

Coarse to Fine: Image Restoration Boosted by Multi-Scale Low-Rank Tensor Completion

Rui Lin, Cong Chen, Ngai Wong

Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong

Email: {linrui, chencong, nwong}@eee.hku.hk

Abstract—Existing low-rank tensor completion (LRTC) approaches aim at restoring a partially observed tensor by imposing a global low-rank constraint on the underlying completed tensor. However, such a global rank assumption suffers the trade-off between restoring the originally details-lacking parts and neglecting the potentially complex objects, making the completion performance unsatisfactory on both sides. To address this problem, we propose a novel and practical strategy for image restoration that restores the partially observed tensor in a coarse-to-fine (C2F) manner, which gets rid of such trade-off by searching proper local ranks for both low- and high-rank parts. Extensive experiments are conducted to demonstrate the superiority of the proposed C2F scheme. A demo is provided in the supplementary, and the complete codes will be made available upon acceptance of this paper.

I. INTRODUCTION

In the past decade, computer vision and image processing have received immense attention due to the proliferation of deep learning and artificial intelligence (AI). Image restoration is one of the vital research topics in this field as the need for large amounts of high quality images is intrinsic to various AI tasks. Image restoration aims to recover the missing pixels in an image based on only partially observed data, thus forming a clearer image. Therefore, deciphering the relations between the known and unknown entries is a key challenge in image restoration tasks.

Existing approaches for image restoration can generally be divided into four categories [1], namely, partial differential equations (PDEs)-based diffusion, exemplar-based techniques, hybrid algorithms, and low-rank tensor completion (LRTC) methods. PDEs-based methods [2]–[4] solve the restoration problem by using smoothness priors to diffuse local structures from the known region to the missing unknown region. Whereas the exemplar-based techniques [5], [6] restore the missing parts by sampling, copying, or stitching together the observed parts of the image. Since PDEs-based diffusion works well when completing missing areas like straight lines and curves but shows cons in texture recovering that exemplar-based techniques are good at, hybrid algorithms [7], [8] are proposed to make the first two categories complementary so as to restore texture while preserving edges or structures. We remark that all these aforementioned schemes require searches to capture the local correlation or dependencies among the pixels. Different from them, LRTC methods [9]–[11] aim to dig out the statistical properties of the image and capture the global structure of the input data.



Fig. 1. There is a global low-rank assumption in most LRTC methods, which can be regarded as a trade-off between restoring the originally low-rank parts and neglecting the potentially high-rank complex objects. For easy understanding and better visualization, we use STDC [12] as an example. Under the global low-rank assumption, it is seen that the restored sky is not smooth while the window loses the details.

Although LRTC breaks through the limitation of relying only on the observed adjacent pixels, and adopts a global low-rank assumption on the completed image, we find the global low-rank setting embodies a *trade-off* between restoring the originally low-rank parts and neglecting the potentially high-rank parts. The trade-off, depicted in Fig. 1, brings about two issues: 1) the restoration of the parts associated with a lower local rank will be hindered with insufficient observed pixels, 2) the potentially high-rank complex objects will suffer the loss of details due to over-smoothing. Therefore, the performance of LRTC can potentially be enhanced by preserving the benefits brought by the global rank setting, while catering for local structures simultaneously.

To this end, we propose a novel multi-scale LRTC strategy that restores the partially observed image in a Coarse-to-Fine (C2F) manner. Succinctly, the image is first completed in the coarse stage, where the whole image is restored by assuming a global low-rank tensor structure. After capturing the overall data structure in the first stage, the partially observed image is divided into smaller patches in the subsequent fine stage, where the patches are completed independently using the same LRTC method. The restored patches then replace their counterparts in the coarse completed image if they fulfill a designated metric. Noticeably, the fine-grained stage can be repeated multiple times to seek a satisfactory restored image, and the experimental results demonstrate the necessity of such *successive* fine-grained completion. We highlight that the performance of existing LRTC methods can be boosted taking advantage of this *divide-and-conquer* strategy. Our main contributions are:

- A general and intuitive C2F strategy is proposed, which effectively enhances the performance of existing LRTC methods by seeking proper local ranks for both the low-

and high-rank parts, respectively.

- No more trade-off that aims to balance the restoration performance of simple and complex parts is required.
- Utilization of the data from both coarse and fine hierarchies, thus capturing both the global and local data structures simultaneously.
- Extensive experiments and ablation study for validating C2F, which demonstrate the superiority of C2F in image completion versus existing algorithms.

II. RELATED WORKS

The methods to determine the global low-rank constraints in the LRTC optimization problem can generally be divided into two categories, namely, by generalizing the low-rank matrix completion techniques or by employing tensor decomposition.

For the first category, Liu et al. [13] are the pioneers that proposed tensor trace norm and defined it as the average of trace norms of all unfolded matrices. The tensor with missing entries is restored by minimizing this metric. However, solving the optimization problem is nontrivial due to the inter-dependency between the unfolded matrices. Therefore, two enhanced methods (viz. FaLRTC and HaLRTC) are proposed [9] wherein a relaxation is employed to separate the dependent relations between the unfolded matrices such that a globally optimal solution is achieved. However, implementing the singular value decomposition needed in the tensor trace norm is computationally expensive. To alleviate the computational burden, a parallel matrix factorization approach [14] is proposed, which expresses each unfolded matrix of the underlying tensor as the product of two low-rank factors that are alternately updated.

