Ghost Imaging by a Proportional Parameter to Filter Bucket Data

Min Tao 1, Xiaobin Gong 1, Jian Guan 1, Junfeng Song 1, Zhixin Song 2,* , Xueyan Li 1, Shuxu Guo 1, Jian Chen 3, Siyao Yu 1,* and Fengli Gao 1,2,*

State Key Laboratory of Integrated Optoelectronics, College of Electronic Science and Engineering, Jilin University, Changchun 130012, China; taomin19@mails.jlu.edu.cn (M.T.); gongxb18@mails.jlu.edu.cn (X.G.); guanjian16@mails.jlu.edu.cn (J.G.); songjf@jlu.edu.cn (J.S.); leexy@jlu.edu.cn (X.L.); guosx@jlu.edu.cn (S.G.)

College of Medical Engineering and Technology, Xinjiang Medical University, Urumqi 830017, China

Shanhe-Gensheng Technology Co., Ltd. of Jilin Province, Jilin 132011, China; qiaojf18@mails.jlu.edu.cn

* Correspondence: yfyymzx@xjmu.edu.cn (Z.S.); yu_sy@jlu.edu.cn (S.Y.); gaofl@jlu.edu.cn (F.G.)

Abstract: Most ghost imaging reconstruction algorithms require a large measurement time to retrieve the object information clearly. But not all groups of data play a positive role in reconstructing the object image. Abandoning some redundant data can not only enhance the quality of reconstruction images but also speed up the computation process. Here, we propose a method to screen the data using two threshold values set by a proportional parameter during the sampling process. Experimental results show that the reserved data after screening can be used in several reconstruction algorithms, and the reconstruction quality is enhanced or at least remains at the same level. Meanwhile, the computing time costs are greatly reduced, and so is the data storage.

Keywords: imaging processing; ghost imaging; photon statistics; threshold selection

1. Introduction

Being different from the conventional imaging technology, Ghost imaging (GI) utilizes two spatially correlated beams to retrieve the object image information [1–7]. The beams are generated when an optical field is divided into an object beam and a reference beam. The target image can be found by calculating the correlation function between the bucket signal from the object beam path and the optical field distribution from the reference beam. When first proposed, ghost imaging was realized by the entangled photon pairs with quantum illumination. Successively, pseudo-thermal light, true thermal light, and even X-rays have been proven to be viable for the GI technique, which allows the experimental principle of ghost imaging to expand from quantum interpretation to the high-order correlation of light, and the diversity of light sources further benefits the development of GI [1,3–8]. Compared with the conventional imaging techniques, GI shows promising potentials in many fields, such as optical encryption, remote sensing, lidar detection, and so on, owing to its separation of imaging and detection [9–14]. However, there are two necessary improvements if GI was to be applied in practice. One is to modify the reconstruction algorithm to improve the reconstruction quality, and the other is to reduce the number of samples so as to save the calculation time. In recent years, new GI reconstruction algorithms have been continually proposed, such as differential ghost image (DGI) [15], normalized ghost imaging (NGI) [16], pseudo-inverse ghost imaging (PGI) [17,18], scalar-matrix-structured ghost imaging (SMGI) [19], iterative pseudo-inverse ghost imaging (IPGI) [20], compressive sensing ghost imaging (CSGI) [21,22], and binomial theorem ghost imaging (BGI) [23], which all improve the imaging quality in different ways. However, better reconstruction qualities are often accompanied by higher system complexity and much longer computation time. In terms of reducing sampling times,
some researchers focus on setting threshold values to screen the sampling data [24–26]. For example, the method using positive and negative correlations introduces the mean value of the speckle field information obtained by the reference arm as a threshold [25], in which all the sampled data are divided into two detailed sets, so each pair of speckle fields is given two labels. Double-threshold time-correspondence imaging (DTTCI) is a method of setting threshold values [26], but it needs an additional bucket detector in the reference arm of the conventional GI experimental system. In addition, computing ghost imaging (CGI) utilizes a spatial light modulator to save the construction of the reference light path, which greatly simplifies the GI experimental system and further facilitates the practical application of GI [27,28]. Recently, deep learning, as a research hotspot, has been applied in many fields, and some researchers apply deep learning in GI [29,30], but it requires vast amounts of data to train its network.

