ICIC Express Letters
Volume **10**, Number **6**, June **2016**
ICIC International ©2016 ISSN 1881-803X
pp. 1457–1460

A BINARY PARTICLE SWARM OPTIMIZATION FOR THE GUIDED SCRAMBLING ENCODING FOR HOLOGRAPHIC STORAGE

Taehyung Park[1,*] and Dae-Hwan Kim[2]

[1]Department of Industrial and Information Systems Engineering
Soongsil University
No. 369, Sangdo-ro, Dongjak-gu, Seoul 156-743, Korea
*Corresponding author: tpark@ssu.ac.kr

[2]Energy IT Convergence Research Center
Korea Electronics Technology Institute
No. 25, Saenari-ro, Bundang-gu, Seongnam 463-816, Korea
kimdh@keti.re.kr

Received December 2015; accepted March 2016

Abstract. *Guided scrambling (GS) encoding evaluates all candidates codewords generated from all possible combinations of control bits to find a balanced code array with the maximum strength value. In this paper, we applied binary particle swarm optimization for finding balanced code array with good strength value. In the computational results, we show average strengths and symbol balance errors from particle swarm optimization heuristic. We find that the average strength value computed from particle swarm optimization heuristic is quite tight to the GS encoding method using only 6% of the candidate codewords.*
Keywords: Guided scrambling, Conservative array, Particle swarm optimization, Holographic storage, Heuristic algorithm

1. **Conservative Array and Guided Scrambling Coding.** In holographic data storage (HDS), interference pattern between an optical representation of data page and a reference beam is stored into the volume of the holographic material and retrieved through the use of a charge-coupled device (CCD) reading 2-D pixel image data [1]. To increase the reliability of the holographic systems, the pattern of '0's and '1's must satisfy modulation constraints. It is required that in each row and column of the encoded array, there are at least t transitions of $0 \rightarrow 1$ or $1 \rightarrow 0$. A binary array of this property is called a conservative array of *strength* t [2,3]. Another constraint is to require that the occurrence of '0' and '1' pixels should be equally likely.

One of the simple multimode encoding techniques is the guided scrambling (GS) approaches [4-6]. In GS, each source word of length n is combined with p control bits, scrambled into $n+p$ bit codeword, and then formatted into $m \times m$ square array. Encoder evaluates the quality of all possible 2^p codewords with respect to modulation constraints, and the best codeword is selected and transmitted. The advantage of the GS coding is its simplicity, but it must evaluate all 2^p candidates to choose the best. In this paper, we apply a binary particle swarm optimization algorithm to finding a good candidate array with almost balanced high strength value.

The rest of this paper is organized as follows. In Section 2, we describe the proposed binary particle swarm optimization for the GS coding and a simple neighborhood search heuristic as well as the performance measures derived from the GS coding. In Section 3, we compare the performance of the proposed particle swarm optimization heuristic and the simple neighborhood search heuristic with respect to GS coding. Section 4 concludes the paper.

2. **Binary Particle Swarm Optimization.** Particle swarm optimization (PSO) is an evolutionary algorithm where sets of particles are moving toward the global optimal solution [7-9]. Each particle $\mathbf{x}_i(t) \in \{0,1\}^d$ and population consist of M particles. At iteration t, each particle $\mathbf{x}_i(t)$, $i = 1, \dots, M$ computes its velocity vector $\mathbf{v}_i(t)$ using its personal best solution $\mathbf{p}_i^{best}(t)$ and the current best global solution $\mathbf{g}^{best}(t)$. Notice that $\mathbf{p}_i^{best}(t) = \arg\max\{f(\mathbf{x}_i(t')) : t' \le t\}$ and $\mathbf{g}^{best}(t) = \operatorname{argmax}\{f(\mathbf{x}_i(t')) : t' \le t, \forall i\}$ with respect to the objective function $f(\mathbf{x}(t))$. At the initialization, $\mathbf{x}_i(t)$ is a random binary vector and $\mathbf{v}_i(t) = 0$.

In the binary PSO, the velocity vector and solution update are described in the following equations.

$$\mathbf{v}_i(t+1) = w\mathbf{v}_i(t) + c_1 R_{i1}\left(\mathbf{p}_i^{best}(t) - \mathbf{x}_i(t)\right) + c_2 R_{i2}\left(\mathbf{g}^{best}(t) - \mathbf{x}_i(t)\right)$$

$$\mathbf{x}_{ij}(t+1) = \begin{cases} 0 & \text{if } U \ge \sigma(\mathbf{v}_{ij}(t+1)) \\ 1 & \text{if } U < \sigma(\mathbf{v}_{ij}(t+1)) \end{cases},$$

where $\sigma(x) = 1/(1+e^{-x})$ is the sigmoid function and U is a uniform $(0,1)$ random number, R_{i1}, R_{i2} are diagonal matrices where each diagonal element is from uniform $(0,1)$. Weights w, c_1, c_2 are further defined as follows.

$$w = 1\Big/\left(\phi - 1 + \sqrt{\phi^2 - 2\phi}\right),$$

$$c_1 = c_2 = \phi w, \ \phi \in [2.01, 2.4].$$

In this paper, $\phi = 2.2$. Notice that although velocity vector $\mathbf{v}_i(t)$ is a general real vector, $\mathbf{x}_i(t+1)$ is still a binary vector.

