# 國立交通大學

## 電子工程學系 電子研究所碩士班

## 碩士論文

適應性電壓調變應用於離散餘弦轉換

## Adaptive Voltage Scaling for Discrete Cosine Transform

研 究 生:劉仲文

指導教授:黄 威 教授

中華民國九十八年七月

## 適應性電壓調變應用於離散餘弦轉換

## Adaptive Voltage Scaling for Discrete Cosine Transform

研究生:劉仲文

Student : Chun-Wen Liu

指導教授:黃 威教授

Advisor : Prof. Wei Hwang

國 立 交 通 大 學 電 子 工 程 學 系 電 子 研 究 所 碩 士 論 文

#### A Thesis

Submitted to Department of Electronics Engineering & Institute of Electronics College of Electrical Engineering and Computer Science National Chiao Tung University in Partial Fulfillment of the Requirements for the Degree of Master in

**Electronics Engineering** 

#### **July 2008**

#### Hsinchu, Taiwan, Republic of China

中華民國九十八年七月

## 適應性電壓調變應用於離散餘弦轉換

學生:劉仲文

#### 指導教授:黃 威教授

#### 國立交通大學電子工程學系電子研究所碩士班

#### 摘要

適應性電壓調變是最有效用的技術在現今低功率積體電路設計 上。論文中提出一個新的可變電壓產生器。這個電壓產生器可以產生 在 0.8V 到 1.2V 之間的五種電壓。我們發展一個適合此可變電壓產生 器的適應性電壓控制器,將這兩者組合成為一個適應性電壓調變系 統。晶片上的可變電壓產生器在這個系統當中扮演很重要的角色,它 取代了在電壓調整時常用的晶片外的直流/直流轉換器。

離散餘弦定理已經成為數位訊號處理中被廣泛應用的轉換技巧之 一。我們將適應性電壓調變的系統應用在離散餘弦定理的處理器上然 後成功的減少了離散餘弦定理處理器最多 45%的電源消耗。所有的模 擬結果都是利用 TSMS 0.13 μm CMOS 的製程下得到。

## Adaptive Voltage Scaling for Discrete Cosine Transform

Student: Chun-Wen Liu

Advisor: Prof. Wei-Hwang

Department of Electronics Engineering & Institute of Electronics National Chiao-Tung University

## ABSTRACT

In the modern digital IC system, adaptive voltage scaling is the most efficient technology for low power design. A new variable voltage generator (VVG) has been proposed in this paper. Five voltage levels ranged from 0.8V to 1.2V can be generated. An adaptive voltage scaling controller has been developed to fit the VVG to form an adaptive voltage scaling control system. In stead of the off-chip DC-DC converter which is often used in voltage regulation, the on-chip VVG takes an important roll in this system.

Discrete Cosine Transform (DCT) has become one of the widely used transform techniques in digital signal processing. The adaptive voltage scaling system has been applied to DCT and reduces at most 45% power consumption of DCT. All simulations are implemented in TSMC0.13- $\mu$ m CMOS technology.

## Acknowledgements

I would like to express my deepest gratitude to my advisor Prof. Wei Hwang for his enthusiastic guidance and encouragement throughout the research. With his support, I have the confidence and energy to stride forward.

Following, I would like to thank all my friends, Po-Tsang Huang and Wei-Chih Hsieh at LPSOC lab. They gave me much support and discussion on my thesis research.

Finally, I give the greatest respect and love to my family, and I want to express my highest appreciation for their support and understanding.

## Contents

1	Intr	roduction
	1.1	Motivation of the Thesis1
	1.2	Research Goal and Contributions1
	1.3	Thesis Organization2
2	Volt	age Scaling Techniques for Low Power
	2.1	Introduction4
	2.2	Voltage Influence on Power4
	2.3	Voltage Scaling Influence in Delay5
	2.4	Voltage Scaling Techniques6
		2.4.1 Multiple Supply Voltage
		2.4.2 Clustered Voltage Scaling (CVS)7
		2.4.3 Multiple Threshold Voltage
		2.4.4 Adaptive Voltage Scaling (AVS)10
		2.4.5 Adaptive Body Bias (ABB)11
	2.5	Dynamic Voltage Scaling12
		2.5.1 Essential Components12
		2.5.2 Improving Energy Efficiency13
		2.5.3 Fundamental Trade off14
	2.6	Conversion Efficiency14
		2.6.1 Limits to Reducing Cdd15
	2.7	Design Constraints Over Voltage17
		2.7.1 Circuit Design Constraints17
		2.7.2 Circuit Delay Variation18
		2.7.3 Noise Margin Variation19
		2.7.4 Delay Sensitivity22
3	Ada	ptive Voltage Scaling
	3.1	Components of Adaptive Voltage Regulation24
	3.2	Voltage Converter25
		3.2.1 Pulsed Width Modulation (PWM) Operation26
		3.2.2 Buck Converter
		3.2.3 Boost Converter
		3.2.4 Buck-Boost Converter
	3.3	Reference Voltage Generator
		3.3.1 Traditional Reference Voltage Generator
		3.3.2 Modified Variable Voltage Generator
	3.4	DC-DC Voltage Converter32

	3.4.1 Output Buffer	
	3.5 Simulation Result	35
	3.6 PLL-based Adaptive Voltage Regulation Using FSM	37
	3.6.1 Reference circuit	
	3.6.2 Finite State Machine (FSM)	
4	Discrete Cosine Transform	
	4.1 Introduction to Discrete Cosine Transform	42
	4.2 Alternative Implementation	44
	4.2.1 Multiplier Implementation	45
	4.2.2 Pure ROM Implementation	46
	4.2.3 Mixed ROM Implementation	47
	4.3 Reducing Power of DCT	47
	4.3.1 Reducing Power Through Pipelining	47
	4.3.2 Reducing Power Through Parallelism	49
	4.3.3 Reducing Power Through Reducing Complexity	50
	4.4 Reconfigurable Architecture	52
	4.4.1 Computational Sharing Multiplier Algorithm (CSHM)	53
	4.4.2 DCT Coefficients for 4-bit Decomposition	55
	4.4.3 DCT Coefficients for 2-bit Decomposition	56
	4.4.4 DCT Architecture Based on CSHM Algorithm	58
	4.4.5 Modified DCT Coefficients	61
	4.5 Simulation Result	62
5	Adaptive Voltage Scaling for Discrete Cosine Transform	
	5.1 The Proposed Architecture	64
	5.2 Reference Circuit	65
	5.2.1 Ring Oscillator	66
	5.2.2 Frequency Detector	67
	5.3 Controller	69
	5.3.1 Control Logic	70
	5.3.2 Selector	73
	5.4 Simulation Result	77
	5.4.1 Adaptive Voltage Scaling System	77
	5.4.2 AVS for Discrete Cosine Transform	79
6	Conclusion & Future Work	81

## **List of Figures**

-	
<b>'</b>	
1.	
_	

3

21		
<i>4</i> •1	Multiple Supply Voltage	6
2.2	Level Converter	7
2.3	Clustered Voltage Scaling	7
2.4	Conventional Level Converting Flip-Flop	8
2.5	V <sub>dd</sub> vs. V <sub>t</sub> for a Fixed Delay	9
2.6	Multiple Threshold Voltage	9
2.7	Frequency vs. Power of AVS and ABB	10
2.8	Scheme of the Conventional AVS System	11
2.9	Adaptive Body Biasing	12
2.10	Energy Efficiency Improvement	13
2.11	Energy Loss Due to Voltage Supply Ripple	16
2.12	Relative CMOS Circuit Delay Variation over Supply Voltage	18
2.13	Noise Margin Degradation	20
2.14	Noise Margin vs. Supply Voltage	21
2.15	Normalized Noise Margin Reduction due to Supply Bounce	21
2.16	Normalized Delay Sensitivity vs. Supply Voltage	23
	1896	
3.1	Adaptive Voltage Regulation	24
3.1 3.2	Adaptive Voltage Regulation	24 26
3.1 3.2	Adaptive Voltage Regulation(a) Switching Mode Buck Converter(b) Waveform of the Switching Mode Buck Converter	24 26 27
<ul><li>3.1</li><li>3.2</li><li>3.3</li></ul>	Adaptive Voltage Regulation(a) Switching Mode Buck Converter(b) Waveform of the Switching Mode Buck ConverterBuck Converter	24 26 27 27
<ul><li>3.1</li><li>3.2</li><li>3.3</li><li>3.4</li></ul>	Adaptive Voltage Regulation(a) Switching Mode Buck Converter(b) Waveform of the Switching Mode Buck ConverterBuck ConverterBoost Converter	24 26 27 27 27
<ul> <li>3.1</li> <li>3.2</li> <li>3.3</li> <li>3.4</li> <li>3.5</li> </ul>	Adaptive Voltage Regulation	24 26 27 27 28 29
<ul> <li>3.1</li> <li>3.2</li> <li>3.3</li> <li>3.4</li> <li>3.5</li> <li>3.6</li> </ul>	Adaptive Voltage Regulation	24 26 27 27 28 28 29 30
<ul> <li>3.1</li> <li>3.2</li> <li>3.3</li> <li>3.4</li> <li>3.5</li> <li>3.6</li> <li>3.7</li> </ul>	Adaptive Voltage Regulation(a) Switching Mode Buck Converter(b) Waveform of the Switching Mode Buck ConverterBuck ConverterBoost ConverterBuck-Boost ConverterTraditional Reference Voltage GeneratorReference Voltage Generator Circuit	24 26 27 27 28 29 30 31
<ul> <li>3.1</li> <li>3.2</li> <li>3.3</li> <li>3.4</li> <li>3.5</li> <li>3.6</li> <li>3.7</li> <li>3.8</li> </ul>	Adaptive Voltage Regulation.(a) Switching Mode Buck Converter.(b) Waveform of the Switching Mode Buck Converter.Buck Converter.Boost Converter.Buck-Boost Converter.Traditional Reference Voltage Generator.Reference Voltage Generator Circuit.Modified Variable Voltage Generator.	24 26 27 27 28 29 30 31 32
<ul> <li>3.1</li> <li>3.2</li> <li>3.3</li> <li>3.4</li> <li>3.5</li> <li>3.6</li> <li>3.7</li> <li>3.8</li> <li>3.9</li> </ul>	Adaptive Voltage Regulation.(a) Switching Mode Buck Converter.(b) Waveform of the Switching Mode Buck Converter.Buck Converter.Boost Converter.Buck-Boost Converter.Traditional Reference Voltage Generator.Reference Voltage Generator Circuit.Modified Variable Voltage Generator.Architecture of On-chip Voltage Converter.	24 26 27 27 28 29 30 31 32 34
3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10	Adaptive Voltage Regulation.(a) Switching Mode Buck Converter.(b) Waveform of the Switching Mode Buck Converter.Buck Converter.Boost Converter.Buck-Boost Converter.Traditional Reference Voltage Generator.Reference Voltage Generator.Reference Voltage Generator.Architecture of On-chip Voltage Converter.Differential Based Output Buffer.	24 26 27 27 28 29 30 31 32 34 35
3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11	Adaptive Voltage Regulation.(a) Switching Mode Buck Converter.(b) Waveform of the Switching Mode Buck Converter.Buck Converter.Boost Converter.Buck-Boost Converter.Traditional Reference Voltage Generator.Reference Voltage Generator Circuit.Modified Variable Voltage Generator.Architecture of On-chip Voltage Converter.Differential Based Output Buffer.The Proposed On-chip Voltage Converter Circuit.	24 26 27 27 28 29 30 31 32 34 35 36
3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11 3.12	Adaptive Voltage Regulation.(a) Switching Mode Buck Converter.(b) Waveform of the Switching Mode Buck Converter.Buck Converter.Boost Converter.Buck-Boost Converter.Traditional Reference Voltage Generator.Reference Voltage Generator Circuit.Modified Variable Voltage Generator.Architecture of On-chip Voltage Converter.Differential Based Output Buffer.The Proposed On-chip Voltage Regulator Using FSM.	24 26 27 27 28 29 30 31 32 34 35 36 37
3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11 3.12 3.13	Adaptive Voltage Regulation (a) Switching Mode Buck Converter (b) Waveform of the Switching Mode Buck Converter Buck Converter Boost Converter Buck-Boost Converter Traditional Reference Voltage Generator Reference Voltage Generator Circuit Modified Variable Voltage Generator Architecture of On-chip Voltage Converter Differential Based Output Buffer The Proposed On-chip Voltage Converter Circuit PLL Based Adaptive Voltage Regulator Using FSM Reference Circuit.	24 26 27 27 28 29 30 31 32 34 35 36 37 38
3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11 3.12 3.13 3.14	Adaptive Voltage Regulation (a) Switching Mode Buck Converter	24 26 27 27 28 29 30 31 32 34 35 36 37 38 39
3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11 3.12 3.13 3.14 3.15	Adaptive Voltage Regulation.(a) Switching Mode Buck Converter.(b) Waveform of the Switching Mode Buck Converter.Buck Converter.Boost Converter.Boost Converter.Traditional Reference Voltage Generator.Reference Voltage Generator Circuit.Modified Variable Voltage Generator.Architecture of On-chip Voltage Converter.Differential Based Output Buffer.The Proposed On-chip Voltage Converter Circuit.PLL Based Adaptive Voltage Regulator Using FSM.Reference Circuit.FSM.Enable Generator.	24 26 27 27 28 29 30 31 32 34 35 36 37 38 39 40