Experiments prove that DTTCI has a good imaging effect in NGI and DGI, but not for other reconstruction algorithms. The reason is that their imaging formulas contain the ratio of the measured data of the object arm bucket detector to the reference arm bucket detector and its average value. Therefore, we propose a novel and simple method by setting double threshold values according to a proportional parameter to screen the data in the sampling process. First, it counts the total light intensity data of the object arm bucket detector with only a few hundred measurements and sets two threshold values (a large one and a small one) in advance. Then, during the sampling process, it only records the total light intensity, which is higher than the larger threshold or lower than the smaller one. Their corresponding light field distribution of the reference arm at the same time and the data that does not meet the threshold condition will be abandoned. The experimental results show that this method can greatly reduce the time cost during the reconstruction procedure and keep the same, or sometimes even higher, levels of reconstruction quality using less screened data. Moreover, this method can also be used in other algorithms, such as PGI, SMGI.

2. Experimental Setup and Principle

The schematic diagram of the experimental system is shown in Figure 1, which consisted of the optical setup part and the data processing part. In the optical setup part, the pseudo-thermal light source was generated through a laser beam illuminating a rotating ground glass. The pseudo-thermal source was divided into the reference beam path and the object beam path by a 50:50 beam splitting mirror. After the object beam illuminated the target object with transmission function \( T(x, y) \), its total light intensity was recorded by a charge coupled device (CCD1), and the \( n \)th measurement value was denoted as \( B_n \). At the same time, the optical intensity distribution on the surface of the object was recorded by another same model CCD(CCD2) located at the same distance in reference path with the object (L1 = L2), and the \( n \)th measurement value was denoted as \( I_n(x, y) \). We defined the \( n \)th measurement value \( B_n \) and \( I_n(x, y) \) at the same moment as a set of sampled data. Since the \( B_n \) and \( I_n(x, y) \) were one-to-one corresponding, when we screened \( B_n \), the \( I_n(x, y) \) was screened as well. In the data processing part, after screened all of the \( B_n \) and \( I_n(x, y) \), these data that met the conditions participate in reconstructing the object.
Assuming that the total measurements time were $N$, the conventional GI algorithm can be expressed as:

$$T_{GI}(x, y) = \frac{1}{N} \sum_{n=1}^{N} (B_n - \langle B_n \rangle) I_n(x, y)$$

(1)

where $\langle B_n \rangle = \frac{1}{N} \sum_{n=1}^{N} B_n$ denotes the ensemble average of $B_n$. It is known from Equation (1) that the group of data had little influence on retrieving the object information when $(B_n - \langle B_n \rangle)$ approximated zero. Therefore, removing these $(B_n)$ and their corresponding $I_n(x, y)$ simplified the calculation and reduced reconstruction time cost.

Now, the question became how to find these data. We know that the mean value approximately equals the median when the sampled data are sufficient. Fortunately, a traditional GI algorithm and most modified reconstruction algorithms require a large amount of sampled data. In other words, those $B_n$ which equal or nearly equal to $\langle B_n \rangle$ are in the exact middle when all $B_n$ are arranged from large to small. To confirm this statistical property, 5000 sets of sampled data used in the experiment part were organized to tabulate the frequency statistics histogram, as shown in Figure 2, where the $B_n$ equals to $\langle B_n \rangle$ ranks 2438th when all of $B_n$ were arranged from large to small, and approximately equals to the median. According to our experiments, DTTCI was effective for NGI and DGI for a similar reason.

As can be seen in Figure 1, those $B_n$ meeting the condition that $(B_n - \langle B_n \rangle)$ equals or nearly equals to zero were located in the very center when all of $B_n$ were arranged from large to small. Since the amount of $B_n$ was rather large, so we could remove a certain number of $B_n$ values distributed around $\langle B_n \rangle$ along with their corresponding $I_n(x, y)$. To implement our method, the major obstacle was how to preselect two suitable intensity
threshold values that were relative to the average of $B_n$ since the average of $B_n$ could only be determined after all the values of $B_n$ had been measured and calculated. Here, we introduced a proportional parameter $C$ to filter the bucket data so as to get fewer and more useful data for the reconstruction, and $C$ was defined as:

$$C = \frac{N_T}{N} \quad (2)$$

where $N$ is the total number of data groups, and $N_T$ is the number of data groups screened from the $N$ sets of data. Hundreds of groups of data were processed according to the frequency statistics histogram of $(B_n - \langle B_n \rangle)$ to determine the proportion parameter $C$ before each experiment, and then we selected two threshold values according to $C$. In the experiments, we combined the screening process with the sampling process to filter each group of the sampled data and put the data corresponding to the $B_n$ within the threshold range into the data pool to participate in the operation. On the contrary, the data out of range were abandoned.

