서브클러스터링을 이용한 홀로그래픽 정보저장 시스템의 비트 에러 보정 기법

Bit Error Reduction for Holographic Data Storage System Using Subclustering

김상훈*, 양현석†,박영필*

Sang-Hoon Kim, Hyunseok Yang and Young-Pil Park

(2009 년 9 월 1 일 접수; 2010 년 3 월 12 일 심사완료; 2010 년 3 월 12 일 게재확정)

Abstract

Data storage related with writing and retrieving requires high storage capacity, fast transfer rate and less access time. Today any data storage system cannot satisfy these conditions, however holographic data storage system can perform faster data transfer rate because it is a page oriented memory system using volume hologram in writing and retrieving data. System can be constructed without mechanical actuating part so fast data transfer rate and high storage capacity about 1Tb/cm3 can be realized. In this research, to correct errors of binary data stored in holographic data storage system, a new method for reduction errors is suggested. First, find cluster centers using subtractive clustering algorithm then reduce intensities of pixels around cluster centers. By using this error reduction method following results are obtained ; the effect of Inter Pixel Interference noise in the holographic data storage system is decreased and the intensity profile of data page becomes uniform therefore the better data storage system can be constructed.

Key Words : Holographic data storage, subtractive clustering, bit error, inter pixel inteference

1. Introduction

Holographic Data Storage System[1], one of the next generation data storage approaches, is a 2-dimensional page oriented memory using a volume hologram in the writing and retrieving process. In contrast to conventional holography, where light scattered from a three-dimensional scene or object is recorded and can be subsequently reconstructed for visual viewing, holographic storage uses the light coming from a very special electronic binary object known as a spatial light modulator (SLM). The SLM is used to display a two-dimensional binary pattern of ‘ones’ and ‘zeros’, which act very much like miniature open and closed shutters representing the information to be stored. A common example of a SLM is the liquid crystal displays commonly used for hand-held television and notebook computer displays.

A laser beam passes through the SLM (or is reflected by it in the case of a reflection SLM), becomes modulated in two dimensions by the displayed pattern, and is directed to a recording medium where it interferes with another beam, called the reference beam, to form the hologram. It is most common to interpose a lens between the SLM and the recording medium so that the Fourier transform of the SLM pattern is recorded rather than the diffraction from the pattern itself (see figure 1). This provides certain advantages in recording and retrieval. A new pattern, or page of information, is set up on the SLM and another hologram is recorded in the same volume of material with a different reference beam. A large number of different holograms are recorded by several multiplexing techniques, spatial, angle, wavelength and phase, etc. To retrieve a particular page

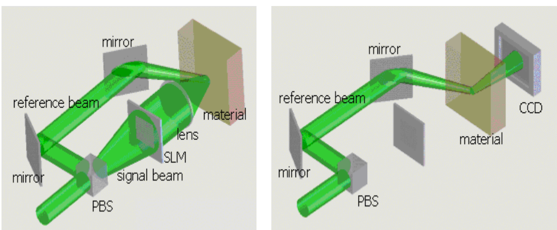
† 연세대학교 기계공학과
E-mail : hsyang@yonsei.ac.kr
TEL : (02)2123-2824
* 연세대학교 기계공학과

of information that has been stored holographically, the recording medium is illuminated with the reference beam that was used to record that page. This light interacts with all of the recorded holograms. However, because they are each highly tuned structure, only light is diffracted from the structure associated with that particular reference beam. A wavefront is created that is identical to the light that was coming from the SLM when that page was displayed, illuminated and recorded. This light is imaged onto an array of photodetectors, converted to an electrical signal and transferred back to the cpu via the electronic channel. Each page can be retrieved independently be using the associated reference beam.

In the Holographic Data Storage system, the data management procedure is performed in parallel so a fast data transfer rate can be realized. The system stores data in binary form (0 or 1), so that computers can use the digital data directly. In the writing procedure, a laser of a specific wavelength passes through a SLM to make 2-dimensional data page. A digital data 0 makes an image of a black pixel (off-pixel) by blocking the light on the SLM and digital data 1 is imaged as a white pixel (on-pixel) on a CCD camera.