4	
-	

5

4.1	Multiplier Implementation	.45
4.2	Architecture of DA	46
4.3	Pure ROM Implementation	.46
4.4	Pure ROM Implementation	.47
4.5	Pipelined DCT Architecture	.48
4.6	Original Data Path	49
4.7	Two Parallel Data Paths	.49
4.8	Changing the Least Sensitive Non Zero Digits to Zeros	50
4.9	CSHM Architecture for 4-bit Decomposition	54
4.10	CSHM Architecture for $c_1 \cdot x$	55
4.11	CSHM Architecture for 2-bit Decomposition of DCT Coefficient	.58
4.12	<b>Architecture of Computing</b> [ <i>z</i> <sub>0</sub> , <i>z</i> <sub>2</sub> , <i>z</i> <sub>4</sub> , <i>z</i> <sub>8</sub> ]	59
4.13	<b>Architecture of Computing</b> [ <i>z</i> <sub>1</sub> , <i>z</i> <sub>3</sub> , <i>z</i> <sub>5</sub> , <i>z</i> <sub>7</sub> ]	60
4.14	Pipelined DCT Architecture	.62
5.1	Architecture of AVS for Discrete Cosine Transform	.64
5.2	Reference Circuit	.66
5.3	Ring Oscillator	.66
5.4	Traditional Frequency Detector	.67
5.5	Frequency Detector	.68
5.6	Waveform of the Frequency detector	68
5.7	Waveform of the Frequency detector	69
5.8	The Proposed Controller	69
5.9	Propose Control Logic	.71
5.10	(a) Multiple Input D Flip-Flop (b) Waveform	.72
5.11	Waveform of pre_sel[4:0]	.73
5.12	Proposed Selector	74
5.13	(a) Waveform of the Selector	.75
	(b) Waveform of the Selector	75
	(c) Waveform of the Selector	.76
5.14	Architecture of the AVS with Voltage Generators	.77
5.15	(a) The Waveform of the adaptive voltage control system	.77
	(b) The Waveform of the adaptive voltage control system	.78
5.16	Architecture of AVS for Discrete Cosine Transform	.79

## **List of Tables**

## 3

	31	Temperature Variation	36
	2.2	Simulation Degult of the Modified Veriable Voltage Constants	27
	3.2	Simulation Result of the Mounteel variable voltage Generator	51
	3.3	Power Efficiency Simulation Results	38
4			
	4.1	8 bits DCT coefficients	43
	4.2	DCT Coefficients Represented by CSD	51
	4.3	PSNR	52
	4.4	8-bit DCT Coefficients and the Alphabets	55
	4.5	8-bit DCT Coefficients and the Alphabets for 2-bit Decomposition	.57
	4.6	Type1 Modified Coefficients	61
	4.7	Type2 Modified Coefficients	62
	4.8	Simulation Result of the Pipelined DCT	.63
5		AND DECK	
	5.1	Value of the Flip-Flops	.72
	5.2	Simulation Result of the AVS System	.78
	5.3	Original DCT	.79
	5.4	Adaptive Voltage controlled DCT	80
	5.5	Original vs. Adaptive	80
6		Contraction of the second s	
	6.1	Fully Digital Power Management System	82

## Chapter 1 Introduction

## 1.1 Motivation of the Thesis

As technology moves into deep submicron feature sizes, power dissipation due to leakage current is increasing at an amazing rate. Supply voltage has not been scaled substantially enough to keep power per unit area constant over technology generations. It's a main trend to integrate computer, communication and consumer electronic (3C). It's emergent to increase battery life and make chips consume as less energy as possible. For the future integrated-circuit (IC) and System-on-Chip (SoC) designs, the need for low-power, high-performance is further prompted by the growing demand for portable devices such as cellular phones, laptops and PDA's. For such portable devices, power consumption is paramount, and performance must somehow be maintained while decreasing power and hence increasing battery life.

Historically, ICs have been designed with a single supply voltage and a single threshold voltage. Process scaling was the primary mechanism by which the exponential growth in integration and performance was realized. While this scaling allowed enormous gains in operating frequencies, transistor count, performance, power consumption and reliability issues forced the supply voltage to be scaled with decreasing feature size. This in turn required threshold voltage scaling in order to maintain performance. This has a dramatic effect on leakage current, as subthreshold current increases exponentially with reduced threshold voltage.

#### **1.2 Research Goals and Contributions**

The goal of this research is to design and implement a on-chip variable voltage generator for low-power and low-voltage applications. This includes the development

of the circuit-level design technique to increase the usefulness of the variable voltage generator in any portable electronic application.

The key contributions of this thesis are listed as follow :

1 A new variable voltage generator which uses the parallel-connected transistors operating has been proposed. It can generate five voltage levels ranged from 0.8V to 1.2V.

2 An adaptive voltage scaling system which is based on the variable voltage generator has been developed. The controller of the system adaptively controls the variable voltage generator to provide the voltage level which is the fittest to the expected performance to the application.

3 The adaptive voltage control system has been applied to the discrete cosine transform processor to reduce the power consumption of it. It successfully reduces at most 45% power consumption of DCT, and only at most 28% power overhead.

### **1.3 Thesis Organization**

The rest of the thesis is organized as follows : the principles of the dynamic voltage scaling design and overview of the voltage scaling techniques in Chapter 2. The reason why the voltage scaling is needed is described in the beginning. The background for voltage scaling is also mentioned in this chapter.

Adaptive voltage scaling is a main stream in recent year. We will introduce three topologies of the switching mode DC/DC converter and linear mode design concept in Chapter 3. We will focus on the on-chip voltage converter. The new variable voltage generator which can provide more than five voltage levels is proposed in this chapter.

The adaptive voltage scaling system has been applied to the discrete cosine transform unit. DCT has become a widely used transform technique in digital signal processing. Sort of the various implementations are described in Chapter 4, we modified one of them which is called Computational Sharing Multiplier Algorithm (CSHM) to implement our DCT unit.

In Chapter 5, the adaptive voltage scaling system would be developed based on the proposed circuit. The system adaptively controls the variable voltage generators and provides the fittest one to the application. And we proposed the adaptive voltage controller which makes the variable voltage generator circuit adaptive to the system frequency. It controls the voltage level to the lowest one which still meets the expected performance. The expected performance is predicted by the reference circuit which is mainly constructed by a fast-lock frequency detector and a ring-oscillator. Finally, we would make the conclusion and the future work about this thesis in Chapter 6.



## Chapter 2 Voltage Scaling Techniques for Low Power

## **2.1 Introduction**

Nowadays, there are more and more requirements for the portable digital products, for example, smaller size, longer run-time and more functional abilities. All these requirements have something to do with power or energy. Therefore, low power issues are more and more important for every product, we need ultra-low power hardware to maximize run-time and to achieve more functional abilities.

## 2.2 Voltage Influence on Power

In CMOS circuits, the average power consumption is defined as follow[1]:

$$P_{av} = P_{dyn} + P_{shor t} + P_{leakage} + P_{static}$$
(2.1)

where  $P_{dyn}$  stands for dynamic power consumption,  $P_{short}$  stands for short-circuit power consumption,  $P_{leakage}$  stands for leakage power consumption, and  $P_{static}$  stands for static power consumption.

In equation (2.1), dynamic power consumption is the dominant component of power dissipation in CMOS circuit[2], which is defined as follow:

$$P_{dyn} = \alpha C f V_{dd}^{2}$$
(2.2)

where  $\alpha$  is the activity factor, f is the switching frequency, C is the effective capacitance fully charged and discharged over voltage swing  $V_{dd}$ , and  $V_{dd}$  is the supply voltage. From equation (2.2), it is clear that the reduction of the supply voltage is an effective way of saving power dissipation since the supply voltage has a quadratic relationship to the dynamic power consumption. Therefore, voltage scaling techniques performs the reduction of supply voltage. Dynamic power may be significantly reduced by scaling down the supply voltage, yet reducing the supply voltage increases the execution delay. We need to minimize the supply voltage for a target performance, which is the goal for all the voltage scaling techniques.

### 2.3 Voltage Scaling Influence on Delay

The energy dissipation per switching event of a properly designed digital CMOS circuit is dominated by the dynamic component. It is clear that a reduction of the power supply voltage yields a quadratic savings in energy dissipation per computational event. However, this comes at the expense of computational throughput as the propagation delay of a digital CMOS gate increases with decreasing  $V_{dd}$ .

Since gate delay increases with decreasing  $V_{dd}$  as indicated in equation(2.3), globally lowering  $V_{dd}$  degrades the overall circuit performance[4].

$$t_d \propto \frac{V_{dd}}{\left(V_{dd} - V_t\right)^{\alpha}} \tag{2.3}$$

where  $V_t$  is the threshold voltage and  $\alpha$  is the velocity saturation parameter. Therefore, it is a trade off between reducing supply voltage ( $V_{dd}$ ) to save power and the best performance[4]. However, if the highest performance is not required, global voltage scaling techniques is applicable. If the power supply provides two fixed voltages, the nominal voltage and the lower voltage, and the delay constraints can not be met using only the lower voltage, multiple supply voltage subject to relaxed delay constraints is the only option[4]. However, if the lower voltage is sufficient to meet the delay constraints, the related delay at the lower voltage compared with the delay at the nominal voltage can be derived from equation (2.3), and is as follow:

$$\frac{t_d(V_{ddl})}{t_d(V_{dd})} = \frac{V_{ddl}}{V_{dd}} \bullet \left(\frac{V_{dd} - V_t}{V_{ddl} - V_t}\right)^a$$
(2.4)

## 2.4 Voltage Scaling Techniques

## 2.4.1 Multiple Supply Voltages

Multiple supply voltage is a voltage scaling approach, which reduces the power consumption while still meeting the timing constraints. Its concept is to operate the speed-critical parts of the circuit at the higher voltage, and parts which are not speed-critical at the lower voltage. This results in the slower speed of the non-critical parts, yet it has no influence on the critical parts. The whole circuit still works on the same performance, but some power is reduced since the voltage on the non-critical parts is lowered. Level converters, illustrated on Figure 2.2, are needed when the signal is propagating from the gates operated on  $V_{ddl}$  to the gates operated on  $V_{ddh}$ , since the high output of the low voltage-gate cannot fully turn off the pmos part of the high voltage-gate which could cause a DC leakage path to increase the power consumption. In Figure 2.1, there are four points (A, B, C, D) inserted level converters.



Figure 2.1 Multiple Supply Voltage[5]



Figure 2.2 Level Converter

## 2.4.2 Clustered Voltage Scaling(CVS)



Figure 2.3 Clustered Voltage Scaling[5]

An advanced scheme of multiple supply voltage, often called clustered voltage scaling[5], is illustrated in Figure 2.3. In Figure 2.3, the gates on the critical path are operated on  $V_{ddh}$  while the gates off the critical path are operated on  $V_{ddl}$ . Since level conversion is required whenever an output from a low  $V_{dd}$  ( $V_{ddl}$ ) gate has to drive an input to a high  $V_{dd}$  ( $V_{ddh}$ ) gate, in order to reduce the overhead of the level converters, clustered voltage scaling which critical and non-critical paths of the design are clustered has been developed. The low  $V_{dd}$  ( $V_{ddl}$ ) clusters are followed by pipeline flip-flops and level conversion is merged into the flip-flops. These flip-flops are called level converting flip-flops(LCFF), which is illustrated in Figure 2.4.



