This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

Maximum a Posterior and Perceptually Motivated Reconstruction Algorithm: A Generic Framework

Vinit Jakhetiya, Weisi Lin, *Fellow, IEEE,* Sunil P. Jaiswal,
Sharath Chandra Guntuku, and Oscar C. Au, *Fellow, IEEE*

***Abstract*—Most of the existing image reconstruction algorithms are application-specific, and have generalization issues due to the need for parameter tuning and an unknown level of signal distortion. Addressing these problems, in this paper, we propose an efficient perceptually motivated and Maximum A-Posterior (MAP)-based generic framework for image reconstruction. This can be applied to several image/video processing applications, where there is a necessity to improve reconstruction accuracy and suppress visible artifacts, such as denoising, deinterlacing, interpolation, de-blocking of Jpeg/Jpeg-2000, demosaicing, etc. The gradient magnitudes are noise insensitive to a moderate levels of noise and we propose to utilize this property for finding the pixels with similar edge semantics in the neighborhood when neighboring pixels are noisy. With this view, we incorporate the gradient magnitude similarity based IQA metric with the MAP estimation and in turn, it can better approximate the variance of the MAP, as compared to non-linear filters. The proposed generic algorithm (without manually tuning any parameters) is shown to produce a better quality of reconstruction when compared to the state-of-the-art application-specific algorithms, for most of the image processing applications.**

***Index Terms*—IQA metrics, Denoising, Visual Perception, Non-linear filters, Generic framework.**

I. Introduction

Image reconstruction plays a crucial role from the stage of capturing data, coding, transmission and decoding, to the point of displaying data on the screen. Image reconstruction requires either the estimation of missing pixels (interpolation, deinterlacing, demosaicing, restoration of missing pixels in a randomly sampled image, etc.) or the rectification of high frequency information in the received data (de-blocking of Jpeg, removing ringing artifacts in Jpeg/Jpeg-2000 decoded images, removal of Gaussian noise, etc.). Reconstruction algorithms can be broadly divided into three categories: 1) geometric duality [14], [20-23], 2) sparse coding/learning [11], [13], [15], [32], [39-40], [50], and 3) non-linear (bilateral, Non-Local Means (NLM) filter or Kernel regression) [2], [6-13], [19-20] filtering based reconstruction.

Many algorithms [2-21], [39-47] have been proposed to address the problem of reconstruction, and these algorithms perform efficiently for the application they were developed for. However, sometimes the cause of the distortion and level of distortion of the signal are unknown, and in such cases it is not beneficial to apply an application specific algorithm to another application. For example, consider a case where a distorted image is received and displayed on a mobile or LCD screen. In this situation, this image can be distorted due to interpolation, sub-pixel rendering, compression, or a combination of these operations, and usually the exact cause of the distortion is unknown. Therefore, an efficient algorithm is required, which can remove the artifacts from the distorted image without having prior knowledge of the cause of the distortion.

Unfortunately, very few algorithms have been proposed in the literature, which can be applied to several image/video processing applications. In this paper, we refer to such algorithms as 'generic frameworks'. The following are a few examples: a) Takeda et al. [2] proposed an algorithm using a steering kernel, which can be applied to several applications, such as interpolation, noise removal, removing blocky artifacts, etc., and b) Liu et al. [5] used a hybrid graph Laplacian regularizer for restoring missing pixels. The algorithm proposed in [2] produced slightly overly smoothed output images, while the computational complexity associated with the algorithm proposed in [5] is quite high.

Our goal is to propose an algorithm which can be a generic framework to enhance the reconstruction accuracy of edges and suppress visible artifacts. These artifacts can be introduced during the phases of data capturing (Bayer's pattern, noise contamination), data transmitting (interlacing, noise contamination, pixels lost), data decoding (de-quantization of Jpeg/Jpeg-2000), displaying it on monitor (interpolation, sub-pixel rendering), or because of any other image/video processing application. While maintaining this generic nature, it is our goal that the proposed algorithm is able to produce better or competitive reconstruction quality as compared to the state-of-the-art application specific reconstruction algorithms.

Mean Square Error-(MSE) based non-linear filters [8]-[9] and their variants are extensively used for the purpose of reconstruction, especially for image denoising. Xiong et al. [6] and Hsing et al. [7] used NLM [9] and a bilateral filter [8], respectively, for removing mixed noise (both Gaussian and RVIN). These algorithms produce satisfactory results when noise density is moderate. However, for high level noise density, the reconstructed image has visible artifacts. In [10], Kunal et al. proposed to replace the mean of the errors with the median of errors to remove Gaussian noise. Recently, Talebi et al. [12] proposed an algorithm which can improve the accuracy of any of the base Gaussian noise removal algorithms, such as bilateral filter, NLM filter, and Kernel regression [2].

Vinit Jakhetiya and Sunil Jaiswal are with Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong E-mail: vjakhetiya@connect.ust.hk, spjaiswal@connect.ust.hk.

Weisi Lin and Sharath Chandra Guntuku are with School of Computer Engineering, Nanyang Technological University, Singapore, 639798 E-mail:wslin@ntu.edu.sg, sharathc001@e.ntu.sg.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

These filters were also used in several image processing and computer vision algorithms. In order to alleviate the situation of "geometric duality" mismatch, in a soft-decision and 2-d autoregressive interpolation algorithm [14], the authors of [23] proposed to incorporate a bilateral filter during the parameter estimation. Similarly, Wang et al. [21] used an NLM filter to efficiently estimate the prediction parameters for deinterlacing. For image interpolation, Hung et al. [41] estimated prediction parameters using a bilateral filter instead of covariance metric based methods. Researchers have also used these non-linear filters for deinterlacing [19-20].

One might argue that because these non-linear filters were applied to several applications they can be considered generic. However, there exist two issues with these non-linear filters. 1) The performance of these filters drops significantly when neighboring pixels are noisy and unreliable, and 2) the performance of these filters is highly dependent upon the noise density and decaying parameter. A decaying parameter with high variance may introduce blurring near the edges in the filtered image, while a decaying function with low variance cannot remove noise efficiently [28], and finding an accurate noise density for several applications is still an open research problem. In these non-linear filters, the MSE between patches/pixels is used to estimate the similarity between neighboring pixels and subsequently, weights are assigned depending upon patch similarity. For such applications, generally the neighborhood is noisy and, especially, the level and cause of degradation in the distorted signal is unknown. This might result in wrong edge direction estimation and/or blurred edges in the restored image, due to the use of non-linear filters. The inclusion of non-linear filters in the post-processing stage does not offer much of an advantage and this argument can be validated by observing the slight improvement in existing algorithms [19], [20].

Due to the above-described problems associated with the MSE and its inability to efficiently represent the signal fidelity ([25], [30]), a lot of Image Quality Assessment (IQA) metrics are proposed in the literature [1], [26], [31-33], [48] to predict image quality. Motivated by the success of SSIM [31] and problems associated with the MSE, a few researchers tried to approximate SSIM and incorporated it within the rate distortion optimization [29], [35-37]. Rehman et al. [34] made the first attempt to use an IQA metric (SSIM) to restore the edges of Gaussian noise contaminated images. Unfortunately, this work is only able to restore the image distorted with Gaussian noise. The above-mentioned algorithms produce visually appealing output images but achieve a lower $PNSR$ than that achieved by using MSE as the optimization criteria.

As previously described, most of the reconstruction algorithms, which include non-linear filters [8-9] in the post-processing stage, can not perform efficiently when neighboring pixels are unreliable and the level and cause of distortion to the signal is unknown. At the same time, very few IQA metrics based optimization algorithms exist, which shows the advantage (only in terms of subjective evaluation) of using IQA for optimization over the MSE. However, none of these algorithms are generic and can be applied to several applications. On the other hand, the gradient magnitudes are noise insensitive to the moderate level of noise [1] and this property can be utilized to find the pixels with similar edge semantics in the neighborhood, even when neighboring pixels are not accurate. This motivates us to incorporate the gradient-magnitude-similarity-based GMSD [1] metric with the image reconstruction and, in turn, the proposed algorithm with the use of GMSD metric does not require the tuning of any parameters to suit the characteristics of the image and specific image processing application.

To this end, we propose an efficient perceptually motivated two stage algorithm, where, in first stage, a simple reconstruction algorithm is used to restore the missing/distorted pixels, and in the next stage, the reconstruction accuracy of these pixels is enhanced using the proposed MAP formulation using the GMSD metric. The main contributions and advantages of the proposed algorithm are as follows.

1) We formulate the proposed reconstruction algorithm as a MAP problem, which has a closed form solution and requires only one iteration to obtain the optimal solution.
2) We combine the GMSD metric with the MAP estimation, which has a very good capability of approximating variances of the MAP and detecting edge direction correctly in complex regions (i.e., in situations when the neighborhood is unreliable).
3) The proposed algorithm is free from manual tuning of the parameters to suit the characteristics of any specific image-processing application.

The rest of the paper is organized as follows. Section II briefly describes the GMSD metric. While section III represents the proposed MAP formulation and approximation of the variance of the MAP estimation. The experimental results for different applications and discussions are provided in Section IV and the concluding remarks are presented in Section V.

II. Brief overview of GMSD metric

The GMSD [1] is a recent IQA algorithm based on gradient magnitude similarity, which is accurate and computationally efficient. The GMSD metric assumes that the gradient magnitudes of two images are close to each other if these images are visually similar. Henceforth, deviation of the gradient magnitude similarity for these images will be quite low, and vice-versa. The gradient magnitude similarity (GMS) between the original image and distorted image (for our algorithm gradient magnitude similarity between the pixel of interest and its neighboring pixels) is calculated as follows:

$$GMS(j) = \frac{2 \times m_{B_i}(j) \times m_{A_{ik}}(j) + c}{m^2_{B_i}(j) + m^2_{A_{ik}}(j) + c} \tag{1}$$

Here, m_{B_i} and $m_{A_{ik}}$ are gradient magnitudes of the pixel of interest and their neighboring pixels, and these gradient magnitudes (m_{B_i} and $m_{A_{ik}}$) reflect the edge strength of the corresponding pixels. In order to calculate the final score, these GMS values are pooled using standard deviation and this is called the gradient magnitude similarity deviation (GMSD):

$$GMSD = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(GMS(j) - GMSM\right)^2} \tag{2}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

Here, $GMSM$ is the mean of all the pixels' GMS values. The lower value of GMSD (close to zero) suggests that these two pixels have quite similar edge structures, and in the proposed algorithm weighting parameters corresponding to the these pixels should attain a higher value, and vice-versa. More details about the GMSD metric can be found in [1].

III. Proposed Algorithm

As aforementioned, we attempt to build a generic computational model, which can enhance the reconstruction accuracy of most of the image processing applications, such as image denoising, interpolation, sub-pixel rendering, deinterlacing, Jpeg/Jpeg-2000, etc. For this purpose, we propose a Maximum A-Posterior (MAP) and perceptually motivated reconstruction algorithm. The success of the MAP estimation lies in the efficient estimation of the variance of the MAP. As the available pixels are not original (restored in the first stage) pixels, it is not possible to estimate the original variance, so, efficient approximation of the variance of the MAP is essentially required for a generic reconstruction algorithm. In the proposed algorithm, we approximated the variance of the MAP using gradient-similarity-based IQA metrics (namely GMS [48] and GMSD [1]), due to its noise insensitivity (to a certain limit) property, [1] and this improvement in the approximation of the variance enables the proposed algorithm to better estimate the edge direction when compared to the commonly used MSE-based bilateral filter [8] and NLM filter [9], especially in the presence of a noisy and unreliable neighborhood.

