

Quantitative characterization of super-resolution infrared imaging based on time-varying focal plane coding

X. Wang

xrwang@mail.xidian.edu.cn

School of Physics and Optoelectronic Engineering, Xidian University, Xi'an, Shaanxi, 710071, China

Y. Yuan

Shaanxi Zhong Tian Rocket Technology Limited Company, Xi'an, Shaanxi, 710025, China

J. Zhang

School of Physics and Optoelectronic Engineering, Xidian University, Xi'an, Shaanxi, 710071, China

Y. Chen

School of Physics and Optoelectronic Engineering, Xidian University, Xi'an, Shaanxi, 710071, China

Y. Cheng

Science and Technology on Low-Light-Level Night Vision Laboratory, Xi'an, Shaanxi, 710065, China

High resolution infrared image has been the goal of an infrared imaging system. In this paper, a super-resolution infrared imaging method using time-varying coded mask is proposed based on focal plane coding and compressed sensing theory. The basic idea of this method is to set a coded mask on the focal plane of the optical system, and the same scene could be sampled many times repeatedly by using time-varying control coding strategy, the super-resolution image is further reconstructed by sparse optimization algorithm. The results of simulation are quantitatively evaluated by introducing the Peak Signal-to-Noise Ratio (PSNR) and Modulation Transfer Function (MTF), which illustrate that the effect of compressed measurement coefficient r and coded mask resolution m on the reconstructed image quality. Research results show that the proposed method will promote infrared imaging quality effectively, which will be helpful for the practical design of new type of high resolution infrared imaging systems.

[DOI: <http://dx.doi.org/10.2971/jeos.2014.14043>]

Keywords: Infrared imaging system, compressed sensing, focal plane coding, super-resolution, modulation transfer function

1 INTRODUCTION

For the restriction of infrared sensors processing technology, compared to visible detector, infrared detector generally have focal plane array of limited number and larger pixel size, which cannot satisfy military and civilian imaging requirement of high resolution [1, 2]. Therefore, it is necessary to introduce new concept, new theory or new imaging mechanism to design new type of infrared imaging system.

Compressive sensing (CS) involves exploiting the sparsity and compressibility of an image in some transform domain so that one can utilize fewer measurements than the ones required for conventional imaging, yet the imaging can be reconstructed with minimal loss information [3, 4]. Focal plane coding (FPC) is a means for intelligent sensing and mapping of optical pixel to enable efficient and faithful digital image reconstruction. Making a combination of CS and FPC in infrared camera, we may obtain a high resolution image from a few measured samples. Many researchers have studied how to improve optical imaging resolution based on the above two points. Andrew D. Portnoy et al. [5, 6] use focal plane coding to produce non-degenerate data between multiple apertures, and then sub-aperture data is integrated to form a single high resolution image. A compressive imaging system using multiplex and multi-channel measurements with static focal plane coding is described in Ref. [7] for compressed sensing, which

demonstrates that this system can achieve up to 50% compression with conventional benchmarking images. The most noteworthy application of compressed sensing is the single-pixel camera designed by Baraniuks team in Ref. [8], the single-pixel camera acquires a recognizable image with a resolution comparable to N pixels. Inspired by the idea of the single-pixel camera, CS has been used for passive millimeter-wave imaging to significantly reduce the imaging time and produce high-fidelity images in Ref. [9]. Refs. [10, 11] present a novel approach for improving infrared imaging resolution by the use of CS. The image sensor measures the compressed samples of the observed image through a coded aperture mask placed on the focal plane of the optical system, the same scene could be sampled repeatedly by using multiplexing technology, and then the image reconstruction can be performed from these samples using an optimal algorithm. Inspired by the above research results, we propose a super-resolution infrared imaging method based on time-varying focal plane coding, and introduce the MTF for quantitative evaluation of image quality.

This paper is organized as follows. Section 2 describes the basic theory of CS briefly. In Section 3 we present a super-resolution infrared imaging method based on time-varying focal plane coding, mainly including the coding

sampling strategy and its corresponding reconstruction method. Experimental results and quantitative evaluation are given in Section 4. Section 5 concludes this paper.

