Formation mechanism of correspondence imaging with thermal light

Jian Leng, Wen-Kai Yu,* and Shuo-Fei Wang
Center for Quantum Technology Research, School of Physics, Beijing Institute of Technology, Beijing 100081, China
and Key Laboratory of Advanced Optoelectronic Quantum Architecture and Measurement of Ministry of Education, School of Physics, Beijing Institute of Technology, Beijing 100081, China

(Received 14 October 2019; revised manuscript received 28 January 2020; accepted 2 March 2020; published 23 March 2020)

Correspondence imaging can achieve positive-negative ghost images by just conditional averaging of partial patterns, without treating bucket intensities as weights. To explain its imaging mechanism, we develop a probability theory assuming the targets are of grayscale and the thermal reference speckles obey an arbitrary independent and identical distribution. By both simulation and experiments, we find that the recovered values in each region of the same original gray value conditionally obey a Gaussian distribution. A cross-point–to–standard-deviation ratio is used as the figure of merit to prove that the patterns with respect to larger bucket values generate a positive image with a higher quality and vice versa for a negative one. This work complements the theory of ghost imaging.

DOI: 10.1103/PhysRevA.101.033835

I. INTRODUCTION

Ghost imaging (GI) provides a way to recover object images via intensity correlation between reference patterns and bucket intensity signals. It was primitively demonstrated by using entangled light [1] and then was also experimentally realized with thermal or pseudothermal (a laser passing through a rotating ground glass) light [2–4] as well as x-ray light [5,6]. As long as the light field of the reference arm and the object arm are conjugated, the lenses in GI with thermal light can be removed [7,8], which makes the imaging setup simpler. Thus, the thermal light GI has been widely used in many fields such as optical encryption [9,10] and lidar [11]. To solve two key problems existing in GI, i.e., the image quality and measurement number, various GI methods have sprung up, such as background-removal GI [12], differential GI (DGI) [13], adaptive GI [14], iterative denoising GI [15], blind GI [16], super sub-Nyquist GI [17]. Among these methods, the bucket values serve as the weights, reflecting the total intensity from the modulated object. Recently, an interesting experimental study found that one could generate the positive and negative ghost images by only conditionally averaging partial reference patterns. This method was named correspondence imaging (CI) [18–20]. It has sparked an upsurge of interest in studying the topic of positive-negative ghost imaging and spawned diverse variants [21–26] and CI-based applications [27,28]. It seemed that the bucket weights no longer participated in the correlation calculations involved in the second- or higher-order correlation functions, but actually they were completely binarized. Confusing questions were raised: Why could CI generate positive-negative images using only a few reference patterns and why could CI work without involving bucket weights in the calculations? Their theoretical explanations have been topics of study in this field for a long time, but after several attempts [29–31], research is ongoing. Recently, a strict explantation based on a probability theory [32] was provided, which regarded the light intensities as stochastic variables and deduced a joint probability density function between the bucket and reference signals, giving us some inspiration. However, this theory was based on a fundamental assumption of the simplified model that consists of the negative exponential distributed light field and binary objects; thus it still had its limitations, especially in universality. The imaging mechanism of CI deserves further research.

In this paper we assume a general model in which the targets are of grayscale (each gray value has a large enough number of pixels), any two thermal speckles in the light field are independent of each other, all following an arbitrary identical distribution, and all the reference speckles constitute a set of independent stochastic variables. The bucket values can be treated as many linear combinations of all pixels, also constituting a random variable. With the above assumptions, we can deduce the joint probability density function between the bucket variable and each reference thermal speckle variable. We then provide the forming formulas of the positive and negative images. Both simulation and experimental results demonstrate the correctness of our derivation. Furthermore, we use this theoretical model to investigate how image quality varies with specific selection intervals used to average reference patterns.

II. PROBABILITY THEORY

A. Statistical model of ghost imaging

As we know, for a continuous random variable X, the probability of $X < x$, i.e., the distribution function, can be written as $F_X(x) = P\{X < x\}$; then we have $F_X(-\infty) = 0$ and $F_X(+\infty) = 1$. Suppose the probability density function

*yuwenkai@bit.edu.cn

 ©2020 American Physical Society

$f_X(x)$ of X is the derivative of $F_X(x)$, i.e., $f_X(x) = F'_X(x)$, then $\int_{-\infty}^{+\infty} f_X(x)dx = F_X(+\infty) - F_X(-\infty) = 1$. Next we will use its two typical mathematical properties of the random variable X: One is the mathematical expectation (also known as the mean) $E(X)$, defined as

$$E(X) = \int_{-\infty}^{+\infty} x f_X(x)dx, \tag{1}$$

and the other is the variance $D(X)$, defined as

$$\begin{aligned} D(X) &= \int_{-\infty}^{+\infty} [x - E(X)]^2 f_X(x)dx \\ &= E(X^2) - E(X)^2. \end{aligned} \tag{2}$$

We assume that the grayscale object has a total of M pixels, with d representing the gray value of a pixel. The gray value of the mth point (pixel) is denoted by d_m, ranging from 0 to 1, with 0 being completely opaque and 1 being completely transparent. Accordingly, each reference pattern can also be divided into M pixels, each of which has a light intensity expressed by I_m. This intensity value can be regarded as a random variable, which obeys an arbitrary identical probability distribution I. For simplicity of mathematics, it is assumed that the intensities of any two thermal speckles (pixels) in the reference spatial light field are statistically independent of each other. Then the distribution function of the mth random variable I_m can be written as $F_{I_m}(i_m)$ and its probability density function can be denoted by $f_{I_m}(i_m)$, where $i_m \in [0, \infty)$. The mth pixel of the object is illumined by the corresponding thermal speckle. On the plane after the thermal light passes through the grayscale object, the mth point will have the value $d_m I_m$. In addition, since there is still a certain distance between the object plane and the bucket detector, along with some existing influence factors such as diffraction and refraction, a certain loss of light intensity should be considered here, expressed by the coefficient factor a. Thus, when the light reaches the sensing surface of the bucket detector, the intensity becomes $Y_m = a d_m I_m$. Then we have the relationship between the mth point in the object arm and the nth point in the reference arm,

$$\begin{aligned} E(Y_m I_n) &= a d_m E(I_m I_n) \\ &= \begin{cases} a d_n E(I^2), & m = n \\ a d_m E(I)^2, & m \neq n, \end{cases} \end{aligned} \tag{3}$$

which we call the independence discrimination separated formula. Now the bucket light intensity can be written as

$$S = \sum_m^M Y_m = a \sum_m^M d_m I_m, \tag{4}$$

whose distribution function and probability density function are denoted by $F_S(s)$ and $f_S(s)$ ($s \in [0, \infty)$), respectively.

For convenience of calculation, suppose the subscript of the point of our interest is n; then we define a physical quantity S_n that is very similar to the bucket value S, but excluding the bucket intensity with the subscript n:

$$S_n = \sum_{m \neq n}^M Y_m = a \sum_{m \neq n}^M d_m I_m. \tag{5}$$

Obviously, S_n is independent of I_n. According to the definition of S_n, we can immediately have

$$S = S_n + Y_n. \tag{6}$$

We let $F_{S_n}(s_n)$ and $f_{S_n}(s_n)$ ($s_n \in [0, \infty)$) denote the distribution function and the probability density function of S_n, respectively.