Figure 2.4 Conventional Level Converting Flip-Flop

## 2.4.3 Multiple Threshold Voltage

Since the energy per computational event ideally scales as  $V_{dd}^2$  while circuit speed is related to  $(V_{dd} - V_t)$  rather than, lower power dissipation can be achieved without compromise of throughput approximately scaling device threshold voltages,  $V_t$ , together with the voltage supply,  $V_{dd}$ . It can be shown that a circuit running at a supply voltage of  $V_{dd} = 1.5V$  with  $V_t = 1.0V$  will have nearly identical performance to the same circuit running at  $V_{dd} = 0.9V$  with  $V_t = 0.5V[6]$ . However, the circuit running at  $V_{dd} = 0.9V$  will consume about one third the power. Voltage scaling with threshold voltage reduction is limited primarily by subthreshold leakage currents in the lower



threshold devices, which increase exponentially with decreasing  $V_t$ . Figure 2.5 shows the relationship that  $V_{dd}$  versus  $V_t$  while keeping the performance constant.

Figure 2.5  $V_{dd}$  vs.  $V_t$  for a Fixed Delay

Multiple threshold voltage can dynamically control leakage currents. Figure2.6 shows a conventional approach to implement the multiple threshold voltage. Low-  $V_t$  transistors are used to implement the logic circuits, while high-  $V_t$  transistors are used as switches between the main circuit to  $V_{dd}$  and between the main circuit to ground. The high- $V_t$  transistors are cut-off during sleep (standby) periods, since the subthreshold leakage can be largely reduced. During active periods, the high- $V_t$  transistors are turned on and the circuits works normally.



Figure 2.6 Multiple Threshold Voltage

## 2.4.4 Adaptive Voltage Scaling (AVS)

In recent years, adaptively control the voltage or the threshold voltage has been emphasized, and gaining more and more attention. Figure2.7 shows the relationship of frequency vs. power of both adaptive voltage scaling (AVS) and adaptive body bias (ABB)[7]. Adaptively control the voltage means that the AVS system can locate the optimal voltage for the operating frequency. Whenever the frequency is changed, it sweeps the power supply for the optimal voltage of the new frequency. We can see in Figure2.7, the AVS operation sweeps the power supply at a reference frequency. Since the frequency indicates the performance which the designer wants, as long as it is not the highest performance of the design, the AVS system can locate a lower optimal voltage for the operating frequency. The difference between full swing voltage and the lower optimal voltage is the power which is saved by the AVS system.



Figure 2.7 Frequency vs. Power of AVS and ABB[7]

The scheme of the conventional AVS system is shown in Figure 2.8. It is consists of a dc-dc converter, which a buck converter is used in the figure, a reference circuit, and a digital controller[8]. The reference circuit indicates the highest frequency at the regulated voltage, which is labeled as V in Figure2.8. the controller compare the reference frequency and the frequency output from the reference circuit, which is labeled as f in Figure2.8, and send the error along with the decision to the dc-dc converter. The buck converter is often used for a good efficiency dc-dc conversion, it converts the voltage to the voltage level that the controller indicates, and then applies the regulated voltage to the digital system.



Figure 2.8 Scheme of the Conventional AVS System[8]

### 2.4.5 Adaptive Body Bias (ABB)

Adaptive body biasing, which is shown in Figure 2.9, changes the threshold voltage dynamically by changing the substrate bias. It uses a lower threshold voltage to operate at low power supply voltage during active periods, and raise the threshold voltage during idle periods. Although the purpose is the same as using multiple threshold voltage, adaptive body biasing controls much better. The limitation of ABB

is that the threshold voltage changes in a square root with respect to source to bulk voltage and therefore a large voltage is required to change  $V_t$ , which also comes along with increased parasitic capacitance.



Figure 2.9 Adaptive Body Biasing[9]

## 2.5 Dynamic Voltage Scaling

### **2.5.1 Essential Components**

The dynamic voltage regulator consists of a detector , a loop filter, and a dc-dc converter. The frequency detector generates a digital error signal in proportion to the frequency error. This error is translated into an update signal for the dc-dc converter through the loop filter. The dc-dc converter provides the supply voltage  $V_{dd}$ , regulating against changes in battery voltage and the supply voltage load current,  $I_{dd}$ .

The voltage-controlled oscillator (VCO) is intergrated together with the circuit, and designed to match its critical path. The loop forces the output frequency of the VCO to equal the commanded frequency, at an input voltage  $V_{dd}$ . The circuit is therefore run at the minimum supply voltage, at which the state request can be met, resulting in the lowest achievable energy per operation while sustaining  $V_{dd}$ .



### 2.5.2 Improving Energy Efficiency

Figure 2.10 Energy Efficiency Improvement

The possible energy efficiency improvement of DVS is illustrated in Figure2.10. Starting at the nominal  $V_{dd}$ , when the clock frequency  $f_{clk}$ , is reduced, there is a proportional decrease in throughput. When this is at constant  $V_{dd}$ , there is no reduction in energy. However, if it is scaled lock-step with  $f_{clk}$ , then the lower curve is traversed, yielding more than a 10x energy reduction at low voltage.

### 2.5.3 Fundamental Trade-off

The digital circuits generally operate at a fixed voltage, and require a regulator to control the supply voltage variation. Sometimes the digital circuit produces large current for which the regulator's output capacitor supplies the charge. Hence, a large output capacitor on the regulator is desirable to minimize the ripple on  $V_{dd}$ . A large capacitor also helps to maximize the regulator conversion efficiency by reducing the voltage variation at the output of the regulator. However, the voltage converter required for DVS is different from a standard voltage regulator because in addition to regulating voltage for a given clock frequency, it must also change the operating voltage when a new clock frequency is request. To minimize the speed and energy consumption of this voltage transition, a small output capacitor on the converter is desirable, in contrast to the supply ripple requirement. Thus, the fundamental trade-off in a DVS system is between good voltage regulation and efficiency dynamic voltage conversion. It is possible to optimize the size of the output capacitor to balance the requirements for good voltage regulation with the requirements for a good dynamic The second voltage conversion.

## **2.6 Conversion Efficiency**

The efficiency of a voltage regulator is defined as:

$$\eta = \frac{\text{Power Delivered to Load}}{\text{Total Power Dissipation}}$$
(2.5)

The buck converter is very efficient at voltage conversion, with efficiencies typically in the 90-95% range[10]. It can be designed methodically for a fixed operating voltage. The converter designed for a large range of voltage and current loads is difficult. Several techniques have been developed for the converter loop design to improve the efficiency over this broad range of operating conditions[10]. In addition to the supply ripple and conversion efficiency performance metrics of a standard voltage regulator, the DVS converter introduces two new performance metrics: transition time and transition energy/ For a large voltage change ( $V_{dd1} \rightarrow V_{dd2}$ ), the transition time is :

$$t_{TRAN} \gg \frac{2 \times C_{dd}}{I_{MAX}} \times \left| V_{dd 2} - V_{dd 1} \right|$$
(2.6)

where  $I_{MAX}$  is the maximum output current of the converter, and the factor of 2 exists because the current is pulsed in a triangular waveform. In practice,  $t_{TRAN}$  will be slightly longer for a low-to-high voltage transition because the actual current changing  $C_{dd}$  is  $I_{MAX} - I_{dd} * V_{dd}$ . The energy consumed during this transition is:

$$E_{TRAN} = h \times C_{dd} \times \left| V_{dd 2}^{2} - V_{dd 1}^{2} \right|$$
 (2.7)

Since both transition time and transition energy are proportional to  $C_{dd}$ , minimizing  $C_{dd}$ , yields a faster and more energy-efficient voltage converter.

# 2.6.1 Limits to Reducing Cdd 1896

Decreasing  $C_{dd}$  reduces transition time, and by doing so increases the speed at which the voltage changes,  $dV_{dd}/dt$ . But decreasing  $C_{dd}$  increases supply ripple, which in turn increases circuit energy consumption as shown in Figure 2.11. The increase is moderate at high  $V_{dd}$ , but begins to increase as  $V_{dd}$  approaches VT because the negative ripple slows down the circuit so much that most of the computation is performed during the positive ripple, which decreases energy efficiency. For values of supply ripple above 10%, the processor can still operate properly, but the increased energy consumption of the processor outweighs the decreased transition energy consumption, degrading overall system energy-efficiency.



Figure 2.11 Energy Loss Due to Voltage Supply Ripple

Loop stability is another limitation on reducing capacitance. As  $C_{dd}$  is reduced the pole frequency increases. As the pole approaches the sampling frequency, interaction with higher-order poles will eventually make the system unstable.

The third limitation is that low-voltage conversion efficiency scales down with  $C_{dd}$ . Since the DVS processor will ideally be operating most of the time at low voltage, it is important to maintain reasonable low-voltage conversion efficiency.

Increasing the converter sampling frequency will reduce the supply ripple and increase the pole frequency due to the sample delay. Thus, these two limits are not fixed, but can be varied. However, increasing the sampling frequency has two negative side-effects. First, low-load converter efficiency will decrease because the converter loop will need to be activated more frequently to maintain the same voltage. Second, the  $f_{CLK}$  quantization error will increase. These side-effects may be mitigated with a variable sampling frequency that adapts to the system power requirements

The maximum  $dV_{dd}/dt$  at which the circuits will still operate properly is a

hard constraint because system failure can be induced, but occurs for a much smaller  $C_{dd}$  than the supply ripple and stability constraints. Low-voltage conversion efficiency is a soft-constraint, but cannot be improved by adjusting the converter sampling frequency.

## 2.7 Design Constraints Over Voltage

A typical circuit targets a fixed supply voltage, and is designed for  $\pm 10\%$  maximum voltage variation. In contrast, a DVS circuit must be designed to operate over a much wider range of supply voltages, which impacts both design implementation and verification time.

## 2.7.1 Circuit Design Constraints

To realize the full range of DVS energy efficiency, only circuits that can operate all the way down to  $V_{Th}$  should be used. NMOS pass gates are often used in low-power design due to their small area and input capacitance. However, they are limited by not being able to pass a voltage greater than  $V_{dd} - V_{thn}$ , such that a minimum  $V_{dd}$  of  $2 \cdot V_{Th}$  is required for proper operation. Since throughput and energy consumption vary by 4x over the voltage range  $V_{Th}$  to  $2 \cdot V_{Th}$ , using NMOS pass gates restricts the range of operation by a significant amount, and are not worth the moderate improvement in energy efficiency. Instead, CMOS pass gates, or an alternate logic style, should be utilized to realize the full voltage range of DVS. As previously demonstrated in Figure 3.1, the delay of CMOS circuits track over voltage such that functional verification is only required at one operating voltage. The one possible exception is any self-timed circuit, which is a common technique to reduce energy consumption in memory arrays. If the self-timed path layout exactly mimics that of the circuit delay path as was done in the prototype design, then the paths will scale similarly with voltage and eliminate the need to functionally verify over the entire range of operating voltages.

### 2.7.2 Circuit Delay Variation

While circuit delay tracks well over voltage, subtle delay variations exist and do impact circuit timing. To demonstrate this, three chains of inverters were simulated whose loads were dominated by gate, interconnect, and diffusion capacitance respectively. To model paths dominated by stacked devices, a fourth chain was simulated consisting of 4 PMOS and 4 NMOS transistors in series. The relative delay variation of these circuits is shown in Figure 2.12 for which the baseline reference is an inverter chain with a balanced load capacitance similar to the ring oscillator.



Figure 2.12 Relative CMOS Circuit Delay Variation over Supply Voltage

The relative delay of all four circuits is a maximum at only the lowest or highest operating voltages. This is true even including the effect of the interconnect's RC delay. Since the gate dominant curve is convex, combining it with one or more of the other effects' curves may lead to a relative delay maximum somewhere between the two voltage extremes. However, all the other curves are concave and roughly mirror the gate dominant curve such that this maximum will be less than a few percent higher than at either the lowest or highest voltage, and therefore insignificant. Thus, timing analysis is only required at the two voltage extremes, and not at all the intermediate voltage values.