In general, interpolated and deinterlaced images/frames are not free from zig-zagging artifacts, while, decoded images using a Jpeg-2000 decoder have ringing artifacts near the edges and these artifacts are clearly visible at low-bit rates. On the other hand, denoised images contaminated with Gaussian noise are either overly smoothed or noise is still visible in the filtered image. In this work, our goal is to propose a generic algorithm that is not application specific and can suppress the above-mentioned artifacts associated with various image processing applications. To efficiently reconstruct these missing/distorted pixels, we propose a two stage algorithm (reconstruction and refinement of reconstruction accuracy). In the first stage, the missing/distorted pixels (if any) are restored using an existing simple reconstruction algorithm (such as, [4], [9], [17], [16] and [38]). In the second stage, the reconstruction done in the first stage is refined using the proposed MAP formulation. In order to build the proposed model we start with the deinterlacing problem, and discuss the application of the proposed formulation on image processing applications in further sections.

In deinterlacing, half of the pixels of a frame are dropped at the encoder side, and at the decoder side, these missing pixels are reconstructed using neighboring pixels [17-20]. Throughout this paper, bold and non-bold notations are used to indicate metrics and point-wise representations, respectively. We propose to formulate the post-processing (second) stage as

$$\underset{\boldsymbol{B}}{arg\,max} \quad P[\boldsymbol{B}|\boldsymbol{A},\tilde{\boldsymbol{B}}] = \quad \underset{\boldsymbol{B}}{arg\,max} \quad \frac{P[\boldsymbol{A}|\boldsymbol{B},\tilde{\boldsymbol{B}}]P[\boldsymbol{B}|\tilde{\boldsymbol{B}}]}{P[\mathbf{A}|\tilde{\mathbf{B}}]} \quad (3)$$

Here, $\mathbf{B}$ and $\tilde{\mathbf{B}}$ represent the image which needs to be restored, and the restored image in the first stage, respectively. $\mathbf{A}$ is the interlaced frame, which contains only half of the pixels of the original frame. As the denominator of (1) does not give any additional information, it can be omitted.

$$\boldsymbol{B}^* = \underset{\boldsymbol{B}}{arg\,max} \quad P[\boldsymbol{A}|\boldsymbol{B},\tilde{\boldsymbol{B}}]P[\boldsymbol{B}|\tilde{\boldsymbol{B}}] \quad (4)$$

In (2), the first and second terms are likelihood and prior, respectively. In the proposed algorithm, we assume that the residual errors (between the restored image in the first stage and original image) follow a Gaussian distribution [21], [23], [46]. So, the prior term is modeled as

$$P[\boldsymbol{B}|\tilde{\boldsymbol{B}}] = \prod_{i=1}^{M} C_i \times \exp(-\frac{(B_i - \tilde{B}_i)^2}{\beta_i}) \quad (5)$$

Here, M, C_i and β_i are the number of missing pixels, scale parameter and variance of the Gaussian distribution, respectively. The same way as the prior, we also modelled the likelihood term using a Gaussian distribution [21], [23], [46]:

$$P[\boldsymbol{A}|\boldsymbol{B},\tilde{\boldsymbol{B}}] = \prod_{i=1}^{M} D_i \times \exp(-\sum_{k\in N_i} \frac{(B_i - A_{ik})^2}{\Delta_{ik}}) \quad (6)$$

Here, A_{ik}, D_i and Δ_{ik} represent the i^{th} missing pixel's neighboring pixels in a window N_i, the scale parameter, and the variances of the difference between the missing pixel and its neighboring pixels, respectively. In the proposed algorithm we use an adaptive window to exploit the relationship between the missing pixel and its neighboring pixels (See Subsection C for more details). After using (5-6) in (4), the MAP formulation results in

$$\boldsymbol{B}^* = \underset{\boldsymbol{B}}{arg\,max} \prod_{i=1}^{M} C_i D_i \times \exp(-\sum_{k\in N_i} \frac{(B_i - A_{ik})^2}{\Delta_{ik}} + \frac{(B_i - \tilde{B}_i)^2}{\beta_i}) \quad (7)$$

For the sake of compact representation, we replace the variances β_i and Δ_{ik} with $1/\lambda_i$ and $1/w_{ik}$, respectively. Through-out this paper, we call w_{ik} as weighting parameters and it represents the variance of the MAP Δ_{ik}. The MAP estimator can be simplified as:

$$\boldsymbol{B}^* = \underset{\boldsymbol{B}}{arg\,min} \sum_{i=1}^{M} (\lambda_i (B_i - \tilde{B}_i)^2 + \sum_{k\in N_i} \omega_{ik} (B_i - A_{ik})^2) \quad (8)$$

It can be observed that the MAP formulation (8) is a convex function without explicit constraint. Hence, the optimal solution in closed form is obtained by differentiating (8) and equating to zero. The solution of the MAP estimator is obtained as

$$B_i^* = \frac{\tilde{B}_i \lambda_i + \sum_{k\in N_i} w_{ik} A_{ik}}{\lambda_i + \sum_{k\in N_i} w_{ik}} \quad (9)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

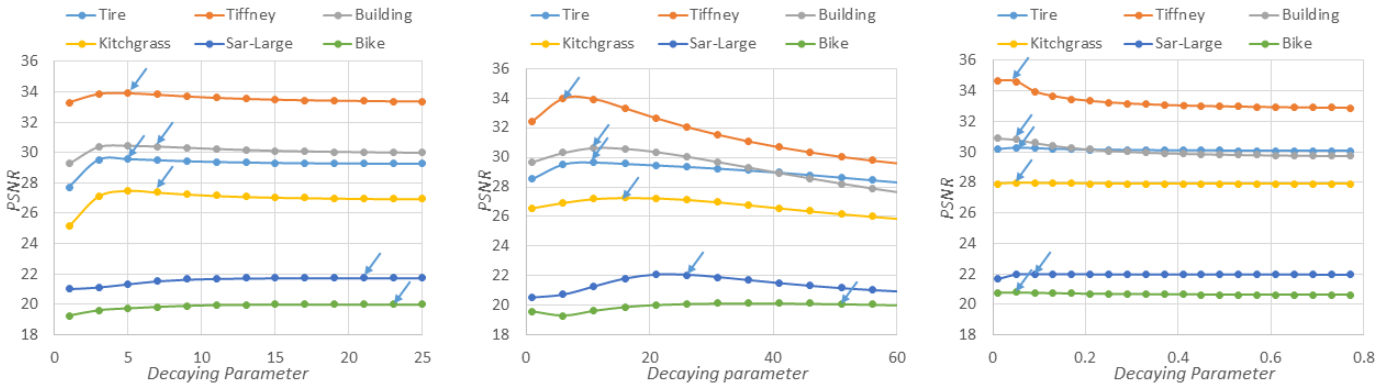

Fig. 1. Dependency of reconstruction accuracy (in terms of $PSNR$) of (a) bilateral filter, (b) NLM filter, and (C) GMSD [1] metric, with a decaying parameter for predicting the missing pixels of the interlaced image. Here, blue arrows represent the best performance.

Here, λ_i is the weighting parameter of the prior term, which controls the contribution of the prior, and w_{ik} represents the variance of the MAP estimation. We used MELA [17] for reconstructing the missing pixel in the interlaced frame, in the first stage. The reconstruction accuracy of the distorted signal highly depends upon the approximation efficiency of these parameters (λ_i and w_{ik}). We propose to approximate these parameters as described below.

A. Approximating the weighting parameter of prior

This parameter λ_i highly depends upon the characteristics of the image [20] and the reconstruction algorithm used in the first stage ($\tilde{B}$). In the proposed algorithm, reconstructed pixels in the first stage are used as the prior information and these pixels are not accurate. Therefore, the contribution of the prior term should be low, when the reconstruction algorithm used in first stage is not able to efficiently reconstruct the missing/distorted pixels, such as edge and texture pixels, and vice-versa.

We assume that any simple reconstruction algorithm, such as bilinear, bicubic, bilateral filter, NLM filter, and MELA [17] can reconstruct a pixel efficiently if it is smooth in nature. This degree of smoothness is estimated using the variance of the pixel of interest, as a higher variance of a pixel suggests it is in a high activity area (edge or texture). On the other hand, these algorithms cannot predict edge and texture pixels accurately. In the situation when the pixel is smooth in nature, there is no need to enhance the reconstruction accuracy as these pixels are already reconstructed efficiently. Restoring these pixels again with the help of unreliable pixels in the neighboring window (N_i) may cause blurring (over-smoothing) or degradation of small details. Therefore, we estimate the weighting parameter λ_i for the prior term as

$$\hat{\lambda}_i = \frac{K}{Var(\tilde{B}_i)} \tag{10}$$

Here, K is a constant, and we chose the value of this constant to be 20 empirically. In the proposed algorithm, we used a 5×5 window to calculate the variance of the reconstructed pixel ($\tilde{B}_i$).

B. Approximation of the variance of the MAP

In (9) w_{ik} are the weighting parameters representing the variance of the MAP, corresponding to the pixels in the neighboring window N_i. In the literature, these weighting parameters [20-21], [23] (w_{ik}) are generally estimated using MSE-based non-linear filters [8-9], or variants of these non-linear filters. The weights of non-linear filters are defined as

$$w_{ik} = \exp\left(-\frac{||\tilde{B}_i - A_{ik}||^2}{2\sigma_1}\right)\exp\left(-\frac{||D - D_k||^2}{2\sigma_2}\right) \tag{11}$$

In these non-linear filters, the first term represents radiometric distance and second term refers to the geometric distance. The radiometric distance in non-linear filters is calculated based upon the pixel difference or MSE. In previous research [29], [35-37], authors have used SSIM [31] or variants of SSIM for rate-distortion optimization. They have highlighted the issue with MSE by giving theoretical arguments [29], [35-37]. In our work, we point out how a few of the concerns associated with MSE as mentioned in previous research, lead to inefficient estimation of weighting parameters which are based upon MSE-centred non-linear filters. Subsequently, these inefficient weighting parameters lead to poor approximation of the variance of the MAP. These drawbacks of non-linear filters due to concerns associated with MSE are as follows:

In (11), σ_1 and σ_2 are the decaying parameters which control the damping, and these parameters highly depend upon the level of signal distortion [8-12], [27] and cause of the distortion in the input image. If these parameters have a small value, then the filtered image is not artefacts free and visible noise patches can be observed. With the increment of distortion level, these filters require to attain a higher value of this parameter (σ_1), as pixels in the neighboring window have a higher degree of unreliability and it leads to blurring in the filtered signal. So, the success of these filters lies in choosing an optimal value of parameter (σ_1). A distortion level can be efficiently estimated for Gaussian noise and in the same way, this decaying parameter can be accurately adjusted based upon the distortion level [28]. So the non-linear filters perform very efficiently for removing Gaussian noise. While for the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

