Abstract—This paper proposes a novel framework called concatenated image completion via tensor augmentation and completion (ICTAC), which recovers missing entries of color images with high accuracy. Typical images are second- or third-order tensors (2D/3D) depending if they are grayscale or color, hence tensor completion algorithms are ideal for their recovery. The proposed framework performs image completion by concatenating copies of a single image that has missing entries into a third-order tensor, applying a dimensionality augmentation technique to the tensor, utilizing a tensor completion algorithm for recovering its missing entries, and finally extracting the recovered image from the tensor. The solution relies on two key components that have been recently proposed to take advantage of the tensor train (TT) rank: A tensor augmentation tool called ket augmentation (KA) that represents a low-order tensor by a higher-order tensor, and the algorithm tensor completion by parallel matrix factorization via tensor train (TMac-TT), which has been demonstrated to outperform state-of-the-art tensor completion algorithms. Simulation results for color image recovery show the clear advantage of our framework against current state-of-the-art tensor completion algorithms.

Index Terms—Color image recovery, tensor completion, ket augmentation, tensor train rank, image concatenation

I. INTRODUCTION

The problem of recovering missing entries from an array of data is known as completion [1]. Often recovery of the missing data will be based on the available elements of the array and its structural properties, but in the event that there is no pattern between elements, then completion would not be possible. Real-world data tends to exhibit correlations, symmetry, repetition and continuity that make completion possible and reliable. The example in Fig. 1 demonstrates a multidimensional array, with entries representing red, green and blue (RGB) values to form the color image. Traditional methods of matrix factorization [2] and nuclear norm (matrix trace norm) minimization [3] can be used to solve matrix completion problems. There has been particular success in applying these methods to many fields of research such as compressed sensing [4], magnetic resonance imaging (MRI) [5], gene expression prediction [6] and image classification [7].

Algorithms that can complete missing entries of tensors have attracted interest in recent years because of the flexibility and potential applications they offer. Tensors are multidimensional arrays that are higher-order generalizations of matrices and vectors. According to tensor nomenclature [8], a vector and matrix are considered first-order and second-order tensors, where the order represents the number of dimensions (also known as ways or modes). Some examples of tensors include a color image, which is a third-order tensor defined by two indices for spatial variables and one index for color mode. A video comprised of color images is a fourth-order tensor with an additional index for a temporal variable. The utilization of tensors and their decompositions can be seen in seminal works on computer vision [9] and data mining [10], [11]. In [12], the authors demonstrated the superiority of tensor completion algorithms to matrix completion algorithms for completion tasks on hyperspectral images, grayscale/color video and brain MRI images. Their work involves the calculation of multiple low rank matrix factorizations for each mode using an unbalanced matricization scheme (one mode versus the rest),

Figure 1: Facade color image exhibiting many structural properties such as repetition, symmetry and continuity that can be used for completion.
which is based on Tucker rank [13], [14]. Alternatively, [13] utilized Tucker rank for matrix nuclear norm minimization on each mode to solve relaxed optimization problems for tensor completion. Specifically, they introduced three tensor completion algorithms SiLRTC, FaLRTC and HaLRTC, which are presently used as baselines for benchmarking new tensor completion algorithms.

Our previous work [15] had shown that methods based on Tucker rank are not ideal for completion due to the underlying use of an unbalanced matricization scheme. It was shown using von Neumann entropy that tensor completion based on tensor train (TT) rank [16], [17] is more ideal for completion problems due to a balanced matricization scheme, i.e. matricization of a tensor along permutations of modes rather than a single mode, which is the foundation of the TT rank. Historically, TT rank has been predominantly used in physics for simulating quantum many-body systems [18], with only recent real-world applications of TT rank-based methods in machine learning [19]–[21] and tensor completion [15]. The proposed TT rank-based algorithm TMac-TT in [15] outperformed SiLRTC and many state-of-the-art tensor completion algorithms such as FBCP [22] and STDC [23] in both color image and video recovery problems. The advantage of TMac-TT is attributed to the utilization of a novel preprocessing tensor augmentation scheme known as ket augmentation (KA) that creates a structured block addressing of a tensor, which is advantageous only for TT rank-based methods. However, the disadvantage of using KA directly on an image is that block-artsfacts [24] are created due to the TT rank optimization from TMac-TT. This effect is minimized for the results on color video completion because the initial fourth-order tensor to complete is reshaped to a third-order tensor by combining the row and temporal indices. This provides different structural properties than completing each frame of the video individually, and gives potential for new patterns to assist in completing the missing entries.

