

정수계획법 모형을 이용한 홀로그래픽 저장장치의 DC-억압 GS코딩의 성능평가

박태형*, 이재진^o

Performance Evaluation of DC-Suppression GS Coding for the Holographic Data Storage Using Integer Programming Models

Taehyung Park*, Jaejin Lee^o

요약

광저장장치의 DC-억압을 위한 멀티모드 코딩 기법 중 Guided Scrambling (GS) 코딩기법이 널리 사용된다. 홀로그래픽 저장장치를 위한 DC-억압 GS코딩에서는 후보코드 선택기준으로 심볼의 균등한 분포 및 심볼간 천이의 최대화 기준이 고려되었다. 본 연구에서는 후보코드행렬의 digital sum value (DSV)의 l_{∞} -norm을 최소화하는 minimum DSV (MDSV) 기준 GS코딩을 정수계획법 모형으로 수식화하고, 제안된 모형을 사용하여 MDSV 기준과 최대천이강도 기준이 적용된 GS코딩의 성능을 제어비트수, 행렬크기 및 스크램블링 다항식들의 조합에 대하여 평가한다.

Key Words : Holographic data storage, Guided Scrambling Encoding, MDSV, Conservative Array, Integer Programming

ABSTRACT

For the DC-free encoding of source data, the Guided Scrambling (GS) technique is widely used as multi-mode coding in the optical data storage system. For DC-suppression GS coding in the holographic data storage system, the conservative array and balanced coding criteria are proposed. In this paper, equivalent integer programming models are developed to determine the optimal control bits for the minimum digital sum value (MDSV), conservative array, and balanced coding criteria. Using the proposed integer programming models, we compare the performance of GS encoding for the various combination of control bit/array sizes and scrambling polynomials.

I. 서론

홀로그래픽 데이터 저장장치(holographic data storage, HDS)는 정보를 매체의 볼륨에 저장하므로 높은 저장 용량(1 Tb/cm^3)을 가지며, 페이지 단위로 데이터를 읽기 때문에 데이터 전송율(>1 Gbps)이 높은 차세대 광학

저장장치이다^[1]. HDS의 장점에도 불구하고, 데이터를 볼륨에 저장하므로 기존의 광학저장장치에서 볼 수 없는 IPI(interpage interference), 2D ISI(two-dimensional intersymbol interference)와 같은 새로운 문제가 발생한다. 홀로그래픽 저장장치에서 변조부호의 역할은 페이지 내에서 광학빔의 불균등 분포 및 인접 심볼간 간섭

* 주저자 겸 교신저자 : 숭실대학교 산업정보시스템공학과, tpark@ssu.ac.kr, 정회원
^o 숭실대학교 정보통신전자공학부 정보저장및통신 연구실, zlee@ssu.ac.kr, 중신회원
 논문번호 : KICS2013-07-300, 접수일자 : 2013년 7월 23일, 최종논문접수일자 : 2013년 8월 26일

을 줄이도록 설계된다. 변조부호가 갖는 제약조건은 예는 ON픽셀과 OFF픽셀의 분포가 균등해야 한다거나, 저역통과 필터링(low-pass filtering)효과를 고려하여 ON픽셀 주변을 OFF픽셀이 둘러싸는 고립된 픽셀이 발생하지 않도록 하는 조건, 런-길이제한(RLL) 코드처럼 인접 심볼 간에 최소, 최대 OFF픽셀의 개수를 제한하는 조건이 있다^{2,4)}. 홀로그래픽 저장장치에 대한 ON픽셀과 OFF픽셀의 최소천이(transition) 회수에 대한 제약식으로 보수행렬 (conservative array)조건이 있다. 보수행렬 조건은 각 행과 열에서 0→1, 1→0의 천이가 최소 t 회수 발생하는 경우, 그 행렬은 천이강도(strength) t 를 갖는다고 정의한다⁵⁾.

