

Error Concealment with Multiscale Patch Clustering and Low-Rank Minimization

Mading Li, Wenhan Yang, Sijie Song, and Zongming Guo*

Institute of Computer Science and Technology, Peking University, Beijing, P. R. China, 100871

Abstract—In this paper, we propose a novel error concealment method based on multiscale patch clustering and low-rank minimization. In order to collect more reliable patches to form a genuine low-rank matrix, an image pyramid is formed utilizing an effective down-sampling process. The classic singular value thresholding (SVT) is modified into a global iteration to solve the low-rank minimization problem. Extensive experimental results on the random pixel loss and the block loss situation validate the effectiveness of the proposed method. The proposed method acquires higher PSNR and better visual quality than the state-of-the-art low-rank based error concealment methods.

I. INTRODUCTION

There are now tons of images and videos being transmitted through all kinds of network everyday. When going through the network, the encoded image/video bit stream is vulnerable to the channel impairment caused by network congestion, signal fading, *etc.* Since the signals are often highly compressed during the transmission, the loss of few bits may cause massive loss of the signal, which severely affects the visual quality of the image/video on the decoder side. The error concealment (EC) technique utilizes the information yielded by the successfully received areas of the image/video to generate a plausible recovery result for the watcher to alleviate the negative effect of the information loss.

In recent years, the EC method has been widely studied. All EC methods exploit the self-similarity or the spatial/temporal correlation within the image/video [1]. In this paper, we focus on the spatial correlation to present the EC method for images. One of the classic methods for EC is utilizing the Bayesian framework [2][3]. These methods attempt to maximize the conditional probability of each individual missing pixel given the available pixels and other recovered pixels. Besag [2] employs the Bayes' rule on the posterior probability to yield an optimization with the prior probability and likelihood, while actually the prior probability model is often not accurate enough or even not available. Under the assumption that images can be locally modeled as a stationary Gaussian process, Li and Orchard [3] estimated the covariance of a local window to characterize the local interpolation coefficients and presented an orientation adaptive interpolation scheme. Nevertheless, the assumption does not always hold for images.

Recently, the low-rank matrix/tensor minimization methods have been drawing attention because of its efficiency in recov-

ering data from few entries. Many matrix/tensor completion methods have been proposed [4][5] to reconstruct images with missing pixels. Under high missing rate, these methods can produce excellent results on images with highly repetitive structures. However, their performances on other images are not so good. In other words, these methods only work well if the whole image is exactly low-rank. To tackle this problem, Chen *et al.* [6] introduced a concept to complete the missing entries and simultaneously capture the underlying model structure, and yet the complexity of such methods is rather high. Liu and Shang [7] introduced a matrix factorization idea into the tensor nuclear norm model in order to achieve a much smaller scale matrix nuclear norm minimization problem. Different from "global low-rankness", Ono *et al.* [8] presented a recovery method by promoting "blockwise low-rankness". They stated that a small block extracted from an image was expected to be low-rank. Nguyen *et al.* [9] proposed the combination of nonlocal grouping of image patches and low-rank tensor approximation. Dong *et al.* [10] presented a low-rank approach toward modeling nonlocal similarity in natural images and discussed its connection with simultaneous sparse coding. Nevertheless, the similar patches these methods collected may not be similar enough to each other, leading to an unsuccessful recovery result.

This paper presents a novel EC method based on multiscale patch clustering and low-rank minimization. The proposed method first builds and initializes an image pyramid by an efficient down-sampling process. After that, similar patches are clustered from the pyramid to form a low-rank matrix. The matrix is then processed by a modified singular value thresholding algorithm to refine the initialized values of the missing pixels. Experimental results demonstrate the effectiveness of the proposed method. The rest of the paper is organized as follows: Section II first formulates the patch clustering under the multiscale modeling. Next, the framework of the algorithm and two applications for the proposed method are presented in the following subsection in details. Experimental results are presented in Section III, followed by the conclusion in Section IV.

II. THE PROPOSED ERROR CONCEALMENT METHOD

A. Multiscale Patch Clustering

One of the most significant issues of low-rank minimization methods is that, the rank of the matrix to be processed should be low enough, otherwise the result would not be satisfactory. Existing methods always focus on manipulating the matrix

*Corresponding author

This work was supported by National High-tech Technology R&D Program (863 Program) of China under Grant 2014AA015205 and Beijing Natural Science Foundation under contract No.4142021.

with various means. In this section, instead of leveraging sophisticated approaches, we are more concerned with the matrix that we are dealing with.