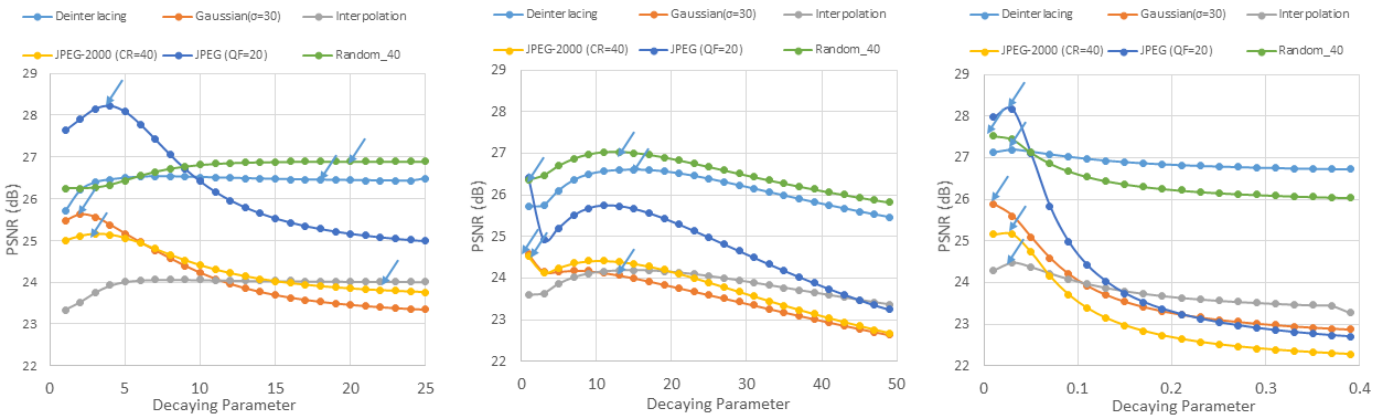

Fig. 2. Dependency of average reconstruction accuracy (in terms of $PSNR$) of (a) bilateral filter, (b) NLM filter, and (c) GMSD [1] metric, with decaying parameter for several image processing applications. Here, blue arrows represent the best performance. In order to show the average reconstruction accuracy for each of these applications, we calculated the $PSNR$ for six images corresponding to the decaying parameter and averaged the $PSNR$.

other image processing applications, accurately estimating the distortion level is still an open research problem [45]. As the level of the signal distortion can not be efficiently estimated, the performance gain achieved by using these non-linear filters is quite limited [19], [20], [23].

The MSE is not a reliable measurement of the signal fidelity, [25], [31], and even MSE is not a perceptually relevant [29] metric to find the correlation among the neighboring pixels that can match with the human visual system [22]. As reconstructed pixels in the first stage are used to estimate the MSE, and these reconstructed pixels are not accurate. So, the calculated MSE will be low between the pixels that have a similar neighborhood pattern, instead of the pixels that have similar edge semantics. In the situation where neighboring pixels are unreliable in the searching window, non-linear filters based upon MSE are not efficient, which leads to poor approximation of the variance of the MAP. The poor approximation of the variance of the MAP, may result in wrong edge direction detection or blurred edges.

In order to justify these arguments and how our proposed algorithm using IQA metrics can overcome these concerns, we have done following empirical studies.

To give insight into two drawbacks associated with MSE-based non-linear filters, we down-sample six frequently used images/frames by a factor of 2 in the vertical direction. The missing pixels of the down-sampled image are first reconstructed (deinterlaced) using the computationally simple MELA [17] algorithm. The proposed model (9) is applied on this reconstructed image, in which weighting parameters (w_{ik}) are estimated based on a bilateral filter, NLM filter and GMSD metric. In the first empirical study, we have shown how an unreliable MSE [29], [35-37] makes the performance of non-linear filters to be highly dependent upon the decaying parameter (σ_1).

In Fig. 1(a) and (b), we have shown the dependency of the reconstruction accuracy of these filters with decaying parameter (σ_1) for predicting the missing pixels of an interlaced image. Similarly, in Fig. 2(a) and (b), we have shown the average performance of these filters with respect to the decaying parameter for several image processing applications (such as deinterlacing, interpolation, denoising, random sampling, and compression) for six images. From figures 1 and 2, we can observe that 1) the performance (in terms of $PSNR$) of these filters highly depends upon the decaying parameter and it is difficult to find a set of optimal decaying parameters for different images and image processing applications; and 2) even with the optimal decaying parameter, the performance of these filters is worse than the proposed algorithm (more details are given in the next paragraph), which shows the inefficient estimation of weighting parameters using these filters.

From these arguments, it can be seen that MSE is not a reliable option for the estimation of weighting parameters, or approximation of the variance of the MAP. It is a well-known fact (in the image quality assessment research community) that the gradient magnitude of the original and its distorted signal does not change significantly with a moderate level of distortion addition in the signal [1]. This fact has been excessively used in image quality assessment research; for example the authors of [48] have used gradient magnitude similarity (GMS) and [1] have utilized deviation of GMS (GMSD), and the performance of these metrics is quite good. At the same time, in the literature, the use of GMSD is missing for the purpose of generic image reconstruction. In image reconstruction it is required to find pixels with similar edge semantics, even in the situation when the neighbouring pixels are distorted. The GMSD values are generally noise-insensitive to a certain level of distortions, due to the use of gradient [1]. So, the GMSD between the distorted pixel and its neighbouring pixels is nearly similar to the GMSD of the original pixel and its neighbouring pixel. It suggests that the use of GMSD can help to find similar pixels in the neighbourhood, even when these pixels are not accurate and in turn, the proposed algorithm with the use of the GMSD metric does not require the tuning of any parameters to suit the characteristics of the image and specific image processing applications. With this view, we have incorporated the gradient magnitude similarity based GMSD metric with the image reconstruction.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

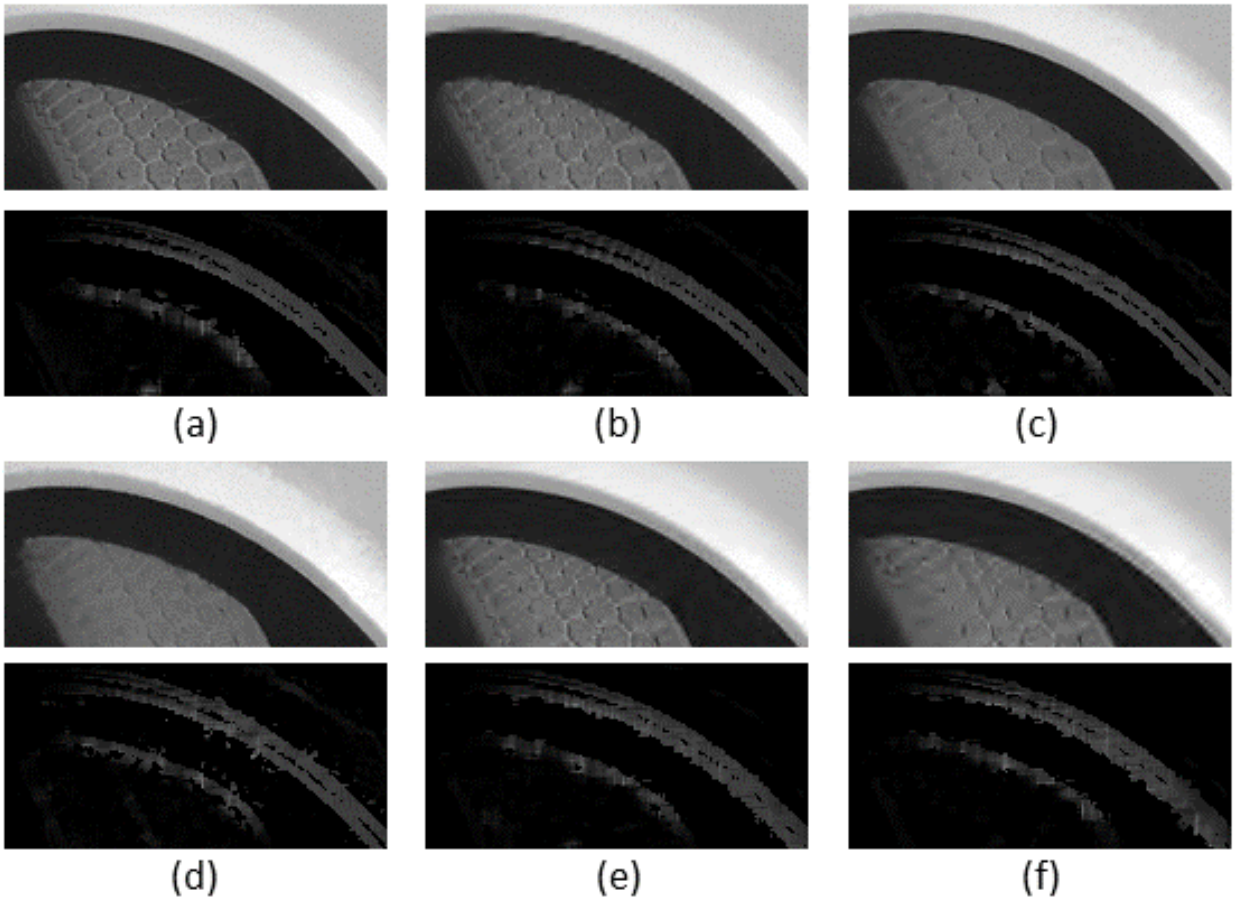

Fig. 3. Dependency of the GMSD with different types and level of distortions such as, (b) deinterlacing, (c) and (d) Gaussian noise with σ = 10 and 20, respectively, (e) and (f) J2k compression with compression ratio = 10 and 20, respectively, for 'Tire' image (a). In each figure, the first row is the distorted image and the second row shows the inverse of GMSD values between the pixel of interest and its neighbouring pixel (in the diagonal direction).

In order to validate these arguments, in Fig. 3, we have shown the invrerse of GMSD between the pixel of interest and its neighbouring pixel (in the diagonal direction) for the original and distorted images, in which distortion is caused by the deinterlacing, Gaussian noise (σ = 10 and 20) and J2k compression (Compression ratio = 10 and 20). From this figure, we can observe that the GMSD values of the pixels of the distorted image (Fig. 3(b) - (f)) are nearly similar to the GMSD value of the pixels of the original image (3(a)). It suggests that the GMSD can identify pixels with similar edge semantics in noisy situations.

Due to the ability of GMSD to find the similar pixels in the neighborhood, the proposed model does not require the tuning of any parameters to suit the characteristics of an image and the cause of distortion, as shown in figures 1(c) and 2(c). In these figures, the optimal decaying parameter corresponding to the best performance (as shown by blue arrows in figures 1(c) and 2(c)) is nearly similar for different image processing applications and images which have different distortion levels. This empirical study validates our assumption and suggests that the proposed model (9) using the gradient-similarity-based GMSD metric does not require adjustment of the decaying parameter (σ_1), to suit the characteristics of the signal and level and cause of the distortion to the signal. Secondly, this accurate finding of pixels with similar edge semantics lead to efficient estimation of weighting parameters (w_{ik}) and subsequently, better approximation of the variance of the MAP (as shown in Fig. 4). This observation also validates the arguments given in [29], [35-37] that MSE is a poor reflector of signal fidelity.

Hence, with the use of IQA metrics, we are able to overcome both of the major problems associated with non-linear filters for weighting parameter estimation. With this view, we propose to incorporate the IQA (GMSD) metric with non-linear filters to estimate the weighting parameters as follows:

$$w_{ik} = \exp(-\frac{GMSD(\tilde{B}_i, A_{ik})}{2\sigma_1}) \exp(-\frac{||D - D_k||^2}{2\sigma_2}) \quad (12)$$

Here, the second term in the proposed filter accounts for the geometric distance while, σ_1 and σ_2 are the decaying parameters, which were set as 0.02 and 2.0 in all the experiments for each application. These weighting parameters are used to approximate the variance of the MAP and subsequently, are used in (9) to reconstruct the final image.

