1

Modulation Codes for Optical Data Storage

Tzi-Dar Chiueh and Chi-Yun Chen
National Taiwan University
Taipei, Taiwan

1. Introduction

In optical storage systems, sensitive stored patterns can cause failure in data retrieval and decrease the system reliability. Modulation codes play the role of shaping the characteristics of stored data patterns in optical storage systems. Among various optical storage systems, holographic data storage is regarded as a promising candidate for next-generation optical data storage due to its extremely high capacity and ultra-fast data transfer rate. In this chapter we will cover modulation codes for optical data storage, especially on those designed for holographic data storage.

In conventional optical data storage systems, information is recorded in a one-dimensional spiral stream. The major concern of modulation codes for these optical data storage systems is to separate binary ones by a number of binary zeroes, i.e., run-length-limited codes. Examples are the eight-to-fourteen modulation (EFM) for CD (Immink et al., 1985), EFMPlus for DVD (Immink, 1997), and 17 parity preserve-prohibit repeated minimum run-length transition (17PP) for Blu-ray disc (Blu-ray Disc Association, 2006). Setting constraint on minimum and maximum runs of binary zeros results in several advantages, including increased data density, improved time recovery and gain control and depressed interference between bits.

In holographic data storage systems, information is stored as pixels on two-dimensional (2-D) pages. Different from conventional optical data storage, the additional dimension inevitably brings new consideration to the design of modulation codes. The primary concern is that interferences between pixels are omni-directional. Besides, since pixels carry different intensities to represent different information bits, pixels with higher intensities intrinsically corrupt the signal fidelity of those with lower intensities more than the other way around, i.e., interferences among pixels are imbalanced. In addition to preventing vulnerable patterns suffering from possible interferences, some modulation codes also focus on decoder complexity, and yet others focus on achieving high code rate. It is desirable to consider all aspects but trade-off is matter-of-course. Different priorities in design consideration result in various modulation codes.

In this chapter, we will first introduce several modulation code constraints. Next, one-dimensional modulation codes adopted in prevalent optical data storage systems are discussed. Then we turn to the modulation codes designed for holographic data storage. These modulation codes are classified according to the coding methods, i.e., *block codes* vs. *strip codes*. For block codes, code blocks are independently produced and then tiled to form a

whole page. This guarantees a one-to-one relationship between the information bits and the associated code blocks. On the contrary, strip codes produce code blocks by considering the current group of information bits as well as other code blocks. This type of coding complicates the encoding procedure but can ensure that the constraints be satisfied across block boundary. We will further discuss variable-length modulation codes, which is a contrast to fixed-length modulation codes. Variable-length modulation codes have more freedom in the design of code blocks. With a given code rate, variable-length modulation codes can provide better modulated pages when compared to fixed-length modulation codes. However, variable-length modulation codes can suffer from the *error propagation* problem where a decoding error of one code block can lead to several ensuing decoding errors.

2. Constraints

Generally speaking, constraints of modulation codes are designed according to the channel characteristics of the storage system. There are also other considerations such as decoder complexity and code rate. In conventional optical data storage systems, information carried in a binary data stream is recorded by creating marks on the disk with variable lengths and spaces between them. On the other hand, information is stored in 2-D data pages consisting of ON pixels and OFF pixels in holographic data storage system. The holographic modulation codes encode one-dimensional information streams into 2-D code blocks. The modulated pages created by tiling code blocks comply with certain constraints, aiming at reducing the risk of corrupting signal fidelity during writing and retrieving processes. Due to the additional dimension, other considerations are required when designing constraints for holographic modulation codes. In the following some commonly adopted constraints are introduced.

2.1 Run-Length Limited Constraint

The run-length limited constraint is widely adopted in optical storage systems. Examples are the eight-to-fourteen modulation (EFM) in CD, EFMPlus in DVD, and the 17 parity preserve-prohibit repeated minimum run-length transition (17PP) in Blu-ray disc. Due to the different reflectivity states, peak detection is the most common receiver scheme. To reliably detect peaks, separation on the order of 1.5-2 mark diameters is required between marks (McLaughlin, 1998). The run-length limited constraint thus sets limits to the frequency of ones in the data stream to avoid the case where large run-length of zeros causes difficulty of timing recovery and streams with small run-length of zeros have significant high-frequency components that can be severely attenuated during readout.

From one-dimensional track-oriented to 2-D page-based technology, the run-length between "1"s (or ON pixels) has to be extended to 2-D. Two run-length limited constraints for 2-D patterns have been proposed. One constraint sets upper and lower bounds to run-length of zeros in both horizontal and vertical directions (Kamabe, 2007). The other one sets upper and lower bounds to 2-D spatial distance between any two ON pixels (Malki et al., 2008; Roth et al., 2001). See Fig. 1 for illustration of this 2-D run-length limited constraint.

2.2 Conservative Constraint

The conservative constraint (Vardy et al., 1996) requires at least a prescribed number of transitions, i.e., 1→0 or 0→1, in each row and column in order to avoid long periodic stretches of contiguous light or dark pixels. This is because a large area of ON pixels results in a situation similar to over-exposure in photography. The diffracted light will illuminate the dark region and lead to false detection. An example of such detrimental pattern is shown in Fig. 2.

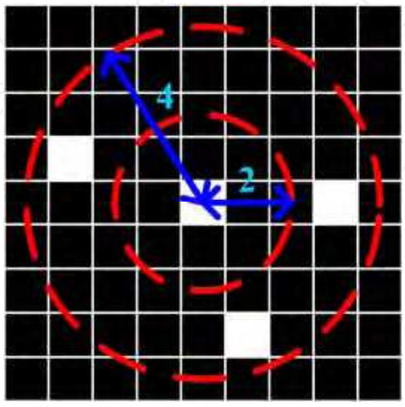


Fig. 1. Illustration of the 2-D run-length limited constraint based on 2-D spatial distance with lower and upper bounds of two and four, respectively.

