

멀티 레벨 셀 낸드 플래시 메모리용 적응적 양자화기 설계

이 동 환*, 성 원 용°

Adaptive Quantization Scheme for Multi-Level Cell NAND Flash Memory

Dong-hwan Lee*, Wonyong Sung°

요 약

본 논문에서는 멀티 레벨 셀 낸드 플래시 메모리에서 연판정 에러 정정을 위한 적응적 비균일 양자화기를 제안한다. 기존의 최대 상호 정보(maximum mutual information) 양자화기는 최적의 연판정 에러 정정 성능을 보이지만, 소모적인 탐색(exhaustive search)으로 인하여 많은 계산량을 요구한다. 본 논문에서 제안된 양자화기는 최대 여섯 개의 파라미터로 표현되는 간단한 구조를 갖고 있어 연산량이 적다. 또한 제안된 양자화기는 쓰기 심볼과 읽기 심볼 사이의 상호 정보를 최대화하는 방향으로 파라미터 값의 최적화시키므로, 최대 상호 정보 양자화기에 근접하는 우수한 연판정 에러 정정 성능을 보인다.

Key Words : NAND flash memory, Quantizer, Maximum mutual information (MMI), Error correcting code (ECC)

ABSTRACT

An adaptive non-uniform quantization scheme is proposed for soft-decision error correction in NAND flash memory. Even though the conventional maximizing mutual information (MMI) quantizer shows the optimal post-FEC (forward error correction) bit error rate (BER) performance, this quantization scheme demands heavy computational overheads due to the exhaustive search to find the optimal parameter values. The proposed quantization scheme has a simple structure that is constructed by only six parameters, and the optimal values of them are found by maximizing the mutual information between the input and the output symbols. It is demonstrated that the proposed quantization scheme improves the BER performance of soft-decision decoding with only small computational overheads.

I. 서 론

낸드 플래시 메모리는 빠른 읽기/쓰기 동작과 낮은 소비 전력, 충격에 강한 특성 등의 장점으로 인하여 휴대폰, USB 메모리, 디지털 카메라, SSD

(solid state drive) 등의 다양한 전자 장치에서 광범위하게 사용되고 있다. 낸드 플래시 메모리의 대용량화는 한 셀에 여러 비트를 저장하는 멀티 레벨 셀(multi-level cell) 방식과 공정 기술의 미세화를 통하여 꾸준히 이루어져왔고, 최근에는 20 nm 대

※ 이 논문은 2013년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2012R1A2A2A06047297).

♦ 주저자 : 서울대학교 전기 정보 공학부 멀티미디어 시스템 연구실, ldh@dsp.snu.ac.kr, 학생회원

° 교신저자 : 서울대학교 전기 정보 공학부 멀티미디어 시스템 연구실, wysung@snu.ac.kr, 정회원

논문번호 : KICS2013-04-157, 접수일자 : 2013년 4월 1일, 최종논문접수일자 : 2013년 6월 5일

의 공정을 이용해 제작된 낸드 플래시 메모리가 이미 양산되고 있다. 그러나 메모리의 집적도가 높아질수록, 셀의 문턱 전압(threshold voltage) 신호는 각종 잡음 및 인접 셀의 간섭(cell-to-cell interference)에 취약해지고, 그로인해 발생하는 비트 에러의 수도 증가하고 있다. 이에 따라 비트 에러를 줄이고 낸드 플래시 메모리의 신뢰도를 향상시키기 위하여 BCH(Bose-Chaudhuri-Hocquenghem) 부호 또는 RS(Reed-Solomon) 부호 기반의 경판정 에러 정정(hard-decision error correction)이 현재까지 널리 이용되고 있다^{11,12}. 그러나 20 nm 이하 공정의 낸드 플래시 메모리에서는 비트 에러의 수가 경판정 에러 정정 방법만으로는 감당할 수 없는 정도로 급격하게 증가한다. 따라서 최근에는 연판정 에러 정정(soft-decision error correction)을 낸드 플래시 메모리에 적용하려는 연구가 활발하게 이루어지고 있다^{3,4}.

연판정 에러 정정을 수행하기 위해서는 복호화기(decoder)의 입력으로 0과 1로 명확히 구분되는 경판정 신호가 아닌 연판정 신뢰도 정보(soft reliability information)가 필요하다. 낸드 플래시 메모리에서 신뢰도 정보는 셀의 컨트롤 게이트(control gate)에 거는 기준 전압(reference voltage)을 바꾸어 가며 메모리를 여러 번 읽고, 이렇게 얻은 멀티 비트(multi-bit)의 신호를 이용하여 계산할 수 있다. 이때, 에러 정정 성능은 메모리 센싱의 횟수와 기준 전압의 위치에 따라 크게 좌우되기 때문에, 최적의 양자화기(quantizer)를 찾는 것이 매우 중요하다. 최근 발표된 연구 결과에서는 낸드 플래시 메모리의 쓰기 심볼과 읽기 심볼 사이의 상호 정보(mutual information)를 최대로 하면 최적의 에러 정정 성능을 얻을 수 있다는 사실이 발표된 바 있다¹⁵. 그러나 이와 같은 양자화기를 소모적인 탐색(exhaustive search) 방식으로 찾는 것은 상당한 계산량을 필요로 한다. 특히, 낸드 플래시 메모리에서 신호 대 잡음비(signal-to-noise ratio, SNR)는 리텐션(retention) 시간과 PE(program-erase) 사이클이 바뀔 때 따라 계속적으로 변화하기 때문에, 신호 대 잡음비에 적응적으로 변화하는 최대 상호 정보(maximizing mutual information) 양자화기를 찾는 것은 매우 큰 계산량을 요구한다.