The second category employs tensor decomposition techniques. For example, Canonical Polyadic (CP) decomposition [15] is applied on the partially observed tensor in [16], [17], which predicts the missing entries through the low-rank constraint. In [18], the underlying tensor structure is assumed to be a low-rank Tensor Train (TT) [19], whose TT-cores are updated by solving the alternating least squares problem. However, manually specifying the suitable tensor ranks for these methods is challenging. To avoid that, the authors of [20] and [21] applied trace norm regularization to the factor matrices of CP and Tucker decomposition [15], respectively. The low-rank constraint is then achieved by adjusting the weight of trace norm regularization term in the optimization problem.

Recently, Xue et al. [22] proposed a completion approach using multilayer sparsity-based tensor decomposition, which further decomposes the CP factors several times. We remark that their algorithm still has the global low-rank assumption for the first CP decomposition, on which all the following decomposition is based. In contrast, the fine-grained completion in our strategy is not entirely dependent on the coarse stage, which instead allows the LRTC methods to correct the global assumption if the ranks are improper in the former stages. Multiscale feature (MSF) is proposed in [23], which aims to build a higher-order tensor containing more image

information, and then applies TT decomposition on it. In fact, MSF-TT is orthogonal to the proposed C2F strategy, whose performance can be further improved by the latter.

In short, the performance of existing LRTC methods belonging to *either* category can be enhanced with C2F, which is able to exploit both the global and local tensor rank information, allowing a more comprehensive utilization of the partially observed data.

III. PRELIMINARY

Tensors are multi-mode arrays, and a d -way tensor is denoted as $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_d}$. The element of \mathcal{A} is represented by $\mathcal{A}(i_1, i_2, \dots, i_d)$, where $1 \leq i_k \leq I_k$, $k = 1, 2, \dots, d$. The numbers I_1, I_2, \dots, I_d are called the dimensions of the tensor \mathcal{A} . We use boldface capital calligraphic letters $\mathcal{A}, \mathcal{B}, \dots$ to denote tensors, italic boldface capital letters $\mathbf{A}, \mathbf{B}, \dots$ to denote matrices, italic boldface letters $\mathbf{a}, \mathbf{b}, \dots$ to denote vectors, and roman letters a, b, \dots to denote scalars.

Definition 1. (*Tensor mode- k matricization* [15], p. 459) The mode- k matricization of a tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_k \times \cdots \times I_d}$ denoted by $\mathbf{A}_{(k)} \in \mathbb{R}^{I_k \times (I_1 \times \cdots \times I_{k-1} \times I_{k+1} \times \cdots \times I_d)}$ and tensor element $\mathcal{A}(i_1, i_2, \dots, i_d)$ maps to matrix element $\mathbf{A}_{(k)}(i_k, j)$, where

$$j = 1 + \sum_{n=1, n \neq k}^d (i_n - 1) J_n \quad \text{with} \quad J_n = \prod_{m=1, m \neq k}^{n-1} I_m. \quad (1)$$

Definition 2. (*Tensor mode- k product* [15], p. 460) Tensor mode- k product involves a multiplication of a matrix with a d -way tensor along one of its d modes. The mode- k product $\mathcal{B} = \mathcal{A} \times_k \mathbf{U}$ of a tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_k \times \cdots \times I_d}$ with a matrix $\mathbf{U} \in \mathbb{R}^{R \times I_k}$ is defined by

$$\begin{aligned} & \mathcal{B}(i_1, \dots, i_{k-1}, j, i_{k+1}, \dots, i_d) \\ &= \sum_{i_k=1}^{I_k} \mathbf{U}(j, i_k) \mathcal{A}(i_1, \dots, i_{k-1}, i_k, i_{k+1}, \dots, i_d), \end{aligned} \quad (2)$$

where $\mathcal{B} \in \mathbb{R}^{I_1 \times \cdots \times I_{k-1} \times R \times I_{k+1} \times \cdots \times I_d}$.

Definition 3. (*Tensor Frobenius norm*) The Frobenius norm of a tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_d}$ is defined as

$$\|\mathcal{A}\|_F = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_d=1}^{I_d} \mathcal{A}(i_1, i_2, \dots, i_d)^2}. \quad (3)$$

Definition 4. (*Tucker decomposition*) The Tucker decomposition decomposes a d -way tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_d}$ into a smaller Tucker core $\mathcal{S} \in \mathbb{R}^{R_1 \times \cdots \times R_d}$ and orthogonal factor matrices $\mathbf{U}_i \in \mathbb{R}^{I_i \times R_i}$ ($i = 1, \dots, d$) as

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}_1 \times_2 \cdots \times_d \mathbf{U}_d. \quad (4)$$

IV. LOW-RANK TENSOR COMPLETION BOOSTED BY COARSE TO FINE STRATEGY

For easy understanding, we now exemplify C2F using a representative LRTC method called LRTC-TV-II [24], which is based on Tucker decomposition. In Section IV-A, we