Diffraction, the nature of light, makes the laser which passes through an on-pixel surrounded by off-pixels on the SLM to affect surrounding pixels by a 2-dimensional Fourier transform of a plane wave and during the retrieving process, this diffracted light effects to surrounding pixels of the on-pixel and are photographed by CCD camera and cause errors to binary data.

In this paper, we analyze the effect of an on-pixel (digital data 1) to surrounding off-pixels (digital data 0) and suggest a new method for reducing errors of binary data using a subtractive clustering method. The error reduction algorithm is ; find centers of on-pixel group and reduce the effect of on-pixels.



(a) (b)

Fig.1 Holographic Storage Principle (a) recording (b) reconstruction

2. Inter Pixel Interference Noise

By the Fraunhofer diffraction theory, the beam from the laser source emits a plane wave. However after crossing the aperture – SLM, the beam propagates as a spherical wave from many point sources and diffracts. The on-pixels of the SLM allow the beam to pass through, but the off-pixels do not. So an on-pixel acts like as an aperture, and the beam through the on-pixel of the SLM will affect the surrounding pixels by a 2-dimensional Fourier Transform[2]. This effect acts as one of the main sources of noise, and is called "Inter Pixel Interference" noise.

Inter Pixel Interference (IPI) degrades the performance of the channel, and it tends to occur when an off-pixel is surrounded by on-pixels or vice versa. Where there are many on-pixels in comparison, errors from IPI noise will appear relatively more numerous than other area of the data page. Therefore if we find a cluster of on-pixels and reduce the intensities of the on-pixels, the possibility of error by IPI noise will decrease. Regarding the entire data page, because the intensity of laser of brighter area will be decreased, a uniform intensity profile will be obtained.

The IPI noise from a 2-dimensional Fourier transform is shown in figure 2. To make simulations with IPI noise, it is necessary to set the numerical values of IPI noise. To obtain the numerical values of IPI noise, a two-dimensional Fast Fourier Transform (function 'fft2' of MATLAB) is used. From the result of MATLAB, the artificial IPI noise similar to the real IPI noise is made. Figure 4 describes the effect of one on-pixel to surrounding pixels with normalized intensity value 1. As shown in figure 3, the effect exists mainly on the horizontal and vertical directions (there is no effect in the diagonal direction in artificial IPI noise). For neighboring pixels, the value of the effect of IPI noise is set to 0.135 and for the next pixel the value is set to 0.075 when the maximum intensity value is normalized to 1.

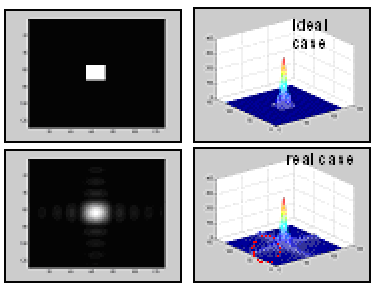


Fig.2 Inter Pixel Interference

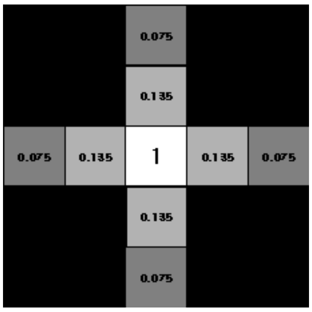


Fig.3 Numerical values of artificial IPI noise

3. Subtractive Clustering Algorithm

To find a region where there are many on-pixels for comparison, a subtractive clustering algorithm is used. By subtractive clustering algorithm, potential values are defined as a function inversely proportional to distances from each datum to the other data. The potential values are calculated about all data and the datum with the largest potential value is determined as the cluster center. Because a potential value is inversely proportional to a summation of distances between data points, the data with many neighbors nearby has a larger potential value. Therefore, the data with largest potential value is determined to be the cluster center and the next cluster center is found without the effect of previous cluster centers. This procedure will be repeated until the calculated potential value becomes smaller than a specific value.[2][3][4]

A potential value is expressed by the next equation.