2 COMPRESSED SENSING

The basic idea of CS theory is that when the image of interest is sparse or highly compressible in some basis (i.e., most basis coefficients are small or zero-valued), relatively few well-chosen observations suffice to reconstruct the most significant nonzero components. In particular, judicious selection of the type of image transformation introduced by measurement systems may dramatically improve our ability to extract high quality images from a limited number of measurements. In this section we review the intuition and theory underlying these ideas. By designing optical sensors to collect measurements of a scene according to CS theory, we can use computational methods to infer critical scene structure and content.

The CS sampling model is described as

$$y = \Phi f = \Phi \Psi x = \Theta x \quad (1)$$

where $f \in R^N$ is an unknown signal, which could be sparsely represented as $f = \Psi x$ in an orthonormal basis Ψ . If there are only K ($K \ll N$) non-zero components of x , f is defined as being K -sparse. Φ denotes a $M \times N$ ($M \ll N$) matrix called measurement matrix. $\Theta = \Phi \Psi$ is a sensing matrix and $y \in R^M$ is an observation vector. The CS theory indicates that, subject to a Restricted Isometry Property (RIP) condition [12, 13] on the observation matrix Θ , high dimensional vectors f can be recovered from a much smaller dimensional observation y with the probability close to 1.

The observation matrix Θ should satisfy the RIP of order $3k$ if, for $T \subset \{1, 2, \dots, n\}$ and Θ_T , a submatrix obtained by retaining the columns of Θ corresponding to the indices in T , there exists a constant $\delta_{3k} \in (0, 1/3)$ such that for all $z \in \mathbb{R}^{|T|}$,

$$(1 - \delta_{3k}) \|z\|_2^2 \leq \|\Theta_T z\|_2^2 \leq (1 + \delta_{3k}) \|z\|_2^2 \quad (2)$$

holds for all subsets T with $|T| \leq 3k$. Intriguingly, many kinds of random matrices meet RIP with high probability. The equivalent condition of RIP, referred to as incoherence, requires that the rows $\{\phi_j\}$ of measurement matrix Φ cannot sparsely represent the columns $\{\psi_i\}$ of orthonormal basis Ψ (and vice versa). Refs. [14]–[17] proved that if Ψ is a Gaussian random matrix, random binary Bernoulli distribution matrix, partial Fourier matrix, local Hadamard matrix, Toeplitz matrix and so on, Θ can meet RIP in all probability. Finally, to solve the l^1 optimization problem described in Eq. (3) will get a high precision estimated value \hat{f} of an original signal f .

$$\begin{aligned} \hat{x} &= \underset{x}{\operatorname{argmin}} \|x\|_1 \quad \text{subject to} \quad \Theta x = y \\ \hat{f} &= \Psi \hat{x} \end{aligned} \quad (3)$$

where $x \in R^N$ is the sparse form of the original signal f , and x only has K ($K \ll N$) non-zero components. Sensing matrix Θ denotes an $M \times N$ matrix and $y \in R^M$ is an incomplete observation vector. The dimension M of y is smaller than the dimension N of x ($M \ll N$). There are many optimal methods

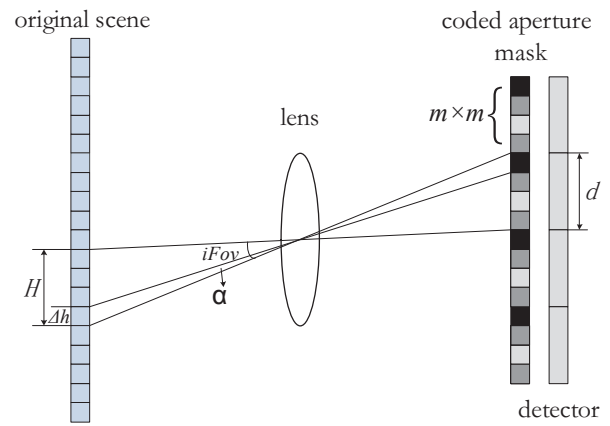


FIG. 1 Time-varying focal plane coding imaging model.

such as Basis Pursuit (BP), Matching Pursuit (MP), Total Variation (TV), and Gradient Projection for Sparse Reconstruction (GPSR) [18] for signal reconstruction.