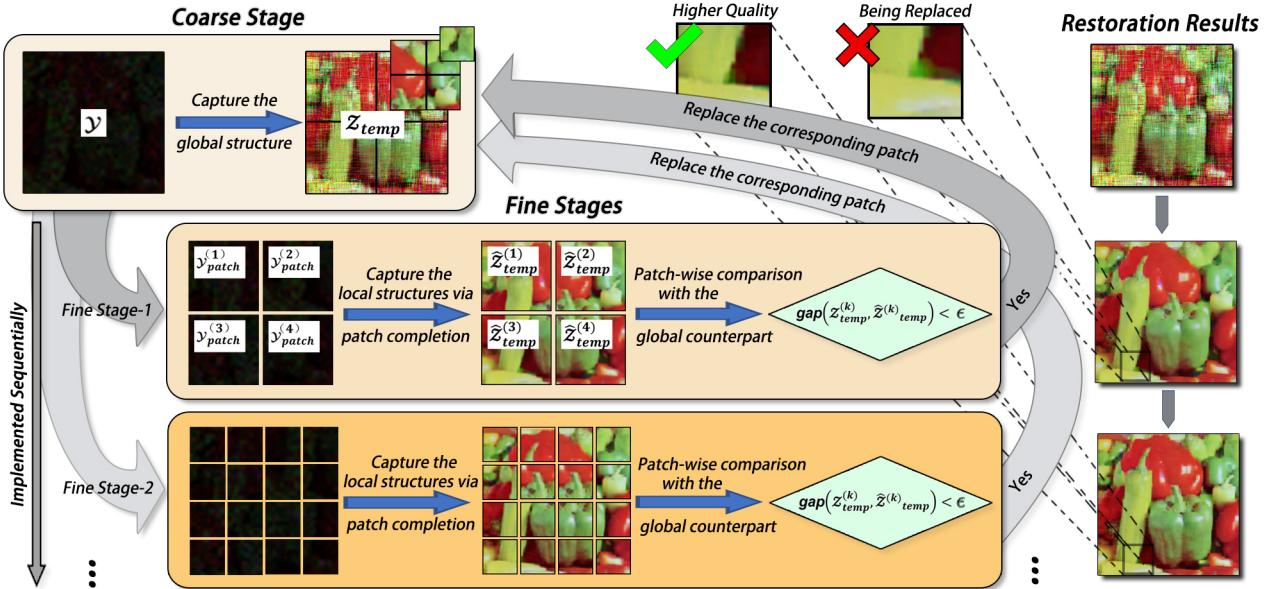


Fig. 2. The workflow of the C2F strategy. First, a coarse completed image \mathcal{Z}_{temp} is obtained. Then, for each fine stage, the original partially observed image \mathcal{Y} is separated into N same-size patches, namely $\mathcal{Y}_{patch}^{(1)}, \mathcal{Y}_{patch}^{(2)}, \dots, \mathcal{Y}_{patch}^{(N)}$. Each patch has overlapping pixels with its neighbors to achieve a smooth patch boundary during subsequent completion. Those patches are then completed separately via LRTC method, and N fine-grained completed patches can be obtained, namely, $\hat{\mathcal{Z}}_{temp}^{(1)}, \dots, \hat{\mathcal{Z}}_{temp}^{(N)}$. The newly obtained patches satisfying the criterion will replace the corresponding patches in the \mathcal{Z}_{temp} image. The number of fine stages can be set according to the restoration requirements.

first introduce the problem formulation. In Section IV-B, we elaborate the details of the proposed scheme and show the C2F algorithm. Finally, we discuss the necessity of the successive fine-grained completion in Section IV-C.

A. Problem Formulation

Given the partially observed tensor $\mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}$, and the observed tensor entry set Ω , the image restoration problem solved by LRTC-TV-II [24] is as follows:

$$\begin{aligned} & \min_{\mathcal{Z}, \mathcal{S}, \{\mathbf{U}_k\}_{k=1}^d} \quad \lambda_1 \sum_{k=1}^d \beta_k |\mathbf{F}_k \mathcal{Z}_{(k)}| + \lambda_2 \sum_{k=1}^d \|\mathbf{U}_k\|_* + \lambda_3 \|\mathcal{S}\|_F^2 \\ & \text{subject to} \quad \mathcal{Z} = \mathcal{S} \times_1 \mathbf{U}_1 \times_2 \dots \times_d \mathbf{U}_d, \\ & \quad [\mathcal{Z}]_\Omega = [\mathcal{Y}]_\Omega, \end{aligned} \quad (5)$$

where \mathcal{Z} is the restored tensor, $\mathcal{Z}_{(k)}$ is the mode- k matricization of \mathcal{Z} . The binary matrix $\mathbf{F}_k \in \mathbb{R}^{(I_k-1) \times I_k}$ has the entries $\mathbf{F}_k(i, i) = 1$ and $\mathbf{F}_k(i, i+1) = -1$, while the remaining elements are all zeros. The term $|\mathbf{F}_k \mathcal{Z}_{(k)}|$ represents the total variation (TV) regularization and is computed by summing over the absolute values of all entries in $\mathbf{F}_k \mathcal{Z}_{(k)}$. The operator $\|\cdot\|_*$ denotes the trace norm of a matrix. The completed tensor \mathcal{Z} is assumed to be in a low-rank Tucker format, namely, $\mathcal{Z} = \mathcal{S} \times_1 \mathbf{U}_1 \times_2 \dots \times_d \mathbf{U}_d$, where $\mathcal{S} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}$ and $\mathbf{U}_k \in \mathbb{R}^{I_k \times I_k}$. Note that instead of setting low Tucker ranks directly, the objective (5) employs a trace norm term to enforce the low-rank constraint in the factor matrix \mathbf{U}_k . The last term in (5) avoids overfitting, and β_k is either 0 or 1 to indicate whether TV normalization is imposed on mode- k of \mathcal{Z} . For color image completion tasks, the setting would

be $\beta_1 = \beta_2 = 1$ and $\beta_3 = 0$, since only spatial modes are expected to be smooth. The above optimization problem can be efficiently solved by alternating direction method of multipliers (ADMM) [25], which introduces auxiliary matrices to separate the inter-dependency between the matrices of (5).

B. C2F Strategy

As shown in Fig. 1, the existing LRTC methods aim to strike a balance between the potentially low- and high-rank parts, which is achieved by imposing a global low-rank assumption. However, our proposed C2F strategy gets rid of such trade-off by utilizing the data from both coarse and fine hierarchies, and focusing more on the local structure (small patches) in a successive manner. Here we present the details of the C2F strategy, which generally consists of two tensor completion stages, namely, the coarse stage and the fine stage.

