, equation 2.21 can be written

$$\Delta W=[(W^T)^{-1}+(1-2y)x^T] \qquad (2.25)$$

Equation 2.25 involves the calculation of inverse matrix, so the computation complexity is high. To solve this problem equation 2.25 is multiplied by W^TW which rescales the result and the new learning equation becomes

$$\Delta W=[(W^T)^{-1}+(1-2y)x^T]W^TW=[I+(1-2y)u^T]W \qquad (2.26)$$

Because it avoids the calculation of matrix inverse, the computation time is reduced significantly.

2.2 Design of the 4-Channel ICA Processor

In physiological electrical signal measurement, the observed signals are always the superposition of independent source signals. In addition, EEG signals are especially vulnerable and easily contaminated by artifacts such as eye movement, eye blink, power line noise and muscle (EMG) noise due to its signal strength of micro volt scale which pose serious problems in analyzing and interpreting the EEG recording [27]. ICA has already shown to be an effective, powerful and applicable method for EEG de-noising, which is able to separate EEG components and artifact components to different channels. Components recognized as the artifact can be removed easily by generating a mixing matrix with the weight to the artifact channel

set 0, and the remixed EEG signals will be clean with only small influence caused by the artifact due to the limitation of the algorithm. To acquire clean EEG signal for observation or analysis, a 4-channel ICA processor for artifact removal is designed and employed in the proposed brain-heart monitoring system. With the designed ICA processor, real-time applications using EEG signals become feasible and more robust.

2.2.1 Overall Architecture of the ICA Processor

Figure 2.5 shows the overall architecture of the four-channel ICA processor. It comprises four main processing units: a first stage buffering and calculation unit (S1), a whitening unit (WU) for calculation of the whitening matrix, an ICA training unit (TU) for unmixing weight training, and an ICA computation unit (CU) for constructing the whitened unmixing weight matrix and the resulting components extracted. Operation pipelining is applied between the data processing (S1, WU and TU) and the ICA output calculation (CU). Therefore, the hardware can be used efficiently.
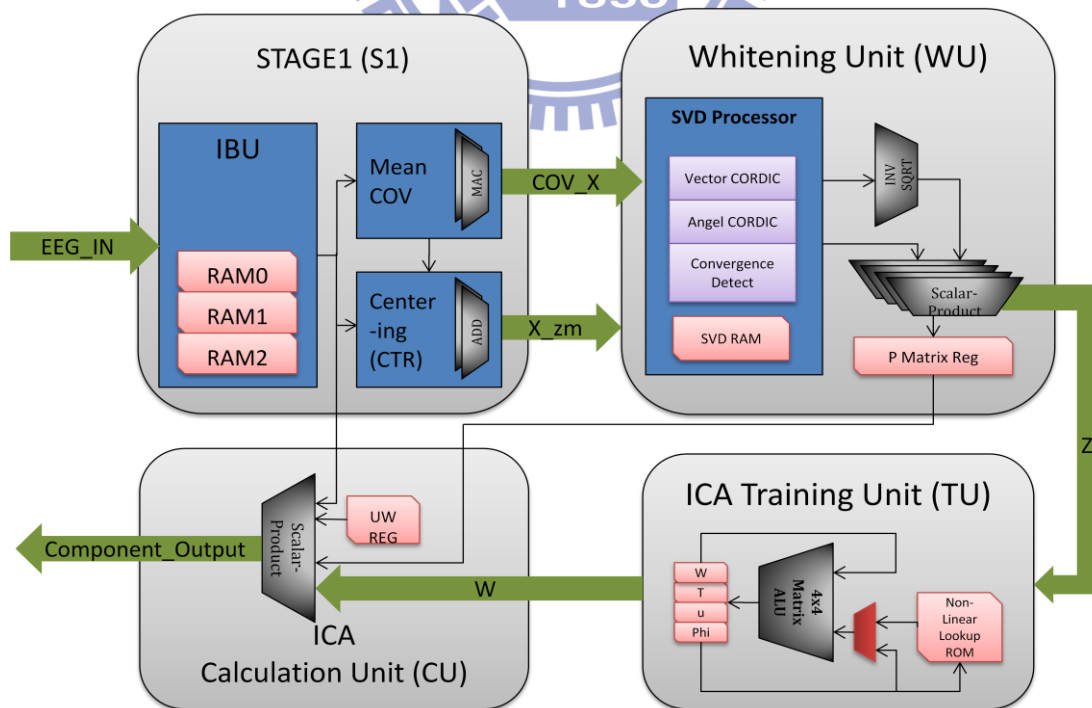
Figure 2.5   Overall architecture of the designed four-channel ICA processor

The algorithm can be divided into three stages:

1. Pre-Processing Stage (performed by S1 and WU)

    In the pre-processing stage, the raw EEG data is pre-processed by data centering and whitening transformation. The whitening transformation is a decorrelation method that converts the covariance matrix COV_X of a set of samples into the identity matrix I. This effectively creates new random variables that are uncorrelated and have the same variances as the original random variables. The method is called the whitening transform because it transforms the input matrix closer towards white noise. After the decorrelation, training iteration need to achieve convergence can be largely decreased. The pre-processing is done in S1 and WU.

2. ICA Unmixing Weight Training Stage (performed by TU)

    The pre-processed data is then used by the TU to find the best unmixing matrix that achieves maximum independence between each component. The TU is designed using the Infomax ICA algorithm described in the section 2.1.3.

3. Component Computation Stage (performed by CU)

    The components are extracted in this stage. The raw EEG data, the P matrix from whitening unit (WU) and the W matrix from ICA training unit (TU) are all required to perform the computation of the resulting components in ICA computation unit (CU). In addition, a handshaking mechanism is implemented to make the output interface flexible.

The calculation and the corresponding module designs are described in the following sections in detailed, and the decision of window size and the design of the training parameters will be explained and analyzed.

2.2.2 Stage 1 Unit and the Data Windowing Technique

The stage1 unit (S1) for buffering and data pre-processing shown in Figure 2.6 comprises an input buffering unit (IBU), a mean and covariance calculation unit (MeanCov), and a data centering unit (CTR).
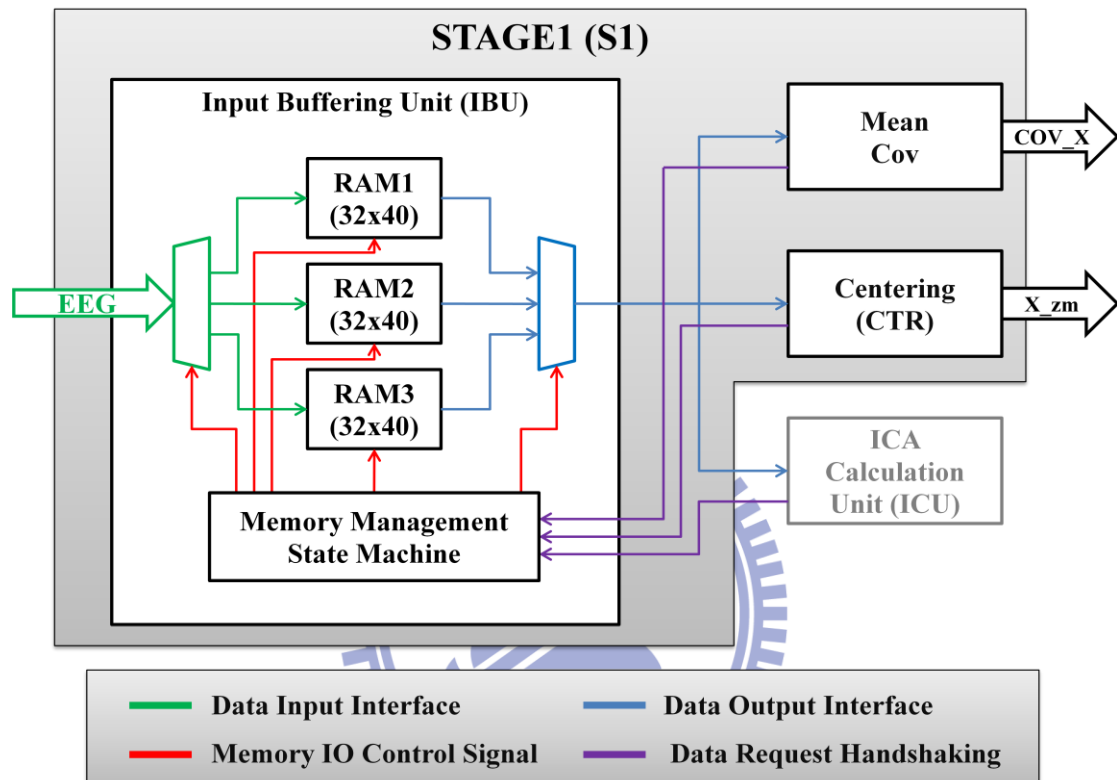
Figure 2.6   The architecture of STAGE1 unit

The input buffering unit employs three interleaved SRAM modules to store and manage the raw EEG sampled data. The three memory modules inside IBU are identical with size equal to 32 words, and each word is 10 bits long. In Figure 2.6, the internal connections of different function are distinguished by their color. The green connections are the data input interface, and the blue connections are the data output interface, while the red ones are the control signal used to perform the pipeline scheduling scheme. The IBU functions as the data controller of the ICA processor that supports the sliding window scheme and pipeline scheduling. The timing plot of the sliding window scheme and pipeline scheduling scheme is shown in Figure 2.7.

28
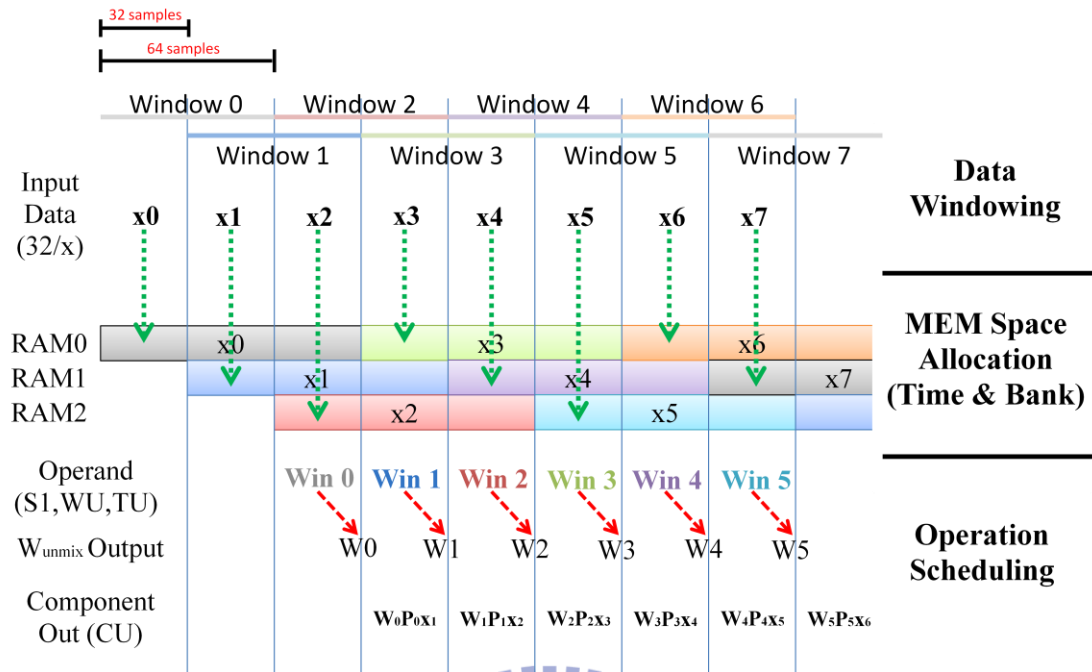
Figure 2.7   The precise timing plot of the sliding window scheme and pipeline scheduling scheme

Figure 2.7 depicts the data windowing technique, memory bank allocation and the operand scheduling used in our design. Data dependency constraints of the ICA algorithm tremendously limit the degree of parallelization in off-line ICA algorithm, so the overlapped sliding window scheme in Figure 2.7 is adopted. Due to the requirements for area and power optimization, memory size reduction is employed to minimize the chip area and power consumption to acceptable levels at minimal performance loss. As shown in Table 2.1, the option with window size of 512 with half overlapping sliding window achieves 0.9208 in correlation coefficient using super-gaussian random pattern sets. However, to implement the corresponding architecture using the same memory management scheme, it takes 30.72 Kilo-bits of memory size. It is not acceptable for portable devices that aims on low-power and low-cost design. Therefore, a window size of 64 with 50% overlap and 0.8401 correlation is chosen. Only 3.84 Kilo-bits memory size is required for the chosen window size, resulting 85% in savings compared with the one with window size of

512, and the corresponding trade-off is correlation degradation from 0.92 to 0.84.

Table 2.1 Memory Complexity Reduction

Window Size	Sliding Window Overlap Sample	Shift Step	Average Correlation	Training Iteration of the first 4 win.	Memory Size(Kb)
512	256 (50%)	256	0.9208	377, 1, 4, 80	30.72
128	112 (87.5%)	16	0.8307	490, 101, 7, 1	5.76
	96 (75%)	32	0.8336	490, 80, 1, 1	6.4
	64 (50%)	64	0.8334	490, 50, 12, 1	7.68
64	32 (50%)	32	0.8401	512, 208, 170, 1	3.84

Therefore, we can find that 32 samples (half-window) are marked a x_i, and the window size is defined to be 64 samples as we chose in Figure 2.7. The operand for all operations including data centering, whitening transform and the ICA training is based on a full window of 64 samples, but for the calculation of output components which is done by the ICA calculation unit (CU), the output size equals to the non-overlapped part that is a half-window of 32 samples. IBU starts the data pre-processing and ICA training after *window 0* filled with sampled data, while the new samples in x_2 are stored into RAM 2 causing no conflicts. Before *x2* is filled up, the calculation for the first unmixing matrix *W0* can be finished using a minimum 0.817 MHz input clock at the worst case condition of 512 iterations limit designed in ICA training unit (TU). In the worst case, it takes 203757 clock cycles to train for an unmixing matrix *W*, and this implies that the data sampled by a maximum sample rate of 9.708K samples per second can be processed in time when 60 MHz clock frequency is applied. After RAM 2 is filled up, *x1* is still left inside RAM 1 before the next half-window is fully allocated. Therefore, the components can be calculated and outputted during the calculation for *W1*.

For off-line ICA algorithms, the resulting component of *x1* is supposed to be extracted using the *W1*, but if we keep *x1* stored inside RAM 1 until *W1* is derived, an

additional memory module will be need, and it also cause a delay of 0.75 seconds from *x1* is sampled to the corresponding component output. To solve the two problems, we first assume that after the convergence, *W* won't change rapidly afterwards. Under this assumption, we can apply the previous *W* for extraction of the current data. Therefore the first component output is calculated as ***W0P0x1***.
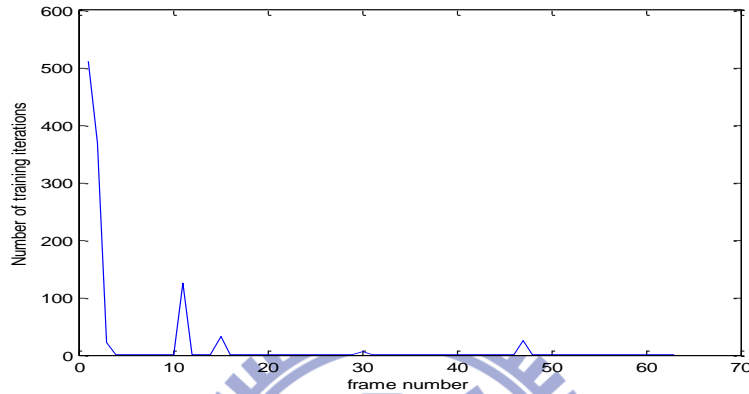
Figure 2.8   Iteration number the training unit takes to achieve conversion in each sliding window

As we can see from the simulation result using the super-gaussian random pattern we used for deciding the window size in Figure 2.8, aside from the first window that requires 512 iterations for exiting the training loop, iteration numbers are usually 1 after the convergence. The correlation difference using this scheme is so small that we can just ignore it. The scheme is also verified using real EEG signal recorded using NeuroScan system shown in section 2.3.2. Using the proposed scheme, memory access conflicts are avoided at window overlaps, while the component can be outputted ahead of time by a half-window that is 0.25 seconds without additional memory module needed.

Aside from data management, STAGE1 also performs the simple calculations needed prior to whitening and ICA training and computation.   The Mean_Cov unit employs shared multiply-accumulate units for calculating the mean and covariance of the EEG samples, while the centering unit utilizes four subtractors to remove the DC

component of each EEG data channel.

The definition of the covariance of two vectors is shown in Equation 2.27.

$$\text{Cov}(X,Y) = \text{E}[(X - \text{E}[X])(Y - \text{E}[Y])] = \text{E}[XY] - \text{E}[X]\text{E}[Y] \qquad (2.27)$$

For the covariance between four vectors, a complete covariance matrix needs to be computed as in Equation 2.28.

$$\text{Cov}\left(\begin{bmatrix} V1(1:64) \\ V2(1:64) \\ V3(1:64) \\ V4(1:64) \end{bmatrix}\right) = \begin{bmatrix} \text{Cov}(V1,V1) & \text{Cov}(V1,V2) & \text{Cov}(V1,V3) & \text{Cov}(V1,V4) \\ \text{Cov}(V2,V1) & \text{Cov}(V2,V2) & \text{Cov}(V2,V3) & \text{Cov}(V2,V4) \\ \text{Cov}(V3,V1) & \text{Cov}(V3,V2) & \text{Cov}(V3,V3) & \text{Cov}(V3,V4) \\ \text{Cov}(V4,V1) & \text{Cov}(V4,V2) & \text{Cov}(V4,V3) & \text{Cov}(V4,V4) \end{bmatrix} \quad (2.28)$$

$$= \begin{bmatrix} \text{Var}(V1,V1) & \text{Cov}(V1,V2) & \text{Cov}(V1,V3) & \text{Cov}(V1,V4) \\ \text{Cov}(V2,V1) & \text{Var}(V2,V2) & \text{Cov}(V2,V3) & \text{Cov}(V2,V4) \\ \text{Cov}(V3,V1) & \text{Cov}(V3,V2) & \text{Var}(V3,V3) & \text{Cov}(V3,V4) \\ \text{Cov}(V4,V1) & \text{Cov}(V4,V2) & \text{Cov}(V4,V3) & \text{Var}(V4,V4) \end{bmatrix}$$

In addition, because $\text{Cov}(X,Y)$ equals to $\text{Cov}(Y,X)$, only the elements in the upper triangle are required to be calculated. Therefore, we calculate the ten elements sequentially using only one shared multiply-accumulate unit (MAC) as shown in Figure 2.9. The colorful small blocks are registers used for accumulation of the product of each combination between channels.
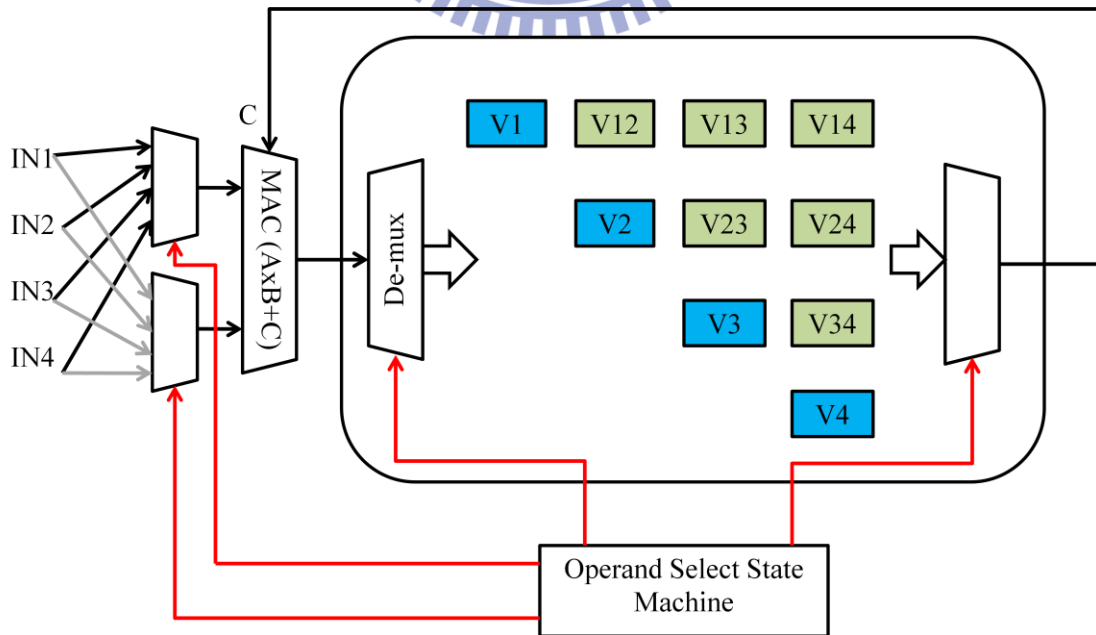
Figure 2.9   Covariance matrix calculation using only one shared MAC operator

After the sums of each combination are accumulated, average is easily acquired by fixed-point shifting by 6 bits that is 64 times smaller. Finally, the ten elements in the covariance matrix are calculated using Equation 2.27.