In this paper we address the problem of block-artsfacts caused by the tensor augmentation of a color image. A novel framework for image completion is introduced that firstly concatenates copies of an image containing missing elements into a third-order tensor, which is inspired by color video completion in [15]. Then, KA is applied on the tensor, followed by the TMac-TT algorithm for tensor completion, and lastly the recovered image is extracted from the tensor. For the remainder of the paper we refer to this framework as concatenated Image Completion via Tensor Augmentation and Completion (ICTAC). Experimental results demonstrate the clear advantage of this framework for image completion problems and we compare the results to KA+TMac-TT (KA combined with TMac-TT) [15] and a recently proposed state-of-the-art algorithm smooth PARAFAC decomposition with quadratic variation (SPC-QV) [25], which has been shown to also outperform FBCP and STDC.

The paper is structured as follows. Section II briefly recalls the existing methods of KA and TMac-TT, which are two components of the proposed ICTAC framework. Section III outlines in detail the proposed ICTAC framework for image completion. In Section IV, color image completion problems are addressed by ICTAC, KA+TMac-TT and SPC-QV. Lastly, Section V concludes the paper.

Most of the notations used in the paper are described here and adopted from [8]. A tensor is a multidimensional array and its order, way or mode is the number of dimensions it contains. Scalars are zero-order tensors denoted by a lowercase letter, e.g. \(x\). Vectors are denoted by a boldface lowercase letter such as \(x\), and matrices are denoted as capital letters, e.g. \(X\). A higher-order tensor (third-order or higher) is denoted by calligraphic letters, e.g. \(\mathcal{X}\). We denote \(\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}\) as a \(N\)th-order tensor with \(I_k (k = 1, \ldots, N)\) representing the size of the \(k\)th mode. The elements of \(\mathcal{X}\) are represented by \(x_{i_1 \ldots i_N}\), where \(1 \leq i_k \leq I_k, k = 1, \ldots, N\). The Frobenius norm of \(\mathcal{X}\) is defined by 

\[||\mathcal{X}||_F = \sqrt{\sum_{i_1} \sum_{i_2} \cdots \sum_{i_N} x_{i_1i_2\ldots i_N}^2}.\]

The mode-\(n\) fiber of a tensor \(\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}\) is denoted by \(x_{i_1 \ldots i_{n-1}i_n \ldots i_N}\), which is a vector defined by fixing all indices but \(i_n\).

Mode-\(n\) matricization (also known as mode-\(n\) unfolding or flattening) of a tensor \(\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}\) is the process of unfolding or reshaping the tensor into a matrix \(X_{(n)} \in \mathbb{R}^{I_n \times (I_1 \cdot I_2 \cdot \cdots \cdot I_{n-1} \cdot I_{n+1} \cdot \cdots \cdot I_N)}\) by rearranging the mode-\(n\) fibers to be the columns of the resulting matrix. Tensor element \((i_1, \ldots, i_{n-1}, i_n, i_{n+1}, \ldots, i_N)\) maps to matrix element \((i_n, j)\) such that

\[j = 1 + \sum_{k=1,k \neq n}^{N} (i_k - 1)J_k \quad \text{with} \quad J_k = \prod_{m=1,m \neq n}^{k-1} I_m. \quad (1)\]

This is an unbalanced matricization scheme because only a single mode represent the matrix row in \(X_{(n)}\).

Mode-(1, 2, \ldots, \(k\)) matricization of a tensor \(\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}\) [17] is defined as \(X_{[k]} \in \mathbb{R}^{m \times n}\) \((m = \prod_{l=1}^{k} I_l, n = \prod_{l=k+1}^{N} I_l)\). The TT rank is defined as the vector \(r = (r_1, r_2, \ldots, r_{N-1})\), where \(r_k\) is the matrix rank of \(X_{[k]}\).

