미지의 선형 순회부호에 대한 복원기법

종신회원 정 하 봉*, 정회원 장 환 석*°, 준회원 조 원 찬**, 정회원 박 철 순***

Reconstruction of Linear Cyclic Codes

Habong Chung* *Lifelong Member*, Hwanseok Jang*° *Regular Member*,
Won-chan Cho** *Associate Member*, Cheol-sun Park*** *Regular Member*

요 약

잡음이 있는 채널을 통한 디지털 통신에서는 채널 잡음에 대항하기 위해 오류정정부호(채널부호)를 사용하게 된다. 만일 송신측의 협조 없이 전송정보를 알아내려면 사용된 채널부호를 복원하는 것이 무엇보다 중요하다. 본 논문에서는 잡음에 오염된 수신 비트열로부터 사용된 채널부호의 여러 파라메타를 추출하여 궁극적으로 채널부호를 복원하는 채널부호 복원기법 중 순회부호(cyclic code)의 복원 기법을 제안한다.

Key Words : 블록부호, 상보부호, 순회부호, 채널부호 복원기법, Channel code recognition

ABSTRACT

In most digital communication systems over the noisy channel, some form of forward error correction scheme is employed for reliable communications. If one wants to recover the transmitted message without any knowledge of the error correcting codes employed, it is of utmost importance to figure out and reconstruct the error correcting codes. In this paper, we propose two algorithms of reconstructing linear cyclic codes from the corrupted received bit sequence, one for general linear binary cyclic codes and the other for Reed-Solomon codes. For two algorithms, we ran computer simulations and the performances are shown to be superior to those with the conventional LWM method.

Ⅰ. 서 론

잡음이 있는 채널을 통한 디지털 통신에서는 통상 정보비트에 잉여의 비트를 추가함으로써 채널 잡음으로 인한 오류를 정정할 수 있는 채널부호를 사용하게 마련이다. 만일 정상적인 수신자가 아닌 제 3의 적대적인 도청자가 송신측의 협조 없이 정보 메시지를 알아내고자 한다면 상대방이 어떤 종류의 채널부호를 사용했는지를 알아내어 이를 복원하는 것이 무엇보다 중요하다. 이처럼 오류가 첨가된 수신 비트열로부터 사용된 채널부호의 여러 파라메타를 추출하여 궁극적으로 사용된 채널부호를 복원하는 기법을 채널 부호의 복원 기법이라 한다.

본 논문에서는 다양한 선형 순회부호(cyclic code)의 복원 기법을 제안한다. 통상 $[n, k]$ 선형블록부호(linear block codes)의 부호화기는 $k \times n$ 생성행렬 G를 이용하여 k 비트의 정보블록 $\boldsymbol{m}$으로부터 n 비트의 부호어 $\boldsymbol{c}$를 $\boldsymbol{c} = \boldsymbol{m}G$로 생성하게 되나, 순회부호의 경우 생성행렬보다는 생성다항식 $g(x)$를 이용한 표현이 좀 더 효율적이다. 즉, n비트의 부호어 $\boldsymbol{c} = (c_0, c_1, \cdots, c_{n-1})$

※ 본 연구는 국방과학연구소 과제의 지원을 받아 수행되었음 (과제번호 : ADD-2009-038-A9D018F)
※ 본 연구는 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구 사업임 (과제번호 : No. 2011-0016682)
* 홍익대학교 전자전기공학부 통신연구실 (habchung@hongik.ac.kr, hsjang@mai.hongik.ac.kr) (° : 교신저자)
** LG 신규개발사업부 (A115235@mail.hongik.ac.kr) *** 국방과학연구소 (csun@add.re.kr)
논문번호 : KICS2011-06-250, 접수일자 : 2011년 6월 13일, 최종논문접수일자 : 2011년 10월 16일

는 부호어 다항식 $c(x)=\sum_{i=0}^{n-1} c_i x^i$ 로 표현되며 모든 부호어 다항식은 반드시 생성다항식 $g(x)$의 배수여야 한다. 따라서 순회부호의 복원기법에서 궁극적으로 찾아내고자 하는 부호의 파라메타는 생성다항식 $g(x)$가 된다. 일반적으로 복원기술의 최초 단계는 부호어의 길이 n과 부호어의 시작점을 알아내는 동기(synchronization) 단계이다. 이후 k 값을 추정하게 되고 나아가 최종의 파라메타를 추출하게 된다. 본 논문에서는 동기단계가 완료된 이후를 가정한다.

본 논문의 구성은 다음과 같다. I장의 서론에 이어 II장에서는 상보(dual)부호를 이용한 기존의 선형블록부호의 복원기술을 알아본다. III장에서는 생성다항식을 이용한 이진(binary) 순회(cyclic)부호의 복원기법을 제안한다. IV장에서는 비이진 순회부호의 대표격인 Reed-Solomon 부호의 복원기술을 제안한다. 마지막으로 V장에서는 향후의 연구방향을 제시하며 결론을 맺는다.