Since the mean values of data from each channel are calculated in MeanCov unit, the mean values stored in V1, V2, V3 and V4 register can be used in data centering unit (CTR). Therefore, after the accumulation, V1 to V4 is provided to CTR using inter-module wire connection with a valid signal *MEAN_VALID* to inform the CTR that mean values are available.

The data centering operation is expressed as Equation 2.29:

$$X\_zm(i) = X(i) - E\{X(m)\} = X(i) - \frac{1}{N}\sum_{j=1}^{N} X(j) \qquad (2.29)$$

Where N is the decided window size that is 64, and i is the values from 1 to 64 representing each elements in a window. A direct and parallel CTR unit is designed using four subtractors shown in Figure 2.10.
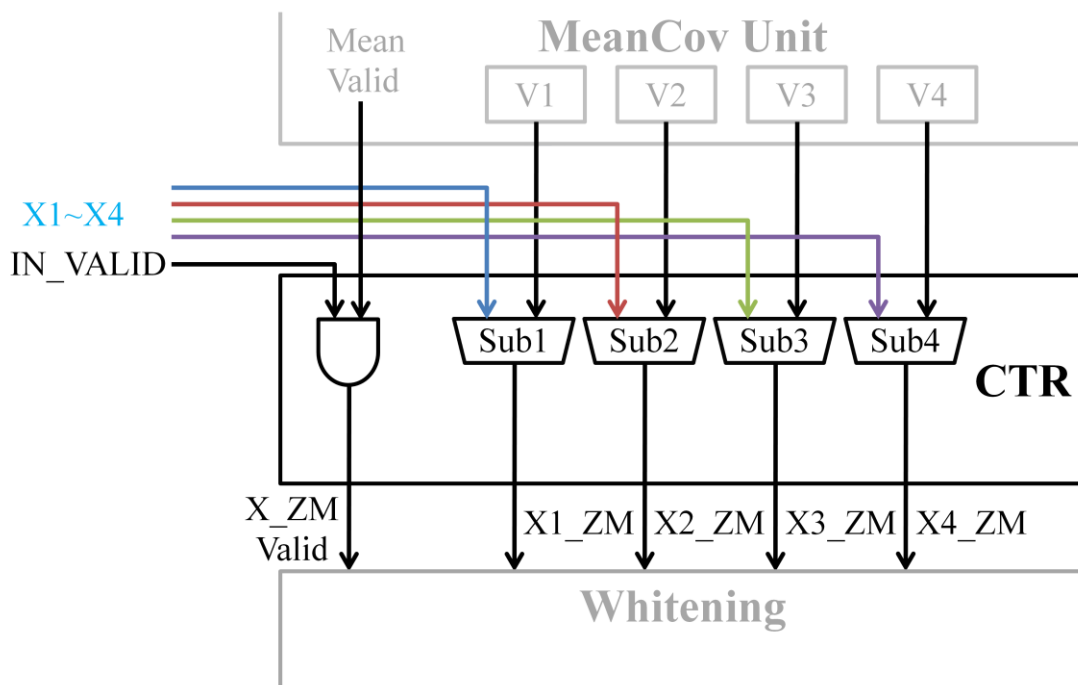
Figure 2.10   The architecture of the centering unit (CTR)

2.2.3 Whitening Unit

The whitening transform is linear operation and always possible. After whitened, the components of new vector z are uncorrelated and their variances equal to unity. In other words, the covariance matrix of z equals to the identity matrix.

$$E\{zz^T\} = I \tag{2.30}$$

One popular method for whitening is to use the eigen-value decomposition (EVD) of the covariance matrix. D is the diagonal matrix of its eigenvalues and E is the orthogonal matrix eigenvectors of covariance matrix of x. The decomposition can be expressed as:

$$E\{xx^T\} = EDE^T \tag{2.31}$$

Where D is

$$\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \tag{2.32}$$

Our goal is to find the whitening matrix $P = E\{xx^T\}^{\frac{-1}{2}} = ED^{\frac{-1}{2}}E^T$, and apply it to the centered data. Therefore, a Jacobi singular value decomposition (JSVD) engine which can be used to perform the EVD is employed in the whitening unit shown in Figure 2.11.
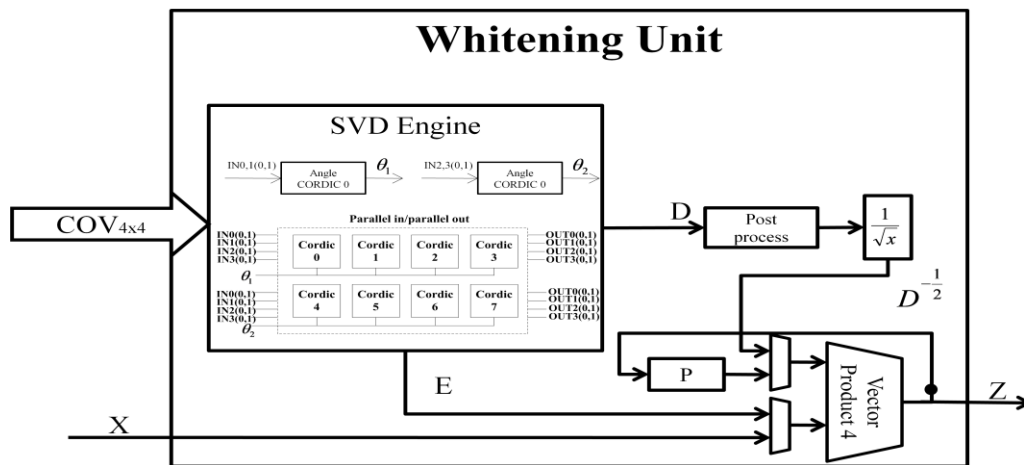
Figure 2.11　　The architecture of the whitening unit (WU)

The SVD engine is designed using CORDICs [28]. There are two parallel angle CORDIC and eight parallel vector CORDIC in the SVD engine. Vectors selected by the state machine are sent to the parallel CORDIC engines. The parallel order of operation is shown as following: $(p,q)=(i,j)=\{(1,2),(3,4);(1,3),(2,4);(1,4);(2,3)\}$. The detailed pipelining scheme is shown in Figure 2.12.

Unit	Clock Cycle												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Angle CORDIC 0	S	F	S	F	S	F	S	F	S	F	S	F	
Angle CORDIC 1	S	F	S	F	S	F	S	F	S	F	S	F	
Vector CORDIC 0		S	F	S	F	S	F	S	F	S	F	S	F
Vector CORDIC 1		S	F	S	F	S	F	S	F	S	F	S	F
Vector CORDIC 2		S	F	S	F	S	F	S	F	S	F	S	F
Vector CORDIC 3		S	F	S	F	S	F	S	F	S	F	S	F
Vector CORDIC 4		S	F	S	F	S	F	S	F	S	F	S	F
Vector CORDIC 5		S	F	S	F	S	F	S	F	S	F	S	F
Vector CORDIC 6		S	F	S	F	S	F	S	F	S	F	S	F
Vector CORDIC 7		S	F	S	F	S	F	S	F	S	F	S	F

Figure 2.12   Operation pipelining for singular value decomposition

In the Figure, "S" represent for the start of calculation while the "F" means the finish of the calculation. Operations marked with different color are using different part of the covariance matrix shown in Figure 2.13.

$$1:\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ C_{41} & C_{42} & C_{43} & C_{44} \end{bmatrix} \quad 2:\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ C_{41} & C_{42} & C_{43} & C_{44} \end{bmatrix} \quad 3:\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ C_{41} & C_{42} & C_{43} & C_{44} \end{bmatrix}$$

$$4:\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ C_{41} & C_{42} & C_{43} & C_{44} \end{bmatrix} \quad 5:\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ C_{41} & C_{42} & C_{43} & C_{44} \end{bmatrix} \quad 6:\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ C_{41} & C_{42} & C_{43} & C_{44} \end{bmatrix}$$

Figure 2.13   Operands for pipelined CORDIC operations

After few iterations for convergence the E and D matrices can be derived, and they are multiplied together to generate the whitening matrix P by a single vector product unit. Afterward, the centered data are also whitened using the same vector product unit.

2.2.4 Infomax ICA Training Unit

The ICA training unit is the most important part used to calculate the unmixing matrix *W*, and most of computational time is consumed here due to the training iteration loop. As the behavior of this unit depends on the input data, we design the training parameters by using the MATLAB simulation result of super-gaussian random pattern as we discussed in section 2.2.2. To recap the Infomax ICA algorithm, the calculation step is listed below.

For initialization, W is first set to 4-by-4 identity matrix as shown below.

$$W_{4\times4}(0) = I_{4\times4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Afterward, when one window of pre-processed data is available, the following calculation will be applied to renew the unmixing matrix *W*. Note that Z is the 4-channel pre-processed data window, and the variable *i* is the iteration number.

1. $i=1$

2. $u_{4\times64}(i) = W_{4\times4}(i)Z_{4\times64}$

3. $y_{4\times64}(i) = g(u) = \frac{1}{1+e^{-u(i)}}$

4. $\Delta W_{4\times4}(i) = R_{\text{learning}}(I_{4\times4} + (1-2y)_{4\times64}u^{\text{T}}{}_{64\times4})W_{4\times4}$

5. $W(i+1) = W(i) + \Delta W_{4\times4}(i)$

6. If$((\text{sum}(\text{abs}(\Delta W_{4\times4}(i))) < \text{threshold}_{\text{convergence}}) \,||\, (i{+}{+}{==}\text{lim}_{\text{ite\_num}}))$
   Exit the loop
   Else
   Go back to step 2

In the equations in calculation steps, there are three constant parameters to be decided. The learning rate R_{learning} controls the speed of convergence. When the value of R_{learning} is too small, the learning time will be long. Furthermore, small R_{learning} also causes the unmixing weight locked at the local minimum in some cases. On the other hand, the threshold$_{\text{convergence}}$ affects the property of convergence. When

threshold$_{convergence}$ is too large, the calculated *W* is not good enough to be used as unmixing matrix, but when threshold is too small, the training may never achieve the convergence.

With intensive simulation in MATLAB, we find that in the working range of threshold$_{convergence}$ and R$_{learning}$, the values of them does not have much influence on the correlation coefficient between the original source and the extracted components. Therefore, the middle values are chosen for threshold$_{convergence}$ and R$_{learning}$. For the limitation of the maximum iteration number lim$_{ite\_num}$, we choose a relatively large value compared with other work in the literature due to the smaller window size in our design. We can see the iterations as the extension of the data length. The values of the parameters are summarized in Table 2.2.

Table 2.2   Parameters for the Infomax ICA training algorithm

Parameter	Symbol	Value
Learning Rate	R$_{learning}$	7.4768x10^{-4}
Convergence Threshold	threshold$_{convergence}$	1.0012x10^{-8}
Limitation of the max. Iteration number	lim$_{ite\_num}$	512

From step 2 to step 4, large matrix additions and multiplications are used. For hardware implementation, the operation for large matrix is decomposed as sum of contributions. If we define the item $(1 - 2y)_{4 \times 64} u^{\mathrm{T}}{}_{64 \times 4}$ in step 4 to be $T_{4 \times 4}$. We can decompose the calculations to be:

$$u_{4 \times 64} = [u_{4 \times 1}^1; u_{4 \times 1}^2; u_{4 \times 1}^3; \dots; u_{4 \times 1}^{64}] \tag{2.33}$$

$$y_{4 \times 64} = [y_{4 \times 1}^1; y_{4 \times 1}^2; y_{4 \times 1}^3; \dots; y_{4 \times 1}^{64}] \tag{2.34}$$

In Equation 2.34, each y^j is derived using Equation 2.35:

$$y_{4 \times 1}^{j} = g(u^{j}) = \frac{1}{1 - e^{-u^{j}}} \tag{2.35}$$

The key step of decomposition is shown in Equation 2.36:

$$T_{4\times4} = (1 - 2y)_{4\times64} u^{\mathrm{T}}{}_{64\times4} = \sum_{j=1}^{64} T_{4\times4}^{j} \quad (2.36)$$

Each T^j in Equation 2.36 equals to:

$$T_{4\times4}^{j} = \left(1 - 2y_{4\times1}^{j}\right) \times u_{1\times4}^{j\,\mathrm{T}} \quad (2.37)$$

The modified Infomax ICA algorithm designed for hardware implementation is written in pseudo code using MATLAB expression and shown below. This snippet is the ICA training algorithm for a window of data assuming W already initialized or keeping the old value calculated from the previous sliding window. The resulting calculation steps for $T_{4\times4}$ from Equation 2.36 are in the internal for-loop from line 3 to line 7.

Modified Infomax ICA Training Algorithm for Hardware Implementation	
1. For i = 1:512	**% lim$_{\text{ite\_num}}$ equals to 512**
2. T = zero(4,4);	
3. For j = 1:64	**% Calculate for T**
4. u = W*Z(:,i);	
5. y = 1/(1-e^(-u));	**% Non-linear function g**
6. T = T+(1-2y)*u';	
7. End	**% T ready**
8. T = ((R*I)+R*T)*W;	**% Store ΔW in T**
9. W=W+T;	**% Update W**
10. If(T<= threshold$_{\text{convergence}}$)	**% Convergence check**
11. Break;	**% Reach covergence**
12. End	

There are eight states in the main state machine of ICA training unit (TU), and the state transfer chart is shown in Figure 2.14.
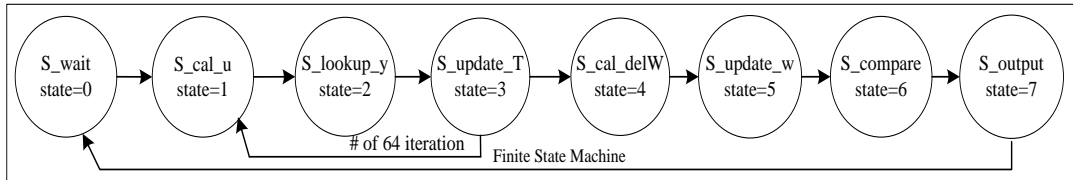
Figure 2.14   State transfer char of the main state machine in ICA training unit (TU)

The resulting architecture for ICA training unit is shown in Figure 2.15. The matrix calculation requires many adders and multipliers, so one shared multiplier

array composed of 16 16-bit scalar multipliers and one shared adder array composed of 16 32-bit scalar adders. The number of adders and multipliers are analyzed and optimized, so the shared operator arrays are utilized efficiently in the training process. Note that a mirrored nonlinear lookup unit is designed to minimize the ROM size for the lookups of non-linear function g(u). Control logic circuits can be classified into two groups: **data updating logic** and **data operation routing logic**. The internal 4-bit counter is used to implement a sub-state control.
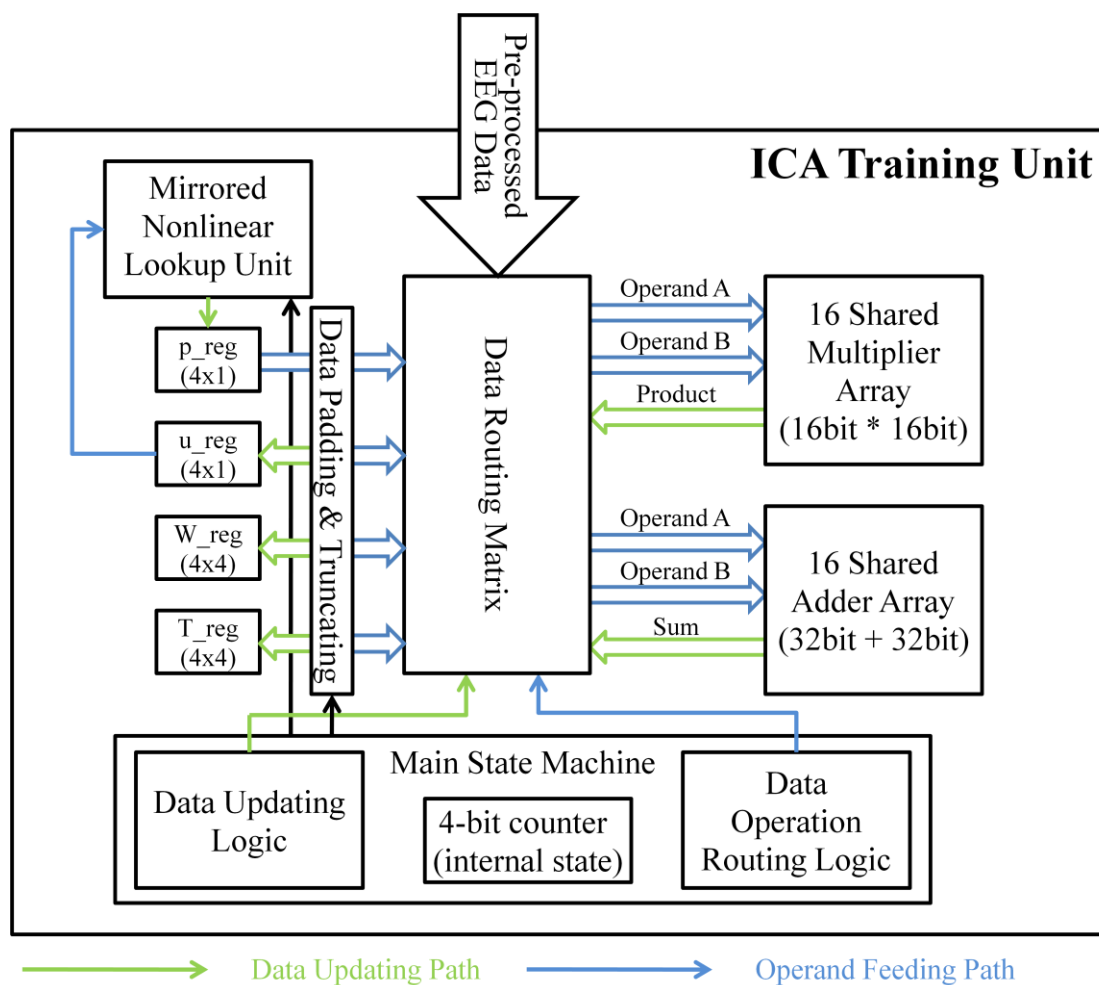
Figure 2.15   Hardware architecture of the ICA training unit

Table 2.3 clearly shows the operation in each state, and the value updates of the variable registers are also indicated.

Table 2.3   State, operation and data control in ICA training unit

State	Adder Operation	Multiplier Operation	Variable Update	Needed #Cycle

			[OpA,OpB]		
0	S_wait	-	-	$T=I_{4x4}$ $u=Z_{4x1}$	1
1	S_cal_u	[mul$_i$_out,mul$_i$_out] [add$_i$_out,add$_j$_out]	[W,u]	u_1=add3_out u_2=add6_out u_3=add9_out u_4=add12_out	1
2	S_lookup_y	-	-	p_1=Looup(u_1) p_2=Looup(u_2) p_3=Looup(u_3) p_4=Looup(u_4)	4
3	S_update_T	[T,mul_out]	[p,u]	$u= Z_{4x1}$ T=add_out	1
4	S_cal_delW	[mul$_i$_out,mul$_j$_out] [add$_i$_out,add$_j$_out]	For counter=0 [R$_{learning}$,T] For counter=1~4 [T(1,:),W] [T(2,:),W] [T(3,:),W] [T(4,:),W]	T=mul_out T(1,:)=add3_out T(2,:)=add6_out T(3,:)=add9_out T(4,:)=add12_out	5
5	S_update_W	[W,T]	-	W=add_out	
6	S_compare	[mul$_i$_out,mul$_j$_out] [add$_i$_out,add$_j$_out]	[T,T]	T= I$_{4x4}$	
7	S_output	-	-	-	16

The ROM size of the mirrored non-linear lookup unit in Figure 2.15 is optimized

using the anti-symmetric property [29]. The non-linear function used in Infomax ICA

algorithm is a sigmoid function $g(u) = \frac{1}{1+e^{-u}}$. The function is plot in Figure 2.16 (a).