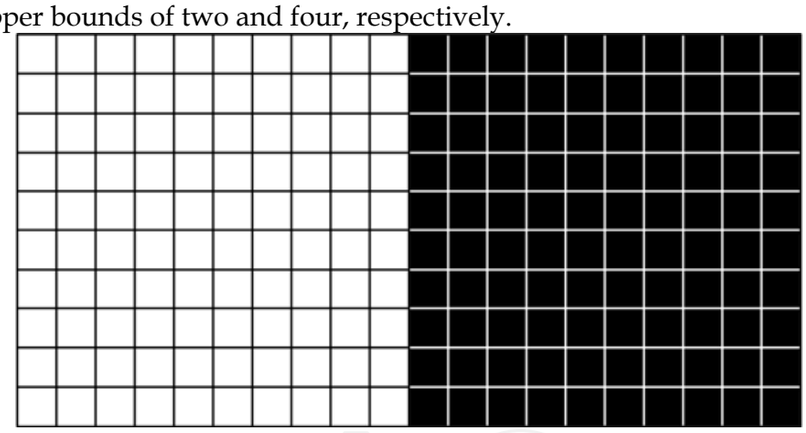


Fig. 2. A pattern forbidden by the conservative constraint (Blaum et al., 1996).

2.3 Low-Pass Constraint

The low-pass constraint (Ashley & Marcus, 1998; Vadde & Vijaya Kumar, 2000) excludes code blocks with high spatial frequency components, which are sensitive to inter-pixel interference induced by holographic data storage channels. For example, an ON pixel tends to be incorrectly detected as "0" when it is surrounded by OFF pixels and similarly an OFF pixel tends to be incorrectly detected as "1" when surrounded by ON pixels. Therefore, such patterns are forbidden under the low-pass constraint.

In fact, the low-pass constraints can be quite strict so that the legal code patterns have good protection against high-frequency cutoff. This is, however, achieved at the cost of lower code rate because few code blocks satisfy this constraint. Table 1 lists five low-pass constraints and examples that violate these constraints.

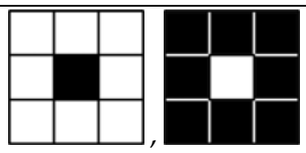
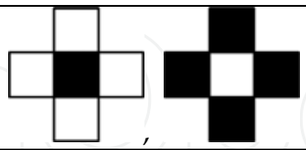
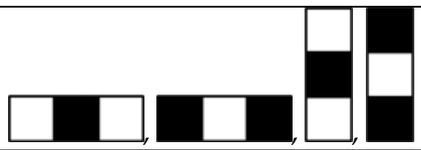
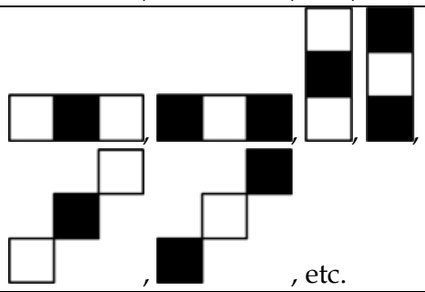
Constraints	Examples of forbidden patterns
At least one ON/OFF pixel exists among eight nearest neighboring pixels of an ON/OFF pixel.	
At least one ON/OFF pixel exists among four nearest neighboring pixels of an ON/OFF pixel.	
At least two ON/OFF pixels exist among four nearest neighboring pixels of an ON/OFF pixel.	 , etc.
No ON/OFF pixels are isolated either horizontally or vertically.	
No ON/OFF pixels are isolated either horizontally, vertically or diagonally.	 , etc.

Table 1. Low-pass constraints and examples of forbidden patterns.

2.4 Constant-Weight Constraint

The weight of a page is defined as the ratio of the number of ON pixels over the number of all pixels. With a constant-weight page, low-complexity correlation detection can be efficiently implemented (Coufal et al., 2000; Burr & Marcus, 1999; Burr et al., 1997). Two kinds of weight distribution garner interests of researchers. One is balanced weight, which makes the numbers of ON pixels and OFF pixels the same. The other is sparse weight, which makes the number of ON pixels less than that of OFF pixels. Balanced-weight modulation codes have higher code rates than sparse-weight modulation codes with the same code block size. However, due to imbalanced interference between ON pixels and OFF pixels, OFF pixels are favored as they cause lower level of interference. Therefore, sparse codes enable more data pages that can be superimposed at the same location on the recording medium by reducing optical exposure and increasing diffraction efficiency (the ratio of the power of the diffracted beam to the incident power of that beam) per pixel (Daiber et al., 2003). In addition, the probability of undesirable patterns can be reduced by decreasing ON pixel density. To be more exact, up to 15% improvement of total memory capacity can be achieved by sparse codes with the page weight decreased to 25% (King & Neifeld, 2000).

2.5 Summary

We have introduced four types of constraints for modulation codes commonly adopted in optical data storage: run-length limited, conservative, low-pass and constant-weight constraints. The run-length limited constraint focuses on the density of ON pixels. The conservative constraint considers the frequency of transitions between ON and OFF pixels. The low-pass constraint helps avoid those patterns vulnerable to inter-pixel interference effects in the holographic data storage systems. As for the constant-weight constraint, it enables a simple decoding scheme by sorting pixel intensities. Sparse modulation codes further decrease the probability of vulnerable patterns and increase the number of recordable pages.

