ICIC Express Letters
Part B: Applications
Volume **7**, Number **3**, March **2016**

ICIC International ©2016 ISSN 2185-2766
pp. 577–581

PERFORMANCE OF THE GUIDED SCRAMBLING CONSERVATIVE ARRAY FOR HOLOGRAPHIC DATA STORAGE

Taehyung Park

Department of Industrial and Information Systems Engineering
Soongsil University
No. 369, Sangdo-Ro, Dongjak-gu, Seoul 156-743, Korea
tpark@ssu.ac.kr

Received August 2015; accepted October 2015

Abstract. *Storing data into a two dimensional pixel image in the holographic data storage (HDS) imposes new constraints in modulation codes. Conservative array constraint requires that in each row and column of a code array, a minimum number of 0↔1 transitions must occur. Several multimode coding for the conservative array for HDS have been proposed. In this paper, balanced conservative guided scrambling (GS) multimode coding and minimum running digital sum (MRDS) GS coding are formulated as integer programming models. Also, simple neighborhood search heuristic replacing integer programming model is introduced. In the simulation, the proposed models are compared using randomly generated user data for various combinations of array size and control bits.*

Keywords: Holographic data storage, Guided scrambling, Integer programming

1. **Introduction.** Holographic data storage (HDS) system is regarded as the next-generation optical storage device with its high storage density (> 1 Tb/cm^3) and fast data access rate (> 1 Gbps). HDS records interference pattern of signal and reference beams as holograms in thick photosensitive medium.

High storage density is achieved by superimposing multiple pages within a three-dimensional medium. Also, data being recorded as a pixilated image, encoding and retrieving data can be parallelized easily, resulting in fast data access speed [1,2].

Unlike the previous optical data storage system where coded sequence is one-dimensional, new interference pattern between neighboring pixels and between pages occurs in HDS. The role of the modulation code in HDS is designed to reduce variations in intensity distribution and inter-symbol interference. Examples of modulation constraints are balanced distribution of ON and OFF pixels on a page, forbidden pattern such as OFF pixel surrounded by ON pixels for low-pass filtering effect, the minimum and maximum numbers of OFF pixels between any ON pixels in two-dimensional direction similar to run length limited (RLL) code in optical storage system [3].

Another constraint for low-pass filtering objective requires that in each row and column of the recorded array, there are at least t transitions of 0→1 or 1→0. A binary array of this property is called conservative array of strength t [4-6]. For balanced conservative array with cyclic strength t it was introduced using a cascaded coding scheme employing two modulators. These algorithms require two $(t-2)$-error correcting input codes which are generated through first order Reed-Muller code or Bose-Chaudhuri-Hochquenghem code.

Guided scrambling (GS) belongs to multimode coding and was applied to balanced conservative arrays, too [7-10]. In GS, n bit source word is augmented with all the possible binary sequences of length p and resulting 2^p augmented sequences are scrambled to generate a selection set consisting of 2^p pseudo random sequences. Encoder then selects best candidate sequence from the selection set for transmission. For conservative array, encoder selects the scrambled array with the maximum strength.

Similar to the strength in the conservative array, disparity in the binary sequence is measured using running digital sum (RDS). For low-pass filtering, optical storages adopt encoding based on minimum RDS (MRDS) or minimum variance of the RDS [11,12].

In this paper, we show that GS coding approach for the conservative array and MRDS encoding can be formulated as integer programming models. Using the proposed model, strength and MRDS values of GS coding are evaluated for various array sizes and control bits. Also, we show that the proposed model can be applied to other RDS related selection criteria GS coding.

This paper is organized as follows. In Section 2, we describe integer programming models equivalent to balanced conservative array GS coding and MRDS GS coding. In Section 3, we compared average performance of the proposed models. Section 4 concludes this paper.

2. Integer Programming Model of the Guided Scrambling Conservative Array.

In balanced conservative array GS coding, a source word of length n is preceded by p control bits, and then $n+p$ bits are scrambled through the self-synchronizing polynomial to generate candidate code sequence. Self-synchronizing polynomial has a form $c_k = d_k \oplus \sum_{i\in A} a_i c_{k-i}$, where d_k is the augmented source word, $a_i \in \{0,1\}$ coefficients, and $\oplus$ is the modulo-2 addition. We describe our approach using polynomial $c_k = d_k \oplus c_{k-2} \oplus c_{k-11}$. Scrambled sequence (c_k) is arranged as $m \times m$ array and encoder determines the best candidate array that has the highest strength with balanced '0' and '1' pixels.