$$P_i = \sum_{k=1}^{N} \exp\left(-\alpha \left\| x_k - x_i \right\|^2\right) \qquad (1)$$

$$P_1^* = \max_i P_i \qquad (2)$$

Then, find the cluster centers after the second by subtracting the previous cluster centers.

$$P_i' = P_i - P_1^* \exp\left(-\beta \left\| x_i - x_1^* \right\|^2\right) \qquad (3)$$

$$P_2^* = \max_i P_i' \qquad (4)$$

To process the subtractive clustering algorithm, several parameters must be determined; a cluster radius, a standard distance between centers and a minimum potential value for recognition as a cluster center. In this paper, a simulation was performed with various parameter values. The cluster radius was determined to 0.1 ; the standard distance between centers was 1, and the minimum potential value was 0.7.

4. Error Reduction Algorithm

Figure 4 is the picture to be stored. The size of it is 120x90 so that it consists of 10800 pixels with 256 level gray-scale. Pixel values of the picture are in the range from 0 to 255 (256 level), from black to white. To store the pixel values, it is necessary to change the decimal values into binary values because the HDS system stores ‘digital data’ (binary data).

When one decimal number is changed into binary, its size becomes to 8 bits. Before changing, one gray-scaled pixel can have a number from 0 to 255, but after changing one binary pixel can have a value of just 0 or 1. Thus the amount of data for representing decimal values for one pixel increases by a factor of eight.

The total size of the picture increases from 120x90 to 960x90. Figure 5 represents the procedure of changing decimal data into binary.

It is necessary to make data pages that will be displayed on a SLM. The size of a data page on the SLM is 120x90. Binary data of the original image has 960x90 pixels so it must be divided into eight data pages of 120x90. Figure 6 shows the first data page which will be displayed on the SLM.

The process for error reduction will be applied to all data pages, however since the process to the first page is equal to the second, to the third and so on, the error reduction process will be evaluated with the first data page (which will be called ‘original data page’) only.

When the blurring of on-pixels by the diffraction effect is applied, the intensity will increase about some range and in the area with comparatively crowded on-pixels, the errors from IPI noise will occur. Therefore reducing intensity of the area by the proper value will be able to correct the errors.

Fig.4 The picture wanted to be stored

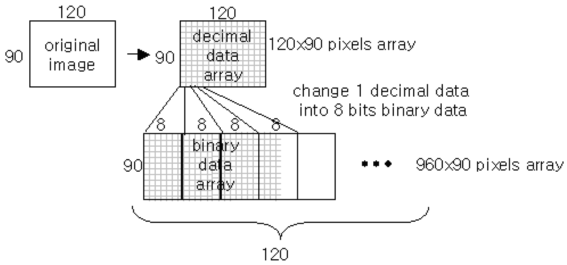


Fig.5 Procedure of changing decimal data into binary

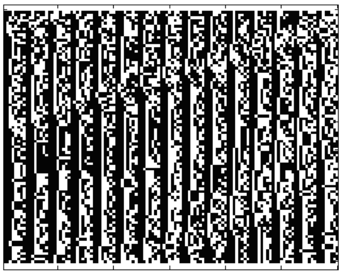


Fig.6 The first data page made

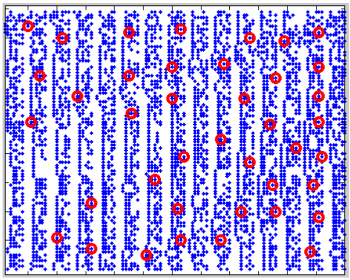


Fig.7 Cluster centers found

Figure 7 is the original data page and cluster centers found (red circles) using the subtractive clustering algorithm. About the cluster centers found, the intensities of 9x9 pixels around the center pixel are reduced by a parameter delta. The ntensity reduction is achieved by subtracting the value of the on-pixel of the SLM with the value of delta. The range of modification is determined by simulation results, the best case among 7x7, 9x9, 11x11, etc.

After applying IPI noise, the number of error pixels in the original data page is 61. The number of error pixels of the modified data page by the bit error reduction algorithm is 17. The value of delta is 0.2, which is chosen arbitrarily.

