

論文97-34C-11-1

Giga Bit급 저전력 Synchronous DRAM 구조에 대한 연구

(A Study on the Low Power Architecture of Multi-giga bit Synchronous DRAM's)

柳 會 峻 *, 李 正 雨 *

(Hoi-Jun Yoo and Jung-Woo Lee)

요 약

DRAM의 시간에 따른 동작 전류 성분을 분석하여 4 Giga bit 급에서는 감지 증폭기 동작 전류와 데이터 Path의 전류 및 DC 누설 전류 성분이 중요함을 밝혔다. 4 Giga DRAM 급 이상에서는 전원 전압이 1.5V 이하로 감소하고 이에 따라 MOS의 문턱 전압도 감소하리라 예상된다. 이때 문제가 되는 Sub-threshold 누설전류를 감소시키기 위해 Self-Timed Dual VT logic이라는 새로운 CMOS 논리를 제안하였다. 이 방식에서는 Pulse 형태로 신호를 전송하며 대기 상태에서의 누설 전류를 High VT MOS로 막으며 Size가 큰 MOS로 고속 Active Transition을 이루고 별도의 Feedback 경로로 대기 Transition을 하도록 한다. 3.3V용 DRAM 회로에서 감지 증폭기 동작 전류를 감소하기 위해 8 bit Prefetch방식을 채택하고 동작 전압을 1.2V로 감소하였다. 8 bit에 대한 Sensing 속도는 두 경우가 모두 동일하지만 전류는 3.3V의 값의 2.3%인 0.7mA만을 갖는 저전력 특성을 보였다. 3 cm 이상의 데이터 경로를 갖으며 RC Delay가 약 23.3ns가 되는 4 Giga bit 급 DRAM의 데이터 경로에서 소모되는 전력을 감소시키기 위해 4단 Pipeline 구조를 채택하였다. 설계의 편의를 도모하고자 Wave Pipeline 방식을 채용하여 5ns Pulse폭을 갖는 4개의 데이터를 전송하도록 하였다. 동작 전압을 3.3V에서 1.2V로 감소함에 따라 전류는 22mA에서 2.5mA로 감소하였지만 Cycle Time이 6ns로 감소하여 오히려 동작속도는 4배 향상됨을 확인할 수 있었다.

Abstract

The transient current components of the DRAM are analyzed and the sensing current, data path operation current and DC leakage current are revealed to be the major current components. It is expected that the supply voltage of less than 1.5V with low VT MOS will be used in multi-giga bit DRAM. A low voltage dual VT self-timed CMOS logic in which the subthreshold leakage current path is blocked by a large high-VT MOS is proposed. An active signal at each node of the self-timed circuit resets its own voltage to its standby state after 4 inverter delays. This pulsed nature speeds up the signal propagation and enables the synchronous DRAM to adopt a fast pipelining scheme. The sensing current can be reduced by adopting 8 bit prefetch scheme with 1.2V VDD. Although the total cycle time for the sequential 8 bit read is the same as that of the 3.3V conventional DRAM, the sensing current is lowered to 0.7mA or less than 2.3% of the current of 3.3V conventional DRAM. 4 stage pipeline scheme is used to reduce the power consumption in the 4 giga bit DRAM data path of which length and RC delay amount to 3 cm and 23.3 ns, respectively. A simple wave pipeline scheme is used in the data path where 4 sequential data pulses of 5 ns width are concurrently transferred. With the reduction of the supply voltage from 3.3V to 1.2V, the operation current is lowered from 22mA to 2.5 mA while the operation speed is enhanced more than 4 times with 6 ns cycle time.

* 正會員, 江原大學校 電子工學科
(Department of Electronic Eng., Kangwon National University)

※ 본 연구는 한국전자통신연구소의 차세대반도체 선
행기초기술연구사업의 지원으로 수행하였습니다.
接受日字:1996年10月31日, 수정완료일:1997年10月1日

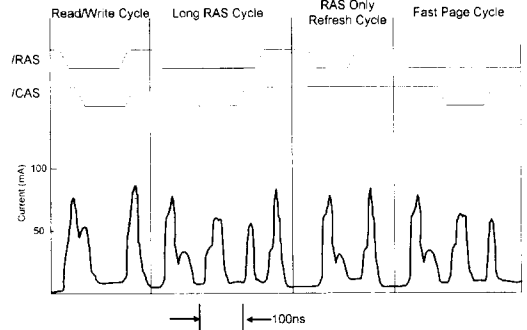
I. 서론

휴대용 컴퓨터의 보급과 휴대용 통신기기의 발달에 따라 전지에 의해서만 구동되며 소모 전력이 적은 부품들의 수요가 나날이 증가하고 있다.^[1] 한편 DRAM은 3년마다 4배씩 그 용량이 증가하고 있으며 최근에는 1 Giga bit DRAM의 개발이 보고되고 있다.^[2-5] 이러한 집적도의 증가는 만일 소자당의 소모 전력에 변화가 없다면 칩의 전력 소모의 급격한 증가를 가져와 DRAM 자체의 성능이나 DRAM을 사용하는 시스템의 성능에 잡음의 증가나 전력 장치의 과부하 등의 악영향을 줄 수가 있다. 따라서 용량이 증가하여도 DRAM을 저전력화함으로써 대용량화에 따른 문제점들을 해결하고 아울러 휴대용 기기에서의 사용을 더욱 확대시킬 필요가 있다^[1].

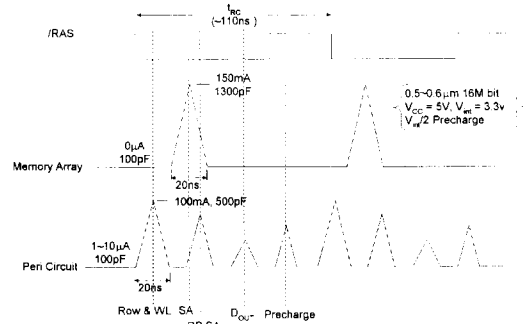
256M bit DRAM 시대부터 DRAM은 범용성의 확보보다는 특정용도에 적합한 형태의 구조를 갖도록 설계되는 경향을 보이고 있다. File Memory나 그래픽 처리를 중심으로 하는 멀티미디어형 Memory 그리고 CPU와의 고속 통신을 위한 DRAM 등이 그 예이다.^[6-12] 이러한 경향이 Giga Bit급 DRAM에서는 더욱 심화될 것으로 예상되며 종래보다 더욱 다양한 DRAM 구조들이 채택될 것이다. 이러한 경향에 맞추어 Giga bit DRAM을 저전력화 관점에서 고찰하여 이에 적합한 회로 및 구조를 연구하고자 한다.

DRAM의 동작 전류 성분은 그림 1(a)에서와 같은 시간 변화를 보인다. 이를 그 성분 별로 분석한 것이 그림 1(b)이다. 이는 /RAS가 하강한 다음 Row Address 경로를 따라 신호가 전달되어 Word Line이 선정될 때까지의 전류, 감지 증폭기에 의해 Cell신호가 증폭될 때 필요한 전류, Column Address에 의해 감지 증폭기가 선택되어 I/O Line에 전달되어 전송될 때 소모되는 전류, 데이터가 출력될 때의 전류 및 /RAS가 Disable되어 DRAM의 내부 회로들이 Pre-charge됨에 따라 소모되는 전류로 나누어 볼 수 있다. 종래의 DRAM에서는 이러한 성분 중 감지 증폭기 동작에서 소모하는 전류가 주를 이루었고 나머지는 상대적으로 무시될 수 있었다. 한편 DRAM의 용량의 증가와 함께 칩 크기가 증가함에 따라 배선의 길이가 증가하므로 배선의 Parasitic Capacitance와 Parasitic Resistance가 더욱 증가한다. 이러한 배선을 V_{DD} 와 0V의 진폭을 갖는 논리 Level로 Full Swing하게 된

다면 총방전용 전력 소모가 증가하며 한편의 길이가 3 Cm 이상이 되리라 예상되는 4 Giga bit DRAM에서는 감지 증폭기 동작 전류만이 아니라 이러한 전류 성분의 비율이 더욱 증가하리라 예상된다.



(a)



(b)

그림 1. (a) DRAM의 동작 Mode에 따른 전류
(b) Read 동작의 전류 성분 분석

Fig. 1. (a) Operating current of DRAM functional mode (b) The current components during read operation.

CMOS DRAM에서는 전류의 소모가 CV_{DD}^2f 로 표현된다. 여기서 C는 각종 노드의 Capacitance 성분이며 V_{DD} 는 논리 Level Swing 전압, f는 동작 주파수이다. 따라서 저전력 소모를 위해서는 Capacitance 값의 감소, 동작 주파수의 감소 및 동작 전압의 감소가 중요하다. 특히 전압의 감소는 V_{DD}^2 에 비례하는 전력 소모를 가져오므로 더욱 효과적임을 알 수 있다. 한편 CMOS의 전류는 MOS의 문턱 전압을 V_T 라 하면 $(V_{DD}-V_T)^2$ 에 비례하며 $i=C(dV/dt)$ 의 식으로부터 CMOS의 동작 시간은

$$t = \int C \frac{dV}{i} dt = K \frac{V_{DD}}{(V_{DD} - V_T)^2}$$

로 표시될 수 있다. 여기서 K 는 비례 상수이다. 따라서 외부 전원 전압 V_{DD} 의 감소에 맞추어 V_T 도 감소시켜 $V_{DD}-V_T$ 를 일정하게 유지하여야 동작 시간의 증가를 막을 수 있다. 즉 V_{DD} 를 종래의 5V에서 3.3V로 그리고 2.5V 또는 1.5V로 감소시킴에 따라 MOS의 V_T 도 동일한 비례율로 감소시켜야 MOS의 크기의 증가없이도 동일한 동작 속도를 얻을 수가 있게 된다. 이때 문제로 되는 것이 Subthreshold 전류에 의한 DC 전류의 증가이다. Subthreshold 전류는 MOS의 폭을 W 라 하고 게이트-소스전압을 V_{GS} 라 하면 $I = I_0 W \exp | (V_{GS}-V_T)/n |$ 의 식으로 증가한다.^[14] 여기서 I_0 는 V_T 를 정의하기 위한 기준 전류이며 n 은 상수로 대략 1.2의 값을 갖는다. 따라서 $V_{GS}=0$ 이어도 상당량의 전류가 MOS의 드레인-소스를 통해 흐를 수 있으며 $W=16\mu\text{m}$ 나 되는 4 Giga bit DRAM 급에서는 동작 전류보다도 훨씬 큰 DC 전류가 Subthreshold 전류로서 흐를 수 있다^[13, 14].

4 Giga Bit급 DRAM에서는 이와 같이 종래에는 큰 비율을 차지하지 못하였던 데이터 경로 전류 성분과 DC 전류 성분 등이 상대적으로 큰 비중을 차지하므로 이의 증가를 억제하는 방법이 절실히 필요하다. 본 논문에서는 이러한 전류를 감소하는 방법에 대해 고찰하여 보고자하며 제 II절에서는 Subthreshold 전류를 감소하기 위한 새로운 고속 CMOS 회로를 다루며 III절에서는 Sensing 전류를 감소시키는 DRAM 구조에 대해 살펴보기로 한다. IV절에서는 Pipeline 구조를 채택하여 데이터 경로의 전류를 감소시키는 방법을 살펴보고 V절에서는 256M 급 또는 초기 1 Giga 급의 MOS parameter를 통해 앞서의 절에서 제안되었던 구조를 채택한 DRAM을 설계한 뒤 SPICE Simulation을 통해 저전력 동작을 증명하며 VI절에서 논문의 결론을 맺고자 한다.

II. 고속 Dual VT CMOS 회로

Giga bit급 DRAM에서는 전력 소모를 감소시키고 소자의 수명을 향상시키기 위해 1.5V 또는 그 이하의 전원 전압이 사용되리라 예상된다. 이러한 저전압 하에서도 DRAM의 동작속도를 고속으로 유지하기 위해서는 MOS의 문턱 전압도 함께 감소되어야만 한다.^[11] 이때 문제로 되는 Subthreshold 전류의 증가를 방지하기 위해 여러 가지 연구가 진행되었다. 소스단의 전

압을 변화시켜 MOS가 논리 "L"에서 NMOS의 게이트-소스전압이 "-값을 갖도록 하는 회로, 출력 단자에 Cross Coupled Latch를 부착하여 논리값을 저장하는 회로 및 Dual V_T CMOS 논리 회로 등이 그것이다^[13-15].

Dual V_T CMOS 논리 회로는 DRAM 회로와 같이 모든 노드에 대한 대기 상태를 미리 알고 있는 경우에 사용되는 것으로 High V_T MOS와 Low V_T MOS 두가지를 사용한다. Inverter회로의 경우 크기가 큰 High V_T MOS를 이용하여 대기 상태의 노드로부터 V_{DD} 나 접지로 Subthreshold 전류가 흐르는 것을 방지한다. 크기가 작은 Low V_T 의 반대 극성 MOS를 이용하여 대기 전압을 출력에 공급한다. 이러한 회로 방식을 통하여 Subthreshold 전류를 감소할 뿐 아니라 대기상태에서 Active 상태로의 천이 시간도 감소시킬 수가 있다. 하지만 이 회로 방식에서는 PMOS/NMOS의 크기비가 충전 방식과 달라 다이내믹 회로에서 반드시 필요한 Precharge 동작, 즉 Active에서 대기로의 천이가 상대적으로 증가하는 것이 문제점으로 지적되고 있다. 따라서 양방향의 천이를 모두 고려하여 최적화된 Low V_T MOS 대 High V_T MOS의 크기비 v 를 구해보면 약 0.26의 값을 얻는다. 그러나 이 값을 이용하여 회로를 설계하면 단순히 Low V_T MOS만을 사용한 논리회로에 비해 동작속도가 현저하게 낮아져 문제로 된다.^[14] 본 논문에서는 Low V_T MOS만을 사용하는 회로의 장점인 고속성과, Dual V_T Logic의 장점인 적은 Subthreshold 전류특성만을 추출하여 새로운 Dual V_T CMOS 논리 회로를 제안하고 그 특성을 살펴보려 한다.

Dual V_T 논리에서는 입력 전압이 변화하면 그 동안 대기 전압을 공급하고 있던 크기가 작은 Low V_T MOS가 OFF되고 반대 극성의 크기가 큰 High V_T MOS가 ON되어 출력이 Active상태로 된다. 따라서 앞단의 Inverter는 다음 단의 Pull-Up MOS와 Pull-Down MOS의 게이트 Capacitance를 모두 충전하여야 한다. 하지만 본 논문에서 제안하는 방식에서는 다음 단의 두 MOS 중 큰 V_T 를 갖는 MOS 1개만이 충전되므로 앞단의 부하가 감소하여 고속 동작이 가능하게 된다. 그림 2는 Self-Timed Dual V_T Logic 회로의 회로 구성을 보여준다. 하강 Edge 검출기 회로를 이용하여 Active Low인 외부 입력 신호를 짧은

폭을 갖는 음 Pulse로 변환한다. Edge 검출기는 회로 전체의 동작 속도에는 무관하므로 Subthreshold 전류를 감소하기 위해 V_T 가 큰 PMOS와 NMOS만을 이용하여 설계된다. 이제 회로 동작을 설명하기 위해 그림 2의 회로의 모든 노드들이 대기상태로 초기화되어 있다고 가정하자. 그림 2에서 홀수로 표시된 노드는 "H", 짝수로 표시된 노드는 "L"의 초기 상태를 갖는다. 초단의 입력 노드의 전압이 하강하면 P_1 이 ON되어 그 출력이 Active 상태인 "H"로 상승한다. 이러한 상승 전압에 의해 다음 단의 N_2 가 ON되며 그 출력은 Active 상태 "L"로 하강한다. 각각의 신호는 MOS 1개만의 게이트를 충전하기만 하면 되므로 각각의 Inverter들이 연속해서 고속으로 Active상태로 천이한다.

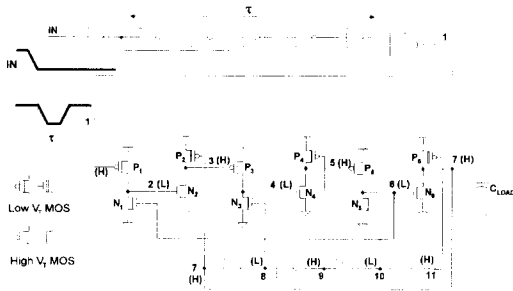


그림 2. Dual VT Self-Timed CMOS Logic 회로
Fig. 2. Circuit diagram of the dual VT self-timed CMOS logic.