The anti-symmetric property of this function allows us to store only half of the values.

As a result, the area of the ROM is also reduced to about half of the original area.
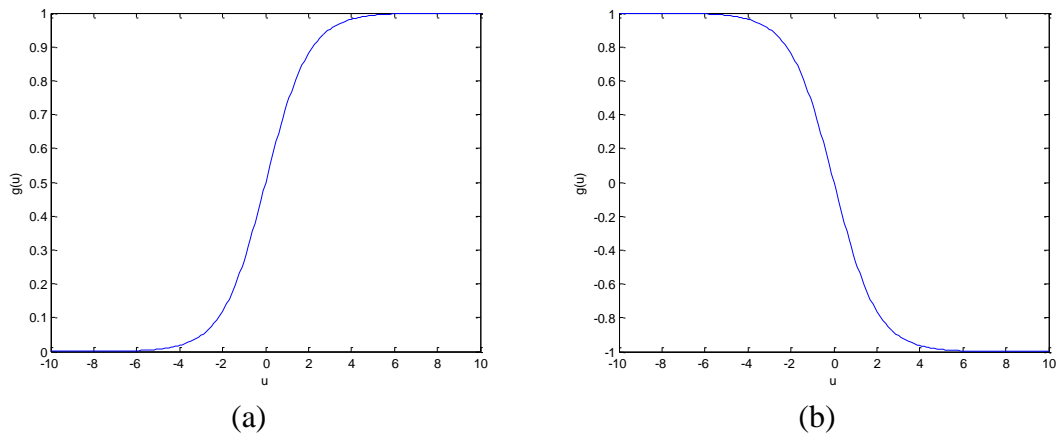
Figure 2.16   (a) The original non-linear function g(u) used in Infomax ICA training
algorithm (b) 1-2*g(u)

In Equation 2.34, the double of the lookup value is subtracted from 1, and the function

f(u)=1-2*g(u) is plot in Figure 2.16 (b). The plot also present us the same property of

anti-symmetric as g(u) has. Therefore, to save the additional subtraction, the value of f(u) is stored in the ROM, instead of g(u).

The size of the ROM is decided from the MATLAB simulation result summarized in Table 2.4. The lookup table ROM size can be reduced by 87.5% from 512 to 64 entries, with almost no loss in performance. As we can see from Figure 2.16 (b), the output is almost saturated when the input is a value larger than 7 or smaller than -7. Therefore a ROM that stores 32 entries is generated, and the lookup range is between +7 and -7. When the input value is out of the range, the mirrored non-linear lookup unit will output the value of saturation that is +1 or -1. The resulting design of the mirrored non-linear lookup unit is shown in Figure 2.17.

Table 2.4   The simulation result of different ROM sizes

ROM Size	L	Step Size	Average Correlation	#Training Iteration
Float	16	-	0.8612	502
512	16	1/16	0.8581	486
256	16	1/8	0.8588	503
128	16	1/4	0.8675	531
64	16	1/2	0.8704	-
128	8	1/8	0.8588	503
64	8	1/4	0.8655	535
32	8	1/2	0.8704	584
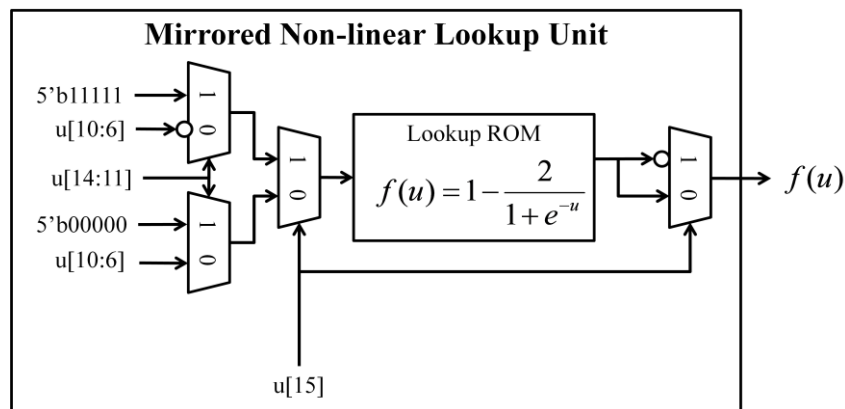
Figure 2.17   The architecture of the mirrored non-linear lookup unit in TU

2.2.5 ICA Computation Unit

The ICA computation unit architecture employs a shared scalar product to

calculate the whitened unmixing matrix *UW* and independent component analysis

output *ICA_OUT*. Equation 2.37 and 2.38 are the calculation step for the final

component outputs.

$$W_{unmixing} = W \times p \tag{2.37}$$

The *W* in Equation 2.37 is the output from the ICA training unit (TU) while the *p* is

whitening matrix previously derived in the whitening unit (WU).

$$ICA\_OUT = W_{unmixing} \times x \tag{2.38}$$

Independent component estimtes are finally calculated by multiplying $W_{unmixing}$ with *x*.

The *x* is consistant with the expression in section 2.2.2 defined to be a half-window

data. Therefore, 32 non-overlapping extracted component samples are outputed each

time. The corresponding calculation flow chart is shown in Figure 2.18. Note that a

handshaking mechanism at the output is added to provide a flexible output interface.
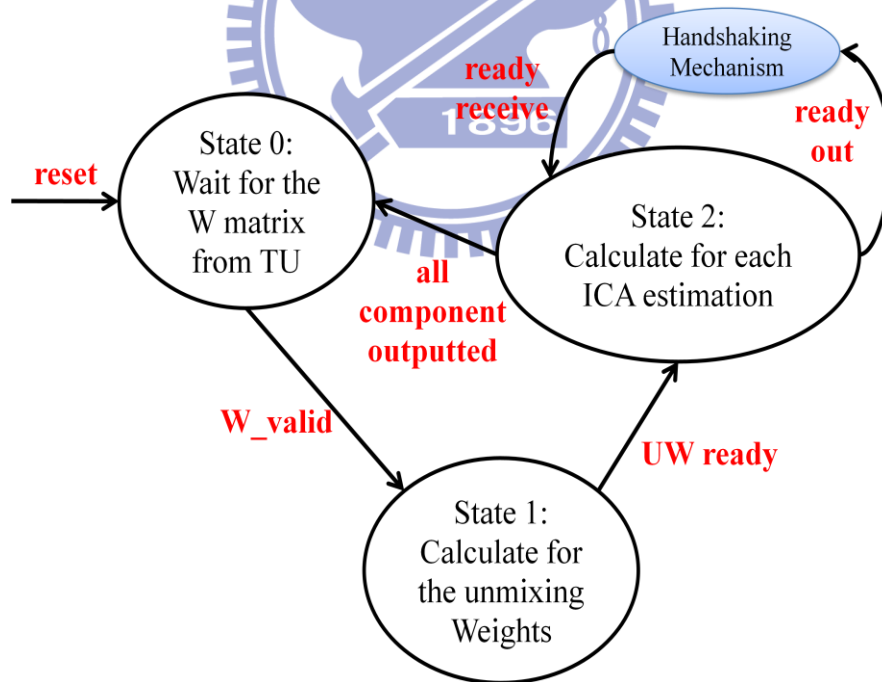
Figure 2.18   Calculation flow char for the final ICA result

2.3 Performance Analysis of the 4-Channel ICA Processor

To analyze the performance of the designed 4-channel ICA processor, a

Super-Gaussian random pattern generator is first created. Before application using real EEG patterns, we have to know if the designed ICA processor can really separate the independent components of Super-Gaussian, but we cannot know the original components inside EEG signal not even the number of components mixed in the measured EEG signals. Therefore, using a set of known pattern can at least confirm the validity of the whole design. After demonstration of one set of know pattern, we use four sets of real EEG pattern recorded using NeuroScan system to analysis the performance compared with EEGLab [30]. Pattern 1 and 2 are relatively clean EEG signals with only few eye-blink artifacts in the signal, and pattern 3 and 4 are contaminated by eye-blink artifacts on an average of one every 2.5 seconds.

2.3.1 Performance Analysis Using Super-Gaussian Pattern

Figure 2.19 (a) shows the 4-channel super-gaussian random sources. In each channel, 1024 samples are generated. To verify the super-gaussianity of the sources, the probability density functions (PDF) are also shown in Figure 2.19 (b).

(a)   (b)
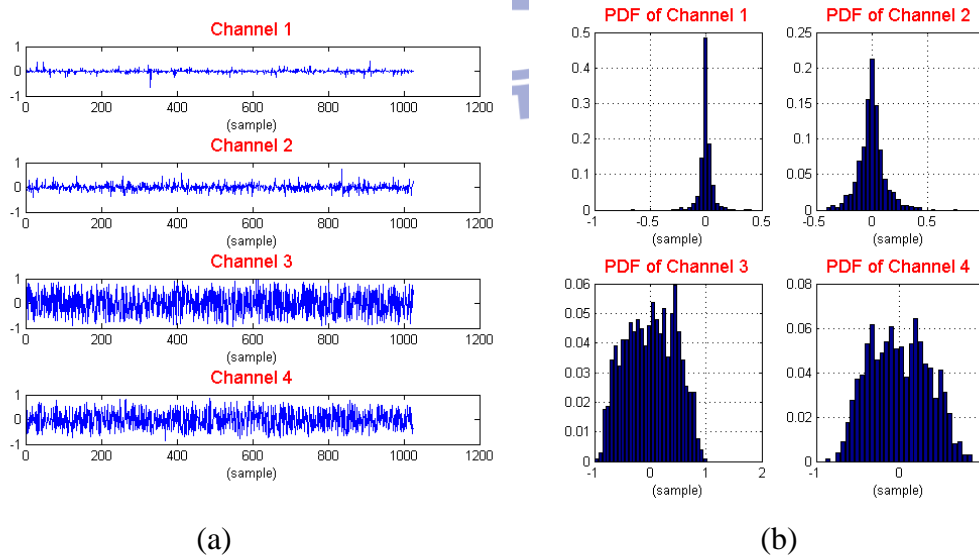
Figure 2.19   (a) 4-channel super-gaussian random sources (b) Probability density functions

The super-gaussian source in Figure 2.19 is mixed by a stationary mixing matrix to generate the input pattern for our design. Figure 2.20 shows the waveforms of the last two windows of the original source (a), mixed input pattern (b) and the extracted

component (c) using the designed processor. We can easily indicate the channel mapping between the original sources and the extracted independent components from Figure 2.20 (a) and (c). For analysis of the non-stationary characteristics of ICA output, the correlation coefficient of the ICA from two neighboring sliding windows is evaluated. The average correlation is 0.86 between the original source signals and extracted ICA components. The correlation variation of each 32 outputs is shown in Figure 2.21, and the red line is the correlation variation of the result using EEGLab.
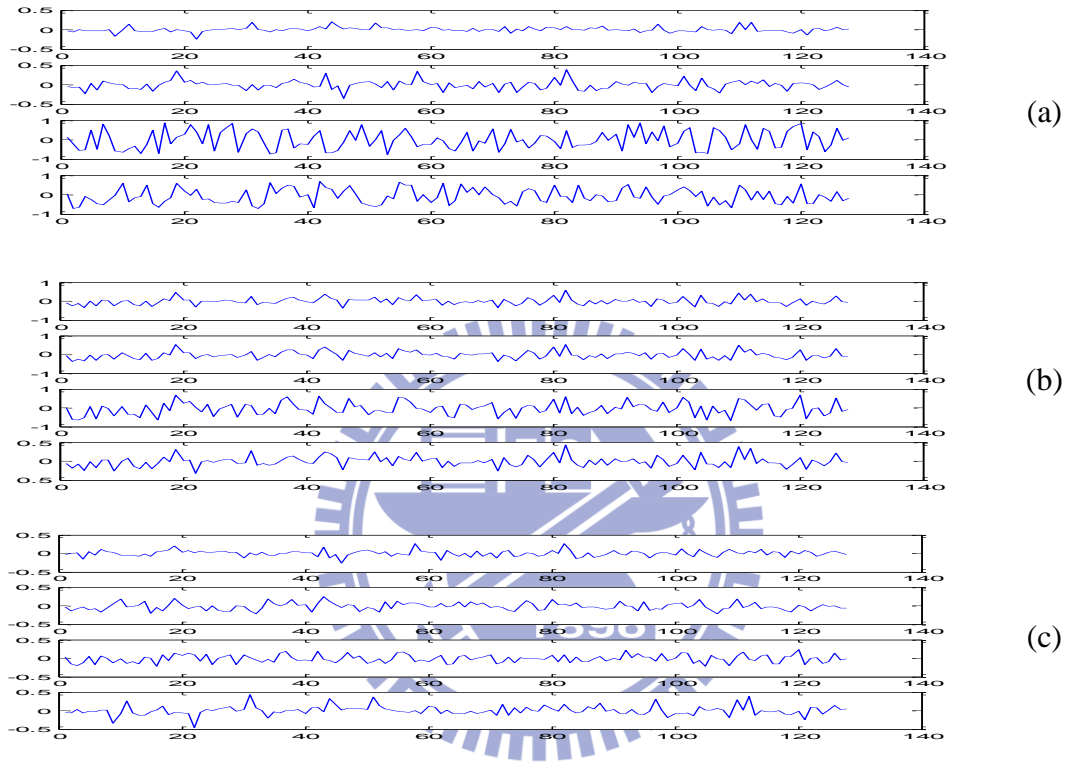
Figure 2.20   (a) Original source signals (b) Mixed signals (c) Extracted ICA signals
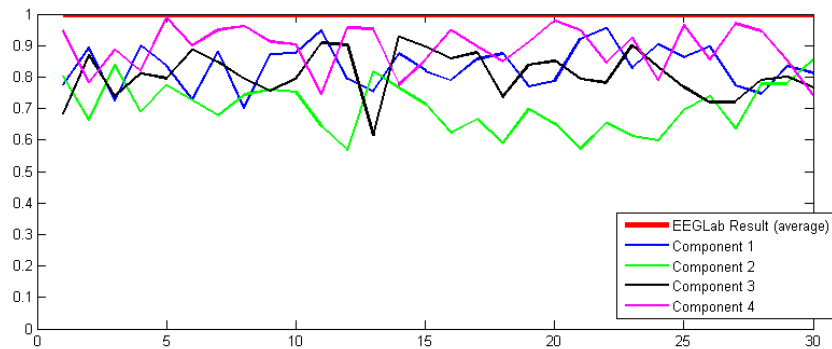
Figure 2.21   Correlation variation of each 32 outputs compared with EEGLab result

2.3.2 Performance Analysis Using Real EEG Patterns

The ability to separate the mixture of super-gaussian random signals does not suggest the ability to find out the artifacts and components in real EEG signals. In this

44

section, 4 real EEG patterns recorded from NeuroScan system are fed as input patterns to the fabricated chip using the setup described in section 2.4.1. The patterns had been recorded using sample rate of 128 Hz for total 24 seconds after stabilization of the experimental subject. A high-pass filter and a low-pass filter of cut-off frequencies equal to 0.15 Hz and 55 Hz are applied to the setup before recording. The detailed information is listed in Table 2.5, and the channel location is mapped using international 10-20 system of electrode placement standard for EEG measuring.

Table 2.5 Detailed pattern information for the 4 EEG patterns recorded

Pattern	HPF (Hz)	LPF (Hz)	Duration (Sec)	Ch1 Loc.	Ch2 Loc.	Ch3 Loc.	Ch4 Loc.
Pattern 1	0.15	55	24	FP1	FP2	FZ	C1
Description	Clean and stable EEG with normal eye-blink period						
Pattern 2	0.15	55	24	M1	CP1	CPZ	CP2
Description	Clean and stable EEG with normal eye-blink period						
Pattern 3	0.15	55	24	FP1	FPZ	FP2	VE0
Description	EEG with eye-blink artifacts on an average of one every 2.5 seconds						
Pattern 4	0.15	55	24	FP1	FPZ	FP2	VE0
Description	EEG with eye-blink artifacts on an average of one every 2.5 seconds						

2.3.2.1 Pattern 1 – Stable EEG without Artifact

In Figure 2.22, the waveforms in the first column are the recorded EEG patterns, and the waveforms in the second column are the extracted components using off-line ICA algorithm performed by the latest EEGLab release. To compare with the off-line ICA result using EEGLab, the components analyzed by the designed chip are shown in the third column. Although the channel order are identical for the results using EEGLab and the chip in pattern 1, the channel orders may be shuffled through the algorithm in the other patterns, so the corresponding components are marked by the same component number, and the correlations between the results from on-line hardware calculation and off-line algorithm are also shown in the figure. Although the data are windowed, the results have shown that each statistical independent

component is locked in the same channel for different window.

Figure 2.22   Time domain comparison between the off-line result from EEGLab and on-line results for pattern 1 with original source EEG shown in the first column
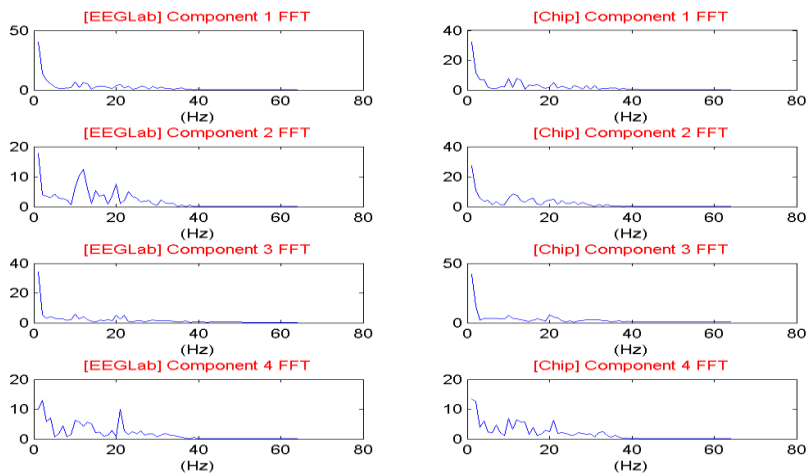
Figure 2.23   Frequency domain comparison between the off-line result from EEGLab and on-line results for pattern 1
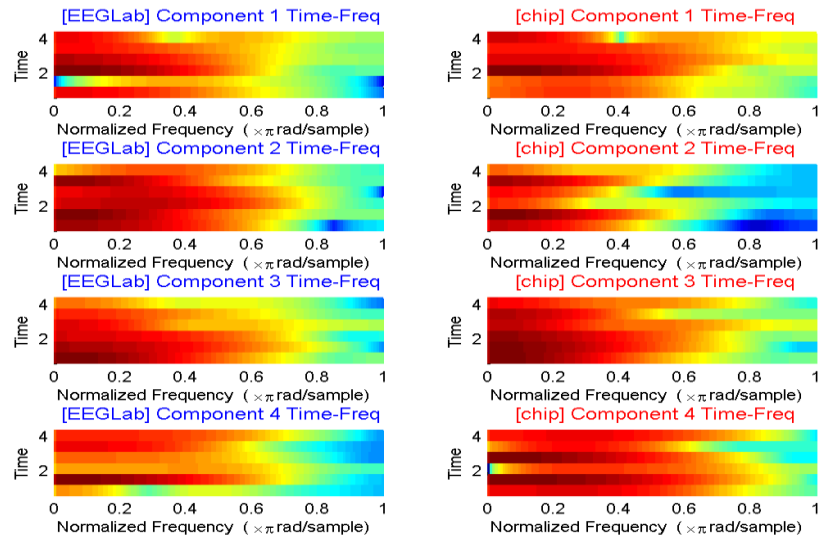
Figure 2.24   Time-frequency analysis comparison between the off-line result from EEGLab and on-line result for pattern 1

Frequency domain comparison and time-frequency comparison are shown in Figure 2.23 and 2.24. An average correlation coefficient of 0.867 is achieved in the first pattern. The time domain, frequency domain and time-frequency results all show consistency between the off-line results and on-line results.

