# High Quality Image Construction from Multiple Low Quality Copies

Lin Ma $^{\dagger \ddagger 1},$  Long Xu $^{\# 2},$  Qian Zhang $^{* 3},$  and King N. Ngan $^{\dagger 4}$ 

<sup>†</sup>Department of Electronic Engineering, the Chinese University of Hong Kong, Hong Kong <sup>‡</sup>Lenovo Corporate Research Hong Kong Branch, Hong Kong

<sup>#</sup>Department of Automatic Engineering, University of Science and Technology Beijing, Beijing, China, 100083

\*School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an, China, 710055

<sup>1</sup>lma@ee.cuhk.edu.hk <sup>2</sup>lxu@jdl.ac.cn <sup>3</sup>qzhangmis@gmail.com <sup>4</sup>knnqan@ee.cuhk.edu.hk

Abstract—In this paper, the authors proposed to construct a high quality image based on multiple low quality input images. The relationship between one pixel and its neighbourhood should be consistent between different degraded images. Therefore, the reconstruction method is proposed by enforcing the pixel consistency property, which is ensured by estimating the parameters of piecewise image model for each pixel. Subsequently, the reconstructed coefficients are regularized within a reasonable range. Experimental results on multiple images (with different distortions of different levels) have demonstrated that the proposed method can effectively alleviate the noises meanwhile preserve the detailed information. Better quality images in terms of both objective and subjective measurements can be generated.

### I. INTRODUCTION

In order to efficiently store and transmit images/videos, compression schemes are employed. In most image and video coding standards, block-based discrete cosine transform (BDCT) is utilized, reducing the inter-pixel statistical redundancy. However, in order to ensure higher compression ratios, BDCT together with the coarse quantization will inevitably introduce artifacts into the compressed images/videos, such as the blocking artifacts in the smooth regions, and the ringing artifacts around the edge boundaries.

In order to remove these artifacts, many research works have been done. In general, the works can be classified into two categories [3]: 1) enhancement-based approach, and 2) restoration-based approach. The former approach tries to alleviate the artifacts introduced by compression as much as possible. The artifacts can be suppressed in transform domain (e.g. DCT [11], overcomplete wavelet representation (OWR) [12]), and spatial domain [13] [14]. In particular, Zeng [11] models the blocking artifact as 2-D step edge and suppresses it by applying the zero-masking scheme. And blocking artifacts are removed in wavelet domain by exploiting cross-scale correlations among wavelet coefficients and protecting the edge information in [12]. In [13] [14], postfiltering in shifted widows (PSW) of image blocks is proposed, which suppresses blocking artifacts by averaging coefficients of neighbouring image blocks in the shifted windows. On the other hand, image compression can be regarded as a distortion process. Therefore,

image reconstruction is viewed as an inverse problem, where iterative algorithms for restoring the original images are proposed. The projection onto convex sets (POCS) algorithms [15] [16] represent the prior information of the original as convex sets, and they converge in the intersection of all the sets through iterating projections. The most commonly used convex sets are quantization constraint sets (QCS) and smoothness constraint sets (SCS). Recently, learning-based image restoration schemes [17]-[20] have been proposed to reconstruct a high-quality image by introducing the learned HF information from pre-designed codebooks into the degraded low quality image. A blind image restoration method [17] is proposed by estimating the HF information of a given blurred image from its low-frequency (LF) information based on the designed multiple codebooks. Liaw et al. [20] propose to restore the image based on the classified vector quantization (CVQ), which employs a codebook to transform the compressed image into a set of indices, and decodes the indices to enhance the compressed image based upon another corresponding codebook. In [18], the HF components are generated by the Markov-chain inference strategy based on the constructed codebook, which is learnt from a training set of natural images.

Nowadays, with the development of Internet, more and more images/videos are shared between different users, such as Flickr and Youtube. Therefore, there may be many copies of different qualities generated from the same single source image appeared on the Internet. How to utilize these different distorted images/videos to generate a high quality image/video are now researched. In [1]-[3], each coefficient of the video in the transform domain is reconstructed by using a narrow quantization constraint set, which is defined by the multiple copies. In this paper, we also deal with the problem of high quality image reconstruction from multiple low quality copies. The reconstruction method is proposed by enforcing the pixel consistency property, which is ensured by estimating the parameters of piecewise image model (PIM) for each pixel followed by a regularization process.

