



FL-MISR: fast large-scale multi-image super-resolution for computed tomography based on multi-GPU acceleration

Kaicong Sun¹ · Trung-Hieu Tran¹ · Jainabalkya Guhathakurta¹ · Sven Simon¹

Received: 9 August 2021 / Accepted: 7 November 2021 / Published online: 29 November 2021
© The Author(s) 2021

Abstract

Multi-image super-resolution (MISR) usually outperforms single-image super-resolution (SISR) under a proper inter-image alignment by explicitly exploiting the inter-image correlation. However, the large computational demand encumbers the deployment of MISR in practice. In this work, we propose a distributed optimization framework based on data parallelism for fast large-scale MISR using multi-GPU acceleration named FL-MISR. The scaled conjugate gradient (SCG) algorithm is applied to the distributed subfunctions and the local SCG variables are communicated to synchronize the convergence rate over multi-GPU systems towards a consistent convergence. Furthermore, an inner-outer border exchange scheme is performed to obviate the border effect between neighboring GPUs. The proposed FL-MISR is applied to the computed tomography (CT) system by super-resolving the projections acquired by subpixel detector shift. The SR reconstruction is performed on the fly during the CT acquisition such that no additional computation time is introduced. FL-MISR is extensively evaluated from different aspects and experimental results demonstrate that FL-MISR effectively improves the spatial resolution of CT systems in modulation transfer function (MTF) and visual perception. Comparing to a multi-core CPU implementation, FL-MISR achieves a more than 50× speedup on an off-the-shelf 4-GPU system.

Keywords Super resolution · Computed tomography · Distributed optimization · Data parallelism · Multi-GPU · Subpixel detector shift

1 Introduction

Super-resolution (SR) is a fundamental task in image processing and has been an attractive research field for decades [1–3]. SR is an algorithm-based image enhancement technique dedicated to improving the spatial resolution of the imaging systems beyond the hardware limit by exploiting the low-resolution (LR) acquisitions and it is widely applied in many applications such as medical diagnostics, surveillance, and remote sensing.

In recent years, we have witnessed tremendous progress of deep learning in multiple image processing and computer vision tasks such as image denoising [4, 5], super-resolution [6, 7], deformable registration [8, 9], and semantic segmentation [10, 11]. Despite of the great success of deep learning in SR, most of the work focuses on single-image SR (SISR) [6, 7, 12–17]. In fact, SR reconstruction can significantly benefit from the available correlated input images which are captured of the same view. Multi-image SR (MISR) exploits the correspondences entailed in the multiple input images and usually outperforms SISR when the relative movements between the reference image and the other input images are well estimated. However, the learning-based MISR approaches in the literature are mainly dedicated to video applications [18–22]. Besides, the quality of the learning-based methods highly depends on the fidelity of the training datasets. In practice, preparing synthetic datasets which adequately resemble the real-world measurements covering diverse imaging conditions would be challenging. Furthermore, although learning-based approaches are able to describe more sophisticated image priors, “hallucinated”

✉ Kaicong Sun
kaicong.sun@ipvs.uni-stuttgart.de

Trung-Hieu Tran
trung-hieu.tran@ipvs.uni-stuttgart.de

Jainabalkya Guhathakurta
jainabalkya.guhathakurta@ipvs.uni-stuttgart.de

Sven Simon
sven.simon@ipvs.uni-stuttgart.de

¹ Institute of Parallel and Distributed Systems, University of Stuttgart, 70569 Stuttgart, Germany

structures can be unpredictably constructed which may impede the employment of the trained models in applications such as metrology and quality control.

Different from the deep learning-based SR methods, optimization-based MISR algorithms [23–27] reconstruct the latent high-resolution (HR) image explicitly based on the real acquisitions but not the training datasets. Nowadays, due to the technological development of sensor manufacturing, sensors or detectors with large resolutions such as 8, 16 Mpixels or even higher are employed in applications such as medical imaging and industrial inspection. Coping with large-scale multi-image input can be computationally expensive and hardware costly. The optimization-based SR methods usually suffer from the iterative manner which leads to undesirable computation time. In this work, we present a multi-GPU accelerated framework for large-scale MISR reconstruction based on distributed optimization. The proposed framework is applied to the computed tomography (CT) imaging system and achieves a real-time SR reconstruction during the CT acquisition without introducing additional computation time. The contribution of this work can be summarized as following:

- We propose a distributed optimization framework for MISR, named FL-MISR, dealing with large sized images based on multi-GPU acceleration. Each GPU accounts for an allocated partition and the latent SR image is obtained by image fusion.
- To obtain a consistent resolution enhancement among all the GPUs, the update of the partitions is synchronized by unifying the local variables of the scaled conjugate gradient (SCG) method. To avoid border effect between neighboring GPUs, an inner-outer border exchange scheme is performed.
- The proposed FL-MISR is applied to real-time CT imaging by super-resolving the projections acquired via sub-pixel detector shift. Extensive evaluation from different aspects demonstrates that FL-MISR not only achieves a significant resolution enhancement for CT systems but also provides very promising results for natural images. Comparing to a multi-core CPU implementation, FL-MISR achieves a more than 50× speedup on a 4-GPU system.

2 Related work

2.1 Optimization-based iterative methods

In the literature, conventional optimization-based iterative SR methods can be traced back to 1980's and they are mainly grouped into two categories: the frequency domain based and the spatial domain based methods [1, 2]. In Huang et

al. [28], firstly address the MISR problem in the frequency domain. Although the frequency domain-based methods have low computational complexity, they behave extremely sensitive to model errors and have limited ability to integrate a priori knowledge as regularization. The majority of the iterative MISR approaches solve the problem in the spatial domain based on the maximum likelihood (ML), the maximum a posteriori (MAP), and the projection onto convex sets (POCS) [23, 25, 27, 29–33]. Most of the work focuses on the reconstruction accuracy and only few concerns the performance in computation time. Specially, Elad et al. [30] propose a fast MISR algorithm concerning the special case of pure translation and space invariant blur. In [23], Farsiu et al. present a robust MISR method based on MAP using the L1 norm data fidelity term and the bilateral total variation regularization. Jens et al. [32] introduce a GPU-accelerated MISR approach for image-guided surgery which supports a 2× SR reconstruction from 4 LR images of size 200×200 in 60 ms. However, due to the GPU memory limit, their method can not handle large sized images. In [34], the authors propose a fast MISR method which composes of registration, fusion, and sharpening for satellite images using high-order spline interpolation. Nevertheless, purely image fusion is performed on a GPU and the rest two steps are on the CPU which results in a degraded performance in runtime.

2.2 Deep learning-based methods

In the last decade, deep learning has been very successfully adopted in SR and has harvested fruitful results. Dong et al. [6] introduce the convolutional neural network (CNN) into SISR which demonstrates the great potential of CNN for feature extraction. Inspired by the distinguished performance of CNN, a series of work from plain CNN to densely connected GAN, from 2D natural image to 3D medical volume, has been successively proposed [7, 12, 13, 15–17]. Comparing to the traditional iterative methods, CNN-based SR approaches focus on super-resolving single LR image by exploiting the relation learned exclusively from the LR-HR image pairs in the external example database. The learning-based MISR methods are mainly proposed to cope with natural video streams [18–21]. Although some work is intended for real-time applications using GPU or FPGA [21, 22, 35], the video SR (VSR) performance highly relies on the fidelity of the synthesized LR-HR frame pairs and the quality of the training datasets. Furthermore, since the supervised learning scheme requires the ground truth (GT) HR images during the training phase, the performance of the trained model will be limited by the available quality of the GT acquired in practice. It is especially true for CT imaging due to the lack of publicly available high-quality HR datasets like DIV8K [36] for natural images.

To the best of our knowledge, the literature on GPU-accelerated MISR methods for large-scale images is very limited despite of its importance. In this paper, we extend our previous work [37] mainly in two folds. First, the locally applied scaled conjugate gradient (SCG) algorithm is adapted to achieve a synchronized convergence rate over multi-GPU systems. Second, instead of performing region averaging, we employ the so-called inner-outer border exchange scheme to preserve the sharpness of the overlapped regions. Particularly, in [37] we introduce a multi-GPU implementation of a MISR method based on data parallelism where each GPU deals with an allocated partition of the latent SR image. Although overlapped regions between neighboring GPUs are exchanged and averaged, we found that the resolved SR image in [37] is not globally optimized and the fused SR image suffers from inhomogeneous resolution enhancement due to the inconsistent convergence rate of the local SCG and region averaging. We address these issues in this work and propose a generalized framework for multi-GPU supported MISR. We conducted extensive experiments to validate the proposed FL-MISR. Experimental results show that the exchange of local SCG variables and overlapped regions among GPUs has limited impact on the overall performance of runtime and leads to a consensus convergence over multi-GPUs without causing border effects. Besides, it is shown that super-resolving four input images of size 4096×4096 by an upscaling of $2\times$ can be achieved within $2.4s$ on a 4-GPU system.