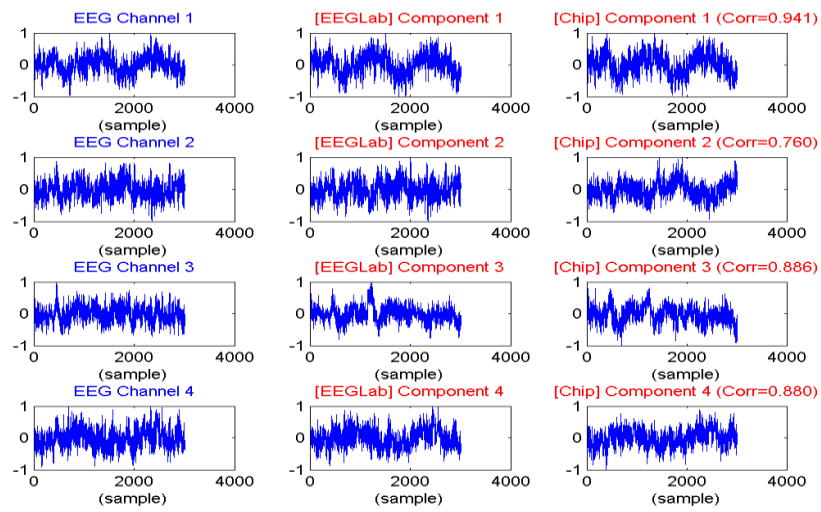## 2.3.2.2   Pattern 2 – Stable EEG without Artifact
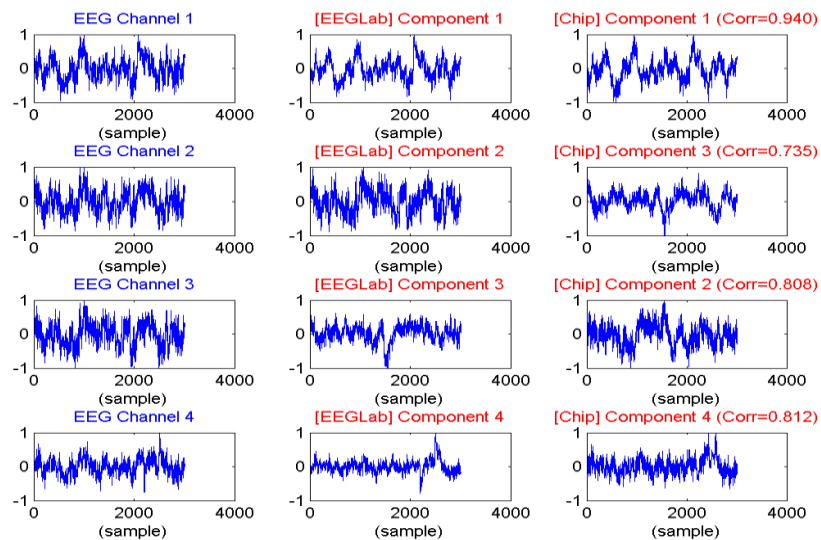
Figure 2.25   Time domain comparison between the off-line result from EEGLab and on-line results for pattern 2 with original source EEG shown in the first column

Figure 2.26　　Frequency domain comparison between the off-line result from EEGLab and on-line results for pattern 2

Figure 2.27　　Time-frequency analysis comparison between the off-line result from EEGLab and on-line result for pattern 2

Figure 2.25 to 2.27 shows the comparison between the off-line result from EEGLab and the on-line result. Similar description is skipped for this pattern. For the conclusion, an average correlation coefficient of 0.824 is achieved in the second pattern. The time domain, frequency domain and time-frequency results all show consistency between the off-line results and on-line results as the first pattern does.

2.3.2.3 Pattern 3 – EEG Contaminated by Eye-Blink Artifact

From Figure 2.28 and Figure 2.32 we can easily tell that the recorded EEG signals for pattern 3 and pattern 4 are heavily contaminated by eye-blink artifacts. EEG signals recorded from location at FP1, FPZ, FP2 and VE0 are chosen because they are influenced more heavily by the eye-blink artifacts due to the spatial locality to the eyes where the eye-blink artifacts originate.

Figure 2.28   Time domain comparison between the off-line result from EEGLab and on-line results for pattern 3 with original source EEG shown in the first column

Figure 2.28 has shown that except the component of eye-blink results in only 0.548 of correlation coefficient, the average correlation between the on-line result and the off-line result for the other three components still achieves 0.84 in correlation. In fact, correlation itself doesn't represent the quality of the algorithm, so an example of artifact removal using EEGLab and the result from the chip is shown in Figure 2.29. The original EEG waveforms are shown in the first column, and the artifact-removed results using EEGLab and the chip is shown in the second and third column. In the recorded EEG, an eye-blink artifact appears from sample 20 to sample 40 marked by

the red dotted lines. As we can see from the processed results, although the influence of the eye-blink event is still obvious, the ratio between the artifact component and the EEG components is definitely decreased. The on-line algorithm and off-line algorithm show similar result for artifact removal, and this also prove the capability of this design for the targeted application.
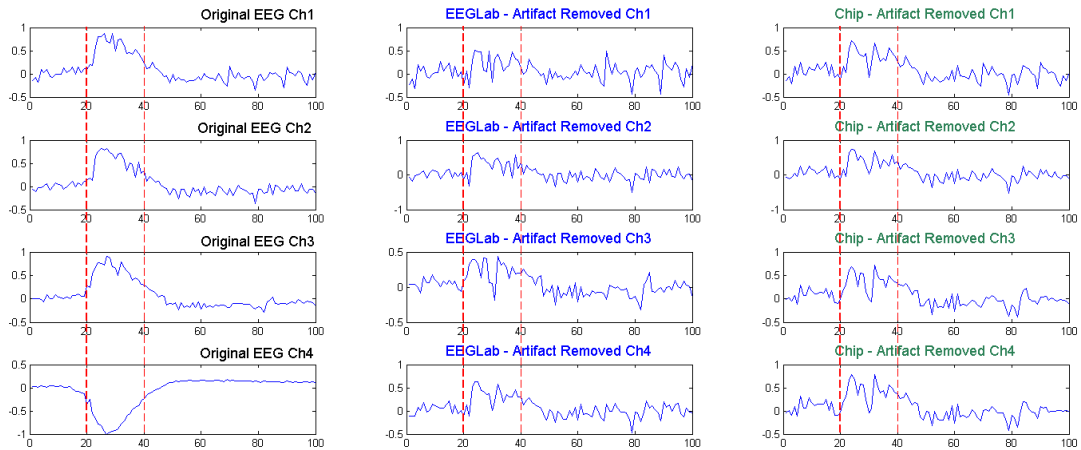
Figure 2.29   An artifact removal example used to compare the off-line performance using EEGLab and the on-line performance using the designed chip
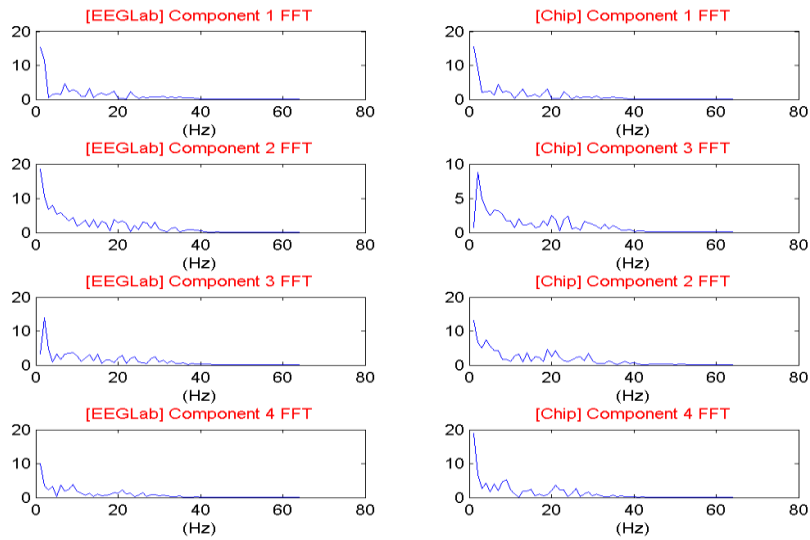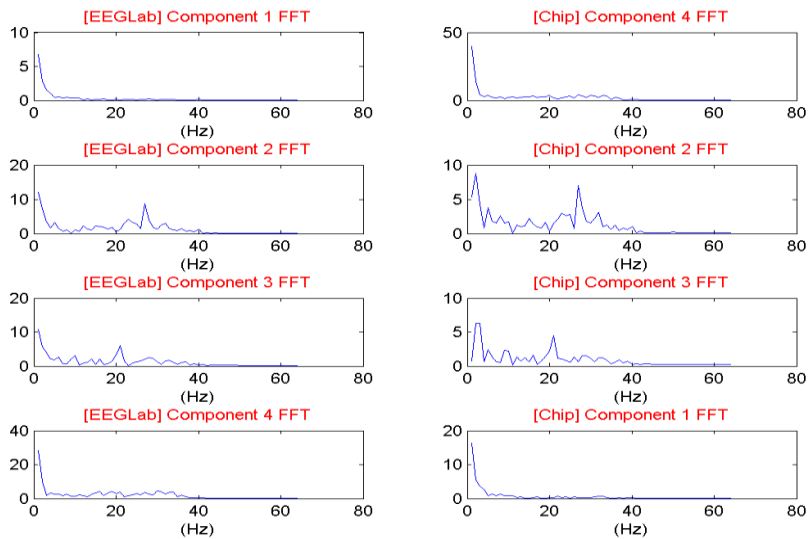

Figure 2.30   Frequency domain comparison between the off-line result from EEGLab and on-line results for pattern 3
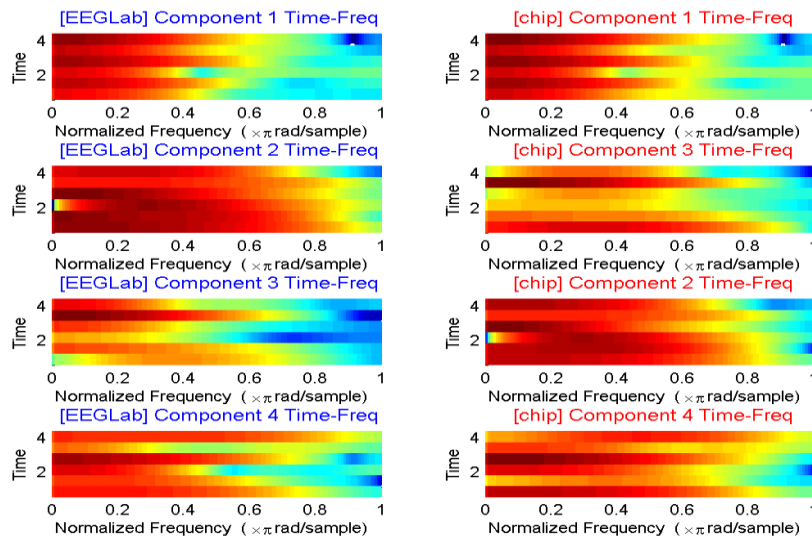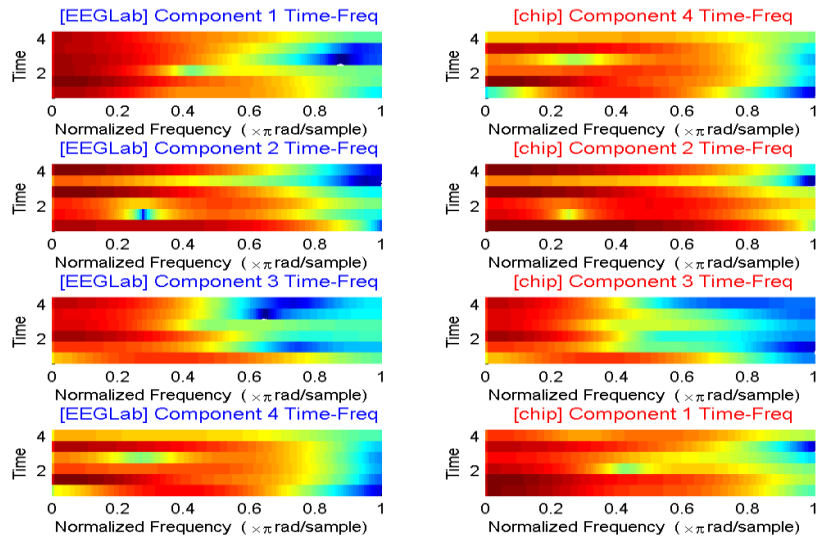
Figure 2.31   Time-frequency analysis comparison between the off-line result from EEGLab and on-line result for pattern 3

Frequency domain and time-frequency analysis are also done for the eye-blink patterns. Frequency domain comparisons shown in Figure 2.30, 2.31, 2.33 and 2.34 demonstrate more similar results between the characteristics of the on-line algorithm and off-line algorithm than it does in pattern 1 and pattern 2.

2.3.2.4   Pattern 4 – EEG Contaminated by Eye-Blink Artifact
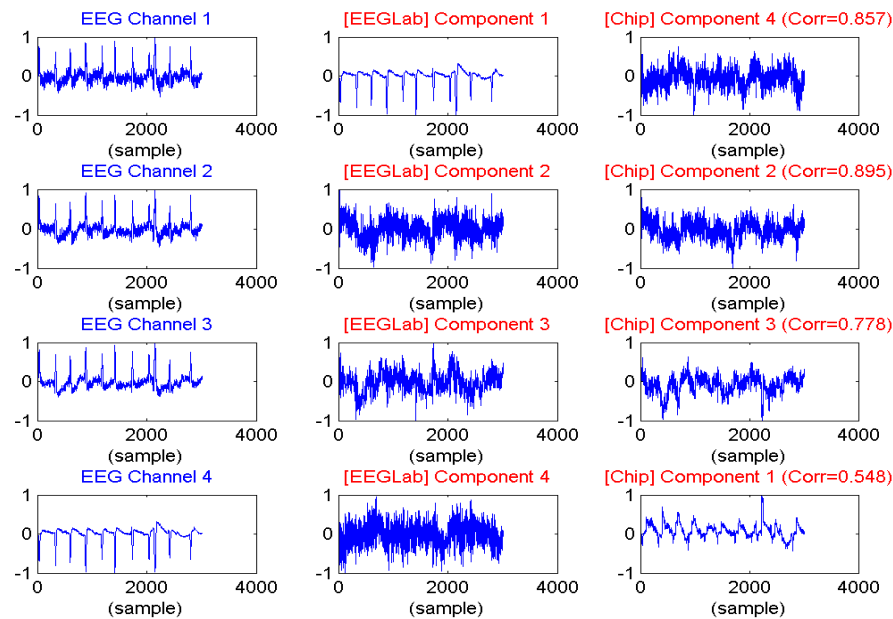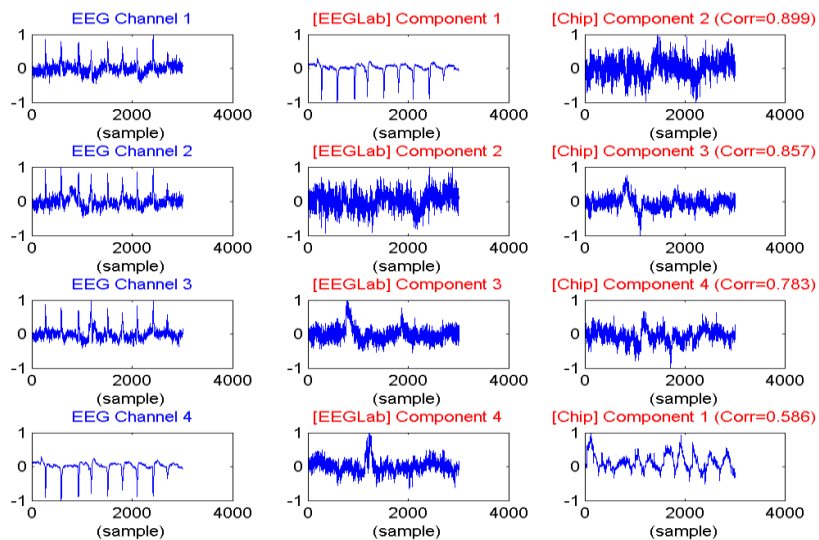

Figure 2.32   Time domain comparison between the off-line result from EEGLab and on-line results for pattern 4 with original source EEG shown in the first column
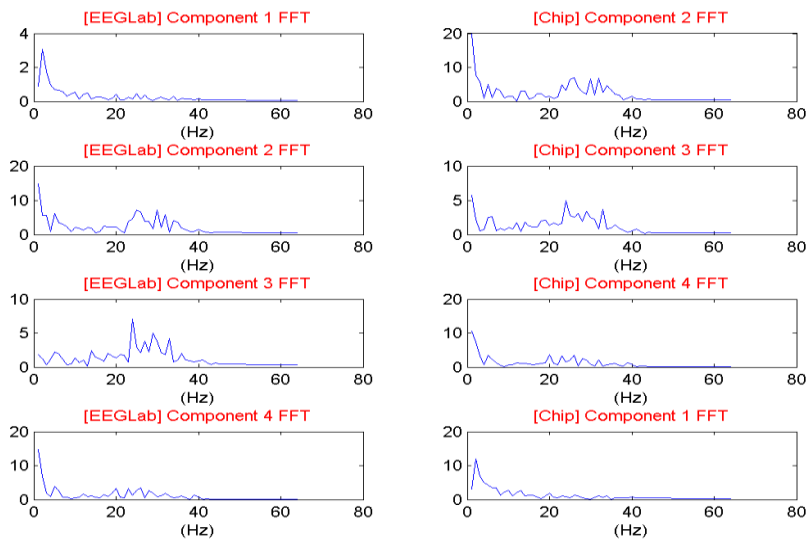
Figure 2.33   Frequency domain comparison between the off-line result from EEGLab and on-line results for pattern 4
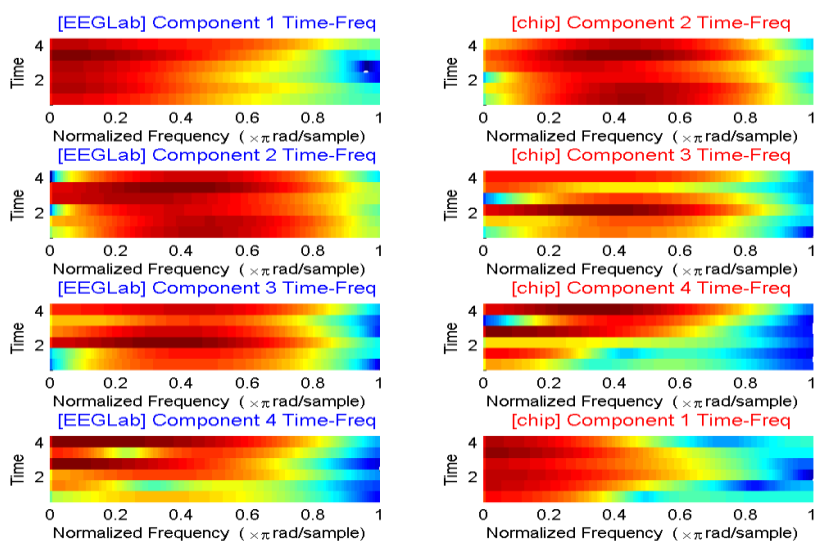
Figure 2.34   Time-frequency analysis comparison between the off-line result from EEGLab and on-line result for pattern 4

To summarize the performance analysis using real EEG signals, following points are concluded:

● Since the 4-channel ICA processor is designed to perform artifact removal from EEG signals, patterns recorded by NeuroScan system with and without eye-blink artifacts are applied to verify the design.

- All results are compared with the off-line algorithm performed by the latest release of EEGLab.

- To compare our work with the other on-line ICA designs, we also applied super-gaussian random patterns to the chip, and the result show that an average correlation of 0.86 is achieved.