이 회로에서는 별도의 신호 경로를 통해 출력 노드가 Reset되므로 높은 V_T 를 갖는 MOS의 크기가 대기 상태로 복귀 시간과 무관하게 되어 MOS의 크기를 임의로 크게 설정할 수 있다. 또한 각 Inverter단이 실효적으로는 1개의 MOS만으로 구성되어 있으므로 종래와 같이 PMOS와 NMOS가 직렬로 연결된 CMOS 논리 회로에 비해 전원 전압을 감소시킬 수 있다.

Active상태에서 대기상태로의 복귀는 반대 극성의 Low V_T MOS(또는 Resetting MOS라고도 부른다)를 통해 실현한다. 현재의 Inverter보다 4번째 뒤의 Inverter의 출력이 Active 상태로 천이하는 전압을 이용하여 Resetting MOS가 동작한다. 예를 들어 노드 6의 활성화 천이는 다음 단의 High V_T NMOS N_6 를 ON으로 할 뿐 아니라 Resetting용 Low V_T NMOS N_1 도 ON으로 한다. 따라서 각 노드는 약 4

Inverter Delay 시간 후에 자동적으로 대기 상태로 Reset된다. 일반적으로 Inverter의 Fan-Out이 1보다 큰 수, 대략 3을 사용하므로 n 번째 단의 Resetting MOS와 $n+4$ 번째 단의 Active MOS와의 크기비, 예를 들어 N_1/N_6 는 매우 작아서 거의 무시할 수 있을 정도로 된다. 더욱이 Low V_T MOS를 이용한다면 High V_T MOS보다 더욱 작은 트랜지스터로도 동일한 전류 구동능력을 갖게 할 수 있으므로 Active용 High V_T MOS의 크기와 Resetting용 Low V_T MOS의 크기비를 크게 할 수 있으며 Active 동작과 Resetting 동작을 각각 따로이 최적화하는 것이 가능하다.

Self-Timed Dual V_T Logic의 동작 속도와 누설 전류를 SPICE Simulation을 통해 조사하여 보았다. 이때 사용된 V_T 값은 NMOS의 경우 High가 0.65V, Low가 0.20V이며 PMOS의 경우 High가 -0.70V, Low가 -0.22V이다. Subthreshold Swing값은 85°C에서 NMOS의 경우 110mV/dec, PMOS의 경우 120mV/dec를 갖는다. 회로의 부하로는 3389 μ m 크기의 MOS를 Capacitor형태로 연결하여 사용하였다. 비교를 위해 동일한 조건에서 Fan-Out이 3으로 고정된 Low V_T 만으로 구성된 회로와 종래의 Dual V_T CMOS회로로서 $v=0.26$ 을 갖는 회로를 Simulation하였다.

한편 본 논문의 Self-Timed Dual V_T Logic에서는 High V_T NMOS의 Fan-Out을 6, High V_T PMOS의 Fan-Out을 3으로 설정하여 Pulse의 Rising Time과 Falling Time을 일치시켰다. 또한 회로의 전체 면적이 3가지 방식의 회로에서 모두 일치 되도록 Design 하였다.

그림 3은 전원 전압 1.5V에서 5 Inverter Chain을 통해 Pulse가 진행되는 것을 Simulation한 파형이다. 입력 신호는 폭이 약 0.8ns인 Pulse로 변환되어 입력 되는데 출력까지의 전파 지연 시간은 1.3ns이다. 한편 Low V_T MOS만으로 구성된 회로의 경우 전파 지연 시간이 1.4ns가 되었으며 종래의 Dual V_T Logic의 경우 약 2.2ns의 전파 지연 시간을 보였다. 또한 본 논문에서 제안한 회로는 20.2pA의 누설전류를 가져 Dual V_T 회로와 동일한 값을 보였지만 Low V_T MOS만으로 구성된 회로에서는 27.5nA나 되는 큰 값을 가져 본 회로를 채용한 경우 누설 전류가 약 1360배 감소되었음을 확인할 수 있었다.

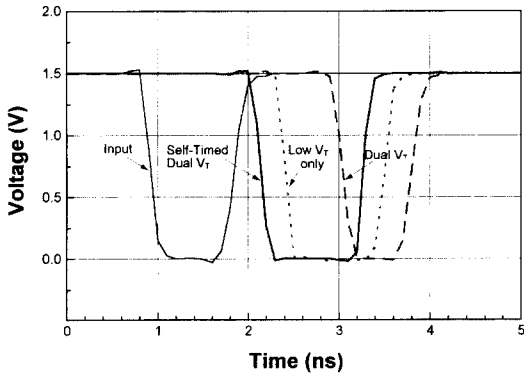


그림 3. 각종 논리회로와 Self-Timed 논리회로의 동작 파형 Simulation의 비교

Fig. 3. Comparison among the self-timed and other logic circuits.

III. 8 Bit Prefetch 저전압 DRAM 구조

Row Address에 의해 Word Line이 선택되고 Cell 데이터가 Bit Line에 실려 감지 증폭기를 통해 증폭되는 Row Address 경로 동작은 다루어지는 신호가 수 100mV 미만의 작은 전압을 가지며 Word Line 이나 Address 배선 등의 Parasitic Capacitance와 Resistance가 커서 고속 동작에도 한계가 있다. t_{RAC} 이 50ns인 64M DRAM의 경우 Row Address 경로 시간에 해당하는 t_{RCD}/RAS to $/CAS$ delay time)가 20ns나 되며 DRAM의 세대가 진전되어도 그리 큰 변화 없이 항상 일정한 시간이 소요된다. 저전력 소모를 위해 외부 전원 전압을 낮출수록 Row Address 경로의 속도 지연이 증가하며 Cell 신호도 감소하여 이의 정확한 감지 및 증폭이 어렵다.

Synchronous DRAM(SDRAM)에서는 고속 동작을 위해 Burst Read 동작 즉, 동일한 Word Line을 갖는 연속된 4 bit 또는 8 bit의 데이터를 순차적으로 읽어내는 동작을 이용한다.^[6] 종래의 SDRAM에서는 Burst Read가 시작되는 Column Address만을 받아들인 뒤 내부 Column Address Counter에 의해 생성된 연속된 Column Address를 매번 Decode하여 데이터를 읽어 낸다^[16]. 그런데 이 때의 동작 시간이 상대적으로 길기 때문에 동작 시간을 단축시키고자 2 bit Prefetch 방식 또는 Full bit Prefetch 방식이 사용되었다^[17, 18].

본 논문에서는 이러한 고속 동작용 Prefetch 개념

을 감지 증폭기의 동작 전류를 감소시키는데 이용하여 저전력 DRAM 구조를 개발하고 그 효율성을 시뮬레이션을 통해 증명하였다. 그림 4는 8 bit Prefetch 방식의 개념도이다. 외부의 Address 단자를 통해 입력된 X-Address에 의해 Word Line이 선택되면 동일 Word Line 상의 Cell 데이터들은 모두 감지 증폭기에 의해 증폭된다. 이후 입력되는 Column Address를 선행 값으로 하는 연속된 8 bit의 데이터가 I/O 감지 증폭기에 의해 증폭된 뒤 8개의 Latch에 저장된다. 따라서 두 번째 데이터부터는 Column Decoder에 의한 선택이나 I/O 감지 증폭기의 증폭 동작에 필요한 시간을 생략하고 Latch로부터 곧바로 데이터를 읽을 수 있으므로 이에 해당하는 시간을 절약한 것이라 생각할 수 있다. 한편 SDRAM이 동일한 동작 속도를 한다고 한다면 8 bit Prefetch 방식에 의해 Column Decode 및 I/O 감지 증폭기의 동작 시간이 8배의 여유를 확보한 것이 된다. 따라서 전원 전압이 낮아져 동작 속도가 1/8로 하강하여도 동일한 동작 속도를 확보할 수 있다.

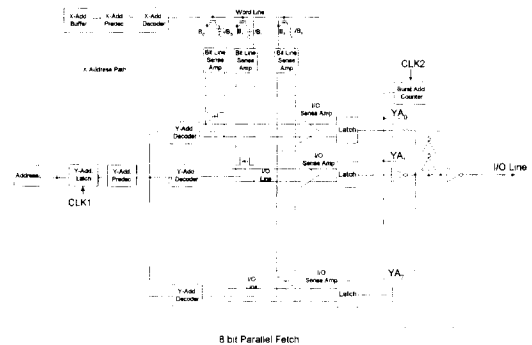


그림 4. 8 bit Prefetch 방식의 회로도

Fig. 4. Circuit diagram of the 8 bit prefetch method.

만일 종래의 방식의 SDRAM을 외부 전원 전압 V_1 에서 동작한다고 하고 8 bit Prefetch 방식의 SDRAM을 V_2 라는 전압에서 동작시켰을 때 종래형과 동작 속도가 일치한다고 한다면 각각의 경우 Access Time인 t_1 과 t_2 는 다음과 같은 식을 만족하여야 한다.

$$t_2 = K \frac{V_2}{(V_2 - V_T)^2} = 8 t_1 = 8K \frac{V_1}{(V_1 - V_T)^2}$$

여기서 K는 비례 상수이다. 이 식으로부터 V_2/V_1 을

구하면

$$\frac{V_2}{V_1} = (\beta + \frac{\gamma}{2}) + \sqrt{\beta\gamma + \gamma^2/4}$$

가 된다. β 와 γ 는 각각 V_T/V_1 및 $(1-\beta)^2/8$ 이다. 여기에 $V_T=0.5V$ $V_1=3.3V$ 를 대입하면 $V_2 = 0.36V_1 = 1.2V$ 를 얻는다.

본 논문에서 다루는 4 Giga bit DRAM은 Cell Array가 2K x 4K Cell로 구성되어 있으며 4K 번에 Bit Line 감지 증폭기가 배치되어 있는 구조이다. Bit Line 감지 증폭기에 의해 증폭된 데이터 중 8개가 동시에 Column Decoder신호에 의해 선택된 뒤 Global I/O Line을 통해 64M Block의 중앙에 위치한 Register Array로 전송된다. 그림 5는 Bit Line 감지 증폭기와 Register회로의 개략도이다. Column Address Decoder에 의해 Bit Line의 데이터가 DB Line으로 전달되고 이 신호가 다시 Global I/O Line을 통해 I/O 감지 증폭기로 입력된다. I/O 감지 증폭기에 의해 증폭된 전압은 일단 Latch되어 Clock에 의해 제어되는 Register에 저장된다. 그림 6은 Column Decoder와 감지 증폭기의 배치 및 Register Array의 배치도이다.

두 개의 8M bit Array가 1개의 Column Decoder와 감지 증폭기 Array를 공유하며 이 둘을 합한 16M bit Array 2개가 1개의 Register Array를 공유하도록 하였다.

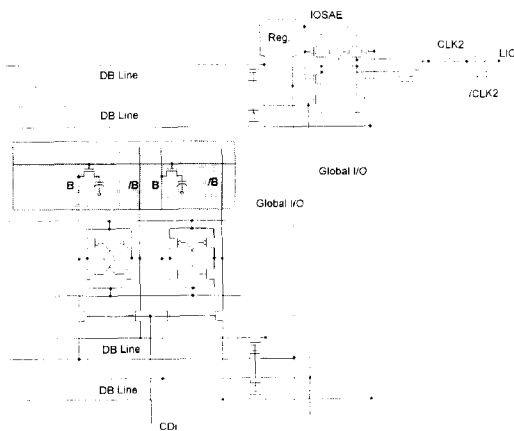


그림 5. Bit Line 감지 증폭기와 Register 회로의 개념도

Fig. 5. Conceptual diagram of the bit line sense amp. and the register circuit.

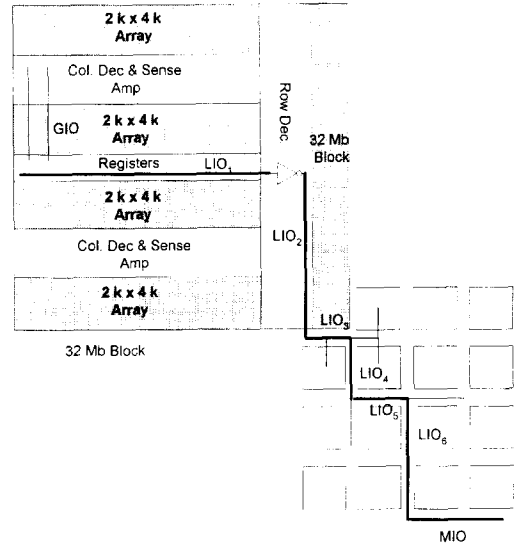


그림 6. Column Decoder, 감지 증폭기 및 Register Array의 배치도

Fig. 6. Floor plan of column decoders, sense amp. and register array.

IV. 저전압 고속 Pipeline 데이터 경로

SDRAM에서는 고속 Burst Read 동작을 위해 데이터 경로를 Pipeline으로 구성한다^[16, 18-21]. 4 Giga급 이상의 대용량 메모리에서는 칩 Size가 커져 Bus Delay가 증가하며 이에 의한 전력 소모도 증가한다. 따라서 그림 7과 같이 데이터 경로를 계층화하고 Latch를 삽입하여 Pipeline 구조를 채택함으로써 고속 동작을 유지할 수 있다^[6].

한편 이러한 Pipeline 구조가 저전력화에서도 유리하다. N-Stage Pipeline을 구성하면 데이터 경로가 1/N으로 감소한 것이 되므로 Access Time t_3 는 t_1/N 으로 감소한다. 이때 Pipeline 데이터 경로를 종래와 동일한 Access Time t_1 을 갖도록 설계한다면 결국 신호의 전송 속도를 종래의 1/N로 감소시켜도 동일한 동작 속도를 얻는다.

따라서 Pipeline경로의 외부 전원 전압 V_3 를 종래의 전원 전압 V_1 보다 낮게 설정할 수가 있다. 이때의 Access Time과 전원 전압은 III절의 식에서 V_2 대신 V_3 를 대입한 경우로 된다.

데이터 경로의 Capacitance를 C_1 이라 하고 Pipeline을 구성하기 위해 삽입된 Latch의 Capacitance를 C_{reg} 라 한다면 소모 전력 P_3 는

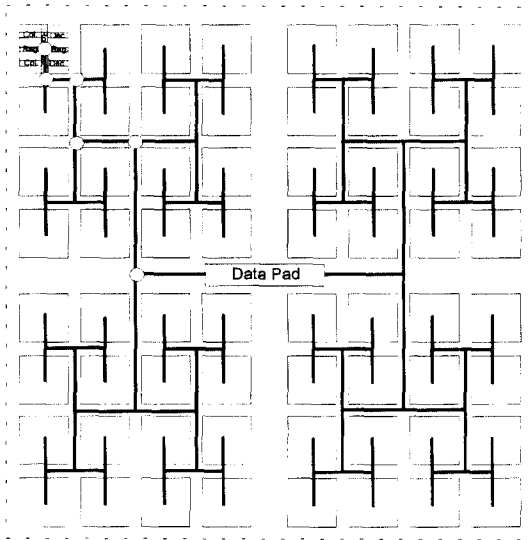


그림 7. 본 연구에서 가정한 4 Giga Bit DRAM의 Pipelined 데이터 경로
 Fig. 7. Pipelined data path of 4 giga bit DRAM assumed in this study.

$$P_3 = [C_1 + (N-1) C_{reg}] V_3^2 f$$

로 표현된다. 이와 Pipeline을 채택하지 않은 종래의 데이터 경로의 소모 전력 P_1 을 비교하면

$$\frac{P_3}{P_1} = \frac{[C_1 + (N-1)C_{reg}] V_3^2 f}{C_1 V_1^2 f} = [1 + \eta(N-1)] \frac{V_3^2}{V_1^2} \cong \left(\frac{V_2}{V_1}\right)^2 \cong \frac{1}{N^2}$$

이 된다. 여기서 $\eta = C_{reg}/C_1$ 로 매우 작은 값을 갖는다고 가정하였다.