With the definitions given above, let us recall the second-order correlation formula of traditional GI

$$G_n^{(2)} = \langle S I_n \rangle, \tag{7}$$

where $\langle \cdot \rangle$ denotes the average operation. Using this second-order correlation function, the reconstructed gray value of the nth pixel can be calculated. It is known from a large number of experimental conclusions that although GI results are accompanied by large fluctuations, we can still visually distinguish the morphological characteristics of the object from ghost images, i.e., the reconstructed pixels in the positions where the original gray value is large are also of large values. Here we will build a statistical model to strictly explain this phenomenon. First, we will calculate the mean $E(SI_n)$ of the GI formula

$$\begin{aligned} E(SI_n) &= E[(S_n + Y_n)I_n] \\ &= E(S_n I_n) + E(Y_n I_n) \\ &= E\left[\left(\sum_{m \neq n}^M Y_m\right) I_n\right] + E(Y_n I_n) \\ &= \sum_{m \neq n}^M E(Y_m I_n) + E(Y_n I_n) \\ &= \sum_{m \neq n}^M a d_m E(I)^2 + a d_n E(I^2) \\ &= a \sum_m^M d_m E(I)^2 + a[E(I^2) - E(I)^2] d_n \\ &= \gamma_1 + \gamma_2 d_n, \end{aligned} \tag{8}$$

where both γ_1 and γ_2 are constants. Since d_n is the gray value of any object point, the physical meaning of the second-order correlation function is to perform the same linear transformation on the gray value of each object point. This is the essential reason why the second-order correlation algorithm can recover the object images. Here Eq. (3) plays a decisive role. Next we will further study the relationship between the CI mean and original grayscale values in a similar way.

B. Approximation of the model

In this section we begin by proving the following theorem to deduce the approximate distribution expressions of S and S_n, which are only related to the mean $E(I)$ and the variance $D(I)$ of the light intensity I, independently of the specific distribution of I.

Theorem 1. When each gray value in the object image has infinite points (pixels), the bucket value S in GI with thermal light strictly obeys a normal distribution.

Proof. Let the arbitrary gray value of the object be $d^{(k)}$ ($k \in \{1, 2, \ldots, K\}$) and its number of points (pixels) be $l^{(k)}$, which tends to infinity. We define the variable $S^{(k)}$ as the sum of all the points with the same gray value $d^{(k)}$ in the object arm as

$$S^{(k)} = \sum_{m=1|\{d_m=d^{(k)}\}}^{l^{(k)}} Y_m = \sum_{m=1|\{d_m=d^{(k)}\}}^{l^{(k)}} a d_m I_m = a d^{(k)} \sum_{m=1}^{l^{(k)}} I_m. \tag{9}$$

Since $l^{(k)}$ tends to infinity, according to the central-limit theorem for independently and identically distributed variables in the probability theory, $S^{(k)}$ follows a normal (Gaussian) distribution with a mean of $\mu^{(k)} = l^{(k)} a d^{(k)} E(I)$ and a variance of $(\sigma^{(k)})^2 = l^{(k)} a^2 (d^{(k)})^2 D(I)$. Therefore, according to the gray value, we can rewrite the definition $S = \sum_m^M Y_m$ of S as

$$S = \sum_m^M Y_m = \sum_k^K \left(\sum_{m=1|\{d_m=d^{(k)}\}}^{l^{(k)}} Y_m \right) = \sum_k^K S^{(k)}. \tag{10}$$

Then S is the sum of k Gaussian distributions. According to the probability theory, S obeys a Gaussian distribution

$$F_S(s) \approx \phi\left(\frac{s-\mu}{\sigma}\right), \tag{11}$$

$$f_S(s) \approx \frac{1}{\sqrt{2\pi}\sigma} e^{-(s-\mu)^2/2\sigma^2}, \tag{12}$$

with a mean of

$$\mu = \sum_k^K \mu^{(k)} = \sum_k^K l^{(k)} a d^{(k)} E(I) = a \sum_m^M d_m E(I) \tag{13}$$

and a variance of

$$\sigma^2 = \sum_k^K (\sigma^{(k)})^2 = \sum_k^K l^{(k)} a^2 (d^{(k)})^2 D(I) = a^2 \sum_m^M d_m^2 D(I). \tag{14}$$

■

Thus, if we suppose each gray value has sufficient points (pixels), the requirements of the above theorem can be satisfied. Then we will have that S approximately follows a normal distribution with a mean of $\mu = a \sum_m^M d_m E(I)$ and a variance of $\sigma^2 = a^2 \sum_m^M d_m^2 D(I)$. Similarly, S_n also approximately fulfills a normal distribution with a mean of $\mu_n = a \sum_{m\neq n}^M d_m E(I)$ and a variance of $\sigma_n^2 = a^2 \sum_{m\neq n}^M d_m^2 D(I)$.

C. Explanation for correspondence imaging

With the obtained distributions of S and S_n, we will start the calculation for CI. The joint probability density function between S and Y_n can be deduced as

$$f_{S,Y_n}(s, y_n) = f_{S_n}(s_n) \otimes f_{Y_n}(y_n) = f_{S_n}(s - y_n) f_{Y_n}(y_n). \tag{15}$$

To average the patterns corresponding to the bucket value above or below its ensemble average, we define

$$s_+ = \begin{cases} 1, & s \geqslant \mu \\ 0, & s < \mu, \end{cases} \tag{16}$$

$$s_- = 1 - s_+. \tag{17}$$

Obviously, there are

$$\lim_{s_{\max}\to\infty} \int_0^{s_{\max}} s_+ f_S(s) ds = \int_\mu^\infty f_S(s) ds = \frac{1}{2}, \tag{18}$$

$$\lim_{s_{\max}\to\infty} \int_0^{s_{\max}} s_- f_S(s) ds = \lim_{s_{\max}\to\infty} \int_0^{s_{\max}} (1 - s_+) f_S(s) ds = 1 - \frac{1}{2} = \frac{1}{2}. \tag{19}$$

To obtain the average of the patterns that correspond to the bucket values above the ensemble average, i.e., $E(s_+ I_n)$, we should first compute

$$E(s_+ Y_n) = \frac{\lim_{s_{\max}\to\infty} \int_0^{s_{\max}} \left[\int_0^s s_+ y_n f_{S,Y_n}(s, y_n) dy_n\right] ds}{\lim_{s_{\max}\to\infty} \int_0^{s_{\max}} s_+ f_S(s) ds} = 2 \int_\mu^\infty \left[\int_0^s f_{S_n}(s - y_n) y_n f_{Y_n}(y_n) dy_n\right] ds. \tag{20}$$