The rest of the paper is organized as follows. Section II will formulate the image reconstruction problem. In Section III,

the proposed image reconstruction method will be introduced. Section IV will present the experimental results. Finally, Section V will conclude the paper.

#### II. PROBLEM FORMULATION

As image is obtained via sensing, storage, and transmission, noise will be inevitably introduced into the image. In [10], Geman *et al.* proposed a simple yet effective image degradation model, which can be viewed as a simple signal attenuation and additive Gaussian noise model:

$$D = G \otimes R + N \tag{1}$$

where  $\otimes$  is the convolution process, *R* is the source image free of artifacts, *D* is the observed distorted image, *N* is the additive Gaussian noise, and *G* is a signal attenuation convolution kernel. Image restoration can be viewed as an inverse problem, which tries to recover an image  $R^*$ , which should be as truthful as possible with the source image:

$$R^* = \underset{R}{\operatorname{argmin}} \{ E(R; D) \}$$
(2)

where *E* measures the reconstruction error between the original and distorted image. If we have multiple distorted images, the restoration problem can be further formulated as:

$$R^* = \underset{P}{\operatorname{argmin}} \{ E(R; D_1, D_2, \dots, D_L) \}$$
(3)

As the distorted images are generated from the same image, the local consistency of each distorted image should be the same as each other. Therefore, Eq. (3) can be further expressed as:  $R^* =$ 

$$\underset{R}{\operatorname{argmin}} \left\{ E_r(R; D_1, D_2, \dots, D_L) + \lambda \left\{ \sum_{q, p \in \{1, \dots, L\}} E_c\{D_p, D_q\} \right\} \right\}$$
(4)

where  $E_r$  denotes the reconstruction error, while  $E_c$  depicts the consistency errors between different distorted images.  $\lambda$ balances the contributions of the two different errors. In this paper, the authors will focus on the consistency error component to reconstruct the high quality image, while the reconstruction error will be minimized by performing a simple regularization method.

## III. PROPOSED IMAGE RECONSTRUCTION METHOD FROM MUL-TIPLE LOW QUALITY COPIES

In order to ensure the consistency between different distorted images, we need first to revisit the piecewise image model (PIM). PIM tries to model the relationship between current pixel and its neighbouring pixels of the image. Detailed information will be introduced in the following.

# A. Piecewise Image Model (PIM)

In order to reconstruct a high quality image based on the low quality copies, a 2D piecewise image model (PIM) for modelling the image is employed:

$$I(i,j) = \sum_{(m,n)\in T} \alpha(m,n)I(i+m,j+n) + \vartheta(i,j)$$
<sup>(5)</sup>

where *T* is a spatial region for considering the PIM model.  $\vartheta(i,j)$  is a random perturbation independent of spatial location (i,j) and image signal.  $\vartheta(i,j)$  can depict both the fractal-



Fig. 1. PIM parameter estimation

like fine details of image signals and measurement noise. From Eq. (5), it can be observed that the PIM hinges on the adjustment of the model parameters  $\alpha(m, n)$  to the local pixel structures. In other words, the PIM model will adaptively model the relationship between current pixel and surrounded pixels in different positions. In [4], it is claimed that the semantically meaningful image constructs, such as edges and surface textures, are formed by spatially coherent contiguous pixels, suggest piecewise statistical stationarity of the image signal. Therefore, the image pixels in a small local region will share similar parameters  $\alpha(m, n)$ , although  $\alpha(m, n)$  may vary significantly in different segments of a scene, which is named as the piecewise auto regressive (PAR) model. In this manner, the local pixel structures such as edges and textures can be learnt by fitting samples of a local window to the PAR model.

The validity of the PIM and PAR model with locally adaptive parameters is corroborated by the success of this modelling technique in image interpolation [4] and lossless image compression, such as CALIC [5], TMW [6], and invertible integer wavelets [7]. Inspired by the successes of PAR in these research areas, image reconstruction method can be further performed by estimating the PIM parameters of each pixel of different degraded images. The truth is that different degraded image should share the same consistency in the colocated local regions, which means that the PIM model of the co-located position should be identical. With this consideration, a novel image construction method from multiple low quality images is proposed by estimating the PIM parameters for each pixel to enforce the local consistency.