3. One-Dimensional Modulation Codes

The modulation codes adopted by current optical data storage systems, i.e., CD, DVD, and Blu-ray disc, are developed according to the run-length limited constraint. As we mentioned previously, short runs result in small read-out signal power which tends to cause errors while long runs cause failure in time recovery. Therefore, a run-length limited code in the non-return-to-zero (NRZ) notation requires the number of "0"s between two "1"s to be at least d and at most k. The real bits recorded on the disc are then transformed from NRZ to non-return-to-zero inverted (NRZI) format consisting of sequence of runs which changes polarity when a "1" appears in the NRZ bit stream, as shown in Fig. 3.

Under the EFM rule adopted by CD, 8-bit information sequences are transformed into 14-bit codewords using a table. The 14-bit codewords satisfy the (2, 10)-run-length limited constraint in the NRZ notation, that is, two "1"s are always separated by at least two "0"s and at most ten "0"s. To make sure bits across two adjacent codewords also comply with the run-length limited constraint, three emerging bits precede every 14-bit codeword. Although two merging bits are sufficient to meet this requirement, one additional bit is used in order to do DC-control, that is, make the *running digital sum*, which is the sum of bit value (±1) from the start of the disc up to the specified position, close to zero as much as possible. Therefore, total 17 bits are needed to store eight information bit in the EFM code.

The EFMPlus code used in DVD translates sequences of eight information bit into 16-bit codewords satisfying the (2, 10)-run-length limited constraint. Unlike EFM, which is simply realized by a look-up table, EFMPlus code exploits a finite state machine with four states. No merging bits are required to promise no violation across codewords. In addition, to circumvent error propagation, the encoder is further designed to avoid state-dependent decoding.

Blu-ray disc boosts data capacity by further shrinking physical bit size recorded on the disc, leading to severer interference than CD and DVD. In Blu-ray disc, the 17PP modulation code follows the (1, 7)-run-length limited constraint and exploits a new DC-control mechanism. The DC-control is realized by inserting DC-bits at certain DC-control points in the information stream. The polarity of the corresponding NRZI bit stream is flipped if a DC-bit is set as "1" while the polarity remains if a DC-bit is set as "0". Therefore, the best DC-free bit stream that makes the running digital sum closest to zero can be obtained by properly setting the DC-bits. The DC-control mechanism is illustrated in Fig. 4.

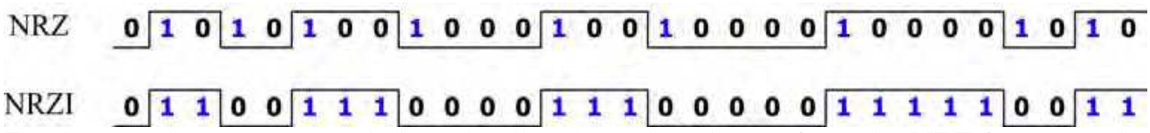


Fig. 3. Actual patterns recorded on the disc through a change from the NRZ format to the NRZI format.

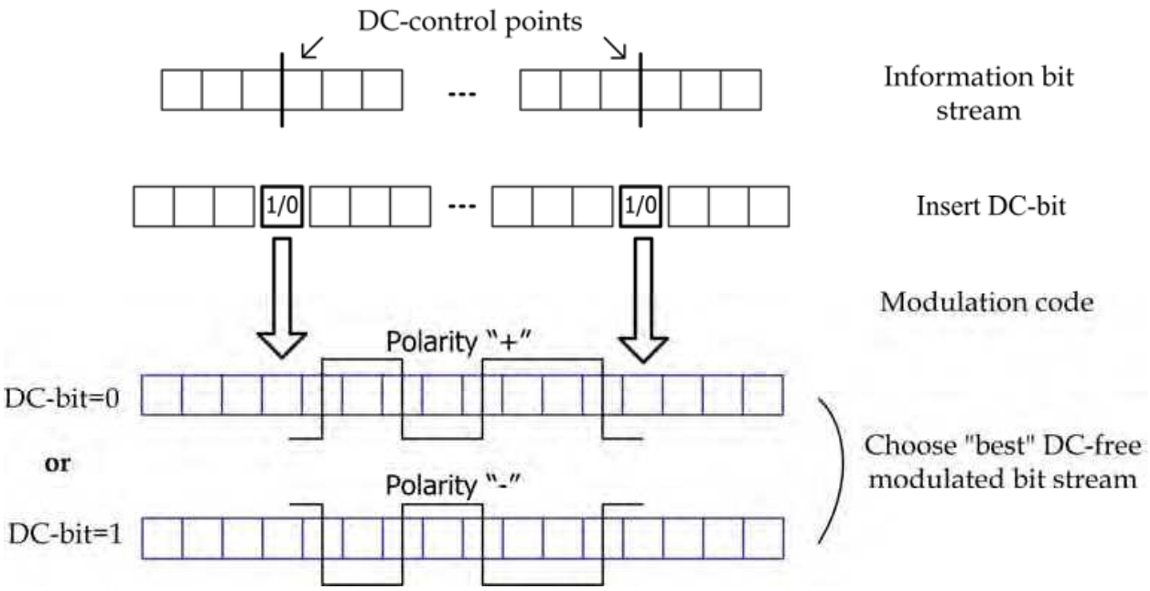


Fig. 4. Illustration of DC control in the 17PP modulation code (Blu-ray Disc Association, 2006).

4. Block Codes for Holographic Data Storage

A block code used in holographic data storage maps a sequence of information bits into a 2-D code block in a one-to-one manner. The encoded blocks are tiled to create a whole page. The ratio of the number of information bits to the number of block pixels is called *code rate*. For example, a simple block code, named *differential code*, uses two pixels to represent one bit and achieves a code rate of 1/2. Its code blocks are illustrated in Fig. 5. Basically, higher code rate is preferred since less redundant pixels are included and more information is carried in the final modulated page. For this purpose, enlarging block size or changing the average intensity is applied (Kume et al., 2001). Fig. 6 illustrates a 5:9 block code with a code rate of 5/9 and a 6:9 block code with a code rate of 2/3. With two of nine pixels in a code block turned ON, there exist $C_2^9 = 36$ possible code blocks. Therefore, these coded blocks are sufficient to represent five information bits, giving a 5:9 block code. Similarly, with three of nine pixels in a code block turned ON, there are $C_3^9 = 84$ possible coded blocks, thus achieving a 6:9 block code.

