Abstract—In this work, high performance Carbon Nanotube (CNT) based Infrared (IR) photodetector was used in single pixel camera, which worked in room temperature without cryogenics cooling. The compressive sensing algorithm limited IR image in low resolution due to iterative computational method. In this paper, a novel model containing smooth and edge components was introduced to reconstruct high resolution IR image using single low resolution IR image. Compared to multiple scenes or video sequences super resolution method, the proposed algorithm is applicable on real-time camera or video system. The overall camera system can get high resolution IR image based on single CNT based IR photodetector, make IR camera better measurement accuracy and observe more details at long distance.

I. INTRODUCTION

The infrared (IR) radiation is invisible radiant energy distinct from visible light which provides distinct information of objects. It covers wide wavelength from sub micrometer to thousands micrometer. The IR imaging can avoid the illumination changes problems and work even in dark night. The thermal radiation distributions of objects are one of most popular research areas [1]. Meanwhile rather than studying an object’s thermal IR emission, there are a great interests in imaging objects reflection of sun or artificial near-infrared (NIR) radiation. It reveals the reflection and absorption of desired object. The range of NIR application is huge from scientific and artistic perspective, including observation of health of vegetation and soil composition [2], private and public properties security, medical visualization of veins, body injuries and identification of diseases [3], astronomical imaging and remote sensing, night vision and navigating units in military. The infrared photodetector can be classified as thermal detector and quantum photon detector. The former response is slow due to the thermal equilibrium of detector, while the latter is temperature dependent. The photon photodetector needs operate at ultra low temperature to reduce thermal noise [4]. The performance of intrinsic IR detectors, HgCdTe photodiodes is characterized by high optical absorption coefficient and quantum efficiency and relatively low thermal generation rate compared to extrinsic photodetectors and man-made structure, QWIP, AlGaAs IR sensors [5]. The extrinsic photon IR detectors require more cooling than intrinsic photon detectors. The low dimensional materials attract more attention since it has distinct electron transport. The carbon nanotube has ballistic transport with no scattering from optical or acoustic phonons in ideally with the assumption of the traveled length greater than the optical phonon mean free path and the electron energy greater than the critical optical phonon emission energy [6]. Meanwhile, the another low dimensional material graphene has a remarkable electron mobility at room temperature which is nearly independent of temperature between 10 K and 100 K [7]. The appearance of nano-materials brings a potential way to overcome the limitations of slow response of thermal detectors and cooling system of photo detectors. Especially, the carbon nanotube and graphene are excellent candidates for infrared sensing from the aspect of thermal noise removement.

The eventually goal of infrared sensor or photodetector is to get infrared imager. In conventional focal plane array detectors (visible, infrared SLR camera), there are at least four factors inducing low resolution (LR) image, including optical distortions (out of focus, diffraction limit), motion blur due to limited shutter speed, noise that occurs within the sensor or during sampling [8]. As shown in Fig. 1, the recorded image usually suffers from distortion, blur, noise, and aliasing effects.

![Fig. 1: LR image acquisition system](image)

In the CNT IR photodetector novel imaging system, it utilizes compressive sensing different from Nyquist sampling.
method to avoid large sensor array fabrication which is hard to reach in nano fabrication [9]. However, the compressive sensing recovery relies on the computational ability of hardware. It leads to relatively low resolution recovery image. IR images or videos with super resolution are desired and often required so as to get more details of object in scientific research and civilian applications. The most directly way to increase resolution is to increase the chip size (together with lens changes) or reduce the pixel size to increase the number of pixels per unit area [8]. Both techniques are dependent on microfabrication manufacturing. The disadvantage of former method is to increase capacitance which makes the readout circuit in low charge transfer rate. It also requires a new high precision optics [10]. For the latter solution, reducing the pixel size will also decrease the amount of light coming to each sensor so that the shot noise will degrade the image quality. The minimum active pixel area is around 40 μm² for a 0.35 μm CMOS process.

The resolution also can be enhanced using signal processing techniques, referred to super resolution. The recovery high resolution image is obtained from multiple or single low resolution images observed. The multiframe SR reconstruction has longest history and it involves three steps, image registration, interpolation and restoration. The most difficulty is to estimate the motion of LR input frames corresponding to reference frame, also named image registration. The typical method is to look up for interest points in the low-resolution image set, then use robust methods to estimate the point correspondences and compute homographies between images. In [11], the iterative method was used to estimate registration parameters, shifts and rotation. The block matching is also applied to register input images in [12]. The Bayesian method searched a continuous space of shift and rotation together with conventional MAP reconstruction algorithm to estimate the high resolution. The registration and estimation could be in a joint framework. In [13], the joint MAP estimation algorithm captures the dependence between LR image registration and HR image estimation although it may introduce overfitting problem. Compared to multiple images based super-resolution, single image super-resolution is more applicable since there is only one low-resolution image required, especially for portable applications.