3 SUPER-RESOLUTION INFRARED IMAGING MODEL

For typical infrared imaging system, the pixel pitch of infrared sensors ranges from 15 m to 40 m [19], which is larger than that of their visible-light counterparts and cannot satisfy the resolution requirements of many science experiments. To improve infrared imaging quality, we design a time-varying coded mask based on CS and set it in front of the focal plane of the optical system closely. Finally, a super-resolution image can be obtained by introducing reconstruction algorithm from a small amount of samples.

As denoted in Figure 1, time-varying focal plane coding imaging model is given for one moment. Conventional infrared imaging resolution is determined by the detector pixel size ($d \times d$) of the focal plane array. Assuming that the angular resolution is $iFov$ and the scene sampling distance is defined as H . After $m \times m$ coded mask sub-array for each pixel and reconstruction algorithm are adopted, final infrared imaging resolution is dependent on the pixel size of the coded mask rather than that of the focal plane array. Theoretically, the angular resolution of the reconstructed image become $\alpha = iFov/m$ and the scene sampling distance of the reconstructed image is $\Delta h = H/m$, thus infrared imaging resolution can be promoted m times.

According to CS theory, the practical design of coded mask choose Gauss random matrix with normal distribution $N(0, 1)$ as the measurement matrix. As illustrated in Figure 2(a), a 12×12 coded mask array corresponds to 3×3 pixels on the sensor. Every pixel is masked with the same 4×4 sub-array pattern. Ref.[20] proved that multi-value mask is superior to conventional binary mask in imaging system. Thus, a multi-value time-varying coded mask is adopted in our infrared imaging system, in which “0” or “1” elements denote the light is fully blocked or passed, and the value between 0 and 1 denotes the light is partly blocked. The white area enables the light to pass through the mask while the black area disables it,

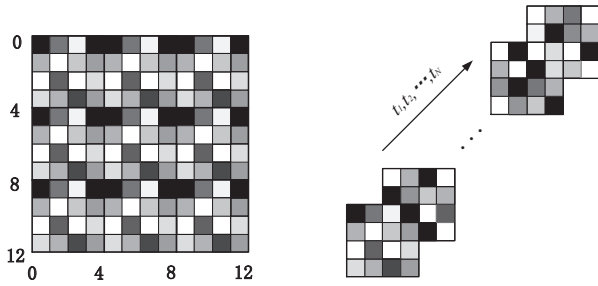


FIG. 2 Schematic of time-varying focal plane coding mask.

the different gray-scale areas represent the light which partly passed in Figure 2. Each sensor pixel records the light intensity of the incoming light beam encoded by a 4×4 mask. In fact, these records are the compressed samples of the observed object.

A measurement matrix corresponds to a compressed measurement, time varying coding pattern can obtain multiple compressed measurement of original scene. Figure 2(b) shows that the variation of single pixels focal plane coding mask with time t . At the same moment, sub-arrays of every pixels coded mask are the same Gauss random matrix, while different random matrices are used for different moments, which obey the same Gaussian distribution. Compared to spatially multiplexing technology used to obtain multiple samples, the time-varying technique in this paper will reduce the volume and mass of an infrared imaging system significantly.

Suppose that Δt is the time step of different moments of coded mask, M is the measurement times, t is the total detection time of all compressed measurements. The variable r is defined as the compressed measurement coefficient which equals to the ratio of M to coded mask resolution ($m \times m$), $r \in [0, 1]$. For m and Δt being fixed, the total detection time t become larger with an increment of r . Its mathematical form is given as follows

$$t = M \cdot \Delta t \quad (4)$$

$$M = r \cdot (m \times m) \quad (5)$$

From Eq. (4) and Eq. (5), M measurements are obtained after time t , thus each detector pixel can achieve a compressed measurement vector y which has M elements. This process can be written as