Since $E(Y) \ll \mu$, we can treat y_n in the above integral as a very small amount $f_{S_n}(s - y_n) \approx f_{S_n}(s) - f'_{S_n}(s) y_n$, where $f'_{S_n}(s)$ represents the derivative of $f_{S_n}(s)$. Since any pixel value in the spatial light field of the object arm will not be greater than the bucket value, i.e., $y_n \leqslant s$, we set the upper and lower limits of the integral of the independent variable y_n to s and 0, respectively. Actually, in the term $\int_0^s f_{S_n}(s - y_n) y_n f_{Y_n}(y_n) dy_n$, given the fact that the value s is much larger than the upper limit of the independent variable y_n, the upper limit of this integral can be equivalent to infinity, i.e., $\int_0^s dy_n \approx \int_0^\infty dy_n$. Then we have

$$\begin{aligned} E(s_+ Y_n) &\approx 2 \int_\mu^\infty \left\{\int_0^\infty [f_{S_n}(s) - f'_{S_n}(s) y_n] y_n f_{Y_n}(y_n) dy_n\right\} ds \\ &= 2E(Y_n) \int_\mu^\infty f_{S_n}(s) ds - 2E\left(Y_n^2\right) \int_\mu^\infty f'_{S_n}(s) ds \\ &= 2E(Y_n)[1 - F_{S_n}(\mu)] - 2E\left(Y_n^2\right)[0 - f_{S_n}(\mu)] \\ &= 2E(Y_n)\{1 - F_{S_n}[\mu_n + E(Y_n)]\} \\ &\quad + 2E\left(Y_n^2\right) f_{S_n}[\mu_n + E(Y_n)] \\ &\approx 2E(Y_n)[1 - F_{S_n}(\mu_n) - F'_{S_n}(\mu_n) E(Y_n)] \\ &\quad + 2E\left(Y_n^2\right)[f_{S_n}(\mu_n) + f'_{S_n}(\mu_n) E(Y_n)]. \end{aligned} \tag{21}$$

Since

$$F_{S_n}(\mu_n) = \frac{1}{2}, \tag{22}$$

$$F'_{S_n}(\mu_n) = f_{S_n}(\mu_n) = \frac{1}{\sqrt{2\pi}\sigma_n}, \tag{23}$$

$$f'_{S_n}(\mu_n) = 0, \tag{24}$$

then

$$E(s_+Y_n) \approx 2E(Y_n)\left[\frac{1}{2} - \frac{1}{\sqrt{2\pi}\sigma_n}E(Y_n)\right] + 2E\left(Y_n^2\right)\frac{1}{\sqrt{2\pi}\sigma_n}$$
$$= E(Y_n) + 2D(Y_n)\frac{1}{\sqrt{2\pi}\sigma_n}, \tag{25}$$

where

$$E(Y_n) = E(ad_nI_n) = ad_nE(I), \tag{26}$$

$$D(Y_n) = D(ad_nI_n) = a^2d_n^2D(I), \tag{27}$$

$$E(s_+Y_n) = E[s_+(ad_nI_n)] = ad_nE(s_+I_n). \tag{28}$$

So we will get

$$E(s_+I_n) = E(I) + \sqrt{\frac{2}{\pi}}\frac{a}{\sigma_n}D(I)d_n. \tag{29}$$

Replacing the standard deviation σ_n of S_n with the standard deviation $\sigma = a\sqrt{\sum_m^M d_m^2 D(I)}$ of S, we can acquire

$$E(s_+I_n) \approx E(I) + \sqrt{\frac{2D(I)}{\pi\sum_m^M d_m^2}}d_n$$
$$= C_2 + C_1d_n, \tag{30}$$

where

$$C_1 = \sqrt{\frac{2D(I)}{\pi\sum_m^M d_m^2}}, \tag{31}$$

$$C_2 = E(I). \tag{32}$$

Similarly, to calculate the average of the patterns that correspond to the bucket values below the ensemble average, i.e., $E(s_-I_n)$, we should first compute

$$E(s_-Y_n) = \frac{\lim_{s_{\max}\to\infty}\int_0^{s_{\max}}\left[\int_0^s s_-y_nf_{S,Y_n}(s,y_n)dy_n\right]ds}{\lim_{s_{\max}\to\infty}\int_0^{s_{\max}}s_-f_S(s)ds}$$
$$= 2\lim_{s_{\max}\to\infty}\int_0^{s_{\max}}\left[\int_0^s(1-s_+)y_nf_{S,Y_n}(s,y_n)dy_n\right]ds$$
$$= 2E(Y_n) - E(s_+Y_n)$$
$$\approx E(Y_n) - 2D(Y_n)\frac{1}{\sqrt{2\pi}\sigma_n}. \tag{33}$$

With the same processing method used for $E(s_+Y_n)$, we get

$$E(s_-I_n) \approx E(I) - \sqrt{\frac{2D(I)}{\pi\sum_m^M d_m^2}}d_n$$
$$= C_2 - C_1d_n. \tag{34}$$

Then we can compute the formula of difference image

$$CI_\pm = E(s_+I_n) - E(s_-I_n)$$
$$= 2C_1d_n. \tag{35}$$

Since C_1 and C_2 are both constants, the positive-negative images and $CI_\pm$ are all the linear transformations of the original object. Because the coefficient C_1 before d_n in $E(s_+I_n)$ is positive, its result presents a positive image, while the coefficient $-C_1$ before d_n in $E(s_-I_n)$ is negative; the result is rendered as a negative image.

III. VERIFICATION FOR CORRESPONDENCE IMAGING

The theoretical averages of the positive and negative images and $CI_\pm$ have been given above, but the gray value of each pixel in the actual reconstructed images generally fluctuates around the mean, following a certain distribution. Below we will focus on this distribution and make a verification. Let us suppose there is a total of T measurements, containing T_+ bucket values $S \geqslant \langle S\rangle$ and T_- bucket values $S < \langle S\rangle$, where $\langle\cdots\rangle$ denotes the ensemble average of the signal. The operators s_+ and s_- still use the definitions mentioned above. We denote the nth point in the tth speckle pattern by I_{nt}. Then the positive and negative image formulas in the actual image reconstruction can be written as

$$\langle s_+I_n\rangle = \frac{1}{T_+}\sum_{t=1}^{T}s_+I_{nt}, \tag{36}$$

$$\langle s_-I_n\rangle = \frac{1}{T_-}\sum_{t=1}^{T}s_-I_{nt}. \tag{37}$$

According to the central-limit theorem for independently and identically distributed variables, when T_+ is large enough, $\langle s_+I_n\rangle$ approximatively obeys a Gaussian distribution with a mean of $E(s_+I_n)$ and a variance of $\frac{D(s_+I_n)}{T_+}$; similarly, when T_- is large enough, $\langle s_-I_n\rangle$ approximatively follows a Gaussian distribution with a mean of $E(s_-I_n)$ and a variance of $\frac{D(s_-I_n)}{T_-}$.