5. Parameter Optimization

5.1 Back Propagation Algorithm

A neural network[5] is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects :

Knowledge is acquired by the network through a learning process.

Interneuron connection strengths known as synaptic weights are used to store the knowledge.

The procedure used to perform the learning process is called a learning algorithm, the function of which is to modify the synaptic weights of the network in an orderly fashion so as to attain a desired design objective.

The modification of synaptic weights provides the traditional method for the design of neural networks. Such an approach is the closest to linear adaptive filter theory, which is already well established and successfully applied in such diverse fields as communications, control, radar, sonar, seismology, and biomedical engineering. However, it is also possible for a neural network to modify its own topology, which is motivated by the fact that neurons in the human brain can die and that new synaptic connections can grow.

The back propagation algorithm[5] makes the network to learn until the output vector resembles the input vector or is classified with the proper input vector. To minimize the mean square error of the network, control weights and biases using back propagation rules. This procedure changes the weights and biases to reduce the error as quickly as possible. The variations of weights and biases are proportional to the effects of the elements about the mean square error of the network.

It often happens that this process does not reach the global minimum but a local minimum when using back propagation learning. For that case to get the global minimum it is necessary to construct the network with more neurons and layers, and then the problem becomes complex. Sometimes using another initial condition can solve the problem.

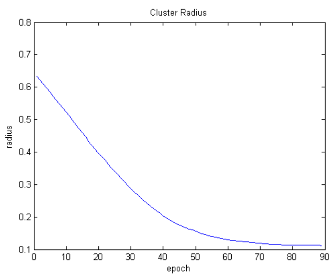
5.2 Optimization Results

The optimization about the cluster radius and intensity reduction value was performed with the back propagation algorithm. The parameters to be optimized are selected by considering the amount of effects to the result of the process. Radius and delta are selected and optimized. The small variation of delta does not affect to the results. Therefore, the value of delta will be changed by the step of 0.01.

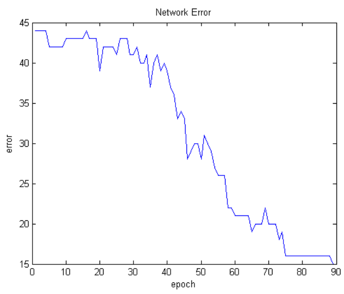
These are the results of optimization of the parameter radius about the first data page using back propagation algorithm. The process started with arbitrary initial value bounded from 0 to 1. Figure 8(a) is variation of the value of the radius versus iteration number(epoch) and figure 8(b) is the number of error pixels with the change of radius.

The optimized value of the radius is 0.1121 and delta is 0.2068 and the number of error pixels is 15. And the number of cluster centers also decreases to 40. A smaller number of cluster centers is better because it makes the subtractive clustering procedure faster than before.

Before optimization the error reduction algorithms were applied with parameter values of 0.1 (radius) and 0.2 (delta). The number of error pixels before optimizing was 17 and number of error pixels is reduced by two.
Optimization results for all data pages are given in table 1. For optimization, the value of the radius is 0.1 and the value of delta is 0.2, same as the first data page.



(a)



(b)

Fig.8 (a) optimization process about radius (b) variation of the number of error pixels

Table 1. Optimization results

data page #	before optimization			after optimization		
	radius	delta	errors	Radius	Delta	errors
1	0.1	0.2	25	0.1121	0.2	15
2	0.1	0.2	38	0.076	0.12	25
3	0.1	0.2	49	0.0576	0.1	21
4	0.1	0.2	46	0.0612	0.1	30
5	0.1	0.2	41	0.0543	0.1	29
6	0.1	0.2	34	0.0822	0.15	22
7	0.1	0.2	33	0.0681	0.1	26
8	0.1	0.2	28	0.0935	0.16	12

6. Simulations and Experimental Results

Figure 9 is a data page made by applying IPI noise from a 2-dimensional Fourier Transform to the original data page shown in figure 6. After reconstruction the digital data from the data page with IPI noise using the threshold method, 61 numbers of pixels are reconstructed differently from the original pixels.