Ⅱ. 이진 블록부호의 복원 기법

오류확률이 τ인 BSC (binary symmetry channel)를 통과하여 경판정(hard decision)된 일련의 수신 비트열이 있다. 사용된 채널부호는 길이가 n인 선형 블록부호이고 일단 동기 단계가 성공적으로 수행된 것을 가정한다. 채널을 통과한 수신 비트열을 N개의 수신워드로 분리한 후 각각의 수신워드가 행이 되도록 $N\times n$ 행렬 X를 만든다. 여기서 수신워드란 전송된 부호어에 채널 오류가 더해진 n 비트 길이의 이진 벡터를 의미한다. 따라서 행렬 X는 부호어로 구성된 행렬 C와 전송오류로 구성된 행렬 E의 합으로 볼 수 있다. 일반적인 블록 부호의 복원은 상보부호(dual code)를 이용한 방법이 사용된다[1].

정의 1. $[n,k]$ 선형블록부호를 C 라고 하면 이 부호의 상보부호 $C^{\perp}$는 다음과 같이 정의된다.

$$C^{\perp}=\{\boldsymbol{h}\boldsymbol{c}^{\boldsymbol{T}}=0,\boldsymbol{h}\mid\forall\ \boldsymbol{c}\in C\} \tag{1}$$

이제 n 비트의 길이를 가진 임의의 이진 벡터 $\boldsymbol{h}$를 생각해보자. 만일 $\boldsymbol{h}$가 상보부호 $C^{\perp}$의 부호어라면,

$$\boldsymbol{h}X^T=\boldsymbol{h}\left(C^T+E^T\right)=\boldsymbol{h}E^T$$

가 되어 $\boldsymbol{h}X^T$의 해밍무게는 비트오류율 τ와 벡터 $\boldsymbol{h}$의 해밍무게의 영향을 받게 될 것이며 실제로 이 경우 $wt(\boldsymbol{h}X^T)$의 평균값은

$$E\{wt(\boldsymbol{h}X^T)\}=\{1-(1-2\tau)^{wt(\boldsymbol{h})}\}\cdot\frac{N}{2} \tag{2}$$

이 된다[1]. 반면 $\boldsymbol{h}\not\in C^{\perp}$ 인 경우에는 $wt(\boldsymbol{h}X^T)$의 평균값은 $N/2$이 될 것이다. 따라서 적당한 임계값을 설정하여 벡터 $\boldsymbol{h}$가 $C^{\perp}$에 속하는지 아닌지를 판별할 수 있다. 이런 방식으로 $(n-k)$개의 선형 독립인 $\boldsymbol{h}$를 찾아내면 $C^{\perp}$를 알 수 있고 따라서 원래의 부호 C를 복원할 수 있게 된다. 이런 방식의 복원기법을 LWM (low weight moment) 기법이라고 한다. LWM 기법은 개개의 입력벡터 $\boldsymbol{h}$를 대상으로 $C^{\perp}$에 속하는지를 판별하는 알고리즘(first-order algorithm) 외에도 복수개의 $\boldsymbol{h}$들을 대상으로 판별하는 향상된 알고리즘의 도입이 가능하다[1]. 예를 들어 두 벡터 $\boldsymbol{h}$와 $\boldsymbol{h}'$을 동시에 고려하는 경우, 만일 $\boldsymbol{h}$와 $\boldsymbol{h}'$ 모두 $C^{\perp}$에 속해있다면 $\boldsymbol{h}+\boldsymbol{h}'$ 역시 $C^{\perp}$에 속해 있기 때문에 $\boldsymbol{h}X^T$와 $\boldsymbol{h}'X^T$ 간의 상관 정보를 추가로 판별에 이용할 수 있다. 여기서는 first-order 알고리즘에 대하여 간략히 살펴보겠다. First-order 알고리즘은 다음에 주어진 B값을 계산하여

$$B=\left(1+(1-2\tau)^{wt(\boldsymbol{h})}\right)\left(\frac{1-(1-2\tau)^{wt(\boldsymbol{h})}}{1+(1-2\tau)^{wt(\boldsymbol{h})}}\right)^{wt(\boldsymbol{h}X^t)} \tag{3}$$

이 값을 기준값과 비교함으로써 $\boldsymbol{h}$가 $C^{\perp}$에 속하는지를 판별하는 알고리즘을 말한다[1]. 이 기준값은 오경보(false alarm) 확률과 검출(detection) 확률의 비와 $p(\boldsymbol{h}\in C^{\perp})$의 함수로 표현할 수 있다. 식 (3)은 채널의 전송오류확률이 τ인 경우이다. 식 (3)에서 만약 τ와 $wt(\boldsymbol{h})$가 주어진다면, 그에 따라 만족해야 하는 $wt(\boldsymbol{h}X^t)$값을 구할 수가 있고, 그 값이 임계값이 된다.

그림 1은 (64, 34) 이진 블록부호를 대상으로 무작위로 선택된 벡터 $\boldsymbol{h}$들의 해밍무게에 따른 $wt(\boldsymbol{h}X^T)$를 도시한 그림이다.

그림에서 보면, 진하게 표시된 점선을 기준으로 하여 왼쪽과 오른쪽으로 구분됨을 확인할 수 있다. 점선은 $wt(\boldsymbol{h})$에 따른 임계값을 나타내고, 이 때 사용된 기준값은 10^5이었으며, 점선의 왼쪽에 위치한 경우 $\boldsymbol{h}$가 상보부호에 속한 것으로 판단하게 된다.