Generally, the rank of a matrix is rather low if most of the entries can be represented by few base entries. In other words, if the entries are similar to each other, the matrix is low-rank and can be well processed by low-rank minimization methods. In image/video denoising area, similar image patches can be used as matrix entries if represented as vectors. Existing methods search for similar patches in a local area (or in the whole image, which might acquire more similar patches but could be really time consuming), using the sum of square differences (SSD) as the distance criterion. Nevertheless, there are not always sufficient similar patches in the local area. As a result, the rank of the matrix formed by these methods are not low enough to be well processed. Giving an adaptive threshold to control the maximum distance is an effective way to limit the similarity between patches, yet it does not guarantee the adequate number of similar patches. A small number of similar patches cannot provide enough information, while too many similar patches affect the processing speed of low-rank minimization methods. Based on these observations, our goal is to collect enough similar patches that are similar enough.

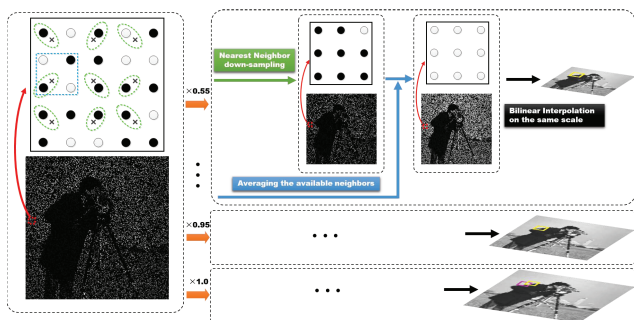


Fig. 1. All levels are down-sampled from the input image and initialized by the method presented in Section II-C. Then, each level is recovered by the proposed method while clustering patches within itself. In the end, the bottom level is recovered utilizing the patches from the whole pyramid. The pink patch in the bottom (original) level represents a reference patch. Yellow patches in different levels represent the collected similar patches.

To this end, we propose to create an image pyramid by down-sampling the input image in different scales and then collect similar patches from the image pyramid (Fig. 1). For a reference patch p , assuming that L similar patches (s_1, s_2, \dots, s_L) have been found in the original input image. Instead of finding similar patches of p in the pyramid, we look for K similar patches for every single $s_k, k = 1, 2, \dots, L$. Finally, the low-rank matrix is formed by all of these $(1 + L + L \cdot K)$ similar patches. Such patch clustering method enhances the result of all the similar patches s_k in addition to the reference patch p . Note that we also apply the SSD in experiments to compute the difference between patches.

B. The Framework of the Algorithm

In this section, the proposed error concealment method is interpreted. The image block containing at least one missing pixel is hereinafter referred to the *reference block*. The missing pixels in the input image are initialized at first, details will be discussed in Section II-C. After that, for each reference block, patches that similar to it are collected. If we represent each similar patch as a vector by concatenating all its columns, the whole set of similar patches can be grouped into a matrix M . Since all the entries of M are similar to each other, the rank of M should be low. Therefore, the missing pixels can be recovered by solving the low-rank minimization problem:

$$\begin{aligned} \min_X \quad & \text{rank}(X), \\ \text{s.t.} \quad & X_{ij} = M_{ij}, (i, j) \in \Omega, \end{aligned} \tag{1}$$

where X is the recovered matrix and Ω is the set of locations corresponding to available pixels.

Unfortunately, the low-rank minimization problem in (1) is an NP-hard problem and cannot be solved efficiently so far. In recent years, the nuclear norm has been shown to be the tightest convex approximation for the rank of matrices. Thus, the nuclear norm has been applied to solve the low-rank minimization problem in this paper and (1) can be represented as:

$$\begin{aligned} \min_X \quad & \sum_i \sigma_i(X), \\ \text{s.t.} \quad & X_{ij} = M_{ij}, (i, j) \in \Omega, \end{aligned} \tag{2}$$

where $\sigma_i(X)$ is the i th largest singular value of X . Plenty of low-rank minimization algorithms have been proposed to solve the problem in (2). In this paper, the singular value thresholding (SVT) [11] is modified and applied for its simplicity and ease of implementation.