3 Methods

3.1 Distributed optimization for MISR

The common formulation of SR model in the pixel domain is presented as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon}(\mathbf{x}) \tag{1}$$

with $\mathbf{x} \in \mathbb{R}^{n \times 1}$, $\mathbf{y} \in \mathbb{R}^{m \times 1}$ being, respectively, the latent and captured image rearranged in lexicographic order. The system matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is usually expressed as $\mathbf{A} = \mathbf{D}\mathbf{B}\mathbf{M}$ with $\mathbf{D} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times n}$, and $\mathbf{M} \in \mathbb{R}^{n \times n}$ describing the decimation, blurring, and motion effects, respectively. The vector $\boldsymbol{\varepsilon}(\mathbf{x}) \in \mathbb{R}^{m \times 1}$ denotes the additive noise existing in the imaging systems. More detailed description of the system model can be found in [27]. To simplify the calculation, in this work we assume $\boldsymbol{\varepsilon}(\mathbf{x})$ is an intensity-independent additive noise and the system matrix \mathbf{A} is known.

Since SR is an ill-posed problem, involving a well-defined image prior can effectively constrain the solution domain. Therefore, MAP estimator is preferably adopted

for SR reconstruction. The posterior probability $P(\mathbf{x}|\mathbf{y})$ of the SR image \mathbf{x} is formulated based on the Bayes' theorem:

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{x})P(\mathbf{x})}{P(\mathbf{y})}. \tag{2}$$

Assuming the noise $\varepsilon_i \in \varepsilon$ in each pixel i is white Gaussian and i.i.d with $\varepsilon_i \sim N(0, \sigma^2)$ and $P(\varepsilon_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\varepsilon_i^2}{2\sigma^2}}$, we yield the likelihood function as

$$P(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^m P(y_i|\mathbf{x}) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^m e^{-\frac{\|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2}{2\sigma^2}} \tag{3}$$

Taking the natural logarithm, the associated negative log-likelihood can be formulated as

$$-\log(P(\mathbf{y}|\mathbf{x})) = \frac{1}{2\sigma^2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + c \tag{4}$$

where c is a constant. For brevity, we will omit the weight $\frac{1}{2\sigma^2}$ and the constant c in the latter formulation.

For MISR with k independent LR images $\mathbf{y}_i, i \in [1 \dots k]$, the posterior probability can be extended as

$$P(\mathbf{x}|\mathbf{y}_1 \dots \mathbf{y}_k) = \frac{P(\mathbf{y}_1 \dots \mathbf{y}_k|\mathbf{x})P(\mathbf{x})}{\prod_{i=1}^k P(\mathbf{y}_i)} = \frac{\prod_{i=1}^k P(\mathbf{y}_i|\mathbf{x})P(\mathbf{x})}{\prod_{i=1}^k P(\mathbf{y}_i)} \tag{5}$$

and the data fidelity term is hence formulated by

$$-\log\left(\prod_{i=1}^k P(\mathbf{y}_i|\mathbf{x})\right) = \sum_{i=1}^k \|\mathbf{A}_i\mathbf{x} - \mathbf{y}_i\|_2^2. \tag{6}$$

It should be noted that in case of additive white Laplacian noise which models the impulse noise (Salt & Pepper noise), we have the L1 norm data fidelity term [38]. Usually, L1 norm data term has better robustness against pixel outliers [23]. Without loss of generality, the data fidelity term can be formulated as

$$-\log\left(\prod_{i=1}^k P(\mathbf{y}_i|\mathbf{x})\right) = \sum_{i=1}^k \|\mathbf{A}_i\mathbf{x} - \mathbf{y}_i\|_p^p \tag{7}$$

with the L_p norm $p \in \{1, 2\}$.

In the literature, there are several well-known handcrafted image priors $P(\mathbf{x})$ including the total variation (TV) [39], Huber-Markov prior [40], bilateral total variation (BTV) [23], nonlocal total variation (NLTV) [41], and the more recent bilateral spectrum weighted total variation (BSWTV) [27]. In this paper, aiming for reducing the computational complexity, we leverage the BTV as the image prior and the regularization term is therefore expressed as

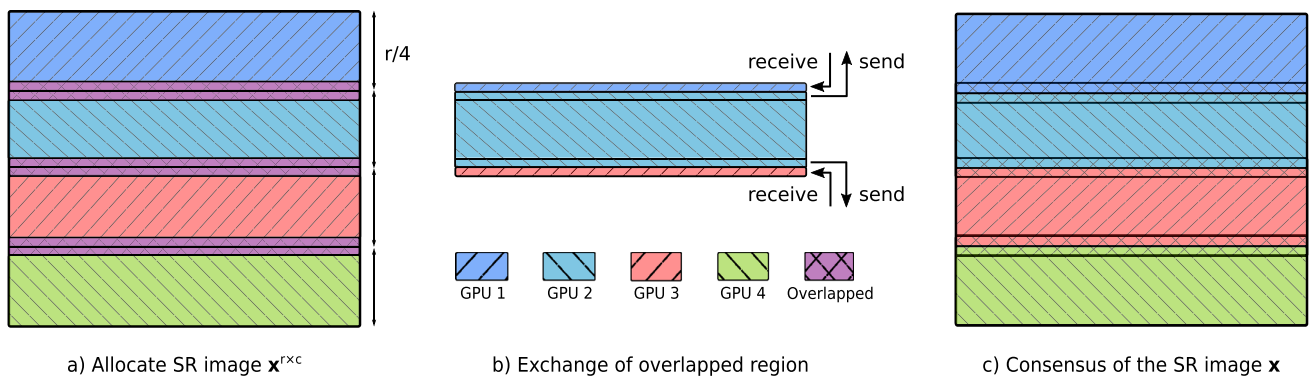


Fig. 1 Demonstration of the inner-outer border exchange scheme of the overlapped regions for 4 GPUs

$$-\log(P(\mathbf{x})) = \sum_{\mathbf{d}} \gamma(\mathbf{d}) \|\mathbf{x} - S_{\mathbf{d}}\mathbf{x}\|_1, \mathbf{d} = (d_x, d_y) \tag{8}$$

where $\mathbf{d} \in \mathbb{N}^2$ with $d_x, d_y \in [0, w - 1]$ and w denotes the window size accounting for the neighbors in the x, y directions. $S_{\mathbf{d}}$ represents the shifting operator along x and y axis by d_x and d_y pixels. $\gamma(\mathbf{d}) := \alpha^{d_x+d_y}$ embodies the spatial decaying effect with constant $\alpha < 1$.

As the denominator of Eq. (5) is independent from the image \mathbf{x} , maximizing the posterior probability $P(\mathbf{x}|y_1 \dots y_k)$ is equivalent to minimizing the negative logarithm of the numerator which is formulated, respectively, in Eqs. (7) and (8). Hence, we yield the overall objective function based on the MAP framework as following:

$$J(\mathbf{x}) = \sum_{i=1}^k \|\mathbf{A}_i\mathbf{x} - \mathbf{y}_i\|_p^p + \lambda \sum_{\mathbf{d}} \gamma(\mathbf{d}) \|\mathbf{x} - S_{\mathbf{d}}\mathbf{x}\|_1 \tag{9}$$

where the scaling factor of the fidelity term $\frac{1}{2\sigma^2}$ in Eq. (4) is actually absorbed into the weighting parameter λ . In the experiments, we have used the L1 norm data term for a better robustness.

To accelerate the computation and alleviate the GPU memory load especially when coping with a sequence of large input images, we distribute the computational demand over multi-GPUs by data parallelism and follow a consensus-based convergence manner to guarantee a centralized solution. The latent SR image \mathbf{x} is finally obtained by data fusion. In particular, Eq. (9) can be rewritten as

$$J(\mathbf{x}) = \sum_{i=1}^k D_i(\mathbf{x}) + \lambda R(\mathbf{x}) \tag{10}$$

with D_i representing the corresponding data term and R being the regularization term. In this regard, the subfunction associated with the h th GPU is expressed as

$$J_h(\mathbf{x}_h) = \sum_{i=1}^k D_i(\mathbf{x}_h) + \lambda R(\mathbf{x}_h), \text{ s.t. } \bigcup_{h=1}^g \mathbf{x}_h = \mathbf{x} \tag{11}$$

where \mathbf{x}_h is a fraction of the latent image \mathbf{x} assigned to the h th GPU and g denotes the number of employed GPUs. To enforce the distributed optimization towards a centralized solution, we allow communication between the local GPU node and the host CPU for a consensus update decision. Specially, we utilize the SCG algorithm [42] to iteratively solve the subproblem described in Eq. (11) in each GPU. Instead of using the handcrafted step size or performing line search, SCG employs a step size scaling mechanism based on an adaptive scalar which achieves a faster and more robust convergence than the widely used approaches such as conjugate gradient with line search (CGL) and Broyden–Fletcher–Goldfarb–Shanno (BFGS).