그림 8은 본 논문에서 가정하고 있는 4 Giga bit DRAM의 Cell Array Block 배치도이다. 64M bit Array Block 64개로 구성되어 있으며 0.15 μ m Design Rule을 사용하였다고 가정하였다. 또한 64M bit Cell Array Block의 크기를 대략 0.60 cm x 0.3 cm로, 전체 칩의 크기를 2.6 cm x 5.2 cm로 가정하였다. III절에서 설명한 Prefetch 동작에 의해 Register Array에 저장되어 있던 데이터가 Column Address Counter 신호에 의해 순차적으로 I/O Line으로 전송된다. 데이터 Pad의 위치를 칩 중앙부로 가정하였으므로 Register Array부터 Pad까지의 I/O Line의 길이는 대략 3 cm~4 cm 정도로 된다. I/O Line이 Metal 2나 Metal 3로 구성된다 하고 Metal 폭이 0.3

μ m인 경우라면 전체 길이는 1×10^5 \square 가 된다. $R_s = 0.4 \Omega / \square$ 이라 가정한다면 저항 값이 40K Ω 가 되며 Si기판으로부터의 거리가 1 μ m라면 Parasitic Capacitance는 0.8pF가 된다. 따라서 RC Delay가 32ns나 되어 고속화에도 장애가 되며 전체 Line을 전원 전압 Level까지 충방전하는데 큰 전력 소모를 가져온다. 따라서 전체 I/O Line을 일정한 길이를 갖는 다수의 Line으로 분할한 뒤 각 경계마다 Latch를 삽입하여 Pipeline을 구성하는 것이 유리하다.

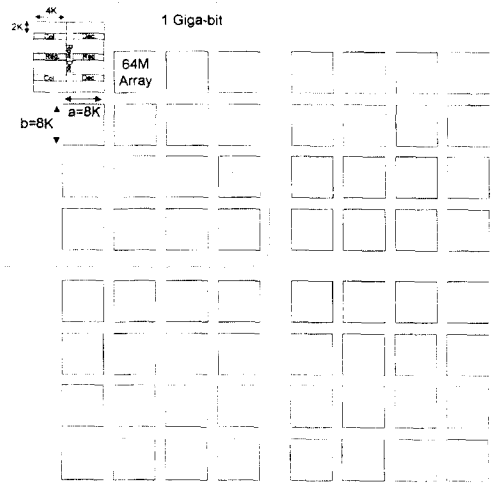


그림 8. 본 연구에서 가정한 4 Giga bit DRAM의 Cell Array Block 배치도
 Fig. 8. Cell array block arrangement of 4 giga bit DRAM assumed in this study.

각 Latch는 Clock 신호에 의해 제어되며 고속 전송을 위해서는 데이터가 Latch에 저장된 뒤 곧바로 데이터 Line을 Precharge시켜 다음 데이터의 전송을 준비하여야 한다. 또한 Pipeline의 각 단은 동일한 전송 지연 시간을 갖도록 분할되어야 하며 이중 가장 긴 Delay 시간을 갖는 단의 전송 속도에 의해 Pipeline Path 제어용 Clock의 속도가 결정된다. 실제의 회로 설계 시에는 전체 Delay를 미리 예상하고 이를 균등하게 분배한다는 것이 쉽지 않다. 또한 다음 동작을 위해 각단을 Precharge시키는데 시간이 소요되고 이의 제어도 복잡하며 삽입된 Latch에 의한 전체 경로의 Delay의 증가도 무시할 수 없다. 이러한 문제점을 해결하기 위해 II절에서 제안된 Self-Timed CMOS Logic을 이용한 Wave Pipeline 개념을 도입하였다^{16, 22)}. Wave Pipeline은 데이터 경로에 Latch를 삽입

하지 않고 데이터를 Pulse 형태로 하여 다수의 Pulse를 시분할하여 전송하는 방식이다. 이 방식에 의하면 Latch의 삽입이 필요치 않으므로 이의 제어 회로와 Latch에 의한 Capacitance부하, 예로 C_{reg} 등이 제거되어 고속 전송이 가능하게 된다. 이를 구현하기 위해서는 단지 작은 Pulse폭을 갖는 데이터의 발생회로와 이를 전송하는 Repeater만 필요하다.

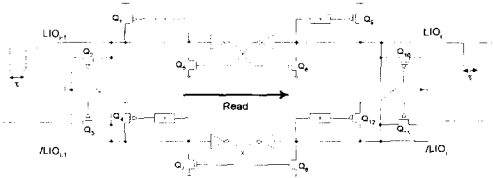


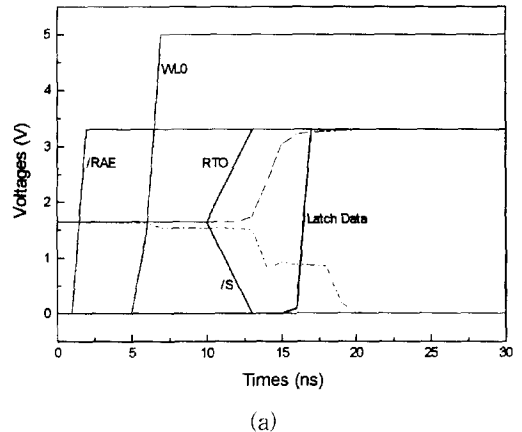
그림 9. Self-Timed Logic 회로를 이용한 I/O Line 구동 회로
Fig. 9. I/O line driving circuit with the self-timed logic.

그림 9의 회로는 Self-Timed Logic 회로를 이용한 I/O Line 구동회로이다. V_{DD} 로 Precharge되어 있던 I/O Line에 외부 Clock에 동기시켜 음 Pulse로 데이터를 전송한다. LIO_i 에 데이터가 실렸다면 데이터 신호의 Falling Edge에 의해 Q_6 가 ON되어 LIO_i 가 "L"로 하강한다. τ 시간 뒤 데이터 Pulse의 Rising Edge에 의해 LIO_{i-1} 이 다시 "H"가 되면 Q_6 는 OFF이 되며 "L"이었던 LIO_i 전압이 Delay를 거쳐 Q_9 의 게이트전압을 방전함에 따라 LIO_i 가 "H"로 된다. Q_2, Q_3 및 Q_{10}, Q_{12} 는 LIO와 /LIO에 확실한 Complementary 신호가 전송시키기 위해 설치하였다. Read 시에는 화살표 방향으로 데이터가 진행되지만 Write 시에는 반대 방향으로 데이터가 진행되는 Bi-Directional Driver회로이다. 이러한 회로를 데이터 경로 상의 전송 지연이 τ 정도 되는 지점마다 삽입하여 Pulse폭의 증가를 막고 고속 전송이 되도록 한다. 전체 경로의 지연 시간을 t_{total} 이라 한다면 이론적으로 최대 Pipeline 단수 N 은 t_{total}/τ 로 되기 때문에 Pulse폭을 조정함에 따라 Pipeline의 단수를 선택할 수 있으며 결국 동일한 동작 속도를 얻을 수 있는 전원 전압 Level을 선정할 수 있다.

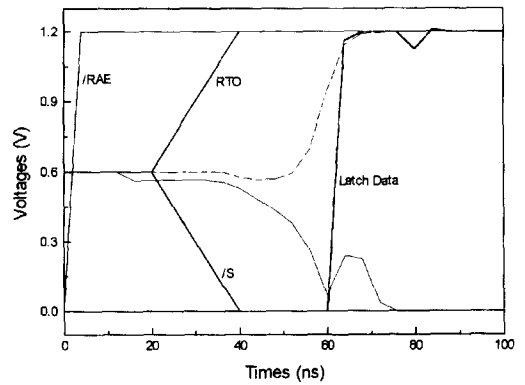
V. Simulation 결과

본 논문에서 제안된 회로와 Architecture들의 유용

성을 검증하기 위해 256M DRAM의 제작에 사용되었던 MOS의 Device Parameter를 이용하여 SPICE Simulation을 통해 검증하여 보았다. 새로운 Architecture를 채택하지 않은 종래형의 SDRAM을 3.3V에서 Simulation하였고 이의 동작 속도를 새로운 구조의 선정에 기준값으로 사용하였다. 새로운 Architecture를 채택한 SDRAM를 설계한 뒤 전원 전압을 낮추면서 기존 SDRAM과 동일한 동작 속도를 갖는 전원 전압 Level을 구하였다.



(a)



(b)

그림 10. (a) 3.3V 단일 감지 증폭기 선택 방식의 내부 파형 (b) 1.2V 8 bit Prefetch 방식의 내부 파형
Fig. 10. (a) Internal waveform of single sense amp. method with 3.3V power supply (b) Internal waveform of 8 bit prefetch method with 1.2V power supply.

그림 10(a)는 전원 전압이 3.3V일 때 /RAS신호 입력부터 I/O Line에 데이터가 실릴 때까지의 각 내부 회로의 Simulation된 파형들이다. Bit Line의

Capacitance는 270fF, Cell Capacitance는 20fF로 가정하였다. Word Line의 전압이 5V라 하면 약 6ns 시간 후 Cell Capacitance와 Bit Line간의 Charge Sharing에 의해 Bit Line의 전압이 변화하며 이 전압이 안정되기를 기다려 약 10ns 시에 Bit Line 감지 증폭기의 활성화 신호 RTO, /S를 ON으로 하면 Cell 데이터가 증폭된다. Bit Line과 /Bit Line의 전압 차이가 약 1.5V로 되었을 때, 즉 14ns 시에 Column Decoder출력을 인가하여 Bit Line의 전압을 I/O Line으로 전송한다. 이후 I/O 감지 증폭기에 의해 I/O Line의 전압이 증폭되고 약 16ns에 Register에 Cell 데이터가 Latch된다. 한편 전원 전압이 1.2V로 되면 그림 10(b)와 같이 모든 동작들이 느려진다. 감지 증폭기의 활성화가 20ns에서 일어나며 약 60ns에서 Column Decoder출력이 연결되어 63ns에 데이터가 Latch된다. 즉 전압을 1.2V로 낮추고 8 bit Prefetch를 사용한 경우 3.3V보다 4배 늦게 데이터가 Latch됨을 뜻한다. 그림 11에서와 같이 이에 필요한 전류의 변화를 보면 3.3V에서는 30mA의 전류가 소모되었지만 1.2V, 8 bit Prefetch 방식에서는 단지 0.7mA만이 소모되므로 97.7%의 감소를 가져왔다. 따라서 전원 전압을 감소하는 방법이 Bit Line 감지 증폭기전류를 감소시켜 저전력 DRAM을 구현하는데 매우 효과적임을 알 수 있었다.

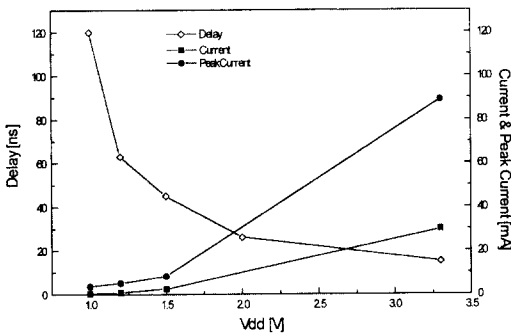
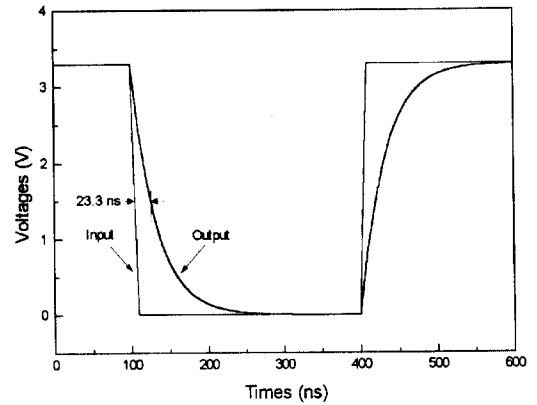
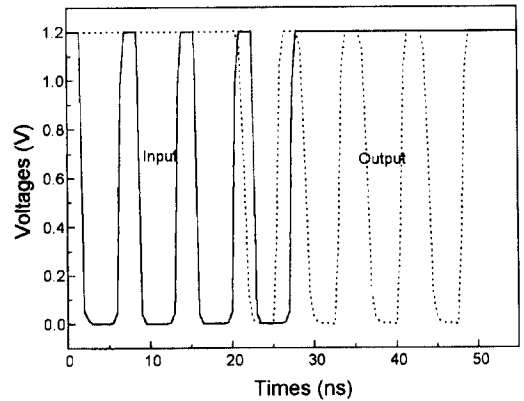


그림 11. 전원 전압에 따른 8 bit Prefetch 방식의 동작 속도와 소모 전류
 Fig. 11. Speed and current of the 8 bit prefetch method with supply voltage variation.

8 bit Prefetch 방식에 의해 Sensing되어 저장된 데이터를 Pipelined 데이터 Path를 통해 외부로 전송시키는 경우의 동작 속도와 전원 전압의 관계를 조사하였다.



(a)



(b)

그림 12. (a) 3.3V에서 동작하는 종래형 데이터 경로의 전압 파형 (b) 1.2V 4 Stage Wave Pipeline 방식의 전압 파형

Fig. 12. (a) Voltage waveform of conventional data path with 3.3V supply voltage (b) Voltage waveform of 4 stage wave pipeline method with 1.2V supply voltage.

그림 12(a)는 40KΩ의 저항과 0.8pF의 용량을 갖는 배선에서의 3.3V 신호 전압의 파형을 보여 준다. Pipeline을 채택하지 않았으므로 신호의 전파 지연은 23.5ns가 되었다. 이 경로를 4단으로 된 Pipeline 경로로 구성하였다. Wave pipeline을 형성하므로 Latch의 삽입 없이 단지 4개의 Short Pulse만을 전송하면 된다. 따라서 Pulse폭을 5ns로 하여 Pulse 간의 중첩을 없애고 경로 전체에서 5ns의 Pulse Width가 유지 되도록 그림 9의 Repeater를 삽입하여 Wave Pipeline 경로를 구성하였다. 이 경우 전원 전압을 낮추어 데이터 경로의 동작 시간이 종래의 23.3ns의 4배인 93ns 정도로 되어도 매 23.3ns 마다 데이터 Pulse

가 외부로 출력되므로 외부에서는 동일한 속도로 동작하고 있는 것으로 보인다. 한편 그림 12(b)에서와 같이 Repeater회로를 삽입하고 전압을 1.2V로 낮추어도 4개의 신호를 22ns 내에 전송할 수 있게 되었다. 즉 전원전압을 낮추었음에도 데이터 경로의 신호 전달 속도가 3.3V의 경우보다 약 4배 증가한 것으로 된다. 이때의 전류는 그림 13에서 보듯이 2.5mA로서 3.3V 값의 11.4%로 감소하였다. 따라서 4 Giga bit급 대용량 DRAM의 데이터 경로에서 동작 속도를 높이고 전류 소모를 감소하기 위해 저전압 Pipeline 구조가 유용함을 확인할 수 있었다.

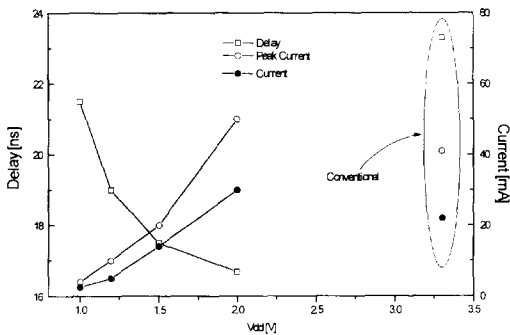


그림 13. 4 Stage Wave Pipeline 방식의 전압 대 Delay Time과 소모 전류 특성

Fig. 13. Delay time and current of the 4 stage wave pipeline with supply voltage variation.

VI. 결 론

Giga bit급 Synchronous DRAM에서 저전력화에 적합한 CMOS 회로와 SDRAM의 구조에 대해 연구하였다. 4 Giga bit 급 이상의 DRAM에서는 감지 전류, 데이터 경로에서의 전류, DC 전류가 큰 비중을 차지함을 밝혔다. 특히 DC 전류는 Subthreshold 누설전류 성분이 주가 되므로 이를 감소하기 위한 새로운 CMOS Logic 회로를 제안하였다. 이 Logic으로 구현된 회로는 Low V_T MOS만으로 구성된 회로보다 동작 속도가 빠르며 누설 전류는 Low V_T 만으로 설계된 회로에 비해 1/1360으로 감소하였다. 4 Giga bit DRAM의 감지 전류를 감소하기 위해 종래 고속 동작용으로 사용되던 8 bit Prefetch 방식을 채택하였다. 연속된 8 bit 전체의 전송 시간은 종래의 3.3V 형과 일치시킨다면 동작 전압을 1.2V까지 하강할 수가

있으며 이때의 동작 전류, 즉 Sensing 전류는 0.7mA에 지나지 않아 3.3V 때 전류 30mA의 2.3%에 불과하였다. 4 Giga bit 급 데이터 경로는 3 cm정도로 길어서 RC Delay가 크며 전체 배선을 전원 전압 Level로 충전전하는데 큰 전력 소모를 가져온다. 데이터 경로의 동작 전압을 1.2V로 낮추고 Pipeline을 채택하여 저전압에서도 고속 동작이 가능하도록 하였다. 데이터 경로에 Repeater회로를 삽입함에 따라 전압을 1.2V로 낮추었음에도 연속된 4개의 데이터를 종래의 3.3V형에서의 1개 데이터의 동작 시간에 해당하는 22ns 내에 전송할 수 있게 되었다. 즉 데이터 경로의 신호 전달 속도가 3.3V의 경우보다 약 4배 증가한 것으로 되며 동작 전류는 2.5mA에 지나지 않아 3.3V 시의 값의 11.4%로 감소하였다.