1) *Coarse Stage*: In the coarse stage, the whole observed image \mathcal{Y} is completed by solving the optimization problem (5) with LRTC-TV-II, and a coarsely completed image \mathcal{Z}_{temp} is obtained. Considering the limits brought by the global low-rank assumption on restoring both the details-lacking (low-rank) parts and the complex (high-rank) parts, we proceed to the fine stage to pursue a higher-quality completion result.

2) *Fine Stage*: The fine stage aims to complete the image from a fine-grained viewpoint, which can be repeated multiple times until a satisfactory restored image is achieved. Every single fine stage comprises three steps, namely, fine-grained completion, patch comparison, followed by replacement.

Step 1: Fine-grained Completion. The original partially observed image \mathcal{Y} is separated into N patches with the same

Algorithm 1 C2F Strategy with LRTC-TV-II [24]

Input: Partially observed image \mathcal{Y} ; observed image entry set Ω ; patch replace threshold ϵ ; number of fine stage implementations F ; weight parameters λ_k ($k = 1, 2, 3$).
Output: The restored image \mathcal{Z} .

```

1:  $\mathcal{Z}_{temp} \leftarrow$  Coarse stage completion;
2:  $f = 1 \leftarrow$  Count the number of fine stages;
3: while  $f \leq F$  do
4:    $\mathcal{Y}_{patch}^{(k)}, k = 1, \dots, 2^{2f} \leftarrow$  Divide  $\mathcal{Y}$  into  $2^{2f}$  patches of the same size;
5:    $\lambda_2 = \mu\lambda_2, \mu > 1 \leftarrow$  Adjust  $\lambda_2$  to impose a local rank assumption;
6:    $\mathcal{Z}_{temp}^{(k)}, k = 1, \dots, 2^{2f} \leftarrow$  Divide  $\mathcal{Z}_{temp}$  according to the same partition used for  $\mathcal{Y}$  in step 4;
7:   for  $k$  from 1 to  $2^{2f}$  do
8:      $\hat{\mathcal{Z}}_{temp}^{(k)} \leftarrow$  Complete the  $k$ -th patch;
9:     if  $\text{gap}(\mathcal{Z}_{temp}^{(k)}, \hat{\mathcal{Z}}_{temp}^{(k)}) < \epsilon$  then
10:       $\mathcal{Z}_{temp}^{(k)} = \hat{\mathcal{Z}}_{temp}^{(k)} \leftarrow$  Replace its counterpart;
11:    end if
12:   end for
13:    $f = f + 1 \leftarrow$  Update the number of fine stages;
14:   Update the threshold  $\epsilon$  (cf. (9));
15: end while
16:  $\mathcal{Z} = \mathcal{Z}_{temp}$ .

```

size, namely $\mathcal{Y}_{patch}^{(1)}, \mathcal{Y}_{patch}^{(2)}, \dots, \mathcal{Y}_{patch}^{(N)}$. To obtain a smooth boundary during the subsequent completion process, we let every single patch have *overlapping* pixels with its *neighbors*. Those patches are then completed separately with the selected LRTC-TV-II technique with a local rank assumption on them, which is achieved by adjusting λ_2 in (5). The local rank is smaller than the global rank we set at the very beginning in the coarse stage and this setting is validated in Section V-C.

Step 2: Patch Comparison. After the first fine stage completion, N completed patches, $\hat{\mathcal{Z}}_{temp}^{(1)}, \hat{\mathcal{Z}}_{temp}^{(2)}, \dots, \hat{\mathcal{Z}}_{temp}^{(N)}$, under local rank assumption are obtained. To compare those patches at the same position in different completion stages, we divide the temporarily restored image \mathcal{Z}_{temp} into N patches following the partitions employed in the current fine stage to obtain $\mathcal{Z}_{temp}^{(1)}, \mathcal{Z}_{temp}^{(2)}, \dots, \mathcal{Z}_{temp}^{(N)}$. We use Relative Squared Error (RSE) to quantitatively compare the difference between the two corresponding patches, which is formulated as below:

$$\text{gap}(\hat{\mathcal{Z}}_{temp}^{(k)}, \mathcal{Z}_{temp}^{(k)}) = \frac{\|\hat{\mathcal{Z}}_{temp}^{(k)} - \mathcal{Z}_{temp}^{(k)}\|_F}{\|\mathcal{Z}_{temp}^{(k)}\|_F}. \quad (6)$$

The greater the value, the greater the difference between the completion results of the selected patch in the different stages.

Step 3: Replacement. Based on the difference between the two corresponding patches, we rely on a pre-defined threshold ϵ to decide whether to replace $\mathcal{Z}_{temp}^{(k)}$ with $\hat{\mathcal{Z}}_{temp}^{(k)}$:

$$\mathcal{Z}_{temp}^{(k)} = \begin{cases} \mathcal{Z}_{temp}^{(k)}, & \text{if } \text{gap}(\hat{\mathcal{Z}}_{temp}^{(k)}, \mathcal{Z}_{temp}^{(k)}) \geq \epsilon, \\ \hat{\mathcal{Z}}_{temp}^{(k)}, & \text{if } \text{gap}(\hat{\mathcal{Z}}_{temp}^{(k)}, \mathcal{Z}_{temp}^{(k)}) < \epsilon. \end{cases} \quad (7)$$

The reason to apply this replacement scheme is that the completed image in the coarse stage captures the overall structure of the underlying ground truth, if the newly completed patch $\hat{\mathcal{Z}}_{temp}^{(k)}$ has a large gap between its counterpart in \mathcal{Z}_{temp} , it risks deviating from the ground truth severely. Therefore, we only replace those eligible newly completed patches into the corresponding positions of \mathcal{Z}_{temp} , where the overlapping region of patches takes the averaged pixel values.