As demonstrated by the series dominant curve, the relative delay of four stacked devices rapidly increases at low voltage. Additional devices in series will lead to an even greater increase in relative delay. As supply voltage increases, the drain-to-source voltage increases for the stacked devices during an output transition. For the devices whose sources are not connected to  $V_{dd}$  or ground, their body-effect increases with supply voltage, such that it would be expected that the relative delay would be a maximum at high voltage. However, the sensitivity of device current and circuit delay to gate-to-source voltage exponentially increases as supply voltage goes down. So even though the magnitude change in gate-to-source voltage during an output transition scales with supply voltage, the exponential increase in sensitivity dominates such that stacked devices have maximum relative delay at the lowest voltage. Thus, to improve the tracking of circuit delay over voltage, a general design guideline is to limit the number of stacked devices, which was four in the case of the prototype design. One exception to the rule is for circuits in non-critical paths, which can tolerate a broader variation in relative delay. Another exception is for circuits whose alternative design would be significantly more expensive in area and/or power (e.g. memory address decoder), but the circuits must still be designed to meet timing constraints at low voltage.

## 2.7.3 Noise Margin Variation

Figure 2.13 demonstrates the two primary ways that noise margin is degraded. The first is capacitive coupling between an aggressor signal wire that is switching and an adjacent victim wire. When the aggressor and victim signals have the same logic level, and the aggressor transitions between logic states, the victim signal can also incur a voltage change. If this change is greater than the noise margin,

the victim signal will glitch and potentially lead to functional failure. Supply bounce is induced by switching current spikes on the power distribution network, which has resistive and inductive losses. If the gate's output signal is the same voltage as the supply that is bouncing, the voltage spike transfers directly to the output signal. Again, if this voltage spike is greater than the noise margin, glitch, and potentially functional failure, will occur.



For the case of capacitive coupling, the amplitude of the voltage spike on the victim signal is proportional to  $V_{dd}$  to first order. As such, the important parameter to analyze is noise margin divided by  $V_{dd}$  to normalize out the dependence on  $V_{dd}$ . Figure 2.11 plots two common measures of noise margin vs.  $V_{dd}$ , the noise margin of a standard CMOS inverter, and a more pessimistic measure of noise margin,  $V_{Th}$ . The relative noise margin is a minimum at high voltage, such that signal integrity analysis to ensure there is no glitch only needs to consider a single value of  $V_{dd}$ . If a circuit passes signal integrity analysis at maximum  $V_{dd}$ , it is guaranteed to pass at all other values of  $V_{dd}$ .



Figure 2.14 Noise Margin vs. Supply Voltage

Supply bounce occurs through resistive (IR) and inductive (dI/dt) voltage drop on the power distribution network both on chip and through the package pins. Figure 2.15 plots the relative normalized IR and dI/dt voltage drop as a function of  $V_{dd}$ . It is interesting to note that the worst case condition occurs at high voltage, and not at low voltage, since the decrease in current and dI/dt more than offsets the reduced voltage swing. Given a maximum tolerable noise margin reduction, only one operating voltage needs to be considered, which is maximum  $V_{dd}$ . to determine the maximum allowed resistance (R) and inductance (L). The global power grid and package must then be designed to meet these constraints on resistance and inductance.



Figure 2.15 Normalized Noise Margin Reduction due to Supply Bounce

## 2.7.4 Delay Sensitivity

Supply bounce has another adverse affect on circuit performance in that it can induce timing violations. Supply bounce decreases a transistor's gate drive, which in turn increases the circuit delay. If this increase occurs within a critical path, a timing violation may result leading to functional failure. A typical microprocessor uses a phase-locked loop to generate a clock frequency which is locked to an external reference frequency and independent of on chip voltage variation. As such, both global and local voltage variation can lead to timing violations if the voltage drops a sufficient amount to increase the critical paths' delay past the clock cycle time. However, in the DVS system, the clock signal is derived from a ring oscillator whose output frequency is strictly a function of  $V_{dd}$ , and not an external reference. As such, global voltage variations not only slow down the critical paths, but the clock frequency as well such that the processor will continue operating properly. Localized supply variation, however, may only effects the critical paths, and not the ring oscillator. These can lead to timing violations if the local supply drop is sufficiently large. As such, careful attention has to be paid to the local supply routing. For the prototype design, a design margin of 5% was included in the timing verification to allow for localized voltage drops. Delay sensitivity is the relative change in delay given a drop in  $V_{dd}$ , and can be calculated as:

$$\frac{\partial Delay}{Delay}(Vdd) = \frac{\partial Delay}{\partial Vdd} \cdot \lim_{\Delta Vdd \to 0} \left(\frac{\Delta Vdd}{Delay(Vdd)}\right)$$
(2.8)

This equation can be analytically quantified using Equation 2.8, and the normalized delay sensitivity is plotted as a function of  $V_{dd}$  in Figure 2.13. For sub-micron CMOS processes, the delay sensitivity peaks at approximately  $2 \cdot V_{Th}$ . Thus, the design of the local power grid only needs to consider one value of  $V_{dd}$ ,  $2 \cdot V_{Th}$ , to ensure that the resistance/inductance voltage drop meets the design margin on delay variation. If the power grid meets timing constraints at this value of  $V_{dd}$ , it is

guaranteed to do so at all other voltages.



## **Chapter 3** Adaptive Voltage Scaling

Adaptive supply voltage regulation reduces power and energy consumption by lowering the supply voltage to the minimum which is required to support the operating frequency. On the other hand, it maximizes the energy efficiency of the circuits. Whenever the maximum performance is not required, the supply voltage can be scaled down so that the critical path can still meet the timing constraints. Hence, the power can be significantly reduced due to the quadratic dependency between the supply voltage and power.



3.1 Components of Adaptive Voltage Regulation

Figure 3.1 Adaptive Voltage Regulation

In order to adaptively control the supply voltage, three essential components are required. A critical path emulator accurately predicts the performance of critical path at the regulated supply voltage, if the regulated voltage is not high enough for the critical path to meet the timing constraints, then the supply voltage is supposed to be raised. On the contrary, if the performance is better than expected, the supply voltage can be scaled down more. The other component of adaptive voltage regulation is a controller. The controller receives the result of the critical path emulator, and then compares the result to the expected performance. The output of the controller is transferred to the voltage converter, it contains the decision, which is scaling up the voltage or scaling down it, after the comparison and it also contains the error which gives voltage converter the information of the factor of scaling operation. A dc-dc converter is often used in voltage regulation, as a voltage converter, good energy efficiency is required to save more power than loss. The scheme is illustrated in Figure 3.1, the regulated voltage is transmitted to both the controller and the application. The controller outputs the up/down signal and the error to the dc-dc converter according to the predicted frequency generated by the critical path emulator which indicates the performance at the regulated voltage. The details of these components will be illustrated in next few sections.

### **3.2 Voltage Converter**

The purpose of a voltage converter is to supply a regulated DC output voltage. dc-dc converters are commonly used in applications requiring regulated DC power, such as computers, medical instrumentation and communication devices.

### 3.2.1 Pulsed Width Modulation (PWM) Operation

Basic dc-dc converters such as buck, boost, and buck-boost converters are similar in that they each have two complementary switches and one inductor. Their conversion ratios may all be adjusted by using PWM to vary the duty cycle. The pulsed width modulation control technique maintains a constant switching frequency and varies the ratio of the charge cycle (time when the switch is on)) and the discharge cycle (time when the switch is off) as the load varies. Duty cycle can be represented as:

$$D = \frac{t_{on}}{T_s} = \frac{V_{control}}{V_{st}}$$
(3.6)

This technique offers high power efficiency. In addition, because the switching frequency is fixed, the noise spectrum is relatively narrow, allowing simple low-pass filter techniques to greatly reduce the peak-to-peak voltage ripple at the output. This is the reason why that PWM is popular in telecommunication application where noise interference is of concern. The circuit in Figure 3.2(a) shows a buck converter with a strictly resistive load and the switch mode, and its waveform is shown in Figure 3.2(b).



Figure 3.2(a) Switching Mode Buck Converter


Figure 3.2(b) Waveform of the Switching Mode Buck Converter

## 3.2.2 Buck Converter

The buck converter circuit, which can produce any arbitrary output voltage, is given in Figure 3.3. In the ideal case, the DC output voltage is given by the product of the input voltage and the duty cycle:

$$V_o = D \cdot V_{in} = \frac{t_{on}}{T_s} \cdot V_{in}$$
(3.7)



Figure 3.3 Buck Converter

## 3.2.3 Boost Converter

The boost converter circuit, which can produce any arbitrary output voltage  $V_o \ge V_{in}$ , is given in Figure 3.4. In one portion of cycle, (1 - D), the NMOS device is on, and the input voltage is applied across L, building up current and thus storing energy in the inductor. When the NMOS switch is turned off, the attempt to interrupt the current in the inductor causes the voltage rise rapidly. The PMOS device is turned on at this point, limiting the voltage produced by this inductive kick to the voltage on the output capacitor. During the fraction of the cycle, D, that the PMOS device conducts, some of the energy stored in the inductor is transferred to the output, along with additional energy flowing from the input. The cycle then repeats. The input and output voltages are related by:





Figure 3.4 Boost Converter

## 3.2.4 Buck-Boost Converter

The operation of the buck-boost converter is similar to that of the buck converter, in that the cycle starts with the input voltage applied across the inductor, in this case through the PMOS device for a duration. However, when the PMOS device is turned off, the circuit produces an output voltage polarity opposite to that of the input Figure 3.5. The energy transferred to C during this portion, 1-D, of the cycle(while NMOS device conducts) is only the energy stored in the inductor, with none coming directly from the input. Setting the average voltage across the inductor equal to zero allows the conversion ratio to be found:



Figure 3.5 Buck-Boost Converter

### **3.3 Reference Voltage Generator**

#### 3.3.1 Traditional Reference Voltage Generator

In order to scale down the supply voltage, we need a reference voltage generator to generate a stable voltage. Therefore, the most important thing in the reference voltage generator is that the reference voltage must independent of temperature and external supply voltage as possible. Figure 3.6 shows the traditional reference voltage generator circuit.



Figure 3.6 Traditional Reference Voltage Generator

In Figure 3.6, there are four stages in this architecture, voltage divider, reference voltage generator, voltage follower and output driver. M1~M3, M5, M7 and M8 are operated at saturation region. M4 and M6 are operated at linear region. Voltage follower is a differential amplifier. Output driver is a large PMOS providing large current to the logic circuit. Figure 3.7 shows the illustration of reference voltage

generator circuit.



According to the Figure 3.6, the transistor M5 and M8 are operated at saturation region, so it is to be a current source which is shown in Figure 3.7. M4, M6 and M7 are operated at linear region, so it is to be a resistor. In this reference voltage generator, there are two important paths, path1 (P1) and path2 (P2). P1 is a supply-independent skill to reduce the dependency between M2 current and supply voltage. P2 is a negative feedback compensation to increase the Vref stability. M5 and M8 are operated in saturation region to be a voltage control current source, so M5 and M8 will pull each other. On the other hand, it should be point out that in this reference voltage generator Vref > Vt (M5) + Vt(M8) must be satisfied.

## 3.3.2 Modified Variable Voltage Generator

In the adaptive voltage regulation, not only one voltage level is needed. Therefore, a modified scheme of the variable voltage generator is proposed. The modified scheme is shown in Figure 3.8. Since the output of the reference voltage generator varies along with the value of R3 in Figure 3.8, we replace the transistor M7 with five paralleled transistors and a 5-bit control signal. After properly sized, these transistors can produce five different values of resistance. As the result, five different voltage levels can be produced. In this case, we set the five voltage levels to be 0.8V, 0.9V, 1.0V, 1.1V and 1.17V.



Figure 3.8 Modified Reference Voltage Generator

## **3.4 DC-DC Voltage Converter**

The linear regulator is the basic building block of nearly every power supply used in electronics. Many efforts have recently been made to incorporate an on-chip dc-to-dc voltage converter into a high-density VLSI chip. This is because the power-supply voltage must be reduced to solve problems such as thermal cooling, power dissipation, and device reliability of a shorter channel MOS transistor. A straightforward approach to solve the above problem is to lower the supply voltage to below the traditional 1.2 V in accordance with the parts of chip that can accept the lower supply voltage but at the same time does not degrade their performance.

A voltage regulator provides this constant DC output voltage and contains circuitry that continuously holds the output voltage at the design value regardless of changes in load current or input voltage (this assumes that the load current and input voltage are within the specified operating range for the part).

There are two kinds of voltage coverters have been widely used in modern VLSI chip. One is op-amp-based on chip dc-dc voltage converter. Another type of voltage converter that has traditionally been used in power systems is the switching mode circuit. Switching-mode voltage converter has high power efficiency but must use an LC filter that requires external parts. This is a main drawback of the switching-mode voltage converter. On the other hand, if the L and C components are integrated into the chip, the layout area will be very large and the accuracy is usually very poor. So, for a fully integrated solution, op-amp-based voltage converter shall be adopted.