Figure 10(a) is the modified data page found by applying the optimal intensity reduction value (delta) to 9 by 9 pixels from the cluster centers in the center. Figure 10(b) is the data page modified by the suggested algorithm with IPI noise.

Only 15 error pixels are found from the reconstructed digital data of the modified data page using the threshold method.

Figure 11(a) is the photographed image of the original data page by a CCD camera and figure 11(b) is the image of the modified data page. It can be perceived that the intensity profile of the modified data page by applying the suggested error reduction algorithm is more uniform than the original data page.

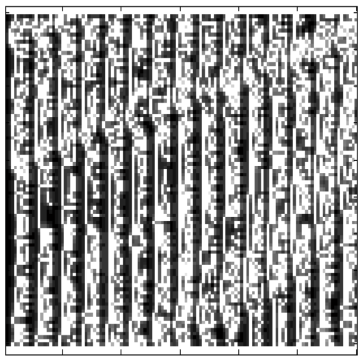
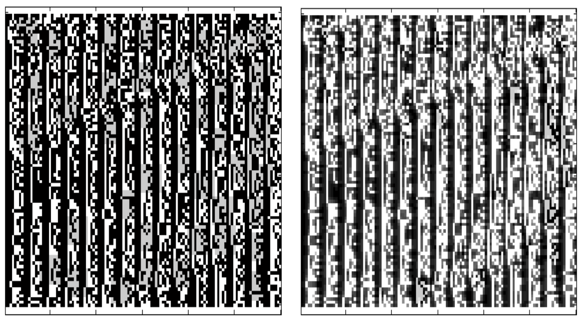
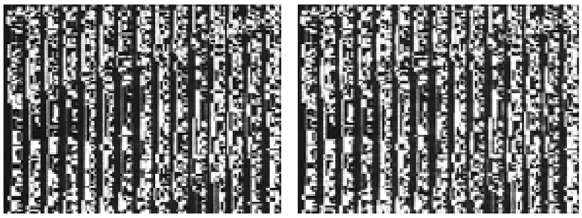


Fig.9 The original data page with IPI noise



(a) (b)

Fig.10 (a) modified data page (b) modified data page with IPI noise



(a) (b)

Fig.11. Photographed images by CCD
(a) original data page (b) modified data page

7. Conclusions

Holographic data storage system (HDSS) is a strong candidate for the next generation data storage system. However, to become a conventional information storage device, the HDSS has to solve several problems, for example noise sources, size and cost, etc.

In this research, the IPI noise – one of the main noise sources of HDSS – is analyzed and a bit error reducing method is suggested. To reduce the effect of IPI noise, finding cluster centers that are a result of groups of on-pixels was performed by using a subtractive clustering algorithm. For the cluster centers found, intensities of 9x9 pixels around the center pixel are reduced by a specific value delta.

The parameter values of the radius in the subclustering algorithm and delta are optimized using a back propagation algorithm of a neural network. Better results are obtained after optimization and the results are shown in table 5.1.

After applying the bit error reduction algorithm with optimized parameters, BER of 1.667×10^{-2} is obtained which is better than the BER of 7.991×10^{-2} before applying the algorithm.

후 기

이 논문은 2008 년도 정부(교육과학기술부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구임 (No. R17-2008-040-01001-0).

참고문헌

[1] Hans J. Coufal, Demetri Psaltis and Glen T.Sincerbox, 2000, *Holographic Data Storage*, Springer, New York, USA.

[2] Won Nah, Joong-Hwan Baek, 2003, "Classification of C.elegans Behavioral Phenotypes Using Clustering" Proceedings of the IEEK Conference, pp.1743-1746.

[3] Min-Soo Kim and Seong-Gon Kong, 1997, "Fuzzy Control System Design by Data Clustering in the Input-Output Subspaces", The Institute of Electronics Engineers of Korea, Vol 34S, No 12, pp.30-40

[4] Su-Hwan Kim, Hyung-Jin Kang and Tae-Won Rhee, 1992, "Image Segmentation Using an Extended Fuzzy Clustering Algorithm", Journal of the Korean Institute of Telematics and Electronics B, Vol 29B, No 3, pp.35-46

[5] Simon Haykin, *Neural Network*