Now we will compute the variances $D(s_+I_n)$ and $D(s_-I_n)$. In a way similar to calculating $E(s_+I_n)$ and $E(s_-I_n)$, we first derive $E(s_+I_n^2)$ and $E(s_-I_n^2)$:

$$E\left(s_+I_n^2\right) \approx E(I^2) + \sqrt{\frac{2}{\pi\sum_m^M d_m^2D(I)}}[E(I^3) - E(I^2)E(I)]d_n, \tag{38}$$

$$E\left(s_-I_n^2\right) \approx E(I^2) - \sqrt{\frac{2}{\pi\sum_m^M d_m^2D(I)}}[E(I^3) - E(I^2)E(I)]d_n. \tag{39}$$

By using the formula $D(X) = E(X^2) - E(X)^2$, the variances can be calculated as

$$D(s_+I_n) = E\left(s_+I_n^2\right) - E(s_+I_n)^2, \tag{40}$$

$$D(s_-I_n) = E\left(s_-I_n^2\right) - E(s_-I_n)^2. \tag{41}$$

Substituting Eqs. (30), (34), (38), and (39) into Eqs. (40) and (41) and omitting minor terms, we will get

$$D(s_+I_n) = E(I^2) + \sqrt{\frac{2}{\pi\sum_m^M d_m^2D(I)}}[E(I^3) - E(I^2)E(I)]d_n$$
$$- \left[E(I) + \sqrt{\frac{2D(I)}{\pi\sum_m^M d_m^2}}d_n\right]^2$$
$$\approx E(I^2) - E(I)^2 = D(I), \tag{42}$$

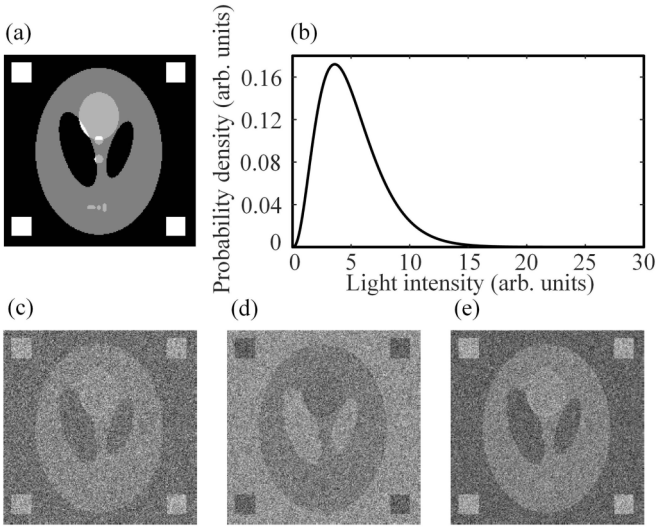

FIG. 1. Simulation results: (a) original image, a modified head phantom image; (b) chosen probability density function curve; and reconstructions of (c) the positive image, (d) the negative image, and (e) their difference image.

$$D(s_{-}I_n) = E(I^2) - \sqrt{\frac{2}{\pi \sum_m^M d_m^2 D(I)}}[E(I^3) - E(I^2)E(I)]d_n - \left[E(I) - \sqrt{\frac{2D(I)}{\pi \sum_m^M d_m^2}}d_n\right]^2 \approx E(I^2) - E(I)^2 = D(I). \quad (43)$$

So far, we can theoretically compute the distribution curve of a certain gray value $d^{(k)}$ (occupying a region that consists of several pixels) after image reconstruction. For the positive image, it follows a Gaussian distribution with a mean of $E(s_{+}I_n)$ and a variance of $\frac{D(s_{+}I_n)}{T_{+}}$, while for the negative image, it obeys a Gaussian distribution with a mean of $E(s_{-}I_n)$ and a variance of $\frac{D(s_{-}I_n)}{T_{-}}$. In both simulation and experiments, we calculate the probability of the recovered pixel values falling in each pixel region where the gray value of the original image equals $d^{(k)}$ and plot the corresponding probability density curves, compared with the Gaussian theoretical curves (obtained from the computed theoretical means and variances) to demonstrate the correctness of the theory.

A. Simulation

Here we choose an object image of 200×200 pixels, as shown in Fig. 1(a); its statistical data of the gray values are given in Table I. We take the speckle variables of the patterns which obey an identical Γ distribution as an example, parameterized in terms of a shape parameter $\alpha = 3.57$ and a scale parameter $\theta = 1.4$. Its probability density function can be expressed as

$$f_I(i) = \frac{i^{\alpha-1}e^{-i/\theta}}{\theta^{\alpha}\Gamma(\alpha)} \quad \text{for } i > 0, \quad (44)$$

TABLE I. Statistical data of gray values in the original image.

Gray value	Total number of pixels	Proportion
0	23353	58.38%
0.5	13147	32.87%
0.7	1733	4.33%
1	1767	4.42%

as plotted in Fig. 1(b). The positive and negative images with a total of 50 000 frames and their difference image $CI_{\pm}$ are given in Figs. 1(c)–1(e).

For both positive and negative images, we separately compute the probability of reconstructed pixel values falling in each pixel region corresponding to the one that consists of pixel positions with the same gray value $d^{(k)}$ of the original image and draw their probability density curves to compare with the Gaussian theoretical curves, as shown in Figs. 2(a) and 2(b). From the graphs, we can clearly see that the recovered pixel value data are highly consistent with presupposed Gaussian distribution.

B. Experiment

For a practical measurement environment, there are many kinds of noise with different distributions, but their superposition will result in a Gaussian distribution with a large

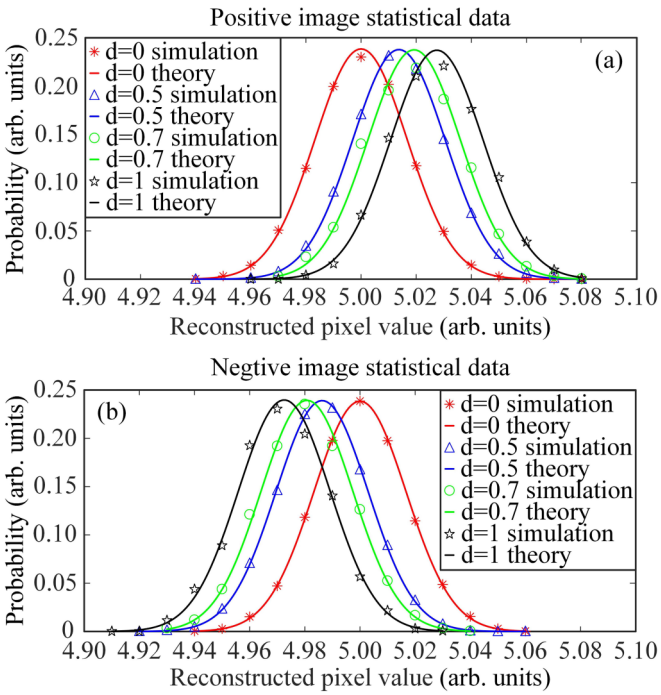

FIG. 2. Probability density function curves for the recovered pixel values, compared with the Gaussian theoretical curves. The probability density distributions and Gaussian theoretical curves of reconstructed pixel values falling in each pixel region where the gray value of the original image equals $d^{(k)}$ are shown for (a) positive and (b) negative images. The abscissa is the reconstructed pixel value and the ordinate indicates the probability of occurrence of these values.