$$y = \Phi \cdot f \quad (6)$$

its corresponding matrix form is expressed as

$$\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_M \end{bmatrix} = \begin{bmatrix} \Phi_1 \\ \Phi_2 \\ \dots \\ \Phi_M \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_{m \times m} \end{bmatrix} \quad (7)$$

where $\phi_i (i = 1, 2, \dots, M)$ denotes the transformed $1 \times (m \times m)$ vector resulting from measurement matrix corresponding to the i th compressed measurement. f is the column vector of light field intensity in a pixel area. The same reconstruction processes of all the pixels are performed independently. For the trade-off between computational

time and the recovery accuracy, the Gradient Projection for Sparse Reconstruction (GPSR) algorithm is used for image reconstruction, Eq. (3) can be formulated as an optimization problem where the objective function is expressed as a combination l^1 and l^2 minimization program, as described in Eq. (8).

$$\begin{aligned} \hat{x} &= \arg \min \frac{1}{2} \|y - \Phi \Psi x\|_2^2 + \tau \|x\|_1 \\ \hat{f} &= \Psi \hat{x} \end{aligned} \quad (8)$$

where τ is a nonnegative parameter and $\tau = 0.02 \|\Theta^T y\|_\infty$ in this paper. Formally, by introducing vectors u and v , make the substitution,

$$x = u - v, u \geq 0, v \geq 0 \quad (9)$$

These relationships are satisfied by $u_i = (x_i)_+$ and $v_i = (-x_i)_+ (i = 1, 2, \dots, n)$, where $(\cdot)_+$ denotes the positive-part operator defined as $(x)_+ = \max\{0, x\}$. So Eq. (8) can be rewritten as the following standard bound-constrained quadratic program(BCQP):

$$\begin{aligned} \min_{u,v} \quad & c^T z + \frac{1}{2} z^T B z \equiv F(z) \\ \text{subject to} \quad & z \geq 0 \end{aligned} \quad (10)$$

$$\begin{aligned} \text{where } z &= [u; v], b = \Theta^T y, c = \tau l_2 n + \begin{bmatrix} -b \\ b \end{bmatrix}, \\ l_n &= [1, 1, \dots, 1]^T, B = \begin{bmatrix} \Theta^T \Theta & -\Theta^T \Theta \\ -\Theta^T \Theta & \Theta^T \Theta \end{bmatrix}. \end{aligned}$$

The two-step gradient projection method defines its iteration z^{k+1} from the previous iteration z^k as,

$$\begin{aligned} w(k) &= (z^{(k)} - \alpha^{(k)} \nabla F(z^{(k)}))_+ \\ z^{(k+1)} &= z^{(k)} + \beta^{(k)} (w^{(k)} - z^{(k)}) \end{aligned} \quad (11)$$

where $w^{(k)}$ is a temporary variable, $\alpha^{(k)} > 0$ and $\beta^{(k)} \in [0, 1]$. Detailed derivation and discussion can see Ref. [18].

4 SIMULATION RESULTS

To verify the feasibility of the proposed super-resolution imaging method based on time varying focal plane coding, a tank image of 512×512 pixels is taken as an original scene. Discrete Cosine Transform(DCT) is used to form a sparse representation, and a 32×32 detector array is adopted to sample the original scene, and each detector element cover 16×16 pixels of an original scene.

4.1 Reconstruction of tank image

First, the effect of compressed measurement coefficient r on the quality of the reconstructed image is discussed. It is assumed that a 8×8 coded mask sub-array for each detector pixel is adopted, the reconstructed images for different $r = 0.8, 0.4$ and 0.2 are obtained and shown in Figures 3(b), (c) and (d), respectively. The resolution of the reconstructed image is 256×256 pixels. The PSNR is selected to evaluate the variation of image quality with the coefficient r , the relationship of the PSNR with r is illustrated in Figure 4. It can be seen

