Gray-scale data pages for digital holographic data storage

Geoffrey W. Burr, Gabriele Barking, Hans Coufal, John A. Hoffnagle, and C. Michael Jefferson

IBM Almaden Research Center, 650 Harry Road, San Jose, California 95120

Mark A. Neifeld

Department of Electrical and Computer Engineering, Optical Sciences Center, University of Arizona, Tucson, Arizona 85721

Received April 16, 1997

The prospects for gray-scale (or multilevel) digital holographic data storage are theoretically and experimentally investigated. A simple signal-to-noise ratio (SNR) partitioning argument shows that when SNR scales as 1 over the number of holograms squared, five gray levels ($\log_2$ 5 bits/pixel) would be expected to result in a 15% capacity increase over binary data pages. However, the additional signal-dependent noise sources present in practical systems create a baseline SNR that reduces both the optimal number of gray levels and the resulting gain in capacity. To implement gray-scale recording experimentally, we adapt the predistortion technique previously developed for binary page-oriented memories [Opt. Lett. **23,** 289 (1998)]. Several new block-based modulation codes for decoding gray-scale data pages are introduced. User capacity is evaluated by an experimental technique using $LiNbO_3$:Fe in the 90° geometry. Experimental results show that a balanced modulation code with three gray levels provides a 30% increase in capacity (as well as a 30% increase in readout rate) over local binary thresholding.

OCIS codes: 210.2860, 210.0210.

Digital holographic data storage is a topic of widespread research interest.[1–3] The ability to store multiple holograms within a small volume of storage material and to retrieve data pages with thousands of bits in parallel provides an attractive combination of density and speed. The tools of digital signal processing have been adapted for holographic storage,[2] and further developments in signal processing,[4] error-correction coding (ECC),[5] and modulation codes[6] have followed. Much of this work, however, used binary encoding, in which pixels encode two distinct states: ON (bright) and OFF (dark), corresponding to binary 1 and 0, respectively. Gray-scale encoding, in which pixels take one of g brightness levels, affects both capacity and readout rate, because each pixel now conveys $\log_2 g$ bits of data. Gray-scale encoding was briefly analyzed in a study of the effects of the Rician noise arising from optical scatter.[7] In this Letter we evaluate gray-scale recording for holographic storage systems.

Gray-scale encoding divides the available signal-to-noise ratio (SNR) into $g - 1$ parts. Because of the finite dynamic range of a holographic storage system, signal levels tend to decrease as 1 over the number of holograms squared. Given the presence of a constant noise floor and the need to stay above some minimum acceptable SNR, the subdivision of the SNR budget into gray levels means fewer stored holograms. Total capacity is the product of two factors that depend on g: bits per pixel ($\log_2 g$) and the number of holograms, which decreases as $1/\sqrt{g-1}$. The resulting relationship between capacity and number of gray levels is shown by the thick solid curve in Fig. 1. At the optimum value of $g = 5$, the capacity is increased by 15% over that derived from binary encoding, and the readout rate (user bits per second) is increased by a factor of $\log_2 5 \sim 2.32$.

In addition to a constant noise floor, however, practical systems exhibit signal-dependent noise sources, such as those associated with the optical imaging of the parallel data array. Since these noise sources increase with increasing signal, they tend to limit the SNR that can be achieved for even a few holograms. These sources also change the scaling between SNR and the number of holograms, favoring a smaller number of gray levels and reducing the gray-scale capacity gain, as shown by the dashed curves in Fig. 1, which are plotted for various ratios between initial SNR (the SNR with one stored hologram) and final SNR (the SNR at which no further holograms can be added). Typically, systems must have high-performance optics for the initial SNR to exceed the final SNR by more than ~15, even when SNR is measured locally (as is