The input matrix M is decomposed via SVD. The singular values smaller than the threshold τ are omitted. Let $M = U\Sigma V^T$ be the SVD for M , then the soft shrinkage operation can be defined as:

$$S_\tau(M) = U\Sigma_\tau V^T, \tag{3}$$

where $\Sigma_\tau = \text{diag}(\max(\sigma(M) - \tau, 0))$. $\sigma(M)$ represents the singular values of M .

Under the assumption of Laplacian prior, the thumb rule for choosing the suitable threshold τ is given in [12], which is $\tau = 2\sqrt{2}\sigma_m^2/(\sigma_M + \varepsilon)$. In our experiment, σ_m is set to be the missing rate of the input image, ε is a tiny number that prevents the division by zero and σ_M denotes the local estimated variance given as follows:

$$\sigma_M = \sqrt{\max(\sigma^2(M)/n - \sigma_m^2, 0)}, \tag{4}$$

where n is the number of similar patches that form M .

After the soft shrinkage operation, the columns of processed matrix $S_\tau(M)$ are reformed and added up back into their original positions. ‘‘Patch-voting’’ is performed to accumulate the pixel values of each overlapping neighbor patch. Until all the reference patches have been processed, all the ‘‘votes’’ of

the missing pixels are averaged to generate a new image, for the next iteration of the global SVT.

C. Different Applications

In this section, we apply the proposed method on two different situations: the random pixel loss situation (the first row of Fig. 3) and the block loss situation (the first column of Fig. 5).

The random pixel loss situation: In this situation, missing pixels are uniformly distributed in the whole image. Therefore the missing pixel can be preliminarily estimated by its available neighbors, which gives an initialization to the patch clustering step. In the upcoming iterations, all of the images in the pyramid will be updated. To make the full use of the available pixels in higher image levels, an effective down-sampling method is applied (Fig. 1):

- 1) To down-sample the original input image to a lower level, first of all, we use the Nearest Neighbor down-sampling to generate a low-resolution image.
- 2) Then, missing pixels in the down-sampled image are estimated by averaging available pixels that enclose the corresponding positions in the higher level.
- 3) At last, if there are still some missing pixels, which means that there are no available pixels around its corresponding position in the higher level, these pixels are interpolated by Bilinear interpolation.

As shown in Fig. 2, such down-sampling method generates low-resolution image of high quality, providing clean patches for the patch clustering step.

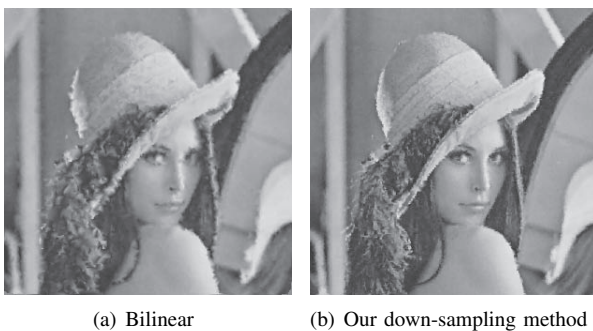


Fig. 2. Comparison of $0.6\times$ down-sampled images using Bilinear and our method. The missing rate of the input image is 85%.

The block loss situation: The experimental configuration for this situation is mostly same with the former one. However, the down-sampling method proposed above cannot be utilized since the missing area is continuous and much larger than the former situation. In our experiment, we use Bilinear interpolation to initialize the missing blocks and use more iterations to generate a plausible recovery. Other differences are listed as follows:

- 1) The patch size in this situation is larger than that in the random pixel loss situation. Larger patches contain more structural information so that the low-rank minimization method can produce more precise results.

- 2) Inspired by [10], we also utilize the deterministic annealing to further improve the performance. The SVT starts with a rather large threshold and then gradually decreases the threshold.

Note that the image pyramid discussed before is not much useful for the block loss situation. The reason is that most of the similar patches found in the original level do not contain missing pixels, which means there are so few missing elements in the matrix M that it does not need extra entries collected from the image pyramid.

III. EXPERIMENTAL RESULTS

In this section, we evaluate the performance of the proposed method for EC tasks on both random pixel loss and block loss situations. All the experiments are performed with MATLAB platform. The state-of-the-art low-rank based methods LRTC [4], STDC [6] and SAIST [10] are used as comparisons. Peak Signal-to-Noise Ratio (PSNR) is selected as the objective evaluation criterion. All PSNR results are computed only on the missing locations.