Dual V_T Self-timed CMOS logic을 채용하고 Prefetch 방식을 더욱 활용한다면 고속, 저전압 DRAM의 설계가 용이하게 되리라 예상되며 Giga bit 급의 큰 면적을 갖는 DRAM에서도 Pipeline 데이터 Path를 구성함으로써 Cycle Time이 6ns 미만인 고속 구동을 달성할 수 있으리라 예측된다.

참 고 문 헌

- [1] K. Itoh, et al, "Limitations and challenges of multi-gigabit DRAM circuits", VLSI Cir. Symp., Digest of Tech. Papers, pp. 2-7, 1996.
- [2] T. Sugibayashi, et al, "A 1Gb DRAM for File Applications", ISSCC, Digest of Tech. Papers, pp. 254-256, Feb. 1995.
- [3] M. Horiguchi, et al, "An Experimental 220MHz 1Gb DRAM", ISSCC, Digest of Tech. Papers, pp. 252-253, Feb. 1995.
- [4] Y. Nitta, et al, "A 1.6GB/s Data-Rate 1Gb Synchronous DRAM with Hierarchical Square-Shaped Memory Block and Distributed Bank Architecture", ISSCC, Digest of Tech. Papers, pp. 376-377, Feb. 1996.
- [5] J.H. Yoo, et al, "A 32-Bank 1Gb DRAM with 1GB/s Bandwidth", ISSCC, Digest of Tech. Papers, pp. 378-379, Feb. 1996.
- [6] H. J. Yoo, et al, "A 150MHz 8-banks 256M Synchronous DRAM with the

- Wave Pipelining Method", ISSCC, Digest of Tech. Papers, pp. 250-251, Feb. 1995.
- [7] T. Hasegawa, et al, "An Experimental DRAM with a NAND Structured Cell", ISSCC, Digest of Tech. Papers, pp. 46-47, Feb. 1993.
- [8] G. Kitsukawa, et al, "256Mb DRAM Technologies for File Applications", ISSCC, Digest of Tech. Papers, pp. 48-49, Feb. 1993.
- [9] T. Sugibayashi, et al, "A 30ns 256Mb DRAM with Multi-Divided Array Structure", ISSCC, Digest of Tech. Papers, pp. 50-51, Feb. 1993.
- [10] M. Asakura, et al, "A 34ns 256Mb DRAM with Boosted Sense-Ground Scheme", ISSCC, Digest of Tech. Papers, pp. 140-141, Feb. 1994.
- [11] H. Kotani, et al, "A 256Mb DRAM with 100MHz Serial I/O Ports for Storage of Moving Pictures", ISSCC, Digest of Tech. Papers, pp. 142-143, Feb. 1994.
- [12] S. Tanoi, et al, "A 32-Bank 256Mb DRAM with Cache and TAG", ISSCC, Digest of Tech. Papers, pp. 144-145, Feb. 1994.
- [13] T. Kawahara, et al, "Subthreshold Current Reduction for Decoded Driver by Self-Reverse Biasing", IEEE J. Solid-State Circuits, vol. 28, no. 11, pp. 1136-1144, 1993.
- [14] D. Takashima, et al, "Standby/Active Mode Logic for Sub-1-V Operating ULSI Memory", IEEE J. Solid-State Circuits, vol. 29, no. 4, pp. 441-446, 1994.
- [15] T. Sakada, et al, "Subthreshold current reduction circuits for multi-gigabit DRAM's," Symp. on VLSI Circuit Dig. Tech. Pap., pp. 45-46, Jun., 1993.
- [16] Y. Takai, et al, "250 Mbyte/sec Synchronous DRAM Using a 3-Stage pipelined Architecture", Symp. on VLSI Circuit Dig. Tech. Pap., pp. 59-60, Jun., 1993.
- [17] Y. Choi, et al, "16 Mbit Synchronous DRAM with 125 Mbyte/sec Data Rate", Symp. on VLSI Circuit Dig. Tech. Pap., pp. 65-66, Jun., 1993.
- [18] T. Sunaga, et al, "A Full Bit Prefetch Architecture for Synchronous DRAM's", IEEE J. Solid-State Cir., vol. 30, no. 9, pp. 998-1005, Sep. 1995.
- [19] A. Fujiwara, et al, "A 200 MHz 16 Mbit Synchronous DRAM with Block Access Mode", Symp. on VLSI Circuit Dig. Tech. Pap., pp. 79-80, Jun., 1994.
- [20] Y. Kodama, et al, "A 150MHz 4 Bank 64 Mbit SDRAM with Address Incrementing Pipeline Scheme", Symp. on VLSI Circuit Dig. Tech. Pap., pp. 81-82, Jun., 1994.
- [21] Y. Sakai, et al, "A Synchronous DRAM with New High-Speed I/O Lines Method for MultiMedia Age", IEICE Trans. Electron., vol. E78C, no. 7, pp. 782-788, Jul., 1995.
- [22] C. T. Gray, W. Liu, and R. K. Cavin III, *Wave Pipelining: Theory and CMOS Implementation*, Kluwer Academic Publishers, Norwell, MA, 1994.

저 자 소 개



柳 會 峻(正會員)

1983年 서울대학교 전자공학과(공학사). 1985年 한국과학기술원 전기·전자공학과(공학석사). 1988年 한국과학기술원 전기·전자공학과(공학박사). 1988年 ~ 1990年 미국 Bell Comm. Research 연구원. 1991年

~ 1995年 현대전자 반도체 연구소 DRAM 설계실장. 1995年 ~ 현재 강원대학교 전자공학과 조교수



李 正 雨(正會員)

1996年 2月 강원대학교 전자공학과 졸업, 공학사 학위 취득. 1996年 3月 ~ 현재 강원대학교 대학원 전자공학과 석사과정. 주관심 분야는 DRAM Architecture, 고속 CMOS Logic, System IC 등임

저전력 논리 회로 설계를 위한 커널에 바탕을 둔 precomputation 알고리즘

(A Kernel-based Precomputation Scheme for Low-power Design of Combinational Circuits)

崔益成*, 柳承鉉**, 黃善泳*

(Ick Sung Choi, Seung Hyun Ryu, and Sun Young Hwang)

요 약

본 논문에서는 저전력 조합 회로 설계를 위한 논리 합성 알고리즘을 제안하고 구현하였다. 제안된 알고리즘은 회로의 동작에 영향을 미치지 않는 불필요한 신호 전이를 제거하여 전력 소모를 감소시킨다. 제안된 알고리즘으로 주어진 회로를 precomputation 구조로 재구성하며, 이 중 prediction logic은 전이 확률을 효과적으로 줄일 수 있는 커널을 선택하여 사용한다. 실험 결과를 통하여 제안된 알고리즘을 이용하여 효율적으로 저전력 논리 합성 과정을 수행함을 보였다.

Abstract

In this paper, we present a logic synthesis algorithm for low power design of combinational circuits. The proposed algorithm reduces power dissipation by eliminating unnecessary signal transitions. The proposed algorithm restructures a given circuit by using a kernel as prediction logic in a precomputation-based scheme such that switching activity of circuit can be minimized. Experimental results show that the system is efficient for low power design of combinational circuits.

I. 서 론

VLSI 제조 및 설계 기술의 발달과 더불어 칩의 고성능 및 고집적화 경향과 함께, 최근에는 회로의 전력 소모에 관한 문제가 대두되어 저전력 설계에 대한 관심이 증가하고 있다. 점차 휴대용 컴퓨터 및 통신 시

스템의 수요가 급증하는 추세로, 고성능 회로의 발열 문제와 배터리 수명 등을 개선하기 위하여 회로의 소모 전력은 ASIC 설계에서 점차 중요한 설계 사양으로 자리잡아 가고 있으며, 다양한 레벨에서 저전력을 위한 최적화 과정을 거치게 된다. 저전력 설계 지원 환경의 일원으로 상위 수준 저전력 합성기 등이 개발되고 있으며, 이들 툴들을 이용하여 설계 및 시뮬레이션이 수행되고 있다^{[1][2]}.

CMOS 게이트로 구현된 VLSI 회로는 내부 노드 값의 전이에 의한 전력 소모가 전체 전력 소모량의 90% 이상을 차지하여 전이 활동을 예측하여 평균 소모 전력을 계산할 수 있으며^[3], 다양한 설계 계층에서 전이 활동을 줄여 소모 전력을 줄이는 설계 방법이 제안되었다. 저전력을 위한 논리 회로 설계는 주로

* 正會員, 西江大學校 電子工學科

(Dept. of Electronic Engineering Sogang Univ.)

** 正會員, LG 精密 研究所

(LG precision)

※ 본 연구는 서울대학교 반도체 연구소의 97년도 교육부 반도체 분야 학술 연구 조성에 의해 수행되었음.

接受日字:1996年4月19日, 수정완료일:1997年10月23日

STG(State Transition Graph) 수준과 논리 게이트 및 F/F 수준에서 이루어지고 있으며, 전이 활동을 줄이기 위하여 don't-care optimization, path balancing, factorization, precomputation logic 등의 설계 방법이 제안되었다^{[4] [5] [6] [7]}. Don't-care optimization은 회로의 설계 시 internal don't care를 사용하여 내부 노드의 신호 전이 확률(signal transition probability)을 감소시켜 소모 전력을 줄이는 설계 방법이다^{[5] [8]}. Spurious transition은 회로의 내부 노드 값이 안정되기 전에 불필요하게 여러 번 전이되는 현상으로 전체 소모 전력 중 10% - 40%를 차지한다^[9]. Path balancing은 데이터 패스 회로의 경로 지연시간을 적절히 조정함으로써 spurious transition을 감소시켜 소모 전력을 줄이는 설계 방법으로, 서로 다른 경로의 지연시간을 동일하게 맞추기 위하여 적절한 지연시간의 게이트를 사용하거나 게이트 입력에 버퍼를 추가하는 방법이 제안되었다^[6]. Factorization은 주어진 논리식의 공통 부분을 재사용함으로써 회로의 면적을 줄이는 기술 독립적(technology-independent) 최적화 방법이다^[8]. Precomputation logic 설계 방법은 입력 조건에 따라 출력에 영향을 미치지 못하는 회로의 영역을 예측하여 이 영역의 상태 전이가 일어나지 않도록 함으로써 전이 활동을 줄이는 방법으로, 주어진 회로에 ODC(Observability Don't-Care set)의 universal quantification 연산을 수행하여 설계한 precomputation 회로를 추가하는 구조들과 회로를 Shannon expansion 구조로 분할하여 입력 값에 따라 선택적으로 수행되는 구조가 제안되었다^[4].

기존의 path balancing 설계 방법은 버퍼나 게이트 capacitance 증가에 따라 전력 소모가 증가하므로 전이 활동의 감소에 의한 소모 전력의 감소 효과가 상쇄되어 효율이 좋지 않고 precomputation 설계 방법은 추가 회로에 의해 지연시간이 증가하고 적절한 ODC가 존재하지 않는 회로에 대하여 비효율적이므로 일반적인 회로에 적용하기 어려운 단점이 있다. 이에 따라 일반적인 회로에 적용이 가능하고 면적 overhead와 소모 전력을 효과적으로 줄일 수 있는 효율적인 저전력 회로 합성 알고리즘의 개발이 요구되었다.

제안된 알고리즘은 기존의 precomputation 구조의 설계 방법과 Shannon expansion 구조 설계 방법의 단점을 보완하여 면적과 전력 소모를 개선하였으며 일

반적인 회로에 적용이 가능한 장점이 있다. 회로 분할 시 ODC나 입력 변수 대신, 커널에 따라 분할된 회로를 선택적으로 수행하는 구조의 회로를 생성하며, 전체 전력 소모를 줄일 수 있는 최적의 커널을 선택하는 알고리즘을 제안하여 저전력 회로를 설계하는 논리 합성 시스템을 개발하였다. 2 장에서는 논문에서 사용되는 용어와 자료 구조를 설명하고, 3 장에서 제안된 저전력 합성 논리 알고리즘을 기술하며, 4 장에서 실험 결과를 보이고, 마지막으로 5 장에서 결론을 맺는다.

II. 논리 회로의 저전력 설계

1. 회로의 소모 전력

전력 예측 방법은 크게 주어진 입력값에 대하여 전력을 예측하는 시뮬레이션 방법과 입력값을 사용하지 않고 확률을 이용하여 전력을 예측하는 두 방법으로 나눌 수 있다. 이 중 시뮬레이션 방법은 실제 가해지는 입력 값에 대한 소모 전력을 직접 계산할 수 있지만, 회로의 평균 소모 전력 예측을 위해 N 개의 입력에 대하여 $O(2^N)$ 번의 시뮬레이션을 수행해야 하는 단점이 있다. 반면, 신호확률(signal probability)과 신호 전이 확률 등의 확률을 이용한 전력 예측 방법은 개개 입력값에 대한 전력 소모나 입력 correlation을 정확히 예측하기 어렵지만 $O(N)$ 시간 내에 수행이 가능하여 저전력 회로 합성 및 전력 예측에 자주 사용되고 있다^{[3] [4] [9]}.

신호 확률은 안정된 논리 high 값을 가지는 clock cycle의 평균 비율을 의미하고, 신호 전이 확률 $P_t(x)$ 는 회로 내부 clock cycle 중 초기값이 전이되는 clock cycle의 평균 비율을 나타낸다. 이 때 신호 확률 = on-set minterm 갯수 / 전체 minterm 갯수 수식으로 나타낼 수 있으며 BDD를 사용하여 계산이 가능하다^{[10] [11] [12]}.

회로의 평균 전력 소모량 P_{av} 는 T_c 를 clock period, C_i 를 노드 x_i 의 총 capacitance, V_{dd} 를 공급 전압, 노드 x 의 신호 전이 확률을 $P_t(x)$, n 을 회로 내의 노드 수라고 할 때 식 (2.1)과 같이 나타낼 수 있다.

$$P_{av} = \frac{1}{2T_c} V_{dd}^2 \sum_{i=1}^n C_i P_t(x_i) \quad (2.1)$$

제안된 저전력 합성 알고리즘은 기술 독립적 레벨에서 노드 전이 횟수를 줄여 전력 소모를 감소시켜 전력

소모를 줄이므로, 합성 시 구체적인 게이트의 지연 시간에 대한 정보가 없어 glitch나 hazard에 현상에 의한 spurious transition의 전력 소모를 고려하기 어렵다. Spurious transition에 의한 전력 소모는 기술 매핑(technology mapping) 과정을 거쳐 게이트의 지연 시간을 결정한 후에 고려하는 방법이 바람직하므로, 제안된 알고리즘으로 합성한 결과 회로를 spurious transition을 줄이는 path balancing 등의 최적화 과정을 통하여 추가로 전력 소모를 줄일 수 있다.

2. 기존의 Precomputation 논리에 바탕을 둔 저전력 설계 알고리즘

그림 1은 레지스터 전송 수준의 기본 회로로 함수 f 를 나타내는 회로 C_i 와 입력 래치 L_i 에 의해 분리되는 회로를 나타낸다. 그림 2는 그림 1의 회로를 pre-computation 구조로 합성한 회로 구조이며, 여기서 g_1, g_2 는 prediction 함수로 동시에 1의 값을 가지지 않으며, $g_1 = 1$ 이면 $f = 1$ 이고, $g_2 = 1$ 이면 $f = 0$ 의 조건을 만족하며 $g_1 + g_2$ 의 값이 1 일 때 블록 C_i 는 노드 전이가 발생하지 않는다^[4].

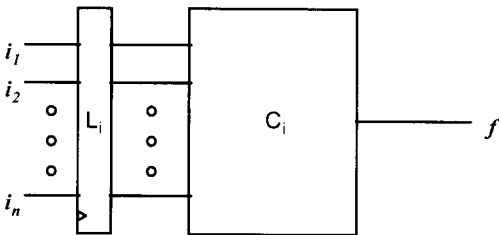


그림 1. 레지스터 전송 수준의 기본회로
Fig. 1. Register-transfer level circuit.