In order to show how an inaccurate MSE leads to a poor estimation of the variance of the MAP, we estimated the variance of the pixel of interest ($\hat{\Delta_{ik}}$) based on three methods (bilateral filter, NLM filter and GMSD metric) for each pixel.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

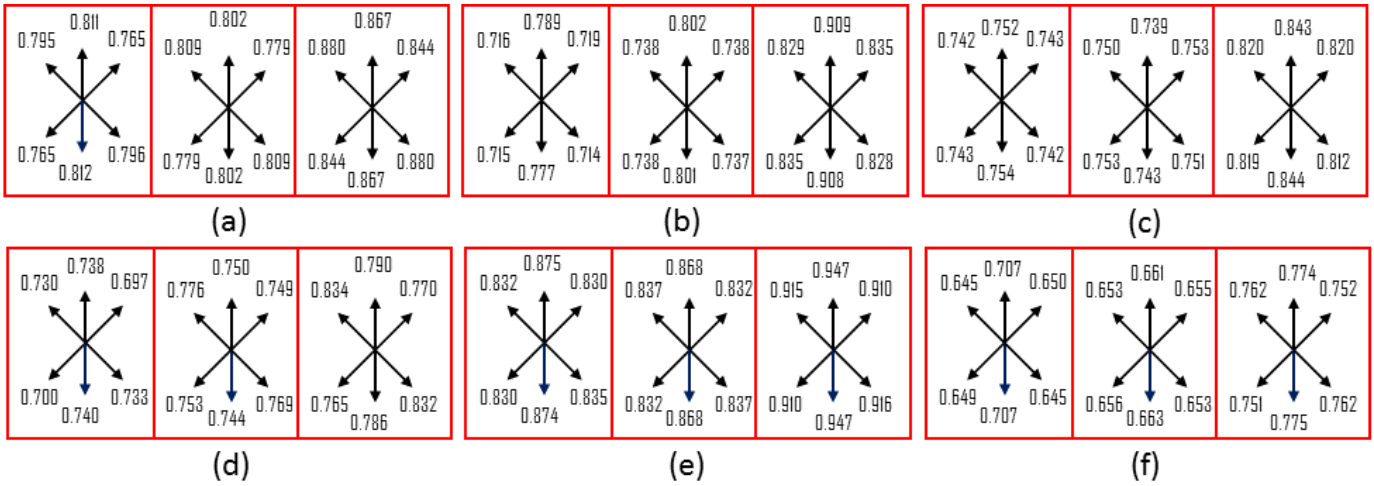

Fig. 4. Comparison of the ability of approximating the variance of the MAP estimation. The first, second and third columns of each figure represent correlation coefficients (ρ) between the variance (Δ_k) and approximated variances ($\hat{\Delta}_k$) with the bilateral filter [8], [20], NLM filter [9] and proposed algorithm, respectively. In figures (a)-(f) correlation coefficient (ρ) is shown for 'Tire', 'Tiffney', 'Building', 'Kitchgrass', 'Sar-Large', and 'Bike' image/frame, respectively, in six directions.

At the same time, we also calculated the original variance of the MAP (Δ_{ik}), assuming that all the pixels are available, and calculated the correlation coefficient between the original variance and approximated variance for these three methods. The correlation coefficient ρ_k is defined as

$$\rho_k = \frac{\sum_{i=1}^{M}(\Delta_{ik} - \bar{\Delta}_{ik})(\hat{\Delta}_{ik} - \bar{\hat{\Delta}}_{ik})}{\sqrt{\sum_{i=1}^{M}(\Delta_{ik} - \bar{\Delta}_{ik})^2 \sum_{i=1}^{M}(\hat{\Delta}_{ik} - \bar{\hat{\Delta}}_{ik})^2}} \qquad (13)$$

Here, $\bar{\Delta}_{ik}$, $\bar{\hat{\Delta}}_{ik}$, k and M represents the mean of the original variance, mean of approximated variance, direction in which the variance of the MAP is estimated and number of missing/distorted pixels, respectively. The correlation coefficient (ρ) between the original variance and approximate variance using a bilateral filter, NLM filter and the proposed method for six images in six directions (k) are shown in Fig. 4. The proposed algorithm achieves the highest value of correlation coefficient (ρ_k) for all six directions, which suggests that the proposed method has a much better capability of approximating the variance of the MAP, as compared to the extensively used non-linear filters [8-9], [20].

In order to show the advantage of having a better approximation of variance of the MAP in the proposed algorithm, we applied it to enhance the reconstruction accuracy of deinteralaced 'Tire' image using the MELA [17] algorithm. The proposed algorithm is compared with the MELA [17], LCID [18], DCS [19] and OBF [20] filter, respectively. The subjective quality comparison for 'Tire' image is shown in Fig. 5, and from this comparison it is clearly visible that the use of the bilateral and NLM filters does not give much improvement (as shown in Fig. 5(d) and (e)), while the proposed algorithm can improve the reconstruction accuracy to a great extent, and suppresses the visible artifacts (zig-zagging).

These previously given arguments and empirical studies validates the replacement of MSE with the GMSD metric in non-linear filters.

C. Adaptive window size

In the proposed algorithm, we have used an adaptive window size which offers a good trade-off in computational requirements and finding a pixel which has similar edge semantics to the pixel of interest. A smaller window might not be able to find such kinds of pixels and a bigger window requires a higher computational power to estimate the weighting parameters.

In the proposed algorithm, we assume that if the GMSD value corresponding to the missing/distorted pixel ($\tilde{B}$) and neighboring pixel is smaller than a threshold (γ), then these pixels have similar edge semantics. In such situations, there is no need to increase the window size. In contrast, if the GMSD value of each pixel in a neighboring window (N_i) is higher than this threshold (γ), it suggests than none of the pixels in the neighboring window have similar edge semantics to the pixel of interest ($\tilde{B}$). In this situation, the window size must be increased to efficiently find the pixels with similar edge semantics. In the proposed algorithm, from experiments, we chose this threshold (γ) to be 0.07. The algorithm for finding the accurate window size for the i^{th} pixel can be summarized as follows.

IV. Experimental Results and Discussion

In this paper, we propose a perceptually motivated and MAP-based generic algorithm. The proposed algorithm can be applied to several image processing applications, where there is a need to enhance reconstruction accuracy and suppress visible artifacts. To evaluate the performance of the proposed algorithm, we applied it on images in which distortion was caused due to interpolation, deinterlacing, denoising, reconstruction of the randomly sampled image and compression (Jpeg-2000 and Jpeg). The proposed algorithm can be applied to very significantly improve the reconstruction accuracy of other image and video processing applications, such as demosaicing, sub-pixel rendering, water-marking, lossless compression, and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

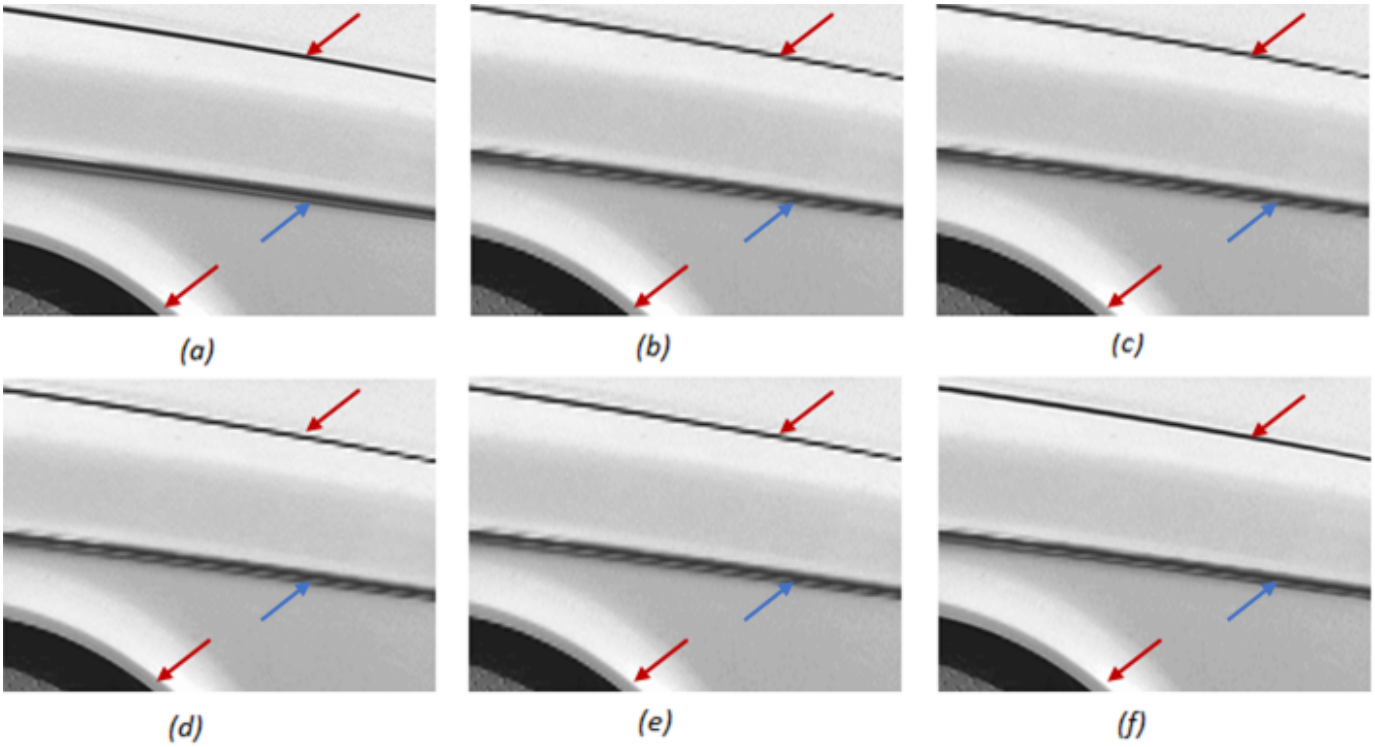

Fig. 5. Comparison of subjective quality for deintelaced 'Tire' image (a) with, (b) MELA [17] (29.347 dB), (c) LCID [18] (29.368 dB), (d) DCS [19] (29.500 dB), (e) OBF [20] (29.812 dB), and (f) the proposed algorithm (30.267 dB).

Input : Pixel of interest ($\tilde{B}_i$),
for $V_i = 3;\ \ V_i < 8;\ \ V_i$++;
do
  $Window\ size\ \ N_i = V_i \times V_i$
  for $Every\ Pixel\ P \in Window\ N_i$;
  do
    if $(GMSD(P, \tilde{B}_i) \leq 0.07)$ **then**
      $Stop$
    end
  end
end
Output : Accurate Window size (N_i)

Algorithm 1: Pseudocode to adaptively determine the window size.

TABLE I
SIMULATION RESULTS COMPARISON FOR INTRA-FRAME DEINTERLACING. THE BEST PERFORMANCE IS HIGHLIGHTED WITH BOLDFACE.

	$PSNR$ (dB)				
Image	[17]	[18]	[19]	[20]	Prop.
Tiffney	33.806	33.740	34.260	34.409	**34.736**
Tire	29.347	29.368	29.500	29.834	**30.267**
Building	30.376	30.366	30.522	30.491	**30.930**
Kitchgrass	27.239	27.254	27.324	27.677	**27.955**
Sar-Large	21.206	21.302	21.206	21.732	**21.800**
Bike	19.738	19.843	19.754	20.296	**20.806**
Average	26.952	26.979	27.094	27.406	**27.749**

deblocking. The proposed algorithm is a two stage algorithm: in the first stage, missing or distorted pixels are reconstructed using the existing algorithms (such as [17], [16], [4], [9] and [38]), and in the later stage, reconstruction accuracy is improved using the proposed perceptually motivated MAP formulation.