## B. Proposed Image Reconstruction Method from Multiple Low Quality Images

For the distorted image, the local consistency property still remains, although some distortions have been introduced into the distorted images. Therefore, PIM is employed for modelling the local relationship of the distorted image in order to remove the distortions for reconstructing a higher visual quality image. The key point of the proposed scheme is to estimate the parameters  $\alpha(m, n)$  for each pixel. In our proposed scheme, the neighbourhood information of each pixel is employed to determine the parameters  $\alpha(m, n)$ . As we have multiple low quality images and they provide different partial information of the original image, not only the neighbouring pixels in the current image, but also the co-located pixels are employed to accurately estimate PIM parameters  $\alpha(m, n)$ .

Detailed information of PIM parameter estimation is illustrated in Fig. 1. Each arrow means that the corresponding pixel in the end has contribution to the head pixel. The red pixel is the current one to reconstruct. The neighbouring pixels in the current distorted image denoted as the green ones, while the pixels lying in the co-located positions are indicated as blue ones. All of these pixels are employed to estimate  $\alpha(m, n)$ . Based on the estimated PIM parameters, the pixel value of the original image can be reconstructed as:

$$\hat{I}(i,j) = \sum_{(m,n)\in T_{d_1}\cup T_{d_2}} \alpha(m,n) I(i+m,j+n)$$
(6)

where  $\hat{I}$  is the reconstructed image pixel value,  $T_{d1}$  is neighbouring pixels in current distorted image, while  $T_{d2}$  is co-located positions in the other distorted images. PIM parameters  $\alpha(m,n)$  depends on the similarity between the neighbourhoods of the pixel I(i,j) and I(i + m, j + n), and satisfies the usual conditions  $0 \le \alpha(m, n) \le 1$ . The Euclidean distance is employed to depict the neighbourhood similarity to obtain the corresponding  $\alpha(m, n)$ , which are defined by:

$$\alpha(m,n) = exp\left(-\frac{\|NS(i,j) - NS(i+m,j+n)\|_2^2}{h}\right) / Z(i,j)$$
(7)

where *h* is a parameter that acts as a degree of filtering, which is empirically set as 300. NS(i, j) denotes the neighbouring pixel values of (i, j). The neighborhood of pixel (i, j) is defined as the 7×7 image block centered at (i, j). Z(i, j) is the normalization constant: Z(i, j) (8)

$$= \sum_{(m,n)\in T_{d_1}\cup T_{d_2}} \left\{ exp\left( -\frac{\|NS(i,j) - NS(i+m,j+n)\|_2^2}{h} \right) \right\}$$
(8)

In this manner, the PIM parameter  $\alpha(m, n)$  can be obtained. It should be noted that the pixels of co-located positions in different image should contribute equally to the current pixel, which means that the weights  $\alpha(m, n)$  in  $T_{d1}$  and  $T_{d2}$  should be exactly the same. However, as aforementioned different images are degraded in different levels with different distortions. Therefore, the distorted pixels may be of different importances for the pixel reconstruction, and accordingly the weights  $\alpha(m, n)$  are different in  $T_{d1}$  and  $T_{d2}$ .

### C. Regularization of Reconstructed Image

Considering the JPEG coding process, coefficients are quantized after BDCT, which introduces the blocking and ringing artifacts. Let *B* denotes the DCT coefficients of the original  $8 \times 8$  image block,  $B_q$  is the corresponding coefficients after quantization, *Q* is the quantization matrix. The quantization process is defined as:

$$B_a(u,v) = round[B(u,v)/Q(u,v)]$$
(9)

Then the original DCT coefficient indexed by (u, v) should have its true value within a known interval:

 $B(u,v) \in \{F(u,v) | | F(u,v) - B_q(u,v) \times Q(u,v) | \le \delta\}$  (10) where  $\delta = Q(u,v)/2$ , which is referred as the quantization constraint set (QCS). With *L* multiple copies of the same image, denoting the QCS of each as  $C_i(u, v)$ ,  $i = \{1, 2, \dots, L\}$ , a narrowed quantization constraint set can be obtained as:

$$C(u, v) = \bigcap_{i=1}^{i=1} C_i(u, v) = [\hat{B}_{low}, \hat{B}_{up}]$$
  

$$\hat{B}_{low} = \max_i \{ (B_q^i(u, v) - 0.5) \times Q^i(u, v) \}$$
  

$$\hat{B}_{up} = \min_i \{ (B_q^i(u, v) + 0.5) \times Q^i(u, v) \}$$
(11)

Therefore, the image coefficients after PIM, as introduced in Section III-B, can be further regularized into a reasonable range by referring to the QCS interval. If the DCT coefficients of  $\hat{I}$  lie in the QCS C(u, v) as defined in Eq. (11), the DCT coefficients are reliable and we will preserve the corresponding coefficients. Otherwise, the DCT coefficients will be clipped by the bounds of Eq. (11). After the regularization process, we finally obtain the high quality image based on the multiple available low quality images.

#### IV. EXPERIMENTAL RESULTS

For simplicity and without losing generality, only two distorted images (with different distortions of different levels) are employed to validate the performance of our proposed method. Three distortions are employed for generating the low quality images, specifically JPEG, J2K, and fast-fading [9]. Several standard testing images are employed for experiments, such as "Lena", "Barbara", "Peppers", "Baboon", and "Boat", and several images from the LIVE image database [9], "Bikes", "Lighthouse", "Monarch", "Sailing2", "Womanhat", and "Parrots"

The objective measurements in terms of PSNR and SSIM are listed in Table I. There are two low quality images of different degradation levels. As illustrated in Table I, the combinations of the two low quality images are different, which is selected to demonstrate the robustness of our proposed method to different given inputs. It can be observed that the proposed method can produce image with higher objective measurements, specifically both PSNR and SSIM values. Compared with the distorted image of better quality, the improvement of PSNR can be up to 1.8dB and the improvement of SSIM can be up to 0.05.

Subjective comparisons are provided in Fig. 2. It can be observed that the proposed method can reconstruct higher quality images. The two input low quality images processed by JPEG, J2K, or fast-fading are of different qualities. It can be observed that the proposed method can remove the artifacts presented in the low quality images. The smooth regions are free of artifacts, compared with the blocking artifacts in JPEG coded images, such as the face of the "Lena" image. Also the ringing artifacts around the edges have been alleviated, such as the edges of the "Peppers" image. Furthermore, the jaggy artifacts which is presented in the fast-fading distorted image has also been effectively removed, meanwhile the high frequency components of the image are well preserved.

#### V. CONCLUSION

In this paper, a novel image reconstruction method based on multiple degraded images is proposed by enforcing the local consistency of image pixels between different distorted images. The reconstruction image is finally obtained by regularizing the corresponding coefficients into a reasonable range. The experimental results demonstrated that the superiority of the proposed method in both objective and subjective measurements.

#### ACKNOWLEDGMENT

The work described in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong SAR, China (Project CUHK415712), and in part by the National Natural Science Foundation of China under Grant 61202242.