The proposed method was tested in SMGI and PGI to verify its effectiveness. Assuming that the object image and the optical fields distribution of the reference beam were both $p \times q$ pixels, the light field distribution of each measurement could be turned into a row vector, then these $N$ row vectors were pieced together into a matrix denoted as observation matrix $\Phi$ as below:

$$\Phi = \begin{bmatrix}
I_1(1, 1) & I_1(1, 2) & \cdots & I_1(p, q) \\
I_2(1, 1) & I_2(1, 2) & \cdots & I_2(p, q) \\
\vdots & \vdots & \ddots & \vdots \\
I_N(1, 1) & I_N(1, 2) & \cdots & I_N(p, q)
\end{bmatrix} \quad (3)$$

The total light intensity $B_n$ in $N$ times of measurements and the transmission coefficient $T(x, y)$, of the object were rewritten into corresponding vector forms. By constructing the observation matrix, the GI formula could also be transformed into a matrix form:

$$T_{GI} = \frac{1}{N} \Phi^T \begin{bmatrix}
B_1 \\
B_2 \\
\vdots \\
B_N
\end{bmatrix} - \frac{1}{N} \Phi^T \Phi \begin{bmatrix}
\langle B_n \rangle \\
\langle B_n \rangle \\
\vdots \\
\langle B_n \rangle
\end{bmatrix} = \frac{1}{N} \Phi^T (B_n) - \frac{1}{N} \Phi^T (\langle B_n \rangle) \quad (4)$$

where $\Phi^T$ is the transpose of $\Phi$. $N$ is the total sampling number. SMGI modifies the observation matrix $\Phi$ by subtracting a matrix $\Phi_x$ to make the calculation result of $(\Phi^T - \Phi_x)\Phi$ more similar to a scalar matrix, expressed as [19]:

$$T_{SMGI} = \frac{1}{N}(\Phi^T - \Phi_x)\Phi \begin{bmatrix}
T(1, 1) \\
T(1, 2) \\
\vdots \\
T(p, q)
\end{bmatrix} - \frac{1}{N}(\Phi^T - \Phi_x) \langle B_n \rangle \quad (5)$$

Similarly, PGI replace $\Phi$ with its pseudo-inverse matrix $\Phi^\dagger$ [17,18] as below:

$$T_{PGI} = \frac{1}{N} \Phi^\dagger \Phi \begin{bmatrix}
T(1, 1) \\
T(1, 2) \\
\vdots \\
T(p, q)
\end{bmatrix} - \frac{1}{N} \Phi^\dagger \langle B_n \rangle \quad (6)$$

From Equations (5) and (6), it was known that all the modifications only change the $I_n(x, y)$ part of the formulas rather than the part containing $(B_n - \langle B_n \rangle)$, so theoretically, our method was equally effective for SMGI and PGI.
3. Experimental Results

All the experiments were carried out on the setup as shown in Figure 1, in which the distance $L_1$ ($L_1 = L_2$) is 200 mm, the laser wavelength $\lambda = 532$ nm, the focal length $f$ of the lens was 150 mm, CCD1 (Stingray F-504B, AVT, Stadtroda, Germany) was identical to CCD2, CCD1 was used to gather the sum of light field distribution modulated by the Object and CCD2 was used to record the light field distribution of reference beam at the surface of the object.

First of all, to test how effective our method was for the conventional GI algorithm expressed by Equations (1) and (4), we respectively screened 2000, 3000, and 4000 groups of data according to the threshold values determined by the corresponding proportional parameter $C$, from the total 5000 groups of sampled data to reconstruct the targets “GI” and “ZHONG”, and the reconstruction results are shown in Figure 3. Using Peak Signal-to-Noise Ratio (PSNR) as the quantitative index to evaluate the quality of the reconstructed image, the results are shown in Figure 4, in which the upper and lower dotted lines represent the PSNR values of reconstruction results when the number of data groups was 5000.

