BURST ERROR COMPENSATION FOR A TWO-DIMENSIONAL CHANNEL

J. Q. Trelewicz

Telecommunications Research Center
Arizona State University
Mail Stop 7206
Tempe, Arizona 85287-7206
Email: jentre@asu.edu

Andrew J. Siefker

Department of Mathematics and Statistics
Murray State University
Post Office Box 9
Murray, Kentucky 42071-0009
Email: siefker@math.mursuky.edu

ABSTRACT

In an optical channel for a digital holographic data storage system, burst errors, in the form of severe amplitude compression, are experienced in the vicinity of dust and other optical aberrations in the channel. The locations of these burst errors are approximately fixed over a range of perturbations in magnification and vertical-horizontal position. In a calibrated channel, lower error rates can be realized by allocating energy in the modulation code based on a measurement of the approximate location of these effects. This paper discusses theoretical bounds on the energy inside the distortion region, and the modulation code which is used to approach this bound.

1. SYSTEM DESCRIPTION

The digital holographic data storage system (DHDSS) is similar in many ways to a communication channel, with the data bits experiencing interference within a page similar to the inter-symbol interference (ISI) seen in a conventional communications channel. However, the DHDSS experiences ISI in two dimensions, with a spatial coordinate instead of a time coordinate. Additionally, the interference is not spatially causal.

1.1. Operation of the System

A block diagram of the holographic data storage system is shown in figure 1. Polarization and other elements are omitted from the picture for simplicity.

Storage of data in the medium is performed as follows: the beam from the laser source is split into two separate beams. One beam is used to illuminate the Spatial Light Modulator (SLM), a 1024×1024 array of pixels on which the image to be stored is formed. The SLM is controlled by a host computer, which sets the states of the pixels to "on" (1) or "off" (0). The beam reflected from the SLM is focused into the storage medium, where it interferes with the reference beam. The interference of the beams creates fringes in the material (here, $LiNbO_3$). The angle of the reference beam, which is used for multiplexing the data, is controlled by a movable mirror.

The data is later read from the medium as follows: only the reference beam is present (the object beam is blocked). If the reference beam is reflected at the same angle as a page of stored data, the fringes reconstruct the stored data, which appears at the charge-coupled device (CCD) camera for recording.

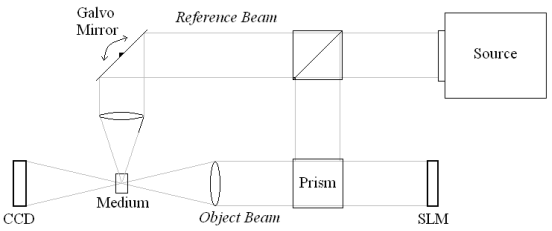

Figure 1: Block Diagram of Digital Holographic Data Storage System

1.2. Channel Characteristics

For the purposes of this study, the system is aligned so that the object beam is centered on the SLM and on the CCD, and the reference beam is approximately uniform across its region of illumination.

The following types of channel perturbation and noise are addressed with the coding:

Experimental Data were obtained with the assistance of Optitek, Inc. All work was performed under NASA GSRP Fellowship NGT 5-22. Our special thanks to Dr. Hosae Lee of Murray State University whose computer programming expertise was invaluable to the timely completion of this paper.

- Optical aberrations on the system components cause fixed "burst" errors in the image plane, with the corresponding diffraction rings causing compression in the gain of the pixel values. A typical example of these rings is shown in figure 2. The localized amplitude change can be approximated by the main lobe of the first order, circularly symmetrical Bessel function $J_0(x, y)$.
- Fixed errors are caused by stuck pixels in the SLM and/or the CCD. These spots can appear as either bright or dark values. Stuck pixels are also seen in figure 2 as a small, very dark region.
- Distortion is caused by the finite aperture of the lens, resulting in a non-uniform focal surface in the image plane. The result, in the alignment used with this system, is a ring of distorted pixels near the inflection points of the Gaussian focal surface shape. This distortion complicates symbol detection.
- Background noise is caused by the non-zero reflectivity of the "off" pixels in the SLM, by stray light from outside the system, and from scattering inside the system. The result is a noise floor at approximately 5% of the maximum pixel value. This noise is reasonably approximated by additive white Gaussian noise (AWGN).