In our experiments, the side length of the reference block is 6 pixels in the random pixel loss situation and 16 pixels in the block loss situation, extracted every 4 pixels from the input image. The missing rates of the random pixel loss situation are set to be 60%, 70%, 80%, 85%, 90% and 95%. The size of the missing block in the latter situation is 16×16 . The image pyramid is formed by down-sampled images with different scales. The scales varies from 0.95 to 0.55, added up to 10 levels, including the original input image. The number of the similar patches found in the original level $L = 45$ and there are $K = 5$ patches collected from the pyramid for each similar patch. In SAIST and the proposed algorithm, the stopping criterion is the maximal number of iterations 30 being reached. Test images are selected from USC-SIPI image database.

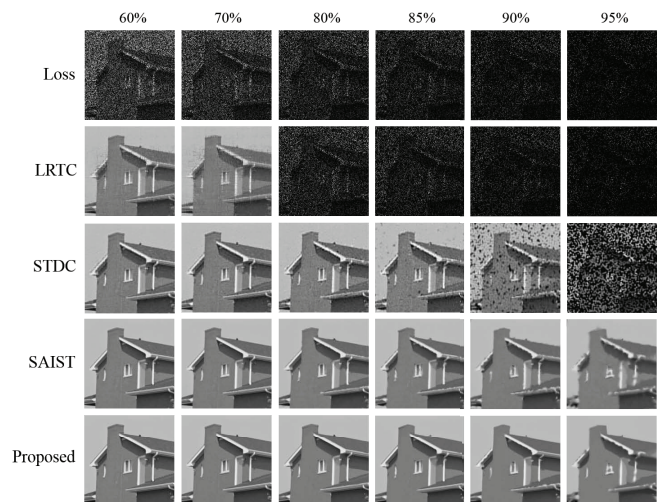


Fig. 3. Comparison of recovery results of LRTC, STDC, SAIST, and the proposed method under different missing rates in the random pixel loss situation.

As illustrated in Fig. 3, LRTC produces burrs on the whole image and it cannot successfully recover the input image under

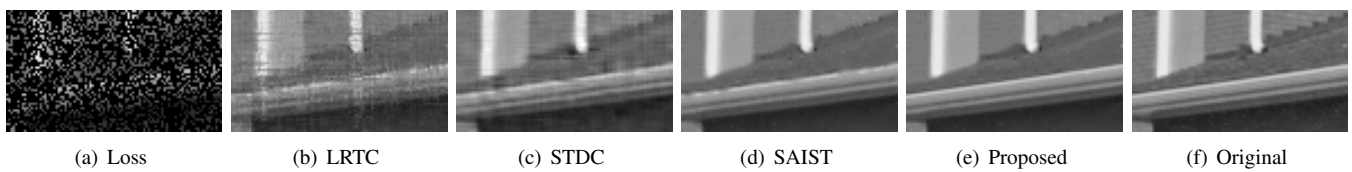


Fig. 4. Magnification of local area of the input image (70% loss), recovery results of LRTC, STDC, SAIST, the proposed method and the original image.

TABLE I
AVERAGE PSNR (DB) RESULTS OF DIFFERENT METHODS UNDER DIFFERENT MISSING RATES.

Missing rates	LRTC	STDC	SAIST	Proposed
60%	23.43	25.22	28.11	29.41
70%	22.22	24.83	26.87	28.05
80%	10.11	23.92	25.45	26.45
85%	5.45	22.95	24.50	25.31
90%	5.45	18.02	23.44	23.96
95%	5.45	8.83	22.01	22.22

high missing rates. STDC presents better recovery results since it tends to capture the underlying model structure, yet there are blurring and ringing artifacts that can be clearly observed. The recovery results of STDC under high missing rates still contain lots of noises. SAIST has rather good performance overall, however, it produces coarse edges (Fig. 4). The proposed method presents the best visual quality. The objective criterion shown in Table I indicates the same result. The recovery results of the block loss situation are shown in Fig. 5. As can be observed, the proposed method also outperforms other competing methods.

IV. CONCLUSION

This paper presents a novel error concealment method based on multiscale patch clustering and low-rank minimization. An image pyramid is formed by an efficient down-sampling process to provide more reliable patches. The classic SVT is modified to apply on the low-rank matrix constructed by similar patches. Extensive experimental results demonstrate the effectiveness of the proposed method on different situations. The proposed method acquires higher PSNR and better visual quality than the state-of-the-art low-rank based error concealment methods.