본 연구에서는 비교적 적은 연산만으로도 구현이 가능한 낸드 플래시 메모리용 적응적 비균일 양자화기의 설계 방법을 제안한다. 본 연구에서 제안된 양자화기는 [3]에서 제시된 양자화기의 구조를 변형

하여, 최대 여섯 개의 파라미터 값으로 결정되는 비교적 간단한 구조를 가지고 있다. 따라서 최대 상호 정보 양자화기보다 훨씬 더 적은 연산량을 요구한다. 또한 제안된 양자화기는 파라미터 값을 [5]에서와 같이 입력 심볼과 출력 심볼간의 상호 정보를 최대화하는 방향으로 최적화하기 때문에, 최대 상호 정보 양자화기에 근접한 수준의 에러 정정 능력을 보인다. 아울러, 본 논문에서는 제안된 양자화기를 적응적으로 이용하여 연판정 에러 정정에 사용하는 방법을 보인다. 제안된 양자화기는 LDPC(low density parity check) 부호 및 TPC(turbo product code) 등의 다양한 부호의 연판정 에러 정정에 사용될 수 있다. 본 논문에서는 EG(Euclidean geometry) LDPC 부호를 이용하여 연판정 에러 정정 모의실험을 반복적으로 수행함으로써 제안된 양자화기의 에러 정정 능력을 측정하였다.

본 논문의 구성은 다음과 같다. 2장에서는 낸드 플래시 메모리를 위한 최대 상호 정보 양자화기와 [3]에서 제시된 단일 파라미터 양자화기에 대하여 설명한다. 3장에서는 기존의 양자화기들의 장점을 결합시킨 적응적 비균일 양자화기를 제안한다. 4장에서는 제안된 양자화기에 대한 모의실험 결과를 제시한 후 5장에서 결론을 맺는다.

II. 기존의 낸드 플래시 메모리용 양자화기

본 장에서는 낸드 플래시 메모리에서 연판정 에러 정정을 위해 제안된 기존의 양자화기에 대해서 설명한다. 우선, [3]에서는 단일 파라미터에 의해 구조가 결정되는 비균일 양자화기가 제안되었다. 이 양자화기는 그림 1에서와 같이 두 심볼의 문턱 전압 분포가 겹치는 구간을 (비트 에러의 발생 확률이 높은 구간) 소거(erasure) 구간으로 정의하고, 이를 그렇지 않은 구간에 비하여 더 조밀하게 나눈다. 즉, 본 양자화기에서는 소거 구간을 $(q_{lk}, q_{rk}]$ 로 ($k = 0, 1, 2$) 정의하고, q_{lk} 과 q_{rk} 사이를 균등하게 구분하여 메모리 센싱의 기준 전압(memory sensing reference voltage)으로 사용한다. 위의 양자화기에서 q_{lk} 와 q_{rk} 는 다음의 식에 의하여 결정된다.

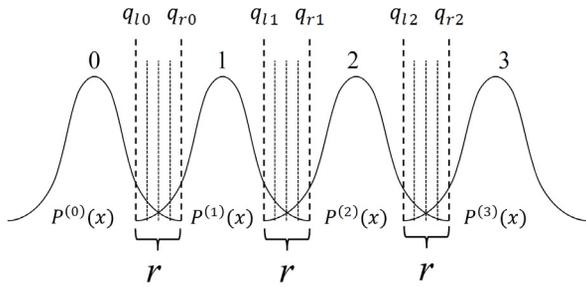


그림 1. [3]에서 제안된 단일 파라미터 비균일 양자화기.
Fig. 1. Uni-parameter non-uniform quantizer proposed in [3].

$$\frac{P^{(k)}(q_{lk})}{P^{(k+1)}(q_{lk})} = \frac{P^{(k+1)}(q_{rk})}{P^{(k)}(q_{rk})} = r \quad (1)$$

여기서 $P^{(k)}(x)$ 는 k 번째 심볼의 확률 분포를 나타내는데 이는 정규분포로 근사된다. 참고로, [6]에서 제시된 방법을 이용하면, $P^{(k)}(x)$ 의 평균과 분산을 비교적 적은 횟수의 메모리 센싱을 통하여 구하는 것이 가능하다. 식 (1)에서 r 은 두 심볼의 확률비를 나타내는데, r 값이 클수록 소거 구간이 넓어진다. 위의 양자화기는 하나의 파라미터 값에 의하여 그 구조가 결정되기 때문에, 단일 파라미터(uni-parameter) 양자화로 볼 수 있고 적은 연산량으로 이를 구하는 것이 가능하다. 이 양자화기를 이용하여 연관성 에러 정정을 수행할 때, 에러 정정 능력은 r 값에 의하여 크게 좌우된다. 따라서 최적의 r 값을 찾는 것이 중요하다. [3]에서는 연관성 에러 정정의 성능 곡선을 구하는 모의실험을 여러 번 수행하여, 평균적으로 좋은 성능을 보이는 값으로 r 을 정하였다. 참고로, 낸드 플래시 메모리에서는 에러 정정 코드의 길이가 매우 길고 (8K에서 64K 비트) 낮은 비트 에러율(bit error rate, BER)을 요구하기 때문에, 에러 정정 성능 곡선을 구하는 모의실험의 수행에 상당히 긴 시간이 소요된다. 한편, 낸드 플래시 메모리에서는 신호 대 잡음비(signal-to-noise ratio)가 데이터 리텐션(data retention)과 PE 사이클(program-erase cycle)에 따라 계속적으로 변화하기 때문에, 적응적(adaptive) 양자화기를 사용하여야 최적의 에러 정정 성능을 얻을 수 있다. 따라서 r 값을 고정하여 사용하는 [3]의 양자화기는 최적화된 에러 정정 성능을 보이기 어렵다.

최근에는 BI-AWGN(binary input additive white Gaussian noise) 채널에서 균일 LDPC(regular low

density parity check) 부호의 다양한 앙상블(ensemble)에 대하여 최적화된 양자화기를 찾는 방법이 제안된 바 있다^[7]. 이 방법에서는 양자화기를 하나 혹은 여러 개의 파라미터를 이용하여 대칭적으로 구성하고, 각 파라미터의 값을 조금씩 바꾸어 가며 에러 정정 성능을 측정한다. 이 때, 성능 측정을 위한 척도로 비트 에러 없이(error free) 복호가 가능한 신호 대 잡음비의 임계값을 (혹은 잡음 신호의 분산의 최댓값) 사용하고, 이를 양자화된 밀도 진화 방법(discretized density evolution)을 이용하여 측정한다. 즉, 이 방법에서는 신호 대 잡음비가 낮은 상황에서도 에러 없이 복호하는 것을 가능하게 하는 양자화기를 찾는 것을 목표로 한다. 참고로, 여기서는 부호의 길이가 무한하고 복호 과정의 횟수의 제한이 없는 이상적인 경우만을 고려한다.