이러한 변조부호 특성을 만족하는 인코딩기법으로 멀티모드 코딩에서는 소스데이터 x 에 L 개의 후보코드워드들의 집합 C_x 로 1-to- L 함수를 정의하고, L 코드워드 중, 조건을 만족하는 최적의 코드워드를 인코딩하는 기법이다. 멀티모드 코딩의 한 종류인 Guided Scrambling (GS) 코딩에서는 소스데이터에 p 비트 제어비트를 추가한 후, linear feedback shift register(LFSR)로 스크램블링한 $L=2^p$ 후보코드워드를 생성한다. 홀로그래픽 저장장치의 GS코딩에서 후보코드워드 중 최적의 코드워드를 선택하는 판단기준으로는 ON/OFF픽셀의 균등분포, 천이강도 조건 및 페이지의 평균 빔세기가 고려되었다^{6,7)}.

본 연구에서는 홀로그래픽 저장장치의 DC-억제 시퀀스를 생성하는 GS코딩절차를 정수계획법 모형으로 수식화하고, MDSV 및 천이강도 선택기준을 만족하는 balanced 시퀀스의 평균 DC-억제 성능을 평가한다. 또한 두 선택기준을 동시에 고려하는 모형의 MDSV 및 천이강도 성능을 평가한다. 마지막으로 스크램블링 다항식과 제어비트의 선택이 MDSV 성능에 끼치는 영향을 평가한다.

본 논문의 구성은 다음과 같다. 2장에서는 MDSV 최소화, 천이강도 최대화 GS코딩의 정수계획법 모형을 소개하고, 천이강도와 MDSV값과의 관계를 분석한다. 3장에서는 제어비트 및 행렬크기에 따른 제안한 모형들의 평균 DC-억제 성능을 평가한다. 4장에서는 본 연구의 결과를 요약한다.

II. MDSV GS코딩의 정수계획법 모형

홀로그래픽 저장장치를 위한 GS코딩에서는 n 소스비트 b_1, \dots, b_n 에 p 제어비트 d_1, \dots, d_p 를 붙여서 확장한 $n+p$ 소스시퀀스 $d = (d_1, \dots, d_{n+p})$ 를 스

크램블링한다. 여기에서 $d_{p+1} = b_1, \dots, d_{n+p} = b_n$. GS코딩에서 스크램블된 코드시퀀스는 self-synchronizing 다항식 $c_k = d_k \oplus \sum_{i=1}^K a_i c_{k-i}$ 을 이용하여 생성한 후, 시퀀스 $c = (c_1, \dots, c_{n+p})$ 를 $m \times m$ 행렬 $(c_{i,j})$ 로 배열한다.

본 연구에서는 11단계 다항식인 $c_k = d_k \oplus c_{k-2} \oplus c_{k-11}$ 를 기본적으로 사용하였다. GS코딩을 정수계획법 모형으로 수식화할 때는 modulo-2 덧셈을 일반적인 덧셈으로 표현해야 한다. 예를 들어 $s_k = d_k \oplus c_{k-2}$ 는 다음과 같은 4개의 부등식으로 표현할 수 있다.

$$\begin{aligned} s_k &\leq b_k + c_{k-2}, s_k \leq 2 - b_k - c_{k-2} \\ s_k &\geq b_k - c_{k-2}, s_k \geq c_{k-2} - b_k \end{aligned} \quad (1)$$

아래에서 (1)의 4개의 부등식을 $s_k = d_k \diamond c_{k-2}$ 로 표시한다. $c_k = d_k \oplus c_{k-2} \oplus c_{k-11}$ 의 경우는 결합법칙을 사용하여 $s_k = d_k \diamond c_{k-2}, c_k = s_k \diamond c_{k-11}$ 로 계산할 수 있다. 따라서 스크램블링된 코드는 다음 식을 만족한다.

$$\begin{aligned} s_k &= d_k \diamond c_{k-2}, k = 1, \dots, m^2 \\ c_k &= s_k \diamond c_{k-11}, k = 1, \dots, m^2 \end{aligned} \quad (2)$$

여기에서 초기값 $c_0 = 0, \dots, c_{-10} = 0$ 으로 가정하였다.