상보부호를 이용한 복원기법에서 가장 중요한 부분은 해밍무게가 작은 후보 $\boldsymbol{h}$를 생성하는 일이다. 식 (2)

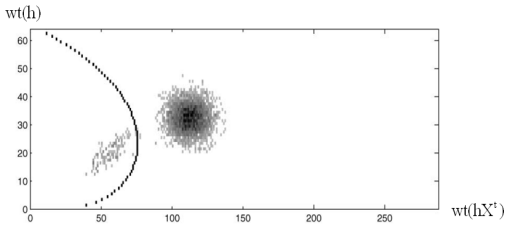

그림 1. [1] (64,34) code, $\tau = 1/64$, $N : 288$

에서 볼 수 있듯이 $wt(\boldsymbol{h})$가 크게 되면 $\boldsymbol{h} \notin C^{\perp}$ 인 경우와 비교하여 평균값에 큰 차이가 없어 그만큼 판별이 어렵기 때문이다. 따라서 이 기법은 상보부호의 최소해밍무게가 비교적 큰 부호에 적용하는 데는 근본적인 어려움이 있다.

Ⅲ. 이진 순회부호의 복원 기법

순회부호란 부호어의 순환 이동된 형태 모두가 역시 부호어가 되는 선형부호를 의미한다. 즉, $(c_0, c_1, \cdots, c_{n-1})$이 부호어라면 $(c_{n-1}, c_0, \cdots, c_{n-2})$ 도 역시 부호어가 되는 부호를 말한다. 부호화기와 복호기 모두가 시프트 레지스터를 이용하여 쉽게 설계할 수 있다는 장점 때문에 널리 사용되는 블록부호들은 대부분 순회부호이다. 순회부호에서는 부호어 $\boldsymbol{c} = (c_0, c_1, \cdots, c_{n-1})$ 를 부호어 다항식 $c(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$ 으로 표현한다. 부호어를 다항식으로 표현하게 되면 한 비트 순환 이동된 부호어 $(c_{n-1}, c_0, c_1, \cdots, c_{n-2})$ 는 $xc(x) \pmod{x^n - 1}$로 쉽게 표현할 수 있다. 순회부호에서 생성 행렬의 역할을 하는 것은 생성다항식 $g(x)$이다. 즉, 순회부호의 모든 부호어 다항식 $c(x)$은 반드시 $g(x)$의 배수여야 하고 그 역도 성립한다. 따라서 순회부호의 최종 복원 파라메타는 생성다항식 $g(x)$가 된다. $[n, k]$ 이진 순회부호의 생성다항식 $g(x)$는 차수가 $n-k$인 다항식으로 $x^n - 1$의 인수인 기약(irreducible) 다항식들 중 일부의 곱으로 표현된다. 이 장에서 제안하는 복원기법의 핵심은 주어진 수신 다항식들이 각각의 기약 다항식을 인수로 가지고 있는가의 여부를 판별하여 궁극적으로 $g(x)$를 복원하고자 하는 것이다.

동기 단계를 거쳐 부호어의 시작점과 부호어의 길이 n를 알아내었다고 가정하면, 생성다항식의 인수로 가능한 기약 다항식들은 $x^n - 1$의 인수들이 된다. 다음의 예제를 통해 생성다항식

$$g(x) = (x^4 + x^3 + 1)(x^4 + x^3 + x^2 + x + 1)$$

를 가진 (15,7) BCH 부호의 경우를 보자. 생성다항식의 인수로 가능한 기약다항식은 $x^{15} - 1$의 인수이므로,

$$q_1(x) = x + 1, \qquad q_2(x) = x^2 + x + 1,$$
$$q_3(x) = x^4 + x + 1, \quad q_4(x) = x^4 + x^3 + 1,$$
$$q_5(x) = x^4 + x^3 + x^2 + x + 1$$

의 5개의 다항식들이다. 아래의 표는 오류를 포함하는 총 14개의 수신다항식들을 위의 가능한 기약다항식들로 나눈 결과를 정리한 표이다.

표 1에서 '1'이 의미하는 것은 해당 열의 기약 다항식이 해당 행의 수신 다항식을 나눈다는 것을 의미하고, '0'은 나누지 못함을 의미한다. 표 1의 수신 다항식들은 오류확률이 약 0.05인 BSC 채널을 통과한 예이다. 표 1과 같은 서브행렬을 이용한 복원 기법은 수직 접근법과 수평 접근법으로 나누어 볼 수 있다.

표 1. 오류가 포함된 경우의 수신 다항식들

	q_1	q_2	q_3	q_4	q_5
수신 다항식 1	1	1	0	1	1
수신 다항식 2	1	0	0	0	0
수신 다항식 3	0	0	1	1	1
수신 다항식 4	1	0	0	1	1
수신 다항식 5	0	1	0	0	0
수신 다항식 6	1	0	0	1	1
수신 다항식 7	0	1	0	1	1
수신 다항식 8	0	0	0	0	0
수신 다항식 9	1	0	0	1	1
수신 다항식 10	0	0	0	1	1
수신 다항식 11	1	0	0	1	1
수신 다항식 12	1	0	0	0	0
수신 다항식 13	0	0	0	1	1
수신 다항식 14	0	0	0	1	1