However, a larger block size may deteriorate decoding performance due to biased intensity throughout the block. Higher intensity also degrades the quality of reconstructed images according to the sparse code principle. On the other hand, increasing the number of possible code blocks (and thus code rate) often leads to more complex decoding. Hence, modulation

code design is a trade-off among higher code rate, simpler decoder and satisfactory BER performance.

Block codes indeed provide a simple mapping in encoding and decoding. However, the risk of violating constraints becomes significant when tiling multiple blocks together since illegal patterns may occur across the boundary of neighboring blocks. To circumvent this problem, additional constraints may be required. Unfortunately, more constraints can eliminate some patterns that would have been legitimate. To maintain the code rate, a larger block size is called for.

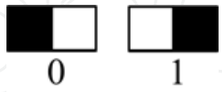


Fig. 5. Differential code.

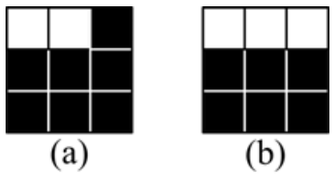


Fig. 6. (a) A code block in the 5:9 block code and (b) a code block in the 6:9 block code.

4.1 6:8 Balanced Block Code

The 6:8 balanced block code (Burr et al., 1997) is one of the most common block modulation codes used in the holographic data storage systems. A group of six information bits is encoded into a 2×4 block with exactly four ON pixels and four OFF pixels, satisfying the balanced constraint. Since $C_4^8 = 70$ is larger than 2^6=64, we choose 64 patterns from 70 code blocks and assign Gray code to these blocks . Correlation detection (Burr et al., 1997) is also proposed to decode such block code. In this method the correlations between the retrieved block and all code blocks are computed and the code block with the maximum correlation to the received block is declared the stored pattern.

4.2 6:8 Variable-Weight Block Code

The constant-weight constraint can have negative impacts on the achievable code rate of a modulation code. For instance, if we apply a sparse constant-weight constraint to a 2×4 block code and require two ON pixels and six OFF pixels in every code block, then there are only 28 legal code blocks, giving four information bits and a low code rate of 1/2.

The 6:8 variable-weight modulation code (Chen & Chiueh, 2007) is a sparse block modulation code with variable weight and relatively high code rate. Similar to the 6:8 balanced block code, the variable-weight code encodes a group of six data bits into a 2×4 block. There are $C_1^8 = 8$ codewords with one ON pixel and $C_3^8 = 56$ codewords with three ON pixels, enough to represent 6-bit information. This design decreases the weight of a modulated page from 50% to 34.38%. Therefore, the variable-weight code provides better BER performance with the same coding scheme as the 6:8 balanced code performance due to lower interference level from fewer ON pixels.

4.3 Codeword Complementing Block Modulation Code

The codeword complementing block modulation code (Hwang et al., 2003) adopts an indirect encoding scheme so as to satisfy the pseudo-balanced constraint and thus achieve uniform spectrum of the modulated pages. The coding procedure consists of two steps: *constant weight encoding* and *selective complementary encoding*. The constant weight encoder maps input information bits into one-dimensional constant-weight codewords. Each one-dimensional N-bit codeword forms a row in an $(N+1)\times N$ code matrix. Then, some rows are flipped according to the elements in the last row of the current code matrix. If a pixel in the last row is "1", the selective complementary encoder flips the corresponding row. In this way, the final $N\times N$ modulated block has constant weight and relatively uniform spectrum in the Fourier plane. An example of the codeword complementing block modulation code encoding process with N=11 is illustrated in Fig. 7.

The decoding procedure is simple. By calculating the average intensity of each row of a retrieved block, rows that have been complemented by the selective complementary encoder are easily identified and the polarities of the pixels are reversed. Moreover, the last row in the code matrix can be decided. Decoding the constant weight encoded blocks is achieved by first sorting the intensity of the received pixels in each codeword and marking proper number of pixels with higher intensity as "ON" and the other pixels of the codeword as "OFF." Then information bits are obtained by inverse mapping of constant weight encoding.