Since the assumed local rank is not necessarily suitable for the patches in the fine stage, we prefer to implement the fine stage several times in a sequential manner. The size of the separated patch decreases along with the sequence of fine stages. By repeating the fine stages, both the complex and details-lacking parts are more likely to meet the suitable tensor rank benefiting from the simultaneously reduced patch size and assumed local rank in a divide-and-conquer manner. The whole procedure of C2F strategy with specific LRTC-TV-II is summarized in Algorithm 1, and the detailed workflow of the *general* C2F strategy that can be combined with any LRTC methods is displayed in Fig. 2.

C. Necessity of Successive Fine-grained Completion

The earlier fine stages with relatively large patches are expected to capture the sub-global structure of the whole image. If all intermediate fine stages are dropped, the transition from global to local viewpoint will be too rapid, which ignores the important sub-global relations between patches and weight too much on the fine-grained neighbors. The omission leads to restored results that are far from the ground truth. In Section V-B, experimental results are provided which compare the proposed C2F with the Short-Cut C2F (which contains only the coarse stage at the beginning and the last fine stage with the smallest patches) to verify this statement.

V. EXPERIMENTAL RESULTS

Extensive numerical experiments are conducted to evaluate the color image restoration performance of the pure LRTC methods and different versions of C2F strategy. In Section V-A, we select four LRTC algorithms with publicly available official implementations, namely, HaLRTC¹ [9], STDC² [12], LRTC-TV-II³ [24], and LRTC-PDS⁴ [10], comparing their performance with and without the proposed C2F strategy. Fig. 3 shows the ground truth of the eight benchmark color images, and the size of each image is $256 \times 256 \times 3$. In Section V-B, we compare the proposed C2F with its short-cut version to demonstrate the effectiveness of the gradual refinement. Additionally, in Section V-C, we demonstrate the rationality of the decreasing local rank setting employed in our experiments. All computations were done on an Intel(R) Core(TM) i5-6500 processor running at 3.2GHz with 16GB RAM, and the implementation platform is MATLAB 2020b.

¹The code of HaLRTC is available at: <https://github.com/halrtc>

²The code of STDC is available at: http://mp.cs.nthu.edu.tw/project_stdc

³The code of LRTC-TV-II is available at: <https://xutaoli.weebly.com>

⁴The code of LRTC-PDS is available at: https://drive.google.com/lrc_pds



Fig. 3. Ground truth of the eight benchmark color images.

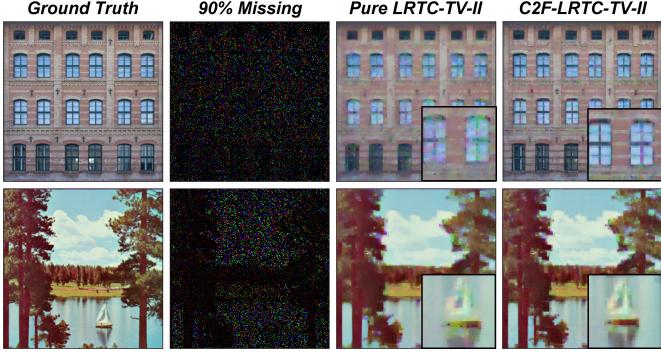


Fig. 4. The C2F strategy helps LRTC methods to restore more details.

A. Pure LRTC Compared with C2F-LRTC

To evaluate the performance of image completion, we employ the RSE and peak signal-to-noise ratio (PSNR) as the evaluation metrics, which are widely used in image restoration tasks. The definition of those two metrics are as follows:

$$\text{RSE} = \frac{\|\mathcal{Z} - \mathcal{Z}_{true}\|_F}{\|\mathcal{Z}_{true}\|_F}$$

$$\text{PSNR} = 10 \log_{10} \frac{\prod_{k=1}^d \hat{\mathcal{Z}}_{true}^2}{\|\mathcal{Z} - \mathcal{Z}_{true}\|_F^2} \quad (8)$$

where \mathcal{Z} , \mathcal{Z}_{true} and $\hat{\mathcal{Z}}_{true}$ represent the completed image, ground truth image, and the maximum value of the pixels in the ground truth image, respectively. The performance of the restoration method is inversely proportional to the value of RSE and proportional to the value of PSNR.

For the C2F implementation, we set the number of fine stage repetitions $F = 3$, and the initial patch replace threshold $\epsilon = 0.15$. For the second and the third fine stages, the replacement threshold will be updated as below:

$$\epsilon = \begin{cases} \frac{3}{2} \max \left(\text{gap}(\hat{\mathcal{Z}}_{temp}^{(k)}, \mathcal{Z}_{temp}^{(k)}) \right), & \text{if } F = 2, \\ \frac{3}{2} \cdot \max \left(\text{gap}(\hat{\mathcal{Z}}_{temp}^{(k)}, \mathcal{Z}_{temp}^{(k)}) \mathbf{r}_k \right), & \text{if } F = 3, \end{cases} \quad (9)$$

where the vector \mathbf{r} is binary with elements 0 or 1, recording whether a completed patch in the second fine stage replaces its counterpart in the coarse stage or not. We remark that the threshold becomes larger as the patch size decreases.

Table III summarizes the comparison results of the pure LRTC and their C2F version under three different missing

TABLE I
PERFORMANCE COMPARISON BETWEEN THREE DIFFERENT COMPLETION STRATEGIES COMBINED WITH LRTC-TV-II [24]. THE MISSING RATIO OF THE EIGHT BENCHMARK COLOR IMAGES IS ALL 90%.