A voltage converter for use in digital logic chips should have the following target specifications:

\$ 1896

1. Low standby current so that the voltage converter consumes little power when on standby.

2. A small layout area.

3. A stable reduced internal supply voltage for a wide range of operation conditions.

The architecture of the on-chip voltage converter is shown in Figure 3.9. The basic blocks include a reference voltage generator, a differential-amplifier-based voltage follower, and an output driver circuit with low output impedance and high-driving capability. The function of the reference voltage generator is to produce a stable voltage that is free from fluctuations of and temperature. Because there are two

possible equilibrium points on the current source circuit, a voltage divider circuit is necessary. The voltage follower consists of a differential amplifier is to work as a gain stage in the voltage converter. Then, a nMOS transistor is used to work as a source follower and the output of the source follower is connected to the input of the differential amplifier to form a negative feedback system. This structure is very suitable for voltage converter when it is used in random logic circuits in terms of both drivability and layout.



Figure 3.9 The Architecture of On-chip Voltage Converter

## 3.4.1 Output Buffer

The voltage follower should have enough current supply capability and low output impedance so that the output voltage is not very much affected by the large loading current fluctuation. In the practical design, two kinds of follower s have been employed: n-type and p-type, both shown in Figure 3.10.



Figure 3.10 Differential Based Output Buffer

N-type has excellent phase margin, and therefore suitable for the logic chip concerning loop stability. The other type has been widely used in memory chips, because the storage capacitances of memory cells can be used to make the follower stable. For digital logic chip application, it should use n-type voltage follower. Since the voltage down converter should supply large current to logic gates which have large fluctuation of loading current, we attach a large-size nMOS transistor at the output as the driver to enhance drivability of this voltage down converter. On the other hands, it should be point out that in n-type voltage follower, Vext > Vout + Vtn should be satisfied. The larger Vref is, the more the range of Vext is limited, though reference voltage generator allows a wide range. Therefore, selection of the voltage follower should be determined by its application and specific parameters.

## **3.5 Simulation Result**

The most important thing in reference voltage generator is that the reference voltage must be independent of temperature. Figure 3.11 shows the proposed voltage converter circuit. Five voltage levels can be produced by switching the "sel[4:0]"





Voltage Level	Vref(-25°C)	$Vref(25^{\circ}C)$	$Vref(100^{\circ}C)$	Variation	
1	1.17V	1.17V	1.15V	-0.02V	
2	1.10V	1.10V	1.07V	-0.03V	
3	1.01V	1.00V	0.96V	-0.05V	
4	0.92V	0.90V	0.82V	-0.10V	
5	0.82V	0.80V	0.70V	-0.12V	

Table 3.1Temperature Variation

Table 3.2 and Table 3.3 show the simulation results of the modified Variable voltage generator.

Simulation Model	TSMC 0.13um
Temperature Variation (at 0.8V)	0.94 mV/°C
Temperature Variation (at 1.17V)	0.16 mV/°C
Average Power Consumption	1.73 mW
Output Loading	1uf/10k

Table 3.2 Simulation Result of the Modified Variable Voltage Generator

Output Voltage	1.17V	1.1V	1.0V	0.9V	0.8V
Power Efficiency	78%	75%	70%	64%	58%

Table 3.3 Power Efficiency Simulation Results

3.6 PLL Based Adaptive Voltage Regulation Using FSM



Figure 3.12 PLL Based Adaptive Voltage Regulator Using FSM

A scheme of PLL based adaptive voltage regulation is shown in Figure 3.12. It has a voltage-controlled ring oscillator as a critical path emulator, a phase lock loop (PLL) with a finite state machine (FSM) as the controller, and a variable reference voltage generator as a voltage converter.

## **3.6.1 Reference Circuit**



Fig. 3.13 Reference Circuit

Figure 3.13 shows the reference circuit which is used in Figure 3.12. The frequency detector is composed of a voltage-controlled ring oscillator, a frequency detector. The ring oscillator operates at the regulated voltage, and indicates the highest performance of the critical path. The frequency detector compares the output of the ring oscillator and the reference frequency, and generates the "fast" signal to the finite state machine (FSM). If the "fast" signal becomes logic "1" means the output of the ring oscillator is higher than the reference frequency. The frequency detector consists of three connected D flip-flops with using the output of the ring oscillator as the clock inputs of the first two D flip-flops, and the reference frequency is used as the clock input of the last D flip-flip. The "fast" signal captured at the second cycle of the reference circuit is the right value. However, there may be some distance between the clock edge of the ring oscillator and the reference frequency at the first clock cycle. Hence, match delay elements in Figure 3.12 are placed between the ring oscillator and the frequency detector.

#### **3.6.2 Finite State Machine (FSM)**



Figure 3.14 FSM

The finite state machine, which is shown in Figure 3.14, is a mechanism to determine if the voltage should be level up or level down according to the "fast" signal and enable. The voltage is determined to be level up if "fast" signal is at logic "0" when the enable rises, and is determined to be level down if "fast" signal is at logic "1" when the enable rises. Moreover, the voltage is determined to be locked in three conditions, if "fast" signal is at logic "0" when the first enable rises and then it changes to logic "1" when the next enable rises, or if "fast" signal is a logic "1" when the first enable rises, the voltage to logic "0" when the next enable rises, the state of the first enable rises and then it changes to logic "0" when the next enable rises, the

other condition is when the voltage level is at the top level (LV6) or at the bottom level (LV1), which can not be level up or level down anymore. In Figure 3.15, four states and the conditions are illustrated. ST0 is the initial state, when a logic "0" is detected, the voltage level count increases by one, and it goes to ST1. At ST1, it remains ST1 when a logic "0" is detected, if a logic "1" is detected or the voltage level count is 5 (LV6), it goes to ST3, which is the lock state. On the other side, when a logic "1" is detected at ST0, the voltage level count decreases by one, and it goes to ST2. At ST2, it remains ST2 when a logic "1" is detected, if a logic "0" is detected or the voltage level count is 0(LV1), it goes to the lock state, ST3. The finite state machine (FSM) determines the voltage level count, and then sends it to the control unit in Figure 3.12. However, and enable generator is needed to produce a pulse when the frequency detector finish detecting a new frequency generated from the ring oscillator. An enable generator circuit is illustrated in Figure 3.15.



Figure 3.15 Enable Generator

The waveform of the finite state machine is illustrated in Figure 3.16.







# Chapter 4 Discrete Cosine Transform

The increase of the demand for high throughput portable digital equipment, with limited improvement in battery topology, has an increasing interest in low power systems. Since more and more functions are related to video information, video compression and decompression become an essential process to be done in digital devices. Discrete cosine transform (DCT) is frequently used in video compression. Many DCT algorithms were proposed in order to achieve high speed DCT.

## 4.1 Introduction to Discrete Cosine Transform (DCT)

The basic computation in a DCT-based system is the transformation of an 8 x 8 image block from the spatial domain to the DCT domain. The 1-D DCT transform is expressed as:

$$z_{k} = \frac{c(k)}{2} \sum_{i=0}^{7} x_{i} \cos \frac{(2i+1)k\pi}{16}, k = 0, 1, 2, 3, \dots, 7$$

$$c(k) = \begin{cases} 1/2, k = 0 \\ 1, otherwise \end{cases}$$
(4.1)

This equation can be represented in vector-matrix form as:

$$z = T \cdot x^t \tag{4.2}$$

where *T* is an 8 x 8 matrix whose elements are cosine functions. x and z are row and column vectors, respectively. The 8 x 8 matrix can be expressed as:

These coefficients approximate to real numbers as Table 4.1 shows.

Coef.	Real Value	Binary number
a	0.4904	0011 1111
b	0.4619	0011 1011
с	0.4157	0011 0101
d	0.3536	0010 1101
e	0.2778	0010 0100
f	0.1913	0001 1000
g	0.0975	0000 1100

Table 4.1 8 bits DCT coefficients

Many algorithms replace the 8 x 8 matrix T with its decomposition matrixes. Since the even rows and the odd rows of the matrix T are symmetric, the 1-D DCT matrix can be rearranged as:

$$\begin{bmatrix} z_{0} \\ z_{2} \\ z_{4} \\ z_{8} \end{bmatrix} = \begin{bmatrix} d & d & d & d \\ b & f & -f & -b \\ d & -d & -d & d \\ f & -b & b & -f \end{bmatrix} \begin{bmatrix} x_{0} + x_{7} \\ x_{1} + x_{6} \\ x_{2} + x_{5} \\ x_{3} + x_{4} \end{bmatrix}$$
$$\begin{bmatrix} z_{1} \\ z_{3} \\ z_{5} \\ z_{7} \end{bmatrix} = \begin{bmatrix} a & c & e & g \\ c & -g & -a & -e \\ e & -a & g & c \\ g & -e & c & -a \end{bmatrix} \begin{bmatrix} x_{0} - x_{7} \\ x_{1} - x_{6} \\ x_{2} - x_{5} \\ x_{3} - x_{4} \end{bmatrix}$$
(4.3)

The seven coefficients in the two matrixes in equation (4.3) are separated in to two groups, which d, b, f are used to compute  $z_0$ ,  $z_2$ ,  $z_4$ ,  $z_8$  and the others are used to compute  $z_1$ ,  $z_3$ ,  $z_5$ ,  $z_7$ . Hence, this decomposition is widely used in most of the DCT algorithms.

## 4.2 Alternative Implementation

There are alternative methods of implementing DCT. Some of them use fast algorithms to achieve high throughput, some use external memory such as ROM to reduce computation time. The only purpose of these methods is to reduce the number of multipliers used in DCT, since multiplication uses most of the computation time of DCT.

## 4.2.1 Multiplier Implementation

While direct implementation of 8-point DCT requires 64 multiplications and 56 additions, many fast algorithms have been proposed to reduce the multipliers required. An example is shown in Fig. 4.1, which needs only 13 multiplications and 29 add/subtract blocks. It's based on the flow graph of the B.G. Lee's algorithm[11].



Fig. 4.1 Multiplier Implementation[11]

# 4.2.2 Pure ROM Implementation



Fig. 4.2 Architecture of DA[11]

This method uses only additions and look-up tables to replace the multiplications in DCT, which is known as distributed arithmetic. Distributed arithmetic is an effective way to compute DCT totally or partially as scalar products. It changes the summing order of the DCT equation and the initial multiplications are distributed to another computation pattern. Fig. 4.2 shows the architecture for the computation of M-bit input inner product[16]. The inverter and the multiplexer are used for inverting the final output of the ROM to compute C0. The pure ROM implementation of 8-point DCT using distributed arithmetic is shown in Fig. 4.3.



Fig. 4.3 Pure ROM Implementation[11]

## 4.2.3 Mixed ROM Implementation



Fig. 4.4 Mixed ROM Implementation[11]

The mixed ROM implementation also uses the distributed arithmetic algorithm. It takes the advantage of the additions and subtractions at first stage and the decomposition of matrix T, which are the matrixes shown in (4.3). In this case, the number of words per ROM can be reduced.

## 4.3 Reducing Power of DCT

## 4.3.1 Reducing Power through Pipelining

As the voltage decreases, the overall delay increases. In order to balance the increase in delay with the decrease in voltage, pipelining techniques are used to divide the data path into smaller stages. By exploiting pipelining techniques, the data path having longer delay can be operated at high voltage, while the data path having shorter delay can be operated at low voltage. Hence, delay of any stages maintains in a correspondent value.

An example of pipelined DCT architecture is illustrated in Fig. 4.5[12].



Fig. 4.5 Pipelined DCT Architecture[12]

The whole DCT data path is divided in to four pipeline stages. Stage 1 operates with a set of 8 pixel values during eight clock cycles. The stages 2, 3, and 4 are executing with different set of values during each clock cycle. This pipelined architecture performs the operation of the two matrixes in equation (4.3) with three addition/subtraction stages and one parallel multiplication stage.

