Journal of Zhejiang University-SCIENCE C (Computers & Electronics)
ISSN 1869-1951 (Print); ISSN 1869-196X (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

A two-dimensional constant-weight sparse modulation code for volume holographic data storage*

Hua-rong GU†, Liang-cai CAO, Qing-sheng HE, Guo-fan JIN
(*State Key Laboratory of Precision Measurement Technology and Instruments, Tsinghua University, Beijing 100084, China*)
†E-mail: ghr@mail.tsinghua.edu.cn

Received May 27, 2010; Revision accepted Sept. 3, 2010; Crosschecked Oct. 12, 2010; Published online Oct. 27, 2010

Abstract: In order to simplify the threshold determination, reduce the inter-pixel cross-talk, and improve the storage density for high-density volume holographic data storage, a two-dimensional constant-weight sparse modulation code is proposed. The evaluation criteria and design rules are investigated based on the page-oriented optical data storage system. Coding parameters are optimized to achieve large channel capacities. An 8:16 modulation code is designed to reduce the raw bit error rate and its performances are experimentally evaluated. A raw bit error rate of the magnitude of 10^{-4} is obtained with a single-data-page storage and 10^{-3} with multiplexing.

Key words: Volume holographic data storage, Modulation codes, Sparse codes, Raw bit error rate, Channel capacity
doi:10.1631/jzus.C1010246 **Document code:** A **CLC number:** O438.1

1 Introduction

Modulation codes are widely used in data storage, such as magnetic storage (Davey *et al.*, 1994) and optical storage (Song *et al.*, 2006). The purpose of a modulation code is to constrain the signal patterns in the time domain or space domain, to forbid certain patterns from appearing that are more likely to be corrupted by the channel, and to demand other patterns that are helpful for a given detection scheme (Coufal *et al.*, 2000). A typical example of a modulation code is the eight-to-fourteen modulation (EFM) used by compact discs (CD) (Immink, 2004).

In order to reduce the inter-pixel cross-talk and improve detection accuracy, modulation codes are usually introduced in volume holographic data storage (VHDS) (Burr *et al.*, 1997; Kim and Lee, 2009). Unlike the modulation codes for traditional data storage which relate to a one-dimensional (1D) channel, the modulation codes for VHDS are two dimensional (2D) which relate to parallel channels. According to the number of gray levels of each pixel, modulation codes are divided into binary modulation codes (Burr *et al.*, 1997) and gray-level modulation codes (Heanue *et al.*, 1995). The former are commonly adopted to simplify the coding. Depending on whether or not the ratio of the white pixels is constant, binary modulation codes are further divided into constant-weight modulation codes and variable-weight modulation codes (Hwang *et al.*, 2002; Chen and Chiueh, 2007). For uniform object beam and simple threshold determination, constant-weight modulation codes are often used in VHDS. This paper focuses on constant-weight modulation codes.

Assuming the codeword is a rectangular block containing q ($=m\times n$) pixels where t pixels are ON, the codeword can represent C_q^t ($=q!/((q-t)!t!)$) different signals, whose information capacity is $p=\log_2\mathrm{C}_q^t$. The coding rate is defined as $r=p/q$, and the code weight is defined as $w=t/q$. The codes are known as

---

[*] Project supported by the National Basic Research Program of China (No. 2009CB724007), the National High-Tech R & D Program (863) of China (No. 2009AA01Z112), and the National Natural Science Foundation of China (No. 60807005)

balanced modulation codes when *w* is equal to 0.5, and the codes are known as sparse modulation codes when *w* is less than 0.5. The codes with *w* larger than 0.5 are less used. In general, a constant-weight modulation code is denoted by its coding rate (*p*:*q*).

Among constant-weight codes, balanced modulation codes, such as the 6:8 modulation code, have the maximal coding rate (Burr *et al.*, 1998). However, sparse modulation codes may achieve higher storage density (King and Neifeld, 2000), since the storage density is affected not only by the single page capacity but also by the number of pages that can be multiplexed in the same volume. In this paper, the evaluation criteria and design rules are investigated based on the page-oriented optical data storage system. Coding parameters are optimized to achieve large channel capacities. An 8:16 2D constant-weight sparse modulation code is designed to reduce the raw bit error rate and its performances are experimentally evaluated.

2 Evaluation criteria

As a kind of channel code, modulation codes can be evaluated by the common criteria used in communication theory, including:

1. Bit error rate (BER). With respect to modulation codes, the BER criterion refers in particular to raw bit error rate, which is defined as the BER of the data before error correction decoding. The purpose of modulation codes is to reduce the raw BER of the channel. As a result, the raw BER after modulation decoding is the primary index of performance under the same conditions.