It is to be noticed that through-out the experiments, we did not tune any parameters. The parameter setting (i.e., threshold (γ) for determining the optimal window size, σ_1 and σ_2) for each experiment and every application is the same. The main advantage of the proposed algorithm is that the image characteristics and reason behind the cause of the distortion does not significantly affect the performance. We classified the experimental results into two categories: 1) where missing pixels are restored (i.e., deinterlacing, interpolation, in randomly sampled images), and 2) where available pixels are distorted and need to be restored (i.e., removal of Gaussian noise and in decoded images using Jpeg-2000 and Jpeg).

To deinterlace, interpolate and restore missing pixels of the randomly sampled image, we used the MELA [17], contrast guided (CG) [16], and IFASDA [4] algorithms, respectively, in the first stage. For deinterlacing, we compared the proposed algorithm with the existing MELA [17], LCID [18], DCS [19], and OBF [20] algorithms. While for interpolation, the performance of the proposed algorithm is compared with the bicubic [24], directional bicubic [49], SAI [14], SME [15] and CG [16] algorithms. We have also shown the experimental results for restoring the missing pixels of the randomly sampled image, and the proposed algorithm is compared with the DPR [42], EPA [43], IFASDA [4], and Hybrid-Graph [5] algorithm.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

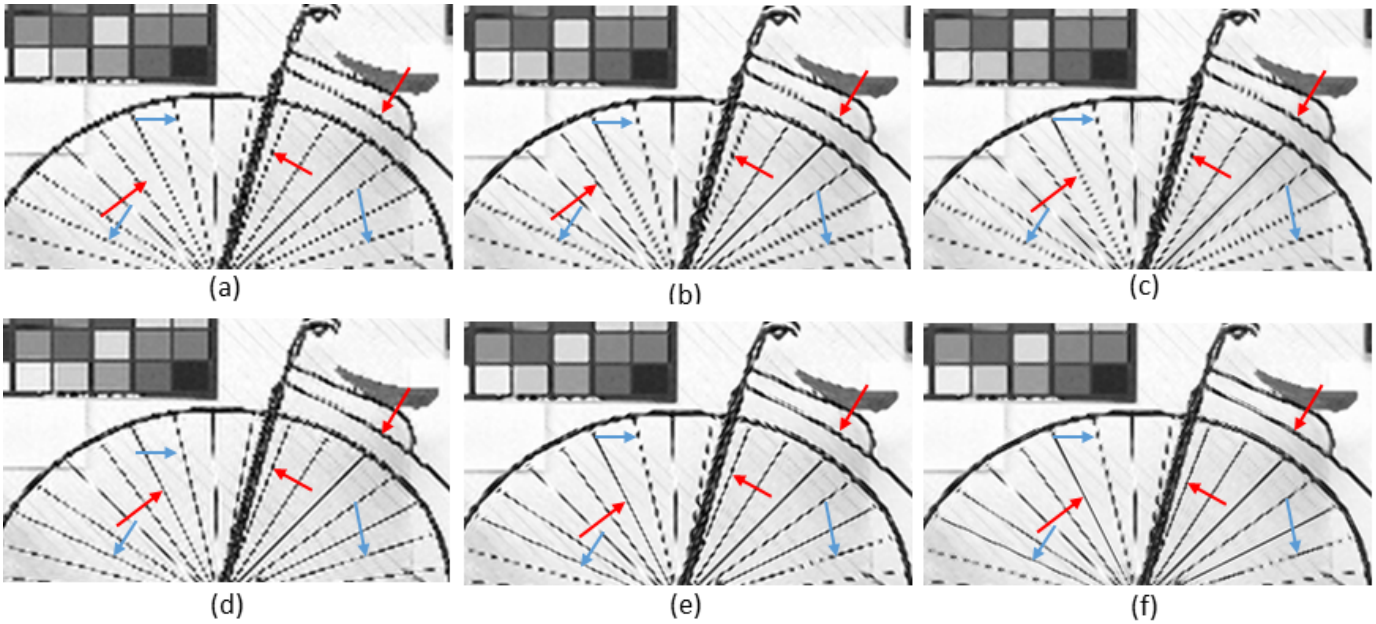

Fig. 6. Comparison of subjective quality for interpolated 'Bike' image with a factor of 2 with, (a) bicubic [24] (17.335 dB), (b) directional bicubic [49] (17.541 dB), (c) SAI [14] (17.656 dB), (d) SME [15] (17.700 dB), (e) contrast guided [16] (17.566 dB) and (f) proposed algorithm (18.110 dB), respectively.

TABLE II
COMPARISON OF RESTORATION RESULTS FOR IMAGE INTERPOLATION. THE BEST PERFORMANCE IS HIGHLIGHTED WITH BOLDFACE.

	$PSNR$ (dB)					
Image	[24]	[49]	[14]	[15]	[16]	Prop.
Tiffney	29.397	29.945	30.076	29.989	29.921	**30.158**
Baboon	22.959	23.051	23.234	23.180	23.041	**23.366**
Sar-Large	18.802	18.971	19.120	18.930	18.899	**19.565**
Aerial2	21.634	21.772	21.866	21.800	21.628	**22.049**
Ruler.512	11.735	11.687	11.217	11.170	11.820	**12.934**
Party	15.237	15.314	15.436	15.280	15.184	**15.620**
Bike	17.341	17.540	17.656	17.700	17.566	**18.110**
Average	19.586	19.754	19.801	19.721	19.723	**20.115**

TABLE III
COMPARISON OF RESTORATION RESULTS FOR IRREGULARLY SAMPLED IMAGES. HERE 'PEPPERS20' REPRESENTS 'PEPPERS' IMAGE WHOSE 20% OF PIXELS WERE RANDOMLY REMOVED. THE BEST PERFORMANCE IS HIGHLIGHTED WITH BOLDFACE.

	$PSNR$ (dB)				
Image	[42]	[43]	[4]	[5]	Prop.
Peppers20	39.738	37.398	39.397	36.433	**39.982**
Peppers30	37.603	35.989	37.257	35.864	**37.880**
Peppers40	35.996	34.518	35.571	35.293	**36.108**
Hat20	39.105	37.568	38.442	36.404	**39.336**
Hat30	36.911	35.823	36.137	35.523	**37.167**
Hat40	34.990	34.522	34.424	33.443	**35.535**
Tiffney20	35.835	34.872	34.724	35.500	**37.033**
Tiffney30	33.639	33.104	32.893	32.933	**34.809**
Tiffney40	32.203	31.807	31.503	32.379	**33.185**
Average	36.224	35.067	35.594	34.863	**36.782**

TABLE IV
COMPARISON OF RESTORATION RESULTS FOR IMAGE ADDED WITH GAUSSIAN NOISE. HERE 'BABOON20' REPRESENTS THE 'BABOON' IMAGE CONTAMINATED WITH GAUSSIAN NOISE WITH NOISE LEVEL σ = 20. THE BEST PERFORMANCE IS HIGHLIGHTED WITH BOLDFACE.

	$PSNR$ (dB)					
Image	[9]	[44]	[2]	[10]	[12]	Prop.
Baboon20	26.07	26.28	25.22	23.50	25.31	**26.31**
Baboon30	24.01	24.23	23.17	21.98	23.11	**24.40**
Baboon40	22.57	22.42	22.07	21.48	21.94	**23.14**
Boat20	26.67	**27.00**	26.72	24.23	26.62	26.97
Boat30	24.72	25.22	24.93	23.06	24.84	**25.23**
Boat40	23.19	**23.83**	23.84	22.45	23.73	23.76
Hat20	30.61	**32.23**	30.72	30.80	31.62	31.90
Hat30	28.01	**30.44**	28.77	28.94	29.70	30.01
Hat40	26.06	**29.03**	27.54	27.52	28.51	28.35
Average	25.77	**26.74**	25.88	24.89	26.15	**26.67**

The objective comparison in terms of $PSNR$ for deinterlacing, interpolation, and restoring the missing pixels of the randomly sampled image is shown in Tables I-III. The proposed algorithm improves the $PSNR$ by 0.817, 0.392 and 1.188 dB as compared to the reconstruction algorithms used in the first stage, such as, MELA [17] for deinterlacing, CG [16] for interpolation, and IFASDA [4] for restoring the missing pixels of the randomly sampled data, respectively. For all three applications, the proposed algorithm achieves a higher $PSNR$ than the recently developed state-of-the-art reconstruction algorithms. These results confirm the effectiveness of including the second stage using the proposed MAP formulation.

To show the visual superiority of the proposed algorithm, the subjective quality of the restored images with the state-of-the-art and the proposed algorithm is shown in figures 5

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

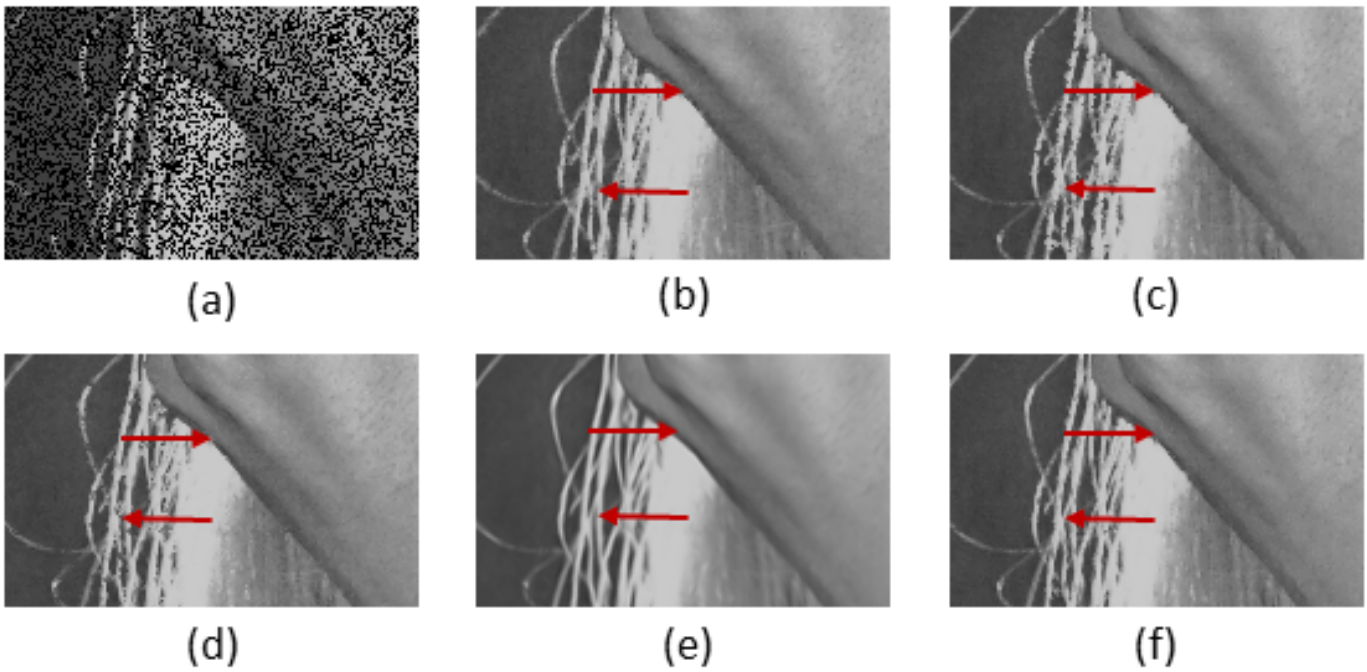

Fig. 7. Comparison of subjective quality for reconstructed 'Tiffney' image whose 40% of pixels were randomly removed. (a) Irregularly sampled image, restored image with (b) DPR [42] (32.203 dB), (c) EPA filter [43] (31.807 dB), (d) IFASDA [4] (31.503 dB), (e) Hybrid-Graph [5] (32.379 dB), and (f) the proposed algorithm (33.185 db).