3.1 수직 접근법

서브행렬의 세로축의 1의 개수에 따라 해당 열의 기약 다항식이 생성 다항식의 인수로 사용되었는지 판단하는 방법이다. 간단히 말해, 많은 개수의 수신다항식들의 인수가 되는 기약 다항식을 생성다항식의 인수로 판정하겠다는 것이다. 수신다항식들은 오류를 포함하기 때문에 판정을 위해서는 각각의 열의 1의 개수에 대한 임계값이 설정되어야 하며, 이 때 임계값은 해당하는 기약 다항식의 차수에 따라 다르게 된다. 기약 다항식 $q(x)$의 차수가 j라 하면, 임의의 다항식

$a(x)$이 $q(x)$의 배수가 될 확률은 2^{-j}이 된다. 그 이유는 $a(x)$가 $q(x)$의 배수가 되려면 $a(x)$를 $q(x)$로 나눈 나머지 다항식 (즉, 임의의 $j-1$ 차 다항식)이 0 가 되어야 하기 때문이다. 따라서 기약 다항식 $q(x)$가 생성다항식 $g(x)$의 인수가 아닌 경우에는 해당 열의 1의 개수의 평균 N_1은 $N_1 = N \cdot 2^{-j}$ 라고 볼 수 있다. (물론 엄밀한 의미에서는 부호어 다항식 $c(x)$와 오류 다항식 $e(x)$의 합인 수신다항식 $r(x)$를 임의의 다항식이라고 가정할 수는 없고, 그 값은 채널오류확률 τ와 사용되는 부호에 따라 다르겠지만 많은 모의실험을 통해 보면 임의의 다항식이라는 가정에 무리가 없음을 알 수 있다.) 반면에 기약 다항식 $q(x)$가 생성다항식 $g(x)$의 인수인 경우, 해당 열의 평균 1의 개수 N_1은 다음과 같이 근사될 수 있다. 기약 다항식 $q(x)$가 오류다항식 $e(x)$를 나누는 사건을 A라고 하자.

$$\begin{aligned} N_1 &= Np(A) \\ &= N\sum_{w=0}^{n} p(A|wt(e)=w)\,p(wt(e)=w) \\ &= N(1-\tau)^n \sum_{w=0}^{n} \binom{n}{w}\left(\frac{\tau}{1-\tau}\right)^w p(A|wt(e)=w) \end{aligned} \tag{4}$$

식 (4)의 정확한 계산을 위해서는 $p(A|wt(e)=w)$를 알아야 한다. 이 확률을 $p(q(x),w)=p(A|wt(e)=w)$ 라고 정의하자. 일반적으로 $p(q(x),w)$ 를 알아낸다는 것은 쉽지 않다. 다만, $q(x)$의 차수나 w값이 작은 몇몇 경우에 대해서는 다음의 결과를 얻을 수 있다.

정리 2 ($w=0,1,2$ 인 경우)

$$\begin{aligned} &p(q(x),0)=1 \\ &p(q(x),1)=0 \\ &p(q(x),2)=\begin{cases} 0, & q(x) \quad is\ primitive \\ \dfrac{n-m}{m(n-1)}, & q(x)\,|\,x^m+1, m(<n)\,|\,n \end{cases} \end{aligned} \tag{5}$$

증명: $w=0,1$의 경우는 자명하므로 $w=2$인 경우만 증명하겠다. $q(x)$가 원시 다항식, 즉 $q(x)$의 근이 1의 원시 n승근인 경우는 $q(x)$의 배수인 최소 차수의 이항식이 x^n+1이므로 차수가 $n-1$ 이하인 그 어떤 이항식도 $q(x)$를 인수로 가질 수 없다. 만일 $q(x)$의 근이 1의 원시 m승근이라면 $x^i(x^{lm}+1),\quad 0\le i\le n-lm-1,\ 1\le l\le \dfrac{n}{m}-1$, 형태의 이항식은 모두 $q(x)$의 배수가 된다. 따라서 이러한 이항식의 총 개수는

$$\sum_{l=1}^{n/m-1}(n-lm)=\frac{n}{2}\left(\frac{n}{m}-1\right)$$

가 되어

$$p(q(x),2)=\frac{n\left(\dfrac{n}{m}-1\right)}{2\binom{n}{2}}=\frac{n-m}{m(n-1)}$$

이다. □

정리 3 ($q(x)=x+1$ 인 경우)

$$p(x+1,w)=\begin{cases} 1, & w \quad even \\ 0, & w \quad odd \end{cases} \tag{6}$$

증명: $x+1$의 임의의 배수는 모두 짝수 개수의 항을 가지므로 자명함. □

정리 4

$n=2^m-1$ 이고 $q(x)$가 원시 다항식이라고 하자.

$$p(q(x),w)=\frac{A_w}{\binom{n}{w}} \tag{7}$$

여기서 A_w는 해밍무게가 w인 부호어의 개수이고 다음의 순환 방정식을 만족한다.

$$\begin{aligned} &A_0=1,\quad A_1=0 \\ &(w+1)A_{w+1}+A_w+(n-w+1)A_{w-1}=\binom{n}{w}, \end{aligned} \tag{8}$$