In the proposed integer programming model, the scrambling process $c_k = d_k \oplus c_{k-2} \oplus c_{k-11}$, $k = 1,\dots,n+p$ are modeled as a set of linear inequalities that are equivalent to modulo-2 addition. Each modulo-2 addition $y = x_1 \oplus x_2$ is substituted by the four inequalities: $y \le x_1 + x_2$, $y \le 2 - x_1 - x_2$, $y \ge x_1 - x_2$, and $y \ge x_2 - x_1$. Using associative law for modulo-2 addition, any scrambling polynomial can be transformed to a set of linear inequalities. In this paper, the linear inequalities representing $y = x_1 \oplus x_2$ are denoted as $y = x_1 \circ x_2$.

In the following, scrambled code sequence c_k, $k = 1,\dots,n+p$ are arranged as $m \times m$ matrix $(c_{i,j})$. Then, a balanced scrambled code sequence satisfies the following constraints

$$\begin{aligned} &s_i = d_i \circ c_{i-2},\ c_i = s_i \circ c_{i-11},\ i = 1,\dots,n+p,\ \sum_{ij} c_{i,j} = m^2/2,\\ &d_i \in \{0,1\},\ i = 1,\dots,p,\ d_{p+i} = b_i,\ i = 1,\dots,n, \end{aligned} \tag{1}$$

where $d_1,\dots,d_p$ are control bits, $b_1,\dots,b_n$ are source data, s_i are auxiliary variables for associate law, and $\sum_{ij} c_{i,j} = m^2/2$ is the symbol balance constraint. Using modulo-2 addition, the number of $0 \leftrightarrow 1$ transitions in row i is $c_{i,1} \oplus c_{i,2} \oplus \dots \oplus c_{i,m}$ and can be computed by a series of modulo-2 additions. Now the balanced conservative array GS coding is formulated as the following integer programming problem.

$$\begin{aligned} z_{strength} = {} & \text{Max } t\\ & \text{s.t. } u_{i,j} = c_{i,j+1} \circ c_{i,j},\ \forall i,j \quad w_{i,j} = c_{i+1,j} \circ c_{i,j},\ \forall i,j\\ & \qquad t \le \sum_j u_{i,j},\ t \le \sum_i w_{i,j},\ \forall i,j, \quad \text{constraints (1).} \end{aligned} \tag{2}$$

Here, variables $u_{i,j}(w_{i,j})$ are auxiliary variables to compute the number of transitions in rows (columns) and $\sum_j u_{i,j}$ and $\sum_i w_{i,j}$ are the number of transitions in row i and column j, respectively. (2) is a mixed integer programming model with 0-1 variables $d_1,\dots,d_p$. (2) with the pseudo-balanced constraint has constraint $\left|\sum_{ij} c_{i,j} - m^2/2\right| \le b$ instead of $\sum_{ij} c_{i,j} = m^2/2$.

Another DC-suppression criteria used in GS coding are minimum running digital sum (MRDS), minimum squared weight (MSW) and minimum threshold overrun (MTO) [10,11]. We denote $s_{i,j}$ as the RDS starting from $c_{i,1}$ to $c_{i,j}$ in row i and $t_{i,j}$ RDS from $c_{1,j}$ to $c_{i,j}$ in column j. Then $s_{i,j}$ and $t_{i,j}$ are computed as

$$s_{i,j} = s_{i,j-1} + 2c_{i,j} - 1,\ j = 1, \ldots, m, \quad t_{i,j} = t_{i-1,j} + 2c_{i,j} - 1,\ i = 1, \ldots, m, \tag{3}$$

where $s_{i,0} = 0$, $t_{0,j} = 0$, $\forall i, j$. The MRDS criteria is to choose candidate array minimizing $\gamma = \max_{i,j}\{|s_{i,j}|, |t_{i,j}|\}$. Thus, γ satisfy the following inequalities

$$\gamma \geq s_{i,j},\ \gamma \geq -s_{i,j},\ \gamma \geq t_{i,j},\ \gamma \geq -t_{i,j},\ \forall i, j. \tag{4}$$