- From the results of pattern 3 and 4, we can tell that the same independent components are locked at the same channel, even though the window is only 64 in length and may not contain any eye-blink artifact at some time. Locked channel order brings tremendous advantage for artifact removal. No matter the identification is done manually or automatically, if the artifact is always locked in the same channel, we need only one time identification.

2.3.3 Comparisons with Other Works

Table 2.6   Comparisons using complexity and average correlation coefficient

	Shyu [31]	Huang [32]	This Work
Application	EEG	Speech	EEG
Channel	4	2	4
Pre-Processing	No	Yes	Yes
Memory (bits)	384,000	24,576	4,352
Equivalent Gate-Count	N/A	315.5K	199.7K
ADC sample rate	64 Hz	16 KHz	128 Hz
ADC resolution	8 bits	N/A	10 bits
Correlation	> 0.8	N/A	0.86
Data format	Floating	Floating	Fixed
Algorithm	Infomax	Fast	Infomax

Table 2.6 shows comparison of the proposed design with other works on ICA. In our proposed design, the memory complexity is much lower while the average correlation coefficient is higher. Target application for [31] and the proposed design are the same, while the work in [32] is for speech separation thus need for much higher sample rate. A big advantage for the proposed design is that only small amount of memory is used result in lower power consumption and small area.

2.4 UMC 90nm 4-Channel ICA Processor Tape-out

The proposed design is fabricated using UMC 90nm SPRVT 1P9M process technology, the physical information of the chip is listed in Table 2.7. The power consumptions of the two boundary working conditions are also included in the table. For high efficiency condition, the system works at 60 MHz and powered by a 1.2 V core voltage, and the chip consumes average 12.24 mW. For low power condition, the system works at 0.817 MHz, the lowest boundary for successful completion of the worst cast 512 training iterations, and powered by a 0.6 V core voltage, and it consumes average 312 uW in this condition. In the low-power condition, the core voltage approaches the threshold voltage of the MOS, and even lower power supply is not possible without the involvement of sub-threshold technology.

Table 2.7   Physical information of the 4-channel ICA chip

Technology	UMC 90nm SPHVT 1P9M
Pad/Core Voltage	2.5V / 1.0V
Die Size	1.068 x 1.068 mm^2
Core Area	0.760 x 0.760 mm^2
Logic Gates	199.7K
On-Chip SRAM	544B
Operating Frequency	Up to 60MHz
Power Consumption	312uW (0.817MHz CLK, 128Hz EEG, 0.6V) 12.24mW (60MHz CLK, 9.708KHz Input, 1.2V)
# PAD	55 pins (functional / power : 31 / 24)
Test Package	68 pin LCC package

Micrographic of the fabricated chip with module partitions marked by their boundary is shown in Figure 2.35 (a). As we can see from Figure 2.35 (a), whitening unit occupies about 50 percent of the core area, while the Infomax training unit takes about 25 percent, and the memory size is optimized, so about 10 percent of core area is filled by the SRAM. The test chips used in section 2.4.1 to 2.4.3 is packaged using 68-pin LCC package shown in Figure 2.35 (b) accompanied by the bonding map in Figure 2.35 (c).

(a)

(b)   (c)
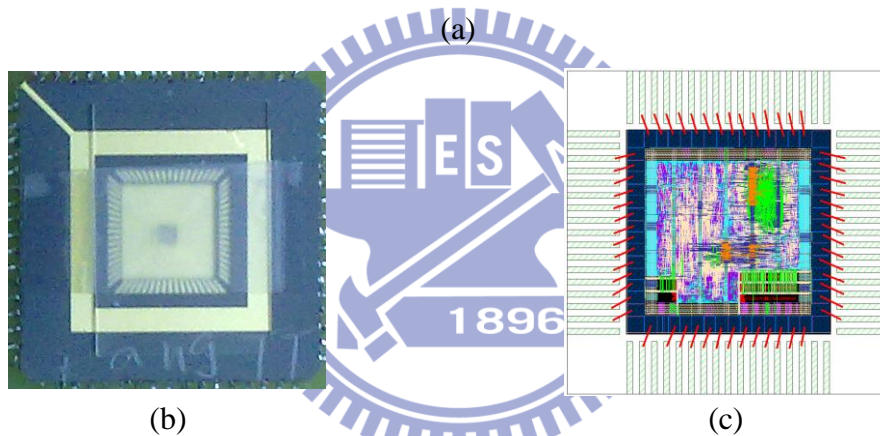
Figure 2.35   (a) Micrograph of the fabricated 4-channel ICA processor (b) Chip packaged by 68-pin LCC package (c) Bonding map

2.4.1 Functional Verification

For the functional test, the only thing we care about is that if the behavior of the chip matches the result of post-layout simulation which is done before the fabrication of the chip. Therefore, the functional test is done by applying the same super-gaussian pattern described in section 2.3.1 to the Agilent 93000 SOC Tester in CIC using normal condition that is applying 2.5 V for pad power and 1.0 V for core power, and the system working frequency equals to 32 MHz as the frequency for RTL synthesis is. The expected output pattern is generated using the original MATLAB code for

architectural exploration and hardware design verification with fix-point format which is identical to the behavior of the designed processor.
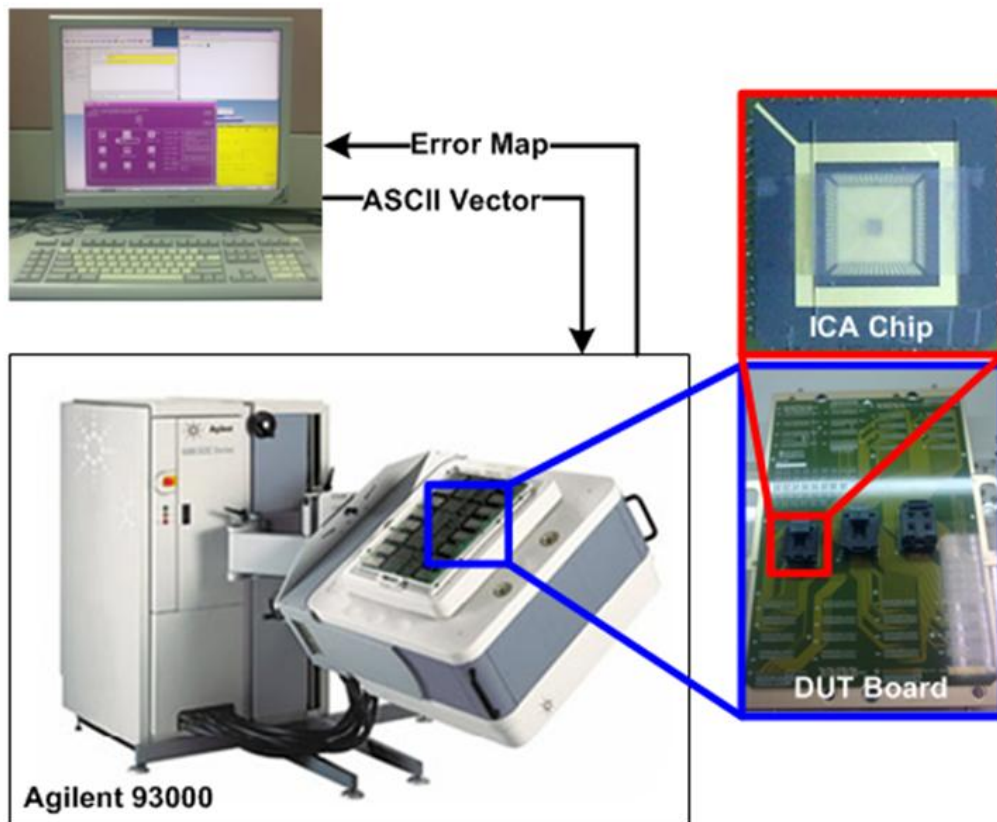
Figure 2.36   Chip testing using Agilent 9300 SOC Tester

The chip packaged by LCC-68 package is first insert to a DUT (design under test) board, and the DUT board is then connected to the Agilent 93000 SOC tester in Figure 2.36, and the pattern setup is done on a workstation connected with the tester. The tester is able to automatically compare the chip output with the expected output pattern, and once the result mismatches, the test process will abort and notice the user the chip fails the specified pattern under current condition. Bypass mode and normal mode for the ICA processor are both verified and successfully passed.

2.4.2 Power Consumption Analysis

In electrical engineering, a Shmoo plot [33] is a graphical display of the response of a component or system varying over a range of conditions and inputs. It's often

used to represent the results of the testing of complex electronic systems such as computers, ASICs or microprocessors. The plot usually shows the range of conditions in which the device under test will operate. For example, when testing semiconductor memory: voltages, temperature, and refresh rates can be varied over specified ranges and only certain combinations of these factors will allow the device to operate. Plotted on independent axes (voltage, temperature, refresh rates), the range of working values will enclose a three-dimensional, usually oddly-shaped volume. Other examples of conditions and inputs that can be varied include frequency, temperature, system- or component-specific variables, and even varying knobs tweakable during silicon chip fabrication producing parts of varying quality which are then used in the process.

- The maximum frequency is 60MHz
- The minimum core voltage is 0.6 V
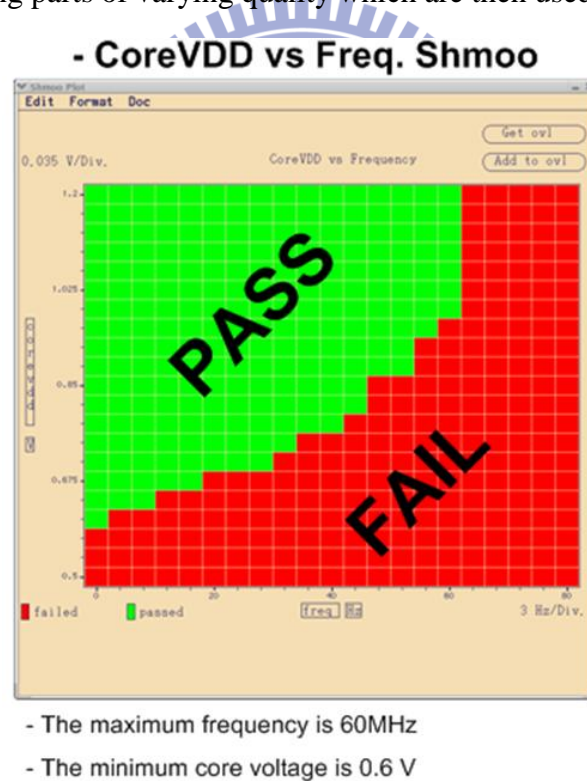
Figure 2.37   Shmoo plot showing the boundary of the working condition using core voltage and frequency as its axes

After the functional test is done, we still are interested by the boundary of the working condition and the corresponding power consumptions. One Shmoo plot shown in Figure 2.37 is scanned by the Agilent 93000 SOC tester with x-axis

representing the system clock rate and y-axis representing the core voltage. The top
boundary of core voltage is limited to 1.2 V because the targeted application for this
chip requires power consumption as low as possible, and the lower boundary of the
system clock rate is limited to 0.5 MHz which merely satisfies the cycle count needed
to perform ICA training for maximum 512 times. The average power consumption of
eight working conditions are measured and listed in Table 2.8 and Figure 2.38 shows
more distinct ratio how the power consumption can be minimized. The measurement
shows minimum 0.312 mW is consumed to perform independent component analysis
using the chip with a lower sample rate of 80 Hz.

Table 2.8   Power consumption table at different working conditions

	0.5MHz	5MHz	30MHz	60MHz
1.2V	4.392 mW	5.028 mW	8.328 mW	12.24 mW
0.9V	1.323 mW	1.629 mW	3.312 mW	-
0.6V	0.312 mW	-	-	-
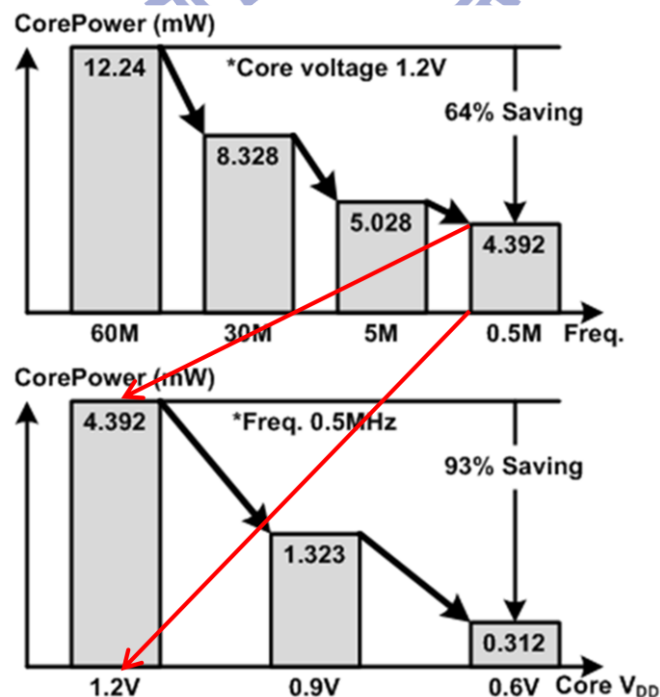
Figure 2.38   Power minimization

2.4.3 A FPGA-Based Testbed

To apply different pattern for performance analysis of the design, we setup a FPGA-based Testbed. One of the packaged chips is soldered to the self-made printed circuit board (PCB) shown in Figure 2.39. The PCB is connected to a SMIMS (北瀚) FPGA board using a standard 40-pin IDE bus, and the described setup is shown in Figure 2.40.
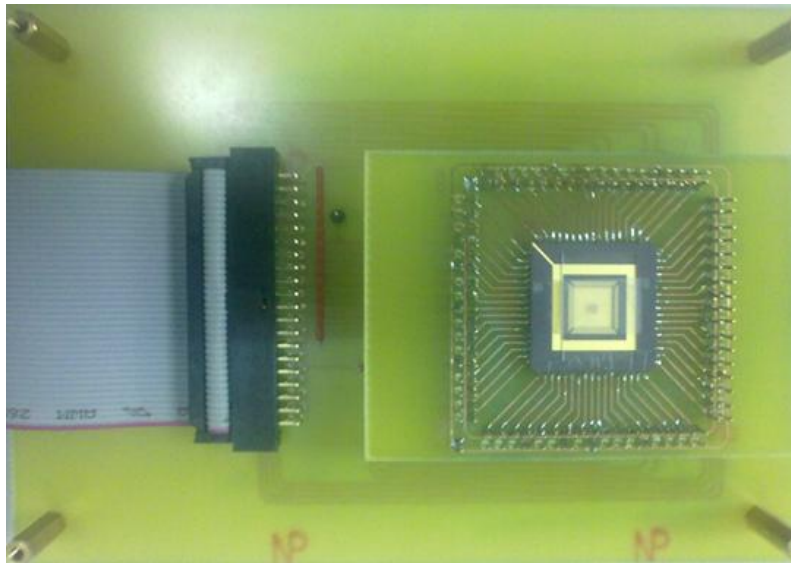
Figure 2.39   The self-made PCB with the packaged chip soldered on it

Figure 2.40   Connection between the FPGA board and the PCB

The FPGA is special and suitable for pattern feeding due to the design of a software controlled data interface shown in Figure 2.41. The FPGA and host computer is connected by a USB cable used for netlist downloading and run-time data exchange.

Engineer can create user interface programs to do whatever he want with the interface, for example input pattern feeding, output result fetching and hardware control.
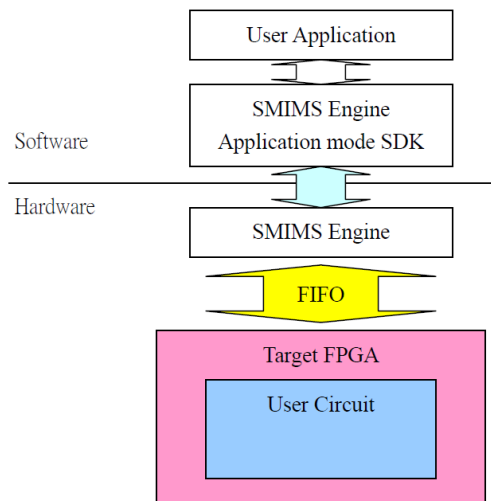
Figure 2.41   Software controlled data exchange interface between the FPGA and a host computer connected by USB

For the user end, SMIMS software development kit (SDK) provides a dynamic link library (DLL) file including the following basic function listed in Table 2.9 for data exchange between the host and FPGA. The argument iBoard in the table all represent for the board number to control and at most 2 SMIMS FPGA boards can be controlled by one host computer at a time.

Table 2.9   The data exchange functions provided by SMIMS SDK

| Function Name and Prototype	Description
bool SMIMS_VEX2_AppOpen (int iBoard, char * Serial)	Initialize the connection between host computer and the FPGA
bool SMIMS_ VEX2_AppClose (int iBoard)	Terminate the connection between host computer and the FPGA
bool SMIMS_ VEX2_AppFIFOReadData (int iBoard, WORD *Buffer, unsigned size)	Read block data from the output FIFO on the FPGA
bool SMIMS_ VEX2_AppFIFOWriteData (int iBoard, WORD *Buffer, unsigned size)	Write block data to the input FIFO on the FPGA
bool SMIMS_VEX2_AppChannelSelector (int iBoard, BYTE channel)	A 8-bit bus for free use
bool SMIMS_VEX2_ProgramFPGA (int iBoard, char * BitFile)	Download a netlist (.bit file) to the FPGA
char * GetLastErrorMsg (int iBoard)	When error occurs user can get the error information using this function

For the FPGA interface, the I/O port list is provided in Table 2.10.