II. KET AUGMENTATION AND TMAC-TT

In this section we review ket augmentation (KA) and TMac-TT, which are two components of the ICTAC framework that will be introduced in the next section.

KA is a tensor augmentation technique that essentially represents a third-order color image tensor \(\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times I_3}\) \((I_3 = 3)\) into a \(K\)th-order tensor \(\tilde{\mathcal{T}} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_K}\) where \(K \geq 3, \prod_{k=1}^{K} J_k = \prod_{k=3}^{3} J_k\), and \(J_k\) represents a unique block structured addressing of the original image. The general form of \(\tilde{\mathcal{T}}\) is given by

\[\tilde{\mathcal{T}}_{[2^n \times 2^n \times 3]} = \sum_{i_n \ldots i_1=1}^{I_3} \sum_{j=1}^{3} c_{i_n \ldots i_1} e_{i_n} \otimes \cdots \otimes e_{i_1} \otimes u_j. \quad (2)\]

Fig. 2 demonstrates the result of KA for color two images. It can be seen that each mode represents different layers of the image, hence by utilizing a TT rank-based algorithm, the correlations of textures can be studied between the modes of
the tensor. For this paper we modify the KA algorithm such that

$$\tilde{\mathcal{T}}_{[3^2 \times 2^2 \times 3]} = \sum_{i_1,\ldots,i_3=1}^{6} \sum_{i_1',\ldots,i_3'=1}^{3} c_{i_1,\ldots,i_3} e_{i_1'} \otimes \cdots \otimes e_{i_3'} \otimes u_j, \quad (3)$$

where each mode is $i_n = 1, \ldots, 6$, rather than $i_n = 1, \ldots, 4$ in (2). This modification is used for image concatenation, which produces a tensor with one mode significantly larger than the others, e.g. $I_1 \gg I_2$ for $\mathcal{T}$, hence the modified KA in (3) caters for rectangular matrices in the subspace $I_1 \times I_2$.

TMac-TT is a tensor completion algorithm that completes a tensor $X \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ containing only a subset of known elements $\Omega$. Specifically, it addresses the following multilinear matrix factorization problem:

$$\min_{U_k,V_k,X[k]} \|U_k V_k - X[k]\|^2_F, \quad (4)$$

for $k = 1, 2, \ldots, N - 1$. This problem is convex when each variable $U_k, V_k$ and $X[k]$ is optimized while keeping the other two fixed. Specifically, the following steps are performed to update each variable:

$$U_k^{l+1} = X[k] (V_k^l)^T, \quad (5)$$

$$V_k^{l+1} = ((U_k^l)^T U_k^{l+1})^T (U_k^{l+1} X[k]), \quad (6)$$

$$X[k]^{l+1} = U_k^{l+1} V_k^{l+1}, \quad (7)$$

where $\cdot^\dagger$ denotes the Moore-Penrose pseudoinverse.

It was shown in our previous paper [15] that KA+TMac-TT outperformed baseline algorithms SiLRTC and TMac [12], state-of-the-art algorithms FBCP and STDC, and another TT rank-based algorithm ALS [26], [27] in color image and video recovery tasks. The next section will introduce a new framework for improving image recovery tasks to minimize block-artifacts caused by the KA on a single image. These artifacts can be seen in the result for completion of the Lena image with 90% missing entries in Fig. 3.

III. A CONCATENATED IMAGE COMPLETION FRAMEWORK

The concatenated image completion via tensor augmentation and completion (ICTAC) framework is outlined in this section. The framework is divided into three main steps:

1) **Image concatenation**: The concatenation of a single image with missing entries into a third-order tensor to prepare it for KA, as discussed in Subsection III-A.

2) **KA+TMac-TT**: The application of KA then TMac-TT on the concatenated third-order tensor to recover missing entries, as discussed in Subsection III-B.

3) **Image extraction**: Extracting a single recovered image from the recovered concatenated tensor, which is discussed in Subsection III-C.