probability. Thus, we may assume that the measurement noise is a random Gaussian variable e with a mean of $E(e) = 0$ and an unknown variance $D(e)$. Then

$$S = S_n + Y_n + e. \tag{45}$$

In previous calculations, we only need to replace S_n with $S_n + e$; $S_n + e$ also satisfies a Gaussian distribution with a mean $\mu_n + E(e)$ and a variance $\sigma_n^2 + D(e)$. Here we directly present the results:

$$\begin{aligned} E(s_+ I_n) &\approx E(I) + \sqrt{\frac{2}{\pi}}\sqrt{\frac{1}{\sum_m^M d_m^2 D(I) + \frac{D(e)}{a^2}}} D(I) d_n \\ &= C_2 + C_3 d_n, \end{aligned} \tag{46}$$

$$\begin{aligned} E(s_- I_n) &\approx E(I) - \sqrt{\frac{2}{\pi}}\sqrt{\frac{1}{\sum_m^M d_m^2 D(I) + \frac{D(e)}{a^2}}} D(I) d_n \\ &= C_2 - C_3 d_n, \end{aligned} \tag{47}$$

$$\begin{aligned} E(s_+ I_n^2) &\approx E(I^2) + \sqrt{\frac{2}{\pi}}\sqrt{\frac{1}{\sum_m^M d_m^2 D(I) + \frac{D(e)}{a^2}}} \\ &\quad \times [E(I^3) - E(I^2)E(I)] d_n, \end{aligned} \tag{48}$$

$$\begin{aligned} E(s_- I_n^2) &\approx E(I^2) - \sqrt{\frac{2}{\pi}}\sqrt{\frac{1}{\sum_m^M d_m^2 D(I) + \frac{D(e)}{a^2}}} \\ &\quad \times [E(I^3) - E(I^2)E(I)] d_n, \end{aligned} \tag{49}$$

where $C_3 = \sqrt{\frac{2}{\pi}}\sqrt{\frac{1}{\sum_m^M d_m^2 D(I) + \frac{D(e)}{a^2}}} D(I)$. Then we have

$$\begin{aligned} D(s_+ I_n) &= E(s_+ I_n^2) - E(s_+ I_n)^2 \\ &= E(I^2) + \sqrt{\frac{2}{\pi}}\sqrt{\frac{1}{\sum_m^M d_m^2 D(I) + \frac{D(e)}{a^2}}} \\ &\quad \times [E(I^3) - E(I^2)E(I)] d_n \\ &\quad - \left[E(I) + \sqrt{\frac{2}{\pi}}\sqrt{\frac{1}{\sum_m^M d_m^2 D(I) + \frac{D(e)}{a^2}}} D(I) d_n \right]^2 \\ &\approx E(I^2) - E(I)^2 = D(I), \end{aligned} \tag{50}$$

$$\begin{aligned} D(s_- I_n) &= E(s_- I_n^2) - E(s_- I_n)^2 \\ &= E(I^2) - \sqrt{\frac{2}{\pi}}\sqrt{\frac{1}{\sum_m^M d_m^2 D(I) + \frac{D(e)}{a^2}}} \\ &\quad \times [E(I^3) - E(I^2)E(I)] d_n \\ &\quad - \left[E(I) - \sqrt{\frac{2}{\pi}}\sqrt{\frac{1}{\sum_m^M d_m^2 D(I) + \frac{D(e)}{a^2}}} D(I) d_n \right]^2 \\ &\approx E(I^2) - E(I)^2 = D(I). \end{aligned} \tag{51}$$

There is only one pending term introduced by noise and light intensity attenuation, i.e., $\frac{D(e)}{a^2}$. It is hard for us to know its specific value. This can only be obtained empirically in order to match the experimental data to the theoretical curve as much as possible.

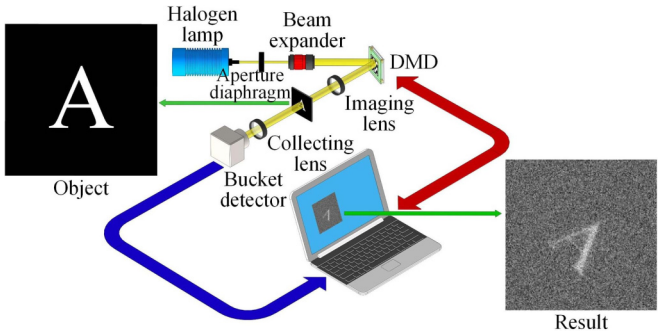

FIG. 3. Optical setup for CI. The thermal light emitted from a halogen lamp passes through an aperture diaphragm and a beam expander and illuminates a DMD. Then the modulated light is projected onto a black-and-white film (the object). The total intensity is recorded by a bucket detector.

Our experiment was based on a widely used computational GI [33,34] setup, as shown in Fig. 3. Unlike double-arm GI, it could modulate the illumination light via a programmable spatial light modulator (SLM) according to the preset patterns without the help of an array detector with spatial resolution. Since the reference patterns could be programmatically encoded onto the SLM, without the need to be experimentally acquired by the pixelated array detector, computational GI could remove the reference arm and simplify the imaging configuration. In our experimental setup, a digital micromirror device (DMD), which consists of 1024×768 micromirrors, each of size 13.68×13.68 μm^2, is used as one kind of SLM here to perform light intensity modulation. Since each of its micromirrors could be oriented either $+12^\circ$ and -12° with respect to the normal of the DMD work plane, corresponding to the bright pixel 1 or the dark pixel 0, the light would be reflected into two directions. In our experiment, the light from a halogen lamp illuminates the DMD through an aperture diaphragm and a beam expander and then the modulated patterns are projected onto an object, which is a black-and-white film printed with "A," as shown in Fig. 4(a). Its statistical data of binary values are provided in Table II. The 0-1 random patterns used occupy the central 160×160 micromirrors (pixels) of the DMD. In each pattern, 0 and 1 have the same probability of occurrence. A 1/1.8-in. charge-coupled device is used as a bucket detector to integrate the gray values of all pixels in one frame. The recovered images with 7761 frames are presented in Figs. 4(b)–4(d). From the curves shown in Fig. 5 it can be seen that the experimental data are in good agreement with the theoretical Gaussian curves.

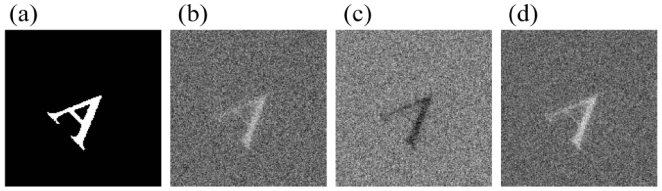

FIG. 4. Experimental results: (a) binarized image obtained by a camera and recovered (b) positive image, (c) negative image, and (d) their difference image.

TABLE II. Statistical data of binary values in the binarized image taken by a camera.

Gray value	Total number of pixels	Proportion
0	24847	97.06%
1	753	2.94%

IV. EXTENSION: CROSS-POINT–TO–STANDARD-DEVIATION RATIO

As mentioned before, the statistical curve of each gray value within a certain pixel region in the positive or negative image corresponds to a Gaussian curve. Figure 6 shows two Gaussian curves obtained from two pixel regions corresponding to two gray values. Obviously, the farther the Gaussian curves of two gray values are separated, the bigger the difference is between the two recovered gray values and the better the image quality is of the reconstruction. We can choose an appropriate measure to describe this distance, e.g., the overlapping area of two curves, denoted by Ω, which can be treated as a criterion for the reconstruction quality.