Table 2.10 I/O port list for the FPGA interface

Port Name	Type	Width	Description
APP_CLK	In	1	System clock (48 Mhz)
APP_CS	In	1	Becomes 1 after the connection is opened
APP_RSTN	In	1	Raises for one cycle after programmed
CH	In	8	A 8-bit bus for free use
APP_RD	Out	1	Read signal to the input FIFO
APP_WR	Out	1	Write signal to the output FIFO
APP_DI	In	16	Input data bus from the input FIFO
APP_DO	Out	16	Output data bus to the output FIFO
APP_FULL	In	1	Indicates if the output FIFO is full
APP_EMPTY	In	1	Indicates if the input FIFO is empty
APP_AlmostFULL	In	1	Indicates the output FIFO is about to be full
APP_AlmostEMPTY	In	1	Indicates the input FIFO is about to be empty

Figure 2.42 and 2.43 are waveform examples for single cycle read/write.
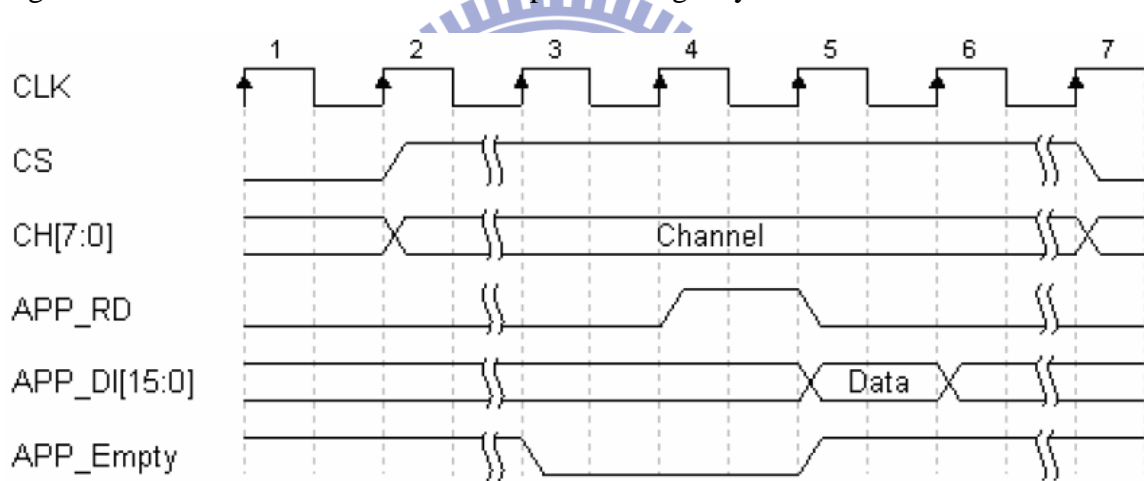
Figure 2.42   Waveform example for single cycle read operation
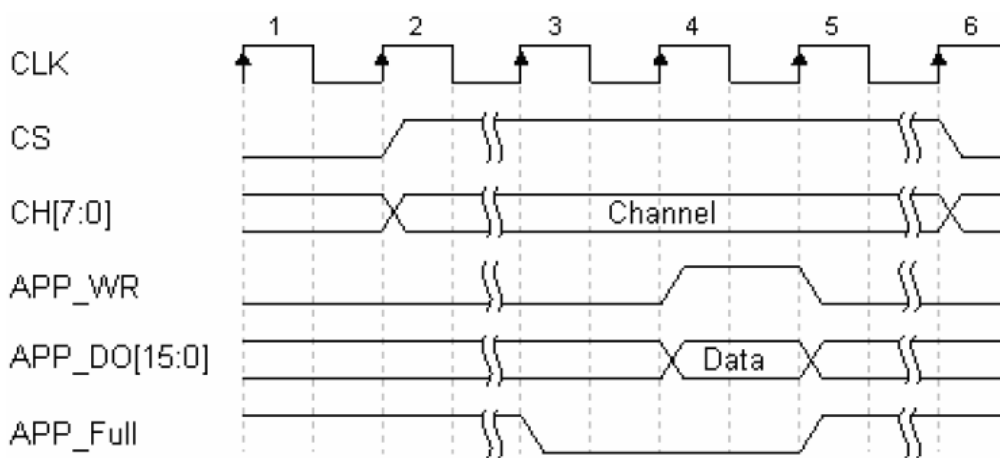
Figure 2.43   Waveform example for single cycle write operation

With SMIMS FPGA and its SDK, a FPGA-based Testbed is built to provde a experiment platform able to perform fast performance analysis of the fabricated chip. The hardware architecture is shown in Figure 2.44.
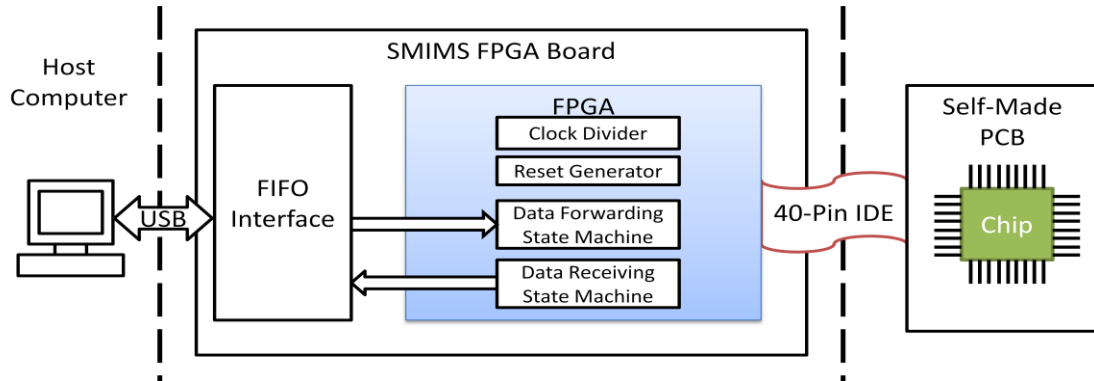
Figure 2.44   Hardware architecture of the FPGA-based Testbed

The experiment flow char and the software flor char are shown in Figure 2.45. On-line result and off-line result are compared and analyed by MATLAB. The generated results are the ones demonstrated in section 2.3.
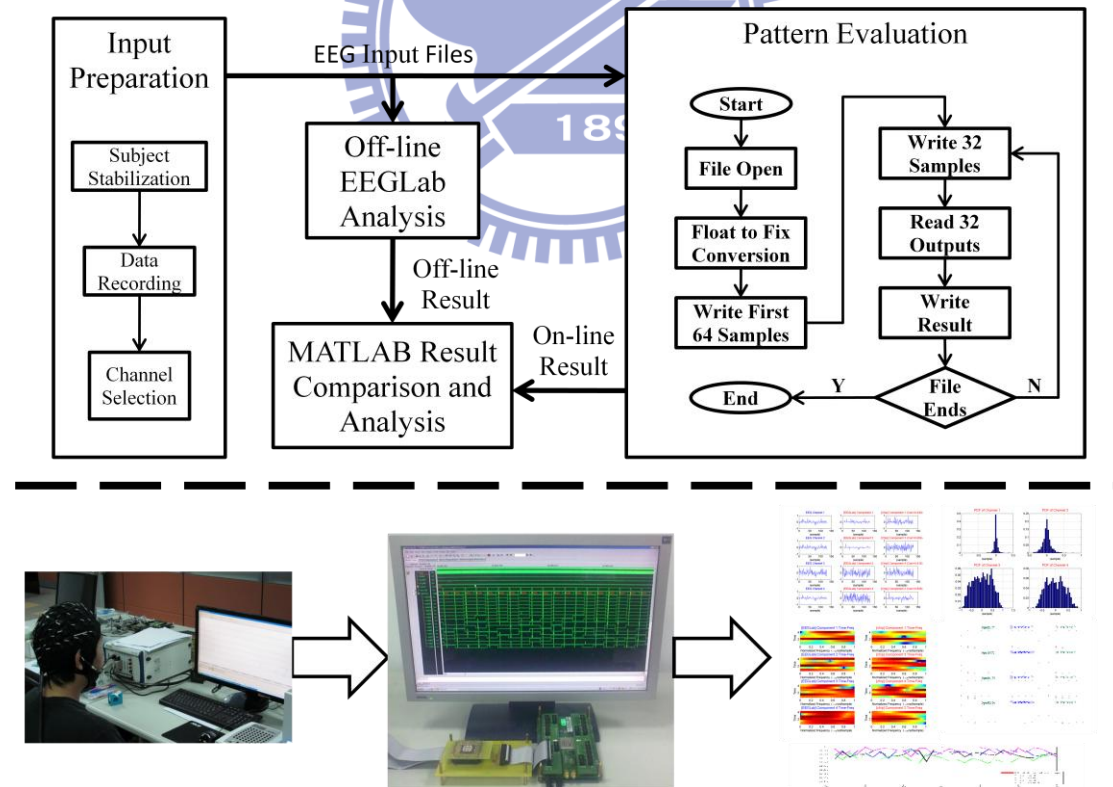
Figure 2.45   Flow char for the experiment for performance analysis using real EEG patterns recorded from the NeuroScan system

Chapter 3   Experimental Brain-Heart Monitoring SoC Tape-out

In chapter 1 the need for integrated health-care systems is already emphasized. Three common human health indicators, EEG, EKG and near-infrared spectrogram, are also pointed out to generate information suitable for joint analysis that is valuable to be used in three potential scenarios. To enable the practical development of such health-care systems, an experimental brain-heart monitoring SoC is integrated as a preliminary version.

3.1 Overall System Architecture

The overall system architecture is shown in Figure 3.1. To start the system function, the science station first sends a trigger command including the system working mode and compression mode to activate the whole system. The command is decoded and evaluated by the system control unit (SCU), and after the reset process, an internal trigger signal is then sent to the front-end interface control unit (FICU). Afterwards, time multiplexed data acquisitions are continuously triggered by two system counters in FICU. Table 3.1 shows the front-end specifications of the system.
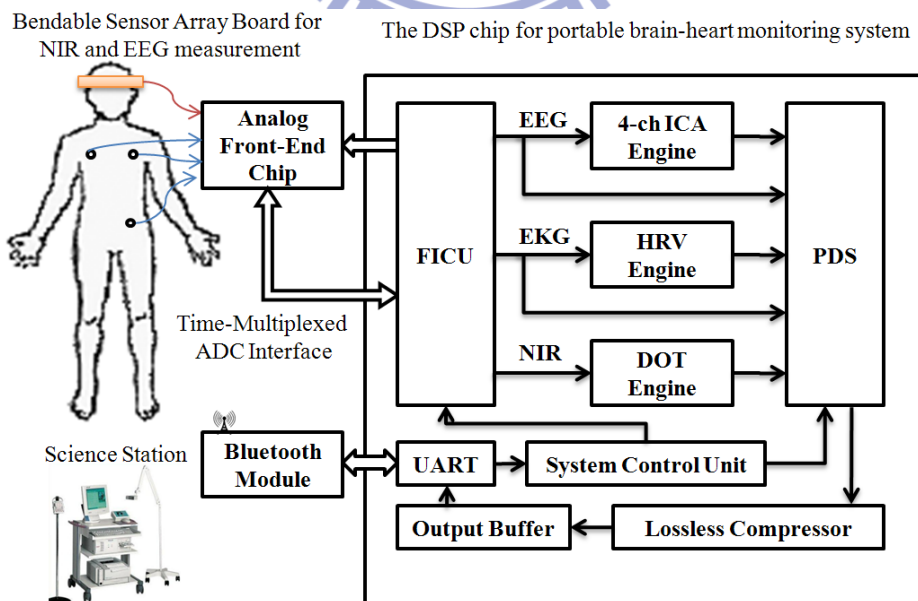
Figure 3.1   Digital signal processing chip architecture for the portable brain-heart monitoring system

Table 3.1 Specifications of the front-end circuits and working modes for the system

Parameter	DOT Sub-System	EEG Sub-System	EKG Sub-System
Sample Rate(Hz)	1 frame (24 sensor values)	128	256
#Sensor(Channel)	12	4	3
#NIR LED	6	-	-
LPF cut-off freq. (Hz)	10	50	100
Gain (dB)	-	250	5000
Output Range (V)	0~2.2 (external sensor)	0~2.5 (built-in IA)	0~2.5 (built-in IA)
ADC Resolution	10 bits		
Digitized Data Range	0~900	0~1023	0~1023
ADC Priority	3	1	2
ADC Order	Only one sample at each time	Ch1-Ch2-Ch3-Ch4	Ch1-Ch2-Ch3
PDS Priority	DOT:4	EEG/ICA:2	EKG : 1 HRV : 3
System Mode	DOT/Off	EEG/ICA/Off	EKG/EKG+HRV/Off
Compression Mode	On/Off	On/Off	EKG:On/Off HRV:Not Supported

Digitized raw bio-signals are sent to both the prioritized data selector (PDS) and the corresponding processors, that is, the ICA, HRV and DOT processor. When the processed data is available from the processors, PDS will check if the lossless compressor is busy or not. When the output buffer and the compressor are available, the queued data is then sent using a fixed priority with their compression mode settings. Compressed data is packed and sent through the Bluetooth wireless link between the system and science station, and further display and signal processing can be performed on the science station.

3.2 The Interface of the Analog Front-End Circuitry

The designed digital integrated SoC is interfaced with an analog front-end IC designed for the biomedical signals, and the specification is already listed in Table 3.1 in the previous section. At the back-end of the AIC, there is a time-multiplexed analog-to-digital convertor (ADC), and its behavior is shown in Figure 3.2. The maximum master clock rate for the ADC is 1200 KHz. Analog-to-digital conversions

are triggered by the *Start_Conversion* signal. The ADC requires constant 12 cycles to finish one conversion including the one which *Start_Conversion* raises. After the conversion is done at the twelfth cycle, the *End-of-Conversion* (EOC) signal will raise for one cycle, and at the same cycle the digitized sample of the bio-medical signal is ready to be read on the *ADC_OUT[9:0]* bus. In Figure 3.2, two complete analog-to-digital conversions are shown with their valid values marked on the waveform. For the digital integrated SoC, the *end-of-conversion* (EOC) signal can be directly used as an *INPUT_VALID* signal.

Figure 3.2   Analog-to-digital conversion (ADC) simulation waveform

The AIC is interfaced to the designed system via the front-end interface control unit (FICU). The 10 KHz and 1.2 MHz clocks are generated and provided to the AIC. The data acquisition scheduling, detailed design and behavior of the FICU will be described in section 3.5. In addition, channel selection table and detailed interface connection is also provided in section 3.5.

3.3 Three Bio-Signal Processing Processors

Besides the function of biomedical signal acquisition and raw data transmission, we also developed three independent modules including the 4-channel independent component analysis (ICA) processor, a heart-rate variability (HRV) analysis processor and a near-infrared (NIR) diffuse optical tomography (DOT) processor for biomedical signal analysis and image reconstruction of the absorption coefficients. The three modules are used as hardware tools for easing the computation load on the science station. The 4-channel ICA processor is already well presented in chapter 2, so only emphasis of the importance to include the ICA processor in this system will be supplemented in this section. Brief descriptions of the other two processors are also given below, but the number of page for them will be limited due to the focus on the system design. Although the detailed designs and key technology of the other two processors will not be fully presented in this thesis, references to the corresponding works will be listed at the end of each sub-section.

3.3.1 4-Channel ICA Processor

As we know from chapter 2, computation complexity of the ICA algorithm is tremendously high due to the dependency between each channel. As a consequence, off-line ICA algorithm is usually applied to the EEG signal for artifact removal on pc-based science station. In practical application, real-time ICA is much more useful than off-line processing. Therefore, the 4-channel ICA processor described in chapter 2 is employed for real-time artifact removal, and it is valuable to be placed in the system as a hardware accelerator.

The EEG signal acquired by the front-end interface control unit is forwarded to the ICA processor if the system mode of ICA processor is activated. Afterward, the ICA processor acts identically as we described in chapter 2, so redundant description

is avoided for the ICA processor. Before we proceed to the monograph for the HRV

and DOT processor, the top module view of the ICA processor is shown in Figure 3.3,

and the I/O behavior of the ICA processor is depicted in the waveform in Figure 3.4.
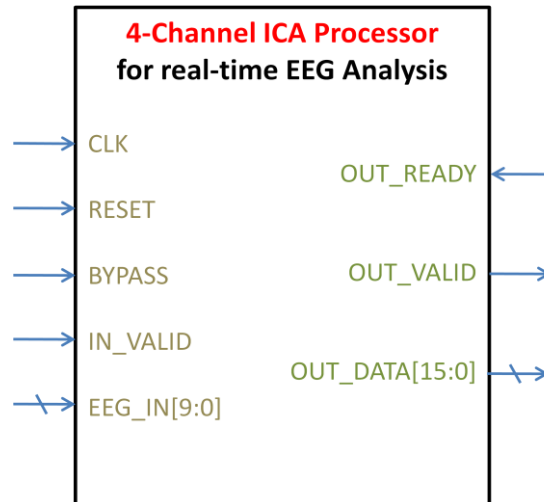

Figure 3.3   The top module view of the ICA processor

IO sequence of the ICA processor is parallelized for port reduction, and the
resulting sequence is shown in Figure 3.4. The basic input unit is 4 samples from each
channel, and the output unit is 128 component values of estimation of one
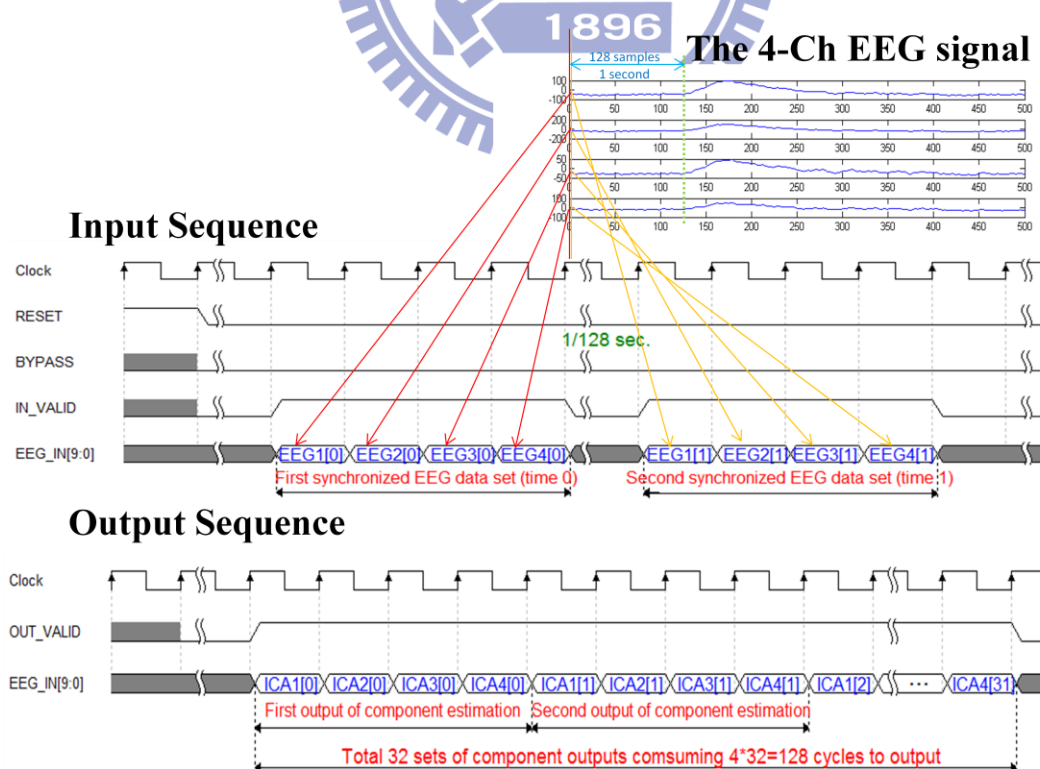half-window.

Figure 3.4   The IO sequence of the ICA processor

3.3.2 Heart-Rate Variability Analysis Processor

Heart rate variability (HRV) is a physiological phenomenon where the time interval between heart beats varies. It is measured by calculation of the variation in the beat-to-beat interval. Methods used to detect heart beats include EKG, blood pressure and photoplethysmograph (PPG). Among these methods, EKG analysis is considered a superior ways to perform HRV analysis, because the signal is relatively clear and stable. The R-peak of EKG introduced in section 1.1.3 is used for detection of the heart beat, because it is easier to be detected.
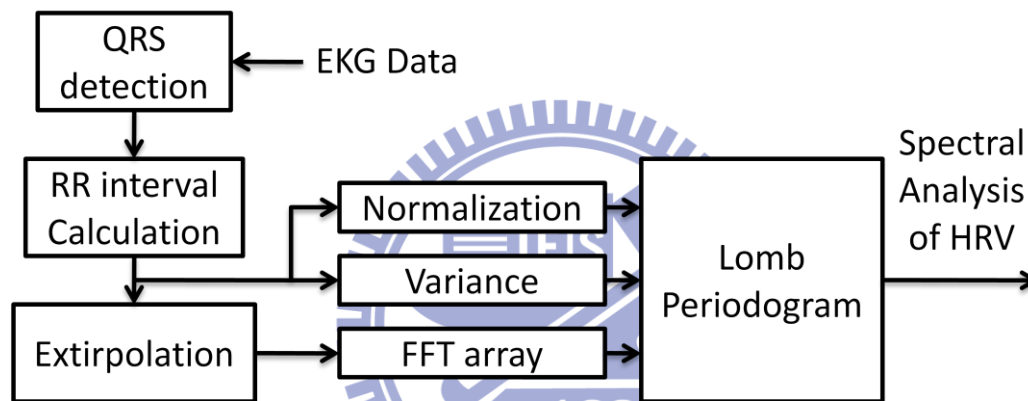
Figure 3.5   Architecture of the HRV analysis processor

A novel frequency domain HRV analysis processor using a fast windowed Lomb periodogram [34] is designed and employed in the system. The Lomb time-frequency distribution (TFD) is suited for spectral analysis of unevenly spaced data and has been applied to the analysis of heart rate variability. The HRV processor in Figure 3.5 comprises the hardware implementation of the Lomb TFD as well as a simple RR interval calculation unit. In consideration of architecture simplicity and real-time properties, the classical derivative-based QRS detection algorithm introduced by Pan and Tompkins [35] was adopted as a baseline for the RR interval calculation unit. In the RR interval calculation unit, EKG signals first pass through a set of linear processes, including a band-pass filter comprising a cascaded low-pass and high-pass,

and a derivative function. Non-linear transformation is then employed in form of a signal amplitude squaring function. Finally, a threshold is applied to detect the R-peaks of the QRS complexes. The RR intervals are then calculated from the detected peaks and HRV analysis is performed. Better time-frequency analysis of HRV is achieved through a de-normalized fast Lomb periodogram with a sliding window configuration similar to the one applied to the ICA processor. RR intervals detected in two minutes are analyzed to generate a 256-point complex spectrum in each window.