증명: $q(x)$가 원시 다항식인 경우, $e(x)$가 $q(x)$의 배수라는 말은 $e(x)$가 $[2^m-1, 2^m-1-m]$ Hamming 부호 C의 부호어 다항식이라는 말이 된다. 따라서 A_w를 부호 C에서 해밍무게가 w인 부호어의 개수라고 하면 $p(q(x),w)$ 는 식 (7)과 같이 표현할 수 있다. C의 상보부호 $C^{\perp}$의 부호어는 모두 해밍무게가 2^{m-1}이라는 점과 MacWilliams identity를 이용하면 A_w가 식 (8)과 같은 순환 방정식을 만족한다는 것을 보일 수 있다[6, Example (4), p. 129]. □

예를 들어 표 1의 경우, $q_4(x)=x^4+x^3+1$가 속해 있는 열을 보자. 만일 $q_4(x)$가 생성다항식 $g(x)$의 인수가 아닌 경우에는 해당 열의 1의 개수의 평균 N_1은 $N_1=14\cdot 2^{-4}=7/8$ 이 될 것이고, $q_4(x)$가 $g(x)$의 인수인 경우는 식 (4)와 (7), 그리고 MacWilliams identity에 의해, 해당 열의 1의 개수의 평균 N_1이

$$N_1 = 14(1-\tau)^{15}\sum_{w=0}^{15} A_w\left(\frac{\tau}{1-\tau}\right)^w = \frac{7}{8}\{1+15(1-2\tau)^8\}$$

이 되어, 예에서와 같이 $\tau\approx 0.05$ 인 경우는 $N_1\approx 6.52$가 된다.

이상에서 본 바와 같이 각 열의 1의 개수는 해당 기약다항식이 $g(x)$의 인수인가 아닌가에 따라 상당한 편차를 보이게 되므로 적당한 임계값(threshold)을 설정하여 판정할 수 있다. 일반적으로 임계값 설정에서의 두 가지 파라메타는 오경보 확률(해당 기약다항식이 $g(x)$의 인수가 아님에도 불구하고 1의 개수가 임계값을 넘을 확률)과 검출 확률(해당 기약다항식이 $g(x)$의 인수이고 1의 개수가 임계값을 넘을 확률)이다. 검출 확률로 임계값을 정하기 위해서는 모든 기약다항식 $q(x)$에 대해 식 (4)에서 사용된 $p(q(x),w)$를 전부 알아야 하기 때문에 본 논문에서는 오경보 확률을 기준으로 임계값을 설정하였다. 기약 다항식의 차수가 j인 서브행렬의 각 열에 대한 오경보 확률(P_{FA})은 다음과 같이 쓸 수 있다.

$$P_{FA}=\sum_{i=T}^{N}\binom{N}{i}2^{-ji}(1-2^{-j})^{N-i} \qquad (9)$$

일반적으로 기약 다항식이 차수가 클수록 임계값은 작아진다. 식 (9)에서 N은 사용된 부호어의 개수, T는 임계값을 의미한다. P_{FA}을 얼마로 할 것인지에 따라 T 값을 정하게 된다. 위 방법은 알고리즘이 간단하고 사용이 용이하나 부호의 오류정정능력이 증가할수록 (즉, 생성다항식의 인수로 사용되는 기약다항식의 개수가 증가할수록) 적용이 힘들어진다는 단점이 있다.

3.2 Fourier 변환을 이용한 수평 접근법[5]

수직 접근법이 생성다항식의 인수로 사용된 기약다항식을 찾는 방법이라면 수평접근법은 한마디로 오류가 없는 수신워드를 찾는 방법이라고 말할 수 있다. 순회부호를 사용하는 경우, n과 k가 결정되면 (즉, 생성다항식의 차수가 정해지면), 일반적으로 생성다항식이 가능한 많은 개수의 연속근을 갖도록 채널부호를 설계하게 된다. 유한체상에서 정의된 Fourier 변환을 순회부호에 적용하면 수신부호로부터 연속근의 개수를 알 수 있다. 부호어 $c(x)$와 퓨리에 변환의 결과 $\lambda(z)$와의 관계는 다음과 같다.

$$c(x)=(c_0,c_1,\ldots,c_{n-1})\leftrightarrow\lambda(z)=(\lambda_0,\lambda_1,\ldots,\lambda_{n-1})$$
$$\lambda_k=\sum_{i=0}^{n-1}c_i\alpha^{ki}=c(x)|_{x=\alpha^k}$$

여기서 α는 유한체 상에서의 1의 원시 n승근이다. 임의의 수신워드 $r(x)$에 대하여, 연속근의 개수가 크면 클수록 $r(x)$가 오류가 없는 부호어일 확률이 크게 된다. 따라서 수평접근법에서는 서브행렬의 각각의 행에 대하여 1이 위치한 지점에서의 기약 다항식들의 근 중 가장 연속근의 개수가 많은 행을 일단 오류가 없을 확률이 가장 큰 수신워드라고 정한다. 이런 방식으로 오류가 없는 부호어로 생각되는 수신워드들을 다수개 찾아줌으로써 생성다항식을 복원할 수 있다. 이 방법에서는 임계값을 사용할 필요가 없다. 위 방법은 부호어의 길이가 길고 오류정정능력이 작은 경우에는 낮은 신뢰도를 보인다는 단점이 있으나 오류정정능력이 증가하면 증가할수록 높은 신뢰도를 보인다는 장점이 있다.