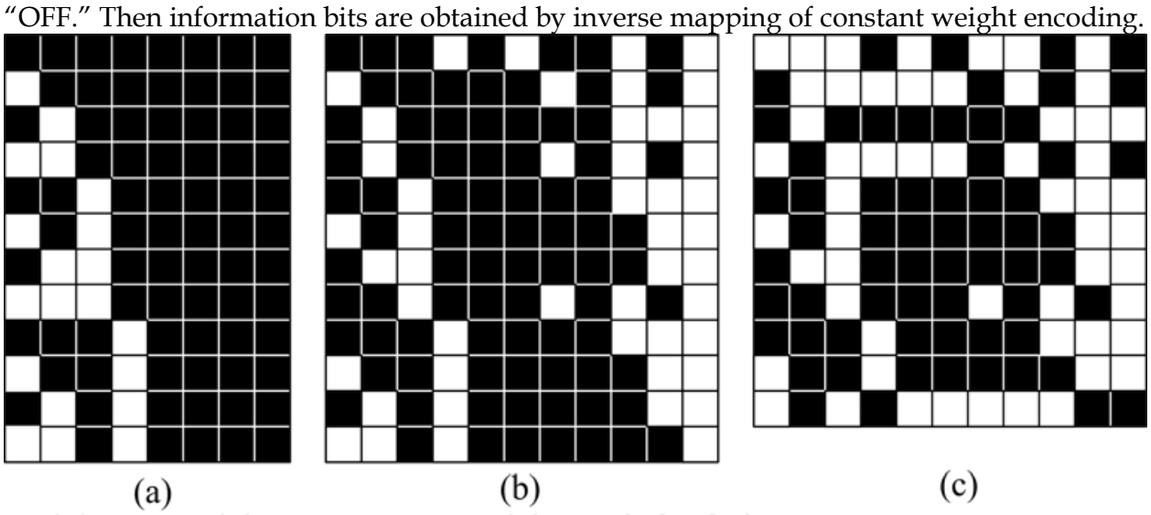


Fig. 7. Illustration of codeword complementing block modulation code encoding process (N=11). (a) Input information block, (b) constant weight encoded block and (c) selective complementary encoded block (Hwang et al., 2003).

4.4 Conservative Codes

Conservative codes are designed to satisfy the conservative constraint discussed previously. The coding scheme consists of a set of modulation blocks which is used as masks for the pixel-wise exclusive-OR operation with input information blocks. The encoded output blocks will be conservative of strength t, meaning that there exist no less than t transitions in each row and column. A method based on an error correcting code having minimum Hamming distance d and $d\geq 2t$-3 is proposed in (Vardy et al., 1996).

First define a mapping φ: $F_n \rightarrow E_n$ as follows:

$$\varphi(\mathbf{x}) = \mathbf{x} \oplus \sigma(\mathbf{x}) \tag{1}$$

where F_n is the set of all binary column vectors and E_n is the subset of F_n consisting of all even-weight vectors; σ(**x**) is the 1-bit cyclic downward shift of a column vector **x**. Given two sets of codewords, C_1 and C_2, produced by two (*t*-2)-error correcting codes of length *m* and *n*, respectively, two sets of column vectors $\{\mathbf{a}_1, \mathbf{a}_2,\ldots, \mathbf{a}_n\}$ and $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m\}$ are obtained by inverse-mapping of φ-1(C1) and φ-1(C2). A theorem in (Vardy et al., 1996) states that for any vector x in Fn, if $\mathbf{x} \oplus \mathbf{a}_i$ has less than t transitions, $\mathbf{x} \oplus \mathbf{a}_j$ will have has at least t transitions for all j≠i.

Next one constructs matrices A_1,A_2,A_n and B_1,B_2,B_m all with the size of $m \times n$, from the aforementioned sets of column vectors, $\{\mathbf{a}_0, \mathbf{a}_1,\ldots, \mathbf{a}_n\}$ and $\{\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_m\}$. $\mathbf{A}_i$ is formed by tiling $\mathbf{a}_i$ *n* times horizontally while B_j is formed by tiling transposed bj m times vertically.

$$\mathbf{A}_i = [\mathbf{a}_i \mid \mathbf{a}_i \mid \cdots \mid \mathbf{a}_i], \quad i = 1, ..., n \tag{2}$$

$$\mathbf{B}_j = \begin{bmatrix} \mathbf{b}_j^T \\ \hline \mathbf{b}_j^T \\ \hline \vdots \\ \hline \mathbf{b}_j^T \end{bmatrix}, \quad j = 1, ..., m \tag{3}$$

Finally, a modulation block is obtained by component-wise exclusive-OR of an $\mathbf{A}_i$ and a $\mathbf{B}_j$. On the total, there are $(m+1)\times(n+1)$ such modulation blocks. With such modulation block construction, it can be shown that pixel-wise exclusive-OR operation of an input information block with a modulation block yields an modulated block that satisfies the conservative constraint with strength *t*. To indicate which modulation block is applied, several bits are inserted as an extra row or column to each recorded page.

4.5 Block Code with 2-D Run-Length-Limited Constraint

In 2-D run-length limited block codes, only the minimum distance (d_{min}) between ON pixels is enforced. Codebook design involves an exhaustive search of all possible binary blocks and prunes any blocks containing two ON pixels with a square distance smaller than d_{min}. A mapping between segments of information bits and these code blocks is then developed.

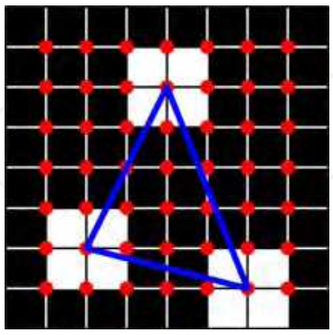


Fig. 8. An example of 8×8 block with three all-ON sub-blocks (Malki et al., 2008).

In (Malki et al., 2008), a fixed number of all-ON sub-blocks are present in each code block. This scheme lets the proposed code comply with the constant-weight constraint automatically. An example of three 2×2 all-ON sub-blocks is shown in Fig. 8. For an 8×8 block without any constraints, there are (8-1)×(8-1)=49 possible positions for each sub-block,

denoted by filled circles. The value of d_{min} determines the available number of legal code blocks.

5. Strip Codes for Holographic Data Storage

In contrast to block codes, which create modulated pages by tiling code blocks independently, strip codes exploit a finite state machine to encode information streams taking adjacent blocks into consideration. The name of "strip" comes from dividing a whole page into several non-overlapping strips and encoding/decoding within each strip independently. To decode such modulation codes, a *sliding block decoder* (Coufal et al., 2000) is required. Such a decoder has a decoding window containing preceding (memory), current and subsequent (anticipation) blocks.

Strip codes provide an advantage of ensuring that the modulated patterns across borders of blocks can still comply with the prescribed constraints if the coding scheme is well designed. However, additional constraints for patterns across borders of strips are still required. These additional constraints are called *strip constraints* (Ashley & Marcus, 1998; Coufal et al., 2000). Another advantage of strip codes is their better error-rate performance due to the fact that adjacent modulated blocks are used to help decode the current block.