Analogously, it is easy to find that for the reconstructed images using the correlation functions, such as $G_2 = \langle SI_n \rangle$, $g_2 = \frac{\langle SI_n \rangle}{\langle S \rangle \langle I_n \rangle}$, and $\mathrm{DGI} = \langle SI_n \rangle - \frac{\langle S \rangle}{\langle S_R \rangle} \langle S_R I_n \rangle$, the conclusion that

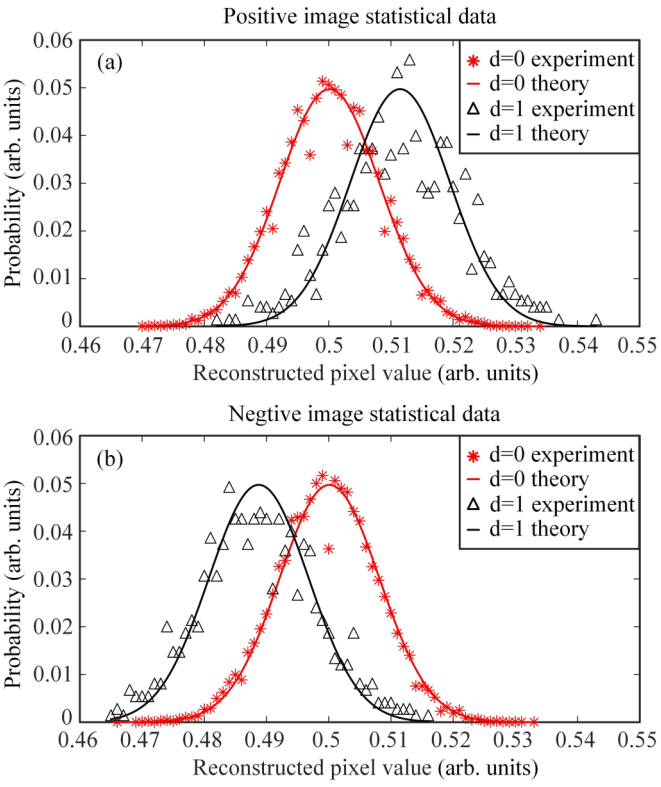

FIG. 5. Probability density function curves for the recovered pixel values, compared with theoretical Gaussian function curves. The probability density distributions and theoretical Gaussian curves of recovered pixel values falling in each pixel region where the original gray value equals 0 or 1 are shown for (a) positive and (b) negative images. Here the value of the pending term $\frac{D(e)}{a^2}$ is set equal to 120.

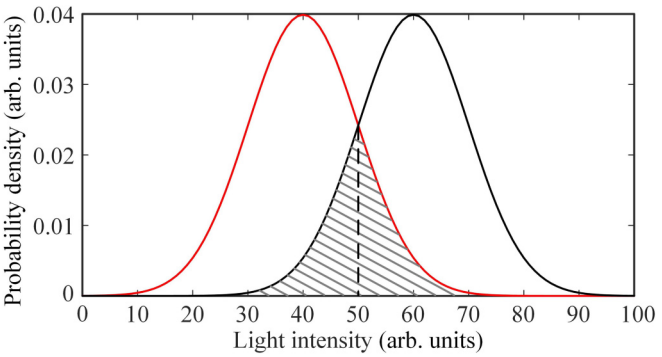

FIG. 6. Schematic diagram of two Gaussian curves corresponding to two gray values.

the reconstructed pixels in each pixel region obey a Gaussian or Gaussian-like distribution is still valid. Thereby, these functions can also use this overlapping area as the image quality measure.

Now let us calculate this overlapping area Ω. In Fig. 6 the two curves that correspond to any two original gray values ς and τ have two means, i.e., μ_1 and μ_2. Generally, as long as the algorithm can reconstruct the object image, it is obvious that there must be a linear relationship between the reconstructed image and the original image, which will be at most affected by noise. For simplicity of mathematics, we suppose the standard deviations are approximately equal, i.e., $\sigma_1 \approx \sigma_2 = \sigma$. Actually, in both simulation and experiments, we also observe that the standard deviations of the Gaussian curves for all different original gray values are very close to each other. Because the original speckle intensities are independent and identically distributed, when the number of pixels contained in each pixel region is large enough, the standard deviations of the average values of the reference patterns inside these pixel regions will also tend to the same value. Without loss of generality, we can set $\mu_1 < \mu_2$. It is easy to calculate the abscissa of the intersection of two curves, i.e., $\frac{\mu_1+\mu_2}{2}$. The shaded area in Fig. 6 is $\Omega = 2\phi(-\frac{\mu_2-\mu_1}{2\sigma})$, where $\phi(x)$ is the standard Gaussian distribution function (the integral of the standard Gaussian probability density function). Note that the area is negatively correlated with $\frac{\mu_2-\mu_1}{2\sigma}$, which is a term related to the original gray values ς and τ. If the standard deviations are assumed to be approximately equal, then the well-known formula of the contrast-to-noise ratio (CNR) [35] differs from this term $\frac{\mu_2-\mu_1}{2\sigma}$ only by a constant factor $\sqrt{2}$. To some extent, for binary objects, the CNR is a special case of the overlapping area and can be derived from the latter; thus the physical meaning of the CNR is manifested here. However, $\frac{\mu_2-\mu_1}{2\sigma}$ is not very suitable as an assessment metric of reconstruction quality for the following reasons. For the same reconstruction image, the value of $\frac{\mu_2-\mu_1}{2\sigma}$ calculated from two distant original grayscale values (such as 0 and 1) is much larger than that of two original gray values which are close to each other (e.g., 0.4 and 0.6), but it does not mean that the former result is much better than the latter. Because they are all obtained from the same recovered image, the former get a larger value since they are calculated from two original grayscale values that are much easier to resolve.

To provide a fair comparison, we will introduce an imaging quality factor named the cross-point–to–standard-deviation ratio (CSR), which is defined as

$$\mathrm{CSR} = \frac{(\mu_2 - \mu_1)/2}{\sigma}\delta, \tag{52}$$

where $\delta = \frac{1}{\tau - \varsigma}$. Since an identical linear relationship is associated with the original gray values and the means, the product between the terms $\frac{1}{\tau-\varsigma}$ and $\mu_2 - \mu_1$ eliminates the effects of the specific gray values so that the CSR values obtained by choosing any two original gray values for the reconstructed images are the same. For any two given original gray values, the larger the CSR value is, the smaller the overlapping area is, and the more obviously the two gray values are separated, the better the imaging quality is.

Next we will further consider the effect of detection noise on the imaging quality of CI in the CSR analysis. For simplicity, here we will only discuss the effect of detection noise on the positive image. The effect of detection noise on the negative image is the same. Based on Eq. (46), for two original gray values ς and τ, we have

$$\begin{aligned} \mu_1 &= C_2 + C_3\varsigma, \\ \mu_2 &= C_2 + C_3\tau. \end{aligned} \tag{53}$$

Then $\mu_2 - \mu_1 = C_3(\tau - \varsigma)$, i.e., $(\mu_2 - \mu_1)\delta = C_3$. In addition, there is

$$\sigma = \sqrt{\frac{D(s_+ I_n)}{T_+}}. \tag{54}$$

According to Eq. (50), we will obtain

$$\sigma = \sqrt{\frac{D(I)}{T_+}}. \tag{55}$$

Substituting σ into the CSR formula, we will have

$$\mathrm{CSR} = \frac{C_3}{2}\sqrt{\frac{T_+}{D(I)}}. \tag{56}$$

Since $C_3 = \sqrt{\frac{2}{\pi}}\sqrt{\frac{1}{\sum_m^M d_m^2 D(I) + \frac{D(e)}{q}}}D(I)$, obviously, the larger the variance $D(e)$ of the noise is, the smaller the CSR value is and the larger the overlapping area is, and the less the two gray values are separated, the poorer the imaging quality is. Since the above formula does not include the parameters ς and τ, i.e., any two original gray values, the same conclusion will be obtained by choosing any two original gray values ς and τ.