Method	Airplane		Baboon		Barbara		Facade	
	PSNR	RSE	PSNR	RSE	PSNR	RSE	PSNR	RSE
LRTC-TV-II	22.80	0.098	21.22	0.165	23.87	0.131	21.79	0.158
C2F-LRTC-TV-II	23.28	0.085	23.05	0.161	24.75	0.119	24.40	0.117
Short-Cut C2F	22.77	0.093	19.64	0.193	20.51	0.193	21.07	0.171

Method	House		Lena		Peppers		Sailboat	
	PSNR	RSE	PSNR	RSE	PSNR	RSE	PSNR	RSE
LRTC-TV-II	24.99	0.094	25.36	0.097	22.04	0.144	20.67	0.162
C2F-LRTC-TV-II	25.37	0.088	25.51	0.095	23.46	0.122	21.61	0.146
Short-Cut C2F	22.34	0.125	23.69	0.118	20.57	0.171	19.72	0.181

TABLE II
THE AVERAGE RELATIVE PATCH RANK (RPR) FOR PATCHES OF DIFFERENT SIZES IN THE GROUND TRUTH. IN EACH COLUMN, THE AVERAGE RPR DECREASES AS THE PATCH SIZE BECOMES SMALLER, WHICH IS CONSISTENT WITH OUR EXPERIMENTAL SETTINGS.

	Airplane	Baboon	Barbara	Facade	House	Lena	Peppers	Sailboat
Coarse	0.2654	0.4889	0.3359	0.3190	0.2405	0.3072	0.2954	0.3647
Fine-1	0.2053	0.4366	0.2953	0.2651	0.2071	0.2572	0.2621	0.3068
Fine-2	0.1621	0.4141	0.2615	0.2266	0.1723	0.2249	0.2286	0.2801
Fine-3	0.1415	0.3867	0.2289	0.2004	0.1561	0.1978	0.2078	0.2529

ratios (viz. 0.7, 0.8, and 0.9). For easy reading, the C2F-LRTC results are marked blue, and are put below the corresponding pure LRTC method. It can be observed that under all missing ratios, the performance of *every* pure LRTC method is enhanced after employing the C2F strategy. In other words, C2F-LRTC can restore images with larger PSNR and smaller RSE values. Furthermore, it can be seen that as the missing ratio increases, C2F has a trend to boost the performance of pure LRTC methods more significantly. For example, when completing the image Barbara, the gaps between the PSNR of STDC and C2F-STDC are 0.13, 1.06, and 4.13 as the missing ratio increases. To sense the improvement brought by C2F more intuitively, Fig. 4 shows the restoration results of facade and sailboat under 90% missing ratio. These two images are representative, since the image facade has *regular patterns* while the image sailboat contains *both the details-lacking parts* (e.g., the sky and the lake) and the *complex objects* (e.g., the trees and the boat). It is noticeable that C2F-LRTC restores both two kinds of images with richer details. In summary, C2F can steadily improve the performance of existing LRTC methods and retain more details.

B. The Effectiveness of Successive Fine-grained Completion

In this section, we validate the effectiveness of the successive fine-grained completion by comparing the restoration results obtained by pure LRTC, C2F-LRTC, and the Short-Cut C2F-LRTC, which only contains the coarse stage at the beginning and the last fine stage with the smallest patches.

The results are displayed in Table I, where we employ the LRTC-TV-II algorithm and set the missing ratio as 90%. For the Short-Cut C2F, we set a larger patch replace threshold $\epsilon = 0.3$ to ensure that the newly completed patches in the last

TABLE III
THE RESTORATION PERFORMANCE OF PURE LRTC AND C2F-LRTC METHODS UNDER DIFFERENT MISSING RATIOS.