이 방법은 신호 대 잡음비가 매우 좋은 경우 (에러 없이 복호가 가능할 정도로), 최적화된 양자화기를 효율적으로 찾을 수 있다. 그러나 이 방법을 낸드 플래시 메모리에 직접 적용하는 데에는 몇 가지 문제점이 있다. 우선, 낸드 플래시 메모리의 문턱 전압 분포는 비대칭적이기 때문에 위 방법을 통해 얻어진 대칭 구조의 양자화기는 최적의 성능을 내기 어렵다. 둘째, 이 방식에서는 특정 신호 대 잡음비에 (에러 없이 전송이 가능한) 대하여 설계한 양자화기를 고정적으로 사용하기 때문에 신호 대 잡음비가 이보다 더 낮은 경우 에러 정정 능력이 열화될 수 있다.

최대 상호 정보(maximum mutual information, MMI)^[5] 양자화기는 에러 정정 부호의 특성에 대해서는 고려하지 않고, 신호 대 잡음비만을 고려하여 최적화된 양자화기를 찾는다. 이 양자화기는 입력 심볼과 (그림 1에서 X) 출력 심볼 (그림 1에서 Y) 사이의 상호 정보를 최대화 하는 양자화기로서, 낸드 플래시 메모리와 같이 신호의 해상도가 낮은 상황에서 최적의 에러 정정 성능을 보인다^[5]. 이 양자화기 역시 에러의 발생빈도가 높은 심볼의 경계면을 높은 정밀도로 읽고, 나머지 구간을 낮은 정밀도로 읽는다. 에러 정정 성능 곡선을 구하는 모의실험을 통하여 평균적으로 좋은 성능을 내는 양자화기를 구한 [3]과는 달리, 이 방식에서는 상호 정보의 크기를 통해 양자화기의 성능을 예측한다. 즉, 상호 정보의 크기가 크면 주어진 양자화기는 우수한 에러 정정 성능을 보일 것이라고 가정한다. 따라서 이 방법에서는 최적화된 양자화기를 찾기 위하여, 경계(boundary)를 조금씩 바꾸어가며 최대 상호 정

보를 내는 조합을 찾는다.

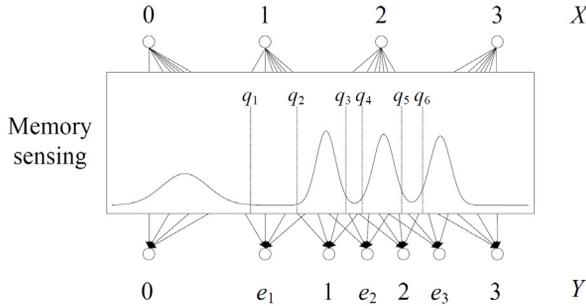


그림 2. 7-level 최대 상호 정보 양자화기.
Fig. 2. 7-level maximizing mutual information quantizer.

그림 2에서는 일곱 개의 전압 구간(7-level)을 가지는 최대 상호 정보 양자화기를 나타낸다. 참고로, 7-level 양자화기에서는 $q_1 \sim q_6$ 위치에 여섯 번의 메모리 센싱이 요구된다. 이 때, X 와 Y 간의 상호 정보는 다음과 같이 구해진다.

$$I(X; Y) = H(Y) - H(Y|X) \quad (2)$$

$$= H(P_0, P_1, \dots, P_6)$$

$$- \frac{1}{4} \sum_{i=0}^3 H(P_{i,0}, P_{i,1}, \dots, P_{i,6})$$

여기서 $H(X)$ 는 확률 변수 X 의 엔트로피(entropy)를 의미한다. $P_{i,j}$ 는 입력 심볼 i 가 심볼 j 로 출력될 확률을 의미하고, P_j 는 $\sum_{i=0}^3 P_{i,j}$ 를 나타낸다. 이 때, $P_{i,j}$ 는 다음과 같이 계산된다.

$$P_{i,j} = Q\left(\frac{q_j - m_i}{\sigma_i}\right) - Q\left(\frac{q_{j+1} - m_i}{\sigma_i}\right) \quad (3)$$

식 (2)에서는 $P_{i,j}$ 를 구하기 위한 Q-함수와 엔트로피의 계산을 위한 로그 함수를 필요로 하기 때문에, 단일 파라미터 양자화기에 비하여 계산이 복잡하다. 전압 구간의 수를 N_L 이라고 가정하고, 식 (2)를 계산하기 위해서 요구되는 연산량을 표 1에 나타내었다.

표 1. 식 (2)를 계산하기 위한 연산량
Table 1. The number of arithmetic operations for Eq. (2)

	$P_{i,j}, P_j$	$I(X; Y)$	Total
Add.	$11N_L$	$5N_L - 1$	$16N_L - 1$
Div.	$4N_L$	$5N_L + 1$	$9N_L + 1$
Q-func.	$4N_L$	0	$4N_L$
Log	0	$5N_L$	$5N_L$

이에 더하여, 상호 정보를 최대화 가지는 양자화기를 얻기 위해서는 $q_1 \sim q_6$ 의 값을 하나씩 바꾸어 가며 상호 정보를 계산하고 그 중 최댓값을 내는 조합을 찾는 과정이 필요한데, 이 과정에서 상당히 많은 계산량이 요구된다. 예를 들어, 각 파라미터를 10번씩만 바꾸어 가며 7-level의 최대 상호 정보 양자화기를 찾는다면, 식 (2)를 10^6 번 계산하는 것이 필요하다. 만일 16-level 최대 상호 정보 양자화기를 위와 같은 과정으로 찾는다면 10^{15} 번의 탐색이 요구된다. 앞서 설명한 대로, 낸드 플래시 메모리에서는 최적의 에러 정정 성능을 내는 양자화기가 신호 대 잡음비에 따라 서로 다르기 때문에, 최소한 수십 개의 양자화기를 찾는 것이 필요하다. 따라서 최대 상호 정보 양자화기를 실제 낸드 플래시 메모리에 적응적으로 적용하는 것은 어렵다.