2. Coding rate. With the same size of data pages, the raw capacity of a single page is determined by the coding rate. Theoretically, a lower coding rate indicates a stronger correction capability, but usually a smaller storage capacity. A trade-off should be made between BER and storage capacity. The coding rate of a modulation code should be appropriately chosen so that the storage density approaches the maximum and the channel capacity is close to the Shannon limit (Shannon, 1948).

3. Complexity. The time costs of modulation encoding and decoding are nonignorable. They should be considered for various modulation schemes. The complexity of encoding and decoding should be reduced so that the modulation code can be easily implemented by software or hardware.

3 Coding rules

Unlike 1D modulation codes, modulation codes for VHDS need to map 1D data streams onto 2D data pages, while meeting the specific limiting conditions in a data page. The state splitting algorithm (Immink *et al.*, 1998) may be employed to design the 2D modulation codes. However, due to the large amount of allowable codewords, the state splitting process and combining process in the code design become too complicated, especially during the design of a high-efficiency 2D modulation code. For simplicity and practicability, a forbidden codeword rejection method is used as follows.

1. To avoid difficulties in threshold determination during the detection stage, and to smooth the object beam, constant-weight modulation codes are used in this study. The ratio of ON pixels to OFF pixels is constant (Fig. 1).

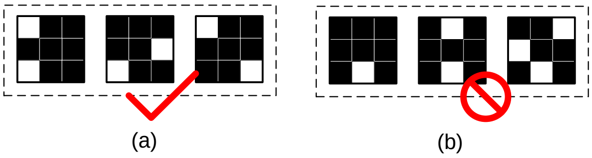

Fig. 1 Basic rules of a constant-weight modulation code
Each 3×3 block represents a codeword, where white squares denote ‘1’ (ON) and black squares denote ‘0’ (OFF). (a) Codewords with the same number of 1’s, allowed; (b) Codewords with different numbers of 1’s, forbidden

2. A variety of noises exist in the VHDS channel. The brightness noise of ‘0’ pixels might exceed that of ‘1’ pixels, which leads to an incorrect decision (Fig. 2).

To solve this problem, another rule is applied and the codewords shown in Fig. 3a are forbidden. According to this rule, Fig. 2b is correctly detected as the codeword represented by Fig. 2a, not Fig. 2c.

3. ‘1’ pixels around ‘0’ pixels should be no more than three to reduce the inter-pixel cross-talk (Fig. 3b). This rule comes from the low-pass feature of VHDS (Ashley and Marcus, 1998). If a codeword depicted in Fig. 3b is stored, then the ‘0’ pixels might be lightened

by the surrounding ‘1’ pixels and a detection error might occur in the retrieved image.

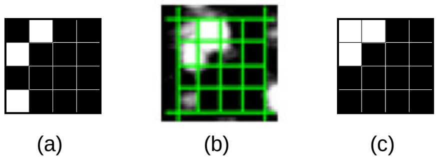

Fig. 2 Ambiguous decision in modulation codes
(a) Recorded codeword; (b) Image acquired by the image sensor; (c) Codeword after binary decision (the brightness of each pixel is sorted in descending order and the first three pixels are marked as 1’s). This result is unexpected

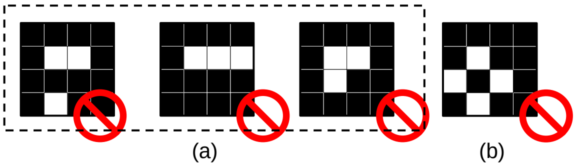

Fig. 3 Forbidden codewords in modulation coding
(a) Adjacent pixels should not be 1’s at the same time; (b) Codeword that has four ‘1’ pixels around ‘0’ pixels is forbidden

4 Coding parameters

After the establishment of the coding rules, the coding parameters are chosen to optimize the raw BER, coding rate, and complexity. There are two main parameters for constant-weight modulation codes: code weight w and the size of the coding block n (the coding block is assumed to be square, $m=n$). First, the choice of code weight w is discussed.