A. Concatenating images for tensor augmentation

Consider an $N$th-order color image tensor $X \in \mathbb{R}^{I_1 \times I_2 \times 3}$ that consists of partially known entries given by a subset $\Omega$. Applying directly the KA scheme to $X$, then subsequently the TMac-TT algorithm for completion, will result in blocking-artifacts as demonstrated in Fig. 3. To circumvent this problem, an initial preprocessing step prior to KA is added that concatenates identical copies of $X$ to form a fourth-order tensor $X_{ci} \in \mathbb{R}^{I_1 \times I_2 \times 3 \times C}$, where $C > 1$ is the number of copies of the tensor $X$. In fact, [15] had naturally formed a tensor similar to $X_{ci}$ for color video recovery, however, rather than have $C$ for the fourth mode, the label $T$ is used to represent the time frames of a color video. Therefore $X_{ci}$ can be considered a motionless color video with $C$ frames.

The next step is to permute and reshape $X_{ci}$ to a third-order tensor $X_{ref} \in \mathbb{R}^{I_1 \times I_2 \times 3}$, where $I_1 = CI_1$ is the combined row mode. Displaying $X_{ref}$ would result in a distorted continuous stream of the original image, hence, its structural properties have completely changed. The motivation to form this type of tensor is that the repetition of the image allows for more potential correlations, symmetry and/or continuity than
a single image would exhibit. Additionally, applying KA on $X_{\text{vst}}$ would result with obvious block-artifacts only on $X_{\text{vst}}$, however, when the recovered image is extracted in the final step of ICTAC, these artifacts are minimized substantially. Fig. 4 demonstrates $X_{\text{vst}}$ for the original Lena image with no missing entries. We can see that there are considerably more patterns in the image compared to the original image itself. Therefore by converting an image to essentially a motionless video allows for new insights and potential for image completion tasks. The formation of a tensor of this type was originally defined as a video sequence tensor (VST) in [15] for preprocessing a color video prior to KA+TMac-TT.

B. KA and TMac-TT

For the next step we apply the modified KA scheme in (3) to augment the third-order tensor $X_{\text{vst}}$ of size $I_1 \times I_2 \times 3$ to a higher-order tensor $\hat{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_K}$, with $K \geq N$. The modified KA scheme is needed due to subspace $I_1 \times I_2$ of $X_{\text{vst}}$ being a large non-square matrix, and the original KA scheme proposed in [15] only catered for tensors of the form $I_1 \times I_2 \times 3$, with $I_1 = I_2$, hence the subspace $I_1 \times I_2$ is a square matrix.

After a block structured addressing via KA has been applied on $X_{\text{vst}}$, then it is ready to be transferred to a tensor completion algorithm. The TMac-TT algorithm can now be utilized to recover the missing entries of $\hat{X}$. For more detail on these steps, the reader can refer to [15].

C. Recovered image extraction

After $\hat{X}$ has been recovered via KA+TMac-TT, to recover a single image, an inverse KA scheme is utilized to obtain a third-order recovered concatenated tensor $\hat{X}_{\text{vst}} \in \mathbb{R}^{I_1 \times I_2 \times 3}$. $\hat{X}_{\text{vst}}$ is subsequently reshaped and permuted back to a fourth-order tensor $\hat{X}_{ci} \in \mathbb{R}^{I_1 \times I_2 \times 3 \times C}$, and from this tensor we can extract a single recovered image, e.g. the image at $\hat{X}_{ci}(; ; ; , 1)$.

IV. EXPERIMENTAL RESULTS

The experiments are conducted for image completion tasks of various missing ratios for the Lena and Peppers color images. The proposed ICTAC framework is benchmarked against current state-of-the-art tensor completion algorithms KA+TMac-TT [15] and SPC-QV [25].

The missing ratio (mr) as a percentage of a test image is defined as

$$mr = \frac{p}{\prod_{l=1}^{N} I_l} \times 100\%,$$

with $p$ being the number of missing entries, which is chosen randomly based on a uniform distribution.

Performance measures include the relative square error (RSE) between an approximately recovered tensor and the original one, which is defined as,

$$RSE = \frac{||X - T||_F}{||T||_F},$$

and the structural similarity index (SSIM) [28].