Aiming for synchronizing the update of the individual \mathbf{x}_h towards a centralized solution, we unify the local SCG scalar variables $\sigma, \lambda, \delta, \mu, \alpha$ by data communication. As these variables are calculated based on the inner product of vectors, we can obtain the consensus variables by the aggregate of the broadcast local ones. By means of consensus variables, the subfunctions can converge synchronically and a homogeneous resolution among multi-GPUs is guaranteed. In Table 1, we list the unified scalar variables and vectors (in bold) of SCG.

Table 1 List of the consensus variables in SCG algorithm

Param.	Description
f_c, f_{c_new}	Consensus of the objective function
\mathbf{p}_c	Consensus of the conjugate weight vector
$\mathbf{r}_c, \mathbf{r}_{c_new}$	Consensus of the steepest descent direction
σ_c, λ_c	Consensus of the scalars
δ_c, μ_c	Consensus of the variables in step size
α_c	Consensus of the update step size

In addition, to avoid border discontinuity of neighboring partitions, region overlapping between neighboring GPUs is required. Instead of the naive averaging of the overlapped regions which sacrifices the sharpness and visual quality, we perform an inner-outer border exchange in each SCG iteration as shown in Fig. 1. A 4-GPU system is demonstrated and each GPU deals with the allocated image partition \mathbf{x}_h . The overlapped regions marked in violet are exchanged between neighboring GPUs. Particularly, since the inner borders can be correctly calculated only in case that the outer borders are consistent with the neighboring GPUs, the outer borders are replaced by the received ones and the inner borders are broadcast to the neighbors as exhibited in Fig. 1b). Consequently, an agreement in the overlapped regions is achieved as shown in Fig. 1c) without compromising the image sharpness. Without loss of generality, assuming g GPU nodes are employed, the architecture of the proposed multi-GPU framework for SR is illustrated in Fig. 2. The local variables and overlapped regions are interchanged in each SCG iteration over the host CPU and updated in a consensus scheme.

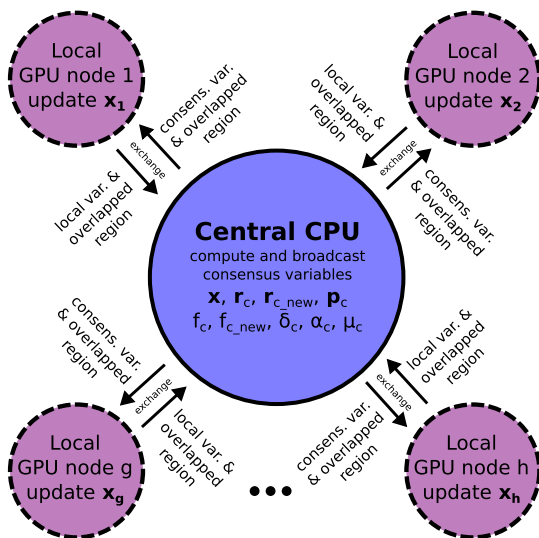


Fig. 2 Architecture of the proposed multi-GPU framework for MISR where g GPU nodes are employed

Algorithm 1 Distributed SR Reconstruction

```

1: Partition the LR images  $\mathbf{y}_i, i \in [1 \dots k]$  for each GPU
   node  $h \in [1 \dots g]$ .
2: Initialize each GPU node with  $\gamma(\mathbf{d}), \lambda, f_h, f_c, \mathbf{p}_h, \mathbf{r}_h, \delta_c,$ 
    $\mu_c, \alpha_c, \sigma, n_{iter}$ .
3: Calculate matrices  $\mathbf{A}_i, \mathbf{A}_i^T, i \in [1 \dots k]$  in each GPU.
4: procedure ESTIMATE LATENT IMAGE  $\mathbf{x}$  ACCORDING TO
   EQS. (9) AND (11) USING SCG [42]
5:   while  $i_{iter} < n_{iter}$  do
6:     Local : Calculate  $\|\mathbf{p}_h\|_2^2, h \in [1 \dots g]$ .
7:     Central: Update  $\sigma_c = \sigma / \|\mathbf{p}_c\|, \|\mathbf{p}_c\|_2^2 = \sum_h^g \|\mathbf{p}_h\|_2^2$ .
8:     Local : Calculate  $\mathbf{x}_{h\_tmp} = \mathbf{x}_h + \sigma_c \mathbf{p}_h$ .
9:     Central: Exchange overlapped regions of  $\mathbf{x}_{h\_tmp}$ 
       with neighboring GPUs.
10:    Local : Calculate  $\delta_h$  according to SCG.
11:    Central: Update  $\delta_c = \sum_h^g \delta_h$ .
12:    Local : Calculate  $\mu_h, \alpha_h$  according to SCG.
13:    Central: Update  $\mu_c = \sum_h^g \mu_h, \alpha_c = \sum_h^g \alpha_h$ .
14:    Local : Calculate  $\mathbf{x}_{h\_new} = \mathbf{x}_h + \alpha_c \mathbf{p}_h$ .
15:    Central: Exchange overlapped regions of  $\mathbf{x}_{h\_new}$ 
       with neighboring GPUs.
16:    Local : Calculate  $f_{h\_new}$  according to Eq. (11).
17:    Central: Update  $f_{c\_new} = \sum_h^g f_{h\_new}$ .
18:    Local : Calculate  $\|\mathbf{r}_{h\_new}\|_2^2$ , inner product
        $\langle \mathbf{r}_h, \mathbf{r}_{h\_new} \rangle$ .
19:    Central: Update  $\|\mathbf{r}_{c\_new}\|_2^2 = \sum_h^g \|\mathbf{r}_{h\_new}\|_2^2,$ 
        $\langle \mathbf{r}_c, \mathbf{r}_{c\_new} \rangle = \sum_h^g \langle \mathbf{r}_h, \mathbf{r}_{h\_new} \rangle$ .
20:    Local : Update  $\mathbf{p}_h$ .
21:    Central:  $i_{iter} = i_{iter} + 1$ .
22:   end while
23:   Central: Fuse  $\mathbf{x}_h, h \in [1 \dots g]$  to reconstruct  $\mathbf{x}$ .
24:   return reconstructed image  $\mathbf{x}$ .

```

In Algorithm 1, we present a detailed description of the proposed distributed optimization framework for MISR based on the SCG approach. The local GPU computation is marked by red and the centralized computation in the host CPU is denoted in blue. The local variables, overlapped regions, and consensus variables are respectively exchanged after the local and central update. The algorithm variables are initialized based on SCG [42] and the calculation of the system matrix \mathbf{A}_i is explained in Sect. 4. The SR image \mathbf{x} is fused when the SCG iterations are complete.

In the implementation, we have used the OpenCL framework. In order to optimize the data deployment on GPU memory, we exploited the local memory in the kernel functions to the most extent. Sparse matrix was employed to calculate the system matrix $\mathbf{A}_i = \mathbf{D}_i \mathbf{B}_i \mathbf{M}_i$, and the transpose \mathbf{A}_i^T due to the sparseness of the downsampling, blurring, and motion matrices. Although memory transfer of local

variables and overlapped regions between the GPU and host CPU is intended to hold the consensus convergence, transfer of large amounts of data is obviated during the SR reconstruction. It is worthy noting that the proposed distributed optimization framework is based on data parallelism and consensus SCG. It can be easily applied to other applications such as SISR and image denoising by replacing the objective function in Eq. 9.

3.2 Real-time MISR for CT

SR is always preferable in CT imaging where spatial resolution plays a determinant role in image quality assessment. We have applied the proposed FL-MISR on the industrial CT scanner as shown in Fig. 3. During the CT acquisition, the object is rotated by 360° and at each rotation angle, four LR projections (X-ray images) are captured via detector shift rightwards, downwards, leftwards, and upwards by half a pixel as illustrated in Fig. 4. As long as all the four LR projections of the same view are collected, SR reconstruction is launched as denoted in green along the scan time axis. The capture-reconstruct fashion repeats until the whole CT acquisition is accomplished. Due to the fact that SR reconstruction usually takes less time than the accumulated time of projection acquisition (in red) and object rotation (in gray), SR can be performed in real-time during

the CT scan without introducing extra runtime. The super-resolved projections are utilized for CT reconstruction and hence, an improved spatial resolution in CT is achieved by the increased detector sampling rate. We demonstrate the experimental results in Sect. 4. It is necessary to note that since the same detector movement pattern is repeated for all the rotation angles during CT scan, the system matrices A_i with $i \in [1, 4]$ are calculated once at the beginning of the CT acquisition and shared by all the rotation angles.