As mentioned above, the positive or negative image is obtained by just averaging partial reference patterns whose bucket values are above or below a threshold. The effect of using different intervals of partial reference patterns on the reconstruction quality has been reported in previous experimental studies [18,19], but without any theoretical explanation. Recently, Yang *et al.* [32] attempted to formulate the visibilities of positive and negative images with the threshold change under a basic assumption of using binary objects and patterns that fulfill Gaussian statistics. Here we will use the CSR to give a more rigorous theoretical explanation based on a more general basic assumption that the objects are of grayscale and the reference speckles obey an arbitrary distribution. For the positive image, we define a logic signal

$$s_\beta = \begin{cases} 1, & s \geqslant \beta\mu \\ 0, & s < \beta\mu. \end{cases} \tag{57}$$

The number of patterns that correspond to the bucket values larger than $\beta\mu$ is $T_\beta = T[\int_{\beta\mu}^{\infty} f_S(s)ds] = T[1 - F_S(\beta\mu)]$, where T is the total number of measurements. Then we can acquire

$$\begin{aligned} E(s_\beta Y_n) &= \frac{\lim_{s_{\max}\to\infty}\int_0^{s_{\max}}\left[\int_0^s s_\beta y_n f_{S,Y_n}(s, y_n)dy_n\right]ds}{\lim_{s_{\max}\to\infty}\int_0^{s_{\max}} s_\beta f_S(s)ds} \\ &= \frac{\int_{\beta\mu}^{\infty}\left[\int_0^s f_{S_n}(s - y_n)y_n f_{Y_n}(y_n)dy_n\right]ds}{\int_{\beta\mu}^{\infty} f_S(s)ds} \\ &\approx \frac{\int_{\beta\mu}^{\infty}\left\{\int_0^{\infty}[f_{S_n}(s) - f'_{S_n}(s)y_n]y_n f_{Y_n}(y_n)dy_n\right\}ds}{1 - F_S(\beta\mu)} \\ &= \frac{E(Y_n)[1 - F_{S_n}(\beta\mu)] - E\left(Y_n^2\right)\left[0 - f_{S_n}(\beta\mu)\right]}{1 - F_S(\beta\mu)} \\ &= \frac{E(Y_n)[1 - F_{S_n}(\beta\mu)] + E\left(Y_n^2\right)f_{S_n}(\beta\mu)}{1 - F_S(\beta\mu)}. \end{aligned} \tag{58}$$

In a similar way, we acquire the formula of $E(s_\beta Y_n^2)$:

$$E(s_\beta Y_n^2) = \frac{E\left(Y_n^2\right)[1 - F_{S_n}(\beta\mu)] + E(Y_n^3)f_{S_n}(\beta\mu)}{1 - F_S(\beta\mu)}. \tag{59}$$

Then there are

$$E(s_\beta I_n) = \frac{E(I)[1 - F_{S_n}(\beta\mu)] + aE(I^2)f_{S_n}(\beta\mu)d_n}{1 - F_S(\beta\mu)}, \tag{60}$$

$$E(s_\beta I_n^2) = \frac{E(I^2)[1 - F_{S_n}(\beta\mu)] + aE(I^3)f_{S_n}(\beta\mu)d_n}{1 - F_S(\beta\mu)}. \tag{61}$$

Thus, the CSR can be written as

$$\mathrm{CSR} = \frac{|E(s_\beta I_n)|_{d_n=\varsigma} - E(s_\beta I_n)|_{d_n=\tau}|}{2\sqrt{\frac{|E(s_\beta I_n^2)|_{d_n=\tau} - E(s_\beta I_n)^2|_{d_n=\tau}|}{T_\beta}}}\frac{1}{|\varsigma - \tau|}. \tag{62}$$

Now we will discuss the generality of the CSR to obtain the trend of the CSR changing with β without pursuing its specific values. In Eq. (62), since each gray value has little effect on the standard deviation, we set τ in the denominator equal to 0; because $E(I) \ll \mu$, the distributions of S and S_n can be considered to be approximately the same and $\beta\mu$ is not much different from μ. Then the CSR formula can be simplified to

$$\mathrm{CSR} = \frac{aE(I^2)f_S(\beta\mu)\sqrt{T}}{2[1 - F_S(\beta\mu)]^{1/2}\sqrt{D(I)}}. \tag{63}$$

Obviously, the larger the total number of measurements is, the higher the CSR value is and the better the reconstruction quality is. Apart from this, the CSR value also depends on the factor

$$g(\beta\mu) = \frac{f_S(\beta\mu)}{[1 - F_S(\beta\mu)]^{1/2}}. \tag{64}$$

We take the derivative of this factor with respect to $\beta\mu$ [there is $f_S'(\beta\mu) = 0$ under a first-order approximation]:

$$g'(\beta\mu) = \frac{\frac{1}{2}f_S^2(\beta\mu)}{[1 - F_S(\beta\mu)]^{3/2}} > 0. \tag{65}$$

It can be concluded that $g(\beta\mu)$ is an increasing function and the CSR value increases gradually as β increases. This means that the patterns which correspond to much larger bucket values (above the mean) will undoubtedly generate a positive image with a much higher quality, and vice versa for the negative image formation, which explains well the phenomenon found in previous experimental work [18,19]. It also helps to explain the inner mechanism of the previous work in super sub-Nyquist single-pixel imaging [17].

V. CONCLUSION

In summary, we have developed a probability theory to explain the formation mechanism of CI whose bucket values are binarized, based on a general model in which the targets are of grayscale, and any two thermal reference speckles are independent of each other, all following an arbitrary identical distribution. By building the joint probability density function between the bucket variable and each reference thermal speckle variable and deducing the related means and variances, we found that the positive-negative images and their difference image are all linear transformations of the object image. Provided each original grayscale value has a large enough number of pixels, then the reconstructed values falling in every pixel region of the same original gray value will obey a Gaussian distribution, no matter what kind of distribution the speckles obey. The measurement noise was also considered in analysis. We have demonstrated the validity of derived formulas through both simulation and experiments. Based on our theory, we introduced an image quality metric, the CSR, which is independent of the two original gray values chosen arbitrarily and takes into account the effect of detection noise. By using the CSR, we have provided a strict theoretical explanation for the experimental phenomenon that the patterns with much larger bucket values (above the mean) will help generate a positive image of much higher quality, and vice versa for the negative one.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant No. 61801022), the National Key Research and Development Program of China (Grant No. 2016YFE0131500), the Civil Space Project of China (Grant No. D040301), and the International Science and Technology Cooperation Special Project of Beijing Institute of Technology (Grant No. GZ2018185101).

J.L. and W.-K.Y. contributed equally to this work.

[1] T. B. Pittman, Y. H. Shih, D. V. Strekalov, and A. V. Sergienko, Optical imaging by means of two-photon quantum entanglement, Phys. Rev. A **52**, R3429(R) (1995).