3.3 제안하는 알고리즘

본 논문에서는 수직 접근법과 수평 접근법을 순차적으로 적용하는 기법을 제안한다. 먼저 수직 접근법은 낮은 SNR에서는 사용된 기약다항식과 사용되지 않은 기약다항식 사이의 차이가 크지 않기 때문에 사용이 어렵다는 단점이 있고, 수평 접근법은 높은 SNR에서는 선택된 모든 워드들이 $g(x)$의 인수로 사용되지 않은 기약다항식을(특히 낮은 차수의) 포함하고 있을 확률이 크게 된다는 단점이 있다. 따라서 낮은 SNR일 때는 수평 접근법이 수직 접근법에 비하여 우월하고, 높은 SNR에서는 수직 접근법이 수평 접근법에 비하여 좋은 성능을 보인다. 따라서 서브행렬로부터 수직 접근법의 사용 가능 여부에 대한 판단을 한 이후에 수평 접근법을 사용함으로써, 위 단점들을 보완할 수 있다.

-Algorithm

1) 기약다항식을 이용하여 N 개의 수신워드에 대한 서브행렬을 생성한다.

2) 각 기약다항식의 차수에 따른 P_{FA}을 이용하여 임계값을 $P_{FA}\times N\times(1+R)$ 으로 하고, 임계값보다 큰 값을 갖는 열의 기약다항식이 후보 기약다항식이 된다. 임계값보다 큰 기약다항식이 존재하지 않는 경우는 가능한 모든 기약다항식이 후보 기약다항식이 된다.

3) 퓨리에 변환을 이용한 수평접근법으로 오류가 없다고 생각되는 M개의 수신워드를 선택한다.

4) 2)에서 구한 후보 기약다항식으로 3)에서 구한 M개의 수신워드를 나누어보고, 만약 M개 중 한번이라도 특정 기약다항식으로 나누어지지 않으면, 그 기약다항식은 사용되지 않은 것으로 생각한다.

5) 4)에서 최종적으로 남은 기약다항식을 이용하여 생성다항식 $g(x)$을 복원한다.

위 알고리즘에서 R은 0보다 큰 실수로서 임계값을 정하는 적정의 값을 의미한다. R을 일반식으로 나타내기는 어렵고, 부호어의 길이와 수신된 부호어의 개수에 따라 좋은 성능을 보이는 R의 값이 다르기 때문에 실제 모의실험에서는 많은 시도 후 경험적으로 적당하다고 생각되는 값으로 정하였다. 그리고 M은 선택하는 수신워드의 개수로서 N 값과 부호어의 길이에 따라 적당한 값을 선택한다.

3.4 모의실험 결과

위 방법을 (63,33) BCH부호에 적용시킨 결과이다. 아래의 모의실험에는 매번 1000개와 5000개의 수신다항식을 사용하였고, 총 100번의 서로 다른 모의실험을 하였다. 아래의 결과는 BSC의 오류확률에 따른 성공횟수를 기록한 것으로 $R=0.1$, $M=4$로 하였다.

II장에서 언급한 상보(dual)부호를 이용하는 일반적인 복원 기법[1]을 이용한 경우에는 1000개의 수신워드를 사용하였을 때 raw BER이 0.03 이하가 되어야 대부분 복원 가능한데 반하여, 표 2에서 보듯이, 제안한 알고리즘을 사용한 경우에는 1000개의 수신워드를 사용한 경우 raw BER이 0.07 이하에서, 5000개의 수신워드를 사용한 경우 raw BER이 0.1 이하에서 대부분 복원 가능함을 확인할 수 있다. 또한 상보부호를 이용한 기법은 N 값의 증가에 따라 큰 이점을 가지지 않는데 반하여, 본 논문에서 제안하는 방법은 N 값을 증가할수록 더 좋은 성능을 보인다는 장점이 있다. 물론 N 값이 증가하고 부호어의 길이가 증가함에 따라, 복원을 위한 연산량 역시 증가하므로 보다 낮은 복잡도를 가지는 알고리즘의 개발이 요구된다.

표 2. Simulation result in (63,33) BCH code

	N= 1000	N= 5000
raw BER	복원횟수	복원횟수
0.13	0	1
0.10	4	92
0.07	97	100
0.04	100	100

Ⅳ. Reed-Solomon 부호의 복원 기법

Reed-Solomon (RS) 부호는 비이진 순회부호로서 주어진 길이 n과 dimension k에 대해 가장 큰 부호의 최소거리, 즉 $d_{\min}=n-k+1$,를 갖는 MDS(Maximum Distance Separable) 부호의 일종이다. 부호의 길이 $n=2^m-1$ 인 RS 부호는 $GF(2^m)$ 상의 심볼로 이루어져 있으며, α를 $GF(2^m)$의 원시원소라고 할 때 생성다항식 $g(x)$는 다음과 같다.

$$g(x)=\prod_{i=s}^{s+n-k-1}(x+\alpha^i)$$

다시 말해 RS 부호의 생성다항식은 α^s 부터 $(n-k)$개의 연속근을 갖는다. 따라서 RS 부호의 복원 파라메타는 연속근의 시작점인 s와 부호의 dimension k가 된다. RS 부호 역시 순회부호이므로 III장에서 제안한 순회부호의 복원기법을 적용할 수 있지만 기약다항식이 모두 일차식이므로 그 효율이 떨어질 수밖에 없다. 반면 다음의 정리에서 볼 수 있듯이 RS 부호의 상보 부호 역시 RS 부호라는 점을 이용한다면 II장의 복원기법을 적용하는 것이 유리하다.