4 Experiments and results

In this section, we conduct extensive experiments to evaluate the performance of the proposed FL-MISR from different aspects, mainly on resolution enhancement and computation acceleration. Specially, FL-MISR is evaluated for real-time CT imaging based on the synthetic and real-world CT measurements. Besides, the application of FL-MISR on natural images is evaluated using the public dataset DIV8K [36].

The CT measurements were carried out on the Nikon HMX ST 225 CT scanner as shown in Fig. 3 which is equipped with a flat panel Varian PaxScan 4030E detector of pixel size $127 \times 127 \mu\text{m}$. The detector is mounted on the controllable linear stages for x - and y -positioning which supports detector displacement with a movement accuracy

Fig. 3 CT scanner equipped with mounted linear stages. **a** Side view; **b** X-ray tube and rotatable object (aluminium cylindrical phantom); **c** X-ray detector mounted on the controllable linear stages

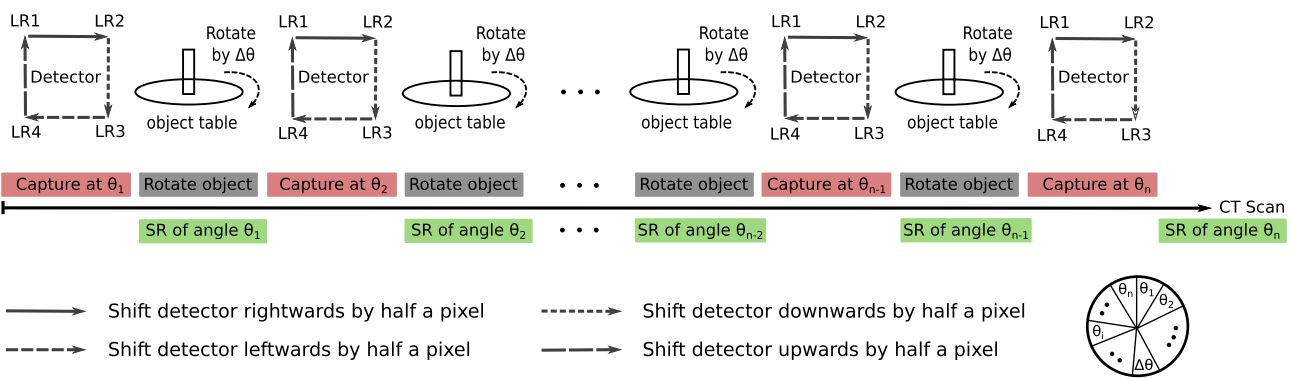
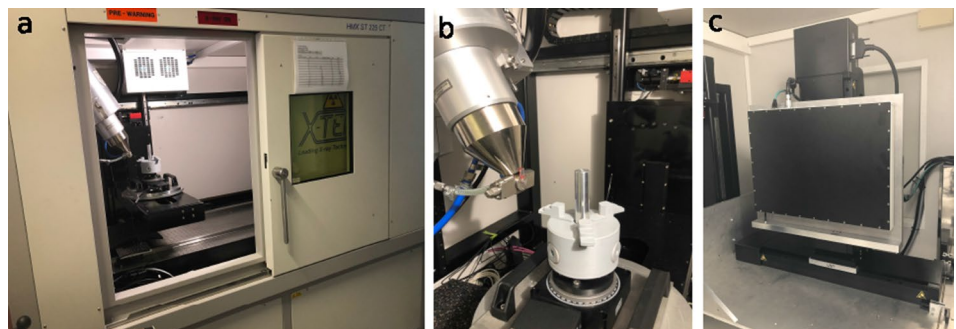


Fig. 4 Schematic illustration of the application of FL-MISR in CT imaging based on the controlled subpixel detector shift

up to $1\ \mu\text{m}$. The focal spot size of the tungsten X-ray tube is power dependent and for the power under 7 W, which was utilized in our experiments, the effective focal spot size is about $6\ \mu\text{m}$ measured by the JIMA RT RC-04 micro chart.

The calculation of the system matrix A_i is thoroughly described in our previous work [26]. For an upscaling of $2\times$ with half pixel detector shift and a 3×3 Gaussian blur for B_i , a 12-row block area in the HR grid is required as the overlapped region between neighboring GPUs. The weighting parameters λ and α were, respectively, set as 0.05 and 0.4. The SCG iteration was limited to 20. In practice, larger λ should be opted in case of strong noise and fewer SCG iterations should be used for fast CT acquisitions. To quantify the resolution enhancement by FL-MISR on CT systems, we adopted the modulation transfer function (MTF) which was measured according to the standard ASTM-E 1695.

4.1 Evaluation of FL-MISR on spatial resolution enhancement

Before we evaluate FL-MISR on CT imaging, we briefly introduce the CT system and the assessment metric. CT scanner mainly consists of two components: the X-ray tube and the X-ray sensitive detector. The spatial resolution of the CT system is hence primarily limited by the focal spot size of the X-ray tube and the detector pixel size. Usually, spatial resolution of imaging systems is assessed by the MTF which is calculated as the normalized magnitude of the Fourier Transform of the point spread function (PSF). The MTF of the CT system is formulated by $MTF_{\text{sys}} = MTF_{\text{fs}} \cdot MTF_{\text{det}} \cdot MTF_{\text{others}}$, where MTF_{fs} and MTF_{det} respectively denote the MTF of the X-ray focal spot and the detector. Other components such as the reconstruction algorithm, X-ray beam hardening, and display monitor are usually of less influence on the overall MTF_{sys} . In this work, we perform subpixel detector shift to achieve a higher detector sampling rate which will lead to an effective improvement of MTF_{sys} when MTF_{det} dominates MTF_{fs} , which is usually the case in many CT applications.

4.1.1 Evaluation on synthetic CT images

To analyze the effectiveness of subpixel detector shift on the spatial resolution enhancement in CT, we first demonstrate the impact of MTF_{det} on the MTF_{sys} . To simplify the system model, we consider only the primary components and therefore, we yield $MTF_{\text{sys}} := MTF_{\text{fs}} \cdot MTF_{\text{det}}$. The MTF_{fs} is modeled by a Gaussian function and the MTF_{det} is represented by a *sinc* function due to the assumed rectangular shape of each pixel. As shown in Fig. 5, the left plot indicates the case where MTF_{fs} dominates MTF_{det} , for instance when the object is extremely close to the X-ray source and the right one depicts the situation where MTF_{det} dominates. The MTF of the detector with full pixel size and with half pixel size is respectively denoted as $Detector_{\text{LR}}$ and $Detector_{\text{HR}}$. The MTF at 10% is usually considered as the visible limit in practice and is marked by the gray dotted line. It is shown that halving the detector pixel size doubles the MTF_{det} and improves the overall MTF_{sys} effectively when MTF_{det} dominates, while for the case MTF_{fs} dominates, MTF_{sys} has a negligible improvement.

Based on the analysis above, we evaluate FL-MISR on the CT images quantitatively and qualitatively. Specially, we conducted CT scans of an aluminium cylindrical phantom with a diameter of 20 mm as shown in Fig. 3b) which was fixed perpendicular to the rotation table and a QRM bar pattern resolution phantom at the magnification of 20. Considering them as the ground truth (GT), we simulated four sets of $0.5\times$ LR projections by shifting the GT projections rightwards, downwards, leftwards, and upwards by one pixel followed by a 2×2 binning. The downsampled LR projections were fused by interpolation and by FL-MISR. As the inter-image offset is assumed to be one pixel and accurate, for interpolation-based fusion we inserted the pixel values of the LR images into the corresponding integer location in the HR grid. The super-resolved projections were then used for CT reconstruction by filter backprojection (FBP). The CT cross sections of the aluminium cylindrical phantom and the associated MTF are demonstrated in Fig. 6. The LR CT was reconstructed by the reference (upper left) set of the downsampled projections.