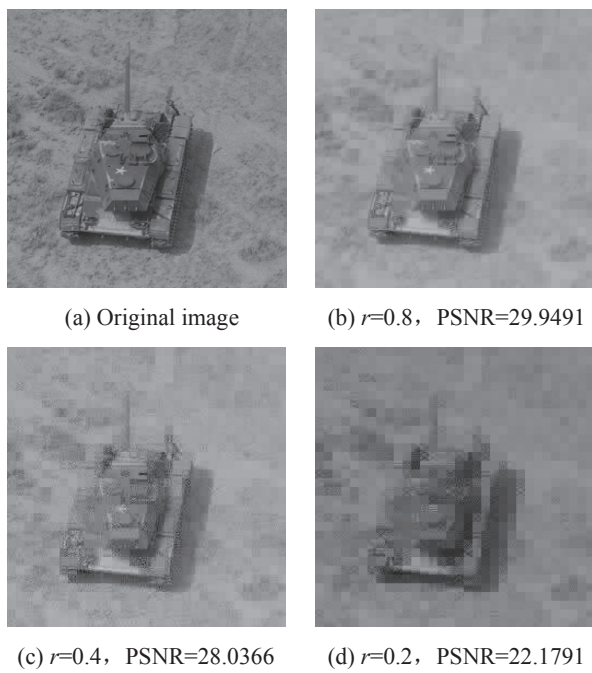


FIG. 3 Reconstructed images for different r .

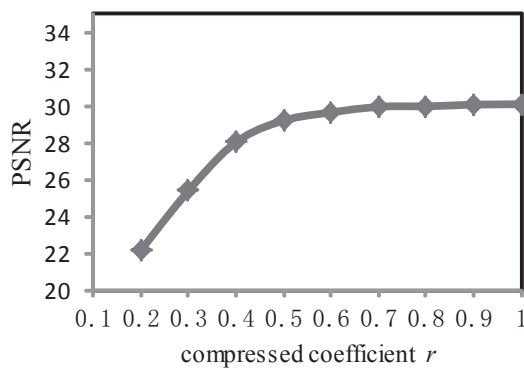


FIG. 4 PSNR value as a function of r for different reconstruction images.

that the reconstructed image quality can be promoted significantly as r become larger. The largest PSNR can be achieved at 30db. However, for a fixed coded mask, an increment of compressive measurement ratio r will lead to longer sampling time.

It is assumed that the coefficient $r = 0.75$ remain unchanged, we set $m = 4$, $m = 8$ and $m = 16$, namely 4×4 , 8×8 and 16×16 coded masks for each pixel are used to sample the original scene, respectively. Their corresponding reconstructed images are shown in Figures 5(b), (c) and (d), respectively. Figure 5(a) is a downsampling image obtained by using a 32×32 detector array without coded mask.

It can be observed that the reconstructed image quality has been improved greatly, and more details of original scene are recovered with m being larger. To quantitatively assess the quality of the reconstructed image, the MTF is introduced.

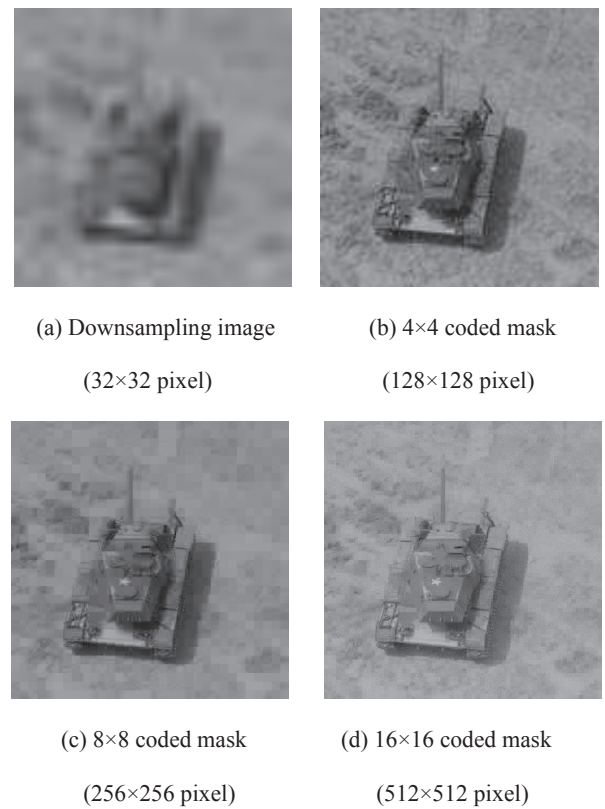


FIG. 5 Reconstructed images for different m .