정리 5

α를 $GF(2^m)$의 원시원소라고 하자. 다항식 $g(x)=\prod_{i=s}^{s+n-k-1}(x+\alpha^i)$를 생성다항식으로 갖는 RS 부호 C의 상보부호 $C^{\perp}$ 역시 RS 부호이고 $C^{\perp}$의 생성다항식 $g^{\perp}(x)$는 다음과 같다.

$$g^{\perp}(x)=x^k h\left(\frac{1}{x}\right) \tag{10}$$

여기서 $h(x)=\frac{x^n+1}{g(x)}$ 이다.

증명: 순회부호의 생성다항식은 가장 낮은 차수의

부호어 다항식이다. $C^{\perp}$는 $[n, n-k]$ 부호이고 $g^{\perp}(x)$의 차수가 k이므로 $g^{\perp}(x)$가 $C^{\perp}$의 부호어 다항식임을 보이면 된다.

정의에 의해 $g(x)h(x)=0 \pmod{x^n+1}$ 이므로, 모든 l에 대해

$$\sum_i g_i h_{l-i}=0 \tag{11}$$

이 성립한다. 식 (10)에 의하면 $g^{\perp}(x)$는 $h(x)$의 역(reciprocal) 다항식이다. 다시 말해, $g^{\perp}(x)$에서 x^i의 계수는 $h(x)$의 x^{k-i}의 계수와 같게 된다. 즉, $g^{\perp}(x)=\sum_i g_i^{\perp} x^i$ 라고 할 때 $g_i^{\perp}=h_{k-i}$, $0 \le i \le k$ 이 성립한다. 고로, (11)에 의해

$$\sum_{i=0}^{n} g_i g_i^{\perp} = \sum_{i=0}^{n} g_i h_{k-i} = 0$$

이 되어 $g^{\perp}(x)$는 $C^{\perp}$의 부호어 다항식이다. 또한 $h(x)$ 역시 k개의 연속근을 가지므로 $C^{\perp}$ 역시 RS 부호가 된다. □

정리 5에서 본 바와 같이 RS 부호의 상보 부호 역시 RS 부호이므로 II장의 상보 부호를 이용한 복원기법에서 가장 어려운 부분으로 간주되었던 상보부호의 후보 부호어 군 생성이 비교적 용이하게 된다. 즉, 일정 개수의 연속근을 갖는 워드를 생성하여 상보부호의 후보 부호어로 사용하는 것이다. 구체적인 알고리즘은 다음과 같다.

-Algorithm

1) 기약다항식을 이용하여 N 개의 수신워드에 대한 서브행렬을 생성한다. 그리고 $g(x)$의 최소 연속근의 개수($r_{\min}$)을 가정한다.

2) $n-r_{\min}$ 개의 연속근만을 근으로 갖는 다항식 $H_i(x)$을 연속근의 시작점을 다르게 하여 모두 n개 생성한다. $H_i(x)$는 아래와 같이 정의된다.

$$H_i(x)=\prod_{k=i}^{i+n-r_{\min}-1}(x+\alpha^k) \quad , 0 \le i \le n-1$$

3) 각각의 $H_i(x)$에 대해, 서브행렬의 N 개의 모든 수신워드와 내적하여 그 결과가 0이 되는 수신워드의 개수를 센다. 그 개수가 임계값 T보다 큰 경우의 i값을 모두 찾는다. 임계값 T는 내적결과가 0인 수신워드의 최대 개수에 1보다 작은 적절한 양의 실수 R을 곱한 값으로 정한다.

4) 적정한 임계값 하에서는, 찾은 i값은 전부 연속된 숫자일 것이다. 연속된 숫자의 개수가 K이라면, 연속된 숫자의 마지막 값이 상보 부호의 연속근의 시작점이 되고, 해당 시작점으로부터 $(n-r_{\min}-K+1)$ 만큼의 연속근을 상보부호가 갖게 된다. 식 (10)를 이용하여 이로부터 생성다항식 $g(x)$을 복원 가능하다.

예를 들어, $n=15$인 RS 부호의 경우, $r_{\min}=4$로 하여 알고리즘을 적용했을 때 선택된 i가 13, 14, 0, 1, 2, 3인 경우에는 i의 끝점이 3이고, $K=6$이므로 상

표 3. Simulation result in (63,53) RS code

	$N=1000$	$N=5000$
raw BER	복원횟수	복원횟수
0.005	100	100
0.006	100	100
0.007	100	100
0.008	100	100
0.009	100	100
0.010	98	100
	$N=1000$	$N=5000$
raw BER	복원횟수	복원횟수
0.011	75	100
0.012	38	92
0.013	14	24
0.014	1	1
0.015	0	0
0.016	0	0

표 4. Simulation result in (127,107) RS code

	$N=1000$	$N=5000$
raw BER	복원횟수	복원횟수
0.001	100	100
0.002	100	100
0.003	100	100
0.004	99	100
0.005	31	100
0.006	2	100
	$N=1000$	$N=5000$
raw BER	복원횟수	복원횟수
0.007	0	100
0.008	0	93
0.009	0	25
0.010	0	1
0.011	0	0
0.012	0	0

표 5. Simulation result in (255,235) RS code

	N= 1000	N= 5000
raw BER	복원횟수	복원횟수
0.001	100	100
0.002	93	100
0.003	1	95
0.004	0	31
0.005	0	0
0.006	0	0

보부호는 $\alpha^3 \sim \alpha^8$까지의 연속근을 갖게 된다.