- 7[1]. From these figures, it can be seen that the proposed algorithm significantly enhances the reconstruction accuracy (as highlighted by the red arrows in these figures). In a few situations, the proposed algorithm fails where either aliasing occurs [47] during down-sampling or edge information is missing in the reconstructed image in the first stage (as shown by blue arrows in Figs. 5 and 6). Moreover, edge structures in the output image of the proposed algorithm are more consistent as compared to other state-of-the-art algorithms.

In order to show the proposed algorithm is generic, we applied the proposed perceptually motivated MAP formulation on images in which distortion is caused due to the removal of Gaussian noise and compression (Jpeg-2000 and Jpeg). In Table IV, the proposed algorithm is objectively (in terms of $PNSR$) compared with the NLM filter [9], BM3D [44], Kernel regression [2], non-local median filter [10] and SAIF [12], for the removal of Gaussian noise. The proposed algorithm achieves the highest $PSNR$ (except for the BM3D [44] algorithm). The proposed algorithm achieves an improvement of 0.90 dB $PSNR$ as compared to the NLM filter [9] used in the first stage. In Table V, $PSNR$ of the Jpeg-2000 [38] and enhanced Jpeg-2000 image using the proposed algorithm is shown for different compression ratios. The proposed algorithm achieves a 0.226 dB better $PSNR$ than Jpeg-2000. The improvement can be attributed to the fact that Jpeg-2000 decoded images predominantly suffer from ringing artifacts in the vicinity of edges. In such a situation, the proposed algorithm is only effective in the vicinity of edges, as shown in Fig. 9. Consequently, $PSNR$ gain using the proposed algorithm is lower for Jpeg-2000 as compared to other applications. In Table VI, a comparison of Jpeg and enhanced Jpeg is made for varying quality factors and the proposed algorithm gains $PSNR$ by 0.718 dB. The results for the visual evaluation of these three applications are shown in figures 8 - 10. From these figures, it can be observed that the proposed algorithm produces visually appealing output images and efficiently suppress the visible artifacts, particularly in the vicinity of edges.

We have also compared the proposed algorithm with the state-of-the-art algorithms for deinterlacing, removal of Gaussian noise and suppressing (blocky and ringing) artifacts in decoded Jpeg images using two perceptual metrics namely, SSIM [31] and GMSD [1]. A higher value of SSIM and a lower value of GMSD reflects better perceptual reconstruction. Experimental results using these metrics for evaluating the reconstruction accuracy for the above-described three applications are reported in Tables VII - IX. From these Tables, it can be observed that the proposed algorithm achieves better SSIM and GMSD values as compared to existing algorithms, which suggests that the proposed algorithm produces perceptually better reconstructed output images. The proposed algorithm typically takes 5-15 seconds to restore the distorted image, in which 50% of the pixels are missing.

In Table X, we have shown the dependency of the proposed algorithm on reconstruction algorithms used in the first stage for interpolation purposes. From Table X, we can see two observations; 1) the proposed algorithm achieves nearly similar performance (in terms of SSIM and GMSD) when different interpolation algorithms (such as bicubic [24], directional bicubic [49], SAI [14], SME [15] and CG [16]) are used in the first stage and 2),the proposed algorithm is able to improve

[1]Full images are available at: http://ihome.ust.hk/ vjakhetiya/Reconstruction

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

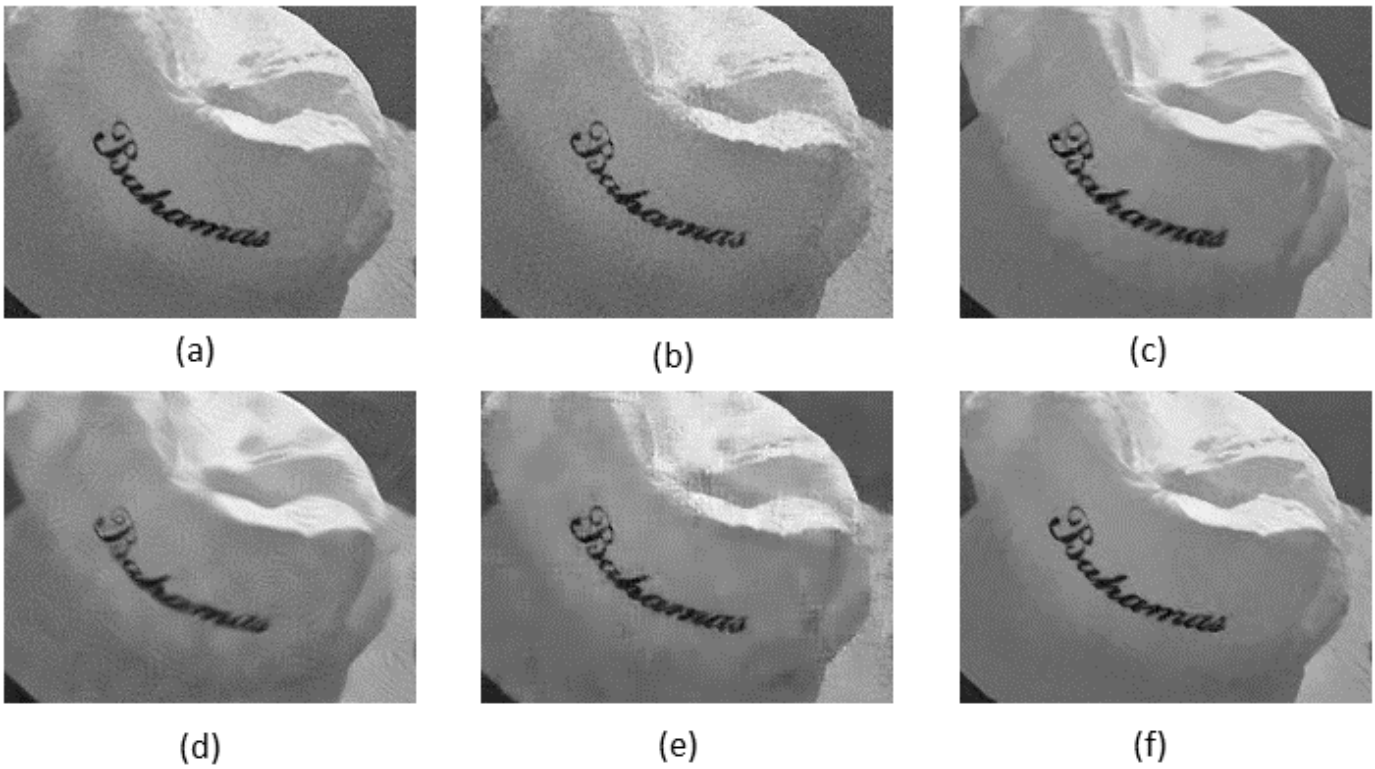

Fig. 8. Comparison of subjective quality for reconstructed 'Hat' image contaminated with Gaussian noise with noise level σ = 30. A Gaussian noise contaminated image, restored with (a) NLM filter [9] (29.166 dB), (b) Kernel regression [2] (28.773 dB), (c) BM3D [44] (30.440 dB), (d) NLEM filter [10] (28.943 dB), (e) SAIF [12] (29.703 dB), and (f) the proposed algorithm (30.093 dB).

TABLE V
COMPARISON OF RESTORATION RESULTS IN TERMS OF PSNR (dB) FOR JPEG-2000 (J2K) DECODED IMAGES. HERE 'CR' REFERS TO COMPRESSION RATIO.

	$PSNR$ (dB)				
Image	Computer	Building	Leaf	Monarch	Bush
J2K (CR=80)	27.974	26.611	23.770	27.718	28.977
Prop. (CR=80)	28.142	26.752	23.937	27.916	29.206
J2K (CR=40)	30.807	29.350	27.241	31.212	32.040
Prop. (CR=40)	30.995	29.510	27.540	31.485	32.333
J2K (CR=26.7)	33.140	31.429	29.650	33.801	34.279
Prop. (CR=26.7)	33.344	31.616	29.991	34.026	34.597

TABLE VI
COMPARISON OF RESTORATION RESULTS IN TERMS OF PSNR (dB) FOR JPEG DECODED IMAGES. HERE 'QF' REFERS TO QUALITY FACTOR.

	$PSNR$ (dB)				
Image	Computer	Building	Leaf	Monarch	Bush
Jpeg (QF=10)	29.267	28.771	28.247	29.155	30.371
Prop. (QF=10)	29.908	29.431	29.014	29.727	31.391
Jpeg (QF=20)	31.726	31.372	31.170	31.678	33.224
Prop. (QF=20)	32.380	31.998	32.029	32.256	34.253
Jpeg (QF=30)	33.093	32.798	32.819	33.051	34.727
Prop. (QF=30)	33.708	33.370	33.643	33.538	35.599

TABLE VII
SIMULATION RESULTS COMPARISON IN TERMS OF SSIM [31] (IN %) AND GMSD [1] (IN %) FOR INTRA-FRAME DEINTERLACING. THE BEST PERFORMANCE IS HIGHLIGHTED WITH BOLDFACE.

Image		[17]	[18]	[19]	[20]	Prop.
Tiffney	SSIM	98.08	98.06	98.20	98.28	**98.33**
	GMSD	1.897	1.907	1.816	1.792	**1.717**
Tire	SSIM	94.55	94.53	94.70	94.65	**94.79**
	GMSD	5.282	5.294	5.189	4.920	**4.783**
Building	SSIM	96.08	96.07	96.17	96.37	**96.38**
	GMSD	6.335	6.340	6.302	6.011	**5.716**
Kitchgrass	SSIM	90.45	90.53	90.86	90.8	**90.91**
	GMSD	5.068	5.079	5.0283	**4.762**	4.840
Sar-Large	SSIM	84.08	84.03	84.27	84.94	**85.26**
	GMSD	10.668	10.747	10.582	9.994	**9.987**
Bike	SSIM	90.38	90.37	90.47	91.00	**91.90**
	GMSD	11.753	11.826	11.753	11.553	**11.121**
Average	SSIM	92.272	92.265	92.444	92.673	**92.931**
	GMSD	6.833	6.865	6.778	6.497	**6.361**

state-of-the-art SAI [14], SME [15] and CG [16] algorithms in terms of both SSIM [31] and GMSD [1].