인코딩된 시퀀스의 DSV(digital sum value)가 유한한 값을 갖는 경우, 시퀀스의 RDS(running digital sum) 값이 상한을 갖고 DC-억제 특성을 갖는 시퀀스를 생성한다. 또한 WRDS(word-end RDS)값이 유한한 경우, RDS는 유한한 값을 갖는다⁸⁾. $m \times m$ 코드행렬 $c = (c_{i,j})$ 에서, 행 i 의 j 열까지의 DSV $s_{i,j}$ 와 열 j 의 행 i 까지의 DSV $t_{i,j}$ 를 다음과 같이 정의한다.

$$\begin{aligned} s_{i,j} &= s_{i,j-1} + 2c_{i,j} - 1, j = 1, \dots, m \\ t_{i,j} &= t_{i-1,j} + 2c_{i,j} - 1, i = 1, \dots, m \end{aligned} \quad (3)$$

여기에서 $s_{i,0} = 0, \forall i, t_{0,j} = 0, \forall j$ 로 정한

다. 후보코드워드를 선택하는 MDSV(minimum DSV) 기준은 DSV $s_{i,j}, t_{i,j}$ 의 l_∞ -norm을 최소화하는 기준으로 $\gamma = \max_{i,j} \{|s_{i,j}|, |t_{i,j}|\}$ 가 최소가 되는 후보코드를 선택한다. MSW(minimum squared weight)기준은 $s_{i,j}, t_{i,j}$ 의 l_2 -norm을 최소화하는 기준이다.

식 (3)에 의해, γ 는 다음 식을 만족한다.

$$\begin{aligned} \gamma &\geq s_{i,j}, \gamma \geq -s_{i,j}, \forall i,j \\ \gamma &\geq t_{i,j}, \gamma \geq -t_{i,j}, \forall i,j \end{aligned} \quad (4)$$

보수행렬의 천이강도 t 는 각 행과 열에서 0→1, 1→0천이의 최소회수를 나타낸다. $m \times m$ 행렬의 i 번째 행 $c_i = (c_{i,1}, \dots, c_{i,m})$ 의 천이회수는 $m-1$ modulo-2 덧셈인 $c_{i,1} \oplus c_{i,2} \oplus \dots \oplus c_{i,m}$ 에 해당한다. 예를 들어, $m=10$, $c_i = (0110110110)$ 인 경우, $s_{i,10} = 2$, 천이의 회수는 6이 된다. γ 가 행 k 에서 발생하고, tr_k 가 행 k 의 천이회수이면, $\gamma \leq m - tr_k$ 이고, 천이강도 t 는 모든 행과 열의 최솟값이므로, γ 와 천이강도 t 는 다음 부등식을 만족한다.

$$\gamma \leq m - t \quad (5)$$

따라서 천이강도가 높은 경우, γ 는 작은 값을 갖게 된다.

각 행렬의 최소천이강도를 계산하기 위해 두 심볼간의 천이를 다음과 같은 modulo-2 연산으로 계산할 수 있다.

$$\begin{aligned} u_{i,j} &= c_{i,j} \diamond c_{i,j+1}, j = 1, \dots, m-1 \\ w_{i,j} &= c_{i,j} \diamond c_{i+1,j}, i = 1, \dots, m-1 \end{aligned} \quad (6)$$

따라서 행렬의 천이강도 t 는 다음 식을 만족한다.

$$t \leq \sum_{i=1}^{m-1} u_{i,j}, \forall i, t \leq \sum_{j=1}^{m-1} w_{i,j}, \forall j \quad (7)$$

코드비트행렬에서 '1'과 '0' 픽셀이 1:1 비율을 갖는 조건과 소스비트와 제어비트에 대한 제약식은 다음과 같다.

$$\begin{aligned} \sum_{i,j} c_{i,j} &= m^2/2 \\ d_{p+1} &= b_1, \dots, d_{n+p} = b_n \\ d_j &\in \{0, 1\}, j = 1, \dots, p \end{aligned} \quad (8)$$

식 (2)-(8)을 이용하여 MDSV기준을 사용하는 홀로그래픽 저장장치의 GS코딩은 다음과 같은 정수계획법 모형으로 수식화된다.

$$z_{MDSV} = Min \{ \gamma | (2), (3), (4), (8) \} \quad (9)$$

모형 (9)는 스크램블된 2^p 개의 후보코드행렬 중, '0', '1' 심볼이 1:1로 균등하고, DSV의 l_∞ -norm이 최소가 되는 후보를 생성하는 최적 제어비트를 결정한다. 마찬가지로 천이강도 최대화 문제는 다음과 같다.

$$z_{strength} = Max \{ t | (2), (6), (7), (8) \} \quad (10)$$

두 문제를 혼합하여 천이강도를 극대화하면서 MDSV는 최소화하는 문제는 다음과 같다.

$$z_p = Min \{ \alpha_1 \gamma - \alpha_2 t | (2-4), (6-8) \} \quad (11)$$

여기에서 상수 α_1, α_2 는 사용자 지정값이다.