In terms of the decoding procedure, strip codes need decoders with higher complexity when compared with block codes. Another concern for strip codes is that using a finite state machine as the encoder may fail to achieve the 2-D capacity C (Ashley & Marcus, 1998) defined as

$$C = \lim_{h,w\to\infty} \frac{\log_2\left(N(h,w)\right)}{hw}, \tag{4}$$

where $N(h,w)$ is the number of legal code blocks on an $h\times w$ block.

Strip codes solve problems of violating constraints across block borders by considering adjacent modulated blocks during encoding/decoding operation of the current block. As such, better conformance to the constraints and BER performance are attained at the cost of higher encoding/decoding complexity.

5.1 8:12 Balanced Strip Code

The 8:12 strip modulation code proposed in (Burr et al.,1997) is a typical balanced strip code. A page is first divided into non-overlapping strips with height of two pixels. Balanced code blocks with size of 2×6 are prepared in advance with minimum Hamming distance of four. During encoding, a finite state machine with four states is used to map 8-bit information into a 2×6 block based on the current state and input information bits, achieving a code rate of 2/3. The process goes on until a complete modulated page is obtained. As for decoding, the Viterbi decoder can be adopted for better performance. The 8:12 modulation code is block-decodable and can be decoded using the correlation detection. However, the error rate performance will not be as good as those using Viterbi decoders.

5.2 9:12 Pseudo-Balanced Strip Code

The 9:12 pseudo-balanced code (Hwang et al., 2002) is a strip code which maps 9-bit information into a 12-bit codeword with a code rate of ¾. It is similar to the 8:12 balanced

modulation code, but with higher code rate by relaxing the balanced constraint to allow certain degree of codeword imbalance. The 9:12 pseudo-balanced code maps information bits to codewords according to a trellis with eight diverging and eight merging states. Each branch represents a set of 64 pseudo-balanced 12-bit codewords with minimum Hamming distance of four. These codewords are divided into 16 groups. The most-significant three information bits of the current data word and the most-significant three bits of the previous data word indicate the states in the previous stage and the current stage, respectively. The codeword set is determined uniquely by these two states. The least-significant six bits of the current information word are used to select a codeword within the set.

The decoding is similar to Viterbi decoder for trellis codes (Proakis, 2001). First, the best codeword with respect to the acquired codeword for each codeword set is found. The index of that best codeword and the associate distance are recorded. Then the Viterbi algorithm finds the best state sequence using those aforementioned distances as branch metrics. In (Hwang et al., 2002), it is shown that the 9:12 pseudo-balanced code, though with a higher code rate, provides similar performance as the 8:12 balanced strip code.

5.3 2-D Low-Pass Codes

Previously, several low-pass constraints which prohibit some particular patterns with high spatial frequency components are introduced. Strip codes are better suited for compliance with these constraints since they can control patterns across block borders more easily.

Current state	Input information bits															
	000		001		010		011		100		101		110		111	
1	00 00	1	10 10	1	01 01	2	11 11	2	11 01	2	01 11	2	00 11	3	11 00	4
2	00 00	1	10 10	1	01 01	2	11 11	2	00 10	1	10 00	1	00 11	3	11 00	4
3	00 00	1	10 10	1	01 01	2	11 11	2	00 10	1	01 11	2	00 11	3	11 00	4
4	00 00	1	10 10	1	01 01	2	11 11	2	11 01	2	10 00	1	00 11	3	11 00	4

Table 2. Encoder for the third constraint in Table 1 (Ashley & Marcus, 1998).

Code block	00 00	10 10	01 01	11 11	00 10	11 01	10 00	01 11	00 11	11 00
Information bits	000	001	010	011	100	100	101	101	110	111

Table 3. Decoder for the third constraint in Table 1 (Ashley & Marcus, 1998).

Besides, the high-frequency patterns may not only appear within strips but also across strips. *Strip constraint* is therefore very important when we require a low-pass property over the whole modulated page. Below is an example of the strip constraint applying to a strip code that satisfies the low-pass constraint. Given a forbidden pattern in Table 1 with height L, a strip constraint banning the top/bottom m top/bottom rows of this forbidden pattern to appear in the top/bottom of the current strip, where m is between $L/2$ and L. With this extra strip constraint, it is guaranteed that the low-pass constraint will be satisfied across strip boundary. Table 2 and Table 3 are an example of encoder/decoder designed for the third

constraint in Table 1. Using a finite state machine with four states, the encoder transforms three information bits into a 2×2 code block based on the current state. The code block and the next state are designated in the Table 2. The three information bits are easily decoded from the single retrieved block. The mapping manner is shown in Table 3.

6. Variable-Length Modulation Codes for Holographic Data Storage

In Sections 4 and 5, we have introduced block codes and strip codes for holographic data storage, respectively. The code rates of the modulation codes we discussed are generally fixed. This kind of modulation codes is called *fixed-length* since they map a fixed number of information bits into a fixed-size code block. To allow for more flexibility in modulation codes, *variable-length* modulation codes were proposed. With such modulation codes, a variable number of information bits are mapped into a variable-size code block. Despite its advantage of more freedom in choosing legal code blocks that comply with certain constraints, variable-length modulation codes need a more elaborate decoding scheme to correctly extract code blocks without causing error propagation.