III. 적응적 비균일 양자화기의 설계

본 장에서는 앞서 설명한 두 양자화기의 특징을 이용하여, 연산량이 적은 적응적 비균일 양자화기의 설계 방식을 제안한다. 본 연구에서 제안하는 양자화기는 단일 파라미터 양자화기를 확장하여, 여러 개의 파라미터로 표현되는 그림 3과 4의 구조를 가진다.

먼저, 그림 3에서는 7-level 양자화기의 구조를 나타낸다. 이 양자화기에서 r_0, r_1, r_2 의 세 파라미터는 다음의 식을 통하여 소거 구간의 경계를 (q_{lk} 및 q_{rk}) 구하는데 이용된다.

$$\frac{P^{(k)}(q_{lk})}{P^{(k+1)}(q_{lk})} = \frac{P^{(k+1)}(q_{rk})}{P^{(k)}(q_{rk})} = r_k \quad (4)$$

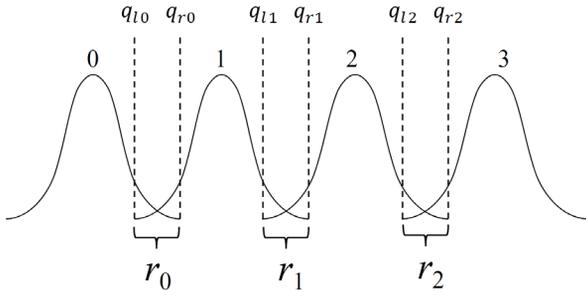


그림 3. 제안된 7-level 비균일 양자화기.
Fig. 3. Proposed 7-level non-uniform quantizer.

앞서 설명한 바와 같이 $P^{(k)}(x)$ 는 정규분포로 근사될 수 있다. 식 (4)에 정규분포 함수를 대입하고 정리하면 다음의 방정식을 얻을 수 있다.

$$-\frac{(q_{l0} - m_0)^2}{\sigma_0^2} + \frac{(q_{l0} - m_1)^2}{\sigma_1^2} = 2\ln\left(\frac{\sigma_0}{\sigma_1} r_0\right) \quad (5)$$

$$-\frac{(q_{r0} - m_1)^2}{\sigma_1^2} + \frac{(q_{r0} - m_0)^2}{\sigma_0^2} = 2\ln\left(\frac{\sigma_1}{\sigma_0} r_0\right) \quad (6)$$

식 (5)와 (6)은 각각 q_{l0} 와 q_{r0} 에 대한 이차 방정식으로, 동일한 방법을 이용하면 q_{l1} , q_{r1} , q_{l2} , q_{r2} 에 관한 방정식을 얻을 수 있다. 식 (5)와 (6)을 계산하기 위해서는 덧셈, 곱셈, 나눗셈, 제곱근 및 로그 연산을 각각 7회, 20회, 2회, 2회, 2회 수행하여야 한다. 10-level 양자화기에서는 각 심볼의 경계 구간에 두 개의 소거 구간이 존재하고, 이를 위해 세 번의 메모리 센싱이 요구된다. 따라서 식 (5)와 (6)을 통하여 결정된 q_{lk} 와 q_{rk} 의 중간 지점을 기준 전압으로 추가함으로써 10-level 양자화기를 구할 수 있다.

13-level 및 16-level 양자화기에서는 각 심볼의 경계 부분에 소거 구간이 각각 3개, 4개씩 존재하고, 이에 따라 각각 4개, 5개의 기준 전압을 정해주어야 한다. 이들 양자화기는 그림 4에서와 같이 여섯 개의 파라미터를 이용하여 구할 수 있고, 기준 전압은 식 (4)을 이용하여 계산된다. 참고로, 낸드 플래시 메모리에서는 16-level 이상의 양자화기를 사용하더라도 여러 정정 능력이 크게 향상되지 않기 때문에^[8], 본 연구에서는 16-level 양자화기까지를 고려하였다. 표 2에서는 제안된 양자화기를 한 파라미터 당 N_s 번의 탐색을 통하여 구할 때의 연산량을 나타낸다.

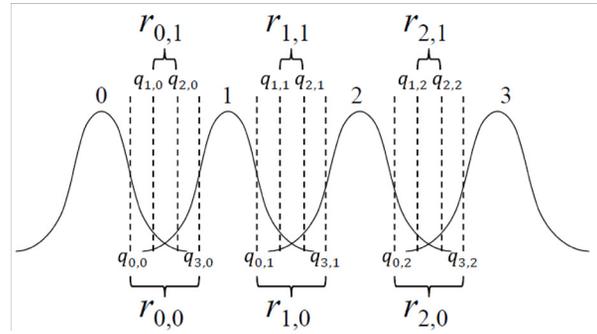


그림 4. 제안된 13-level 비균일 양자화기.
Fig. 4. Proposed 13-level non-uniform quantizer.

표 2. 제안된 양자화기의 연산량
Table 2. The number of arithmetic operations for the proposed quantizers

	7- and 10-level	13- and 16-level
Add.	$7N_s^3$	$7N_s^6$
Mul.	$20N_s^3$	$20N_s^6$
Div.	$2N_s^3$	$2N_s^6$
Log	$2N_s^3$	$2N_s^6$
Sqrt.	$2N_s^3$	$2N_s^6$

단일 파라미터 양자화기의 경우와 마찬가지로, 제안된 양자화기에서는 파라미터의 값에 따라 여러 정정 성능이 크게 좌우된다. 본 연구에서는 최적의 여러 정정 성능을 내는 파라미터 값을 결정하기 위한 척도로 입력 심볼과 출력 심볼간의 상호 정보를 이용한다. 즉, 최대 상호 정보를 내는 파라미터의 조합을 찾고, 이를 이용하여 식 (4)에 따라 양자화기를 생성한다.

최대 상호 정보를 내는 파라미터 조합을 찾기 위한 가장 간단한 방법으로 소모적인 탐색을 생각할 수 있다. 이 방식에서는 각 파라미터를 조금씩 바꾸어 가며 식 (4)를 이용하여 다중 파라미터 양자화기를 구한다. 이렇게 얻어진 양자화기에서의 상호 정보의 양은 식 (2)와 (3)을 이용하여 구한다. 제안된 양자화기는 이러한 계산을 반복적으로 수행하여, 여러 조합 중에서 가장 큰 상호 정보를 내는 파라미터 조합을 찾는다. 이 방식에서는 파라미터 수가 최대 여섯 개로 제한되기 때문에 소모적인 탐색으로도 최적의 파라미터 조합을 찾는 것이 가능하다.