#### REFERENCES

- G. Yang, C. Wang, and Y. P. Tan, "Regularized dequantization for image copies compressed with different quantization parameters", *Proc. ICIP*, 2008.
- [2] V. A. Nguyen, Y. P. Tan, and Z. Chen, "On the method of multicopy video enhancement in transform Domain", *Proc. ICIP*, 2009.
- [3] C. Wang, G. Yang, and Y. P. Tan, "Reconstructing videos from multiple compressed copies", *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 19, No. 9, Sep. 2009.
- [4] X. Zhang, and X. Wu, "Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation", *IEEE Tran. Image Process.*, vol. 17, no. 6, pp. 887-896, Jun. 2008.
- [5] X. Wu, and N. Memon, "Context-based adaptive lossless image coding", *IEEE Trans. Commun.*, vol. 45, no. 4, pp. 437-444, Apr. 1997.
- [6] B. Meyer, and P. Tischer, "TMW a new method for lossless image compression", *Proc. PCS*, 1997.
- [7] A. Said, and W. Pearlman, "An image multiresolution representation for lossless and lossy compression", *IEEE Tran. Image Process.*, vol. 5, no. 9, pp. 1303-1330, Sep. 1996.
- [8] E. Y. Lam, and J. W. Goodman, "A mathematical analysis of the DCT coefficient distributions for images", *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1661-1666, Oct. 2000
- [9] H. R. Sheikh, Z. Wang, L. Cormack and A. C. Bovik, "LIVE Image Quality Assessment Database Release 2", http://live.ece.utexas.edu/research/quality.
- [10] S. Geman, and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images", *IEEE Trans. Pattern Analysis* and Machine Intelligence, vol. 6, no. 6, pp. 721–741, Nov. 1984.
- [11] B. Zeng, "Reduction of blocking effect in DCT-coded images using zero-masking techniques", *Signal Process.*, vol. 79, no. 2, pp. 205–211, Dec. 1999.
- [12] Z. Xiong, M.T. Orchard, and Y.Q. Zhang, "A deblocking algorithm for JPEG compressed images using overcomplete wavelet representations", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 2, pp. 433–437, Apr. 1997.
- [13] G. Zhai, W. Zhang, X. Yang, W. Lin, and Y. Wu, "Efficient image deblocking based on postfiltering in shifted windows", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 1, pp. 122–126, Jan. 2008.
- [14] G. Zhai, W. Zhang, X. Yang, W. Lin, and Y. Wu, "Efficient deblocking with coefficient regularization, shape-adaptive filtering, and quantization constraint", *IEEE Trans. Multimedia*, vol. 10, no. 5, pp. 735– 745, Aug. 2008.
- [15] J. J. Zou, and H. Yan, "A deblocking method for BDCT compressed images based on adaptive projections", *IEEE Trans. Circuits Syst. Vid*eo Technol., vol. 15, no. 3, pp. 430–435, Mar. 2005.
- [16] S. H. Park, and D. S. Kim, "Theory of projection onto the narrow quantization constraint set and its application", *IEEE Trans. Image Process.*, vol. 8, no. 10, pp. 1361–1373, Oct. 1999.
- [17] R. Nakagaki, and A. K. Katsaggelos, "A VQ-based blind image restoration algorithm", *IEEE Trans. Image Process.*, vol. 12, no. 9, pp. 1044–1053, Sep. 2003.

- [18] L. Ma, D. Zhao, and W. Gao, "Learning-based image restoration for compressed images", *Signal Processing: Image Communication*, vol. 27, no. 1, pp. 54-65, Jan. 2012.
- [19] L. Ma, F. Wu, D. Zhao, W. Gao, and S. Ma, "Learning-based image restoration for compressed image through neighboring embedding", *Proc. PCM*, 2008.
- [20] Y. Liaw, W. Lo, and J. Z. Lai, "Image restoration of compressed image using classified vector quantization", *Pattern Recognition*, vol. 35 pp. 329–340, 2002.