Fig. 5 Influence of improved detector MTF on the system MTF based on one-dimensional analysis. **a** When MTF_{fs} dominates, MTF_{sys} rarely improves; **b** in case of MTF_{det} dominating, MTF_{sys} improves significantly

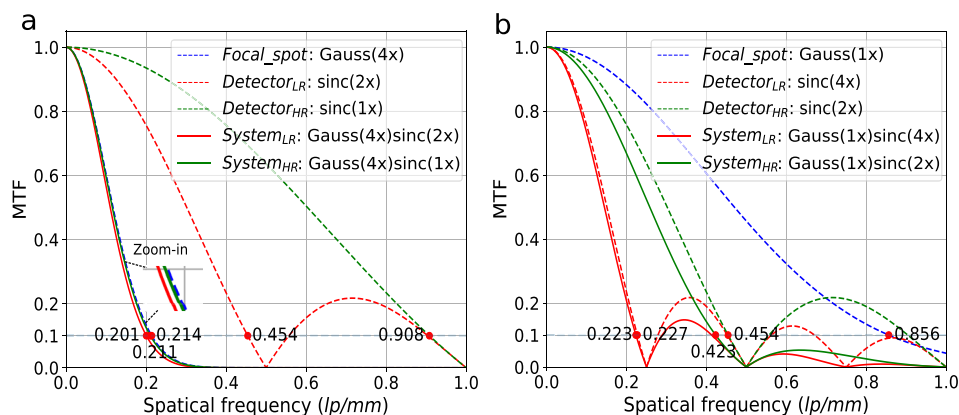


Fig. 6 Evaluation of MTF on the CT cross section of an aluminium cylindrical phantom. Left: **a** LR, **b** multi-image interpolation, **c** FL-MISR, **d** GT; Right: MTF

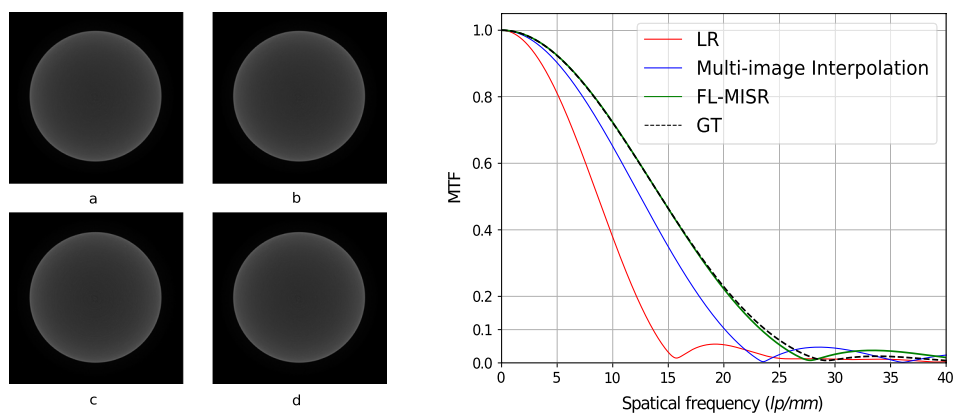
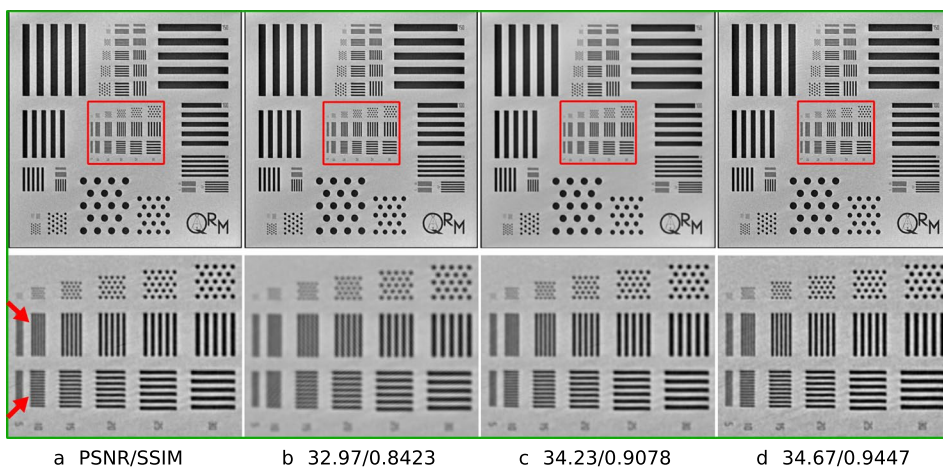


Fig. 7 CT images of the QRM bar pattern phantom. The ROIs are marked by red rectangle and zoomed in. **a**) GT; **b**) Bilinear interpolation; **c**) multi-image interpolation; **d**) FL-MISR



As we can clearly see that FL-MISR resembles the MTF of the GT extremely well and almost doubles the MTF of the LR image. To illustrate the performance of FL-MISR visually, we present the CT images of the QRM bar pattern target in Fig. 7. It is shown that FL-MISR provides a more pleasant result with sharper structures and better visual quality.

4.1.2 Evaluation on real-world CT images

As the spatial resolution of CT systems depends on the magnification, we evaluate FL-MISR on the real-world CT scans at different magnifications. Particularly, we conducted CT measurements of aluminium cylindrical phantoms with diameters of 10 mm and 20 mm, QRM bar pattern phantom with spatial resolution ranging from 3.3 to 100 lp/mm, QRM bar pattern nano phantom which covers resolution from 50 to 500 lp/mm, and a cylindrical dry concrete joint with a diameter of 50 mm. The aluminium cylindrical phantoms and the QRM resolution targets were both scanned at magnifications of 5 (voxel size of $25.4 \mu\text{m}$), 10 (voxel size of $12.7 \mu\text{m}$), and 25 (voxel size of $5.08 \mu\text{m}$) and the concrete joint was acquired at magnifications of 3 (voxel size of $42.3 \mu\text{m}$) and 5. The detailed measurement setup is summarized in Table 2. As

illustrated in Fig. 4, the X-ray detector was repeatedly displaced clockwise by half a pixel in a precisely controlled way. The projection at each detector position took 3 s, namely at each rotation angle 4×3 s was required for the acquisition. The object table rotated over 360° with 0.1 degree resolution following a stop-move manner and hence in total 4×3600 projections were taken. Aluminium filters were utilized to absorb the soft X-ray beam and suppress the beam hardening artifact. We compare FL-MISR with multi-image interpolation and the standard CT without detector shift where the exposure time was set as 12 s, the same as FL-MISR.

In Fig. 8, we demonstrate the MTF measured by the aluminium cylindrical phantoms at different magnifications according to the standard ASTM-E 1695. It is shown that FL-MISR performs significantly better than the standard CT at all the investigated magnifications covering voxel size up to $5.08 \mu\text{m}$. The multi-image interpolation behaves worse than FL-MISR as expected due to the naive manner of fusion.

The CT images of the QRM bar pattern phantom and QRM bar pattern nano phantom are illustrated in Fig. 9 with the corresponding closeup views. Comparing to the standard CT images, we can observe that FL-MISR and multi-image interpolation both improve the spatial

Table 2 Parameter setup for CT measurements

Test phantoms	Voltage (kV)	Current (μA)	# of Angles	Exposure (s)	Subpixel shift	Magnification	Filter (mm)
Aluminium cylinder	200	34	3600	3	0.5	5, 10, 25	Al 2.5
QRM bar pattern	80	86	3600	3	0.5	5, 10, 25	None
Dry concrete joint	180	110	3600	3	0.5	3, 5	Al 0.1

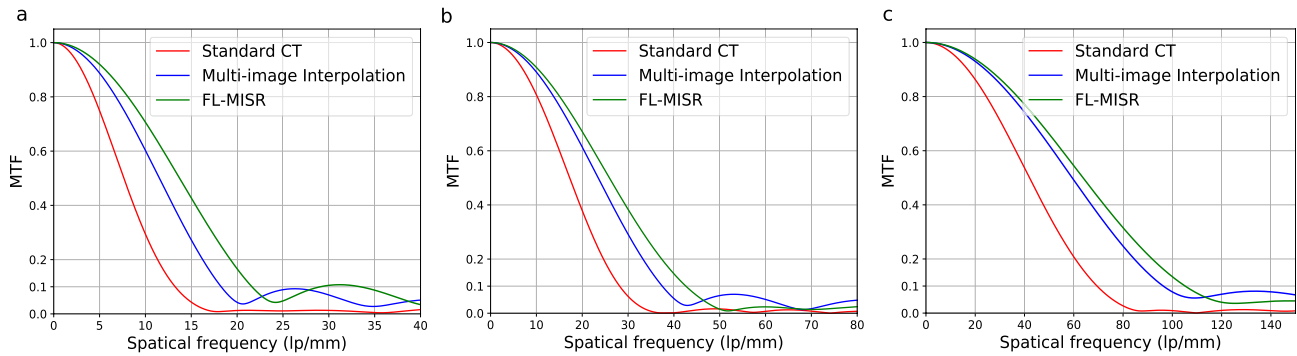


Fig. 8 Evaluation of MTF at different magnifications. **a** Magnification of 5; **b** magnification of 10; **c** magnification of 25