Instead of searching better particles in the swarm, simple neighborhood search algorithm can be applied to GSO coding. The following steps summarize the neighborhood search heuristic. For $\mathbf{x} \in \{0,1\}^d$, define 1-dimensional neighbor $\tilde{\mathbf{x}}_k = e_k \oplus \mathbf{x}$, where e_k is the kth unit vector and $\oplus$ is modulo-2 addition. Similarly, 2-dimensional neighbor $\tilde{\mathbf{x}}_{jk} = \tilde{\mathbf{x}}_k \oplus e_j$, $j \ne k$. So, the neighbor vector is the one with bit-flipped in one or two positions from the original binary vector $\mathbf{x}$. The improvement measure between two binary vectors $\mathbf{x}^{pre}$ and $\mathbf{x}$ is defined as $\Delta(\mathbf{x}, \mathbf{x}^{pre}) = c_1(t(\mathbf{x}) - t(\mathbf{x}^{pre})) + c_2(\tau(\mathbf{x}) - \tau(\mathbf{x}^{pre}))$, where c_1, c_2 are positive constants, function $t(\mathbf{x})$ is the *strength* value in the vector $\mathbf{x}$ represented as $m \times m$ scrambled array $(c_{i,j})_{i,j=1}^m$ and $\tau(\mathbf{x})$ is the symbol balance penalty function. We assume $\tau(\mathbf{x}) = \mathbf{0}$, if $|\sum_{i,j=1}^m c_{i,j} - m^2/2| < 2$, and $\tau(\mathbf{x}) = -2|\sum_{i,j=1}^m c_{i,j} - m^2/2|$. Otherwise, one-dimensional neighborhood search heuristic is as follows.

Step 1. Iteration $n = 0$, begin with $\mathbf{x}(\mathbf{0}) = (x_1, \dots, x_d)$ where x_i are random binary number.

Step 2. $\bar{k} = \operatorname{argmax}\{\Delta(\tilde{\mathbf{x}}(n)_k, \mathbf{x}(n))\}$

Step 3. If $\Delta(\tilde{\mathbf{x}}(n)_{\bar{k}}, \mathbf{x}(n)) > 0$, $\mathbf{x}(n+1) = \tilde{\mathbf{x}}(n)_{\bar{k}}$, $n = n+1$, go to Step 2. Otherwise, stop.

3. **Computational Results.** We applied binary PSO algorithm to GS encoding for finding the balanced conservative array with the maximum strength. For this PSO, each vector $\mathbf{x}_i(t)$ is the control bit vector in $\{0,1\}^p$ and the performance measure is defined as $f(\mathbf{x}_i(t))$ = strength − max{symbol balance error −2, 0}. The symbol balance error is computed as $|\sum_{i,j=1}^m c_{i,j} - m^2/2|$, where $c_{i,j}$ is the scrambled code in $m \times m$ array. Notice that if one applies GS coding for $m = 20$, $p = 10$, encoder needs to evaluate all possible 2^{10} candidate arrays. We set the population size $M = \lceil 2^p/k \rceil$, where $k \in \{5, 7, 10, 14\}$, that is, the population size is approximately 7%, 10%, 14%, and 20% of the 2^p GS coding candidates. For example, when the number of control bit is 10, the candidates in the GS coding is 1,024, while in the PSO, the smallest population we use is 74. Figure 1 shows the average strength values for $m = 10, 12, 14, 16, 18, 20$. For each m, we considered cases for control bit size $p = 6, 7, 8, 9, 10$ and for each (m, p) combination, we generated 20