From Figure 3, it is obvious that the visual effects using 2000, 3000, and 4000 groups of data were almost the same as all the groups of data (5000 groups). From Figure 4, we can see that when the $C$ was no less than 0.6, the PSNR of the reconstructed results remained the same quality or slightly improved if some data were abandoned. It can also be seen
from Figure 4 that when \( C = 0.6 \) to 0.7, the reconstruction quality was better than that of GI using all sampled data. Multiple experimental results showed that \( C = 0.65 \) was applicable to all binary images tested. In other words, the parameter \( C \) between 0.6 and 0.7 was a general optimal solution. Therefore, when reconstructing a target, we could choose the minimum and maximum of the total light intensity in the rejection data when \( C = 0.65 \) as two thresholds. Moreover, the reason why the PSNR values go slightly down with the increase in the \( C \) value (The amount of data used for reconstruction was also increasing) was that the effect of the deleted data groups might bring more noise so as to affect the improvement in reconstruction quality, which also proved that if the data were not filtered, there would be a considerable amount of useless data in the sampled data and they would have a side effect on the reconstruction quality. The experiment showed that our method can keep the quality of reconstruction basically unchanged when it deletes a certain number of the sampled data. Meanwhile, the computing time is greatly reduced since the screening process does not take much time. For example, when \( C = 0.6 \), using the filtered data could save about 15 s computing time (the computer environment: CPU i5-8300H, memory 8G, Lenovo, Beijing, China) under the premise of ensuring the reconstruction quality. From the above, we demonstrated the effectiveness of our method on the GI algorithm.

According to the experiments above, the optimum range of \( C \) was 0.6–0.7, so in the subsequent experiments (procedure of sampling and storing data), we selected the maximum and minimum values of \( Bn \) corresponding with \( C = 0.65 \) as the two thresholds and carried out the sampling procedure combined with the filtering process. Respectively, we sampled a total of 1200 sets of data with and without filtering process for the two different images, “GI” and “ZHONG”, which were processed by the PGI and SMGI algorithms to further test the effectiveness of our method. PGI and SMGI using the data filtered by the two thresholds were recorded as T-PGI and T-SMGI, respectively, for short, and all the results are shown in Figures 5–8, where 300, 600, 900, and 1200 in Figures 5 and 7, respectively, represent the amount of data used from the total 1200 sets of data, and Figures 6 and 8 are the PSNR curves corresponding to Figures 5 and 7 respectively.

![Image](image_url)

**Figure 5.** Experimental results for the “GI” and “ZHONG” by the pseudo-inverse ghost imaging (PGI) and T-PGI methods with 300, 600, 900, and 1200 sets of data.
Figure 6. PSNR curves of the reconstruction result “GI” and “ZHONG” of pseudo-inverse ghost imaging (PGI) and T-PGI; (a) the PSNR curves of the “GI” of PGI and T-PGI; (b) the PSNR curves of “ZHONG” of PGI and T-PGI.

Figure 7. Experimental results for the “GI” and “ZHONG” were obtained by the scalar-matrix-structured ghost imaging (SMGI) and T-SMGI methods with 300, 600, 900, and 1200 measurements.

Figure 8. PSNR curves of the reconstruction result “GI” and “ZHONG” of SMGI and T-SMGI; (a) the PSNR curves of the “GI” of SMGI and T-SMGI; (b) the PSNR curves of SMGI and T-SMGI of “ZHONG”.

In Figure 5, the visual effects of reconstructed images of T-PGI and PGI enhanced along with the increase in data sets, and under the same amount of data, the visual effect of
T-PGI was also better than that of PGI. From Figure 6, we can see that the PSNR increased along with the data sets when the PSNR of T-PGI was much higher than that of PGI all the time, which were consistent with the visual effects in Figure 5.