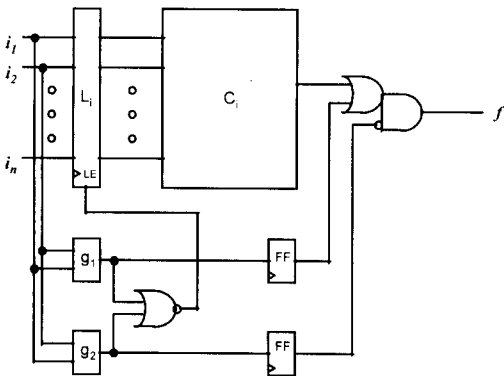


그림 2. Precomputation 논리를 이용한 회로 구조
Fig. 2. The architecture based on precomputation scheme.

그림 3은 다른 precomputation 구조를 나타낸 것

으로 클럭 주기 t 동안 g_1 또는 g_2 가 1이 되면 래치 L_i 의 load-enable 신호는 0이 되어 다음 클럭 주기 $t+1$ 동안 L_i 의 출력은 값이 바뀌지 않고, L_i 의 값에 의한 스위칭 활동만 일어나므로 블록 C_i 의 전력 소모는 감소된다. 함수 g_1 과 g_2 는 회로의 ODC와 입력 변수 집합을 적절히 선택하여 설계한다. 그림 3에서 전체 입력 집합을 X , 함수 g_1, g_2 에서 사용되는 입력의 부분 집합을 $S = \{x_1, \dots, x_m\}$ ($m < n$)라고 가정할 경우, 확률 $P(g_1 + g_2 = 1)$ 이 최대인 식 (3.2)를 만족하는 g_1, g_2 함수는 universal quantification 연산을 사용하여 결정할 수 있다. 다수의 출력이 있는 회로에서 최적의 precomputation 회로의 입력 부분 집합을 선택하는 알고리즘이 참고 문헌 [4]에서 제안되었으며, precomputation 논리에 사용되는 입력의 수는 전체 입력의 수보다 적을 수록 양호한 결과를 얻을 수 있다.

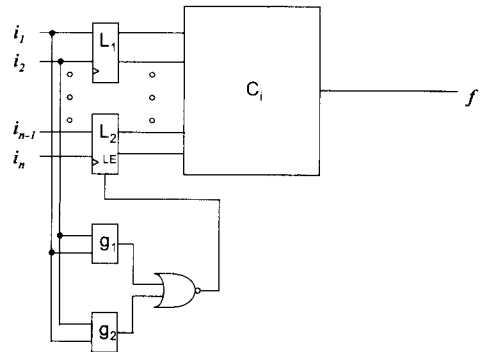


그림 3. Precomputation 논리를 이용한 다른 회로 구조
Fig. 3. The another architecture based on pre-computation scheme.

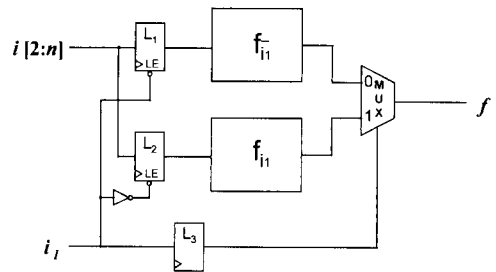


그림 4. Shannon expansion을 이용한 회로 구조
Fig. 4. The architecture based on Shannon expansion.

그림 4는 Shannon expansion을 사용한 구조의 회

로이며 i_1 은 선택된 입력 변수를 나타내며, $i_1 = 1$ 일 때 f_{i_1} 불럭을, $i_1 = 0$ 일 때 \bar{f}_{i_1} 불럭을 선택하여 수행한다^[41].

III. 제안된 저전력 회로 합성 알고리즘

1. 커널에 바탕을 둔 precomputation 회로 설계 알고리즘

기존의 precomputation 설계 방법과 Shannon expansion에 바탕을 둔 저전력 설계 방법은 내부 노드의 평균 전이 횟수를 감소시킴으로써 동적 전력 소모를 줄이는 방법이다. 이 중 precomputation 설계 방법은 입력의 신호 전이 확률을 변경하여 평균 노드 전이 횟수를 줄이며, Shannon expansion에 바탕을 둔 저전력 설계 방법은 회로를 2개로 분할하고 항상 회로의 일부만 수행하여 평균 노드 전이 횟수를 감소시킨다. 제안된 저전력 회로 합성 알고리즘은 후자의 방법을 사용하였으며, 회로의 분할 시 입력 대신 공유되는 부분 회로로 회로를 분할하여 전체 회로의 크기와 전력 소모를 효율적으로 개선하였다.

회로의 분할 시 둘 혹은 그 이상의 함수에서 공유하는 공통 부분 회로를 효율적으로 찾기 위하여 커널을 사용한다^[13]. 커널은 회로 내에서 공통되는 부분을 나타낼 수 있는 가장 기본적인 단위이므로, 회로를 나누는 효율적인 divisor가 될 수 있어 저전력 소모 합성 시 전력과 면적의 개선이 가능하다^[13]. 회로의 분할하는 공통 인수를 어떻게 선택하느냐에 따라 면적, 전력과 속도에 영향을 미치므로, 커널의 선택에 사용되는 비용 함수를 적절히 설계해야 한다. 제안된 저전력 설계 방법은 precomputation 함수 계산에 ODC가 필요하지 않으므로 합성에 적절한 ODC가 존재하지 않는 회로도 합성할 수 있는 장점이 있다.

제안된 저전력 논리 합성 시스템의 전체 구성을 그림 5에 개략적으로 나타내었다. 저전력 논리 회로 합성 과정은 크게 BDD 생성 과정, 커널 추출 및 선택 과정, 최종 회로 합성 과정으로 나눌 수 있다. 저전력 합성 과정은 먼저 게이트 레벨 회로를 입력으로 받아들여 회로로부터 BDD를 생성한 후, 이로부터 커널을 추출하고, 제안된 알고리즘을 이용하여 저전력 합성에 적합한 커널을 선택한다. 그리고 선택된 커널과 회로 BDD를 이용하여 결과 회로를 합성하여 게이트 레벨 회로를 출력한다. 제안된 합성 알고리즘은 선택된 커

널 k 를 합성한 회로, 주어진 회로(f)를 커널 k 로 나눈 회로(f_k)와 커널 \bar{k} 로 나눈 회로($f_{\bar{k}}$)를 출력한다. 커널로 나눈 f_k 와 $f_{\bar{k}}$ 회로는 커널을 합성한 회로 k 의 출력 신호 값에 따라 선택되어 수행된다.

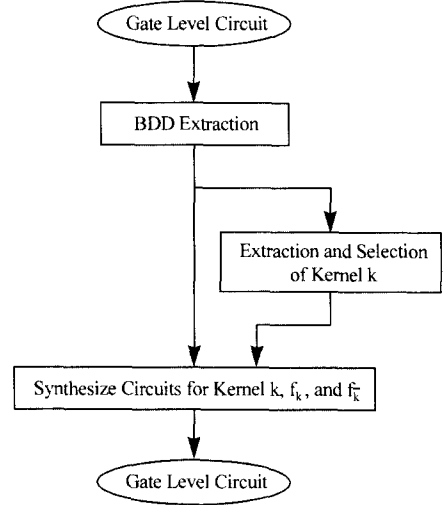


그림 5. 저전력 합성 시스템의 전체 흐름도
Fig. 5. Flow of low power synthesis system.

제안된 알고리즘에 의해 합성된 저전력 회로의 구조를 그림 6에 보였으며, I_1 과 I_2 는 각각 회로의 전체 입력 집합과 선택된 커널을 합성한 회로의 입력 집합을 나타낸다. 회로의 래치의 enable 신호는 커널을 합성한 회로(K)의 출력값에 따라 결정되며, 출력은 f_k 와 $f_{\bar{k}}$ 회로 출력단에 mux 혹은 transmission gate를 이용하여 선택된다. 출력단의 mux를 transmission gate를 사용하여 wired-OR 로직 구조로 대체하는 경우, 각 출력 당 하나의 transmission gate로 구현이 가능하여 mux를 게이트로 구현한 경우보다 전력 소모량이 줄어드는 장점이 있다.

합성된 회로에서 소모되는 전체 전력은 커널을 합성한 회로의 출력 노드가 i 이고, 이 노드의 신호 확률을 $P(i)$ 라고 할 때 식 (3.3)과 같이 나타낼 수 있다.

$$\begin{aligned}
 \text{전체 전력 소모} &= \text{커널 } k \text{를 합성한 회로의 전력 소모} + \\
 &P(i) * f_k \text{ 회로의 전력 소모} + \quad (3.3) \\
 &(1 - P(i)) * f_{\bar{k}} \text{ 회로의 전력 소모}
 \end{aligned}$$

그림 7 (a)는 benchmark 회로 중 b1 회로의 예제

를 보인 것이다. 이 회로는 sis 1.2^[14]에서 script.rugged 를 이용하여 최적화한 회로로, 입력의 신호 확률을 0.5 로 가정할 때 44.8 μ W 의 전력을 소모한다.

이 예제 회로를 입력 집합 선택 알고리즘을 사용하여 그림 2에 보인 precomputation 구조로 합성할 경우 선택한 입력 집합이 {a, b} 일 때 precomputation 함수가 $g = a b + a' b'$ 일 때 소모 전력이 34.6 μ W 인 회로를 그림 7 (b)와 같이 얻을 수 있다. 예제 회로를 그림 4의 Shannon expansion 구조로 합성한 경우 그림 7 (c)와 같이 합성된 회로를 얻을 수 있으며, 입력 a를 선택했을 때 소모 전력이 22.8 μ W 로 감소한다.

그림 7 (a)의 예제 회로에 대하여 제안된 알고리즘을 사용하여 합성할 경우 커널 (a+b)를 선택할 때 최적의 결과를 얻을 수 있으며, 이 회로를 그림 7 (d)에 나타내었다. 이 때 소모되는 전력은 식 (3.3)에 의하여 $5.9 + 0.25 * 1.24 + (1-0.25) * 12.1 = 15.3 \mu$ W 이므로 전력이 효율적으로 감소하였음을 확인할 수 있다. 면적을 나타내는 리터럴의 갯수는 그림 7 (a)의 합성 전 회로가 10 개, 그림 7 (b)의 precomputation 구조 회로가 27 개, 그림 7 (c)의 Shannon expansion 구조 회로가 18 개, 그림 7 (d)의 제안된 구조 회로가 9 개이며, 이 중 제안된 합성 알고리즘이 가장 효율적임을 알 수 있다.

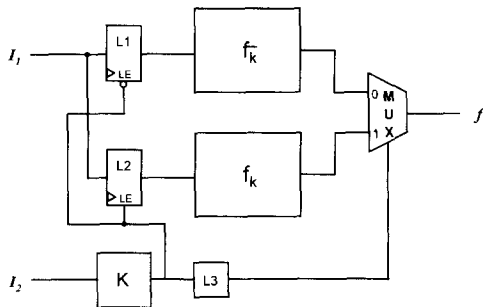


그림 6. 제안된 구조
Fig. 6. The proposed architecture.

2. 저전력 소모를 위한 커널 선택 알고리즘

저전력 논리 합성 시 전력 소모를 최대한 줄일 수 있는 커널이 선택하여 회로를 분할하여야 한다. CMOS 게이트로 구현된 VLSI 회로는 전이 활동에 의한 전력 소모가 전체 전력 소모량의 90 % 이상을 차지하므로 식 (3.3)에서 전력 소모를 회로 내부 노드

의 전이 활동으로 근사하여 사용하였다.

함수 f_k 와 f_k 의 전력 소모는 BDD를 변환한 게이트 레벨 회로에서 각 내부 노드의 신호 전이 확률을 합하여 전이 활동을 예측하였다. 제안된 시스템은 합성된 회로의 전력 소모를 반영하는 커널의 비용 함수를 식 (3.4)과 같이 사용하였다.

$$\begin{aligned} \text{커널 비용 함수} &= \text{커널 회로의 전이 활동} + \\ &P(i) * f_k \text{ 회로의 전이 활동} + \quad (3.4) \\ &(1 - P(i)) * f_k \text{ 회로의 전이 활동} \end{aligned}$$

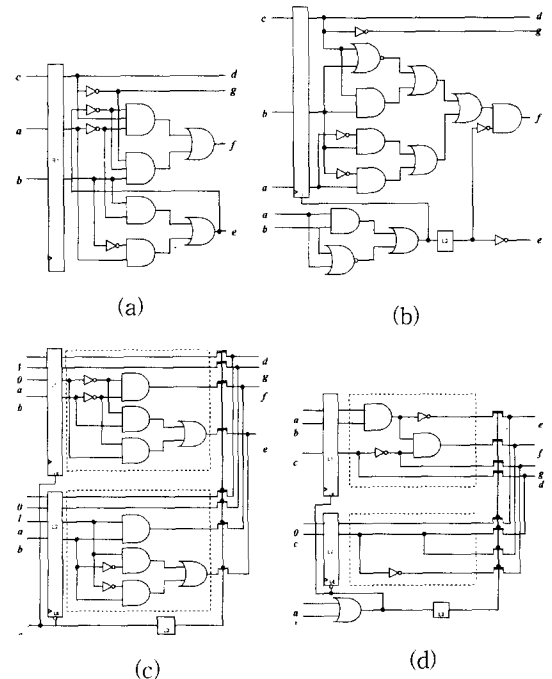


그림 7. 조합 회로의 저전력 설계 예제
(a) 조합 회로 예제 (b) Precomputation 논리를 이용한 회로 (c) Shannon expansion을 이용한 회로 (d) 제안된 구조로 변환한 회로

Fig. 7. Examples of combinational circuit design for low power.

(a) An example combinational circuit (b) Transformed circuit based on precomputation scheme (c) Circuit based on Shannon expansion (d) Transformed circuit based on the proposed scheme.

제안된 저전력 논리 회로 합성 시스템은 회로의 전력 소모를 줄일 수 있는 커널을 선택하여 회로를 합성하는 것이 바람직하며, 전체 전이 활동을 반영하는 커널 비용 함수의 값이 가장 작은 커널을 선택하여 합성을 수행하도록 설계하였다.

IV. 실험 결과

개발된 저전력 합성 알고리즘은 UNIX 환경에서 C 언어로 구현되었으며, 성능 평가를 위해 MCNC benchmark 회로에 대하여 실험을 수행하였다. 표 1 에 제안된 저전력 합성 시스템의 결과를 나타내었으며, 표에서 차례대로 원래 회로와 참고 문헌 [4] 에서 제안한 precomputation 구조의 회로와 각각 결과를 비교하였다.

표 1. 전력 소모 비교

Table 1. Comparison of power consumption.