In this paper, several single image based super resolution approaches were compared firstly. A model based single image super resolution was proposed to enhance image quality. We modeled the smooth components of an image using kernel Hilbert space (RKHS) and the edges using approximated Heaviside functions. In order to reduce computation, selected image patches were used for high resolution computing. The reconstruct high resolution images were also compared to low resolution in experimental results.

II. SINGLE IMAGE SUPER RESOLUTION

A. Observation Modeling

The digital imaging system suffers form hardware limitations, acquiring images with various kinds of degradations, the optical distortion, the motion blur due to aperture time and the sensor noise. Finally the frames captured by the low resolution imaging system are blurred, aliasing and noisy versions of the underlying true scene. As shown in Eq. 1, \( \mathbf{X} = [x_1, x_2, \ldots, x_N] \) denote the desired high resolution image, which is sampled above Nyquist sampling frequency. \( \mathbf{Y} \) is the LR images observed. \( M \) describes the motion information for LR image. \( B \) is the blur models effect and \( D \) is the down sampling operator. \( n \) represents the noise model. The LR image is denoted as \( Y(s,t), (s = 1, \ldots, N_1, t = 1, 2, \ldots, N_2) \). \( N_1, N_2 \) are horizontal and vertical direction resolution of LR images. The parameters \( L_1, L_2 \) are down sampling factors for each direction respectively and \( N = N_1 N_2 L_2 \). In mathematically, the motion warp matrix size is \( L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2 \), blur matrix \( B \) is \( L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2 \), while subsampling matrix \( D \) is \( (N_1 N_2)^2 \times L_1 N_1 L_2 N_2 \) size.

\[
\mathbf{Y} = \text{DBM} + \mathbf{n}, \\
\mathbf{X} = [x_1, x_2, \ldots, x_N]
\]

B. Direct Interpolation and Learning based Super Resolution

There are two kinds of basic interpolation to get high resolution image using low resolution observation. The nearest neighbor interpolation models the unknown point by its nearest neighbor point [14]. For each point on the HR grid, the closest known LR pixel is selected and the value of that pixel is simply used as the value at the point in HR grid. The disadvantage is jaggy effect on the HR images although it is a fast and direct algorithm. Bicubic interpolation utilized a cubic kernel to interpolate but create blur effect [15]. The edge directed interpolation method is presented in [16] for super resolution. It estimated the local covariance coefficients from a single low resolution image, and then applied the same coefficients to reconstruct high-resolution image. The contourlet transform and wavelet based linear interpolation is proposed in [17]. The directional filtering and data fusion was used to edge guided nonlinear interpolation to preserve sharp edges and reduce ring artifacts in [18]. The another single image based super resolution is popular statistics based methods. It utilized statistical edge dependency information as in low and high resolution images in [19]. The learning based method operates by building a model from example, learns from and make predictions on data. It is also a powerful tool for image super resolution, although it requires two large training data sets. In order to get super resolution, a novel profile of image gradient was described as the shape and sharpness in recovery algorithm [20].

C. Heaviside Model and RKHS Image Smoothing

Heaviside function is also referred to Heaviside step function, which is defined as Eq. 2. There is a singular point at \( x = 0 \) which is not applicable to programming and calculating. The approximated Heaviside function is used for practical, shown in Eq. 3. The Fig. 2 compares the Heaviside function and two approximated Heaviside functions under \( \xi = 0.7 \) and \( \xi = 0.03 \). In Eq. 3, where \( \xi \) controls the smoothness and the smaller \( \xi \) has shaper edge [21].

\[
\phi(x) = \begin{cases} 
0 & x < 0 \\
1 & x \geq 0
\end{cases} 
\]

\[
\psi(x) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x}{\xi}\right)
\]
In [22], it theoretically proves that any function in $L([0,1]^d)$, has a best approximation by linear combinations of $m$ characteristic functions of half-space. For any 2D image signal, the underlying image intensity function (Grayscale image processing) $f$ is defined on $[0,1]^2$ and $f \in L_p([0,1]^2)$. The $f$ function can be approximated by Eq. 4, where $w_j \in R, v_j \in R^2$ and $c_j \in R$. The $v_j$ is discretized to denote different directions. The $c_j$ represents discrete positions. For an image, it is a discretization of intensity function $f$ on $[0,1]^2$. The image function can be rewritten as matrix vector form as $L = \Psi w$. After computing the coefficient $w$, the image super resolution can be reconstructed with upscaling factor by equation $\hat{w}$, where $\Psi$ is $N = s^2 n_1 n_2$ ($n_1 n_2$: the low resolution dimension) [23].