## 4.3.2 Reducing Power through Parallelism



Fig. 4.7 Two Parallel Data Paths

Parallelism is another method to maintain the throughput while the voltage decreases. In order to compensate the increase in the data path delay, the data path is replicated several times. The input samples are split among those data paths, and the outputs of these data paths are multiplexed into a single output stream. By exploiting parallelism, even if the overall delay increases, the same throughput can be achieved since more data paths compute more inputs at a time. However, the drawback of this method is that more hardware is required when replicating the data path. In order to two times throughput than the original data path, two times hardware than the original is required. Therefore, it becomes a trade off between hardware cost and the increase in delay as the voltage decreases. For example, an original data path is shown in Fig. 4.6, and Fig. 4.7 illustrates the data paths replicated from the original one. Since parallelism greatly increases the area, the power dissipation of a system consisting of N parallel data paths and having the same throughput rate is given as follow[12]:

$$P_{parallel} = \frac{P_{original}}{N^2}$$
(4.4)

where  $P_{original}$  is the power dissipation for a single data path, and  $P_{parallel}$  is the power dissipation for a system consisting of *N* parallel data paths.

## 4.3.3 Reducing Power though Reducing Complexity

Some algorithms are proposed to reduce the complexity of DCT[13], with the reduction of the complexity, the power consumption can be proportionally reduced. However, some approximation has to be made in order to reduce the complexity. Therefore, it becomes a trade-off between image quality and power consumption. An example is illustrated as follow. Since the number of multipliers is decided by the number of non zero digits of the DCT coefficients, if some non zero digits can be

changed to zeros, some multipliers can be reduced. However, changing the non zero digits to zeros affects the quality of the image, the degree of the impacts is related to the sensitivities which is the PSNR degradation of an image when the non zero digit is modified to zero. At first, the DCT coefficients are represented by the Canonical Signed Digit (CSD), which is shown in Table 4.2, instead of the binary numbers. The non zero digits are first reduced without influencing on the image quality. For some non zero digits of the DCT coefficients in CSD format, changing them to zeros leads to large PSNR degradation of an image while for other non zero digits, changing them to zeros leads to negligible PSNR degradation. Based on the sensitivities of each non zero digits, the least sensitive non zero digits are found and are changed to zeros. This concept is shown in Fig. 4.8.



Coef.	Real Value	Binary number	CSD number
а	0.4904	0011 1111	0100 000i
b	0.4619	0011 1011	0100 1011
с	0.4157	0011 0101	0011 0101
d	0.3536	0010 1101	0010 1101
e	0.2778	0010 0100	0010 0100
f	0.1913	0001 1000	0001 1000
g	0.0975	0000 1100	0001 0100

Table 4.2 DCT Coefficients Represented by CSD



Fig. 4.8 Changing the Least Sensitive Non Zero Digits to Zeros

The PSNR degradation versus the power consumption of changing the least sensitive non zero digits to zeros is shown in Table 4.3. In Table 4.3, the PSNR between original version and modified version does not degrees much, but the power consumption reduces a significant amount.

	Original	Modified
PSNR	33.0	32.9

Table 4.3 PSNR

## 4.4 Reconfigurable Architecture

In modern society, the customers prefer more and more optional functions in digital devices, more and more battery is desirable. However, with the scale down of the size of digital devices, it's hard to have more battery and functions at the same time. Therefore, reconfigurable architecture becomes a good choice for both battery saving and more optional functions. For example, if the best image quality is not required, the digital devices can use the low power mode, which the image quality may be lower but saves the battery; if the best image quality is required, the digital devices can use the normal mode.

## 4.4.1 Computation Sharing Multiplier Algorithm (CSHM)

Computation Sharing Multiplier (CSHM) algorithm targets computation re-use in vector-scalar products and is effectively used in DCT[14], since multipliers are essential components in DCT. This algorithm reduces the redundant computation in multiplication by using vector scaling operation. For example, when computing  $C \cdot x[n]$  with C is a vector and  $C=[c_0, c_1, c_2,..., c_{M-1}]$ . If  $c_0=01100111$ , and  $c_1=10001011$ ,  $c_0 \cdot x$  and  $c_1 \cdot x$  can be decomposition in to vector-scalar form as follow:

$$c_0 \cdot x = 2^5 \cdot (0011) \cdot x + 2^0 \cdot (0111) \cdot x$$
 (4.5)

$$c_1 \cdot x = 2^7 \cdot (0001) \cdot x + 2^0 \cdot (1011) \cdot x \tag{4.6}$$

which is called 4-bit decomposition. If (0001)x,  $(0011) \cdot x$ ,  $(0111) \cdot x$ ,  $(1011) \cdot x$  are available, the entire multiplication  $c_0 \cdot x$  and  $c_1 \cdot x$  can be implemented by only adders and shifters. If these bit basic sequences is referred as alphabets, an alphabet set is a set of alphabets that spans all the coefficients in vector *C*, which is  $c_0$ ,  $c_1$ ,  $c_2$ ,...,  $c_{M-1}$ . The alphabet set of  $c_0 \cdot x$  and  $c_1 \cdot x$  is {0001, 0011, 0111, 1011}. The CSHM architecture for 4-bit decomposition is illustrated in Fig. 4.9 and the CSHM architecture for  $c_1 \cdot x$  is illustrated in Fig. 4.10. Fig. 4.9 shows that the CSHM architecture is composed of a pre-computer, an adder, multiplexers and shifters. The pre-computer computes the multiplication of alphabets and x. When the pre-computer finishes its operations, the 4-bit sequence coefficient after decomposing the coefficient is transferred to the shifter to remove the zeros of the ride side of the 4-bit sequence. The select decision is made and also the number of zeros which are shifted is recorded. The multiplexers choose the right element from the results of the pre-compute according to the select decision and send the value to the I-shifter. The I-shifters shift the zeros in inverse direction of the shifters according to the value recorded by the shifters. The two results of the I-shifters are added after properly shifted. Finally, the multiplication is done with reduction of redundant computation.



Fig. 4.9 CSHM Architecture for 4-bit Decomposition[14]



4.4.2 DCT Coefficients for 4-bit Decomposition

Coef.	Real Value	Binary number	Alphabet • x
a	0.4904	0011 1111	0011x, 1111x
b	0.4619	0011 1011	0011x, 1011x
с	0.4157	0011 0101	0011x, 0101x
d	0.3536	0010 1101	0001x, 1101x
е	0.2778	0010 0100	0001x
f	0.1913	0001 1000	0001x
g	0.0975	0000 1100	0011x

Table 4.4 8-bit DCT Coefficients and the Alphabets[14]

The DCT coefficients are represented by 8-bit binary numbers. Table 4.4 shows the binary numbers and the corresponding alphabets for 4-bit decomposition. The alphabet set of DCT coefficients is {0001, 0011, 0101, 1011, 1101, 1111}, which means six products need to be pre-computed.

### 4.4.3 DCT Coefficients for 2-bit Decomposition

According to 4-bit decomposition of CSHM algorithm, 2-bit decomposition can be derived. For example, when computing  $C \cdot x[n]$  with *C* is a vector and  $C=[c_0, c_1, c_2, ..., c_{M-1}]$ . If  $c_0=01100111$ , and  $c_1=10001011$ ,  $c_0 \cdot x$  and  $c_1 \cdot x$  can be decomposition in to vector-scalar form as follow:

$$c_0 \cdot x = 2^6 \cdot (00) \cdot x + 2^4 \cdot (11) \cdot x + 2^2 \cdot (01) \cdot x + 2^0 \cdot (11) \cdot x_{(4.7)}$$

$$c_1 \cdot x = 2^6 \cdot (00) \cdot x + 2^4 \cdot (01) \cdot x + 2^2 \cdot (10) \cdot x + 2^0 \cdot (11) \cdot x_{(4.8)}$$
which is called 2-bit decomposition.

Table 4.5 shows binary numbers and the corresponding alphabets for 2-bit decomposition of the DCT coefficients. The alphabet set of DCT coefficients is {01, 11}, which means only two products need to be pre-computed. Compare Table 4.4 to Table 4.5, six products needs to be pre-computed for 4-bit decomposition, while only two products needs to be pre-computed for 2-bit decomposition. Since the number of the alphabets in an alphabet set is fewer, the multiplication is reduced. However, the number of multiplexers and shifters required in 2-bit decomposition is twice than that in 4-bit decomposition, since the 8-bit coefficient is now divided into four groups and 2-bit for each group. Luckily, the first two bits of the seven DCT coefficients are all zeros, hence, only one multiplexer and two shifters are required more.

Coef.	Real Value	Binary number	Alphabet • x
a	0.4904	00 11 11 11	11x
b	0.4619	00 11 10 11	01x, 11x
с	0.4157	00 11 01 01	01x, 11x
d	0.3536	00 10 11 01	01x, 11x
e	0.2778	00 10 01 00	01x
f	0.1913	00 01 10 00	01x
g	0.0975	00 00 11 00	11x

Table 4.5 8-bit DCT Coefficients and the Alphabets for 2-bit Decomposition

The architecture for 2-bit decomposition of DCT coefficients is illustrated in



Fig. 4.11 CSHM Architecture for 2-bit Decomposition of DCT Coefficients

## 4.4.4 DCT Architecture Based on CSHM algorithm

Since 1-D DCT matrix T can be expressed as equation (4.3), by decomposing and rearranging equation (4.3), it gives the equation as follow:

$$\begin{bmatrix} z_{0} \\ z_{2} \\ z_{4} \\ z_{6} \end{bmatrix} = \begin{bmatrix} d & d & d & d \\ b & f & -f & -b \\ d & -d & -d & d \\ f & -b & b & -f \end{bmatrix} \cdot \begin{bmatrix} x_{0} + x_{7} \\ x_{1} + x_{6} \\ x_{2} + x_{5} \\ x_{3} + x_{4} \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} z_{0} \\ z_{2} \\ z_{4} \\ z_{6} \end{bmatrix} = (x_{0} + x_{7}) \begin{bmatrix} d \\ b \\ d \\ f \end{bmatrix} + (x_{1} + x_{6}) \begin{bmatrix} d \\ f \\ -d \\ -b \end{bmatrix} + (x_{2} + x_{5}) \begin{bmatrix} d \\ -f \\ -d \\ b \end{bmatrix} + (x_{3} + x_{4}) \begin{bmatrix} d \\ -b \\ d \\ -f \end{bmatrix} (4.9)$$

During this operation, the multiplications with the same set of coefficients are combined together in order to reduce the number of pre-computers. According to equation (4.9), the architecture of computing  $[z_0, z_2, z_4, z_6]$  is illustrated in Fig. 4.12. The select-adder unit in Fig. 4.12 means the architecture shown in Fig. 4.10. Similarly, the equation for computing  $[z_1, z_3, z_5, z_7]$  is given as follow:

$$\begin{vmatrix} z_{1} \\ z_{3} \\ z_{5} \\ z_{7} \end{vmatrix} = \begin{vmatrix} a & c & e & g \\ c & -g & -a & -e \\ e & -a & g & c \\ g & -e & c & -a \end{vmatrix} \cdot \begin{vmatrix} x_{0} - x_{7} \\ x_{1} - x_{6} \\ x_{2} - x_{5} \\ x_{3} - x_{4} \end{vmatrix}$$
$$\Rightarrow \begin{vmatrix} z_{1} \\ z_{3} \\ z_{5} \\ z_{7} \end{vmatrix} = (x_{0} - x_{7}) \begin{vmatrix} a \\ c \\ e \\ g \end{vmatrix} + (x_{1} - x_{6}) \begin{vmatrix} c \\ -g \\ -a \\ -e \end{vmatrix} + (x_{2} - x_{5}) \begin{vmatrix} e \\ -a \\ g \\ c \end{vmatrix} + (x_{3} - x_{4}) \begin{vmatrix} g \\ -e \\ c \\ -a \end{vmatrix} (4.10)$$

And the corresponding architecture of equation (4.10) is illustrated in Fig. 4.13.



Fig. 4.12 Architecture of Computing [z<sub>0</sub>, z<sub>2</sub>, z<sub>4</sub>, z<sub>8</sub>]



Fig. 4.13 Architecture of Computing  $[z_1, z_3, z_5, z_7]$ 

## 4.4.5 Modified DCT coefficients

By exploiting the 2-bit decomposition of CHSM algorithm, if we modify the DCT coefficients, the complexity of the computation can be reduced. The way we modify the DCT coefficients is focus on the least sensitive bits which are the last two bits of each coefficient. Type1 and Type2 of modified coefficients are shown in Table 4.6 and Table 4.7, respectively. We properly adjust the original values of the coefficients into approximate values as a result that the last two bits of the coefficients are "11" and "00" for [ $z_0$ ,  $z_2$ ,  $z_4$ ,  $z_6$ ] and [ $z_1$ ,  $z_3$ ,  $z_5$ ,  $z_7$ ], which becomes Type1 modified coefficients. Type1 can replace the select-adder for the last two bits with a multiplexer only. The method of generating Type2 of the coefficients is to change the last two bits of all the coefficients into "00". As a result, the select-adder for the last two bits can be reduced.