The algorithms ICTAC and SPC-QV recover color images represented by third-order tensors of size $243 \times 512 \times 3$ to cater for the modified KA scheme of ICTAC as discussed in the previous sections, whereas KA+TMac-TT uses images of size $256 \times 256 \times 3$ because the traditional KA scheme works only on the condition $I_1 = I_2$. To compare the algorithms in a fair manner, the $RSE$ and $SSIM$ is calculated based on the same image sizes for the initial tensor with missing entries and the final recovered tensor, and there is no image distortion which can change the $mr$ throughout the runtime of the completion algorithms. The only distortion that may happen is after the algorithms have completed their calculations, where the final recovered tensors of size $243 \times 512 \times 3$ from ICTAC and SPC-QV is reshaped to $256 \times 256 \times 3$ so that a visual comparison can be made to the results of KA+TMac-TT. The number of copied images $C = 81$ for all benchmarks, and simulations are conducted using a Matlab environment.

The ICTAC framework contains several tensor size transformations. For clarity, an outline of these changes is given for an initial tensor $X \in \mathbb{R}^{243 \times 512 \times 3}$ to the higher-order tensor $\hat{X}$ in Subsection III-B:

1) Image concatenation: $X \in \mathbb{R}^{243 \times 512 \times 3} \rightarrow X_{ci} \in \mathbb{R}^{243 \times 512 \times 3 \times 81}$.
2) Obtaining a VST: $X_{ci} \in \mathbb{R}^{243 \times 512 \times 3 \times 81} \rightarrow X_{\text{vst}} \in \mathbb{R}^{19683 \times 512 \times 3}$.
3) Applying the modified KA: $X_{\text{vst}} \in \mathbb{R}^{19683 \times 512 \times 3} \rightarrow X \in \mathbb{R}^{6 \times 6 \times 6 \times 6 \times 6 \times 6 \times 6 \times 6 \times 6 \times 3}$, i.e. $n = 9$ in (3).

Fig. 5 presents the results for the Lena completed images using ICTAC, KA+TMac-TT and SPC-QV. In the case of 90% missing entries, the KA+TMac-TT algorithm performs the worst in terms of $RSE$ and $SSIM$ compared to the other two algorithms. This is a significant result because it was shown recently in [15] that KA+TMac-TT outperformed state-of-the-art algorithms in color image recovery. Comparing
ICTAC and SPC-QV, it is interesting to see that although both algorithms had similar performance in \( RSE \) and \( SSIM \), with ICTAC slightly better in both, there are quite striking visual differences in their image recovery results. Specifically, SPC-QV tends to have a smoother uniform textures of the Lena image, especially around edges, however, a slight blur and fading can be seen that affects detail such as the hair strands when comparing it to the original image. The recovered ICTAC image demonstrates an attempt to detail the finer textures of the original image, which can be clearly seen on the hair strands, however, some block-artifacts and errors from the image recovery can be seen. For 80% missing entries, ICTAC provides a recovered image almost completely similar to the original image with an SSIM of 0.983 and RSE of 0.048. SPC-QV has a slight image blur and still does not provide detail on more complicated parts of the image such as the hair and lips. The KA+TMac-TT results still has the worst performance with an RSE of 0.08, and block-artifacts can still be easily observed.

### V. Conclusion

In this paper, we proposed a novel framework known as concatenated image completion via tensor augmentation and completion (ICTAC). The framework formulates a tensor from a concatenation of identical copies of a single color image with missing entries, which provides additional patterns to support image completion algorithms. It then utilizes tensor augmentation based on modified KA, a TT rank-based tensor completion algorithm TMac-TT to impute the missing entries, and finally, an image extraction method to recover the completed image. Our method was shown to outperform recently proposed state-of-the-art tensor completion algorithms KA+TMac-TT and SPC-QV for color image completion tasks. ICTAC is still at its infancy and there are many future endeavours to further improve the efficiency of the algorithm. A major direction is to investigate the concatenation of images in different tensor structures and copies \( C \), as there may be a minimum or maximum \( C \) that could determine how good a result we can obtain for tensor completion. The exploration
of different KA schemes would also benefit the framework as well as other TT rank-based tensor completion algorithms.

ACKNOWLEDGEMENT

The authors would like to thank Tatsuya Yokoto and Andrzej Cichocki for generously providing their Matlab code of the SPC-QV algorithm.

REFERENCES