III. 계산 결과

모형 (9), (10), (11)은 정수계획법 모형의 algebraic modeling language인 AMPL과 ILOG Cplex 12.4로 구현되었다^{9,10}. 실험을 위해 행렬의 크기는 10x10, 12x12, 14x14, 16x16, 18x18, 20x20, 제어비트는 6에서 10사이의 값으로 정하고, 각 제어비트, 행렬크기 별로 20개의 임의의 소스시퀀스 $b = (b_1, \dots, b_n)$ 를 생성하였다. 표 1은 제어비트, 행렬크기별 20개 문제에 대한 모형 (9)의 z_{MDSV} 의 평균 (' γ '로 표시)과 (10)의 $z_{strength}$ 의 평균('t'로 표시)을 기록하였다. 표 1에서 천이강도의 평균은 행렬과 제어비트가 커짐에 따라 증가함을 알 수 있다. MDSV값의 평균은 행렬크기와 상관없이 대체로 일정한 값을 가진다. 하지만 MDSV는 제어비트가 많아짐에 따라 모든 행렬 크기에서 일정한 비율로 감소하고 있음을 알 수 있다.

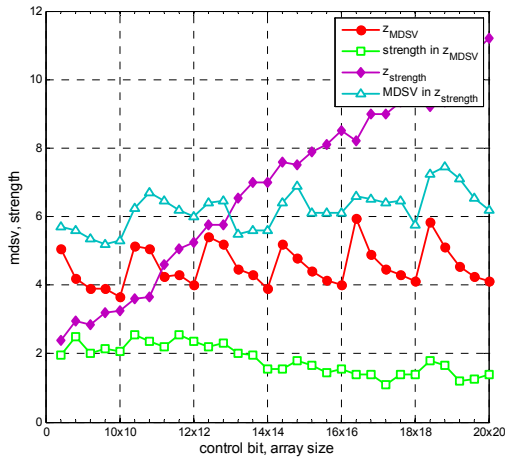


그림 1. 제어비트/행렬크기별 MDSV와 천이강도
Fig. 1. MDSV and strength result for various control bit, array sizes.

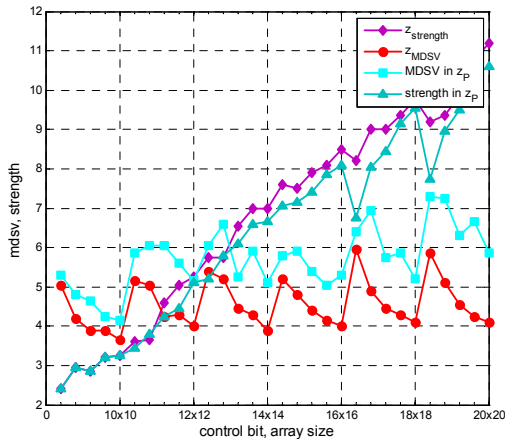


그림 2. z_p 의 MDSV 및 천이강도와 z_{MDSV} , $z_{strength}$ 의 비교.
Fig. 2. Comparison of MDSV and strength from z_p with z_{MDSV} , $z_{strength}$.