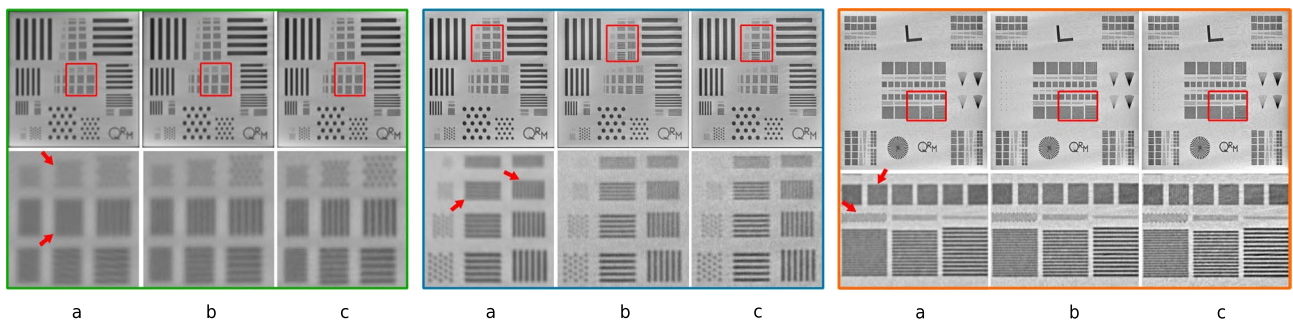


Fig. 9 CT images of QRM bar pattern phantom. Left (marked in green): magnification of 5; Middle (marked in blue): magnification of 10; Right (marked in orange): magnification of 25. **a** Standard CT without detector shift; **b** multi-image interpolation; **c** FL-MISR

resolution by exploiting the additional information captured via subpixel detector shift. However, multi-image interpolation is less robust than the optimization-based FL-MISR. FL-MISR generates sharper edges and provides more pleasant results in visual perception. In fact, the spatial resolution estimated by the visibility of the QRM bar patterns coincides with the MTF measured by the cylindrical phantoms.

In Fig. 10, we illustrate the CT images of a dry concrete joint with the zoomed-in region of interest (ROI). Fig. 10a and b represent respectively the results of the standard CT without detector shift and FL-MISR at the magnification of 3. Fig 10c exhibits the results of standard CT at the magnification of 5 which is considered as the reference image. It is shown that comparing to the standard CT with a voxel size of $42.3 \mu\text{m}$ at the magnification of 3, FL-MISR generates sharper contours with more detailed structures which resembles the CT measurement at the magnification of 5 better.

4.1.3 Evaluation on border effect and consensus convergence

As explained in Fig. 1, we exchange the overlapped regions between neighboring GPUs to avoid border discontinuity. In Fig. 11, we demonstrate the super-resolved projections and the associated CT images of the synthetic (top row) and the real-world measurements (bottom row). For the synthetic image, we employed four GPUs and for the real-world one, two GPUs were in use. The individual \mathbf{x}_h of each GPU is partitioned by the red dotted line. As we can observe that the overlapped regions, a 12-row block surrounding the borders (the red dotted lines), are of inherent sharpness without intensity discontinuity and the border effect is fundamentally obviated. Besides, to avoid inhomogeneous resolution in different partitions, we synchronize the update of the partitioned \mathbf{x}_h among all the GPUs by exchanging the local variables of SCG. In Fig. 12, we illustrate the convergence

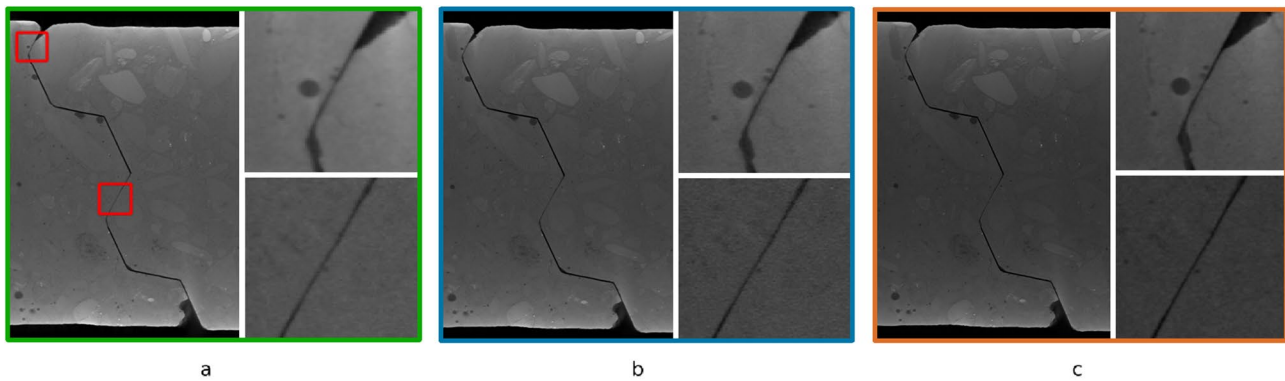


Fig. 10 CT images of a dry concrete joint with the ROI in the closeup views. **a** Standard CT without detector shift at the magnification of 3; **b** FL-MISR with an upscaling of 2× at the magnification of 3; **c** standard CT without detector shift at the magnification of 5

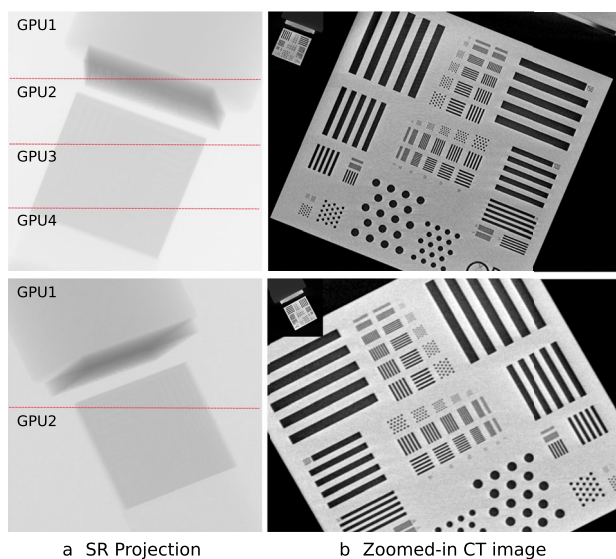


Fig. 11 Evaluation on the border effect. First row: on the synthetic volume as utilized in Fig. 7; second row: on the real-world volume as used in the middle graph of Fig. 9. Red dotted line marks out the border of the partitions allocated to the GPUs

curve of the centralized objective of Eq. 9 running on a single GPU and the distributed objective of Eq. 11 running on four GPUs. The consensus convergence is reflected in two aspects. First, the four GPUs have exactly the same convergence trend, where they are almost overlaid, due to the share of the SCG variables. Second, the distributed objective follows the same convergence trend as the centralized one and moreover, the sum of the four distributed objectives equals the centralized one by resorting to the scheme we adopt for the calculation of the consensus variables of SCG as described in Sect. 3.1. In addition, we can observe that the objective function is almost converged after 5 SCG iterations.

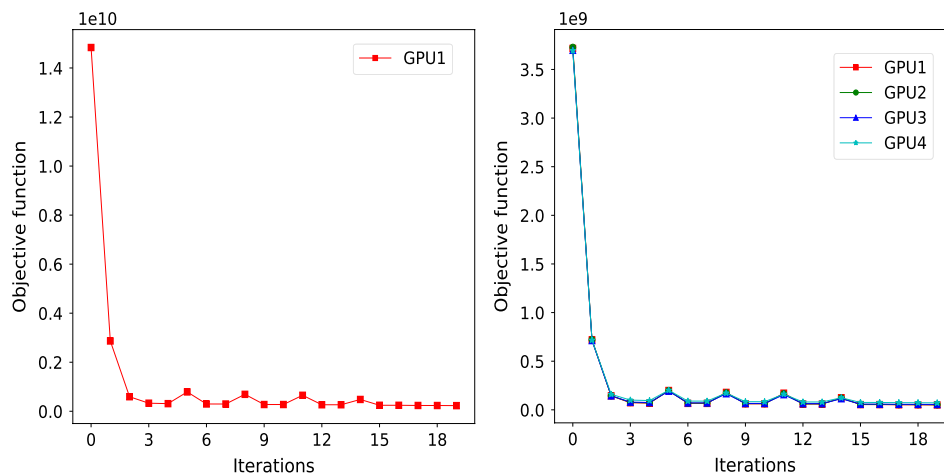
4.1.4 Evaluation on natural images

Since the distributed optimization of FL-MISR is based on data parallelism, FL-MISR is not limited to a certain application. We evaluate the proposed FL-MISR on natural images using the public dataset DIV8K [36]. Particularly, we randomly selected 7 natural images with the vertical or horizontal resolution ranging from 1920 to 5760 pixels as the GT. For each GT image, 4 and 9 LR images were respectively generated for upscaling factors of 2× and 3× according to Eq. (1) with $\varepsilon \sim N(0, 1)$ and translational movement of $\frac{1}{2}$ and $\frac{1}{3}$ pixel. We performed SR reconstruction only for the luminance channel on 4 GPUs and set the SCG iterations as 10. The SR performance is assessed by PSNR, SSIM, and runtime. Quantitative evaluation is summarized in Table 3. As we can see, the proposed FL-MISR outperforms the multi-image interpolation by a large margin in PSNR and SSIM. Although the iterative FL-MISR requires 2~5 × runtime as the naive interpolation one, it supports an SR output of 5760 × 5760 resolution within 1.3 s for the upscaling of 2× and 1.93 s for the upscaling of 3×. It is interesting to find that the runtime of SCG depends not only on the image size but also on the count of iterations with successful reduction in the objective function as expressed in [42]. In Fig. 13, we illustrate the closeup views of images 0002, 0027, and 0055 of DIV8K. The top two rows demonstrate the results for an upscaling of 2× and the bottom row is for the upscaling of 3×. We can observe that FL-MISR provides pleasant results with significantly better visual quality than the multi-image interpolation.