Circuit	Original Benchmark Circuits				Precomputation Logic [4]		Proposed Algorithm		성능 비교(%)	
	#Ins	#Outs	#Lits	Power (uW)	#Lits	Power (uW)	#Lits	Power (uW)	ΔArea	ΔPower
bl	3	4	17	45	19	40	9	15	- 32.6	- 61.8
cc	21	20	60	182	123	182	72	95	- 41.5	- 47.9
cht	47	36	167	1835	168	1537	237	418	- 41.1	- 72.8
cml38	6	8	35	286	51	57	29	19	- 43.1	- 67.0
cml50	21	1	61	744	117	121	96	281	- 17.9	- 132.3
cmb	16	4	62	620	56	174	58	49	- 3.6	- 71.9
comp	32	3	185	1372	198	627	189	195	- 4.5	- 68.9
cordic	23	2	194	1049	175	299	92	123	- 47.4	- 59.0
count	35	16	143	588	165	560	217	427	- 31.5	- 23.8
cu	14	11	60	234	89	150	48	81	- 46.1	- 46.0
duke2	22	29	458	906	448	1328	544	781	- 21.4	- 41.2
majority	5	1	12	173	24	39	13	39	- 45.8	- 0.6
misex2	25	18	113	976	129	828	101	236	- 21.7	- 71.5
misex3	25	14	626	2300	628	1908	843	1750	+ 34.2	- 8.0
mux	21	1	54	715	54	557	49	202	- 71.7	- 63.7
f51m	8	8	109	491	-	-	199	426	+ 82.6	- 13.2
pcl	19	9	71	682	74	486	94	148	- 78.6	- 69.5
pcler8	27	17	95	917	98	571	126	166	+ 28.6	- 71.0
sao2	10	4	270	1191	272	422	200	348	- 26.5	- 17.5
sct	19	15	92	372	335	347	375	204	- 5.1	- 41.2
tcon	17	16	48	220	83	204	64	141	- 22.9	- 30.9
term1	34	10	625	3065	639	2133	1399	637	+118.9	- 70.1
unreg	36	16	144	1499	146	1234	272	528	+ 86.3	- 57.2
x2	10	7	47	204	70	313	44	107	- 37.1	- 65.8
z4ml	7	4	54	154	59	214	51	116	- 13.6	- 45.9
Average	20	10	178	856	178	597	217	301	- 8.8	- 43.4

표에서 #Ins/#Outs는 회로의 입출력 갯수, #Lits 은 회로의 리터럴 수를 나타낸다. 실험에서 sis 1.2 를 이용하여 5 V 전압과 20 MHz 주파수에서 zero-delay model 환경 동작 게이트를 가정하고 소모 전 력을 측정하였다. 실험 결과에서 f51m 회로는 pre-

computation 구조로 회로 설계가 불가능하여 원래 회 로에 대한 전력 감소량을 나타내었다. 회로 표 1에서 보인 바와 같이 제안된 시스템의 결과에서 원래 회로 보다 60.6 %, 참고 문헌 [4] precomputation 구 조보다 43.4 % 전력 소모가 줄어들어 성능이 향상되 었음을 알 수 있다. 제안된 시스템의 결과 회로를 sis 1.2의 mcnc 라이브러리를 이용하여 기술 매핑 후 general delay model 환경하에서 시뮬레이션 방법을 이용하여 전력 소모를 측정한 결과를 표 2에 보였 으며, 원래 회로에 비해 전력 소모가 평균 69.8 % 감소 되었다. 실험 결과로부터 제안된 시스템이 저전력 회 로를 효율적으로 합성할 수 있음을 확인할 수 있다.

표 2. 기술 매핑 후 측정된 전력 소모 비교

Table 2. Comparison of power consumption after technology mapping.

Circuit	Original Benchmark Circuits				Proposed Algorithm		성능 비교(%)	
	#Ins	#Outs	#Lits	Power (uW)	#Lits	Power (uW)	ΔArea	ΔPower
bl	3	4	17	86	9	31	- 47.1	- 64.1
cc	21	20	60	394	72	99	+ 20.0	- 75.0
cht	47	36	167	1316	237	488	+ 41.9	- 62.9
cml38	6	8	35	129	29	22	- 17.1	- 83.2
cml50	21	1	61	544	96	191	+ 57.4	- 64.8
cmb	16	4	62	329	58	25	- 6.5	- 92.4
comp	32	3	185	1468	189	442	+ 2.2	- 69.9
cordic	23	2	194	678	92	123	- 52.6	- 81.9
count	35	16	143	829	217	351	+ 51.7	- 57.6
cu	14	11	60	438	48	11	- 20.0	- 97.6
duke2	22	28	458	5181	544	509	+ 18.8	- 90.2
majority	5	1	12	54	13	5	+ 8.3	- 91.2
misex2	25	18	113	710	101	329	- 10.6	- 53.6
misex3	25	14	626	12265	843	1045	+ 34.7	- 91.5
mux	21	1	54	666	49	214	- 9.3	- 67.8
f51m	8	8	109	1049	199	385	+ 82.6	- 63.3
pcl	19	9	71	325	94	204	+ 32.4	- 37.4
pcler8	27	17	95	387	126	251	+ 32.6	- 35.3
sao2	10	4	270	1558	200	137	- 25.9	- 91.2
sct	19	15	92	795	375	54	+307.6	- 93.3
tcon	17	16	48	204	64	162	+ 33.3	- 20.7
term1	34	10	625	3310	1399	376	+123.8	- 88.6
unreg	36	16	144	710	272	662	+ 88.9	- 6.7
x2	10	7	47	408	44	40	- 6.4	- 90.3
z4ml	7	4	54	463	51	118	- 5.6	- 74.6
Average	20	10	118	733	217	251	+ 29.4	- 69.8

V. 결론

제안된 논문에서는 저전력 회로 설계를 위한 논리 합성 알고리즘을 제안하고 구현하였다. 제안된 알고리즘은 동작에 불필요한 전이를 감소하여 전력 소모를 줄일 수 있도록 한 클럭 전에 수행할 세부 회로를 결정하고, 이에 따라 다음 클럭에 두 세부 회로 중 하나가 선택되어 수행되는 구조의 회로를 합성한다. 효율적인 저전력 합성을 위하여 전력을 효과적으로 줄일 수 있는 커널을 선택하여 회로를 합성하는 알고리즘을 제안하였다.

제안된 설계 과정 중 수행될 세부 회로를 결정하는 부분은 커널을 사용하여 회로를 합성하며, 저전력 회로 합성 과정에서 전력 소모를 위해 증가된 면적과 수행 속도의 overhead를 줄이기 위해 커널의 면적과 지연시간이 되도록 작게 설계하는 것이 바람직하다. 이는 설계 작업 중 커널의 레벨을 제약 조건으로 반영하여 적절한 설계 작업을 수행할 수 있다. 합성 결과에서 생성된 회로의 출력단에 mux 회로는 wired-OR 구조의 transmission gate로 대체함으로써 추가 회로에 의한 부담을 줄일 수 있다. 제안된 시스템은 출력 함수에 많이 공유된 커널로 합성한 경우보다 양호한 결과를 얻었으며, 회로의 입력이 대칭적이거나 반복되는 구조를 가지는 회로의 경우 precomputation-based scheme과 비슷한 성능을 보인다.

실험 결과에서 제안된 알고리즘은 기존의 precomputation logic 구조보다 소모 전력이 줄어들어 보다 효율적임을 알 수 있다. 제안된 저전력 논리 합성 시스템을 사용하여 설계 작업을 수행하는 경우, 빠른 시간 내에 적절한 저전력 회로를 합성할 수 있어 저전력을 위한 설계 자동화의 효율을 높일 수 있을 것으로 기대된다.

참 고 문 헌

- [1] S. Devadas, S. Malik, "A Survey of Optimization Techniques Targeting Low Power VLSI Circuits," in Proc. 32nd DAC, pp. 242-247, June 1995.
- [2] A. Chandrakasan, "Low Power Digital CMOS Design," PhD Thesis, Univ. of California, Berkeley, UCB/ERL Memorandum no. M94/65, August 1994.
- [3] C. Tsui, M. Pedram, and A. Despain, "Exact and Approximate Methods for Switching Activity Estimation in Sequential Logic Circuits," in Proc. 31st DAC, pp. 18-23, June 1994.
- [4] M. Alidina, J. Monteiro, S. Devadas, and A. Ghosh, "Precomputation-Based Logic Optimization for Low Power," in Proc. ICCAD, pp. 74-81, Nov. 1994.
- [5] A. Shen, S. Devadas, A. Ghosh, and K. Keutzer, "On Average Power Dissipation and Random Pattern Testability of Combinational Logic Circuits," in Proc. ICCAD, pp. 402-407, Nov. 1992.
- [6] C. Lemonds, S. Shetti, "A Low Power 16 by 16 Multiplier Using Transition Reduction Circuitry," in Proc. Int'l Workshop on Low Power Design, pp. 139-142, April 1994.
- [7] K. Roy, S. Prasad, "SYCLOP: Synthesis of CMOS Logic for Low Power Applications," in Proc. ICCD, pp. 464-467, Oct. 1992.
- [8] S. Iman, M. Pedram, "Multi-Level Network Optimization for Low Power," in Proc. ICCAD, pp. 371-377, Nov. 1994.
- [9] A. Ghosh, S. Devadas, K. Keutzer, and J. White, "Estimation of Average Switching Activity in Combinational and Sequential Circuits," in Proc. 29th DAC, pp. 253-259, June 1992.
- [10] F. Najm, "Power Estimation Techniques for Integrated Circuits," in Proc. ICCAD, pp. 492-499, Nov. 1995.
- [11] R. Bahar, F. Somenzi, "Boolean Techniques for Low Power Driven Resynthesis," in Proc. ICCAD, pp. 428-432, Nov. 1995.
- [12] S. Akers, "Binary Decision Diagrams," IEEE Trans. on Computers, vol. C-27, no. 6, pp. 509-516, June 1978.
- [13] S. Iman, M. Pedram, "Logic Extraction and Factorization for Low Power," in Proc. 32nd DAC, pp. 248-253, June 1995.
- [14] A. Aziz, F. Balarin, R. Brayton, A. Sangiovanni-Vincentelli, "Sequential Sy-

thesis Using SIS," in Proc. ICCAD, pp. 612-617, Nov. 1995.

- [15] D.Lui, C. Svensson, "Power Consumption Estimation in CMOS VLSI Chips," IEEE Journal of Solid-State Circuits, vol. 29, no. 6, pp. 663-670, June 1994.
- [16] A. Bellaouar, M. Elmasry, *Low-power*

Digital VLSI Design, Kluwer Academic Publishers, 1984.

- [17] L. Lavagno, P. McGeer, A. Saldanha, A. Sangiovanni-Vincentelli, "Timed Shannon Circuits: A Power-Efficient Design Style and Synthesis Tool," in Proc. 32nd DAC, pp. 254-260, June 1995.

저 자 소 개



崔 益 成(正會員)

1992 년 2 월 서강대학교 전자공학과 졸업. 1994 년 2 월 서강대학교 전자공학과 공학 석사 취득. 1994 년 2 월 ~ 현재 동 대학원에서 박사 과정 재학 중. 주관심분야는 CAD 시스템, Synthesis for Low Power,

Computer Architecture 및 VLSI Testability 등임



柳 承 鉉(正會員)

1988 년 2 월 서강대학교 전자공학과 졸업. 1996 년 8 월 서강대학교 전자공학과 공학 석사 취득. 1988 년 1 월 ~ 현재 LG 정밀 연구소에서 근무 중. 주관심분야는 CAD 시스템, Logic Synthesis for Low Power,

차량항법 장치

黃 善 泳(正會員)

1976년 2월 서울대학교 전자공학과 졸업. 1978년 2월 한국과학원 전기 및 전자공학과 공학석사 취득. 1986년 10월 미국 Stanford 대학 공학박사 학위 취득. 1976년 ~ 1981년 삼성 반도체 주식회사 연구원. 1986년 ~ 1989년 Stanford대학 Center for Integrated Systems 연구소 책임연구원, Fairchild Semiconductor Palo Alto Research Center 기술자문. 1989년 ~ 1992년 삼성전자(주) 반도체 기술자문. 1989년 3월 ~ 현재 서강대학교 전자공학과 부교수. 주관심분야는 CAD 시스템, Computer Architecture 및 Systems Design, VLSI 설계 등임

論文97-34C-11-3

비디오 인코더용 양자화 및 역양자화기(Q_IQ Unit) 모듈의 설계

(The Design of Quantization and Inverse Quantization Unit (Q_IQ Unit) module with Video Encoder)

金銀源*, 趙源敬*

(Eun Won Kim and Won Kyung Cho)

요 약

본 논문에서는 MPEG-2 동영상 압축 시스템의 구성 요소로서 양자화 및 역양자화기를 설계하였다. 양자화 과정에 있어서, 양자화 행렬 및 스케일 계수를 SD (Signed Digit) code로 표시하고, 이를 이용하여 연산 회로를 설계하는 방법을 제안하였다. 또한, 역양자화 과정에 있어서, 양자화 스케일 코드를 제어bit(2bit)와 데이터 bit(3bit)로 분할하여 양자화 스텝 크기를 구하는 방법을 제안하였다. 설계는 VHDL로 부호화하였으며, 합성 결과로 약 6,110gates가 소요되었으며, 동작 속도로 52MHz를 얻었다.

ABSTRACT

In this paper, Quantization and Inverse Quantization unit, as a component of MPEG-2 moving picture compression system, are designed. In the processing of Quantization, this design adopted newly designed arithmetic units in which quantization matrices and scale code was expressed with SD(Signed-Digit) code. In the arithmetic unit of Inverse Quantization, quantization scale code, which has 5-bits length, is splitted into two pieces; 2-bits for control code, 3-bits for quantization data, and the method to devise quantization step size is proposed. The design was coded with VHDL and synthesis result is that it consumed about 6,110 gates, and operating speed is 52MHz.

I. 서 론

비디오 인코더에서 양자화란 DCT 처리된 결과를 정의된 값으로 나누어 보다 적은 digit로 표현함으로써 부호량을 줄이기 위한 방법이다^[1-5]. 동영상 처리 표준안에서, 양자화 및 역양자화는 양자화 행렬과 양자화 스케일에 의하여 수행되어진다. 또한, 부호화 정보량을 줄이기 위한 목적으로 인하여 인코더에 대한 규정을 두지않고, 디코더에 대한 규정만을 두고 있다. 따라서, 디코더는 역양자화에서의 포화(Saturation) 상태와 불일치(Mismatch)에 대하여 만족하여야 한다

^[1-35]. 한편, 양자화 및 역양자화기의 H/W 설계시, 양자화를 위한 제산 과정들과 양자화 스케일 값을 발생시키기 위한 과정들이 회로의 복잡성을 가중시킨다^[5-10].

본 논문은 이러한 문제점을 해결하기 위하여 양자화 과정에 있어서, 양자화 행렬 계수 및 양자화 스케일 계수를 SD(Signed Digit) code로 표시하고, 이들 SD 정보를 사용하여 양자화 연산을 수행하는 방법을 제안하였다. 이와 같은 방법은 기존의 방법에서와 같은 LUT(Look-Up Table)를 필요로 하지 않는다. 한편, 역양자화 과정에 있어서, 양자화 스케일 코드를 제어 bit(2bit)와 데이터 bit(3bit)로 분할하여 양자화 스텝 크기를 구하는 알고리즘을 제안하였다.

또한, 본 논문에서는 양자화 및 역 양자화기에 대한

* 正會員, 慶熙大學校 電子工學科

(Dept. of Electronic Eng., Kyunghee Univ)

接受日字:1997年8月29日, 수정완료일:1997年10月14日

설계를 C-언어 모델을 통하여 검증한 후, VHDL로 기술하여 설계하였다.

II. 양자화 및 역양자화

양자화(Quantization : Q) 과정은 DCT 처리된 데이터에서 실제 화질에 영향이 적은 고주파 성분(AC-level)들을 제거함으로써 보다 높은 영상 압축률을 얻기 위한 처리 과정이며, 역 양자화(Inverse Quantization) 과정은 양자화된 결과를 이전의 값들로 복원하여 IDCT Unit에 입력하기 위한 처리 과정이다. 따라서, 양자화 과정은 입력 데이터를 양자화 스텝 크기로 나누어 주고, 역 양자화 과정은 스텝 크기를 곱하게 된다. 그러므로, 양자화는 DCT 처리된 데이터를 양자화 Parameter에 따라 (1)식과 같이 양자화하고, 결과로써 발생된 데이터를 Inverse Quantizer와 Scan unit로 출력한다.

$$\text{Intra DC} = \text{dct_out_dc} // 8$$

$$\text{Intra AC} = [16 \times \text{dct_out_ac} // W_I(i,j) + \text{sign}(16 \times \text{dct_out_ac} // W_I(i,j))] // \text{qscal_cal}$$

$$\text{Non Intra} = \text{dct_out} // \text{qscal_cal} \quad (1)$$

dct_out_dc : 양자화기에 입력되는 DC계수
 dct_out_ac : 양자화기에 입력되는 AC계수
 dct_out : Non-intra인 경우의 dct계수
 Intra_dc_mult : intra의 경우에 DC값에 대한 스케일 값으로 8로 고정

$W_I(i,j)$: Intra 양자화 행렬 값
 $\text{Sign}(16 \times \text{dct_out_ac} // W_I(i,j))$:
 $16 \times \text{dct_out_ac} // W_I(i,j) > 0, 1 ; < 0, -1 ; = 0, 0 ;$
 qscal_cal : 양자화 스케일 값

또한, 역양자화는 동일한 Parameter를 가지고 양자화 이전의 데이터로 복원하며, 역 양자화를 위한 처리 과정은 (2)식과 같다.

$$\text{Intra DC} = \text{q_out} \times 8$$

$$\text{Intra AC} = (2 \times \text{q_out} + K) \times W_I(i,j) \times \text{qscal_cal}$$

$$\text{Non Intra} = (2 \times \text{q_out} + K) \times W_N(i,j) \times \text{qscal_cal} \quad (2)$$

q_out : 양자화 결과치
 K : Intra block인 경우 0,
 Non Intra block인 경우 Sign(q_out)의 값

$W_I(i,j)$: Intra 양자화 행렬 값

$W_N(i,j)$: Non Intra 양자화 행렬 값 "16"으로 균일

qscal_cal : 양자화 스케일 값

양자화 과정에서 Intra 블록인 경우, DC 계수는 고정된 상수 "8"로 나누며, AC 계수들은 양자화 스케일 코드와 양자화 스케일 타입에 의하여 결정된 양자화 스텝 크기와 양자화 행렬에 의하여 양자화된다. 양자화 행렬은 Intra블록에 대하여 표 1과 같은 행렬을 사용하며, Non Intra 블록의 행렬은 모두 "16"으로 구성된다.