그림 1에는 표 1의 MDSV의 평균 (z_{MDSV}), 모형 (9)의 최적제어비트에서 선택된 행렬의 천이강도의 평균('strength in z_{MDSV} '), (10)의 천이강도의 평균($z_{strength}$) 및 모형 (10)에서 선택된 행렬의 MDSV 평균 ('MDSV in $z_{strength}$ ')을 보여준다. 그림 1의 x-축은 각 행렬크기별로 제어비트개수 6에서 10까지 5가지 값이 기록되었다. 그림 1에서 MDSV의 평균은 제어비트 크기에 따라 일정하게 감소함을 알 수 있다. MDSV 기준으로 선택한 코드시퀀스의 천이강도값은 모형(10)에서 계산한 값에 비해 현저하게 낮은 값을 보여주고 있다. 이는 천이강도를 높게 하

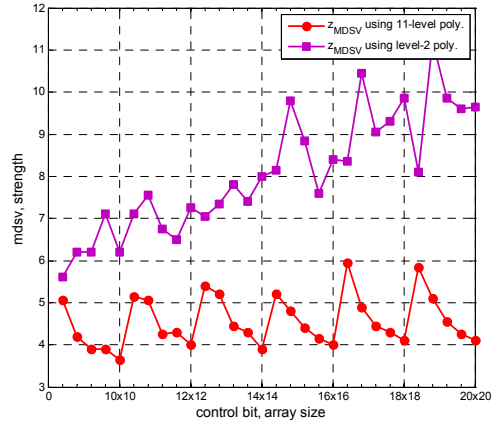


그림 3. 다항식과 제어비트에 따른 MDSV 비교.
Fig. 3. MDSV performance with respect to scrambling polynomial and control bit sizes.

지 않고도 MDSV값을 낮게 유지할 수 있음을 보여 준다. 모형 (10)을 이용하여 천이강도를 최대화하였을 경우, 최적해에서의 MDSV값은 z_{MDSV} 값 보다 평균 1.64 높은 값을 가지지만, 모든 행렬크기에서 우수한 성능을 보였다.

그림 2에는 (11)의 $\alpha_1\gamma - \alpha_2t$ 를 목적함수로 사용한 z_p 에서 계산한 MDSV 및 천이강도와 모형 (9), (10)을 비교한다. (11)에서 계산한 MDSV는 (9)의 MDSV 평균에 비해 25.4% 높은 값을 기록하였고, 천이강도의 평균은 (10)보다 6.2% 낮은 값을 기록하였다. 모형 (11)에서 $\alpha_1 = 1, \alpha_2 = 10$ 을 사용하였다. 그림 3에는 스크램블링 다항식으로 11단계 다항식 $c_k = d_k \oplus c_{k-2} \oplus c_{k-11}$ 을 사용한 경우 계산한 MDSV 값의 평균 (' z_{MDSV} using 11-level poly. ')과 2 단계 다항식 $c_k = d_k \oplus c_{k-2}$ 를 사용하여 계산한 MDSV값의 평균 (' z_{MDSV} using 2-level poly. ')을 보여준다. 모든 제어비트, 행렬크기 조합에서 2단계 다항식을 사용한 경우의 MDSV 평균이 높은 값을 가졌다. 식 (5)에 의해, 2단계 다항식을 사용하는 경우, 11단계 다항식을 사용하는 경우보다 천이강도는 낮은 값을 갖는다. 따라서 GS코딩을 적용하는 경우, 스크램블링 다항식과 제어비트의 선택이 DC-억압 GS코딩에서 중요한 요소임을 알 수 있다.

IV. 결론

본 연구에서는 홀로그래픽 저장장치의 DC-억압을

표 1 제어비트와 행렬크기에 따른 MDSV, 천이강도.
Table 1. z_{MDSV} and $z_{strength}$ for various control bit/array sizes.

array size	# control bits									
	6		7		8		9		10	
	γ	t	γ	t	γ	t	γ	t	γ	t
10x10	5.05	2.4	4.2	2.95	3.9	2.85	3.9	3.2	3.65	3.25
12x12	5.15	3.6	5.05	3.65	4.25	4.6	4.3	5.05	4	5.25
14x14	5.4	5.75	5.2	5.75	4.45	6.55	4.3	7	3.9	7
16x16	5.2	7.6	4.8	7.5	4.4	7.9	4.15	8.1	4	8.5
18x18	5.95	8.2	4.9	9	4.45	9	4.3	9.35	4.1	9.75
20x20	5.85	9.2	5.1	9.35	4.55	9.85	4.25	10.8	4.1	11.2