4.2 Evaluation of FL-MISR on acceleration

To demonstrate the performance of FL-MISR in acceleration, we conducted SR reconstruction of different sized inputs ranging from 512 × 512 to 4096 × 4096 for an

Fig. 12 Evaluation on consensus convergence based on the objective function. Left, convergence curve obtained using single GPU; right, convergence curves obtained using four GPUs



upscaling factor of 2× on a multi-core CPU, single GPU, and multi-GPU systems. In particular, the CPU experiments were performed on the Intel Xeon Gold 5120 CPU with 755GB memory which contains two nodes and each is equipped with 28 cores. The GPU experiments were carried out on the Nvidia GeForce GTX 1080 GPUs with 11GB memory. Since FL-MISR is based on the iterative SCG algorithm, we evaluated the runtime also with regard to the number of SCG iterations. Besides, we also demonstrate the runtime of the multi-image interpolation as the baseline. The performance of different configurations was calculated based on an average of 100 runs and is summarized in Table 4 where N/A denotes not applicable due to the large GPU memory footprint. As we can see, comparing to the 56-core CPU variant, the single GPU implementation accelerates the computation by more than 25× for LR images of size 2048 × 2048 and the multi-GPU implementation which uses 4 GPUs achieves a speedup up to 50×. For large-scale images of size 2300 × 3200 and 4096 × 4096, FL-MISR running on 4 GPUs obtains a more than 55× speedup than the CPU implementation, while single GPU can not fulfill the memory requirement. For small sized inputs like 512 × 512 and 1024 × 1024, single GPU implementation has similar performance as

multi-GPU and achieves a 20× speedup comparing to the multi-core CPU. Although the iterative FL-MISR requires more runtime than the naive interpolation one, FL-MISR has much better SR performance and the runtime difference becomes less as the image dimension increases.

In addition, we analyzed the runtime distribution for the local and central computation on a 4-GPU system where the data communication time is aggregated into the central computation. We exhibit the average runtime distribution over 100 runs for input images of different sizes in Fig. 14. It is shown that the consumed time for consensus computing is almost negligible comparing to the local computation, while it is fundamentally necessary to avoid border effects between neighboring GPUs and guarantee a consensus convergence over multi-GPU systems.

5 Conclusion

In this paper, we propose a multi-GPU accelerated large-scale multi-image super-resolution (MISR) framework based on data parallelism. Specially, each GPU node accounts for a designated region of the latent high-resolution (HR) image by applying an adapted scaled conjugate gradient (SCG)

Table 3 Evaluation of FL-MISR on 8-bit natural images in DIV8K dataset

Image Index		#0001	#0002	#0007	#0027	#0055	#0066	#0084
Resolution of GT		5376×5760	5568×5760	1920×2880	2112×2880	5760×5760	1920×2880	5760×3840
Upscaling 2×								
MI Interp.	PSNR/SSIM	30.49/0.9215	28.44/0.8677	33.68/0.8810	28.37/0.8988	33.80/0.9018	35.21/0.9296	29.11/0.8277
	Runtime (s)	0.51	0.52	0.11	0.20	0.53	0.11	0.36
FL-MISR	PSNR/SSIM	37.11/0.9620	32.99/0.9360	35.09/0.9111	33.21/0.9417	38.03/0.9564	37.12/0.9452	34.13/0.9410
	Runtime (s)	1.50	1.29	0.69	0.71	1.3	0.66	1.21
Upscaling 3×								
MI Interp.	PSNR/SSIM	26.74/0.8460	25.65/0.7749	32.03/0.8395	25.15/0.8212	30.79/0.8153	32.65/0.8968	26.19/0.6883
	Runtime (s)	1.00	0.99	0.11	0.13	0.55	0.11	0.38
FL-MISR	PSNR/SSIM	33.24/0.9446	29.43/0.8941	33.99/0.8941	30.17/0.9139	35.90/0.9379	36.06/0.9398	30.54/0.8764
	Runtime (s)	1.78	1.73	0.32	0.38	1.93	0.35	1.65

MI Interp. denotes multi-image interpolation

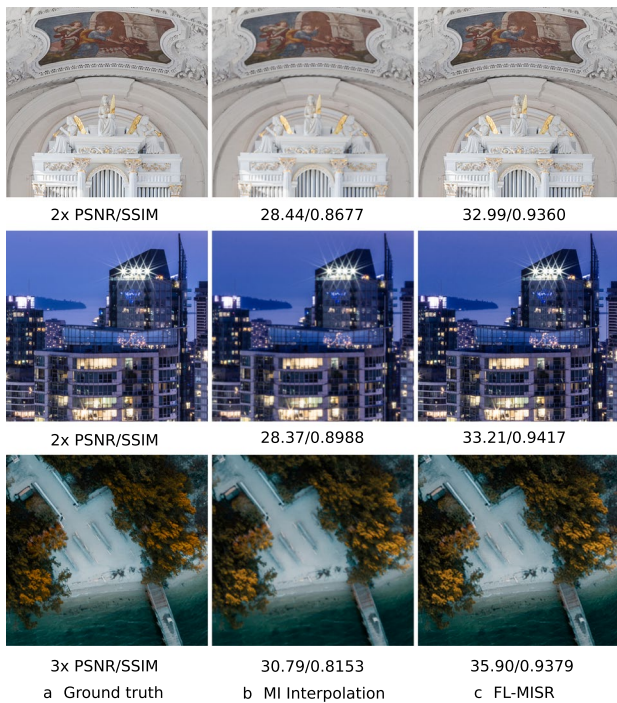


Fig. 13 Evaluation on natural image dataset DIV8K. **a** GT; **b** multi-image interpolation; **c** FL-MISR. The first two rows are for the upscaling of 2× and the bottom row is for the upscaling of 3× (Better viewed in color)

algorithm to the distributed subproblem. The local variables of the SCG algorithm are broadcast and aggregated in each iteration to synchronize the convergence rate over multi-GPUs towards a centralized optimum and consistent resolution. Furthermore, an inner-outer border exchange mechanism is performed in the overlapped regions of neighboring GPUs to avoid border effect without compromising the sharpness.

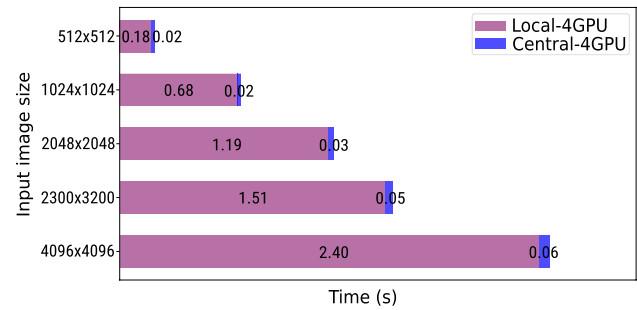


Fig. 14 Runtime distribution for the local and central computation for super-resolving images of different sizes by an upscaling of 2× under 20 SCG iterations on 4 GPUs

The proposed FL-MISR is seamlessly integrated into the computed tomography (CT) systems by super-resolving projections of the same view captured via subpixel detector shift. The SR reconstruction is performed on the fly during the CT acquisition such that no additional computation time is induced. Extensive experiments were conducted based on simulated data and real-world CT measurements of cylindrical phantoms, QRM bar pattern resolution targets, and cylindrical dry concrete joints to quantitatively and qualitatively evaluate the proposed FL-MISR. Experimental results demonstrate that the spatial resolution of CT systems is significantly improved in modulation transfer function (MTF) and visual perception by the application of FL-MISR. Moreover, comparing to a multi-core CPU implementation, the multi-GPU accelerated FL-MISR achieves a more than 50× speedup on a 4-GPU system and it is shown that the exchange of local SCG variables and overlapped regions between GPUs has limited impact on the overall runtime. Last but not least, evaluation on public dataset DIV8K shows that FL-MISR is not confined to CT imaging but also provides very promising results for natural images.