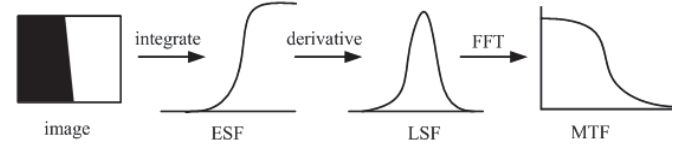


FIG. 6 Schematic of MTF measurement method based on the slanted-edge.

4.2 Quantitative evaluation of image quality based on MTF

MTF is one of the key indicators to characterize the signal transfer characteristics of an imaging system as a function of spatial frequency in terms of linear response theory [21]. Various methods have been proposed to determine the MTF of an imaging system, the slanted-edge measurement method to calculate the MTF is used in this paper [22]. The schematic of MTF by the slanted-edge measurement method is shown in Figure 6.

MTF gives an idea of a deeper and more objective conception about resolution and falls from one to zero by increasing spatial frequency. In this paper, we will measure the spatial frequency by cycles per pixel, taking one dark and one light line as a cycle. In order to connect the MTF with the usual limiting resolution, the spatial frequency corresponding to minimum contrast ratio of human eye 0.05 is taken as the limiting resolution.

A slanted-edge image of 1024×1024 pixel is taken as an original scene. Assume that $m = 8$ remain unchanged, then the original scene is sampled for $r = 0.25$, $r = 0.50$ and $r = 0.75$, respectively. Their corresponding reconstructed images are

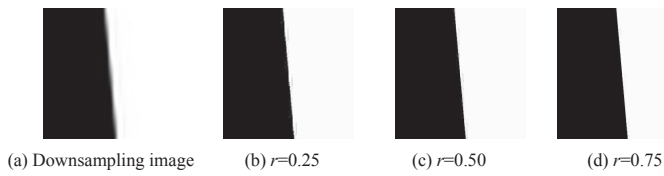


FIG. 7 Slanted-edge images with different r .

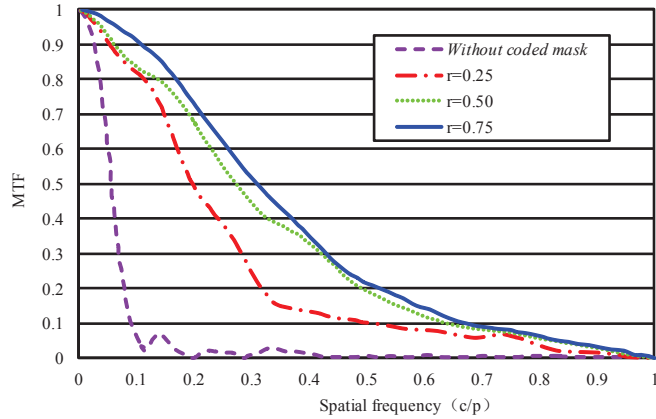


FIG. 8 MTF curves measured for different r based on slanted-edge images.

shown in Figures 7(b), (c) and (d), respectively. Figure 7(a) is the downsampling image obtained by using a 128×128 detector array without coded mask. The MTF curves of the slanted-edge images with different r values are shown in Figure 8. In the case of conventional imaging without coded mask, the sampled image is degraded seriously, and the image contrast quickly falls off with an increment of the spatial frequency. For $r = 0.25$, the recovered image contrast is well reproduced compared with the downsampling image. Further, the quality of the reconstructed image can be improved for $r = 0.5$ and $r = 0.75$. When $r = 0.25$, $r = 0.50$ and $r = 0.75$, the limiting resolution are $0.775c/p$, $0.797c/p$, $0.825c/p$, and the MTF at Nyquist frequency are 0.1121, 0.2008, 0.2175, specifically listed in Table 1. For a fixed m , the quality of reconstructed image improves with the increase of r . Both the limiting resolution and the MTF at Nyquist frequency improves significantly compared with conventional imaging mode without coded mask.