In existing IQA-metrics-based video coding algorithms [29], [35-37], IQA metrics are mainly used for the purpose of rate-distortion optimization. These algorithms produce visually appealing output frames but the $PSNR$ between the original and restored frames is lower (as compared to MSE-based optimization). In contrast to these algorithms, the proposed algorithm is a post-processing step and it can enhance both

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

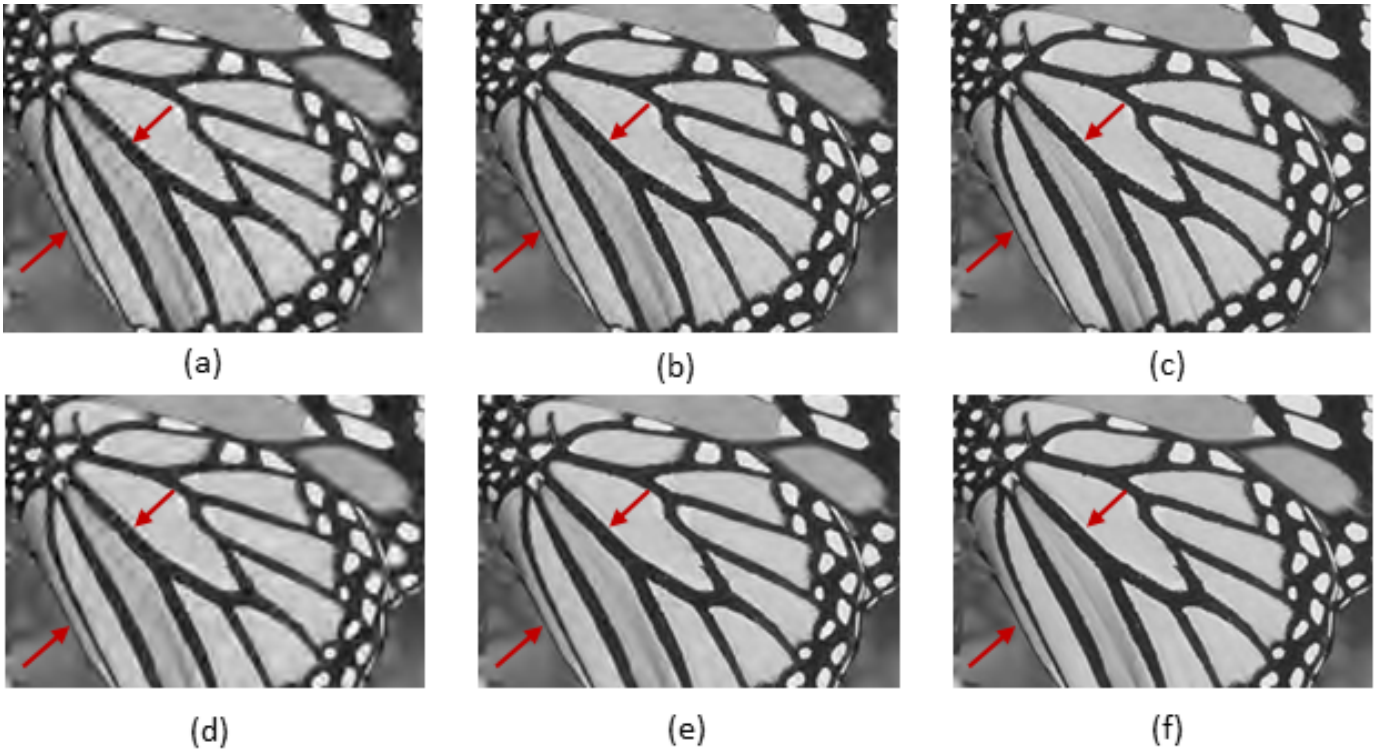

Fig. 9. Comparison of subjective quality for reconstructed 'Monarch' image, (a) Jpeg-2000 (compression ratio (CR) = 80, 27.718 dB), (b) Jpeg-2000 (CR = 40, 31.212 dB), (c) Jpeg-2000 (CR = 26.7, 33.801 dB), (d) proposed algorithm (CR = 80, 27.916 dB), (e) proposed algorithm (CR = 40, 31.485 dB), and (f) proposed algorithm (CR = 26.7, 34.026 dB).

TABLE VIII
COMPARISON OF RESTORATION RESULTS IN TERMS OF SSIM [31] (IN %) AND GMSD [1] (IN %) FOR IMAGE ADDED WITH GAUSSIAN NOISE. HERE 'BABOON20' REPRESENTS THE 'BABOON' IMAGE CONTAMINATED WITH GAUSSIAN NOISE WITH NOISE LEVEL σ = 20. THE BEST PERFORMANCE IS HIGHLIGHTED WITH BOLDFACE.

Image		[9]	[44]	[2]	[10]	[12]	Prop.
Baboon20	SSIM	89.17	88.97	88.48	74.85	**89.18**	89.02
	GMSD	8.11	8.21	7.73	15.23	7.64	**7.54**
Baboon30	SSIM	81.69	81.73	80.22	65.73	80.24	**82.22**
	GMSD	11.19	11.73	**10.83**	18.4	11.21	11.62
Baboon40	SSIM	74.61	73.81	72.76	61.07	71.03	**76.79**
	GMSD	13.34	14.9	**12.99**	19.93	14.63	13.25
Boat20	SSIM	89.96	89.81	90.76	75.61	**90.82**	89.71
	GMSD	8.11	8.19	**7.35**	16.20	7.40	7.64
Boat30	SSIM	83.39	83.13	**84.57**	69.21	83.98	83.68
	GMSD	11.13	11.83	**10.36**	18.47	10.86	11.59
Boat40	SSIM	76.75	76.63	**78.93**	65.18	76.96	77.93
	GMSD	13.57	14.92	**12.43**	19.63	14.01	13.99
Hat20	SSIM	88.06	**91.39**	81.36	87.93	88.27	90.91
	GMSD	6.39	**6.80**	10.57	8.77	7.16	7.05
Hat30	SSIM	80.02	**87.15**	73.79	82.01	84.12	86.72
	GMSD	8.73	9.38	14.18	11.4	9.42	**8.69**
Hat40	SSIM	71.99	**83.39**	68.10	76.01	80.64	81.52
	GMSD	11.16	11.00	16.48	11.87	11.33	**10.24**
Average	SSIM	81.74	84.00	79.88	73.06	82.80	**84.28**
	GMSD	10.19	10.77	11.43	15.54	10.40	**10.18**

TABLE IX
COMPARISON OF RESTORATION RESULTS IN TERMS OF SSIM [31] (IN %) AND GMSD [1] (IN %) FOR JPEG DECODED IMAGES.

Image		Computer	Building	Leaf	Monarch	Bush
Jpeg QF=10	SSIM	92.00	91.27	92.76	93.07	91.07
	GMSD	7.45	7.35	7.87	6.71	7.06
Prop. QF=10	SSIM	93.99	93.14	94.33	95.28	93.40
	GMSD	6.74	6.78	7.21	5.98	6.40
Jpeg QF=20	SSIM	96.48	96.12	96.60	96.49	96.04
	GMSD	3.35	3.46	3.74	3.13	3.21
Prop. QF=20	SSIM	97.43	96.93	97.32	97.78	97.16
	GMSD	2.96	3.21	3.35	2.63	2.90
Jpeg QF=30	SSIM	97.79	97.64	97.91	97.76	97.65
	GMSD	2.00	2.14	2.27	1.93	1.88
Prop. QF=30	SSIM	98.13	97.90	98.08	98.17	98.09
	GMSD	1.84	2.09	2.22	1.71	1.80

the visual quality and objective quality (in terms of $PSNR$). It also gives an indication that it is beneficial to use IQA metrics in the post-processing stage, instead of using them in the optimization process. We will investigate this issue in detail in our future work.

From above-described experimental results, it can be seen that the proposed algorithm is generic, since it works for diverse situations without the need for parameter tuning. The proposed algorithm also does not depend on the cause of the distortion and can efficiently reduce many types of artifacts, such as blurring, zig-zagging, clustering of edges, and ringing artifacts.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2016.2609419, IEEE Transactions on Multimedia

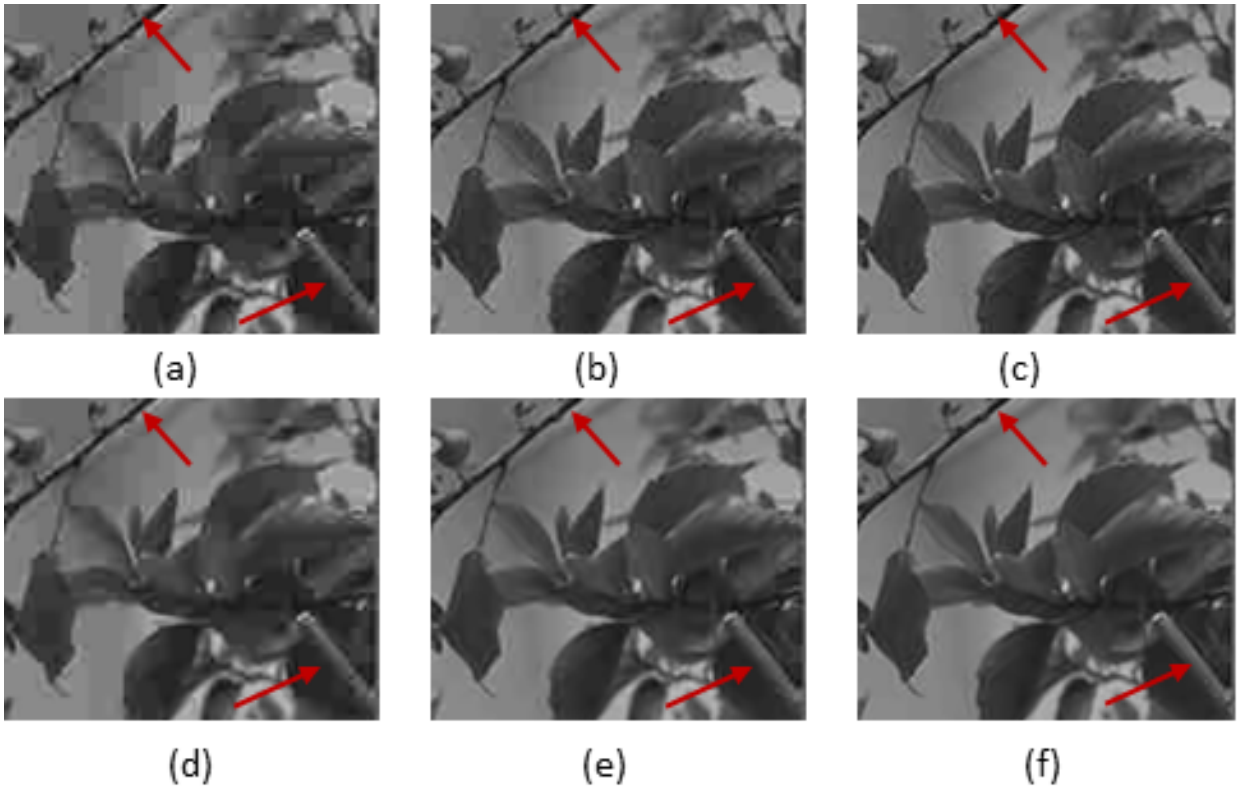

Fig. 10. Comparison of subjective quality for reconstructed 'Building' image, (a) Jpeg (quality factor (QF) = 10, 28.771 dB), (b) Jpeg (QF = 20, 31.372 dB) (c) Jpeg (QF = 30, 32.798 dB), (d) proposed algorithm (QF = 10, 29.431 dB), (e) proposed algorithm (QF = 20, 31.998 dB), and (f) proposed algorithm (QF = 30, 33.370 dB).