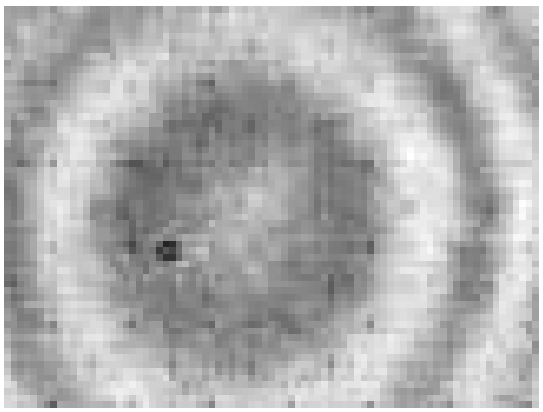

Figure 2: Typical Diffraction Pattern Burst Error

Unlike the conventional communications channel, the DHDSS channel cannot accurately be modeled as AWGN since the effects of the burst errors and distortion far outweigh the effects of the background noise.

1.3. Calibrated Channel

The precise amplitude compression in the burst errors cannot be predicted deterministically in the channel; instead, the burst error locations have some probability distribution because of the small changes in the alignment that occur over time.

The storage crystal and optics are considered to be a "fixed" system; i.e., the crystal will be used with the optics without removing the crystal (like a hard disk drive, where the platters are not removed from the heads). In this case, the system can be calibrated for the areas of highest distortion and other fixed phenomena (e.g., aberrations in the optics). The information from calibration is used to allocate the storage capacity in the image plane, and all data are stored encoded for this capacity distribution.

1.4. Admissible Modulation Codes

Since the most severe amplitude compression from the diffraction rings of aberrations occurs for aberrations close to the SLM and CCD, and since the "main lobe" of the rings measures approximately 40 to 60 pixels in diameter, a diffraction ring main lobe measuring 49 pixels in diameter is considered for analysis. Call this circular region $\mathcal{R}$.

For synchronization of the detector, the symbols are pixel run-length limited (RLL) to four off pixels. Near the center of $\mathcal{R}$, low-energy codes that are RLL to 1 on pixel work well for reducing ISI. Near the edges (outside half the region radius), 2-3 on pixels are required to produce enough energy after the point-spread function (PSF) and the intensity attenuation so that the pixel values do not drop below the noise floor. These restrictions have been found by experimentation to facilitate acceptable performance in the detector. Call the class of all such modulation codes $\mathcal{C}$.

In $\mathcal{R}$, the maximum energy after modulation coding could be bound by the signal which uses the "all-one" signal, Z. However, this bound is not reasonable since Z contains no information. On the other hand, it can be shown by exhaustive search that more than 99.8% of the members of $\mathcal{C}$ correspond to non-negative trigonometric polynomials, as described below. A similar result holds for main lobes of different diameters. Thus, a better energy bound for $\mathcal{C}$ is found by analyzing these non-negative trigonometric polynomials.

2. BOUNDING THE POWER

The total image plane signals considered in this paper may be thought of as deriving from the uniform, 1024×1024, square lattice sampling of light intensity fields. They are therefore discrete, two-dimensional, spatial signals. Assume that the diameter of the circular region $\mathcal{R}$ is $2N+1$. Let $\mathcal{A}$ be the set of two-dimensional indices over $\mathcal{R}$, such that the $(0, 0)$ point is at the center of $\mathcal{R}$.

Denote by $I(n_1, n_2)$ the amplitude signal on $\mathcal{R}$. Then $I(n_1, n_2)$ can be represented in matrix form as

$$I(n_1, n_2) = \sum_{(n_1,n_2)\in\mathcal{A}} B(n_1, n_2) r(n_1, n_2) \delta(n_1, n_2)$$

where n_1 is the row index, n_2 is the column index, $\delta(n_1, n_2)$ is the unit sample, $r(n_1, n_2) \geq 0$ for all n_1 and n_2, and $B(n_1, n_2)$ represents the amplitude compression function. Note that $B(n_1, 0) = J_0(n_1)$ where $J_0(x)$ is the first order Bessel function scaled so that $J_0(-N) = J_0(N) \approx 0$. The energy stored in I is

$$E_I = \left\{ \sum_{(n_1,n_2)\in\mathcal{A}} B(n_1, n_2)^2 r(n_1, n_2)^2 \right\}^{\frac{1}{2}}$$