The diffraction efficiency η is in inverse proportion to the square of the number of data pages M_{page} stored in a VHDS system (Mok *et al.*, 1996):

$$\eta \propto \left(\frac{\text{M}/\#}{M_{\text{page}}}\right)^2, \tag{1}$$

where M/# is the dynamic range of the material. If divided equally by each ‘1’ pixel, the diffraction efficiency that each ‘1’ pixel could obtain is

$$\eta^* \propto \left(\frac{\text{M}/\#}{M_{\text{page}}}\right)^2 \frac{1}{t}. \tag{2}$$

The strength of the readout signal depends on the magnitude of the diffraction efficiency. Suppose that each pixel could be read out correctly only when its average diffraction efficiency is larger than a certain threshold η_{th}^*:

$$\eta_{\text{th}}^* \propto \left(\frac{\text{M}/\#}{M_{\text{page}}}\right)^2 \frac{1}{wq}. \tag{3}$$

For a specific image sensor, η_{th}^* should be a constant, and thus the number of data pages that can be stored in a VHDS system is

$$M_{\text{page}} \propto \frac{\text{M}/\#}{\sqrt{\eta_{\text{th}}^* wq}}. \tag{4}$$

The channel capacity could be expressed as

$$C(w) = M_{\text{page}} \cdot I_p(w) \cdot q, \tag{5}$$

where $I_p(w)$ is the information that each pixel could represent. According to the information theory (Shannon, 1948), it should meet the following equation:

$$I_p(w) \le R_0(w) = -w\log_2 w - (1-w)\log_2(1-w), \tag{6}$$

where $R_0(w)$ is the information entropy. Thus, the limit of the channel capacity could be expressed as

$$C_{\max}(w) \propto \frac{\sqrt{q}\cdot \text{M}/\#}{\sqrt{\eta_{\text{th}}^* w}}\left[-w\log_2 w - (1-w)\log_2(1-w)\right]. \tag{7}$$

With respect to balanced modulation codes (w=0.5), the limit of the channel capacity $C_{\max}(w)$ changed with the code weight w (Fig. 4). When the code weight is around 20%, the channel capacity reaches its maximum. It means that although spare modulation codes make the coding rate less than 1 and the capacity of a single data page is somewhat decreased, the number of multiplexed pages in the same volume will increase. In consideration of the dynamic range of the material for VHDS, the total capacity of the system will increase instead.

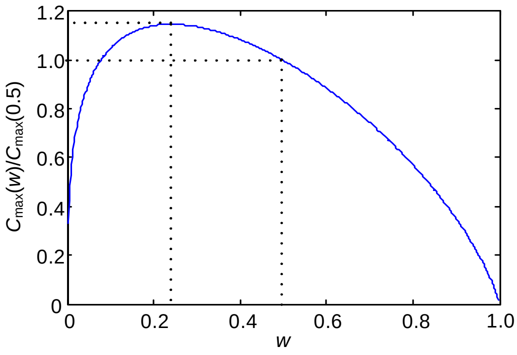

Fig. 4 The relationship between the limit of the channel capacity and code weight

In order to determine *n* and *w* under the current coding rules, the relative capacities of VHDS systems ($C(w)/C(0.5)$) with n=3–6 and various *w* are calculated here (Fig. 5). The corresponding code weight where the relative capacity reaches its maximum is smaller than that of the theoretical maximum capacity limit as shown in Fig. 4, because modulation codes have deleted partial codewords. The bigger the coding block, the larger the maximum value of relative capacity. However, too big a coding block leads to the increase of complexity and too small a coding block leads to the decrease of the coding rate. According to known research (King and Neifeld, 2000), the coding block size *n* is chosen ranging from 3 to 5, and n=4, w=3/16 are selected here (that is, 3 pixels are set as '1' pixels out of 16 pixels). It facilitates the encoding and decoding with a higher coding rate 8:16.

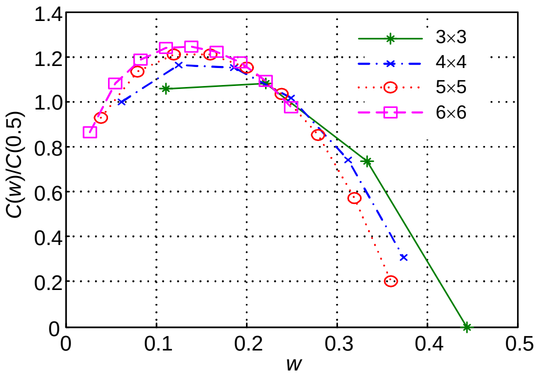

Fig. 5 Influences of block size and code weight on the relative capacity

The modulation code described above is the proposed '2D constant-weight sparse modulation code'. '2D' means that the codeword has a 2D structure. 'Constant-weight' indicates that the number of '1' pixels in the codeword is constant. 'Sparse' specifies that the code weight is less than 0.5.