A variable-length modulation code designed to obey the low-pass constraint is proposed in (Pansatiankul & Sawchuk, 2003). It eliminates patterns which contain high spatial frequency not only within blocks but also across block borders. With this strict constraint and a given code rate, variable-length modulation codes produce better modulated pages than fixed-length modulation codes. Instead of preventing forbidden patterns such as those listed in Table 1, (Pansatiankul & Sawchuk, 2003) sets maximum numbers of ON pixels surrounding an OFF pixel. The notation of this variable-length modulation code is ([m_1, m_2, …], [n_1, n_2, …]; [k_1, k_2, …]; a, β), where m_i and n_i are the size of 2-D blocks, k_i is the length of one-dimensional input information bits, a is a maximum number of ON pixels in the four connected neighbor positions surrounding an OFF pixel and β *is* another maximum number of ON pixels in the rest four of eight nearest neighbor positions surrounding an OFF pixel. Parts of the modulated page produced by (4,[1,2,3];[2,4,6];2,3) and (1,[4,6,8]; [2,3,4];3,0) variable-length modulation codes are given in Fig. 9. Rectangles indicate 3×3 blocks with highest allowable number of ON-pixels around an OFF-pixel according to the respective (a, β) constraints.

The encoding process is the same as those in conventional block codes, giving a one-to-one manner between information bits and code blocks. In contrast to the straightforward encoding scheme, the decoding has the challenge of correctly determining the size of the current code block to do inverse-mapping. A decoding scheme for the (4,[1,2,3];[2,4,6];2,3) variable-length modulation code is provided in (Pansatiankul & Sawchuk, 2003). We first grab a 4×1 retrieved block and compare it to all 4×1 code blocks corresponding to 2-bit information. Note that in (4,[1,2,3];[2,4,6];2,3) variable-length modulation code, all 4×2 and 4×3 coded patterns are designed to have an all-OFF last column. If we cannot find a 4×1 code block that is close to the retrieved block and the next column contains only OFF pixels, we enlarge the retrieved block size from 4×1 to 4×2 and compare it to all 4×2 code blocks corresponding to 4-bit information. If still we cannot find any good 4×2 code block, then an error is declared. Similarly, we can check the second column to the right of the current 4×2 retrieved block for all OFF pixels and enlarge the retrieved block to 4×3 if the check turns out positive. Then compare the extended retrieved block to all 4×3 code blocks

corresponding to 6-bit information. Similarly, an error is declared when no code block matches the retrieved block.

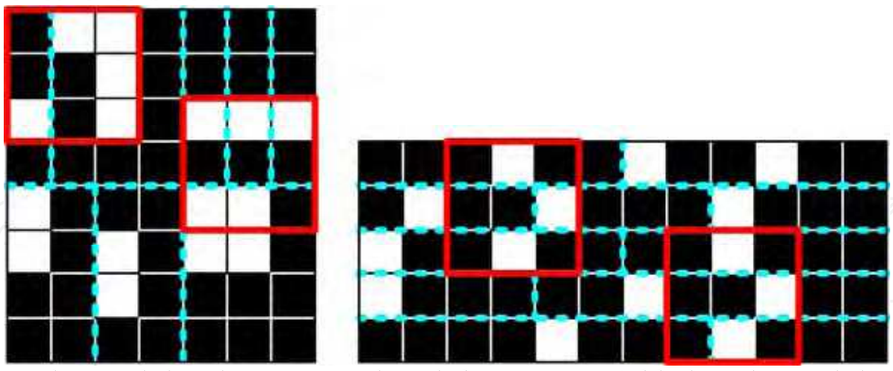


Fig. 9. Sample modulated page produced by two variable-length modulation codes (Pansatiankul & Sawchuk, 2003).

7. Conclusion

In this chapter, modulation codes for optical data storage have been discussed. At first, four types of constraints are introduced, including run-length limited, conservative, low-pass and constant-weight constraints. Since high spatial frequency components tend to be attenuated during recording/reading procedures and long runs of OFF pixels increase difficulty in tracking, the former three types of constraints are proposed to avoid these adverse situations as much as possible. On the other hand, the constant-weight constraint gives modulated pages that are easier to decode. In addition, experiments indicate that better performance can be obtained for modulation codes that have sparse weight.

Based on the constraints, several modulation codes are discussed. The one-dimensional modulation codes adopted in current optical storage systems, i.e., EFM for CD, EFMPlus for DVD and 17PP for Blu-ray disc, are first introduced. All of these modulation codes are developed for the run-length limited constraint.

Next, we focus on 2-D modulation codes for holographic data storage systems. They are classified into block codes and strip codes. Information bits and code blocks have a one-to-one relationship in block codes whose encoder/decoder can be simply realized by look-up tables. However, block codes cannot guarantee that patterns across block borders comply with the required constraints. This shortcoming can be circumvented by strip codes, which produce code blocks based on not only the input information bits but also neighboring modulated blocks. A finite state machine and a Viterbi decoder are typical schemes for the encoding and decoding of the strip codes, respectively.

Variable-length modulation codes, in contrast to fixed-length modulation codes, do not fix the number of input information bits or the code block size. The relaxed design increases the number of legal patterns and provides better performance than the fixed-length modulation codes with the same code rate. However, error propagation problems necessitate a more elaborated decoder scheme.

Finally, comparisons among different types of modulation codes introduced in this chapter are listed in Table 4 and Table 5.

	Block code	Strip codes
One-to-one manner between information bits and code blocks	Yes	No
Constraint satisfaction across block borders	Difficult	Simple
Encoding/decoding complexity	Low	High

Table 4. Comparison between block codes and strip codes.