It can be seen from Figures 7 and 8 that SMGI and T-SMGI had similar results. These results confirm that our data screening method is suitable for PGI and SMGI algorithm enhancement. In addition, the PSNR and visual effect of SMGI reconstruction results were better than those of PGI, which is consistent with our previously published experimental results [19]. In fact, the threshold effect of counting hundreds of sets of data was similar to that of 5000 groups of data, because the values of Bn were basically concentrated near the mean value of Bn. Because the process of data screening was directly completed in the sampling process, and the filtering process only involved simple addition or subtraction calculation, the filtering process hardly increased the sampling process time, but also saved the storage space occupied by the original redundant data and greatly reduced the time of the following reconstruction computation.

The experimental results above were all based on the reconstruction of binary images, so this method is suitable for GI, PGI, and SMGI when processing binary images.

4. Conclusions

In conclusion, we chose the proportion parameter C value corresponding to the best reconstruction effect to select the double thresholds to complete the data filtering in the sampling process. Before sampling the data, we verified the effectiveness of our method for the GI algorithm and first determined the two threshold values by the proportion parameter C. In the following experiments, by comparing the reconstruction results of PGI and SMGI using filtered data and unfiltered data, experimental results showed that our method can be applied to a variety of algorithms. Besides, combining the filtering procedure with the sampling process, the storage space occupied by the data that are not beneficial to the reconstruction quality can be saved. The next reconstruction calculation time is reduced greatly, and the reconstruction quality remained basically unchanged or even increased slightly. Based on the above advantages, we believe that this method will be useful for the practicability of ghost imaging.

Author Contributions: Methodology, M.T.; project administration, X.G.; validation, J.G.; formal analysis, J.S.; investigation, Z.S.; resources, J.C.; data curation, S.Y.; writing—original draft preparation, X.G.; writing—review and editing, F.G.; supervision, X.L.; funding acquisition, S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key R&D Program of China (2016YFE0200700), National Natural Science Foundation of China (NSFC) (61627820, 61934003), Industrial Technology Research and Development Project of Development and Reform Commission of Jilin Province (2020C019-2), Jilin University Science and Technology Cooperation Special Project of Science and Technology Innovation Development Plan of Jilin City (201830797).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study.