Airplane	0.7		0.8		0.9		Baboon	0.7		0.8		0.9	
	PSNR	RSE	PSNR	RSE	PSNR	RSE		PSNR	RSE	PSNR	RSE	PSNR	RSE
HaLRTC	24.50	0.074	21.97	0.099	18.97	0.140	HaLRTC	21.94	0.148	20.48	0.175	18.60	0.220
C2F-HaLRTC	24.76	0.072	22.26	0.096	19.45	0.133	C2F-HaLRTC	22.16	0.144	20.69	0.171	18.62	0.217
STDC	22.77	0.090	18.67	0.145	15.25	0.236	STDC	17.19	0.256	16.50	0.277	14.48	0.350
C2F-STDC	23.19	0.086	21.16	0.137	20.13	0.123	C2F-STDC	20.19	0.176	19.50	0.197	17.19	0.256
LRTC-TV-II	26.91	0.056	27.33	0.068	22.80	0.098	LRTC-TV-II	23.30	0.129	22.32	0.143	21.22	0.165
C2F-LRTC-TV-II	27.84	0.051	27.35	0.067	23.28	0.085	C2F-LRTC-TV-II	23.47	0.126	22.52	0.141	23.05	0.161
LRTC-PDS	25.42	0.072	23.55	0.089	20.72	0.144	LRTC-PDS	23.17	0.136	22.07	0.154	20.72	0.245
C2F-LRTC-PDS	25.79	0.066	23.83	0.084	21.15	0.140	C2F-LRTC-PDS	23.25	0.131	22.15	0.150	20.75	0.239
Barbara	0.7		0.8		0.9		Facade	0.7		0.8		0.9	
	PSNR	RSE	PSNR	RSE	PSNR	RSE		PSNR	RSE	PSNR	RSE	PSNR	RSE
HaLRTC	25.27	0.112	22.65	0.151	19.06	0.229	HaLRTC	28.58	0.083	26.09	0.110	22.57	0.144
C2F-HaLRTC	25.72	0.106	22.87	0.147	19.26	0.223	C2F-HaLRTC	29.74	0.063	27.54	0.081	24.87	0.111
STDC	22.04	0.162	19.84	0.209	16.28	0.315	STDC	26.54	0.091	24.54	0.115	20.95	0.153
C2F-STDC	22.17	0.160	20.90	0.207	20.41	0.196	C2F-STDC	26.64	0.090	24.56	0.114	22.62	0.143
LRTC-TV-II	27.88	0.085	26.06	0.100	23.87	0.131	LRTC-TV-II	27.19	0.092	25.76	0.099	21.79	0.158
C2F-LRTC-TV-II	28.46	0.078	26.69	0.097	24.75	0.119	C2F-LRTC-TV-II	27.89	0.068	25.80	0.091	24.40	0.117
LRTC-PDS	26.86	0.107	25.13	0.132	22.13	0.185	LRTC-PDS	24.01	0.130	22.21	0.159	19.93	0.246
C2F-LRTC-PDS	27.00	0.098	25.38	0.124	22.34	0.178	C2F-LRTC-PDS	24.41	0.123	22.58	0.153	20.09	0.239
House	0.7		0.8		0.9		Lena	0.7		0.8		0.9	
	PSNR	RSE	PSNR	RSE	PSNR	RSE		PSNR	RSE	PSNR	RSE	PSNR	RSE
HaLRTC	25.79	0.084	23.72	0.107	20.54	0.154	HaLRTC	25.86	0.093	23.19	0.125	19.77	0.185
C2F-HaLRTC	25.92	0.830	24.07	0.103	21.05	0.145	C2F-HaLRTC	25.99	0.091	23.28	0.124	20.15	0.177
STDC	24.28	0.100	22.50	0.123	17.54	0.218	STDC	23.22	0.125	20.45	0.171	16.64	0.266
C2F-STDC	24.73	0.095	22.88	0.118	21.91	0.132	C2F-STDC	22.64	0.118	21.56	0.169	20.84	0.164
LRTC-TV-II	28.27	0.065	27.37	0.074	24.99	0.094	LRTC-TV-II	29.62	0.059	27.52	0.072	25.36	0.097
C2F-LRTC-TV-II	28.34	0.064	27.40	0.073	25.37	0.088	C2F-LRTC-TV-II	30.04	0.058	27.92	0.070	25.51	0.095
LRTC-PDS	27.19	0.075	25.63	0.096	22.77	0.146	LRTC-PDS	28.48	0.086	26.39	0.106	22.40	0.141
C2F-LRTC-PDS	27.30	0.066	25.74	0.089	22.85	0.142	C2F-LRTC-PDS	28.55	0.077	26.49	0.099	22.46	0.135
Peppers	0.7		0.8		0.9		Sailboat	0.7		0.8		0.9	
	PSNR	RSE	PSNR	RSE	PSNR	RSE		PSNR	RSE	PSNR	RSE	PSNR	RSE
HaLRTC	22.82	0.132	20.31	0.176	16.86	0.262	HaLRTC	22.78	0.127	20.60	0.164	17.81	0.226
C2F-HaLRTC	23.59	0.121	20.72	0.168	16.86	0.261	C2F-HaLRTC	23.09	0.123	20.87	0.159	17.93	0.223
STDC	21.37	0.156	18.61	0.214	16.31	0.347	STDC	20.58	0.184	18.10	0.245	14.87	0.317
C2F-STDC	21.48	0.154	20.64	0.210	18.42	0.219	C2F-STDC	20.85	0.178	18.23	0.185	19.12	0.194
LRTC-TV-II	27.75	0.076	25.25	0.103	22.04	0.144	LRTC-TV-II	25.21	0.103	23.42	0.118	20.67	0.162
C2F-LRTC-TV-II	28.37	0.072	26.17	0.096	23.46	0.122	C2F-LRTC-TV-II	26.61	0.092	23.76	0.112	21.61	0.146
LRTC-PDS	27.05	0.102	24.06	0.131	20.09	0.194	LRTC-PDS	23.95	0.123	22.13	0.153	19.37	0.209
C2F-LRTC-PDS	27.33	0.092	24.31	0.122	20.35	0.188	C2F-LRTC-PDS	24.33	0.117	22.46	0.147	19.58	0.205

fine stage can replace their counterparts in the coarse stage. For the remaining hyper-parameters, we keep them the same as in the original C2F. It is seen that instead of improving the performance of pure LRTC, the Short-Cut C2F *degrades* the restoration ability of the pure LRTC, which demonstrates the importance of the successive refinement in our C2F procedure.

C. Local Rank Analysis

We then validate the correctness of our assumption that smaller local ranks go along with decreasing patch sizes. We compute the average relative patch ranks (RPR) for patches of different sizes in the ground truth. The RPR is defined as the ratio between the number of singular values (SVs), which accounts for 90% of the total singular values summation, and the patch size. Along this side, the patch size can be ignored when comparing the patch ranks. The average RPR in the coarse and three fine stages are listed in Table II. For all eight images, it is seen that the smaller the patch size, the smaller RPR, which we adopt as a general rule applicable to most real-life images. We emphasize that the decreasing local ranks in the fine stages will not degrade the global rank of

the restored image. For instance, the global ranks (number of SVs that takes up 90% of the summation of the total SVs) of the image facade restored by the pure LRTC-TV-II and C2F-LRTC-TV-II in Fig. 4 are 49 and 50, respectively. Therefore, we conclude that it is insufficient to set a global rank only to complete partially observed images, and appropriate local ranks are important to restore local structures.