[2] D. Zhang, Y.-H. Zhai, and L.-A. Wu, Correlated two-photon imaging with true thermal light, Opt. Lett. **30**, 2354 (2005).

[3] F. Ferri, D. Magatti, A. Gatti, M. Bache, E. Brambilla, and L. A. Lugiato, High-Resolution Ghost Image and Ghost Diffraction Experiments with Thermal Light, Phys. Rev. Lett. **94**, 183602 (2005).

[4] J. Xiong, D.-Z. Cao, F. Huang, H.-G. Li, X.-J. Sun, and K. Wang, Experimental Observation of Classical Subwavelength Interference with a Pseudothermal Light Source, Phys. Rev. Lett. **94**, 173601 (2005).

[5] H. Yu, R. Lu, S. Han, H. Xie, G. Du, T. Xiao, and D. Zhu, Fourier-Transform Ghost Imaging with Hard X Rays, Phys. Rev. Lett. **117**, 113901 (2016).

[6] A.-X. Zhang, Y.-H. He, L.-A. Wu, L.-M. Chen, and B.-B. Wang, Tabletop x-ray ghost imaging with ultra-low radiation, Optica **5**, 374 (2018).

[7] D.-Z. Cao, J. Xiong, and K. Wang, Geometrical optics in correlated imaging systems, Phys. Rev. A **71**, 013801 (2005).

[8] G. Scarcelli, V. Berardi, and Y. H. Shih, Phase-conjugate mirror via two-photon thermal light imaging, Appl. Phys. Lett. **88**, 061106 (2006).

[9] P. Clemente, V. Duran, V. Torres-Company, E. Tajahuerce, and J. Lancis, Optical encryption based on computational ghost imaging, Opt. Lett. **35**, 2391 (2010).

[10] W.-K. Yu, Cryptographic key distribution over a public network via variance-based watermarking in compressive measurements, Appl. Opt. **58**, 5294 (2019).

[11] W. Gong, C. Zhao, H. Yu, M. Chen, W. Xu, and S. Han, Three-dimensional ghost imaging lidar via sparsity constraint, Sci. Rep. **6**, 26133 (2016).

[12] A. Gatti, E. Brambilla, M. Bache, and L. A. Lugiato, Ghost Imaging with Thermal Light: Comparing Entanglement and Classical Correlation, Phys. Rev. Lett. **93**, 093602 (2004).

[13] F. Ferri, D. Magatti, L. A. Lugiato, and A. Gatti, Differential Ghost Imaging, Phys. Rev. Lett. **104**, 253603 (2010).

[14] W.-K. Yu, M.-F. Li, X.-R. Yao, X.-F. Liu, L.-A. Wu, and G.-J. Zhai, Adaptive compressive ghost imaging based on wavelet trees and sparse representation, Opt. Express **22**, 7133 (2014).

[15] X.-R. Yao, W.-K. Yu, X.-F. Liu, L.-Z. Li, and G.-J. Zhai, Iterative denoising of ghost imaging, Opt. Express **22**, 24268 (2014).

[16] A. M. Paniagua-Diaz, I. Starshynov, N. Fayard, A. Goetschy, R. Pierrat, R. Carminati, and J. Bertolotti, Blind ghost imaging, Optica **6**, 460 (2019).

[17] W.-K. Yu, Super sub-Nyquist single-pixel imaging by means of cake-cutting Hadamard basis sort, Sensors **19**, 4122 (2019).

[18] L.-A. Wu and K.-H. Luo, in *S. N. Bose National Centre for Basic Sciences Silver Jubilee Symposium on 75 Years of Quantum Entanglement–Foundations and Information Theoretic Applications, Kolkata, India, 2011*, edited by D. Home, G. Kar, and A. S. Majumdar, AIP Conf. Proc. No. 1384 (AIP, Melville, 2011), p. 223.

[19] K.-H. Luo, B.-Q. Huang, W.-M. Zheng, and L.-A. Wu, Nonlocal imaging by conditional averaging of random reference measurements, Chin. Phys. Lett. **29**, 074216 (2012).

[20] R. E. Meyers, K. S. Deacon, and Y. Shih, Positive-negative turbulence-free ghost imaging, Appl. Phys. Lett. **100**, 131114 (2012).

[21] M.-J. Sun, M.-F. Li, and L.-A. Wu, Nonlocal imaging of a reflective object using positive and negative correlations, Appl. Opt. **54**, 7494 (2015).

[22] S.-M. Zhao and P. Zhuang, Correspondence normalized ghost imaging on compressive sensing, Chin. Phys. B **23**, 054203 (2014).

[23] G.-L. Li, Y. Zhao, Z.-H. Yang, and X. Liu, Positive-negative corresponding normalized ghost imaging based on an adaptive threshold, Laser Phys. Lett. **13**, 115202 (2016).

[24] H. Wu, X. Zhang, Y. Shan, Z. He, H. Li, and C. Luo, Adaptive differential correspondence imaging based on sorting technique, AIP Adv. **7**, 045121 (2017).

[25] Y.-X. Li, W.-K. Yu, J. Leng, and S.-F. Wang, Pseudo-thermal imaging by using sequential-deviations for real-time image reconstruction, Opt. Express **27**, 35166 (2019).

[26] H.-C. Liu, H. Yang, J. Xiong, and S. Zhang, Positive and Negative Ghost Imaging, Phys. Rev. Appl. **12**, 034019 (2019).

[27] M.-F. Li, Y.-R. Zhang, X.-F. Liu, X.-R. Yao, K.-H. Luo, H. Fan, and L. A. Wu, A double-threshold technique for fast time-correspondence imaging, Appl. Phys. Lett. **103**, 211119 (2013).

[28] M.-F. Li, Y.-R. Zhang, K.-H. Luo, L.-A. Wu, and H. Fan, Time-correspondence differential ghost imaging, Phys. Rev. A **87**, 033813 (2013).

[29] J. Wen, Forming positive-negative images using conditioned partial measurements from reference arm in ghost imaging, J. Opt. Soc. Am. A **29**, 1906 (2012).

[30] W.-K. Yu, X.-R. Yao, X.-F. Liu, L.-Z. Li, and G.-J. Zhai, Ghost imaging based on Pearson correlation coefficients, Chin. Phys. B **24**, 054203 (2015).

[31] X.-R. Yao, X.-F. Liu, W.-K. Yu, and G.-J. Zhai, Correspondence imaging based on correlation coefficient, Chin. Opt. Lett. **13**, 010301 (2015).

[32] H. Yang, S. Wu, H.-B. Wang, D.-Z. Cao, S.-H. Zhang, J. Xiong, and K. Wang, Probability theory in conditional-averaging ghost imaging with thermal light, Phys. Rev. A **98**, 053853 (2018).

[33] J. H. Shapiro, Computational ghost imaging, Phys. Rev. A **78**, 061802(R) (2008).

[34] Y. Bromberg, O. Katz, and Y. Silberberg, Ghost imaging with a single detector, Phys. Rev. A **79**, 053840 (2009).

[35] K. W. C. Chan, M. N. O'Sullivan, and R. W. Boyd, Optimization of thermal ghost imaging: High-order correlations vs. background subtraction, Opt. Express **18**, 5562 (2010).