표 3. 최대 상호 정보 양자화기와 제안된 양자화기의 연산량 비교 (N_L : 전압 레벨 수, N_s : 한 파라미터 당 탐색 횟수)
 Table 3. The number of arithmetic operations for the MMI and the proposed quantizers (N_L : number of voltage levels, N_s : number of iterations per parameter)

	MMI quantizer	proposed quantizer (7- and 10-level)	proposed quantizer (13- and 16-level)
Add.	$(16N_L - 1)N_s^{N_L - 1}$	$(16N_L + 6)N_s^3$	$(16N_L + 6)N_s^6$
Mul.	0	$20N_s^3$	$20N_s^6$
Div.	$(9N_L + 1)N_s^{N_L - 1}$	$(9N_L + 3)N_s^3$	$(9N_L + 3)N_s^6$
Q-func.	$4N_L N_s^{N_L - 1}$	$4N_L N_s^3$	$4N_L N_s^6$
Log	$5N_L N_s^{N_L - 1}$	$(5N_L + 2)N_s^3$	$(5N_L + 2)N_s^6$
Sqrt.	0	$2N_s^3$	$2N_s^6$

표 3은 최대 상호 정보 양자화기와 제안된 양자화기의 연산량을 나타낸다. 최대 상호 정보 양자화기의 경우 파라미터의 수가 증가함에 따라 연산량이 지수적으로 증가한다. 이에 반하여, 제안된 양자화기에서는 최대 N_s^6 에 비례하는 연산량을 보인다. 예를 들어, 7-level 양자화기에서 한 파라미터 당 10 번의 탐색을 수행하는 경우를 살펴보자. 모든 연산에 동일한 수행시간이 걸린다고 가정하면, 최대 상호 정보 양자화기는 제안된 양자화기에 비하여 약 600배 많은 연산을 필요로 한다. 16-level 에서 최대 상호 정보 양자화기의 연산량은 제안된 양자화기의 약 7.7×10^8 배에 달하게 된다. 이에 더하여, 파라미터 당 탐색 수를 늘릴수록 두 양자화기 간의 연산량의 차이는 더욱 커지게 된다. 예를 들어, 16-level의 경우 N_s 를 20으로 증가시키면, 최대 상호 정보 양자화기는 제안된 양자화기에 비하여 약 4.0×10^{11} 배의 연산량을 필요로 한다.

제안된 양자화기의 연산량은 경사추적법(gradient descent method)과 같은 최적화 방법을 이용하여 추가적으로 감소시킬 수 있다. 그림 5와 6은 각각 7-level 및 10-level 양자화기에서 r_0 및 r_2 의 값을 변화시켜 가면서 측정한 상호 정보의 크기를 나타낸다. 이 때, r_1 값은 최적화된 값으로 고정되었다. 그림 5와 6에서 볼 수 있듯이, 제안된 양자화기에서는 상호 정보가 볼록 함수(convex function)의 형태를 하고 있기 때문에, 최대 상호 정보를 내는 파라미터 조합을 유일하게 결정하는 것이 가능하다. 따라서 경사추적법 및 뉴턴 방법(Newton's method)과 같은 최적화 방법을 이용하여 최대 상호 정보를 내는 파라미터의 조합을 유일하게 결정할 수 있다.

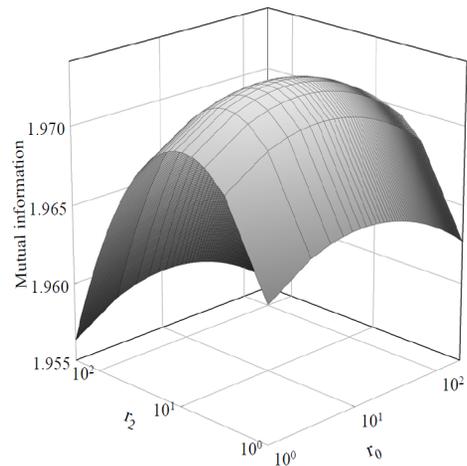


그림 5. r_0 및 r_2 의 변화에 따른 상호 정보의 크기 (7-level 양자화기).
 Fig. 5. Mutual information while changing the values of r_0 and r_2 (7-level quantizer).

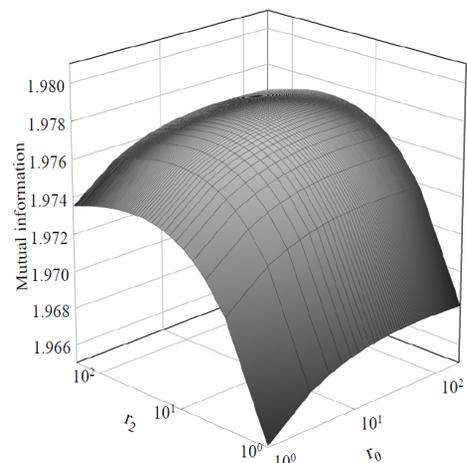


그림 6. r_0 및 r_2 의 변화에 따른 상호 정보의 크기 (10-level 양자화기).
 Fig. 6. Mutual information while changing the values of r_0 and r_2 (10-level quantizer).

앞서 설명한 소모적인 탐색이나 최적화 방법을 사용하면, 특정 신호 대 잡음비에 최적화된 양자화기를 비교적 적은 연산량으로 찾는 것이 가능하다. 따라서 본 논문에서 제안한 양자화기를 이용한다면 신호 대 잡음비를 변화시켜가며 각 상황에 최적화된 다수의 양자화기를 찾는 것이 가능하다. 이렇게 얻은 양자화기의 파라미터 값들은 낸드 플래시 메모리 내부의 별도의 저장 공간에 참조표(look-up table) 형태로 저장된다. 낸드 플래시 메모리 컨트롤러(controller)는 메모리 읽기를 수행할 때마다 이 참조표를 이용하여 현재 신호 대 잡음비에 최적인 양자화기를 얻을 수 있다.