The HRV processor is verified using the MIT/BIH database [36] and results of the QRS detection algorithm were compared with offline simulations. The output of the QRS detection algorithm and the power spectrum of the RR intervals are shown in Figure 3.6 and 3.7 respectively.
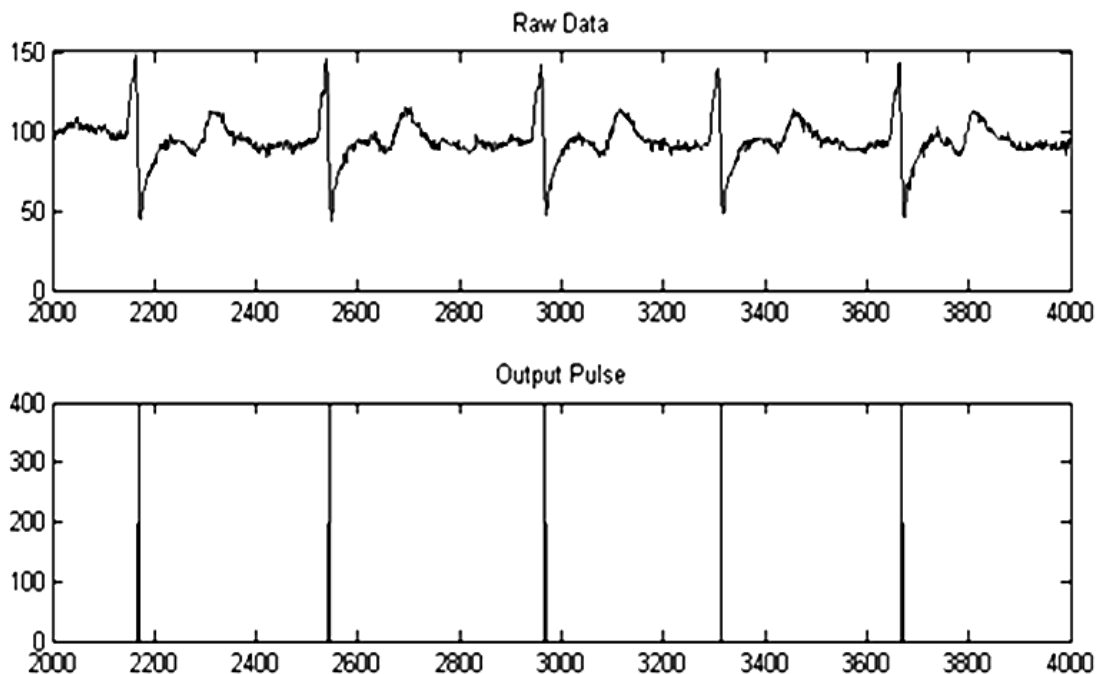
Figure 3.6   Input and the resulting output from the QRS detection unit
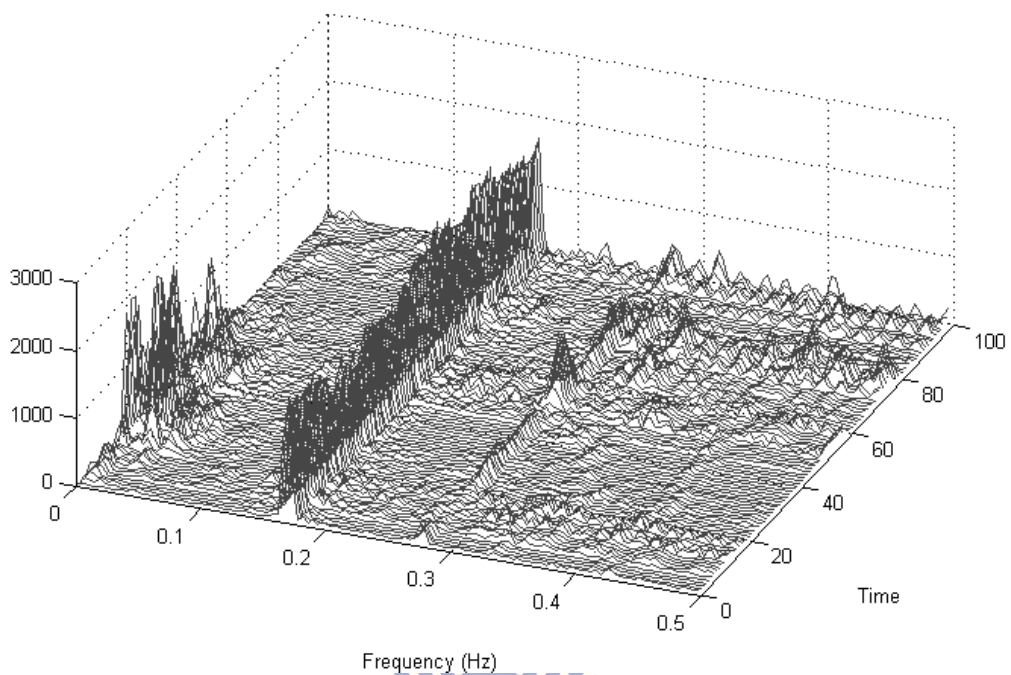
Figure 3.7   Time-frequency HRV analysis of EKG data from MIT-BIH arrhythmia
database using Lomb TFD

The top level module view is shown in Figure 3.8. One sample will be put on the
EKG_DATA bus every 1/128 seconds synchronous to the *IN_VALID* signal. For every
minute, continuous 256 point of complex spectrum values are output sequentially.
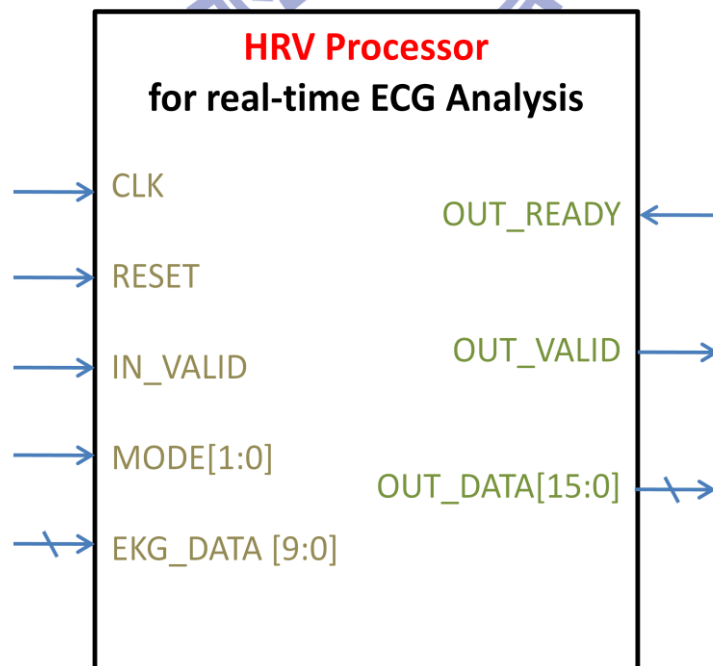
Figure 3.8   Top level module view of the HRV analysis processor

3.3.3 Near-Infrared Diffuse Optical Tomography Processor

The DOT processor is designed using Continuous Wave (CW) algorithm [37] comprising a sub-frame operation control circuit [38], a DOT reconstructor and an image post-processor, and its overall architecture is shown in Figure 3.9 (a) while a top level view is shown in Figure 3.9 (b). The CW DOT algorithm can be divided into a forward model and its inverse problem. Forward model includes abundant optical parameters and mathematic equations. Indeed, when the depth of the surface to be calculated is fixed, the inverse matrix is always the same. Therefore, the pre-calculated inverse matrices are stored in a look-up table. The DOT reconstructor controlled by the sub-frame operation control circuit is mainly used to perform matrix operations. Pixels reconstructed represent absorption coefficient variance $\Delta\mu_{x,y}^a$. Normally, $\Delta\mu_{x,y}^a$ are too small to be observed, so an image post-processor is employed to perform linear mapping, contrast enhancement and color mapping so that we can observe clear images on the science station monitor.

(a)	(b)
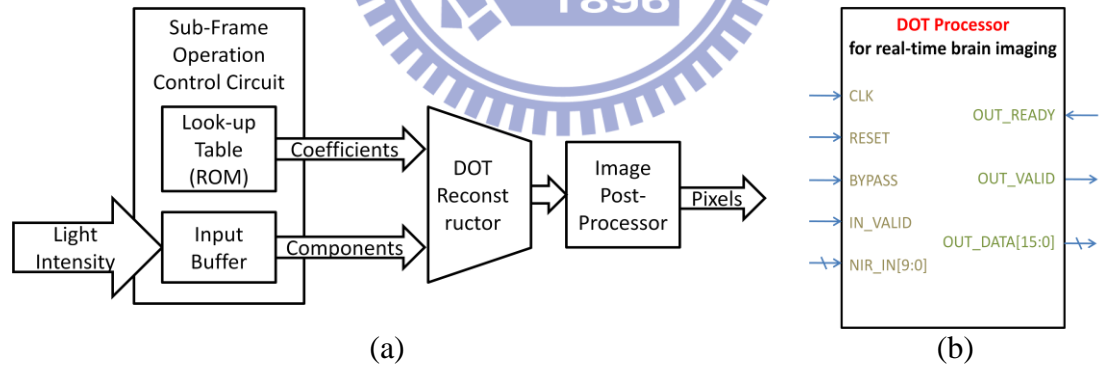
Figure 3.9   (a) The architecture of the DOT processor (b) Top level module view of the DOT processor

To verify the proposed DOT processor, we use the experimental model shown in Figure 3.10. The frame area is 4x6cm^2 and the volume of each voxel is (0.25cm)^3. The background medium is homogenous with $\mu_a^{bg} = 0.05 cm^{-1}$ and the reduced scattering coefficient $\mu_s^{'bg} = 10 cm^{-1}$. Two kinds of inhomogeneous media were embedded at depth of 0.5cm below the surface. The absorption coefficients of the inhomogeneous

mediums A and B (yellow) are 0.21cm^{-1} and 0.5cm^{-1} respectively. The reconstructed colored image using the processor is shown in Figure 3.10 (b).

(a)   (b)
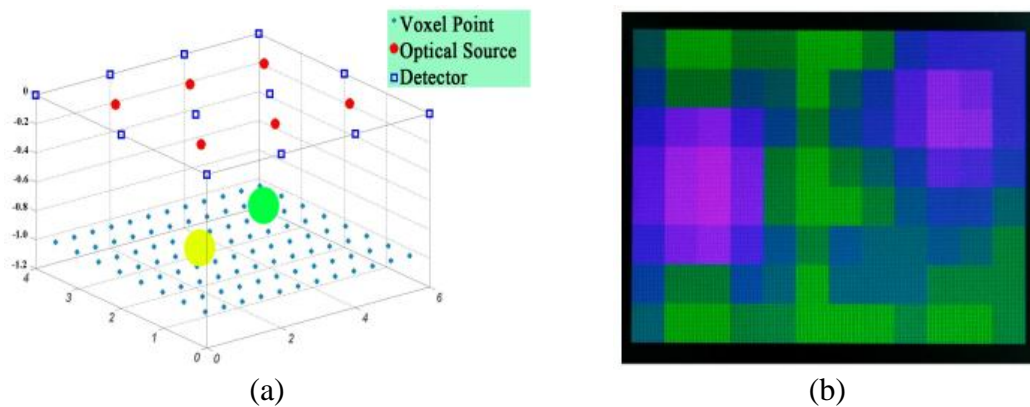
Figure 3.10   (a) An experimental model for the DOT processor (b) The reconstructed image on the LCD of the development platform

To make sure the near-infrared doesn't leak through the space between the LED and the measuring surface, we designed the DOT sensor array board using bendable printed circuit board (PCB). Aside from the NIR sensors and the bi-wavelength LEDs, there is an analog multiplexer chip for selecting the channel to be digitized and a decoder chip for selecting the LED to emit near-infrared. The timing chart of IO sequence is shown in Figure 3.11.
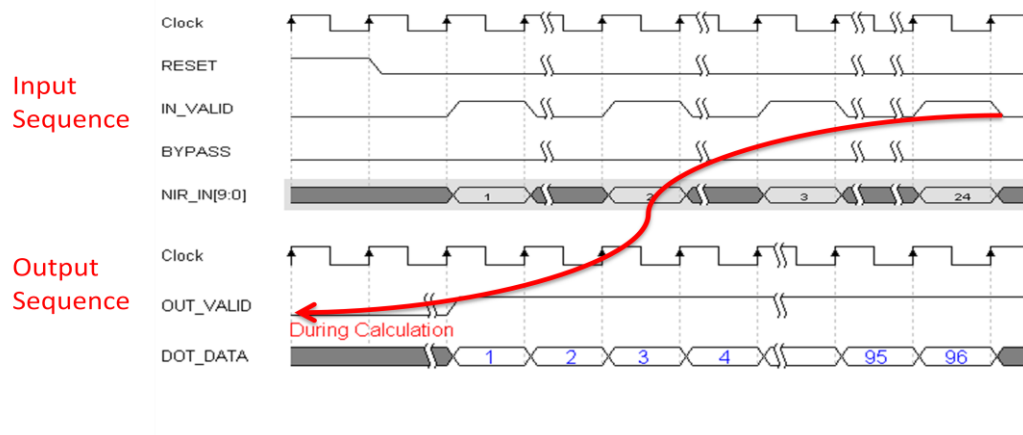
Figure 3.11   The timing chart of IO sequence of the DOT processor

3.4 System Control Unit

The system control unit (SCU) is responsible for the system initialization and the generation of the gated clocks for the processors, and the flow chart of the system initialization is shown in Figure 3.12.
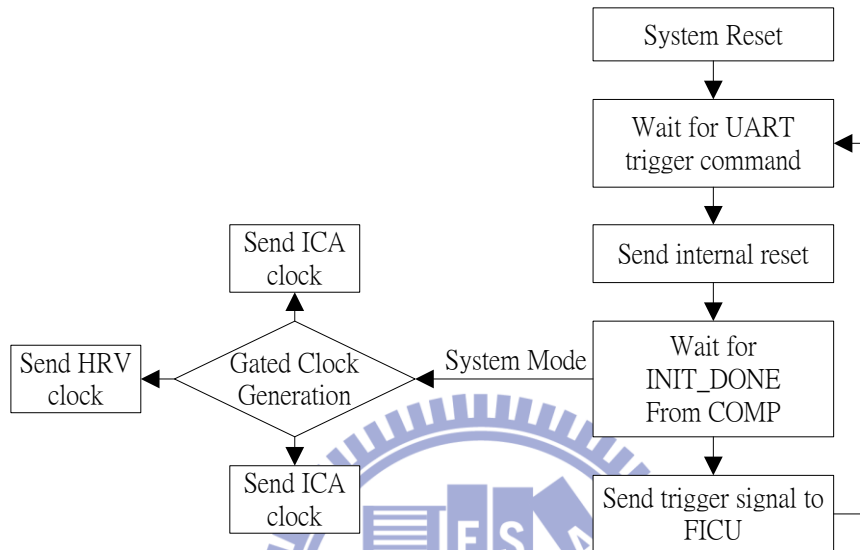
Figure 3.12   Flow chart of the system initialization

When the external system reset is sent from outside of the chip, the chip enters inactive state. An activation command from the UART module needs to be received to activate the system, and the command includes the system modes and the compression modes listed in Table 3.2.

Table 3.2   System activation command

Bit	Mode Selection	Clock Gating
0	Activate EEG analog to digital conversion	ICA
1	Activate 4-channel ICA processor	ICA
2	Activate EKG analog to digital conversion	HRV
3	Activate HRV Processor	HRV
4	Activate NIR ADC and DOT Processor	DOT
5	Bypass EEG compression	-
6	Bypass EKG compression	-
7	Bypass DOT compression	-

After the activation command, an internal reset is sent to every other module. The reset process takes only one cycle for all modules inside the chip except the compression module that requires 96 cycles for initialization, and the waveform in the

duration is shown in Figure 3.13.
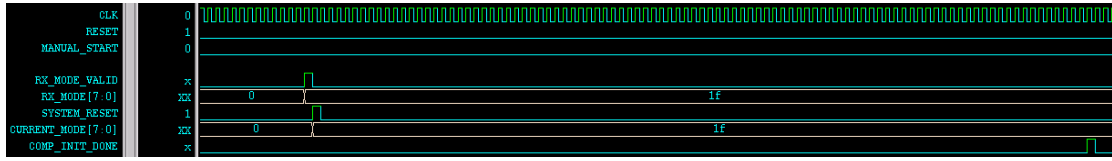
Figure 3.13　　The duration from the system reset is sent to the end of initialization inside the compression module

When the signal *INIT_DONE* from the compression is received, the SCU starts to generate the clocks for the three processors according to the system mode received from the UART module. If any one of the bio-signal acquisition is not activated or is set to transmit the raw data, the clock to that processor will be turned off, so the redundant power consumption can be saved by clock tree trimming. In the meanwhile, a trigger signal is sent to the front-end interface control unit (FICU), and then the data acquisitions are automatically scheduled afterwards.

3.5 Front-End Interface Control Unit

The front-end interface control unit (FICU) is designed to actively acquire the bio-medical signal from the analog front-end IC. The connection between the analog front-end IC and the proposed digital chip is shown in Figure 3.14.
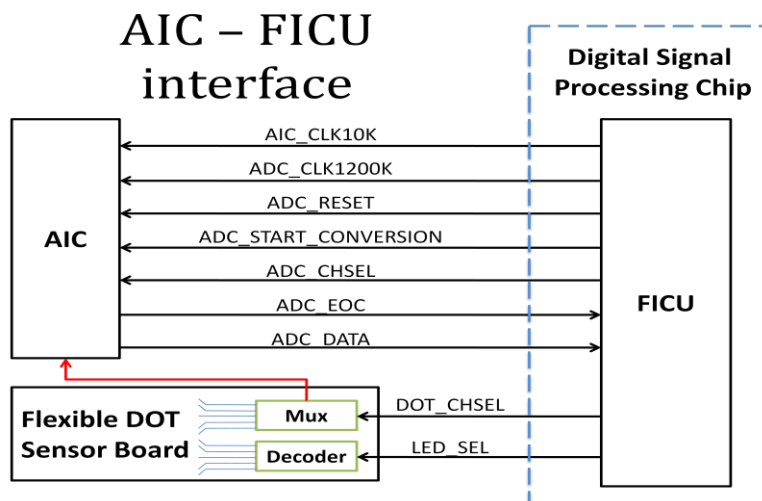
Figure 3.14　　Interface connection between the front-end interface control unit and the analog front-end integrated chip

Two clock dividers are built inside the FICU, and they are used to provide the 10 KHz and 1200 KHz clocks to the AIC. The 10 KHz clock is needed for the internal chopper-stabilized differential difference instrumentation amplifier (CHDDA) and a low-pass switched-capacitor filter (SC LPF). The 1200 KHz clock is provided as the master clock of the analog-to-digital converter (ADC) for the AIC, and an active-high reset signal which is synchronous to the master clock is also generated when the system is reset.