Then, MRDS GS coding is formulated as the following integer programming problem

$$\begin{aligned} z_{MRDS} = {} & \text{Min } \gamma \\ & \text{s.t. constraints (1), (3), (4).} \end{aligned} \tag{5}$$

When the number of transitions in any row or column is t_k, then, $\gamma \leq m - t_k$. Since strength t is the minimum in any row or column, we have

$$\gamma \leq m - t. \tag{6}$$

3. **Computer Simulation.** We applied our model to randomly generated user data for several array sizes ranging from 10×10 to 20×20 and control bit sizes from 6 to 10 bits. For each array size and control bit, we generated 20 cases and in each case, the source data are randomly generated from discrete uniform $(0, 1)$. In Table 1, we record the average strength values of 20 random instances for various combinations of control bit and array size, with balanced and pseudo-balanced constraints. Strength value from the algorithm in [4] guarantees the existence of strength $t \leq m/4$ for $m \times m$ array if $m = 2^d$ using $(m+1)^2$ control arrays. For $d = 4$, average strength in Table 1 is higher than $4 = 16/4$. Also, in algorithm from [6], the probability of balanced t-conservative array in the selection set is about 0.5 and for $\left|\sum_{ij} c_{i,j} - m^2/2\right| \leq 2$ pseudo-balanced case, it approaches 1.0. These estimates are obtained from simulation of 10^5 random instances. For each problem, the strength from the balanced case is a lower bound for the pseudo-balanced case. For 10×10 balanced array, the average strength for the balanced and pseudo-balanced case is about 3 and as the number of control bits increases, the strength also increases. For the same size array with pseudo-balanced case, the strength values are higher than the balanced case. This trend is similar for every array size. The guaranteed minimum strength of the pseudo-balanced 16×16 array is 4 in [5], while in our simulation, the observed average strength is 7 in Table 1.

In Figure 1, we recorded average strength and z_{MRDS} for array sizes ranging from 10×10 to 20×20 and control bits ranging from 6 to 10 for each array size. Also, we recorded MRDS value computed from the optimal solution of model (2) (MRDS in $z_{strength}$). Notice that MRDS value in the optimal solution of balanced conservative array is an upper bound

Table 1. Average strength for balanced and pseudo balanced arrays

array size	balanced					pseudo balanced				
	number of control bits					number of control bits				
	6	7	8	9	10	6	7	8	9	10
10×10	2.4	2.95	2.85	3.2	3.35	3	3.35	3.45	3.6	3.8
12×12	3.5	3.8	4.2	4.9	5.2	4.2	4.25	4.55	5	5.45
14×14	5.45	5.6	6.55	6.9	7.05	5.5	5.85	6.7	6.95	7.1
16×16	7.35	7.3	7.75	8.1	8.5	7.4	7.4	7.95	8.2	8.6
18×18	8.05	8.55	8.95	9.35	9.75	8.6	8.85	9.05	9.35	9.9
20×20	8	9.1	10	10.8	11.2	9.05	9.4	10.1	10.9	11.3