IV. 실험 결과

본 논문에서는 제안된 적응적 비균일 양자화기의 성능을 평가하기 위하여, 그림 7과 같은 낸드 플래시 메모리 모델^{[9][10]}을 이용하여 모의실험을 수행하였다. 심볼 1, 2, 3의 문턱 전압은 0.4 V, 1.9 V, 3.5 V로 가정하였고, 지우기 과정 후 문턱 전압의 확률 분포는 평균이 -2.0 V 이고, 표준편차가 0.40 V 인 정규 분포로 가정하였다. 프로그래밍 과정은 ISPP(incremental step pulse programming)^[11]을 이용하여 수행되었고, 이 때 ΔV_{pp} 는 0.4 V로 설정되었다. 셀 간 간섭(cell-to-cell interference)은 주변 셀의 문턱 전압 변화의 선형 조합(linear combination)으로 모델링되었다. 낸드 플래시 메모리에서는 시간이 지남에 따라 플로팅 게이트의 전하가 조금씩 누설되는데, 이에 따른 잡음(data retention noise)은 정규 분포를 따르는 확률 변수로 가정하였고, 그 평균과 표준편차는 데이터 리텐션 시간의 로그 값에 비례하도록 설정하였다.

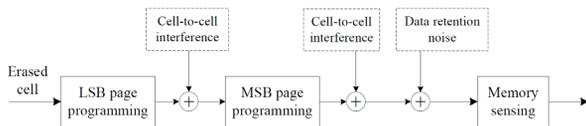


그림 7. 모의실험에 사용된 낸드 플래시 메모리 모델.
Fig. 7. NAND flash memory model used for simulations.

그림 8에서는 7-level 및 10-level 양자화기에서 신호 대 잡음비(E_b/N_0) 값에 따라 최대 상호 정보를 내는 r_k 값의 변화를 나타낸다. 그림 8의 정보는 앞서 설명한 바와 같이 낸드 플래시 메모리내의 참조표의 형태로 저장되는데, 이를 이용하면 특정 신호 대 잡음

비에 맞는 최적의 파라미터 값을 쉽게 얻을 수 있다. 예를 들어, 현재 신호 대 잡음비가 9.0 dB 이고 10-level 양자화기를 사용하고자 한다면, r_0 , r_1 , r_2 를 각각 55, 32, 29로 잡으면 에러 정정 성능을 최대화하는 양자화기를 얻을 수 있다.

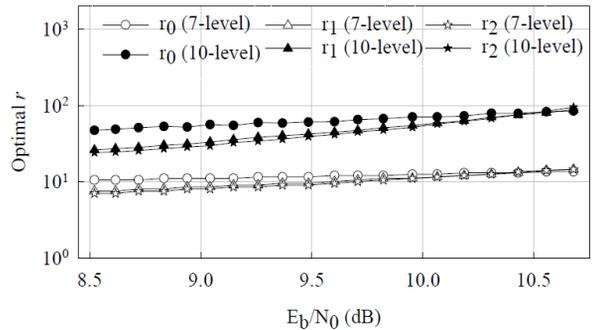


그림 8. E_b/N_0 에 따른 최대 상호 정보를 내는 r_k 값의 변화 (7-level 및 10-level 양자화기).

Fig. 8. Optimal values of r_k according to E_b/N_0 (7- and 10-level quantizers).

제안된 양자화기에서의 최적의 파라미터 값은 신호 대 잡음비에 따라 달라진다. 즉, 신호 대 잡음비가 작아지면, r 값을 작게 하여 좁은 전압 영역을 소거 구간으로 설정하는 것이 더 나은 에러 정정 능력을 보인다. 반대로, 신호 대 잡음비가 커지면 소거 구간의 폭을 넓게 할 때, 더 나은 에러 정정 성능을 얻을 수 있다. 결국, 제안된 양자화기에서는 소거 구간에 해당되는 셀의 개수가 거의 일정하도록 소거 구간의 넓이가 결정된다고 생각할 수 있다. 7-level 과 10-level의 두 경우를 비교해 보면, 후자 경우에 r_k 가 더 큰 값을 갖는 것을 관찰할 수 있다. 7-level 양자화기의 경우에는 각 심볼의 경계면에 오직 하나의 소거 구간만이 존재하고, 10-level의 경우에는 두 개의 소거 구간이 존재한다. 따라서 10-level 양자화기 경우에는 소거 구간을 좀 더 넓게 설정하여 더 많은 셀을 소거 비트(eraser bit)로 할당하더라도, 연판정 에러 정정 통하여 비트 에러를 정정하는 것이 가능하다.

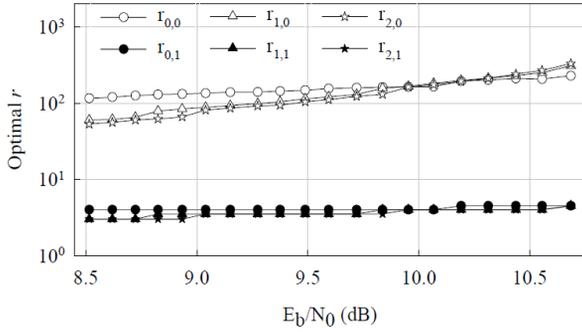


그림 9. E_b/N_0 에 따른 최대 상호 정보를 내는 r_k 값의 변화 (13-level 양자화기).

Fig. 9. Optimal values of r_k according to E_b/N_0 (13-level quantizer).

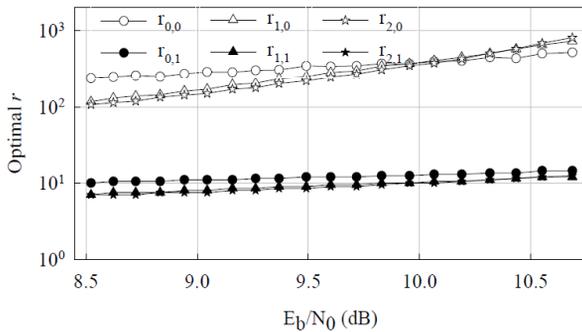


그림 10. E_b/N_0 에 따른 최대 상호 정보를 내는 r_k 값의 변화 (16-level 양자화기).