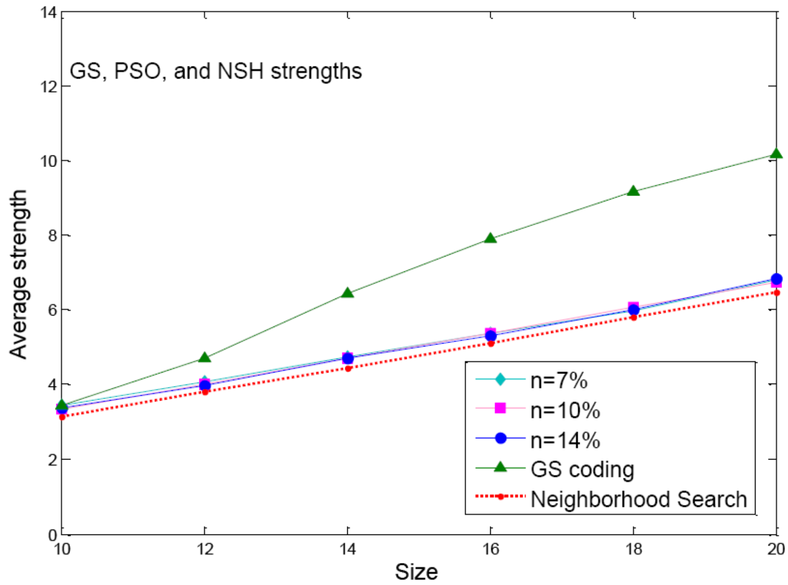

FIGURE 1. GS and PSO, neighborhood search strengths

TABLE 1. Average strength and symbol balance error

Population size	Control bit size	Average strength	Symbol balance error
7%	6	6.45	1.25
	7	6.65	1.25
	8	6.9	1.25
	9	7.05	0.7
	10	7.1	1.15
10%	6	6.35	1.15
	7	6.45	1.2
	8	6.75	1.1
	9	7	1.1
	10	7.05	1.3
14%	6	6.35	1
	7	6.85	1.3
	8	6.85	1.15
	9	6.95	1.25
	10	7.05	0.95
20%	6	6.4	1.35
	7	6.7	1.1
	8	6.8	1
	9	7.05	0.85
	10	7.05	0.85

random cases. The average shown in Figure 1 is the average across all control bit sizes, so that for each m, it is an average of 100 cases.

Comparing with the GS coding value, we find that the PSO lower bound is ranging from 97% to 67% as the size m increases. The PSO strength values between different population sizes are all very close. So using only 7% of the GS coding candidates, PSO finds good candidate array with enough strength value. Also, PSO lower bound always outperforms the bound from neighborhood search heuristic. In neighborhood search heuristic, we set $c_1 = 8$, $c_2 = 1$. Previously, different multimode coding for the conservative array suggests strength $t = 4$ for 16×16 sized arrays [2]. Our computation shows that in GS coding, average strength in 16×16 cases is approximately 7 while the average strength from PSO heuristic provides $t \approx 5$. In Table 1, we show detailed results for $m = 20$, $p = 6, 7, 8, 9, 10$. We also show the symbol balance error of the best solution. From Table 1, we can see that in PSO, increasing the population size has minor effects on the strength values and the symbol balance error. Zero symbol balance error means there are equal numbers of 0's and 1's in $m \times m$ array. The symbol balance error in the PSO optimal solution is less than 1.25 in most cases.

4. **Conclusion.** Holographic storage is a promising candidate of the next-generation data storage system. To increase capacity and reliability of the holographic system, achieving the target bit error rate is essential and careful engineering of optical system combined with the signal processing and modulation coding is required. This paper shows that a binary PSO algorithm is an effective multimode coding scheme to obtain balanced conservative array. Future research topic includes effective coding scheme for the two-dimensional modulation constraints such as prohibiting isolated pixel or run-length condition and the computation of asymptotic coding capacity of these modulation constrained codes.

Acknowledgment. This work is partially supported by the IT R&D Program of MKE/KEIT. [2015-10054762, Development of Business Continuity Support Model and System for Crisis and Risk Management of Logistic Supply Chain].

REFERENCES

[1] V. Vadde and B. V. Kumar, Channel modeling and estimation for intrapage equalization in pixel-matched volume holographic data storage, *Applied Optics*, vol.38, pp.4374-4386, 1999.
[2] 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.
[3] 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.
[4] 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*, vol.4090, pp.191-196, 2000.
[5] 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.
[6] T. Park and J. Lee, Worst case performance assessment of DC-free guided scrambling coding by integer programming model, *IEEE Trans. Magnetics*, vol.50, no.7, pp.1-4, 2014.
[7] J. Kennedy and R. C. Eberhart, A discrete binary version of the particle swarm algorithm, *Proc. of Conference on System, Man, Cybernetics*, pp.4104-4108, 1997.
[8] S. Lee, S. Soak, S. Oh, W. Pedrycz and M. Jeon, Modified binary particle swarm optimization, *Progress in Natural Science*, vol.18, pp.1161-1166, 2008.
[9] H. Banka and S. Dara, A hamming distance based binary particle swarm optimization (HDBPSO) algorithm for high dimensional feature selection, classification and validation, *Pattern Recognition Letters*, vol.52, pp.94-100, 2015.