V. Conclusion

In this paper, we have proposed an efficient perceptually motivated and Maximum A Posterior (MAP)-based reconstruction algorithm. The proposed algorithm can be applied to several image processing applications (such as, denoising, deinterlacing, demosaicing, interpolation, removing compression noise in Jpeg-2000 and Jpeg decoded images, and restoring the lost pixels), where there is a need to enhance the reconstruction accuracy or to remove visible artifacts. In this work, we have utilized the fact that gradient magnitudes are insensitive to a moderate level of noise. With this view, we formulated the proposed algorithm as a MAP estimation and approximated the variance of the MAP using a gradient-magnitude-based Image Quality Assessment (IQA) metric. With the use of gradient-similarity-based IQA metric enabled the proposed algorithm to better approximate the variance of the MAP, when compared to commonly used mean-square-error-based bilateral and non-local means filters. It is especially true when neighboring pixels have an enormous amount of distortion and the level and cause of distortion to the signal is unknown. Results from the experiments showed that the proposed algorithm produces better reconstruction quality (both visually and objectively), when compared to the state-of-the-art algorithms, and more importantly, without having to manually tune any parameters for different image processing applications.

References

[1] W. Xue, L. Zhang, X. Mou and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. on Image Processing,* pp. 684-695, 2014.

[2] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. on Image Processing,* pp. 349-366, 2007.

[3] T. Chen and H. R. Wu, "Adaptive impulse detection using center-weighted median filters," *IEEE Signal Proc. Letters,* pp. 1-3, 2001.

[4] J. Cai, R. H. Chan, and M. Nikolova, "Fast two-phase image deblurring under impulse noise," *Journal Math. Imaging Vis.,* pp. 46-53, 2010.

[5] X. Liu et al., "Progressive image denoising through hybrid graph laplacian regularization: A unified framework," *IEEE Trans. on Image Processing,* pp. 1491-1503, 2014.

[6] B. Xiong and Z. Yin, "A universal denoising framework with a new impulse detector and nonlocal means," *IEEE Trans. on Image Processing,* pp. 1663-1675, 2012.

[7] L. C. Hsing, T. J. Shiuan and C. C. Te, "Switching bilateral filter with a texture/ noise detector for universal noise detector," *IEEE Trans. on Image Processing,* pp. 2307-2320, 2010.

[8] C. Tomasi and R. Manduchi, "Bilateral filter for color and gray images," in *Proc. of Int. Conf. Computer Vision (ICCV),* pp. 839-846, 1998.

[9] A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," in *Proc. of Computer Vision and Pattern Recognition, 2005 (CVPR),* pp. 60-65, 2005.

[10] K. N. Chaudhury and A. Singer, "Non-Local Euclidean Medians," *IEEE Signal Proc. Letters,* pp. 745-478, 2012.

[11] D. A. Huang et al., "Self-learning based image decomposition with applications to single image denoising," *IEEE Trans. on Image Multimedia,* pp. 83-93, 2014.

[12] H. Talebi, X. Zhu and P. Milanfar, "How to SAIF-ly boost denoising performance," *IEEE Trans. on Image Processing,* pp. 1470-1485, 2013.

[13] H. Bhujle and S. Chaudhuri, "Novel speed-up strategies for non-local means denoising with patch and edge patch based dictionaries," *IEEE Trans. on Image Processing,* pp. 356-365, 2014.

TABLE X
SIMULATION RESULTS COMPARISON IN TERMS OF SSIM [31] (IN %) AND GMSD [1] (IN %) FOR INTERPOLATION AND DEPENDENCY OF THE PROPOSED ALGORITHM ON RECONSTRUCTION ALGORITHM USED IN THE FIRST PASS. THE BEST PERFORMANCE IS HIGHLIGHTED WITH BOLDFACE.

Image		[24]	[49]	[14]	[15]	[16]	[24]+ Prop.	[49] + Prop.	[14]+ Prop.	[15] + Prop.	[16] +Prop.
Tiffney	SSIM	84.47	84.71	85.80	85.75	84.71	85.66	84.91	85.93	86.00	85.93
	GMSD	5.24	11.75	5.59	4.84	4.76	5.03	11.69	5.46	4.66	4.52
Baboon	SSIM	72.09	71.86	73.08	73.02	72.50	72.18	71.04	71.92	71.82	71.82
	GMSD	9.88	12.61	9.76	9.76	9.98	9.60	12.43	9.61	9.60	9.67
Sar-Large	SSIM	50.80	52.23	52.80	51.55	51.74	53.92	54.49	54.55	54.27	54.33
	GMSD	16.48	17.07	16.24	16.29	16.60	14.64	15.32	14.51	14.49	14.74
Aerial2	SSIM	65.97	66.24	66.75	66.54	66.34	66.83	66.40	66.82	66.79	66.94
	GMSD	10.84	12.37	10.70	10.73	10.69	10.32	11.98	10.31	10.29	10.20
Ruler.512	SSIM	68.42	69.06	66.79	61.49	69.77	69.66	69.23	67.60	62.31	69.93
	GMSD	38.71	41.22	39.67	38.73	40.64	39.26	41.05	39.87	38.94	40.54
Party	SSIM	60.58	61.28	61.59	60.84	61.04	62.29	62.09	62.35	61.93	62.23
	GMSD	15.95	16.86	15.77	15.93	16.03	15.53	16.52	15.49	15.58	15.62
Bike	SSIM	67.93	68.47	69.48	69.65	69.58	70.89	70.04	71.04	71.23	71.18
	GMSD	15.07	18.49	14.79	14.74	16.24	14.38	18.05	14.25	14.18	14.46
Average	SSIM	**67.18**	**67.69**	**68.04**	**66.98**	**68.10**	**68.78**	**68.31**	**68.60**	**67.76**	**68.91**
	GMSD	**16.02**	**18.63**	**16.08**	**15.86**	**16.24**	**15.54**	**18.15**	**15.64**	**15.39**	**15.68**

[14] X. Zhang and X. Wu, "Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation," *IEEE Trans. on Image Processing,* pp. 887-896, 2008.

[15] S. Mallat and G. Yu, "Super-resolution with sparse mixing estimators," *IEEE Trans. on Image Processing,* pp. 2889-2900, 2010.

[16] Z. Wei and K. K. Ma, "Contrast-guided image interpolation," *IEEE Tans. on Image Processing,* pp. 4271-4285, 2013.

[17] W. Kim, S. Jin and J. Jeong, "Novel intra deinterlacing algorithm using content adaptive interpolation," *IEEE Trans. on Cons. Elec.,* 53(3), pp. 1036-1043, 2007.

[18] P. Yin et al., "A low-complexity interpolation method for deinterlacing," in *IEICE Trans. on Information and Systems,* pp. 606-608, 2007.

[19] J. Wang, G. Jeon and J. Jeong, "Efficient adaptive deinterlacing algorithm with awareness of closeness and similarity," *Journal of Optical Engineering,* 51(1), pp. 606-608, 2012.

[20] V. Jakhetiya et al. "Fast and efficient intra-frame deinterlacing using observation model based bilateral filter," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014* pp. 5819 - 5823.

[21] J. Wang, G. Jeon and J. Jeong, "Deinterlacing algorithm using weighted least squares," *IEEE Trans. on circuit and systems for video technology,* pp. 39-48, 2014.

[22] J. Wu et al. "Just noticeable difference estimation for images with free-energy principle," in *IEEE Trans. on Multimedia,* 1705-1710, 2013.

[23] K. W. Hung and W. C. Siu, "Robust soft-decision interpolation using weighted least squares," *IEEE Trans. on Image Processing,* pp. 1061-1069, 2012.

[24] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. on Acoustics, Speech and Signal Processing,* pp. 1153-1160, 1981.

[25] Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it? A new look at signal fidelity measures,". *IEEE Signal Processing Magazine,* pp. 98-117, 2009.

[26] K. Gu et al., "Using free energy principle for blind image quality assessment," *IEEE Trans. on Multimedia,* pp. 50-63, 2015.

[27] D. Lu et al., "Exploiting entropy masking in perceptual graphic rendering," *Signal Processing: Image Communication,* pp. 1-13, 2015.

[28] T. Thaipanic et al., "Improved image denoising with adaptive nonlocal means (ANL-means) algorithm," *IEEE Trans. on Consumer Electronics,* pp. 2623-2630, 2010.

[29] H. L. Tan et al., "A perceptually relevant MSE-based image quality metric," *IEEE Trans. on Image Processing,* pp. 4447-4459, 2013.

[30] W. Lin and C. C. J. Kuo, "Perceptual visual quality metrics: A survey," *Journal of Visual Communication and Image Representation,* pp. 297-312, 2011.

[31] Z. Wang et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Processing,* pp. 600-612, 2004.

[32] M. Narwaria et al., "SVD-based quality metric for image and video using machine learning," *IEEE Trans. on Systems, Man, and Cybernetics, Part B,* pp. 347-364, 2012.

[33] J. Wu et al., "Perceptual quality metric with internal generative mechanism," *IEEE Trans. on Image Processing,* pp. 43-54, 2013.

[34] A. Rehman et al. "SSIM-based non-local means image denoising," in *Proc. IEEE International Conference on Image Processing (ICIP), 2011*.

[35] S. Wang et al., "SSIM-motivated rate-distortion optimization for video coding," *IEEE Trans. on circuit and systems for video technology,* pp. 516-529, 2012.

[36] C. Yeo et al., "On rate distortion optimization using SSIM," *IEEE Trans. on circuit and systems for video technology,* pp. 1170-1181, 2013.

[37] D. Wei et al., "SSIM-based rate-distortion optimization in H.264," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014* pp. 7343-7347.

[38] A. Skodras et al., "The JPEG 2000 still image compression standard,". *IEEE Signal Processing Magazine,* pp. 36-58, 2001.

[39] J. Xie et al., "Joint super resolution and denoising from a single depth image," *IEEE Trans. on Multimedia,* pp. 1525-1537, 2015..

[40] L. W. Kang et al., "Learning-based joint super-resolution and deblocking for a highly compressed image," *IEEE Trans. on Multimedia,* pp. 921-934, 2015.

[41] K. W. Hung and W. C. Siu, "Fast image interpolation using bilateral filter." *IET Image Processing,* pp. 877-890, 2012.

[42] R. H. Chan, C. W. Ho and M. Nikolova, 'Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization," *IEEE Trans. Image Processing*, pp. 1479-1485, 2005.

[43] P. Y. Chen and C. Y Lien, "An efficient edge-preserving algorithm for removal of salt and pepper noise," *IEEE Signal Processing Letters*, pp. 833-836, 2008.

[44] K. Dabov et al, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Processing*, pp. 2080-2095, 2007.

[45] X. Zhang et al., "Compression noise estimation and reduction via patch clustering," in *Proc. Proceedings of APSIPA Annual Summit and Conference, 2015*.

[46] X. Zhang et al., "Compression artifact reduction by overlapped-block transform coefficient estimation with block similarity," *IEEE Trans. Image Processing*, pp. 4613-4626, 2013.

[47] F. Lu et al., "Joint demosaicing and subpixel-based down-sampling for Bayer images: A fast frequency-domain analysis approach," *IEEE Trans. Multimedia*, pp. 1359-1369, 2012.

[48] A. Liu et al., "Image quality assessment based on gradient similarity," *IEEE Trans. Image Processing*, pp. 1500-1512, 2012.

[49] D. Zhou, X. Shen and W. Dong, "Image zooming using directional cubic convolution interpolation," *IET Image Processing,* pp. 627-635, 2012.

[50] M. Aharon et al., "K-SVD: An Algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Singal Processing*, pp. 4311-4322, 2006.