III. 제안된 양자화 및 역양자화기

1. 기존 양자화 및 역양자화기 설계의 문제점

양자화 및 역양자화기 설계에 있어서, 양자화를 위한 계산 회로는 부호 및 크기에 대한 비교 회로등 부가적인 회로를 필요로 함으로 보다 복잡한 회로를 요구하게 된다. 따라서, 이러한 계산 회로 대신에 승산 회로를 사용하기 위하여 일반적으로 사전에 계산된 양자화 행렬 계수들과 스케일 계수들의 역수를 LUT (Look-Up Table)에 저장하여 사용함으로써 양자화 과정의 구현이 용이해진다는 장점이 있다.

표 1. 양자화 행렬

Table 1. Quantization matrices.

(a) Intra 블록에 대한 양자화 행렬

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

(b) Non Intra 블록에 대한 양자화 행렬

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

그러나, LUT의 사용에 따른 문제점들은 사전에 계산된 양자화 행렬 계수들 및 스케일 계수들의 역수들이 모두 부동 소숫점으로 이루어져 있음으로, 이를 부호화하여 LUT에 저장함으로 LUT의 크기가 매우 커지며, 국제 표준 기구(ISO) 규정안에 따라 일반적으로 LUT에 저장된 양자화 행렬 계수들을 Zig-Zag addressing하여 그 내용을 읽어내어 사용함으로 인하여 Zig-Zag address 변환 회로가 요구되고, 이 변환 회로로 인한 소요 시간이 증가하게 된다^{16 10)}.

표 2. 규정안의 양자화 스케일 코드와 양자화 스케일 타입
Table 2. Quantization Scale Code and Scale Types of ISO Draft.

q_scal_code	quantiser_scale [q_scal_type]	
	q_scal_type = 0	q_scal_type = 1
0	(forbidden)	
1	2	1
2	4	2
3	6	3
4	8	4
5	10	5
6	12	6
7	14	7
8	16	8
9	28	10
10	20	12
11	22	14
12	24	16
13	26	18
14	28	20
15	30	22
16	32	24
17	34	28
18	36	32
19	38	36
20	40	40
21	42	44
22	44	48
23	46	52
24	48	56
25	50	64
26	52	72
27	54	80
28	56	88
29	58	96
30	60	104
31	62	112

한편, 양자화 스케일 값 즉, 양자화 스텝 크기는 외부에서 입력되는 양자화 스케일 타입(q_scal_type)과

양자화 스케일 코드(q_scal_code)의 조합에 의하여 결정되어지며, 표 2와 같이 양자화 스케일 타입은 0과 1의 값을 갖고, 양자화 스케일 코드는 0에서 31사이의 값을 갖는다. 표 2를 고찰하면, 양자화 스케일 타입이 0일 경우, 단순히 양자화 스케일 코드값에 2를 곱한 형태이므로 1bit만 shift하여 양자화 스텝 크기를 구할 수 있다. 그러나, 양자화 스케일 타입이 1인 경우에는 양자화 스텝 크기가 비선형적으로 증가하며, 4가지 유형으로 구분된다. 따라서, 그림 1과 같이 회로를 구성되어지며, 첫번째 유형을 제외한 각 유형에 따라 지정된 값(4, 10, 17)을 감산한 후, 이 스케일 코드를 다시 1 ~ 3 bit shift시킴으로써 실제의 스텝 크기를 구하게 된다. 그러나, 이와 같은 방법으로 양자화 스텝 크기를 구할 경우, 먼저 양자화 스케일 코드가 4가지 유형중 어느 부류에 속하는지 판별하기 위한 비교 회로가 필요하며, 판별된 스케일 코드에 대하여 지정된 값을 감산함으로써 스텝 크기를 구하기 위한 감산 회로와 보정 회로 및 1 ~ 3bit shift operation이 필요하고, 부가적으로 멀티플렉서 선택 회로 및 이에 따른 부가 회로가 필요하다. 특히, 양자화 과정에 있어서 제산 과정을 수행하여야 함으로 이에 따른 회로가 복잡하게 된다.

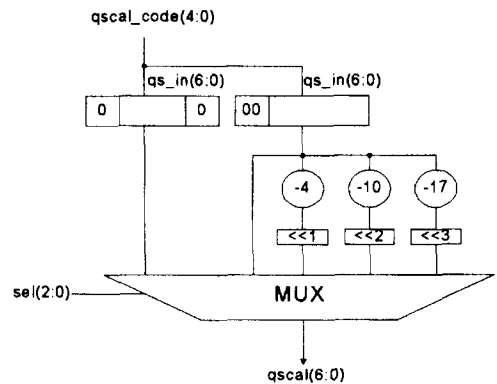


그림 1. 양자화 스케일 연산부
Fig. 1. Quantization Scale Arithmetic Unit.

2. 제한된 양자화 행렬

기존의 방식에서는 양자화 처리를 위하여 양자화 행렬 계수들의 역수를 최소 12bit이상으로 부호화하고 있으며, DCT 처리된 결과(12bit)와 승산을 위하여 12bit X 12bit의 승산기가 필요하다. 이와 같은 경우에 있어서, 승산기는 0.5 μ m의 공정하에서 약 1,250 gates가 소요되며, 동작 속도는 약 45MHz를 얻을 수

있고, 이때 LUT(Look-Up Table)은 12bit X 64의 크기를 갖는다.

그러나, 본 논문에서는 실제 사용되는 양자화 행렬의 AC 계수들이 19종류로 제한된다는 사실과 (1)식에서의 양자화 행렬 연산 부분에 대한 성질을 이용하여 양자화 행렬 계수들의 역수를 4디지트 이하의 SD(Signed Digit) code로 부호화하고, 이를 이용하여 그림 2와 같은 효과적인 양자화 행렬 연산부의 설계 방법을 제안한다.

본 논문에서 제안된 방식은 동일한 공정하(0.5 μ m tech.)에서 기존의 방식보다 적은 약 632gates 정도로 양자화 행렬에 대한 연산을 수행할 수 있으며, 4디지트 이하로 부호화된 SD code는 표 3에서의 같이 8bit의 2진수 표현보다 정확하다. 표 3에서, 양자화 과정에서 사용된 행렬 계수들은 AC 계수 16을 제외하고 19 ~ 83까지 18종류의 계수들로 구성하였으며, SD code에 의한 표현은 (1)식으로 부터 다음과 같은 (3)식과 (4)식에 의하여 구하였다.

$$Q'_{mat} = \frac{16}{Q_{mat}} \quad (3)$$

Q_{mat} ; 규정한에 대한 양자화 행렬 계수

Q'_{mat} ; SD code로 표현될 양자화 행렬 계수

$$Q'_{mat(SD)} = \sum S_k 2^{-k}, \quad S_k \in \{-1, 0, 1\} \quad (4)$$

$Q'_{mat(SD)}$; SD code로 표현된 양자화 행렬 계수

AC 계수 16은 (3)식으로 부터 "1"의 값을 갖으므로 SD code를 사용하지 않고 해당 DCT 출력을 양자화 스케일 연산부에 전송하게 된다.

제안된 양자화 행렬 연산부는 그림 2와 같이 QMG (Quantization Matrix Generator)와 BS (Barrel Shifter), 그리고 가/감산기로 구성된다. Control logic으로 구성된 QMG는 양자화 행렬 계수에 대한 SD code를 발생시킨다. 또한, 발생된 SD 정보에 대하여, digit 정보를 각각의 해당 BS에 전달하고, sign 정보는 가/감산기에 제공한다. 각각의 BS는 QMG에서 제공된 해당 digit 정보를 이용하여 DCT 출력을 지정된 bit만큼 shift operation을 수행하며, digit 정보가 "0"인 경우의 해당 BS는 "0" 출력을 수행하도록 변형된 BS를 구성하였다. 또한, 가/감산기에서는 각각의 BS에서 전달받은 출력을 QMG로 부터 제공된 sign 정보에 의하여 연산을 수행한다. 양자화 행렬 연

산부의 최종단에 있어서는 연산된 최종 결과를 12bit 로 출력하여 양자화 스케일 연산부로 전달한다.

표 3. 양자화 행렬 계수에 대한 SD code
Table 3. The Signed Digit code for Quantization matrix coefficients.

AC 계수	16/Qmat	8Bit 표현	Signed Digit 표현
19	0.84211	0.83984	$2^{-1}+2^{-2}+2^{-4}+2^{-5}=0.84375$
22	0.72727	0.72656	$2^{-1}+2^{-2}+2^{-5}+2^{-7}=0.72656$
24	0.66667	0.66406	$2^{-1}+2^{-3}+2^{-5}+2^{-7}=0.66406$
26	0.61539	0.61328	$2^{-1}-2^{-3}-2^{-7}=0.61719$
27	0.59259	0.58984	$2^{-1}+2^{-4}+2^{-5}=0.59375$
29	0.55172	0.55078	$2^{-1}+2^{-4}-2^{-7}=0.55469$
32	0.5	0.5	$2^{-1}=0.5$
34	0.47059	0.46875	$2^{-1}-2^{-5}=0.46875$
35	0.45714	0.45703	$2^{-1}-2^{-5}-2^{-7}=0.46094$
37	0.43243	0.42969	$2^{-1}-2^{-4}-2^{-8}=0.43359$
38	0.42105	0.41797	$2^{-1}-2^{-4}-2^{-6}=0.42188$
40	0.4	0.39844	$2^{-2}+2^{-3}+2^{-6}+2^{-7}=0.39844$
46	0.34783	0.34766	$2^{-2}+2^{-4}+2^{-5}+2^{-8}=0.34766$
48	0.33333	0.33203	$2^{-2}+2^{-4}+2^{-6}+2^{-8}=0.33203$
56	0.28571	0.28516	$2^{-2}+2^{-5}+2^{-8}=0.28516$
58	0.27586	0.27344	$2^{-2}+2^{-5}-2^{-8}=0.27734$
69	0.23188	0.23047	$2^{-2}-2^{-6}=0.23438$
83	0.19277	0.19141	$2^{-2}-2^{-4}+2^{-8}=0.19141$

산, 역양자화 과정에서의 양자화 행렬 계수들 역시 SD code로 부호화하여 사용하였으며, 양자화 과정의 양자화 행렬 연산부와 동일한 구조를 갖는다.

3. 제안된 양자화 스케일

본 논문에서 제안한 양자화 과정에서의 양자화 스케일 연산부는 그림 2와 유사한 구조를 갖으며, QMG 대신 QSG (Quantization Scale Generator)를 사용한다. 양자화 행렬 연산부에서의 마찬가지로 QSG는 양자화 스케일 계수들에 대한 SD code를 발생시킨다. 양자화 스케일 계수들에 대한 SD code의 예는 표 4에 나타내었으며, (5)식에 의하여 구하였다.

$$Q'_{scal} = \frac{1}{Q_{scal}}$$

$$Q'_{scal(SD)} = \sum S_k 2^{-k}, \quad S_k \in \{-1, 0, 1\} \quad (5)$$

- Q_{scal} ; 규정에 대한 양자화 스케일 계수
- Q'_{scal} ; SD code로 표현될 양자화 스케일 계수
- $Q'_{scal(SD)}$; SD code로 표현된 양자화 스케일 계수

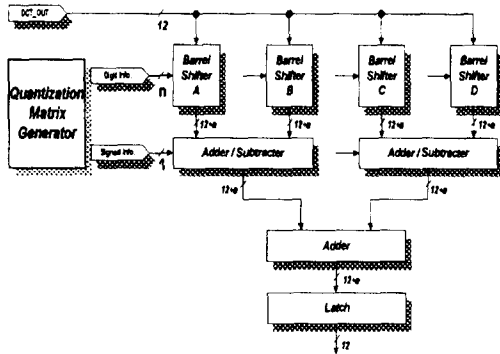


그림 2. 양자화 행렬 연산부
Fig. 2. The Arithmetic unit of Quantization matrices.

양자화 스케일 연산부에 있어서, QSG (Quantization Scale Generator)는 외부로 부터 양자화 스케일 타입과 양자화 스케일 코드를 입력받아 양자화 스케일 계수에 대한 SD code를 발생시키며, III-2절에서와 동일한 연산 과정을 수행하게 된다. 다만, 양자화 스케일 타입 1에서의 첫번째 계수 "1"에 대한 SD code는 사용하지 않으며, 양자화 과정의 최종 출력단에서 처리하도록 구성하였다.

또한, 역양자화 과정에 있어서, 본 논문에서 제안된 방식은 양자화 스케일 타입과 양자화 스케일 코드의 구조적 특성을 이용하여 양자화 스텝 크기를 구하였다. 즉, 양자화 스케일 타입이 0인 경우에는 기존의 방법대로 양자화 스케일 코드를 1bit씩 shift시킴으로써 양자화 스텝 크기를 구하였으나, 양자화 스케일 타입이 1인 경우, 제안한 방식에서는 양자화 스케일 코드 5bit 중 상위 2bit를 제어 bit로 사용하여 4가지 유형을 분류하고, 유형에 따라 하위 3bit를 데이터 bit로 이용하여 1 ~ 3bit의 bit shift operation을 수행한 후, 지정된 값을 가산함으로써 용이하게 양자화 스텝 크기를 구하며, 표 5와 같다. 그림 3은 역양자화 과정에 있어서의 제안된 양자화 스케일 발생부를 보여준다. 예로서, 양자화 스케일 타입이 1이고 양자화 스케일 코드가 20(10100)인 경우는 상위 2bit가 "10(2)"이므로 하위 3bit를 왼쪽으로 2bit shift시킨 후(10000), 지정 값 24 (11000)을 가산하면 스텝 크기 40 (101000)을

얻을 수 있다. 따라서, 양자화 스텝 크기는 7bit로 출력된다. 또한, 지정된 값에 대한 설정 역시 양자화 스케일 코드의 상위 2bit에 의하여 정하여진다. 즉, 양자화 스텝 크기에 대한 출력 7bit와 동일한 크기의 bit로 구성하되 설정 bit의 위치를 4번째 bit에 설정하여 표 6과 같이 bit 설정하면 지정되어야 할 값을 용이하게 구할 수 있다.

표 4. 양자화 스케일 타입 1에 대한 스케일 계수들의 SD code
Table 4. The SD code of scale coefficients for Quantization scale type 1.