REFERENCES

[1] G. Zhai, X. Yang, W. Lin, and W. Zhang, "Bayesian error concealment with DCT pyramid for images," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 9, pp. 1224-1232, Sep. 2010.

[2] J. Besag, "On the statistical analysis of dirty pictures," *J. Roy. Statist. Soc. B*, vol. 48, no. 3, pp. 259-302, 1986.

[3] X. Li and M.T. Orchard, "Novel sequential error-concealment techniques using orientation adaptive interpolation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 10, pp. 857-864, Oct. 2002.

[4] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208-220, Jan. 2013.

[5] Y. Hu, D. Zhang, J. Ye, X. Li, and X. He, "Fast and accurate matrix completion via truncated nuclear norm regularization," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2117-2130, Sep. 2013.

[6] Y. Chen, C. Hsu, and H.-Y.M. Liao, "Simultaneous tensor decomposition and completion using factor priors," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 36, no. 3, pp. 577-591, Mar. 2014.

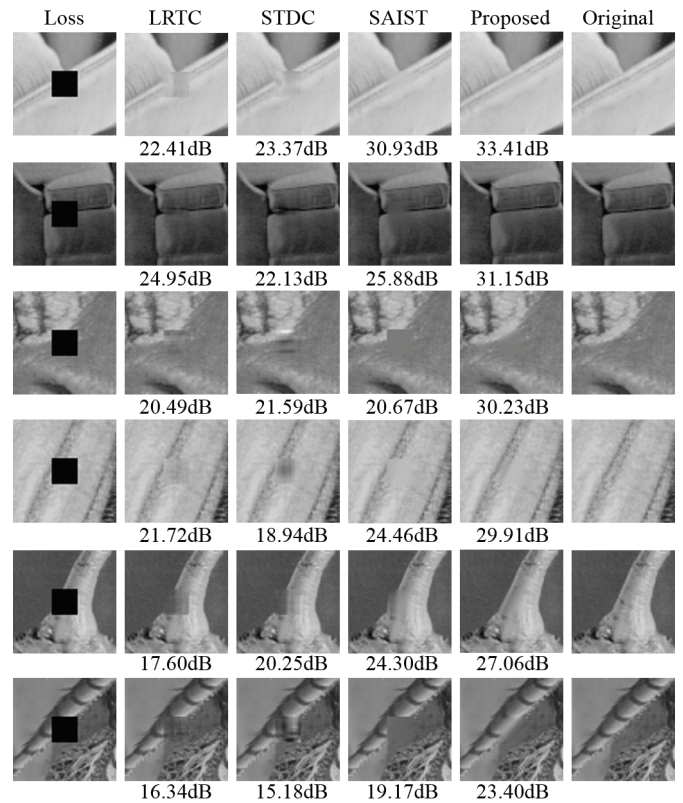


Fig. 5. Portions of the block loss images, the recovery results of LRTC, STDC, SAIST, the proposed method and the original image (from left to right). From top to bottom: Lena, Couple, Baboon, Baboon, Pepper, Man.

[7] Y. Liu and F. Shang, "An efficient matrix factorization method for tensor completion," *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 307-310, Apr. 2013.

[8] S. Ono, T. Miyata, I. Yamada, and K. Yamaoka, "Missing region recovery by promoting blockwise low-rankness," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 1281-1284, Kyoto, Japan, Mar. 2012.

[9] D.T. Nguyen, M.D. Dao, and T.D. Tran, "Error concealment via 3-mode tensor approximation," in *Proc. IEEE Int. Conf. Image Process.*, pp. 2081-2084, Brussels, Belgium, Sep. 2011.

[10] W. Dong, G. Shi, and X. Li "Nonlocal image restoration with bilateral variance estimation: a low-rank approach," in *IEEE Trans. Image Process.*, vol. 20, no. 2, pp. 700-711, Feb. 2013.

[11] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956-1982, Mar. 2010.

[12] S.G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532-1546, Sep. 2000.

[13] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using Low rank matrix completion," in *Proc. IEEE Conf. Comput. Vision Pattern Recogn.*, pp. 1791-1798, San Francisco, USA, Jun. 2010.