위한 코드시퀀스를 찾는 MDSV최소화, 천이강도 최대화 GS코딩을 정수계획법 모형으로 수식화하였다. 개발된 모형을 제어비트수/코드행렬크기 조합별로 20개의 임의의 데이터를 생성하여 MDSV판단기준과 천이강도 판단기준의 GS코딩의 성능을 비교하였다. MDSV는 제어비트의 크기에 직접적으로 영향을 받으며, 천이강도 최대화 기준을 사용할 경우 선택된 후보코드의 MDSV성능이 MDSV최소화 기준으로 선택한 후보코드워드의 MDSV값과 큰 차이가 없음을 알 수 있었다. 또한 두 가지 판단기준을 결합하는 경우, MDSV 및 천이강도 두 가지 모두 만족하는 코드워드를 찾을 수 있었다. GS코딩을 효과적으로 적용하기 위해서는 스크램블링 다항식과 제어비트의 결정이 중요한 요소임을 두 가지 다항식의 MDSV 비교를 통하여 확인하였다.

References

[1] L. Hesselink, S. S. Orlov, and M. C. Bashaw, "Holographic data storage systems," *Proc. IEEE*, vol. 92, no. 8, pp. 1231-1280, Aug. 2004.

[2] J. J. Ashley and B. H. Marcus, "Two-dimensional low-pass filtering codes," *IEEE Trans. Commun.*, vol. 46, no. 6, pp. 724-727, June 1998.

[3] G. W. Burr, J. Ashley, H. Coufal, R. K. Grygier, J. A. Hoffnagle, C. M. Jefferson, and B. Marcus, "Modulation coding for pixel-matched holographic data storage," *Optical Letters*, vol. 22, no. 9, pp. 639-641, May 1997.

[4] K. Yang, J. Kim, and J. Lee, "Mis-alignment channel performance of error correcting 4/6 modulation codes for holographic data storage," *J. KICS*, vol. 35, no. 12, pp. 971-976, Dec. 2010.

[5] A. Vardy, M. Blaum, P. H. Siegel, and G. T. Sincerebox, "Conservative arrays: multidimensional modulation codes for holographic recording," *IEEE Trans. Inf. Theory*, vol. 42, no. 1, pp. 227-230, Jan. 1996.

[6] W. Y. H. Wilson, K. A. S. Immink, X. B. Xi, and C. T. Chong, "An efficient coding technique for holographic storage with the method of guided scrambling," *Proc. SPIE*, vol. 4090, pp. 191-196, Sep. 2000.

[7] N. Y. Kim, J. Lee, Y. Hong, and J. Lee, "Optimal number of control bits in the guided scrambling method for holographic data storage," *Japanese J. Applied Physics*, vol. 44, pp. 3449-3452, May 2005.

[8] Y. Zhu and I. J. Fair, "Modified minimum sum squared weight selection criteria for DC-free multimode codes," *Electron. Lett.*, vol. 41, no. 17, pp. 973-975, Aug. 2005.

[9] R. Fourer, D. M. Gay, and B. W. Kernighan, *AMPL*, Scientific Press, 1993.

[10] IBM, *IBM ILOG CPLEX 12.4 User's Manual*, 2012.

박 태 형 (Taehyung Park)



1986년 2월 고려대학교 산업공
학과 학사

1989년 2월 고려대학교 산업공
학과 석사

1998년 6월 Virginia Tech 산업
시스템공학과 박사

2001년~현재 송실대학교 산

업정보시스템공학과 부교수

<관심분야> 정수계획법, 통신네트워크

이 재 진 (Jaejin Lee)



1983년 2월 연세대학교 전자
공학과 학사

1984년 12월 U. of Michigan,
Dept. of EECS 석사

1994년 12월 Georgia Tech.
Sch. of ECE 박사

1995년 1월~1995년 12월

Georgia Tech. 연구원

1996년 1월~1997년 2월 현대전자 정보통신 연구
소 책임 연구원

1997년 3월~2005년 8월 동국대학교 전자공학과
부교수

2005년 9월~송실대학교 정보통신전자공학부 교수

<관심분야> 통신이론, 채널코딩, 기록저장 시스템