Coef.	Real Value	Binary number	Alphabet $\cdot x$
a	0.4904	00 11 11 11	3x
b	0.4688	00 11 11 00	3x
с	0.3984	00 11 00 11	3x
d	0.3438	00 10 11 00	1x, 3x
e	0.2734	00 10 00 11	1x. 3x
f	0.1913	00 01 10 00	1x
g	0.0860	00 00 10 11	1x, 3x

Table 4.6 Type1 Modified Coefficients

Coef.	Real Value	Binary number	Alphabet $\cdot x$
a	0.4688	00 11 11 00	3x
b	0.4375	00 11 10 00	1x, 3x
с	0.4063	00 11 01 00	1x, 3x
d	0.3438	00 10 11 00	1x, 3x
e	0.2778	00 10 01 00	1x
f	0.1913	00 01 10 00	1x
g	0.0975	00 00 11 00	3x

Table 4.7 Type2 Modified Coefficients

# 4.5 Simulation Result

We pipelined the DCT architecture into two stages in order to fit the performance requirement. The pipelined architecture is illustrated in Fig 4.14.



Fig. 4.14 Pipelined DCT Architecture

The Simulation result of the pipelined DCT is shown in Table 4.8.
Simulation Model	TSMC 0.13um
Temperature	25°C

Frequency (VDD=1.2V)	435MHz	399MHz	350MHz	300MHz	222MHz
Power Consumption (8bit)	53.45mW	50.38mW	45.33mW	38.09mW	25.71mW

Table 4.8 Simulation Result of the Pipelined DCT



# Chapter 5 Adaptive Voltage Scaling for Discrete Cosine Transform

Discrete cosine transform has been widely used in mobile products recent years. Therefore, the performance and the power consumption of the discrete cosine transform unit take a more and more important roll in digital systems. In order to minimize the energy-delay-product (EDP) of the discrete cosine transform, we apply the proposed adaptive voltage scaling technique on it. The adaptive voltage scaling system adaptively controls the operating voltage to the fittest level on the operating frequency. As the result, the power consumption of the discrete cosine transform unit can be reduced yet the timing requirement is still met.

## 5.1 The Proposed Architecture



Figure 5.1 Architecture of Adaptive Voltage Scaling System

#### for Discrete Cosine Transform

The proposed architecture of adaptive voltage scaling system for discrete cosine transform is shown in Figure 5.1. It is composed of two variable voltage generators (VVG), a reference circuit, a controller and the application which is the discrete cosine transform unit in this case. One of the variable voltage generators (VVG1) is to provide the regulated voltage to the ring oscillator, the other one (VVG2) is connected to the output buffer for the use of discrete cosine transform unit. The reference circuit compares the operating frequency to the oscillated frequency which is the predicted performance of the critical path with the regulated voltage. The output of the reference circuit is transferred to the controller to determine the value of the "pre sel". The "pre sel" is a 5-bit signal which controls the VVG1 to scale up or scale down the regulated voltage. If the "slow" signal, which is the output of the reference circuit, is logic low, it means that the predicted frequency is quicker than the operating frequency, so that the regulated voltage should be scale down. In other hand, if the "slow" signal is logic high, the regulated voltage should be scale up. Since we scale the voltage from 1.2V at first, when the "slow" signal turns to logic high, the "pre sel" signal is locked at the same time, which means the fittest voltage level is determined. At last, the "sel" signal is transferred to VVG2 to provide the final voltage to the discrete cosine transform unit.

#### 5.2 Reference Circuit

The reference circuit, which is shown in Figure 5.2, is a critical path emulator, which is also called a critical path replica. The critical path of the system can be duplicated to form a fan-out-4 (FO4), ring oscillator, or a delay line which adaptive responds to environment and process variation. The reference circuit in our design contains two blocks, one is the ring oscillator, the other one is the frequency detector. The ring oscillator operates at the regulated voltage, and indicates the highest performance of the critical path. The frequency detector compares the output of the ring oscillator and the reference frequency.



Figure 5.2 Reference Circuit

#### 5.2.1 Ring Oscillator

The ring oscillator circuit is shown in Figure 5.3. The first stage of the ring oscillator is replaced with an NAND gate so that the "ctrl" signal enables or disables the oscillation. The delay element is formed by two inverters. At first, the "vco\_out" over the 1.2V is set to be the critical path of the application, which is the discrete cosine transform unit. The predicted critical path has been compared to the simulation result of the critical path of the discrete cosine transform over the five voltage level in order to make sure the accuracy of the predicted model.



Figure 5.3 Ring Oscillator

### **5.2.3 Frequency Detector**

The frequency detector detects the difference between the operating frequency and the "vco\_out". If the "vco\_out" is slower than the operating frequency, the "slow" signal turns to logic high. The traditional frequency detector, which is shown in Figure 5.4, finishes the detection in one reference frequency cycle. In the closed loop of the adaptive voltage scaling, there are at most five iterations. The frequency detector is used one time each iteration. Therefore, there are at most five cycles that we would waste on the frequency detector. Different from the traditional frequency detector, the frequency detector, which is shown in Figure 5.5, used in our design can finish the detection in half reference frequency cycle. In this way, we can get the decision in the first half reference cycle, and then the other half cycle can be used for the controller. Therefore, the next iteration goes on in the next reference cycle, no cycles would be wasted then.



Figure 5.4 Traditional Frequency Detector



Figure 5.5 Frequency Detector

In Figure 5.5, the reversed "vco\_out" is connected to a D flip-flop and the output is compared to the delayed and reversed frequency. Figure 5.6 and Figure 5.7 shows how the frequency detector works, where "freq" means the operating frequency, "d\_freq" means the delayed frequency. It is obvious that when the "vco\_out" is faster than the frequency, the "slow" signal is always at logic "0", only if the "vco\_out" is slower than the frequency, the "slow" signal turns to logic "1". Since only half reference cycle is required for this kind of frequency detector, it improves the adaptive voltage scaling system to save half cycle time compared to the traditional one.



Figure 5.6 Waveform of the Frequency detector



Figure 5.7 Waveform of the Frequency detector

#### **5.3 Controller**

The controller controls the two variable voltage generators, it decides the 5-bit "pre\_sel" signal and the 5-bit "sel" signal. During the adaptive voltage scaling iterations, the controller determines the "pre\_sel" signal according to the "slow" signal from the frequency detector. The proposed controller is shown in Figure 5.8. It is consist of a control logic and a selector. The control logic controls the 5-bit "pre\_sel" signal, and the selector controls the 5-bit "sel" signal which is connected to the second variance voltage generator in order to scale the voltage level of the application.

ANILLER,



Figure 5.8 The Proposed Controller

#### **5.3.1 Control Logic**

The architecture of the control logic is shown in Figure 5.9. It is consist of eleven D flip-flops and one multiple input D flip-flops (MDFF). The multiple input D flip-flop, which is shown in Figure 5.10(a), has four inputs and one output. When the "CLK" turns to logic "1", it triggers the MDFF to transfer either input "D0" or "D1" to the output "Q" according to the value of "S0". If "S0" is at logic "0", the MDFF transfers "D0" to "Q"; else if "S0" is at logic "1", the MDFF transfers "D1" to "Q". An example for how the MDFF works is illustrated in Figure 5.10(b). In Figure 5.9, the operating frequency of the discrete cosine transform is taken as the reference frequency (ref). At the first reference cycle, the "reset" remains at logic "1" so that a logic "0" is transferred to the next flip-flop. In the second reference cycle, the "reset" turns to logic "0" so that a logic "1" is transferred to the next flip-flop. Therefore, the logic "1" will be transferred stage by stage. The value of all the flip-flops in one iteration of the controller is shown in Table 5.1. In this way, the 5-bit "pre sel" can be produced by Q4a, Q4b, ...., Q0a, and Q0b. The waveform of the "pre sel" is shown in Figure 5.11. 2 Manute



Figure 5.9 The Propose Control Logic



Figure 5.10 (a) Multiple Input D Flip-Flop (b) Waveform

ALLING.

				-SC		SIN	13				
Cycle	FF	Q4a	Q4b	Q3a	Q3b	Q2a	Q2b	Q1a	Q1b	Q0a	Q0b
0	0	0	0	0	0	<sup>39</sup> 0	0	0	0	0	0
1	1	0	0	0	0	10.11	0	0	0	0	0
2	1	1	0	0	0	0	0	0	0	0	0
3	1	1	1	0	0	0	0	0	0	0	0
4	1	1	1	1	0	0	0	0	0	0	0
5	1	1	1	1	1	0	0	0	0	0	0
5	1	1	1	1	1	1	0	0	0	0	0
6	1	1	1	1	1	1	1	0	0	0	0
7	1	1	1	1	1	1	1	1	0	0	0
8	1	1	1	1	1	1	1	1	1	0	0
9	1	1	1	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	1	1

Table 5.2 Value of the Flip-Flops



Figure 5.11 Waveform of pre\_sel[4:0]

The variable voltage generator scales the voltage according to the "pre\_sel[4:0]". Since the "pre\_sel[4:0]" varies every two reference cycles, the voltage scales down one level once the "pre\_sel[4:0]" changes. There are at most ten reference cycles for the controller to lock the fittest voltage level for the application according to the decision of the reference circuit. Once the "vco\_out" is slower than the operating frequency, the "pre\_sel[4:0]" is locked.

#### 5.3.2 Selector

The functionality of the selector is to select the correct signal for the 4-bit "sel" signal when the "pre\_sel[4:0]" is locked. Since when the "lock" signal is at logic "1", which means that the "vco\_out" is slower than the operating frequency, the correct voltage level we should apply on the application is the previous "pre\_sel[4:0]" before it is locked so that the regulated voltage can fit the required performance. However, there are two exceptions. One of them is when the voltage is at the highest level, which is 1.17V, there is no more higher level for the variable voltage generator to scale, thus the selector would select the same "pre\_sel[4:0]" when it is locked. The

other exception is when the voltage is at the lowest level, there is no more lower level for the variable voltage generator to scale, thus the selector would select the same "pre\_sel[4"0] when it is locked. The proposed selector is shown in Figure 5.12.



Figure 5.12 The Proposed Selector



Figure 5.13 (b) Waveform of the Selector



Figure 5.13(a), (b) and (c) illustrate the behavior of the selector. Figure 5.13(a) shows the general case that the selector chooses the former value of the "pre\_sel[4:0]", which is the correct one we need in general cases. Figure 5.13(b) shows the first exception case, when the "vco\_out" is still slower than the reference frequency at the highest voltage level, the selector would choose the same value of the "pre\_sel[4:0]" instead of the former one. Figure 5.13(c) shows the other exception case, when the "vco\_out" is still faster than the reference frequency at the lowest voltage level, the selector would choose the same value of the selector case, when the "vco\_out" is still faster than the reference frequency at the lowest voltage level, the selector would choose the same value of the "pre\_sel[4:0]".

### **5.4 Simulation Result**

#### 5.4.1 Adaptive Voltage Scaling System



Figure 5.14 Architecture of the AVS with Voltage Generators

Figure 5.14 shows the architecture of the adaptive voltage scaling system. The VVG1 is controlled by "pre\_sel[4:0]" which is determined by the controller, and it generates the reference voltage "vref" to the reference circuit. The VVG2 is controlled by "sel[4:0]" which is also determined by the controller, and it generates the output voltage to the application. The Waveform of the adaptive voltage control system is shown in Figure 5.15(a) and Figure 5.15(b).



Figure 5.15(a) The Waveform of the adaptive voltage control system



Figure 5.15(b) The Waveform of the adaptive voltage control system

In Figure 5.15(a), it's clear to see that the output of the ring oscillator (dco\_out) is slower than the reference frequency (ref10) at the fifth cycle of the "ctrl" signal. Which means the fittest voltage level is at the fourth cycle of the "ctrl" signal. Therefore, the "lock" signal raises at that cycle and outputs the 5-bit digital signal "sel[4:0]" to the VVG2. As the result, the output voltage is at 0.9V which is the fourth voltage level we designed. In Figure 5.15(b), the output of the ring oscillator (dco\_out) is never slower than the reference frequency (ref10). Consequently, the control system would lock the voltage level at the lowest one which is 0.8V in this case.