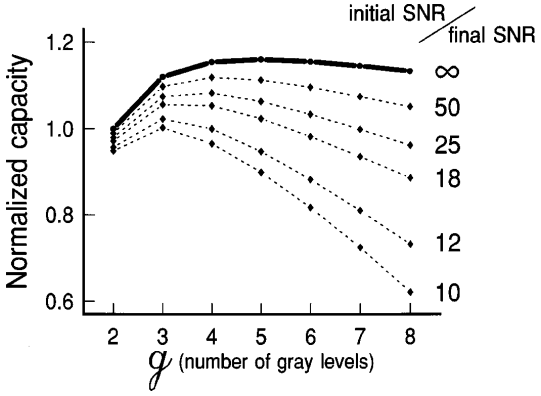

Fig. 1. Normalized capacity as a function of gray level (g). For the solid curve, SNR drops as 1 over the number of holograms squared. The dashed curves indicate the effect when signal-dependent noise sources increase in importance relative to the background noise floor, as quantified by the ratio between the initial and the final SNR.

0146-9592/98/151218-03$15.00/0

appropriate for any local decoding–decision method). Thus, Fig. 1 indicates that three gray levels ($g = 3$) will most likely be the optimal choice for a practical system. One could perform a more-elaborate analysis by considering the raw bit-error rate (BER) seen by the ECC or by optimizing either the level separations or the probability of occurrence of each gray level. However, Fig. 1 serves its purposes of justifying an experimental implementation of gray-level holographic storage and of emphasizing the importance of high initial SNR.

A recently developed technique for improving initial SNR is the predistortion technique.[8] By manipulation of each pixel's exposure (through recording time or pixel brightness), deterministic variations among pixels one can be suppressed (increasing the initial SNR). The appropriate recording exposure for each pixel is obtained by measurement of the ratio between the actual and the desired pixel brightness. Modifying the predistortion technique for gray scale is simply a matter of using a set of $g - 1$ desired pixel brightness levels. This modification was experimentally implemented on the IBM DEMON holographic storage platform. The presence of a small amount of intersymbol interference necessitated the use of several iterations, with practice holograms, for optimization of the individual pixel exposures. A histogram of a six-level hologram made possible by the predistortion technique is shown in Fig. 2. The SNR of this histogram was 6.39, where the SNR for g levels is defined as

$$\mathrm{SNR} = \frac{1}{g-1} \sum_{i=0}^{g-2} \frac{\mu_{i+1} - \mu_i}{\sqrt{{\sigma_{i+1}}^2 + {\sigma_i}^2}} \qquad (1)$$

and μ_i and ${\sigma_i}^2$ are the mean and the variance, respectively, of the ith gray level.

To obtain measurements of the raw BER of gray-level data pages with practical decoding methods, we developed a general local-thresholding method and several balanced modulation codes. The min–max local-thresholding method operates on a small block of pixels, sacrificing one pixel from each block to guarantee at least one occurrence of both the darkest (zero) and the brightest ($g - 1$) level. (Given the presence of ECC, the probability that both of these levels are absent is small enough to be ignored.) On decoding, the $g - 1$ thresholds are derived from the minimum and maximum values within the block and from look-up information about the level separations. We chose to encode over blocks of 4×4 pixels, to decode over blocks of 8×8 pixels, and to use uniformly spaced gray levels. We anticipate that the performance upon optimizing these choices will be slightly improved. On the other hand, we have not considered how to encode binary user data into gray-scale pixels efficiently, and we expect that this conversion process will reduce the effective code rate (except in the case of four gray levels).