Recall that a function T which can be represented as

$$T(x) = \sum_{k=M_1}^{M_2} a(k)\sin(kx) + b(k)\cos(kx)$$

is called a *trigonometric polynomial*. Let

$$Coef = \left\{ J_0\left(\frac{2.4}{N}k\right) : 0 \leq k \leq N \right\} \cup \{0\}$$

and let S_N represent the set of all nonnegative, trigonometric polynomials of the form

$$Q(x) = 1 + \sum_{k=1}^{N} a(k)\sin(kx) + b(k)\cos(kx)$$

where $a(k), b(k) \in Coef$. In this section, a bound on

$$\begin{aligned} I_N &= \sup\left\{ \frac{1}{2\pi}\int_0^{2\pi} Q^2(x)dx : Q(x) \in S_N \right\} \\ &= 1 + \frac{1}{2}\sum_{k=1}^{N} a(k)^2 + b(k)^2 \end{aligned}$$

is approximated. The reason for considering such a bound was discussed in the previous section, as was the choice of $N = 24$. A related problem was addressed by Brown *et al.* in [1] where they found that

$$\sup\left\{ \frac{1}{2\pi}\int_0^{2\pi} Q^2(x)dx : Q(x) \in P_N \right\} \in [(N+1)C_1, (N+1)C_1 + 1]$$

where P_N represents the set of all nonnegative, trigonometric polynomials of degree less than or equal to N. An algorithm for calculating $C_1 = 0.686981293\ldots$ was given by Garsia *et al.* in [2]. This result implies that

$$17 \approx (24+1)C_1 \leq I_{24} \leq (24+1)C_1 + 1 \approx 18$$

however a simple calculation shows that

$$I_{24} \leq 1 + \frac{1}{2}\sum_{k=1}^{N} 2J_0(0.1k)^2 \leq 11.9 \ .$$

This implies that S_{24} is a relatively low energy subset of P_{24}. A nonexhaustive computer search for high energy members of S_{24} was performed in order to judge the tightness of 11.9 as a bound for this one-dimensional slice through $\mathcal{R}$. This search produced the following member of S_{24}:

$$w(x) = 1 + \sum_{k=4}^{24} J_0(0.1k)\cos[(k-3)x]$$

The (one-sided, or cosine-only) L_2^2 norm of w is approximately 4.98. (There are no units on the energy, since it is a function of pixel value.) This is approximately 48% of the maximum possible energy for any member of S_{24}. However, the high degree of w leads one to suspect that 11.9 is not a tight bound for I_{24}.

Now consider the whole of $\mathcal{R}$, rather than the one-dimensional slices. Then the maximum energy possible for Z is 22.1. The maximum-energy, non-negative trigonometric polynomial w was used to generate a two-dimensional function w_2, which has energy 17.8. The maximum energy for a member of $\mathcal{C}$, found by exhaustive search, is 16.3. However, the bound on w_2 found by nonexhaustive search and analytical methods is only 9% higher, making this bound useful and more computationally tractable.

3. ENERGY DISTRIBUTION

When the channel is calibrated, the modulation codes can be adjusted in the vicinity of burst errors to operate with an energy distribution more suited to the amplitude compression.

Reed-Solomon codes are used by Chiarulli *et. al.* [3] in holographic memories. However, in this particular system, because of the large amount of distortion experienced, some type of modulation coding is needed to assist detection. Reed-Solomon codes are generally used for correcting burst errors, where the burst is constrained to only a few Reed-Solomon symbols (a collection of several bits, not to be confused with the modulation code symbols in this system). By the nature of the Reed-Solomon code, its performance tends to be superior only when the interleaving tends to group bit

errors together so that several occur within the same Reed-Solomon symbol, and are not spread uniformly over the Reed-Solomon symbols in a single word.