Item	limiting resolution (c/p)	MTF at Nyquist
without mask	0.105	0.0047
$r = 0.25$	0.775	0.1121
$r = 0.5$	0.797	0.2008
$r = 0.75$	0.825	0.2175

TABLE 1 Image quality characterization for different r

When $r = 0.75$ remain unchanged, then the original scene is sampled for $m = 4$, $m = 8$ and $m = 16$, respectively. Their corresponding reconstructed images are shown in Figures 9(b), (c) and (d), respectively. Figure 9(a) is the downsampling image obtained by using a 64×64 detector array without coded mask. MTF curves of slanted-edge images with different m values are shown in Figure 10. Similarly, in the case of conventional imaging without coded mask, the sampled image

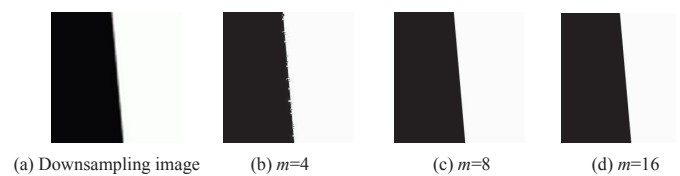


FIG. 9 Slanted-edge images for different m .

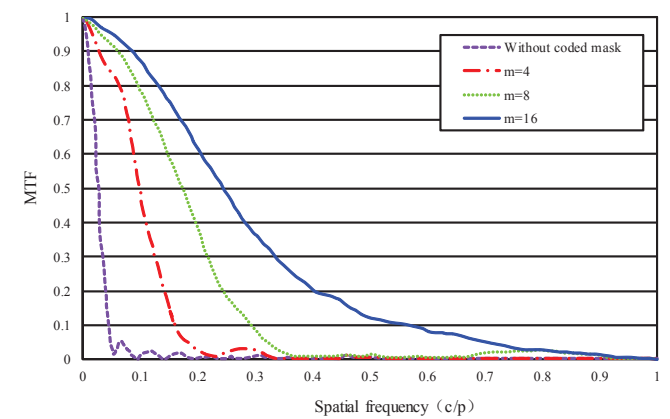


FIG. 10 MTF curves measured for different m based on slanted-edge images.

is degraded seriously, and the image contrast quickly falls off with an increment of the spatial frequency. For $m = 4$, the recovered image contrast is well reproduced compared with the downsampling image. Further, the quality of the reconstructed image can be improved for $m = 8$ and $m = 16$. When $m = 4$, $m = 8$ and $m = 16$, the limiting resolution are $0.245c/p$, $0.330c/p$, $0.697c/p$ and the MTF at Nyquist frequency are 0.0089, 0.0163, 0.1220, specifically shown in Table 2. For a fixed r , the quality of reconstructed image improves with the increase of m . Both the limiting resolution and the MTF at Nyquist frequency improves significantly compared with conventional imaging mode without coded mask.

Item	limiting resolution (c/p)	MTF at Nyquist
without mask	0.050	0.0062
$m = 4$	0.245	0.0089
$m = 8$	0.330	0.0163
$m = 16$	0.697	0.1220

TABLE 2 Image quality characterization for different m

5 CONCLUSION

We have presented a super-resolution infrared imaging method based on time-varying focal plane coding for increasing image resolution. This method can obtain high quality images by using low resolution detector array. Time-varying focal plane coding scheme makes it possible to sample enough compressed measurement by one detector and thereby reduce the volume and mass of an infrared image system. Moreover, because the coded mask is not fixed and would be usable in many applications by changing the coding strategy. Both the PSNR and the MTF have been introduced to quantitatively evaluate the reconstruction image quality, and simulation

results indicate that the proposed method can obtain more information of an original scene and improve image quality significantly compared with conventional imaging mode without coded mask. The optical experiments based on the proposed method will be an important part of our further research plan. The practical exposure time will be obtained in next step. Exposure time will be shorter and the Signal to Noise Ratio (SNR) will be reduced in the case of non-steady objects.

6 ACKNOWLEDGEMENTS

This research is supported by the National Natural Science Foundation of China (61007014, 61377007). We express our sincere appreciation for reviewers valuable comments.