In addition to this thresholding scheme, several multilevel block-based modulation codes were developed. For $g = 3$, a 15:12 code was implemented (15 bits of user data decoded from 12 detector pixels). The enforced constraint was that exactly 4 of these 12 pixels be of the 0 level, 4 be level 1, and 4 be 2 pixels. We implemented decoding by sorting the 12 received signal values. The complexity involved in even this three-level modulation code led us to implement the remaining gray-level codes by multiple application of binary modulation codes. For instance, a 24:16 decoder for four gray levels was implemented by application of a 6:8 decoder for binary data[6] twice. On the first pass, the 6:8 decoder operates on two discrete 8-pixel blocks—within each, the decoder identifies the four brightest pixels, which serves to classify the 0's and 1's separately from the 2's and 3's and produces 12 bits of user data. The eight 0 and 1 pixel values are then collected from across both blocks, creating a virtual 8-pixel block that is separated into 0 and 1 by a second pass through the 6:8 decoder. Decoding the 2 and 3 pixels similarly produces the final six user bits. For $g = 6$ gray levels, a 48:24 code was created by use of the 15:12 decoder (to divide the 24 pixels into three classes: 0/1, 2/3, and 4/5) followed by the 6:8 decoder (to complete the classification into six states). By omitting one of the virtual 6:8 blocks, we modified this code into a 42:24 code for $g = 5$. Note that since these codes are built to encode from binary data there is no loss of efficiency in encoding directly from user data. These modulation codes and the min–max thresholding method are summarized in Table 1.

The storage capacity with these decoding methods was evaluated with an experimental procedure for capacity estimation.[9] All experimental conditions were identical to those in Ref. 9. After performing the predistortion iteration procedure we recorded three holograms to obtain a reliable measure of raw BER and hologram strength versus recording time. Then a single hologram was measured periodically during erasure. All holograms were angularly separated by >20 Bragg nulls. The capacity-estimation procedure uses the measured raw BER and exposure data to derive a recording schedule[10] that strives to equalize raw BER instead of diffraction efficiency.[9] This experimental technique quantifies the relationship between

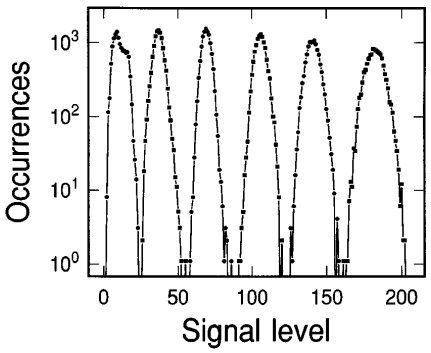

Fig. 2. Histogram of six-level hologram recorded in the IBM DEMON system by use of the predistortion technique.[8]

Table 1. Gray-Scale Decoders

	Code Rate (r_{detect}) for g Gray Levels				
Code	$g = 2$	$g = 3$	$g = 4$	$g = 5$	$g = 6$
Min–max	0.94	1.49	1.88	2.18	2.42
Modulation code	–	1.25	1.5	1.75	2.0

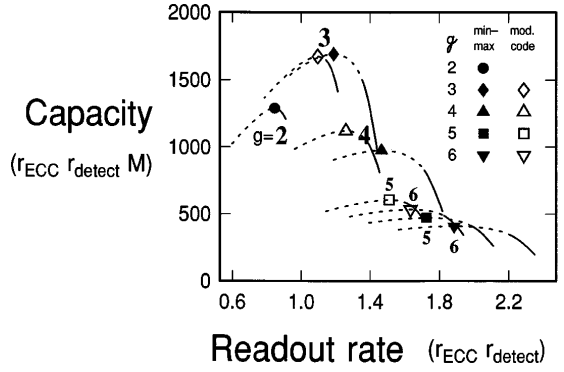

Fig. 3. Capacity ($\propto r_{\mathrm{detect}} r_{\mathrm{ECC}} M$) versus readout rate ($\propto r_{\mathrm{detect}} r_{\mathrm{ECC}}$) for modulation codes and the min–max thresholding method at varying numbers of gray levels g. Each curve indicates a range of choices of 8-bit/symbol Reed–Solomon codes, progressing left to right from stronger to weaker codes; the symbols indicate the maximum capacity points.

the number of holograms that can be stored, M, and the raw BER. In general, as the raw BER of the system increases, M increases slowly. For instance, with min–max decoding of binary data pages, each factor of 10 in raw BER allows only 5.5% more holograms to be stored.