The modulation codes chosen from the system were restricted to 2×2 symbols for compatibility with previous research, and because an integral number of 2×2 symbols can be fit into the image plane. Low-energy modulation codes perform well for detection in this system over most of the image plane (this work is currently being summarized for publication). Two of the codes analyzed in the work included the symbols shown in figure 3. Call the top-row symbol set V_1 and the bottom-row symbol set V_2.

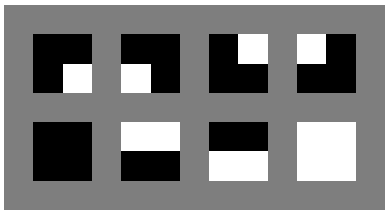

Figure 3: Modulation Code Symbols

However, the performance of low-energy modulation codes in $\mathcal{R}$ is poor because of the amplitude compression: near the edges of $\mathcal{R}$, the attenuation from the PSF, coupled with the amplitude compression, pushes isolated "one" pixels below the noise floor. The trade-off for locating effective modulation codes is that the higher the energy of the symbol, the higher the level of ISI resulting from the symbol.

A modulation code from $\mathcal{C}$ is generated when the top row from figure 3 is used inside the center half-radius ($\mathcal{R}_1$), and when the second row is used for the outer half-radius ($\mathcal{R}_2$). Call this code v. (Exclusive-or with a PN sequence may be required to ensure that the RLL restrictions are met.) This code offers an energy of 14.9, which is 91% of the $\mathcal{C}$ bound. It is not certain if the bound can actually be achieved, since the energy bounds found in the previous section do not necessarily correspond to 2×2 pixel symbols in a rate-1/2 modulation code.

The modulation codes are used in conjunction with conventional channel coding, such as algebraic or convolutional codes. The raw symbol error rate (RSER) will not be zero because of distortion and local errors. However, with intelligent interleaving, tandem coding, and an RSER less than about 10^{-3}, the composite symbol error rate can reach zero.

In a typical run of the detector, V_1 experiences an RSER of 10^{-3} in $\mathcal{R}$, with all of the errors in $\mathcal{R}_2$. The set V_2 experiences an RSER of approximately 3×10^{-2} overall, with better performance than V_1 in $\mathcal{R}_2$. The combined performance of v is approximately 5×10^{-4} over $\mathcal{R}$. Because v does not depend on the use of an accurate inverse of the amplitude attenuation, shifts in the position of $\mathcal{R}$ in the image plane have only a small effect, if any, on the resulting RSER.

4. CONCLUSIONS

The maximum energies are summarized in table 1. It can be seen that the energy bound, as found by analysis of non-negative trigonometric polynomials, is much tighter than the coarse bound found from Z. It is interesting to note that the bound given on w_2 without the restriction to $\mathcal{C}$ is only 9% higher than the bound found by exhaustive search, while the coarse bound is 36% higher.

The v modulation code achieves 91% of the bound energy. As discussed above, v performs better than the higher energy code obtained by using only V_2, since V_2 does not correspond to a member of $\mathcal{C}$.

Table 1: Summary of Signal Energies

Signal	Maximum Energy
Z	22.1
w_2	17.8
$\sup\{E_I : I \in \mathcal{C}\}$	16.3
v	14.9

Finally, since the members of $\mathcal{C}$ correspond 99.8% to non-negative trigonometric polynomials, the theory developed in this paper can be used to select a modulation code from $\mathcal{C}$ and to judge its effectiveness by considering its energy. This measure must be used since not all elements of $\mathcal{C}$ necessarily correspond to modulation codes, and exhaustive search of $\mathcal{C}$ for the "optimal" modulation code is computationally intractable.

5. REFERENCES

[1] J. Brown, M. Goldstein, and J. McDonald, "A sequence of extremal problems for trigonometric polynomials," *Journal of Mathematical Analysis and Applications*, vol. 130, no. 2, pp. 545–551, 1988.

[2] A. Garsia, E. Rodemich, and H. Rumsey, "On some extremal positive definite functions," *Journal of Mathematics and Mechanics*, vol. 18, no. 9, pp. 805–834, 1996.

[3] D. M. Chiarulli, R. Hofmann, S. P. Levitan, and M. Pandey, "Error detection and correction for an optoelectronic cache memory," *SPIE*, vol. 3468, 1998.