Table I. Experimental results in terms of PSNR and SSIM

Distortion Type	Image	Low quality copy 1		Low quality copy 2		Reconstructed image	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
JPEG and JPEG	Barbara	25.6992	0.771	27.0546	0.8223	28.2510	0.8491
		25.6992	0.771	28.2538	0.8559	29.6170	0.8797
		23.8608	0.6642	25.6992	0.771	26.7080	0.8023
	Lena	30.4109	0.8184	31.9496	0.8529	33.3299	0.8798
		30.4109	0.8184	32.9631	0.8733	34.1898	0.8940
		28.2395	0.7613	31.0948	0.8341	32.5392	0.8662
	Peppers	30.1788	0.7898	31.5834	0.8234	32.6918	0.8508
		31.5834	0.8234	32.4697	0.8418	33.4827	0.8630
		29.3065	0.7681	30.1788	0.7898	31.5123	0.8293
	Baboon	22.8303	0.6403	23.4267	0.6808	24.0763	0.6908
		24.5065	0.7451	25.0028	0.7712	25.6059	0.7786
		23.4267	0.6808	24.5065	0.7451	25.1998	0.7577
	Boat	26.2536	0.6842	28.1346	0.758	29.1797	0.7894
		27.3174	0.7267	28.1346	0.758	29.2402	0.7905
		28.1346	0.758	29.5252	0.8036	30.5987	0.8298
		29.5252	0.8036	30.4935	0.8301	31.4676	0.8499
JPEG and J2K	Barbara	25.6992	0.771	26.4503	0.7791	28.0729	0.8214
		27.0546	0.8223	27.5398	0.8131	29.3675	0.8572
		23.8608	0.6642	24.7636	0.7053	25.7294	0.7308
		25.0792	0.7368	25.4479	0.737	27.0713	0.7872
	Lena	29.4666	0.7949	30.3348	0.8309	31.6545	0.8511
		30.4109	0.8184	31.0973	0.8479	32.3723	0.8656
		31.9496	0.8529	32.4015	0.8672	33.6280	0.8842
		28.2395	0.7613	29.211	0.8092	30.5278	0.8302
	Peppers	28.0375	0.7333	28.9889	0.79	30.2632	0.8105
		30.1788	0.7898	29.8681	0.8053	31.6241	0.8340
		31.5834	0.8234	31.1559	0.8244	32.867	0.8531
		31.5834	0.8234	30.2888	0.8111	32.6676	0.8509
	Baboon	22.4722	0.6134	21.6929	0.514	23.2334	0.6221
		24.5065	0.7451	24.3925	0.6891	25.3986	0.7520
		22.8303	0.6403	23.0993	0.6273	24.0630	0.6719
	Boat	26.2536	0.6842	26.9545	0.7098	28.1494	0.7448
		28.1346	0.7580	28.7442	0.7626	29.8988	0.7985
		30.4935	0.8301	30.1469	0.8063	31.6738	0.8491
JPEG and Fast- fading	Bikes	25.6255	0.778	24.4189	0.7792	26.7302	0.8245
		24.2759	0.7108	24.4189	0.7792	25.8569	0.7952
	Lighthouse	23.2924	0.6134	23.5705	0.6665	24.7939	0.6891
	Monarch	26.314	0.7923	26.8655	0.8827	28.6716	0.8953
		29.4754	0.87	26.8655	0.8827	30.9392	0.9196
	Sailing2	26.0093	0.7404	26.8118	0.7840	28.3312	0.8064
	Womanhat	29.3424	0.743	29.3316	0.7755	30.7847	0.7996
		30.3158	0.7786	29.3316	0.7755	31.3582	0.8126
	Parrots	31.9091	0.8532	30.1501	0.8888	33.6115	0.9019





JPEG: PSNR: 30.4935; SSIM:



J2K: PSNR: 30.1469; SSIM: 0.8063



Reconstructed image: PSNR: 31.6738; SSIM: 0.8491

Reconstructed image: PSNR: 26.7302; SSIM: 0.8245



JPEG: PSNR: 28.2395; SSIM:



JPEG: PSNR: 30.1788; SSIM: 0.7898



JPEG PSNR: 22.8303; SSIM: 0.6403







JPEG: PSNR: 31.0948; SSIM: 0.8341



J2K: PSNR: 29.8681; SSIM: 0.8053



J2K PSNR: 23.0993; SSIM: 0.6273



JPEG: PSNR:28.2538; SSIM: 0.8559



Reconstructed image: PSNR: 32.5392; SSIM: 0.8662



Reconstructed image: PSNR: 31.6241; SSIM: 0.834



Reconstructed image: PSNR: 24.063; SSIM: 0.6719



Reconstructed image: PSNR: 29.617; SSIM: 0.8797



0.8301

JPEG: PSNR: 25.6255; SSIM: 0.778



JPEG: PSNR: 29.4754; SSIM:0.87



JPEG: PSNR: 31.9091; SSIM: 0.8532



Fast-fading: PSNR: 30.1501; SSIM: 0.8888

Reconstructed image: PSNR: 33.6115; SSIM: 0.9019

JPEG: PSNR: 25.6992; SSIM: 0.771

Fig. 2. Subjective quality results of our proposed method, which reconstruct one high quality image from two low quality copies





Reconstructed image: PSNR: 30.9392; SSIM: 0.9196