$$f(z) = \sum_{j=1}^{m} w_j \psi(v_j : z + c_j)$$  

(4)

The Heaviside functions mostly model the edge. However, the smooth components are also critical in IR images, especially in thermal IR and long wavelength IR camera. The splines based RKHS model can solve the smoothing problems using optimization problem. For a real valued function $f(t) = f : f \in [0,1], f^{(m)} \in [0,1]$, it can be expanded by Taylor series at $t = 0$ as Eq.5. Let $\phi_i(t) = \frac{t^{v-1}}{(v-1)!}, v = 1, 2, ... m, \mathcal{H}_0 = \text{span}\{\phi_1, \phi_2, ..., \phi_m\}$. For a function $f_0 = \sum_{v=0}^{m-1} \frac{t^{v}}{v!} f^{(v)}(0)$, it can be expressed using the basis of $\mathcal{H}_0$ (RKHS).

$$f(t) = \sum_{v=0}^{m-1} \frac{t^{v}}{v!} f^{(v)}(0) + \int_{0}^{1} (t - u)^{m-1} \frac{f^{(m)}(u)}{(m-1)!} du$$  

(5)

The 2D image can be modeled by 2D thin plate spline based RKHS. Let $f$ be the intensity function of a image defined in a continuous domain $E^2 = [0,1]^2$, and it belongs in a RKHS so that $\hat{f} = (f(t_1), f(t_2), ... , f(t_n))$ be its discretization on grids $i \in [0,1]^2, i = 1, 2, ..., n$. The whole noisy image can be described by Eq.6 with noise $\eta$. In order to get smoothing of the problem, it can be estimated by $f$ using the minimizing model of Eq.7, where $J_m$ is the null space of the penalty function. Since the low resolution $L$ can be formulated by $L = DBH, H$ : the high resolution image. By using smoothing model of $H$ into Eq.8, $c, d$ can be solved. In the this work, the image intensity function is the sum of smooth components and edges. The proposed model is shown in Eq.9.

$$\hat{g} = \hat{f} + \eta$$  

(6)

$$\min \frac{1}{n} \| \hat{g} - \hat{f} \|^2 + \lambda J_m(f)$$  

(7)

$$\min \frac{1}{n} \| L - DB(T^h d + K^h c) \|^2 + \lambda c^T K^l c$$  

(8)

$$\min \frac{1}{n} \| L - DB(T^h d + K^h c + \Phi^h \beta) \|^2 + \lambda c^T K^l c + \alpha \| \beta \|_1$$  

(9)

III. EXPERIMENT SETUP AND RESULTS

A. Algorithm Simulation

Some stationary IR images were captured as algorithm test, shown in Fig. 3. The low resolution image is 320X240 pixels. The algorithm can reconstruct three times resolution higher with smooth background and non-smooth component. Fig. 3 shows the comparison between bicubic single image super resolution and proposed method super resolution. Based on the ground truth value, the root-mean-square error (RMSE) for quantitative comparisons is calculated to characterize image quality. The RMSE of universe image are 4.91 and 4.63 for bicubic and proposed method respectively, while they are 3.02 and 2.79 for medical IR image. The proposed novel method has higher image quality compared to general bicubic method.

B. Experimental Results

In order to test developed single image based super resolution method, a single pixel IR imaging system was used to verify the sampling and recovery. As shown in Fig. 4, known as single pixel camera configuration. The camera will consist of a group of lenses, a Digital Micromirror Device (DMD) and a photodetector, along with some associated processing.
hardware. After the light is captured by the photodetector, the information flows to a readout circuit and analog-to-digital converter (ADC), which will convert analog photocurrent into digital values. Those digital values are then used in image recovery, which is the next step of the process and will be run on a stand-alone computer for recovery.

Fig. 5 shows the experimental results on IR camera system. The sampling recovery is 96X96 pixels (the left), while the high resolution images has noisy removal during calculation, the sharpness goes better than LR image. The middle has factor of 2 for super resolution (192X192) and the right has factor of 3 (288X288). The image details and edge are improved together with single image super resolution in the prototype camera image recovery.

Fig. 5: Single pixel super resolution image recovery on developed IR camera system.

IV. CONCLUSIONS

In this paper, a novel model based single image super resolution was introduced to compressive IR imaging system. The desired IR image was described as smooth components and edge. The smooth components were modeled by 2D thin plate spline based RKHS and the edge was represented by approximated Heaviside functions. By computing the coefficients of the redundant basis of low resolution image, it was applied as same in high resolution image for computing. The iterative scheme was proposed to preserve more image details and image can be divided into patches to reduce computation and storage. This image reconstruct method together with carbon nanotube based IR photodetector will make single pixel super resolution IR camera available.

REFERENCES