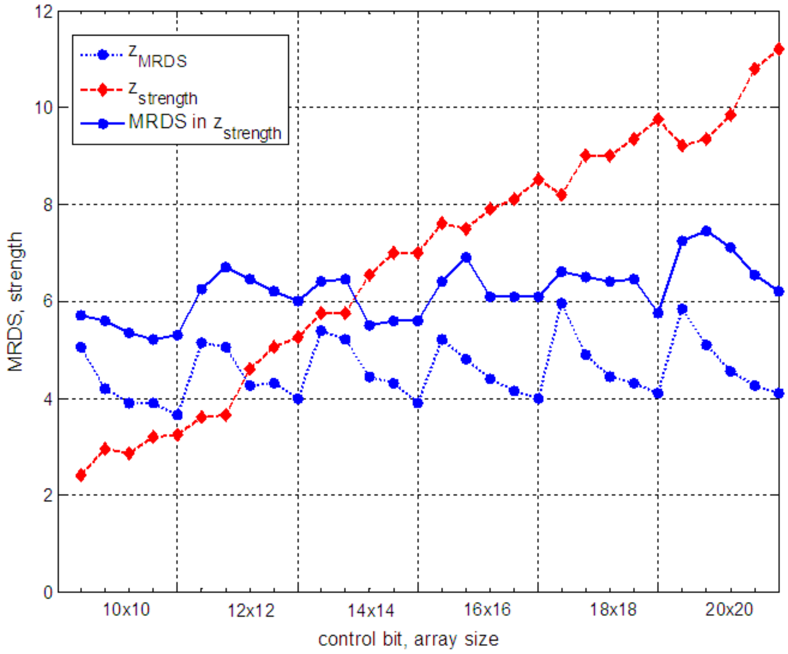

FIGURE 1. $z_{strength}$, z_{MRDS} and MRDS values for various control bit/array sizes

to z_{MRDS} and the difference between these two values is limited within two. As the size of the array increases, strength t increases almost linearly while z_{MRDS} decreases with the number of control bits but z_{MRDS} values for same number of control bits are almost similar for different array sizes. Similar trend can be found in MRDS values from model (2).

4. **Conclusions.** In this paper, scrambling process using self-synchronizing polynomial is formulated as set of linear inequalities and using these formulations, balanced conservative array GS coding and MRDS GS coding for HDS are formulated as integer programming models.

We find that the strength bound from the balanced conservative array GS coding increases with the size of array and control bits, while the MRDS bound decreases with the number of control bits. The proposed integer programming model does not require $(t-2)$-error correcting code as input and also the average strength bounds obtained are higher than the lower bound using cascaded coding in the previous algorithm [4]. As shown in our simulation, conservative arrays show strong performance in terms of MRDS. With respect to the run length limited (RLL) constraint, we found that the conservative arrays have poor performance. We observed that there are long consecutive '0's and '1's in the optimal conservative arrays. Note that in n bits with strength t, there is a possibility of at most $n-t-1$ runs of the same symbol similar to Equation (6).

Based on the proposed integer programming model, one can evaluate the effect of choosing different scrambling polynomial on the GS performance.

REFERENCES

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

[2] H. Coufal, D. Psaltis and G. T. Sincerbox, *Holographic Data Storage*, Springer-Verlag, Berlin, 2000.

[3] B. Marcus, Modulation codes for holographic recording, in *Holographic Data Storage*, H. Coufal, D. Psaltis and G. T. Sincerbox (eds.), Berlin, Springer-Verlag, 2000.

[4] A. Vardy, M. Blaum, P. H. Siegel and G. T. Sincerbox, Conservative arrays: Multidimensional modulation codes for holographic recording, *IEEE Trans. Inf. Theory*, vol.42, pp.227-230, 1996.

[5] R. Talyansky, T. Etzion and R. M. Roth, Efficient code constructions for certain two-dimensional constraints, *IEEE Trans. Inf. Theory*, vol.45, pp.794-799, 1999.

[6] Y. Zhu and I. J. Fair, Multimode two-dimensional balanced conservative codes for holographic storage, *IET Commun.*, vol.4, no.15, pp.1809-1816, 2010.

[7] I. J. Fair, W. D. Grover, W. A. Krzymien and R. I. MacDonald, Guided scrambling: A new line coding technique for high bit rate fiber optic transmission systems, *IEEE. Trans. Commun.*, vol.39, pp.289-297, 1991.

[8] 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. of SPIE*, pp.191-196, 2000.

[9] N. Y. Kim, J. Lee, Y. Hong and J. Lee, Optimal number of control bits in the guided scrambling method for holographic data storage, *Jpn. J. Appl. Phys*, vol.44, pp.3449-3452, 2005.

[10] S. W. Phillips and I. J. Fair, Sparse binary-to-ternary encoding for holographic storage, *IEICE Trans. Fundamentals*, vol.E97A, pp.1231-1239, 2014.

[11] K. A. S. Immink, *Codes for Mass Data Storage Systems*, Shannon Foundation Publisher, Eindhoven, 2004.

[12] K. A. S. Immink and L. Pátrovics, Performance assessment of dc-free multimode codes, *IEEE Trans. Commun.*, vol.45, pp.293-299, 1997.