Fig. 10. Optimal values of r_k according to E_b/N_0 (16-level quantizer).

그림 9와 10은 13-level 및 16-level 양자화기에서 신호 대 잡음비에 따른 최적의 파라미터 값을 나타낸다. 이들 양자화기는 총 여섯 개의 파라미터들에 의하여 구조가 결정되는데, $r_{k,0}$ 는 소거 구간의 바깥쪽 경계를, $r_{k,1}$ 은 안쪽 경계를 결정하는데 이용된다. 그림 9와 10에서 볼 수 있듯이, 최적의 $r_{k,0}$ 값은 신호 대 잡음비에 따라 수십에서 수백까지 상당히 다양한 값을 가진다. 이에 반하여, 최적의 $r_{k,1}$ 값은 거의 일정하게 유지되는 것을 볼 수 있다. 13-level과 16-level 양자화기를 비교해 보면, 16-level의 경우 13-level에 비하여 각 심볼의 경계면당 하나의 소거 구간을 더 갖고 있기 때문에, 소거 구간의 넓이를 더 넓게 하는 것이 유리하다.

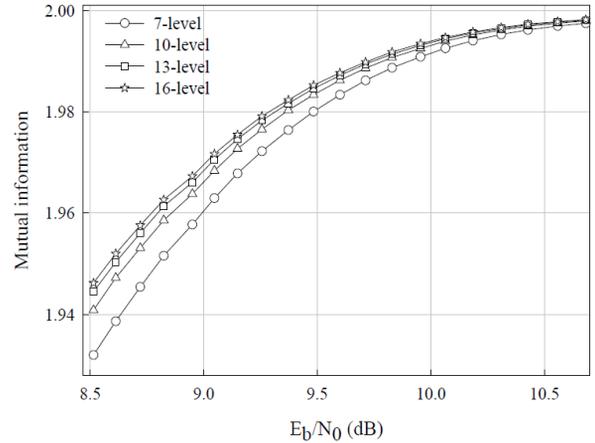


그림 11. E_b/N_0 에 따른 최대 상호 정보의 크기.

Fig. 11. Maximum mutual information value according to E_b/N_0 .

그림 11은 신호 대 잡음비를 변화시켜가며, 제안된 양자화기로 얻은 최대 상호 정보의 크기를 나타낸다. 신호 대 잡음비가 큰 경우에는 모든 양자화기가 2.0에 가까운 비슷한 값을 가진다. 이에 반하여, 신호 대 잡음비가 작아질수록 최대 상호 정보의 크기 차이가 커지며, 특히 7-level 양자화기의 경우에는 다른 양자화기에 비하여 훨씬 더 낮은 최대 상호 정보 값을 보인다. 이러한 최대 상호 정보의 차이는 결국 에러 정정 능력의 차이로 이어진다.

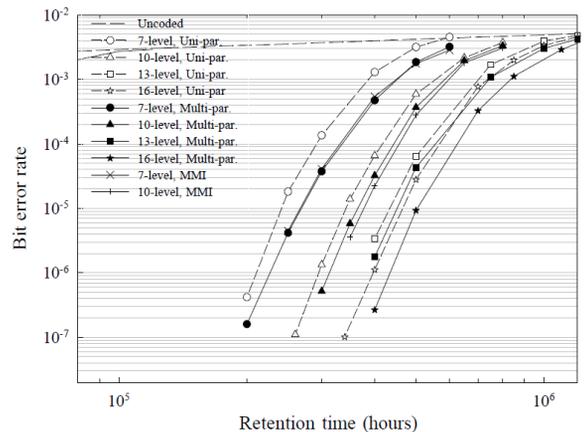


그림 12. 연관정 에러 정정 성능곡선.

Fig. 12. Bit error rate performances for soft-decision error correction.

본 논문에서는 제안된 적응적 비균일 양자화기를 이용하여 연관정 에러 정정을 수행하였고, 그 결과를 그림 12에 나타내었다. 본 실험에서는 가상의 낸드 플래시 메모리의 MSB(most significant bit) 페이지를 (68254, 65536) EG(Euclidean geometry)

LDPC 코드를 이용하여 복호하였다. 또한, 에러 정정 능력의 비교를 위하여 동일한 조건에서 최대 상호 정보 양자화기 및 단일 파라미터 양자화기를 이용하여 연판정 에러 정정을 수행한 결과도 그림 12에 함께 도시하였다. 단일 파라미터 양자화기에서는 각 신호 대 잡음비에 대하여 가장 좋은 에러 정정 성능을 내는 결과를 나타내었다. 최대 상호 정보 양자화기와 다중 파라미터 양자화기의 소모적인 탐색 과정에서는 각 파라미터 당 10회의 탐색을 수행하였다. 앞서 살펴봤듯이, 최대 상호 정보 양자화기에서는 파라미터 수의 지수 승에 비례하여 연산량이 증가한다. 따라서 파라미터의 수가 늘어날수록 최대 상호 정보 양자화기를 찾기 위해서는 매우 긴 시간이 소요되기 때문에, 본 논문에서는 10-level 이하의 최대 상호 정보 양자화기에 대해서만 실험을 수행하였다. 참고로, 3.2 GHz의 동작 주파수를 가지는 CPU로 하나의 10-level 최대 상호 정보 양자화기를 구할 때 약 3시간 정도가 소요되었고, 13-level 최대 상호 정보 양자화기를 얻기 위해서는 약 3,000 시간 정도가 소요될 것으로 예상되었다.

그림 12에서, 본 논문에서 제안된 양자화기는 최대 상호 정보 양자화기 보다 열화된 에러 정정 능력을 보였다. 최대 상호 정보 양자화기는 파라미터의 모든 조합을 통하여 상호 정보를 최대로 하는 조합을 찾기 때문에 전역 최대(global maximum) 점을 찾는 과정으로 볼 수 있다. 반면에, 제안된 방법은 특정 구조를 만족시키는 양자화기 중에서 상호 정보가 최대인 점을 찾는 방법이다. 따라서 이는 국부 최대(local maximum) 점을 찾는 과정과 동일하다. 이러한 이유로 최대 상호 정보 양자화기는 제안된 방법보다 우수한 성능을 낸다. 그러나 그림 12에서는 제안된 양자화기의 성능이 최대 상호 정보 양자화기의 성능에 매우 근접하여 성능 열화의 정도가 크지 않음을 관찰할 수 있다. 예를 들어, 7-level에서는 제안된 양자화기와 최대 상호 정보 양자화기가 거의 동일한 에러 정정 성능을 보였고, 10-level에서도 성능의 차이가 미미하였다. 반면에, 제안된 방법으로는 13-level 이상의 양자화기도 적은 연산만으로 구현할 수 있었다.