Type 1 Scale 계수	1/Qscal	12Bit 표현	Signed Digit 표현
2	0.5	0.5	$2^{-1}=0.5$
3	0.33333	0.33325	$2^{-2}+2^{-4}+2^{-6}+2^{-8}+2^{-10}-2^{-12}=0.33325$
4	0.25	0.25	$2^{-2}=0.25$
5	0.2	0.19995	$2^{-2}-2^{-4}-2^{-6}-2^{-8}-2^{-10}-2^{-12}=0.19995$
6	0.16667	0.16650	$2^{-3}+2^{-5}+2^{-7}+2^{-9}+2^{-11}-2^{-12}=0.16675$
7	0.14286	0.14282	$2^{-3}+2^{-5}+2^{-9}+2^{-12}=0.14282$
8	0.125	0.125	$2^{-3}=0.125$
10	0.1	0.09985	$2^{-4}+2^{-5}+2^{-8}+2^{-9}+2^{-11}=0.10009$
12	0.08333	0.08325	$2^{-4}+2^{-6}-2^{-8}+2^{-10}-2^{-12}=0.08325$
14	0.07143	0.07129	$2^{-4}-2^{-7}+2^{-10}-2^{-12}=0.07153$
16	0.0625	0.0625	$2^{-4}=0.0625$
18	0.05555	0.05542	$2^{-4}-2^{-7}-2^{-10}=0.05566$
20	0.05	0.04981	$2^{-4}-2^{-6}+2^{-8}-2^{-11}-2^{-12}=0.05005$
22	0.04545	0.04541	$2^{-4}-2^{-6}-2^{-9}+2^{-11}=0.04541$
24	0.04167	0.04150	$2^{-5}+2^{-7}+2^{-9}+2^{-11}+2^{-12}=0.04175$
28	0.03571	0.03565	$2^{-5}+2^{-8}+2^{-11}=0.03565$
32	0.03125	0.03125	$2^{-5}=0.03125$
36	0.02778	0.02759	$2^{-5}-2^{-8}+2^{-11}=0.02783$
40	0.025	0.02490	$2^{-6}+2^{-7}+2^{-10}+2^{-11}=0.02490$
44	0.02273	0.02271	$2^{-6}+2^{-7}-2^{-11}-2^{-12}=0.02271$
48	0.02083	0.02075	$2^{-6}+2^{-8}-2^{-10}+2^{-12}=0.02075$
52	0.01923	0.01904	$2^{-6}+2^{-8}-2^{-12}=0.01929$
56	0.01786	0.01782	$2^{-6}+2^{-9}+2^{-12}=0.01782$
64	0.01563	0.01563	$2^{-6}=0.015625$
72	0.01389	0.01367	$2^{-6}-2^{-9}+2^{-12}=0.01392$
80	0.0125	0.01245	$2^{-6}-2^{-8}+2^{-10}-2^{-12}=0.01245$
88	0.01136	0.01123	$2^{-6}-2^{-8}+2^{-12}=0.01147$
96	0.01042	0.01025	$2^{-7}+2^{-9}+2^{-11}+2^{-12}=0.01049$
104	0.00962	0.00952	$2^{-7}+2^{-9}-2^{-11}=0.00952$
112	0.00892	0.00879	$2^{-7}+2^{-10}+2^{-12}=0.00903$

표 5. 역양자화 과정에서 제안된 제어 신호에 따른 양자화 스케일

Table 5. The Proposed Quantization scale with control signal in IQ.

qscale_code	q_scale value (type=1)	qscale_code Digit(5Bit)		q_scale value	비고
		ctrl Bit	Data Bit		
0	0	00	000	0000 000	① Data Bit<<0
1	1		001	0000 001	
2	2		010	0000 010	
3	3		011	0000 011	
4	4		100	0000 100	
5	5		101	0000 101	
6	6		110	0000 110	
7	7		111	0000 111	
8	8	01	000	00x 000 0	① Data Bit<<1 ② Data Bit+8 (1000)
9	10		001	00x 001 0	
10	12		010	00x 010 0	
11	14		011	00x 011 0	
12	16		100	00x 100 0	
13	18		101	00x 101 0	
14	20		110	00x 110 0	
15	22		111	00x 111 0	
16	24	10	000	0x 000 00	① Data Bit<<2 ② Data Bit+24 (11000)
17	28		001	0x 001 00	
18	32		010	0x 010 00	
19	36		011	0x 011 00	
20	40		100	0x 100 00	
21	44		101	0x 101 00	
22	48		110	0x 110 00	
23	52		111	0x 111 00	
24	56	11	000	x 000 000	① Data Bit<<3 ② Data Bit+56 (111000)
25	64		001	x 001 000	
26	72		010	x 010 000	
27	80		011	x 011 000	
28	88		100	x 100 000	
29	96		101	x 101 000	
30	104		110	x 110 000	
31	112		111	x 111 000	

표 5: "X"는 shift연산 후, 지정된 값과의 가산에 따른 자리 올림으로 "0" 또는 "1"을 설정함

표 6. 지정된 값과 양자화 스케일 코드 상위 2bit와의 비교

Table 6. The comparison of setting value and MSB 2bit with Quantization scale code.

qscale_code 상위 2bit	Bit 설정	10진수 표현
00 (0)	0000000	0
01 (1)	0001000	8
10 (2)	0011000	24
11 (3)	0111000	56

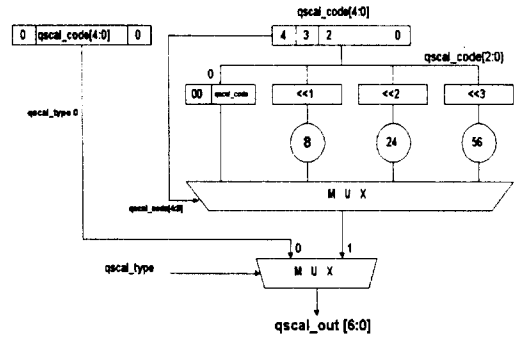


그림 3. 제안된 양자화 스케일 발생부
Fig. 3. The Proposed Quantization Scale Generator.

IV. Q_IQ Unit : 양자화 및 역 양자화기의 설계

양자화 및 역 양자화기 (Q_IQ_UNIT : Quantizer & Inverse Quantizer)는 2차원 DCT된 데이터를 양자화 처리한다. q_iq_unit의 Top-Cell은 그림 4와 같다. dct_unit으로부터 12bit 데이터(dct_out)를 입력 받아 양자화 처리한 후, 주사기로 12bit 데이터(q_out)를 출력하며, 역 양자화 출력(iq_out)은 idct_unit 과 인터페이스 된다. 제어 신호로서 dct_unit으로부터 DCT의 유효 데이터 입력을 나타내는 dct_oen과 양자화 처리된 유효 데이터의 출력을 표시하는 qout_oen을 포함하여 reset, sck2, 그리고 Macro block의 시작을 표시하는 mb_ck를 갖는다. 입출력 데이터와 제어 신호들은 시스템 clock sck2의 Falling-Edge에 따라 작동한다. 그림 5는 양자화 및 역 양자화 unit을 구성하는 각 component들의 연결을 보여준다.

양자화 과정의 parameter(intra, overf, qscal_code, qscal_type)들을 Macro Block 제어 신호에 의하여 Latch한다. 양자화 및 역 양자화 연산에 필요한 latency는 각각 4 Clock 이 소요된다.

V. 양자화 및 역 양자화기의 기능 검증

양자화 및 역 양자화기(q_iq_unit)의 설계와 기능 검증을 위하여 먼저 q_iq_unit의 기능을 C언어로 모델링하였다. 또한, VHDL을 이용하여 모듈을 설계하였고, VHDL모델과 C모델 출력을 비교하여 설계된 q_iq_unit의 기능을 검증하였다. 시험에는 2047 ~

-2048 범위내의 난수를 발생하여 테스트 벡터를 생성한 후, 이를 C모델과 VHDL모델에 적용하고, 출력 결과를 비교함으로써 설계된 양자화 및 역 양자화기의 기능을 검증하였으며, VHDL 모델 검증을 위하여 테스트 벤치를 구성하였다. 양자화 및 역 양자화기 기능 검증 구성도는 그림 6과 같다. 그림 7은 q_iq_unit의 VHDL Test-Bench인 tb_q_iq_unit의 출력 파형이며, 시험에 사용한 parameter들은 다음과 같다.

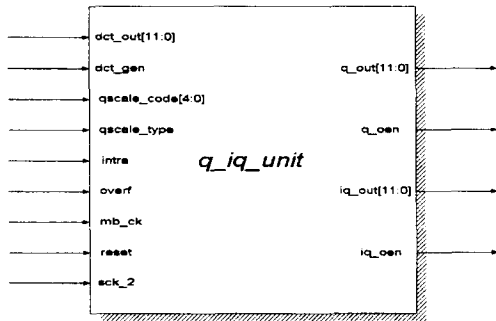


그림 4. q_iq_unit의 Top-Cell
Fig. 4. Top-Cell for q_iq_unit.

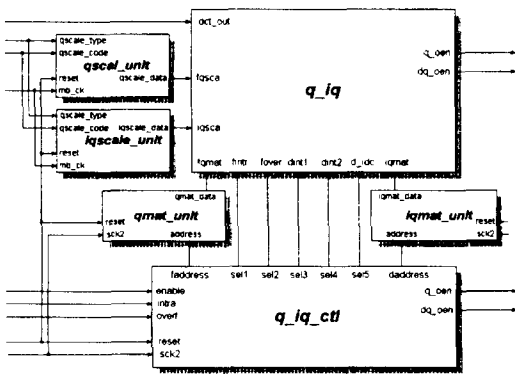


그림 5. 양자화 및 역 양자화 Unit을 구성하는 각 component들의 연결
Fig. 5. Interface of each component with construction of q_iq_unit.

난수 발생 범위 = 2047 ~ -2048
intra = 1
overf = 0
q_scal_code = 11
q_scal_type = 0

그림 7의 VHDL Test-Bench의 출력 신호 중 err_q, err_iq는 C모델의 출력과 VHDL모델의 출력을 비

교한 후, 그 차이를 나타내며, 이들 두 신호의 값이 0인 것은 C-Model과 설계된 하드웨어의 VHDL Model이 일치함을 보여준다. 그림 8은 Reset과 Macro Block(mbck) 제어 신호와 양자화 과정 입력 값의 enable이 발생된 후, 4clock후에 양자화 계산 결과가 출력되며, 8clock후에 역 양자화 결과가 출력되는 초기 Latency를 보여주는 상세 파형이다.

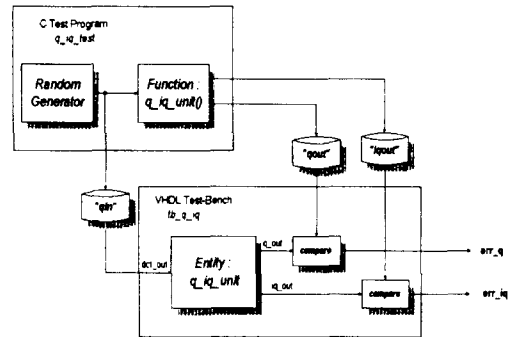


그림 6. 양자화 및 역 양자화기 기능 검증 구성도
Fig. 6. The scheme of functional verification with q_iq unit.

기능 검증된 양자화 및 역 양자화기의 VHDL 모델은 COMPASS Tool을 사용하여 합성하였으며, 사용된 Technology는 VTI 0.5μm 3중 메탈(vcmn5a3)이다. 설계된 양자화 및 역 양자화기의 VHDL 합성 결과, 약 6,110 gate가 소요되었으며, 52MHz의 동작 주파수를 얻었다.

또한, 그림 7과 8에서와 같이 설계된 양자화 및 역 양자화기는 MPEG 규정을 만족하였으며, 표 7에서 보는 바와 같이 참고문헌 [6], [7]에서 제시한 방법보다 우수함을 볼 수 있다.

표 7. 제안된 방법과 기존의 방법에 대한 비교

Table 7. The comparison of proposed and reported method.

	Tr. 수	비고
Fujiwara의 6 [6]	81K	Q, IQ, Zigzag Scan controller, Frame rate controller
Jain의 2 [7]	31K	VQ
제안된 방법	24K	Q, IQ, Controller

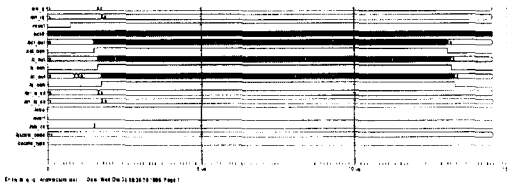


그림 7. 양자화 및 역 양자화기VHDL Test Bench 출력 파형
 Fig. 7. The output wave of VHDL test bench with q_iq unit.

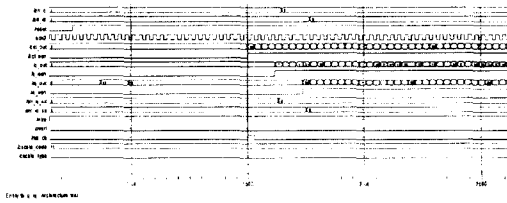


그림 8. 양자화 및 역 양자화기의 초기 Latency 상세 파형
 Fig. 8. The detailed wave of initial latency with q_iq unit.

```
#####
#   TIMING REPORT   #
#####
!top cell "Q_IQ_UNIT_P"   derating: 1.00 |
=====|
|Type  |Node(s) |Required|Initial|Current|
|-----|-----|-----|-----|-----|
|MaxDel| *->*   |10.00 ns| NA    |19.30 ns|violated +9.30
|-----|-----|-----|-----|-----|
```

그림 9. 제안된 방법에 대한 동작 주파수 결과
 Fig. 9. The result of Operating frequency with proposed method.

VI. 결론

양자화 및 역 양자화기의 H/W 설계에 있어서, 양자화를 위한 제산 과정들과 양자화 스케일 값을 발생시키기 위한 과정들이 회로의 복잡성을 가중시킨다.

본 논문에서는 양자화 과정에 있어서, 양자화 행렬 및 스케일 계수를 SD (Signed Digit) code로 표시하고, 이를 이용하여 연산 회로를 설계하는 방법을 제안하였다. 이와 같은 방법을 사용하여 본 논문에서는 기존의 방법에서의 LUT를 계수 발생기(QMG, QSG)로 대체하였으며, 또한 계수들은 4디지트 이하의 SD code로 표시될 수 있는 것을 이용하여 양자화 행렬

및 스케일 연산부의 승산기를 4개의 변형된 BS (Barrel Shifter)와 3개의 가/감산기로 대체함으로써 보다 고속의 연산 회로를 적은 gate로 실현하였다. 한편, 역양자화 과정에 있어서, 양자화 스케일은 양자화 스케일 타입과 양자화 스케일 코드의 구조적인 특성을 이용하는 알고리즘을 제안하였다. 즉, 양자화 스케일 코드의 상위 2bit를 제어 bit로 사용하여 하위 3bit에 대한 1 ~ 3bit shift operation을 수행한 후, 지정된 값만을 가산함으로써 용이하게 양자화 스텝 크기를 구할 수 있도록 하였다. 또한, 지정된 값의 설정 역시 양자화 스케일 코드의 상위 2bit에 의하여 설정되도록 하였다.

또한, 본 논문에서는 양자화 및 역 양자화기에 대한 설계를 C 언어 모델을 통하여 기능 검증을 수행한 후, C 언어 모델과 동일하게 VHDL 모델로 기술하여 설계하였다. 기능 검증된 VHDL 모델은 COMPASS Tool을 사용하여 합성하였으며, 사용된 Technology는 VTI CMOS 0.5 μ m 3중 메탈(vcmn5a3)이다. 합성한 결과, 약 6,110gate가 소요되었고, 약 52MHz의 동작 주파수를 얻었으며, MPEG 규정을 만족하였다.

참고 문헌

- [1] ISO/IEC JTC1/SC29/WG11, MPEG Committee Draft CD11172, 1991.
- [2] ISO/IEC JTC1/SC29/WG11, MPEG Committee Draft CD13818-2, 1995.
- [3] CCITT Recommendation H.261, 1990.
- [4] K.R.Rao, P.Yip, Discrete Cosine Transform, Academic Press, 1990.
- [5] V.Bhaskaran, K.Konstantinides, Image and Video Compression Standards Algorithms and Architectures, Kluwer Academic Publishers, 1995.
- [6] Hiroshi Fujiwara 외 6, "An All-ASIC Implementation of a Low Bit-Rate Video Codec", IEEE Trans. Circuits Syst. For Video Tech., vol. 2, no. 2, pp. 123-134, June, 1992.
- [7] Rajeev Jain 외 2, "An Integrated Circuit Design for Pruned Tree Search Vector Quantization Encoding with an Off-Chip Controller", IEEE Trans. Circuits Syst.

- For Video Tech., vol. 2, no. 2, pp. 147-158, June, 1992.
- [8] J.H. Kim 외 3, "A Pipelined Architecture for VLSI Implementation of Quantizer / Inverse Quantizer Processor for Video CODEC", ICSPAT'94, pp. 1737-1742, Oct. 18-21, 1994.
- [9] Peter A. Ruetz 외 4, "A High-Performance Full-Motion Video Compression Chip Set", IEEE Trans. Circuits Syst. For Video Tech., vol. 2, no. 2, pp. 111-122, June, 1992.
- [10] CCITT Video Compression Databook, LSI Logic, Sep., 1991.

 저 자 소 개



金 銀 源(正會員)

1962년 1월 20일생. 1985년 2월 경희대학교 전자공학과 졸업(공학사). 1990년 2월 경희대학교 대학원 전자공학과 졸업(공학석사). 1993년 8월 경희대학교 대학원 전자공학과 박사과정 수료. 주요관심분야는 동영상

신호 처리 시스템 및 VLSI 설계, 멀티미디어 및 컴퓨터 통신망, 고속 연산 처리 시스템 및 회로 설계, 문자 인식 시스템 및 패턴 인식

趙 源 敬(正會員) 第 33卷 B編 第 11號 參照