5 Implementation and experiments

In order to verify the performances of the proposed sparse modulation codes, an experimental setup is built (Fig. 6). The light emitted by a diode pump laser (DPL) with a wavelength of 532 nm is reflected by a first polarization beam splitter (PBS) which filters out infrared light, and then the light is divided into two beams by a second PBS. Each beam is controlled by a shutter, followed by a spatial filter and beam expander. The object beam enters the polymer after modulation by a reflecting spatial light modulator (SLM) whose resolution is 1280×768. The reference beam enters the polymer after reflection by two mirrors. The polymer is inclined for substantially symmetrical incidence of the object beam and the reference beam, reducing the impact due to shrinkage. The resolution of CCD is 768×576.

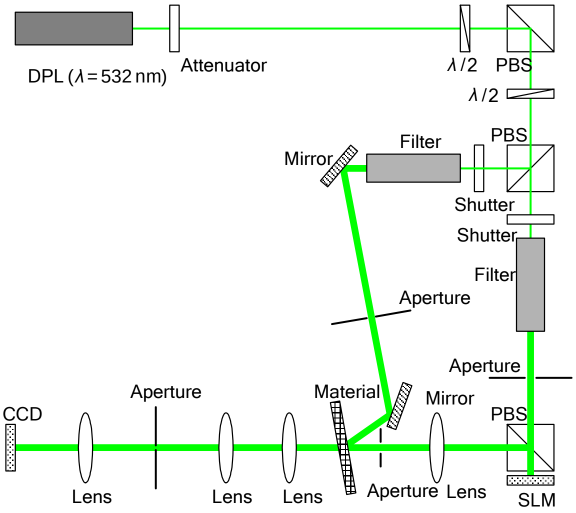

Fig. 6 Optical setup to verify the performances of modulation codes

DPL: diode pump laser; PBS: polarization beam splitter; SLM: spatial light modulator

Both the encoding and decoding of the 2D constant-weight sparse modulation code are created experimentally by software. The coding table (8 bits user data→4×4 codeword) is constructed as follows: (1) screen out 276 possible codewords according to coding rules in Section 3; (2) delete those codewords that have a Hamming distance of 2 from other codewords at most times (there are 4 in total); (3) recalculate the Hamming distances, and delete the codewords that have a Hamming distance of 2 from other codewords at most times again (there are 16 in total), and the number of the remaining codewords is 256. The encoding process could be completed by querying the coding table, and the decoding process could be accomplished by binary sorting and minimum distance decoding. Partial coding table of the 2D constant-weight sparse modulation code is shown in Fig. 7.

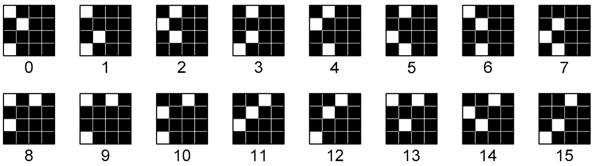

Fig. 7 Partial coding table of the two-dimensional constant-weight sparse modulation code

Encoding and decoding procedures are as shown in Fig. 8. Data is stored and recovered by encoding, storage, readout, division, and decoding.

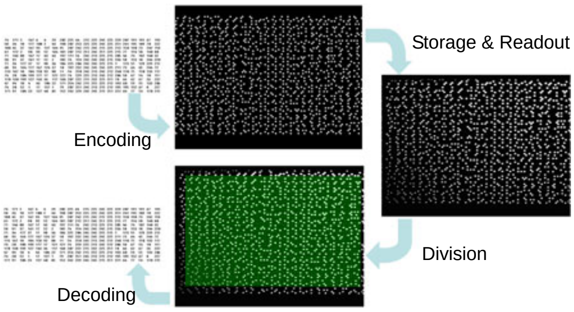

Fig. 8 Encoding and decoding procedures of the proposed modulation code

For single-page storage, the results are shown in Table 1. The signal-to-noise ratio (SNR) of the retrieved page is 4.03. If decoded as none-modulated data, it could represent 4928 bits data, with 69 bits error, and the raw BER is 0.014. If decoded as modulated data, 2464 bits data could be represented. Although the data capacity drops to one half, no error is detected, and the raw bit error is less than 4×10^{-4}.

Random speckle modulation is used in the reference path. Multiplexing is achieved through moving the storage material by computer-controlled translational stages. The multiplexing interval is 200 μm, and 20 data pages are multiplexed. The SNR and the raw BER with and without modulation codes of the 4th, 7th, 10th, 14th, and 17th data page are shown in Table 2. Generally speaking, the raw BER could be decreased by one order of magnitude using 2D constant-weight sparse modulation codes, and drop to 10^{-3} with the SNR around 3.8, which ensures subsequent interleaving and error correction.