Acknowledgments: We are grateful to the anonymous reviewers for constructive comments that improved the quality of the work.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Pittman, T.B.; Shih, Y.H.; Strekalov, D.V.; Sergienko, A.V. Optical imaging by means of two-photon quantum entanglement. Phys. Rev. A 1995, 52, 3429–3432.
2. Strekalov, D.V.; Sergienko, A.V.; Klyshko, D.N.; Shih, Y.H. Observation of two-photon “ghost” interference and diffraction. Phys. Rev. Lett. 1995, 74, 3600–3603. [PubMed]
3. Bennink, R.S.; Bentley, S.J.; Boyd, R.W. “Two-photon” coincidence imaging with a classical source. *Phys. Rev. Lett.* 2002, 89, 113601. [PubMed]
4. Erkmen, B.; Shapiro, J. Ghost imaging: From quantum to classical to computational. *Adv. Opt. Photon.* 2010, 2, 405–450.
5. Padgett, M.J.; Boyd, R.W. An introduction to ghost imaging: Quantum and classical. *Philos. Trans. Roy. Soc. A* 2017, 375, 20160233.
6. Ferri, F.; Magatti, D.; Gatti, A.; Bache, M.; Brambilla, E.; Lugliato, L.A. High-resolution ghost image and ghost diffraction experiments with thermal light. *Phys. Rev. Lett.* 2005, 94, 183602.
7. Gatti, A.; Brambilla, E.; Bache, M.; Lugliato, L.A. Ghost imaging with thermal light: Comparing entanglement and classical correlation. *Phys. Rev. Lett.* 2004, 93, 093602.
8. Pelliccia, D.; Rack, A.; Scheel, M.; Cantelli, V.; Paganin, D.M. Experimental X-ray ghost imaging. *Phys. Rev. Lett.* 2016, 117, 113902.
9. Cheng, J.; Lin, J. Unified theory of thermal ghost imaging and ghost diffraction through turbulent atmosphere. *Phys. Rev. A* 2013, 87, 043810.
10. Zhang, P.; Gong, W.; Shen, X.; Han, S. Correlated imaging through atmospheric turbulence. *Phys. Rev. A* 2010, 82, 033817.
11. Zhang, Y.; Li, W.; Wu, H.; Chen, Y.; Su, X.; Xiao, Y.; Wang, Z.; Gu, Y. High-visibility underwater ghost imaging in low illumination. *Opt. Commun.* 2019, 44, 45–48.
12. Erkmen, B. Computational ghost imaging for remote sensing. *J. Opt. Soc. Am. A* 2012, 29, 782–789.
13. Gong, W.; Zhao, C.; Yu, H.; Chen, M.; Xu, W.; Han, S. Three-dimensional ghost imaging lidar via sparsity constraint. *Sci. Rep.* 2016, 6, 26133. [PubMed]
14. Moreau, P.A.; Toninelli, E.; Gregory, T.; Padgett, M.J. Ghost imaging using optical correlations. *Laser Photon Rev.* 2018, 12, 1700143.
15. Ferri, F.; Magatti, D.; Lugliato, L.A.; Gatti, A. Differential ghost imaging. *Phys. Rev. Lett.* 2010, 104, 253603. [PubMed]
16. Sun, B.; Welsh, S.; Edgar, M.; Shapiro, J.; Padgett, M. Normalized ghost imaging. *Opt. Express.* 2012, 20, 16892–16901.
17. Zhang, C.; Guo, S.; Cao, J.; Guan, J.; Gao, F. Object reconstitution using pseudo-inverse for ghost imaging. *Opt. Express.* 2014, 22, 30063–30073.
18. Gong, W.L. High-resolution pseudo-inverse ghost imaging. *Photonic Res.* 2015, 3, 234–237.
19. Yang, C.; Wang, C.; Guan, J.; Zhang, C.; Guo, S.; Gong, W.; Gao, F. Scalar-matrix-structured ghost imaging. * Photon. Res.* 2016, 4, 281–285.
20. Lv, X.; Guo, S.; Wang, C.; Yang, C.; Zhang, H.; Song, J.; Gong, W.; Gao, F. Experimental investigation of iterative pseudoInverse ghost imaging. *IEEE Photonics J.* 2018, 10, 1–8.
21. Katz, O.; Bromberg, Y.; Silberberg, Y. Compressive ghost imaging. *Appl. Phys. Lett.* 2009, 95, 131110. [CrossRef]
22. Huang, H.; Zhou, C.; Tian, T.; Liu, D.; Song, L. High-quality compressive ghost imaging. *Opt. Commun.* 2018, 412, 60–65. [CrossRef]
23. Yue, C.; Chen, P.; Lv, X.; Wang, C.; Guo, S.; Song, J.; Gong, W.; Gao, F. Object Reconstruction Using the Binomial Theorem for Ghost Imaging. *IEEE Photonics J.* 2018, 10, 1–13. [CrossRef]
24. Luo, K.; Huang, B.; Zheng, W.; Wu, L. Nonlocal imaging by conditional averaging of random reference measurements. *Chin. Phys. Lett.* 2012, 29, 074216. [CrossRef]
25. Sun, M.; Li, M.; Wu, L. Nonlocal imaging of a reflective object using positive and negative correlations. *Appl. Opt.* 2015, 54, 7494–7499. [CrossRef] [PubMed]
26. Li, M.; Zhang, Y.; Liu, X.; Yao, X.; Luo, K.; Fan, H.; Wu, L. A double-threshold technique for fast time-correspondence imaging. *Appl. Phys. Lett.* 2013, 103, 211119. [CrossRef]
27. Shapiro, J.H. Computational ghost imaging. *Phys. Rev. A* 2008, 78, 061802. [CrossRef]
28. Bian, L.; Suo, J.; Dai, Q.; Chen, F. Experimental comparison of single-pixel imaging algorithms. *J. Opt. Soc. Am. A* 2018, 35, 78–87. [CrossRef]
29. Komuro, K.; Nomura, T.; Barbastathis, G. Deep ghost phase imaging. *Appl. Opt.* 2020, 59, 3376–3382. [CrossRef]
30. Wu, H.; Wang, R.; Zhao, G.; Xiao, H.; Liang, J.; Wang, D.; Tian, X.; Cheng, L.; Zhang, X. Deep-learning denoising computational ghost imaging. *Opt. Lasers Eng.* 2020, 134, 106183. [CrossRef]