위 방법을 $g(x)=\prod_{i=1}^{10}(x+\alpha^i)$ 인 (63, 53) RS 부호, $g(x)=\prod_{i=1}^{10}(x+\beta^i)$ 인 (127, 107) RS 부호, 그리고 $g(x)=\prod_{i=1}^{10}(x+\gamma^i)$ 인 (255, 235) RS 부호에 적용시킨 결과이다. 이 때, α, β, γ는 각각 $GF(2^6), GF(2^7), GF(2^8)$의 원소를 지칭한다. 모의실험에는 매번 1000개와 5000개의 수신벡터가 사용되었고, 총 100번의 모의실험을 하였다. 아래의 결과는 raw BER에 따른 성공 횟수를 기록한 것으로 (63, 53) RS 부호의 경우, $R=0.8$, $r_{\min}=8$로, 나머지 경우에는 $R=0.8$, $r_{\min}=14$로 하였다.

V. 결 론

본 논문에서는 BCH 부호와 같은 이진 순회부호의 새로운 복원기법을 제안하고 더불어 비이진 순회부호인 RS 부호의 복원기법도 제안하였고 모의실험을 통해 기존에 알려진 복원기법보다 훨씬 좋은 성능을 보임을 보였다. 본 연구에서는 부호의 동기단계가 완료된 것을 가정하였으나 제안된 기법이 실제적으로 적용되기 위해서는 동기 기법에 대한 연구가 선행되어야 한다. 또한 채널부호의 복원이라는 주제는 민간에서보다는 주로 군사적으로 연구되는 분야여서 발표된 연구결과가 미약한 실정이다. 그 중요성을 감안한다면 다양한 채널부호에 대한 효율적인 복원기법 연구가 앞으로 절실하다고 하겠다.

참 고 문 헌

〔1〕 A. Valembois. "Detection and recognition of a binary linear code," Discrete Applied Mathematics, 111(1-2): pp.199-218, July 2001.

〔2〕 M. Cluzeau and M. Finiasz, "Recovering a code's length and synchronization from a noisy intercepted bit stream," Proceedings of International Symposium on Information Theory (ISIT '09), pp. 2737-2741, Seoul, Korea, June 28 - July 3, 2009.

〔3〕 J. Barbier, G. Sicot, and S. Houcke, "Algebraic approach for the reconstruction of linear and convolutional error correcting codes," International Journal of Applied Mathematics and Computer Sciences 2, pp.113-118, summer 2006.

〔4〕 M. Cluzeau, "Block code reconstruction using iterative decoding techniques," Proceedings of International Symposium on Information Theory (ISIT '06), pp.2269-2273, 2006.

〔5〕 정하봉, 장환석, 조원찬, 김창구, "미지의 채널부호에 대한 파라메타 추출 및 복원" *제1회 신소재 • 에너지 무기 학술대회 논문집*, pp.316-319, 2009년 9월.

〔6〕 F.J. MacWilliams and N. J. A. Slone, *The theory of error correcting codes*, North-Holland Publishing Company, 1977.

정 하 봉 (Habong Chung) 종신회원

1981년 2월 서울대학교 전자공학과 공학사

1985년 2월 미국 University of Southern California, 전기공학과 공학석사

1988년 미국 University of Southern California, 전기공학과 공학박사

1988~1991년 미국 뉴욕주립대 전기공학과 조교수

1991년~현재 홍익대학교 전자전기공학부 교수

<관심분야> 부호 이론, 조합수학, 시퀀스 설계, 협력통신, 시공간 부호

장 환 석 (Hwanseok Jang) 정회원

2008년 2월 홍익대학교 전자전기공학부 졸업
2010년 2월 홍익대학교 전자정보통신공학과 석사
2010년 3월~현재 홍익대학교 전자통신공학과 박사과정
<관심분야> 부호 이론, 채널 코딩, 소스 코딩

조 원 찬 (Cho-won Chan) 준회원

2009년 2월 홍익대학교 전자전기공학부 졸업
2011년 2월 홍익대학교 전자통신공학과 석사
2011년 3월~현재 LG 전자 연구원
<관심분야> 부호 이론, 채널 코딩, 소스 코딩

박 철 순 (Cheol-sun Park) 정회원

1989년 2월 경기대학교 전자계산학과 졸업
1991년 2월 인하대학교 전자계산공학과 석사
1991년~현재 국방과학연구소 선임연구원
1997년 5월 전자계산조직응용기술사
2007년 2월 충남대학교 정보통신공학과 박사
<관심분야> 신호처리, 통신응용