	Fixed-length	Variable-length
Freedom of choosing legal pattern	Low	High
Error propagation problem during decoding	No	Yes
BER performance with the same code rate	Poor	Good
Code rate with the same BER performance	Low	High

Table 5. Comparison between fixed-length and variable-length modulation codes

8. References

Ashley, J. J. & Marcus, B. H. (1998). Two-dimensional low-pass filtering codes. *IEEE Trans. on Comm.*, Vol. 46, No. 6, pp. 724-727, ISSN 0090-6778

Blaum, M.; Siegel, P. H., Sincerbox, G. T., & Vardy, A. (1996). Method and apparatus for modulation of multi-dimensional data in holographic storage. *US patent* (Apr. 1996) 5,510,912

Blu-ray Disc Association (2006). *White paper Blu-ray disc format: 1. A physical format specifications for BD-RE*, 2nd edition

Burr, G. W.; Ashley, J., Coufal, H., Grygier, R. K., Hoffnagle, J. A., Jefferson, C. M., & Marcus,B. (1997). Modulation coding for pixel-matched holographic data storage. *Optics Letters*, Vol. 22, No. 9, pp. 639-641, ISSN 0146-9592

Burr, G. W. & Marcus, B. (1999). Coding tradeoffs for high density holographic data storage. *Proceedings of the SPIE*, Vol. 3802, pp. 18–29, ISSN 0277-786X

Coufal, H. J.; Psaltis, D. & Sincerbox, G. T. (Eds.). (2000). *Holographic data storage,* Springer-Verlag, ISBN 3-540-66691-5, New York

Chen, C.-Y. & Chiueh, T.-D. A low-complexity high-performance modulation code for holographic data storage, *Proceedings of IEEE International Conference on Electronics, Circuits and Systems*, pp. 788-791, ISBN 978-1-4244-1377-5, Dec. 2007, Morocco

Daiber, A. J.; McLeod, R. R. & Snyder, R. (2003). Sparse modulation codes for holographic data storage. *US patent* (Apr. 2003) 6,549,664 B1

Hwang, E.; Kim, K., Kim, J., Park, J. & Jung, H. (2002). A new efficient error correctible modulation code for holographic data storage. *Jpn. J. Appl. Phys.*, Vol. 41, No. 3B, pp. 1763-1766, ISSN 0021-4922

Hwang, E.; Roh, J., Kim, J., Cho, J., Park, J. & Jung, H. (2003).A new two-dimensional pseudo-random modulation code for holographic data storage. *Jpn. J. Appl. Phys.*, Vol. 42, No. 2B, pp. 1010-1013, ISSN 0021-4922

Immink, K. A.; Nijboer, J. G., Ogawa, H. & Odaka, K. (1985). Method of coding binary data. *U.S. Patent* (Feb. 1985) 4,501,000

Immink, K. A. (1997). Method of converting a series of m-bit information words to a modulated signal, method of producing a record carrier, coding device, decoding

device, recording device, reading device, signal, as well as record carrier. *U.S. Patent* (Dec. 1997) 5,696,505

Kamabe, H. (2007). Representation of 2 dimensional RLL constraint, *Proceedings of IEEE International Symposium on Information Theory (ISIT),* pp. 1171-1175, ISBN 978-1-4244-1397-3, Jun. 2007, Nice

King, B. M. & Neifeld, M. A. (2000). Sparse modulation coding for increased capacity in volume holographic storage. *Applied Optics*, Vol. 39, No. 35, pp. 6681-6688, ISSN 0003-6935

Kume, T.; Yagi, S., Imai, T. & Yamamoto, M. (2001). Digital holographic memory using two-dimensional modulation code. *Jpn. J. Appl. Phys.*, Vol. 40, No. 3B, pp. 1732-1736, ISSN 0021-4922

Malki, O.; Knittel, J., Przygodda, F., Trautner, H. & Richter, H. (2008). Two-dimensional modulation for holographic data storage systems. *Jpn. J. Appl. Phys.*, Vol. 47, No. 7, pp. 5993-5996, ISSN 0021-4922

McLaughlin, S. W. (1998). Shedding light on the future of SP for optical recording. *IEEE Signal Processing Magazine*, Vol. 15, No. 4, pp. 83-94, ISSN 1053-5888

Pansatiankul, D. E. & Sawchuk, A. A. (2003). Variable-length two-dimensional modulation coding for imaging page-oriented optical data storage systems. *Applied Optics*, Vol. 42, No. 26, pp. 5319-5333, ISSN 0003-6935

Proakis, J. G. (2001) *Digital Communications*, McGraw-Hill Science/Engineering/Math; 4th edition, ISBN 0-07-232111-3, International

Roth, R. M.; Siegel, P. H. & Wolf, J. K. (2001). Efficient coding schemes for the hard-square model. *IEEE Trans. on Info. Theory*, Vol. 47, No. 3, pp. 1166-1176, ISSN 0018-9448

Vadde, V. & Vijaya Kumar, B.V.K. (2000). Performance comparison of equalization and low-pass coding for holographic storage, *Conference Digest Optical Data Storage*, pp. 113-115, ISBN 0-7803-5950-X, May 2000

Vardy, A.; Blaum, M., Siegel, P. H., & Sincerbox, G. T. (1996). Conservative arrays: multidimensional modulation codes for holographic recording. *IEEE Trans. on Info. Theory*, Vol. 42, No. 1, pp. 227-230, ISSN 0018-9448

IntechOpen

Data Storage

Edited by Florin Balasa

ISBN 978-953-307-063-6

Hard cover, 226 pages

Publisher InTech

Published online 01, April, 2010

Published in print edition April, 2010

The book presents several advances in different research areas related to data storage, from the design of a hierarchical memory subsystem in embedded signal processing systems for data-intensive applications, through data representation in flash memories, data recording and retrieval in conventional optical data storage systems and the more recent holographic systems, to applications in medicine requiring massive image databases.

InTech Europe

University Campus STeP Ri

Slavka Krautzeka 83/A

51000 Rijeka, Croatia

Phone: +385 (51) 770 447

Fax: +385 (51) 686 166

www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai

No.65, Yan An Road (West), Shanghai, 200040, China

中国上海市延安西路65号上海国际贵都大饭店办公楼405单元

Phone: +86-21-62489820

Fax: +86-21-62489821

IntechOpen