Table 4 Evaluation of computation time in terms of input image size, number of SCG iterations, and CPU/GPU platforms for the upscaling of 2× where four 16-bit input images were utilized. (N/A indicates not applicable)

Input image size	512×512			1024×1024			2048×2048			2300×3200			4096×4096		
Non-iterative method															
Multi-image interp.	0.03			0.07			0.26			0.45			2.07		
Proposed iterative FL-MISR															
SCG iterations	5	10	20	5	10	20	5	10	20	5	10	20	5	10	20
CPU* (s)	1.06	2.24	4.48	4.08	8.64	17.37	16.21	34.89	69.02	23.86	50.68	113.96	49.67	105.71	250.82
1 GPU (s)	0.08	0.13	0.25	0.22	0.42	0.78	0.70	1.30	2.43	N/A	N/A	N/A	N/A	N/A	N/A
4 GPU (s)	0.07	0.12	0.22	0.25	0.44	0.79	0.52	0.76	1.32	0.79	1.20	2.33	2.38	3.02	4.33

*CPU experiments were conducted on the Intel Xeon Gold 5120 CPU equipped with 56 cores

Acknowledgements This work was supported by the German Research Foundation (DFG, Germany) under the DFG-project SI 587/18-1 in the priority program SPP 2187.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Park, S., Park, M., Kang, M.G.: Super-resolution image reconstruction: a technical overview. *IEEE Signal Process. Mag.* **20**(5), 21–36 (2003)
- Nasrollahi, K., Moeslund, T.B.: Super-resolution: a comprehensive survey. *Mach. Vis. Appl.* **25**(6), 1423–1468 (2014)
- Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J.: Deep learning for single image super-resolution: a brief review. *IEEE Trans. Multimedia* **21**(12), 3106–3121 (2019)
- Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L.: Beyond a gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.* **26**(7), 3142–3155 (2017)
- Mildenhall, B., Barron, J.T., Chen, J., Sharlet, D., Ng, R., Carroll, R.: Burst denoising with kernel prediction networks. In: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 2502–2510 (2018)
- Dong, C., Loy, C.C., He, K., Tang, X.: Learning a deep convolutional network for image super-resolution. In: *Proc. Eur. Conf. Comput. Vis.*, pp. 184–199 (2014)
- Lim, B., Son, S., Kim, H., Nah, S., Lee, K.M.: Enhanced deep residual networks for single image super-resolution. In: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshop*, pp. 136–144 (2017)
- Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V.: VoxelMorph: a learning framework for deformable medical image registration. *IEEE Trans. Image Process.* **38**(8), 1788–1800 (2019)
- Sun, K., Simon, S.: FDRN: a fast deformable registration network for medical images. *Med. Phys.* **2021**, 1–11 (2021)
- Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: NNU-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods* **18**(2), 203–211 (2021)
- Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J.: UNet++: redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Trans. Med. Imaging* **39**(6), 1856–1867 (2020)
- Dong, C., Loy, C.C., Tang, X.: Accelerating the super-resolution convolutional neural network. In: *Proc. Eur. Conf. Comput. Vis.*, pp. 391–407 (2016)
- Kim, J., Lee, J.K., Lee, K.M.: Accurate image super-resolution using very deep convolutional networks. In: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 1646–1654 (2016)
- Zhang, K., Zuo, W., Zhang, L.: Deep plug-and-play super-resolution for arbitrary blur kernels. In: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 1671–1681 (2019)
- Wang, X., et al.: ESRGAN: Enhanced super-resolution generative adversarial networks. In: *Proc. Eur. Conf. Comput. Vis.*, pp. 1–16 (2018)
- Pham, C.H., et al.: Multiscale brain MRI super-resolution using deep 3D convolutional networks. *Comput. Med. Imaging Graph.* **77**, 101647 (2019)
- Chen, Y., Shi, F., Christodoulou, A.G., Zhou, Z., Xie, Y., Li, D.: Efficient and accurate MRI super-resolution using a generative adversarial network and 3d multi-level densely connected network. In: *Proc. Int. Conf. Med. Imag. Comp. Comput. Assist. Interv.*, pp. 91–99 (2018)
- Kappeler, A., Yoo, S., Dai, Q., Katsaggelos, A.K.: Video super-resolution with convolutional neural networks. *IEEE Trans. Comput. Imaging* **2**(2), 109–122 (2016)
- Caballero, J. et al.: Real-time video super-resolution with spatio-temporal networks and motion compensation. In: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 2848–2857 (2017)
- Haris, M., Shakhnarovich, G., Ukita, N.: Recurrent back-projection network for video super-resolution. In: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 3897–3906 (2019)
- Sajjadi, M.S.M., Vemulapalli, R., Brown, M.: Frame-recurrent video super-resolution. In: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 6626–6634 (2018)
- Sun, K., Koch, M., Wang, Z., Jovanovic, S., Rabah, H.: An FPGA-based residual recurrent neural network for real-time video super-resolution. *IEEE Trans. Circ. Syst. Video Technol.* **2021**, 1–12 (2021)
- Farsiu, S., Robinson, M.D., Elad, M., Milanfar, P.: Fast and robust multiframe super-resolution. *IEEE Trans. Image Process.* **13**(10), 1327–1344 (2004)
- Yue, L., Shen, H., Yuan, Q., Zhang, L.: A locally adaptive l1–l2 norm for multi-frame super-resolution of images with mixed noise and outliers. *Signal Process.* **105**(1), 156–174 (2014)
- Köhler, T., Huang, X., Schebesch, F., Aichert, A., Maier, A., Hornegger, J.: Robust multiframe super-resolution employing iteratively re-weighted minimization. *IEEE Trans. Comput. Imaging* **2**(1), 42–58 (2016)
- Sun, K., Tran, T., Krawtschenko, R., Simon, S.: Multi-frame super-resolution reconstruction based on mixed Poisson-Gaussian noise. *Signal Process. Image Commun.* **82**, 115736 (2020)
- Sun, K., Simon, S.: Bilateral spectrum weighted total variation for noisy-image super-resolution and image denoising. *IEEE Trans. Signal Process.* 1–13 (2021). [arXiv:2106.00768](https://arxiv.org/abs/2106.00768)
- Huang, T., Tsai, R.: Multiframe image restoration and registration. *Adv. Comput. Vis. Image Process* **1**, 317–339 (1984)
- Stark, H., Oskoui, P.: High-resolution image recovery from image-plane arrays, using convex projections. *J. Opt. Soc. Am. A* **6**(11), 1715–1726 (1989)
- Elad, M., Hel-Or, Y.: A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur. *IEEE Trans. Image Process* **10**(8), 1187–1193 (2001)
- Tipping, M.E., Bishop, C.M.: Bayesian image super-resolution. In: *Adv. Neural. Inf. Process. Syst.* **2033**, 1303–1310 (2003)

32. Wetzl, J., Taubmann, O., Haase, S., Köhler, T., Kraus, M., Hornegger, J.: GPU-accelerated time-of-flight super-resolution for image-guided surgery. In: *Bildverarbeitung für die Medizin 2013*, pp. 21–26. Springer, Berlin (2013)
33. Xu, J., Liang, Y., Liu, J., Huang, Z., Liu, X.: Online multi-frame super-resolution of image sequences. *EURASIP J. Image Video Process.* **2018**(1), 1–10 (2018)
34. Anger, J., Ehret, T., de Franchis, C., Facciolo, G.: Fast and accurate multi-frame super-resolution of satellite images. *ISPRS J. Photo. Rem. Sens.* **5**(1), 1–8 (2020)
35. Kim, Y., Choi, J., Kim, M.: A real-time convolutional neural network for super-resolution on FPGA with applications to 4K UHD 60 fps video services. *IEEE Trans. Circ. Syst. Video Technol.* **29**(8), 2521–2534 (2019)
36. Gu, S., Lugmayr, A., Danelljan, M., Fritsche, M., Lammour, J., Timofte, R.: DIV8K: diverse 8k resolution image dataset. In: *IEEE Int. Conf. Comput. Vis. Workshop*, pp. 3512–3516 (2019)
37. Sun, K., Kieß, S., Sven, S.: Spatial resolution enhancement based on detector displacement for computed tomography. In: *Proc. Conf. Industrial Computed Tomography*, pp. 1–8 (2019)
38. Rodríguez, P.: Total variation regularization algorithms for images corrupted with different noise models: a review. *J. Electr. Comput. Eng.* **2013**, 5 (2013)
39. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Phys. D: Nonlinear Phenom.* **60**(1–4), 259–268 (1992)
40. Huber, P.J.: Robust estimation of a location parameter. *Ann. Math. Stat.* **35**(1), 73–101 (1964)
41. Gilboa, G., Osher, S.: Nonlocal operators with applications to image processing. *Multiscale Model. Simul.* **7**(3), 1005–1028 (2009)
42. Møller, M.F.: A scaled conjugate gradient algorithm for fast supervised learning. *Neural Netw.* **6**(4), 525–533 (1993)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.