The simulation result of the adaptive voltage scaling system is illustrated in Table 5.3.

Simulation Model	TSMC 0.13um
Temperature Variation (at 0.8V)	0.94 mV/°C
Temperature Variation (at 1.17V)	0.16 mV/°C
Average Power Consumption	2.26 mW

Table 5.2 Simulation Result of the AVS System

#### 5.4.2 AVS for Discrete Cosine Transform



Figure 5.16 Architecture of AVS for Discrete Cosine Transform

The adaptive voltage scaling system has been applied to the discrete cosine transform to adaptively control the power consumption of it. According to the critical path of the discrete cosine transform, the ring oscillator predicts the performance under the regulated voltage. After the decision which is made by the controller, the select unit chooses the fittest voltage and then applies it to the discrete cosine transform. Table 5.4 shows the original performance/power data and the performance/power data of low power discrete cosine transform which has been adaptively controlled is shown in Table 5.5. The adaptive control system reduces at most 45% power consumption of the discrete cosine transform. The overhead of the control system is at most 28%. Table 5.6 shows the comparison of the original DCT and the adaptive voltage controlled DCT. All the simulation is done with TSMC 0.13um technology.

Frequency(MHz)	435	399	350	300	222
Power(mW)	53.45	50.38	45.33	38.09	25.71

Table 5.3 Original DCT

Vdd(V)	1.2	1.1	1.0	0.9	0.8
Frequency(MHz)	435	399	350	300	222
Power(mW)	57.32	49.05	40.84	28.63	13.91

Table 5.4 Adaptive Voltage controlled DCT



Table 5.5 Original vs. Adaptive

# **Chapter 6 Conclusion & Future Work**

#### 6.1 Conclusion

For SoCs era, the on-chip voltage converter becomes more and more important in circuit design. Take the loading for example, the driving ability and loading are in direct ratio. The direct advantage of on chip variable voltage converter is that can reduce the supply voltage fast and directly for low power design. As the result, the on-chip variable voltage converter is necessary component for feature low power digital IC design.

We proposed two schemes in this thesis. First we proposed the variable voltage generator to provide a stable voltage source for other application. The circuit can change the output value by adjust the transistor size in the allowing range and has temperature dependency of 0.93mV per each Celsius degree. And we apply this voltage source to generate an adaptive voltage control system. Second we develop an adaptive voltage controller based on the variable voltage generator. Along with the reference circuit, it forms an adaptive voltage scaling system. We can considerate frequency and voltage at once by using this circuit.

In the rest of the thesis, we combine the discrete cosine transform and the variable voltage generators to become a power-aware discrete cosine transform design. We have five voltage levels which are controlled by a 5-bit select signal. The adaptive voltage control system has reduced at most 45% of the power consumption of DCT and its power overhead is at most 28%.

#### **6.2 Future Work**

For further more research, two aspects are worth to be developed. One of them is a fully digital power management system with dynamic frequency scaling. The architecture of the fully digital power management is shown in Figure 6.1. An all digital phase-lock-loop (ADPLL) and the Adaptive voltage control system are combined together. The ADPLL divides the reference frequency (Fref) by N. After the adaptive voltage control system defined the predicted performance and the corresponding regulated voltage, the predicted frequency (Fpredict) is feedback to the ADPLL. At last, the ADPLL multiplies the predicted frequency by N. In this way, the lock time of the control system can be reduced since the reference frequency is divided first. Additionally, the frequency scaling can be made by the ADPLL.



Figure 6.1 Fully Digital Power Management System

Another one is Inverse Discrete Cosine Transform design. IDCT is often used in digital signal processing, even more frequently than DCT does. The IDCT design is simple since the operation is almost the same as the DCT, the only difference is the coefficient matrix. The IDCT operation is described by the equation 6.1.

$$\begin{bmatrix} x_{0} + x_{7} \\ x_{1} + x_{6} \\ x_{2} + x_{5} \\ x_{3} + x_{4} \end{bmatrix} = \begin{bmatrix} d & b & d & f \\ d & f & -d & -b \\ d & -f & -d & b \\ d & -b & d & -f \end{bmatrix} \begin{bmatrix} z_{0} \\ z_{2} \\ z_{4} \\ z_{8} \end{bmatrix}$$

$$\begin{bmatrix} x_{0} - x_{7} \\ x_{1} - x_{6} \\ x_{2} - x_{5} \\ x_{3} - x_{4} \end{bmatrix} = \begin{bmatrix} a & c & e & g \\ c & -g & -a & -e \\ e & -a & g & c \\ g & -e & c & -a \end{bmatrix} \begin{bmatrix} z_{1} \\ z_{3} \\ z_{5} \\ z_{7} \end{bmatrix}$$
(6.1)

According to the IDCT coefficient matrix, the CSHM architecture still works except the coefficients changed.



# **Bibliography**

- [1] N. Chabini, I. Chabini, E.M. Aboulhamid, Y. Savaria, "Methods for minimizing dynamic power consumption in synchronous designs with multiple supply voltages," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Page(s):346 – 351, Volume 22, Issue 3, March 2003.
- [2] C. Yeh, M.-C.Chang, "Gate-level voltage scaling for low-power design using multiple supply voltages," IEEE Proceedings Circuits, Devices and Systems, Page(s):334 – 339, Volume 146, Issue 6, Dec. 1999.
- [3] A. Chandrakasan, S. Sheng, R. Brodersen, "Low-Power CMOS Digital Design", IEEE Journal of Solid-State Circuits, 27(4), Apr. 1992.
- [4] T. Mahnke, S. Panenka, M. Embacher, W. Stechele, W. Hoeld, "Efficiency of dual supply voltage logic synthesis for low power in consideration of varying delay constraint strictness," 2002. 9th International Conference on Electronics, Circuits and Systems, Page(s):701 - 704, Volume 2, 15-18 Sept. 2002.
- [5] K. Usami, M. Igarashi, "Low-power design methodology and applications utilizing dual supply voltages," Proceedings of the ASP-DAC 2000. Asia and South Pacific Design Automation Conference, 2000, Page(s):123 – 128, 25-28 Jan. 2000.
- [6] D. Liu and C. Svensson, "Trading Speed for Low Power by Choice of Supply and Threshold Voltages", IEEE Journal of Solid-State Circuits, 28(1), Jan. 1993.
- [7] M. Meijer, F. Pessolano, J. Pineda de Gyvez, "Technology Exploration for Adaptive Power and Frequency Scaling in 90nm CMOS," Proceedings of the 2004 International Symposium on Low Power Electronics and Design, 2004. ISLPED '04, Page(s):14 – 19, 2004.
- [8] Jaeha Kim, M.A. Horowitz, "An efficient digital sliding controller for adaptive power-supply regulation," IEEE Journal of Solid-State Circuits, Page(s):639 – 647, Volume 37, Issue 5, May 2002.
- [9] A. Dancy, A. Chandrakasan, "Techniques for aggressive supply voltage scaling

and efficient regulation [CMOS digital circuits]," Proceedings of the IEEE 1997 Custom Integrated Circuits Conference, 1997, Page(s):579 – 586, 5-8 May 1997.

- [10] A. Stratakos, S. Sanders, and R.W. Brodersen, "A Low-voltage CMOS DC-DC Converter for Portable Battery-operated Systems", Proceedings of the Twenty-Fifth IEEE Power Electronics Specialist Conference, pp. 619-626, June 1994.
- [11] Min Jiang, Yuan Luo, Yiling Fu, Bing Yang, Baoying Zhao, Xin-an Wang, Shimin Sheng, Tianyi Zhang, "A low power 1D-DCT processor for MPEG-targeted real-time applications," ISCIT 2004. IEEE International Symposium on Communications and Information Technology, 2004, Page(s):682 – 687, Volume 2, 26-29 Oct. 2004.
- [12] Yeong-Kang Lai, and Han-Jen Hsu, "A cost-effective 2-D discrete cosine transform processor with reconfigurable datapath," Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS '03, Pg.II-492 -II-495, Vol.2, 25-28 May 2003.
- [13] J.Park and K.Roy, "A low power reconfigurable DCT architecture to trade off image quality for computational complexity," Processing, 2004. Proceedings.
  (ICASSP '04). IEEE International Conference on Acoustics, Speech, and Signal, Pg.V - 17-20, Vol.5, 17-21 May 2004.
- [14] Jongsun Park, Soonkeon Kwon and Kaushik Roy, "Low power reconfigurable DCT design based on sharing multiplication." IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings., pp. III/3116-III/3119, Vol. 3, 2002.
- [15] Yi-Jong Yeh, Sy-Yen Kuo, Jing-Yang Jou, "Converter-free multiple-voltage scaling techniques for low-power CMOS digital design," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Page(s):172 – 176, Volume 20, Issue 1, Jan. 2001.
- [16] "A system level memory power optimization technique using multiple supply and threshold voltages," T. Ishihara, K. Asada, Proceedings of the ASP-DAC

2001. Asia and South Pacific Design Automation Conference, 2001, Page(s):456 – 461, 30 Jan.-2 Feb. 2001.

- [17] Yi-Jong Yeh, Sy-Yen Kuo, "An optimization-based low-power voltage scaling technique using multiple supply voltages," The 2001 IEEE International Symposium onCircuits and Systems, 2001. ISCAS 2001, Page(s):535 - 538, Volume 5, 6-9 May 2001.
- [18] J.T. Kao, M. Miyazaki, A.R. Chandrakasan, "A 175-MV multiply-accumulate unit using an adaptive supply voltage and body bias architecture," IEEE Journal of Solid-State Circuits, Page(s):1545 – 1554, Volume 37, Issue 11, Nov. 2002.
- [19] M. Elgebaly, M. Sachdev, "Efficient adaptive voltage scaling system through on-chip critical path emulation," ISLPED '04. Proceedings of the 2004 International Symposium on Low Power Electronics and Design, 2004, Page(s):375 – 380, 9-11 Aug. 2004.
- [20] S. Dhar, D. Maksimovic, "Low-power digital filtering using multiple voltage distribution and adaptive voltage scaling," ISLPED '00. Proceedings of the 2000 International Symposium on Low Power Electronics and Design, 2000, Page(s):207 – 209 ,2000.
- [21] M. Meijer, J.P. de Gyvez, R. Otten, "On-chip digital power supply control for system-on-chip applications," ISLPED '05. Proceedings of the 2005 International Symposium on Low Power Electronics and Design, 2005, Page(s):311 – 314, 8-10 Aug. 2005.
- [22] S. Dhar, D. Maksimovic, "Switching regulator with dynamically adjustable supply voltage for low power VLSI," IECON '01. The 27th Annual Conference of the IEEE Industrial Electronics Society, 2001, Page(s):1874 1879, Volume 3, 29 Nov.-2 Dec. 2001.
- [23] Shoei-Chuen Lin and Ching-Chih Tsai, "Adaptive voltage regulation of PWM buck DC-DC converters using backstepping sliding mode control," Proceedings of the 2004 IEEE International Conference on Control Applications, 2004, Page(s):1382 – 1387, Volume 2, 2-4 Sept. 2004.

- [24] Martin Yeung-Kei Chui, Wing-Hung Ki, Chi-Ying Tsui, "An integrated digital controller for DC-DC switching converter with dual-band switching," 2003 Symposium on VLSI Circuits, 2003. Digest of Technical Papers, Page(s):45 – 48, 12-14 June 2003.
- [25] Hong Mao, J. Abu-Qahouq, Shiguo Luo, I. Batarseh, "Zero-voltage-switching half-bridge DC-DC converter with modified PWM control method," IEEE Transactions on Power Electronics, Page(s):947 – 958, Volume 19, Issue 4, July 2004.
- [26] Kyeounsoo Kim, P.A. Beerel, "A high-performance low-power asynchronous matrix-vector multiplier for discrete cosine transform," The First IEEE Asia Pacific Conference on ASICs, 1999. AP-ASIC '99, Page(s):135 – 138, 23-25 Aug 1999.
- [27] Min Jiang, Yuan Luo, Yiling Fu, Bing Yang, Baoying Zhao, Xin-an Wang, Shimin Sheng, Tianyi Zhang, "A low power 1D-DCT processor for MPEG-targeted real-time applications," IEEE International Symposium on Communications and Information Technology, 2004. ISCIT 2004, Page(s):682 – 687, Volume 2, 26-29 Oct 2004.