References

- [1] F. Jinxiang, and Y. Yanjun, "Development in new concepts and new schemes for military infrared imaging systems," *Infrared Laser Eng.* **40**, 1-6 (2011).
- [2] F. Jinxiang, and Y. Yanjun, "Development trends of infrared detecting technology," *Infrared Laser Eng.* **41**, 3145-3153 (2012).
- [3] D. L. Donoho, "Compressed sensing," *IEEE T. Inform. Theory* **52**, 1289-1306 (2006).
- [4] Y. Tsaig, and D. L. Donoho, "Extensions of compressed sensing," *Signal Process* **86**, 549-571 (2006).
- [5] A. D. Portnoy, N. P. Pitsianis, D. J. Brady, J. Guo, M. A. Fiddy, M. R. Feldman, and R. D. Te Kolste, "Thin digital imaging systems using focal plane coding," *Proc. SPIE* **6065**, 60650F (2006).
- [6] A. D. Portnoy, N. P. Pitsianis, X. Sun, and D. J. Brady, "Multichannel sampling schemes for optical imaging systems," *Appl. Optics* **47**, B76-B85 (2008).
- [7] N. Pitsianis, D. Brady, A. Portnoy, X. Sun, T. Suleski, M. Fiddy, M. Feldman, and R. TeKolste, "Compressive imaging sensors," *Proc. SPIE* **6232**, 62320A (2006).
- [8] R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Proc. Mag.* **25**, 83-91 (2008).
- [9] N. Gopalsami, S. Liao, T. W. Elmer, E. R. Koehl, A. Heifetz, A. C. Raptis, L. Spinoulas, and A. K. Katsaggelos, "Passive millimeter-wave imaging with compressive sensing," *Opt. Eng.* **51**, 091614-1 (2012).
- [10] L. Xiao, K. Liu, D. Han, and J. Liu, "Focal plane coding method for high resolution infrared imaging," *Infrared Laser Eng.* **40**, 2065-2070 (2011).
- [11] L.-L. Xiao, K. Liu, D.-P. Han, and J.-Y. Liu, "A compressed sensing approach for enhancing infrared imaging resolution," *Opt. Laser Technol.* **44**, 2354-2360 (2012).
- [12] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *C. R. Math.* **346**, 589-592 (2008).
- [13] J. Haupt, and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE T. Inform. Theory* **52**, 4036-4048 (2006).
- [14] E. J. Candes, and T. Tao, "Decoding by linear programming," *IEEE T. Inform. Theory* **51**, 4203-4215 (2005).
- [15] E. J. Candes, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun Pur. Appl. Math.* **59**, 1207-1223 (2006).
- [16] G.-M. Shi, D.-H. Liu, D. Gao, Z. Liu, J. Lin, and L.-J. Wang, "Advances in theory and application of compressed sensing," *Acta Electron.* **37**, 1070-1081 (2009).
- [17] W. U. Bajwa, J. D. Haupt, G. M. Raz, S. J. Wright, and R. D. Nowak, "Toeplitz-structured compressed sensing matrices," in *Proceedings of Statistical Signal Processing, 2007. SSP'07. IEEE/SP 14th Workshop on*, 294-298 (IEEE, Madison, 2007).
- [18] M. A. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Top. Signa.* **1**, 586-597 (2007).
- [19] C. Bo-liang, "Development state of IRFPA imaging device [J]," *Infrared Laser Eng.* **1**, 000 (2005).
- [20] X.-P. Shao, C. Zhong, J. Du, and C.-C. Rao, "Super-resolution imaging method using multi-value compressed coded aperture," *J. Optoelectronics Laser* **6**, 032 (2012).
- [21] Y. Li, and B. He, "Quantitative evaluation of image quality of CCD subpixel imaging using MTF," *Infrared Laser Eng.* **42**, (2013).
- [22] F. B. Xu Baoshu, and S. Zelin, "Modulation Transfer Function Measurement Method of Electro-optical Imaging System," *Acta Optica Sin.* **31**, 1111004 (2011).