Table 1 Performances of single page storage

SNR	Modulation	Capacity (bit)	Error (bit)	Raw BER
4.03	No	4928	69	0.014
	Yes	2464	0	$<4.0\times10^{-4}$

SNR: signal-to-noise ratio; BER: bit error rate

Table 2 Performances of multiplexing storage

No.	SNR	Modulation	Capacity (bit)	Error (bit)	Raw BER
4	3.82	No	3328	38	0.011
		Yes	1664	2	1.2×10^{-3}
7	3.78	No	3328	46	0.014
		Yes	1664	1	6.0×10^{-4}
10	3.88	No	3328	44	0.013
		Yes	1664	4	2.4×10^{-3}
14	3.70	No	3328	50	0.015
		Yes	1664	2	1.2×10^{-3}
17	3.62	No	3328	47	0.014
		Yes	1664	2	1.2×10^{-3}

SNR: signal-to-noise ratio; BER: bit error rate

6 Conclusions

In order to simplify the threshold determination, reduce the inter-pixel cross-talk, and decrease the raw BER, modulation codes are introduced in the channel processing of the volume holographic data storage. It is discovered from theoretical analysis that a larger capacity could be achieved using sparse modulation codes whose code weight is less than 0.5 when the dynamic range of the material is constant. Accordingly, a 2D constant-weight sparse modulation code is proposed. Corresponding coding rules are established, and proper coding parameters are chosen to maximize the channel capacity. An 8:16 modulation code is designed to reduce the raw BER and its performances are experimentally evaluated. A raw BER of the magnitude of 10^{-4} is obtained with a single-data-page storage and 10^{-3} with multiplexing.

References

Ashley, J.J., Marcus, B.H., 1998. Two-dimensional low-pass filtering codes. *IEEE Trans. Commun.*, **46**(6):724-727. [doi:10.1109/26.681399]

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. *Opt. Lett.*, **22**(9):639-641. [doi:10.1364/OL.22.000639]

Burr, G.W., Chou, W.C., Neifeld, M.A., Coufal, H., Hoffnagle, J.A., Jefferson, C.M., 1998. Experimental evaluation of user capacity in holographic data-storage systems. *Appl. Opt.*, **37**(23):5431-5443. [doi:10.1364/AO.37.005431]

Chen, C.Y., Chiueh, T.D., 2007. A Low-Complexity High-Performance Modulation Code for Holographic Data Storage. 14th IEEE Int. Conf. on Electronics, Circuits and Systems, p.788-791. [doi:10.1109/ICECS.2007.4511109]

Coufal, H.J., Psaltis, D., Sincerbox, G.T. (Eds.), 2000. Holographic Data Storage. Springer-Verlag, Berlin.

Davey, P.J., Donnelly, T., Mapps, D.J., 1994. Two dimensional coding for a multiple-track, maximum-likelihood digital magnetic storage system. *IEEE Trans. Magn.*, **30**(6): 4212-4214. [doi:10.1109/20.334038]

Heanue, J.F., Bashaw, M.C., Hesselink, L., 1995. Channel codes for digital holographic data storage. *J. Opt. Soc. Am. A*: *Opt. Image Sci. Vis.*, **12**(11):2432-2439. [doi:10.1364/JOSAA.12.002432]

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. Part 1-Regul. Pap. Short Notes Rev. Pap.*, **41**(3B):1763-1766.

Immink, K.A.S., 2004. Codes for Mass Data Storage Systems. Shannon Foundation Publishers, Eindhoven, the Netherlands.

Immink, K.A.S., Siegel, P.H., Wolf, J.K., 1998. Codes for digital recorders. *IEEE Trans. Inf. Theory*, **44**(6):2260-2299. [doi:10.1109/18.720539]

Kim, J.Y., Lee, J.J., 2009. Two-dimensional 5:8 modulation code for holographic data storage. *Jpn. J. Appl. Phys.*, **48**(3):03A031. [doi:10.1143/JJAP.48.03A031]

King, B.M., Neifeld, M.A., 2000. Sparse modulation coding for increased capacity in volume holographic storage. *Appl. Opt.*, **39**(35):6681-6688. [doi:10.1364/AO.39.006681]

Mok, F.H., Burr, G.W., Psaltis, D., 1996. System metric for holographic memory systems. *Opt. Lett.*, **21**(12):896-898. [doi:10.1364/OL.21.000896]

Shannon, C.E., 1948. A mathematical theory of communication. *Bell Syst. Tech. J.*, **27**:379-423, 623-656.

Song, J., Pei, J., Xu, D.Y., Xiong, J.P., Chen, K., Pan, L.F., 2006. Pit depth and width modulation multilevel run-length limited read-only optical storage. *Chin. Phys. Lett.*, **23**(6):1504-1506. [doi:10.1088/0256-307X/23/6/041]