The scheme of FICU and the main state machine for data conversion are shown in Figure 3.15. Due to the fact that an analog multiplexer is used to perform channel selection, after the *ADC_CHSEL* changes, the analog signal routed to the input of ADC takes few cycles to get stable not to mention the *ADC_CHSEL* signal is connected from outside of the chip that may introduce much more delay time. As a matter of fact, when we immediately start a conversion after *ADC_CHSEL* is changed, invalid digital data will be converted. Therefore, an additional dummy conversion is invited to solve this problem.
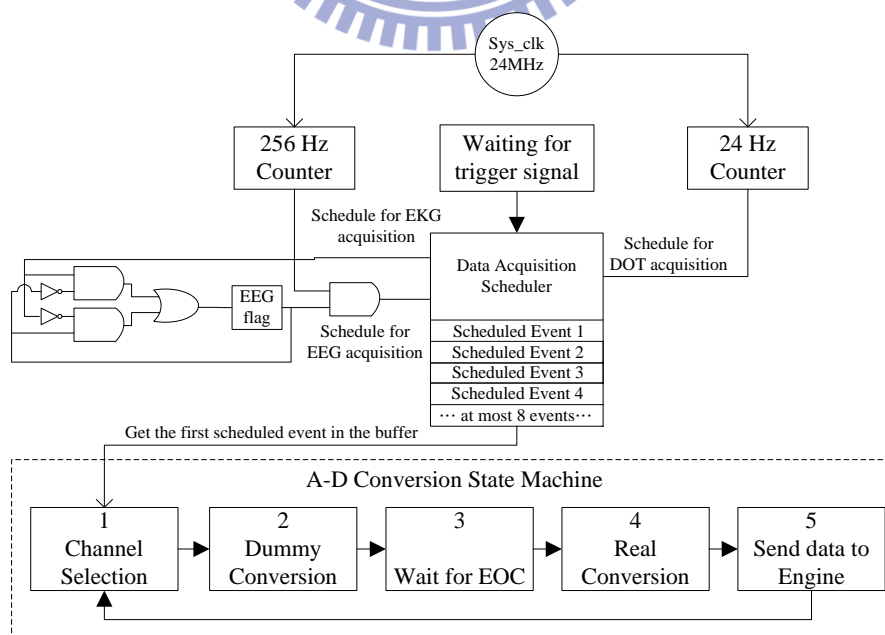
Figure 3.15   The scheme and state machine of FICU

When a scheduled conversion is accepted, FICU first changes *ADC_CHSEL* to select the specified channel to be converted, and then the dummy conversion is started by sending *ADC_START_CONVERSION*. An analog-to-digital conversion takes total 12 cycles to complete. Right after the dummy conversion, a real conversion is started, and a same process is taken to derive the valid data. When the conversion finishes, the converted data is available on the 10-bit *ADC_DATA* bus and it is synchronous to the *ADC_EOC* (end-of-conversion) as we previously discussed in section 3.2.

A model of the DOT sensor board is shown in Figure 3.16. Note that there are 6 near-infrared LEDs and twelve light intensity sensor for near-infrared on the sensor board. Only one LED is turned on at each time to avoid interference not from the nearby LED. Each time an LED is turned on, the four light intensity values from nearby sensors are acquired and digitized by the ADC. Therefore, before the image reconstruction of one frame can be launched, total 24 sensor values are converted.
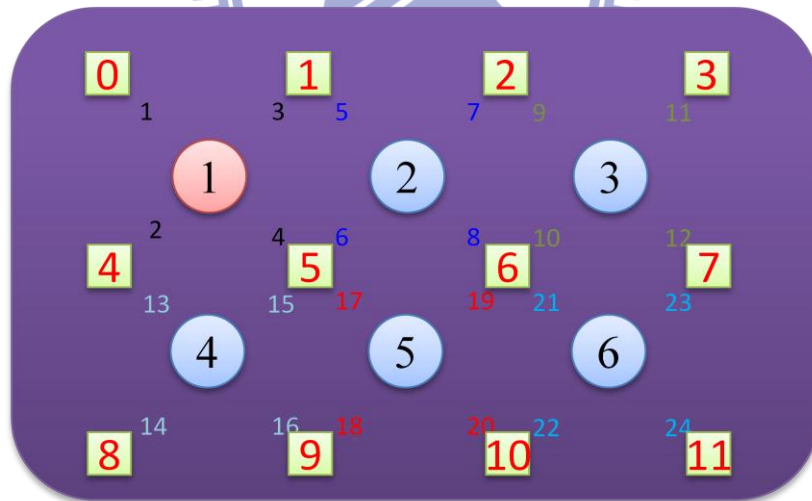
Figure 3.16   A model of the DOT sensor board

Each time the last conversion for conversion numbers marked with the same color is done for one LED, the *LED_SEL* is switched to the next LED in advance to avoid possible unstable conversions after switching LEDs. For example, after conversion eight is done, FICU will immediately switch the LED to the third one.

The full channel mapping table of the *ADC_CHSEL* and *DOT_CHSEL* signal provided to the analog multiplexer and the bendable DOT sensor board in the AIC is listed in Table 3.3. After the conversions of near-infrared sensor values with conversion numbers marked in red, the LED will be switch to the next one.

Table 3.3 Channel and conversion mapping controlled by *ADC_CHSEL*, *DOT_CHSEL* and *LED_SEL*

ADC_CHSEL	Selected Channel		
000 (Ch1)	EEG Channel 1		
001 (Ch2)	EEG Channel 2		
010 (Ch3)	EEG Channel 3		
011 (Ch4)	EEG Channel 4		
100 (Ch5)	EKG Channel 1		
101 (Ch6)	EKG Channel 2		
110 (Ch7)	EKG Channel 3		
111 (Ch8)	*LED_SEL*	*DOT_CHSEL*	Selected DOT Conversion
	000 (LED 1)	0	1
		4	2
		1	3
		5	4
	001(LED 2)	1	5
		5	6
		2	7
		6	8
	010 (LED 3)	2	9
		6	10
		3	11
		7	12
	011 (LED 4)	4	13
		8	14
		5	15
		9	16
	100 (LED 5)	5	17
		9	18
		6	19
		10	20
	101 (LED 6)	6	21
		10	22
		7	23
		11	24

3.6 Three-Stage Backward Handshaking Mechanism

The processed and raw data flow in the designed system is controlled by a three-stage backward handshaking mechanism shown in Figure 3.17. Data flows forward while the handshaking control propagates backward. When the output buffer in the UART module is not full, the compressed data is forwarded to the output buffer in UART for wireless transmission. Otherwise, the first stage handshaking mechanism will hold the compression module from outputting compressed data. The second stage handshaking works similarly to the first stage. The two conditions that activate the second stage handshaking are listed below:

● The compression is busy compressing the previous accepted data.

● The output buffer in UART module is full causing the data packing buffer in compression module also full. In this condition, input data from the PDS is not allowed, because no buffer space is available in the compression module.
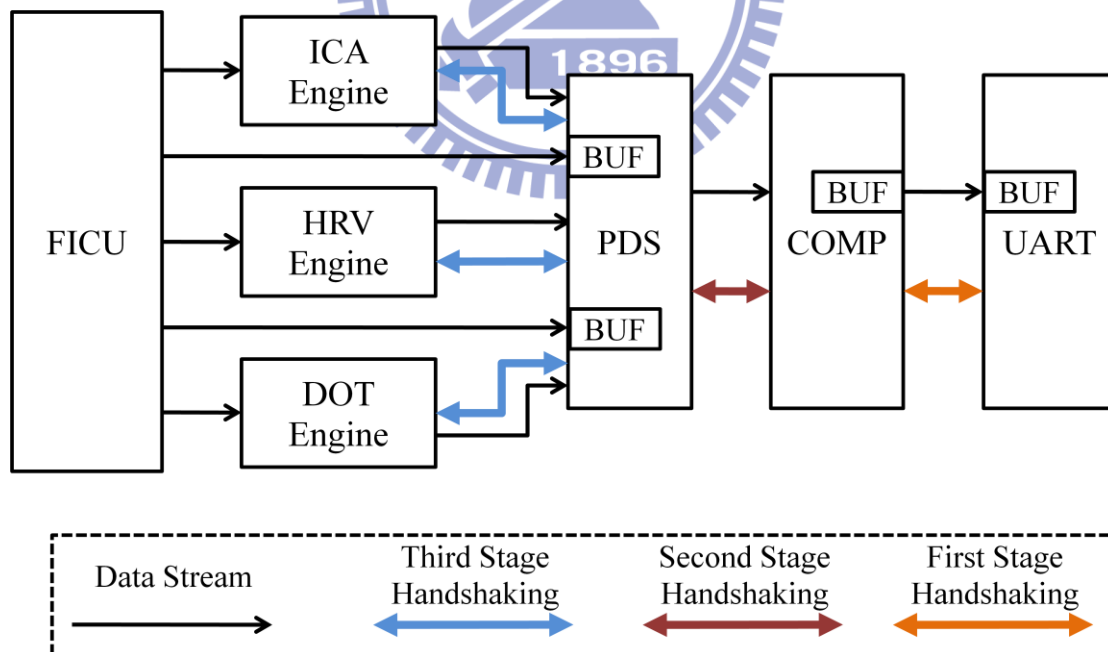

Figure 3.17   Three-stage backward handshaking mechanism

Similar to the first and second stage handshaking, the third stage handshaking doesn't allow data to pass when the second stage handshaking is active.

Chapter 4   Conclusion and Future Works

4.1 Conclusions

In this thesis, a hardware and power efficient 4-channel ICA processor is implemented using various techniques. Optimized data windowing, 3-bank circular memory allocation and an optimized mirrored non-linear lookup unit have been employed to reduce memory usage and power consumption. Operation pipelining between the ICA training and component extraction not only shortens the output delay by 0.25 seconds but also increases the hardware efficiency. A hardware efficient ICA training unit which comprises shared multiplier and adder arrays and a data routing matrix capable of pipelining the large matrix operations is designed. All the employed techniques and optimizations result in a 95 percent reduction of power, an 85 percent reduction of memory usage and an 87.5 percent reduction of ROM size. In addition, for the targeted application using 128 Hz sample rate, a 0.817 MHz clock frequency is needed under the worst case condition of 512 iterations in each window. On the other hand, when using 60 MHz clock frequency, a maximum data rate produced by 9.708 K sample rate can be processed under the worst case condition.

Both super-gaussian random signals and real EEG signals are applied to evaluate the performance of the designed processor. The result shows an average 0.86 correlation between the original super-gaussian source and the extracted ICA components is achieved. Over 0.8 average correlation between off-line result and the on-line processing using real EEG signals with or without eye-blink artifacts is also demonstrated. An example of artifact removal has shown that similar results are produced by our designed chip and the EEGLab.

The designed ICA processor is fabricated using UMC 90 nm technology. The size of the core area is 780x780 μ m^2. Functional verification and power

measurements are done using Agilent 93000 SoC tester. Power measurements show that a minimum 0.312 mW is consumed to perform independent component analysis using the chip with sample rate of 80 Hz. In addition, the stable convergence property of the ICA algorithm also provided us clues for designing a power efficiency improved version of ICA processor that consumes only 20 percent of the original power.

A preliminary portable brain-heart monitoring system comprises an fNIR-DOT processor, a 4-channel ICA processor that achieves 0.86 of correlation and a HRV analysis processor using Lomb periodogram is designed and integrated. Signals acquired from the front-end sensor modules are processed in real-time or bypassed according to user configurations, and are then losslessly compressed and packaged by a biomedical signal compressor achieving an average 2.5 CR before being wirelessly sent to a base-station with a commercial Bluetooth module, and the packaging protocol adds only additional 5 percent overhead. Internal data flow is controlled by a prioritized data selector that ensures the output buffer utilizations are not wasted in the three processors so that the output buffer in UART module can be reduced by using a three stage handshaking mechanism. By integrating three biomedical systems into a single chip, bulk associated with external circuitry is reduced. The ICA and HRV processor are verified by real EEG and EKG signals while the DOT processor is verified by an experimental model.

4.2 Future Works

For the ICA processor, an automatic artifact removal scheme is an essential function. The artifact removal process includes one step of artifact recognition. Traditionally, the artifact component is recognized manually or semi-manually. This process degrades the EEG system to an off-line analysis system. To overcome this problem, automatic artifact recognition can be done using statistic analysis or wavelet [39]. The number of channel in the ICA processor should also be raised for higher precision, because there are definitely more than four components in the EEG signal.

For the integrated system, further power and cost efficiency can be achieved by further integration of the AIC, DIC and the wireless transmission module by using system-in-package (SIP) technology. ICA design with more channels is possible by using external memory chip. By integration of the system and the external memory chip using SIP technology, the connection path of massive amount of data exchange can be shortened. Improvements for the next stage also include the adoption of more advanced low-power techniques like power shut-off (PSO) and dynamic voltage and frequency scaling (DVFS) that are suitable for this system. With these advanced techniques, the operation time of the system can be further prolonged.

Reference

[1] Malarvili, M.B.; Mesbah, M., "Combining newborn EEG and HRV information for automatic seizure detection,"*Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pp.4756-4759, 20-25 Aug. 2008

[2] Abdullah, H.; Holland, G.; Cosic, I.; Cvetkovic, D., "Correlation of sleep EEG frequency bands and heart rate variability," *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pp.5014-5017, 3-6 Sept. 2009

[3] Fidopiastis, Cali; Hughes, Charles; , "Workshop 1: Use of psychophysiological measures in virtual rehabilitation, " *Virtual Rehabilitation, 2008*, pp.xi-xi, 25-27 Aug. 2008

[4] Rankine, L.; Stevenson, N.; Mesbah, M.; Boashash, B., "A Nonstationary Model of Newborn EEG," Biomedical Engineering, IEEE Transactions on , vol.54, no.1, pp.19-28, Jan. 2007

[5] Ashtiyani, M.; Asadi, S.; Birgani, P.M.; , "ICA-Based EEG Classification Using Fuzzy C-mean Algorithm," 3rd International Conference on Information and Communication Technologies: From Theory to Applications, 2008. ICTTA 2008. , pp.1-5, 7-11 April 2008

[6] R. Vigario, "Extraction of ocular artifacts from EEG using independent component analysis," Electroencephalogr. Clin. Neurophysiol., vol. 103, pp. 395–404, 1997.

[7] R. Vigario, J. Sarela, V. Jousmaki, M. Hamalainen, and E. Oja, "Independent component approach to the analysis of EEG and MEG recordings," IEEE Trans. Biomed. Eng., vol. 47, no. 5, pp. 589–593, May 2000.

[8] Hosna Ghandeharion, Abbas Erfanian "A fully automatic ocular artifact suppression from EEG data using higher order statistics: Improved performance by wavelet analysis" Medical Engineering & Physics 32 (2010) 720–729.

[9] Muhammad Tahir Akhtar,Christopher J. James" Focal Artifact Removal from Ongoing EEG − A Hybrid Approach Based on Spatially-Constrained ICA and Wavelet De-noising" 31st Annual International Conference of the IEEE EMBS Minneapolis, Minnesota, USA, September 2-6, 2009.

[10] Boas, D.A.; Brooks, D.H.; Miller, E.L.; DiMarzio, C.A.; Kilmer, M.; Gaudette, R.J.; Quan Zhang; , "Imaging the body with diffuse optical tomography," Signal Processing Magazine, IEEE , vol.18, no.6, pp.57-75, Nov 2001

[11] G. Strangman, D. A. Boas, and J. P. Sutton, "Non-invasive neuroimaging using near-infrared light," Biological Psychiatry, vol. 52, pp. 679-693, 10/1 2002.

[12] A. Siegel, J. J. Marota, and David Boas, "Design and evaluation of a continuous-wave diffuse optical tomography system," Opt. Express **4**, 287-298 (1999)

[13] Lin, Yuanqing; Lech, Gwen; Nioka, Shoko; Intes, Xavier; Chance, Britton; , "Noninvasive, low-noise, fast imaging of blood volume and deoxygenation changes in muscles using light-emitting diode continuous-wave imager," *Review of Scientific Instruments* , vol.73, no.8, pp.3065-3074, Aug 2002

[14] Rajendra Acharya, U., K. Paul Joseph, et al. (2006). "Heart rate variability: a review." Medical and Biological Engineering and Computing 44(12): 1031-1051.

[15] Roche F, Pichot V, Sforza E, Court-Fortune I, Duverney D, Costes F, Garet M, Barthe´le´my J-C (2003) Predicting sleep apnoea syndrome from heart period: a time-frequency wavelet analysis. Eur Respir J 22:937–942

[16] M. M. Ohayon, "Epidemiology of insomnia: what we know and what we still need to learn," *Sleep Medicine Reviews*, vol. 6, pp. 97-111, 2002.

[17] Kenneth Revett; Sergio Tenreiro de Magalhães, "Cognitive Biometrics: Challenges for the Future," *Communications in Computer and Information Science, 2010*, Volume 92, 79-86

[18] Joseph P. Culver; Bradley L. Schlaggar; Hamid Dehghani; Benjamin W. Zeff; , "Diffuse Optical Tomography for Mapping Human Brain Function," *Life Science Systems and Applications Workshop, 2006. IEEE/NLM* , pp.1-2, July 2006

[19] A. Villringer, J. Planck, C. Hock, L. Schleinkofer, and U. Dirnagl, "Near infrared spectroscopy (NIRS): A new tool to study hemodynamic changes during activation of brain function in human adults," *Neurosci. Lett*., vol. 154, pp. 101–104, 1993.

[20] G. Strangman, D.A. Boas, and J.P. Sutton, "Non-invasive neuroimaging using near-infrared light," *Biol. Psych*., vol. 52, no. 7, pp. 679–693, 2002.

[21] U. Beese, H. Langer, W. Lang, and M. Dinkel, "Comparison of near- infrared spectroscopy and somatosensory evoked potentials for the detection of cerebral ischemia during carotid endarterectomy," *Stroke*, vol. 29, pp. 2032–2037, 1998.

[22] K. Sakatani, S. Chen, W. Lichty, H. Zuo, and Y.P. Wang, "Cerebral blood oxygenation changes induced by auditory stimulation in newborn infants measured by near infrared spectroscopy," Early Hum. Dev., vol. 55, pp. 229–236, 1999.

[23] H. Sato, T. Takeuchi, and K.L. Sakai, "Temporal cortex activation during speech recognition: An optical topography study," *Cognition*, vol. 73, pp. B55–B66, 1999.

[24] Y. Hoshi and M. Tamura, "Dynamic multichannel near-infrared optical imaging of human brain activity," *J. Appl. Physiol., vol*. 75, pp. 1842–1846, 1993.

[25] E. Watanabe, A. Maki, F. Kawaguchi, Y. Yamashita, H. Koizumi, and Y. Mayanagi, "Noninvasive cerebral blood volume measurement during seizures using multichannel near infrared spectroscopic topography," J. Biomed. Opt., vol. 5, pp. 287–290, 2000.

[26] Anthony J. Bell, Terrence J. Sejnowski, "A Non-linear Information Maximization Algorithm that Performs Blind Separation"

[27] Hong Bo, Tang Qingyu, Yang Fusheng **"ICA IN THE SINGLE-TRIAL ESTIMATION AND ANALYSIS OF VEP"** Department of Electrical Engineering, Tsinghua University, Beijing 100084

[28] Joseph R. Cavallaro, Franklin T. Luk, CORDIC arithmetic for an SVD processor, Journal of Parallel and Distributed Computing, Volume 5, Issue 3, June 1988, Pages 271-290

[29] 黃煒忠，"四通道即時 EEG 訊號獨立事件分析之 FPGA 實現"，國立交通大學，碩士論文，民國九十七年。

[30] EEGLAB - Open Source MATLAB Toolbox for Electrophysiological Research URL: http://sccn.ucsd.edu/eeglab/

[31] Kuo-Kai Shyu and Ming-Huan Li," *FPGA Implementation of FastICA based on Floating-Point Arithmetic Design for Real-Time Blind Source Separation,*" 2006 International Joint Conference on Neural Networks Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada July 16-21, 2006.

[32] Wei-Chung Huang; Shao-Hang Hung; Jen-Feng Chung; Meng-Hsiu Chang; Lan-Da Van; Chin-Teng Lin,"*FPGA implementation of 4-channel ICA for on-line EEG signal separation,*" Biomedical Circuits and Systems Conference, Page(s):65 – 68,20-22 Nov. 2008

[33] Shmoo Plot – From Wikipedia URL: http://en.wikipedia.org/wiki/Shmoo_plot

[34] T. Thong, J. McNames, and M. Aboy, "Lomb-Wech periodogram for non-uniform sampling," in Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE, 2004, pp. 271-274

[35] P. E. Trahanias, "An approach to QRS complex detection using mathematical morphology," *Biomedical Engineering, IEEE Transactions on,* vol. 40, pp. 201-205, 1993.

[36] MIT/BIH – ECG Database http://ecg.mit.edu/

[37] X. Cheng and D. Boas, "Diffuse optical reflection tomography with continuous-wave illumination," Med. Biol., vol. 42, pp. 841-854, 1997.

[38] Y-H Hsu; C-C Fu; W-C Fang; T-H Sang, "A VLSI-inspired image reconstruction algorithm for continuous-wave diffuse optical tomography systems," Life Science

Systems and Applications Workshop, 2009. LiSSA 2009. IEEE/NIH , pp.88-91, 9-10 April 2009

[39] Nazareth P. Castellanos, Valeri A. Makarov "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis" Journal of Neuroscience Methods 158 (2006) 300–312.