VI. CONCLUSION

This work presents a general and easily implementable coarse-to-fine framework called C2F, which can effectively boost the performance of any existing LRTC approach with the mindfully designed progressive refinement process and appropriate local ranks. To the best of our knowledge, it is the first time that global and local ranks are both utilized for completing images. By employing the proposed C2F strategy, both the global and local tensor structures are utilized, leading to a higher completion performance. Extensive experiments are done to demonstrate the practicality of C2F, which sheds light on boosting images restoration results and LRTC techniques performance in a divide-and-conquer manner.

REFERENCES

- [1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, “Image inpainting,” in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, 2000, pp. 417–424.
- [2] F. Voci, S. Eiho, N. Sugimoto, and H. Sekibuchi, “Estimating the gradient in the perona-malik equation,” *IEEE Signal Processing Magazine*, vol. 21, no. 3, pp. 39–65, 2004.
- [3] A. Telea, “An image inpainting technique based on the fast marching method,” *Journal of graphics tools*, vol. 9, no. 1, pp. 23–34, 2004.
- [4] B. R. Kumar and A. Halim, “A linear fourth-order pde-based gray-scale image inpainting model,” *Computational and Applied Mathematics*, vol. 38, no. 1, p. 6, 2019.
- [5] J.-F. Aujol, S. Ladjal, and S. Masnou, “Exemplar-based inpainting from a variational point of view,” *SIAM Journal on Mathematical Analysis*, vol. 42, no. 3, pp. 1246–1285, 2010.
- [6] O. Le Meur, J. Gautier, and C. Guillemot, “Examplar-based inpainting based on local geometry,” in *2011 18th IEEE international conference on image processing*. IEEE, 2011, pp. 3401–3404.
- [7] J.-L. Starck, M. Elad, and D. L. Donoho, “Image decomposition via the combination of sparse representations and a variational approach,” *IEEE transactions on image processing*, vol. 14, no. 10, pp. 1570–1582, 2005.
- [8] A. Bugeau, M. Bertalmío, V. Caselles, and G. Sapiro, “A comprehensive framework for image inpainting,” *IEEE transactions on image processing*, vol. 19, no. 10, pp. 2634–2645, 2010.
- [9] J. Liu, P. Musialski, P. Wonka, and J. Ye, “Tensor completion for estimating missing values in visual data,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 1, pp. 208–220, 2012.
- [10] T. Yokota and H. Hontani, “Simultaneous visual data completion and denoising based on tensor rank and total variation minimization and its primal-dual splitting algorithm,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3732–3740.
- [11] J. D. Hauenstein, L. Oeding, G. Ottaviani, and A. J. Sommese, “Homotopy techniques for tensor decomposition and perfect identifiability,” *Journal für die reine und angewandte Mathematik (Crelles Journal)*, vol. 2019, no. 753, pp. 1–22, 2019.
- [12] Y.-L. Chen, C.-T. Hsu, and H.-Y. M. Liao, “Simultaneous tensor decomposition and completion using factor priors,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 3, pp. 577–591, 2013.
- [13] J. Liu, P. Musialski, P. Wonka, and J. Ye, “Tensor completion for estimating missing values in visual data,” in *2009 IEEE 12th International Conference on Computer Vision (ICCV)*. Los Alamitos, CA, USA: IEEE Computer Society, oct 2009, pp. 2114–2121. [Online]. Available: <https://doi.ieeecomputersociety.org/10.1109/ICCV.2009.5459463>
- [14] Y. Xu, R. Hao, W. Yin, and Z. Su, “Parallel matrix factorization for low-rank tensor completion,” *arXiv preprint arXiv:1312.1254*, 2013.
- [15] T. G. Kolda and B. W. Bader, “Tensor decompositions and applications,” *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.
- [16] G. Tomasi and R. Bro, “Parafac and missing values,” *Chemometrics and Intelligent Laboratory Systems*, vol. 75, no. 2, pp. 163–180, 2005.
- [17] E. Acar, D. M. Dunlavy, T. G. Kolda, and M. Mørup, “Scalable tensor factorizations for incomplete data,” *Chemometrics and Intelligent Laboratory Systems*, vol. 106, no. 1, pp. 41–56, 2011.
- [18] M. Imaizumi, T. Maehara, and K. Hayashi, “On tensor train rank minimization: Statistical efficiency and scalable algorithm,” in *Advances in Neural Information Processing Systems*, 2017, pp. 3930–3939.
- [19] I. V. Oseledets, “Tensor-train decomposition,” *SIAM Journal on Scientific Computing*, vol. 33, no. 5, pp. 2295–2317, 2011.
- [20] Y. Liu, F. Shang, H. Cheng, J. Cheng, and H. Tong, “Factor matrix trace norm minimization for low-rank tensor completion,” in *Proceedings of the 2014 SIAM International Conference on Data Mining*. SIAM, 2014, pp. 866–874.
- [21] Y. Liu, F. Shang, W. Fan, J. Cheng, and H. Cheng, “Generalized higher-order orthogonal iteration for tensor decomposition and completion,” in *Advances in Neural Information Processing Systems*, 2014, pp. 1763–1771.
- [22] J. Xue, Y. Zhao, S. Huang, W. Liao, J. C.-W. Chan, and S. G. Kong, “Multilayer sparsity-based tensor decomposition for low-rank tensor completion,” *IEEE Transactions on Neural Networks and Learning Systems*, 2021.
- [23] H. Zhang, X.-L. Zhao, T.-X. Jiang, M. K. Ng, and T.-Z. Huang, “Multiscale feature tensor train rank minimization for multidimensional image recovery,” *IEEE Transactions on Cybernetics*, 2021.
- [24] X. Li, Y. Ye, and X. Xu, “Low-rank tensor completion with total variation for visual data inpainting,” in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- [25] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.