제안된 양자화기와 단일 파라미터 양자화기를 비교했을 때, 모든 경우에 제안된 양자화기가 더 우수한 에러 정정 성능을 보였다. 비록 단일 파라미터 양자화기가 상대적으로 더 적은 연산량을 요구하지만, 두 경우 모두 일반적인 PC를 이용하여 구현하는 것이 가능하다. 따라서 이 경우에는 더 우수한 성능을 보이는

다중 파라미터 양자화기를 사용하는 것이 바람직하다.

V. 결 론

본 논문에서는 낸드 플래시 메모리의 연판정 에러 정정에 필수적인 적응적 양자화기의 최적화방법을 제안하였다. 제안된 양자화기는 최대 여섯 개의 파라미터로 표현되는 복잡도가 낮은 구조를 가지고 있고, 입력 심볼과 출력 심볼 간의 상호 정보를 최대화하도록 설계되었다. 낸드 플래시 메모리에서는 데이터 리텐션 및 PE (program erase) 사이클에 의하여 신호 대 잡음비가 계속적으로 변화한다. 따라서 최대의 에러 정정 성능을 내기 위해서는 메모리 블록의 상태에 따라 양자화기를 다르게 사용하여야 하는데, 본 논문에서 제시한 저복잡도의 양자화기는 이러한 상황에 매우 유리하다. 본 논문에서는 가상의 낸드 플래시 메모리를 이용하여 제안한 양자화기의 연판정 에러 정정을 수행하였다. 본 논문에서 제안한 다중 파라미터 양자화기는 적은 연산량만으로 최대 상호 정보 양자화기에 근접한 수준의 에러 정정 능력을 보였다.

References

- [1] W. Liu, J. Rho, and W. Sung, "Low-power high-throughput BCH error correction VLSI design for multi-level cell NAND flash memories," in *Proc. IEEE Workshop Signal Process. Syst. Design Implementation (SIPS)*, pp. 303-308, Banff, Canada, Oct. 2006.
- [2] B. Chen, X. Zhang, and Z. Wang, "Error correction for multi-level NAND flash memory using Reed-Solomon codes," in *Proc. IEEE Workshop Signal Process. Syst. (SiPS)*, pp. 94-99, Washington D.C., U.S.A., Oct. 2008.
- [3] G. Dong, N. Xie, and T. Zhang, "On the use of soft-decision error correction codes in NAND flash memory," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, no. 2, pp. 429-439, Feb. 2011.
- [4] C. Yang, Y. Emre, and C. Chakrabarti, "Product code schemes for error correction in MLC NAND flash memories," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20,

- no. 12, pp. 2302-2314, Dec. 2012.
- [5] J. Wang, T. Courtade, H. Shankar, and R. Wesel, "Soft information for LDPC decoding in flash: Mutual-information optimized quantization," in *Proc. IEEE Global Commun. Conf. (GLOBECOM 2011)*, pp. 1-6, Houston, U.S.A., Dec. 2011.
- [6] D. Lee and W. Sung, "Estimation of NAND flash memory threshold voltage distribution for optimum soft-decision error correction," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 440 - 449, Jan. 2013.
- [7] J. Cho, J. Kim, and W. Sung, "Optimal output quantization of binary input AWGN channel for belief-propagation decoding of LDPC codes," in *Proc. IEEE Workshop on Signal Process. Syst. (SiPS)*, pp. 282-287, Quebec City, Canada, Oct. 2012.
- [8] J. Kim, D. Lee, and W. Sung, "Performance of rate 0.96 (68254, 65536) EG-LDPC code for NAND flash memory error correction," in *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 7029 - 7033, Ottawa, Canada, June 2012.
- [9] G. Dong, Y. Pan, N. Xie, C. Varanasi, and T. Zhang, "Estimating information-theoretical NAND flash memory storage capacity and its implication to memory system design space exploration," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20, no. 9, pp. 1705-1714, Sep. 2012.
- [10] D. Lee and W. Sung, "Monte Carlo simulation of NAND flash memory channel in the presence of cell to cell interference," in *Proc. KICS Winter Conf.*, pp. 407-407, Yongpyoung, Korea, Feb. 2011.
- [11] K.-D. Suh B.-H. Suh, Y.-H. Lim, J.-K. Kim, Y.-J. Choi, Y.-N. Koh, S.-S. Lee, S.-C. Kwon, B.-S. Choi, J.-S. Yum, J.-H. Choi, J.-R. Kim, and H.-K. Lim, "A 3.3 V 32 MB NAND flash memory with incremental step pulse programming scheme," *IEEE J. Solid-State Circuits*, vol. 30, no. 11, pp. 1149-1156, Nov. 1995.

이 동 환 (Dong-hwan Lee)



릴 구현

2009년 2월 서울대학교 전자공학학사
 2009년 3월~현재 서울대학교 전기정보공학부 석박통합과정 <관심분야> 낸드 플래시 메모리용 에러 정정 시스템 개발, 신호처리 알고리즘의 병

성 원 용 (Wonyong Sung)



1978년 2월 서울대학교 전자공학학사
 1980년 2월 한국 과학 기술원 전기공학석사
 1987년 University of California, Santa Barbara 전기공학박사
 1989년~1993년 서울대학교 반도체 공동 연구소 조교수
 1993년~1999년 서울대학교 전기공학부 부교수
 1999년~현재 서울대학교 전기공학부 교수
 1997년~1999년 서울대학교 반도체 공동 연구소 집적 시스템 설계 센터 센터장
 <관심분야> 고정 소수점 최적화 툴 개발, 멀티미디어 VLSI 개발, 낸드 플래시 메모리용 에러 정정 시스템 개발, 신호처리 알고리즘의 병렬 구현. 음성 인식 시스템 개발