To maintain a low-output user BER (say, 10^{-12}) as the raw BER operating point increases, however, one must increase the redundancy of the ECC code. So, although the number of holograms increases, the ECC code rate decreases. These two opposing trends create an optimal raw BER, at which the user capacity is maximized.[5] To compare the different decoders fairly, we took care to convert all measured error data into a consistent raw BER that is representative of the BER that would be seen by the ECC decoder. This conversion is necessary, for instance, when a modulation code outputs raw data in blocks (or symbols) of 6 bits while the ECC code corrects raw data by use of symbols of 8 bits.

Figure 3 shows a parametric plot of capacity versus readout rate for the gray-level modulation codes and the min–max algorithm, over the range $g = 2$–6. To concentrate on the relative comparison between gray levels and decoding methods, we did not scale by the number of pixels per page or the number of pages per second. Each data point indicates the maximal capacity point; the curve through it indicates the effects of varying the ECC code. The ECC for Fig. 3 utilized t-symbol-correcting Reed–Solomon codes with 255 eight-bit symbols, using $2t$ symbols of redundancy to achieve a corrected user BER specification of 10^{-12}. Each curve in Fig. 3 indicates the same range of ECC choices, from $t = 47$ (raw BER, $\sim 10^{-2}$) on the left to $t = 4$ (raw BER, $\sim 3 \times 10^{-5}$) on the right. Since distance along the curves is not linear in t, we indicate the point $t = 16$ (raw BER, $\sim 10^{-3}$) by changing the curve from dotted to solid. Roughly speaking, the solid curves indicate the range of high-speed ECC chips available off the shelf; the dotted curves indicate where the complexity of the ECC decoder may have a negative effect on throughput.

We have shown that gray-scale holographic data pages provide an advantage over binary encoding in both capacity and readout rate. As predicted by our simple SNR model, three gray levels produce the highest capacity in the presence of both signal-dependent noise and a constant noise floor. However, the SNR model is accurate only in indicating the general trends, making the experimental capacity-estimation procedure invaluable in quantifying the trade-off between higher bits per pixel and a smaller number of stored holograms. The experimental results in Fig. 3 show that use of three gray levels provides a 30% increase in both capacity and readout rate over conventional binary data pages. Implementing three gray levels with a 15-bit–12-pixel modulation code is convenient because it expressly encodes from and decodes back to binary data and requires ECC coding of only moderate complexity. The similar capacity performance obtained with a local-thresholding technique, however, indicates that the preference for three gray levels is independent of the particular choice of decoding scheme.

References

1. D. Psaltis and F. Mok, Sci. Am. **273**(5), 70 (1995).
2. J. F. Heanue, M. C. Bashaw, and L. Hesselink, Science **265,** 749 (1994).
3. J. H. Hong, I. McMichael, T. Y. Chang, W. Christian, and E. G. Paek, Opt. Eng. **34,** 2193 (1995).
4. J. F. Heanue, K. Gurkan, and L. Hesselink, Appl. Opt. **35,** 2431 (1996).
5. M. A. Neifeld and M. McDonald, Opt. Lett. **19,** 1483 (1994).
6. G. W. Burr, J. Ashley, H. Coufal, R. K. Grygier, J. Hoffnagle, C. M. Jefferson, and B. Marcus, Opt. Lett. **22,** 639 (1997).
7. J. F. Heanue, M. C. Bashaw, and L. Hesselink, J. Opt. Soc. Am. A **12,** 2432 (1995).
8. G. W. Burr, H. Coufal, J. A. Hoffnagle, and C. M. Jefferson, Opt. Lett. **23,** 289 (1998).
9. G. W. Burr, W.-C. Chou, M. A. Neifeld, H. Coufal, J. A. Hoffnagle, and C. M. Jefferson, "Experimental evaluation of user capacity in holographic data storage systems," Appl. Opt. (to be published).
10. D. Psaltis, D. Brady, and K. Wagner, Appl. Opt. **27,** 1752 (1988).