A Bayesian Hyperprior Approach for Joint Image Denoising and Interpolation, with an Application to HDR Imaging

Cecilia Aguerrebere, Andrés Almansa, Julie Delon, Yann Gousseau, Pablo Musé

A Bayesian Hyperprior Approach for Joint Image Denoising and Interpolation, with an Application to HDR Imaging

Cecilia Aguerrebere, Andrés Almansa, Julie Delon, Yann Gousseau and Pablo Musé

***Abstract*—Bayesian approaches assuming Gaussian models for the image patches have recently achieved impressive denoising performance. First, fixed models were used for all patches and then, performance was significantly improved by using per-patch models. Unfortunately, the local estimation implied by such an approach is particularly unstable for most inverse problems beyond denoising. In this work, we advocate the use of a hyperprior to model image patches, in order to stabilize the estimation procedure. Two characteristics of the proposed restoration scheme are: first that it is adapted to diagonal degradation matrices, and in particular to missing data problems (e.g. inpainting of missing pixels or zooming), and second that it can deal with generic noise models, particularly suited to digital cameras. As such, the scheme is especially adapted to computational photography purposes. In order to illustrate this point, we provide in the experimental section an application to single image high dynamic range imaging showing the effectiveness of the proposed scheme.**

***Index Terms*—Non-local patch-based restoration, Bayesian restoration, Maximum a Posteriori, Gaussian Mixture Models, hyper-prior, conjugate distributions, high dynamic range imaging, single shot HDR.**

I. Introduction

DIGITAL images are subject to a wide variety of degradations, which in most cases can be modeled as

$$\mathrm{Z} = \boldsymbol{D}\mathrm{C} + \mathrm{N}, \tag{1}$$

where Z is the observation, $\boldsymbol{D}$ is the degradation operator, C is the underlying ground-truth image and N is additive noise. Different settings of the degradation matrix $\boldsymbol{D}$ model different problems such as zooming, deblurring or random missing pixels. Different characterizations of the noise term N describe noise degradations, ranging from the classical additive Gaussian noise to more complicated and realistic models such as multiplicative or signal dependent noise. These degradations are often combined in practice. For instance, raw images captured with regular digital cameras combine signal dependent noise, limited spatial resolution and limited dynamic range, among other problems [1].

C. Aguerrebere is with the Department of Electrical and Computer Engineering, Duke University, Durham NC 27708, US (e-mail: cecilia.aguerrebere@duke.edu)

A. Almansa and Y. Gousseau are with the LTCI CNRS, Télécom ParisTech, Université Paris-Saclay, 75013 Paris, France (e-mail: gousseau,almansa@telecom-paristech.fr).

J. Delon is with MAP5 (CNRS UMR 8145), Université Paris Descartes, 75270 Paris Cedex 06 (e-mail: julie.delon@parisdescartes.fr)

P. Musé is with the Department of Electrical Engineering, Universidad de la República, 11300 Montevideo, Uruguay (e-mail: pmuse@fing.edu.uy)

Due to the inherent ill-posedness of such inverse problems, the standard approach has been to impose some prior on the image, let it be in a variational or a Bayesian approach. Popular image models have been proposed through the total variation [2], wavelet decompositions [3] or the sparsity of image patches [4]. Inspired by the patch-based approach for texture synthesis proposed by Efros and Leung [5], Buades et al. [6] introduced the use of patches and the self-similarity hypothesis to the denoising problem leading to a new era of patch-based image restoration techniques. This hypothesis can actually be interpreted as a non-local prior [7]. A major step forward in fully exploiting the potential of patches was taken with the introduction of patch prior models. Recent state-of-the-art methods make use of patch models in a Bayesian framework to restore degraded images. Some of them are devoted to the denoising problem [8], [9], [10], [11], while others propose a more general framework for the solution of image inverse problems [12], [13], including for instance inpainting, deblurring and zooming. The work by Lebrun et al. [14], [10] presents a thorough analysis of several recent restoration methods, revealing their common roots and their relationship with the Bayesian approach.

Among the state-of-the-art restoration methods, two noticeable approaches are the patch-based Bayesian approach by Yu et al. [13], namely the piece-wise linear estimators (PLE), and the non-local Bayes (NLB) algorithm by Lebrun et al. [10]. PLE is a general framework for the solution of image inverse problems under Model (1), while NLB is a denoising method ($\boldsymbol{D} = Id$). Both methods use a Gaussian patch prior learnt from image patches through iterative procedures. In the case of PLE, patches are modeled according to a Gaussian Mixture Model (GMM), with a relatively small number of classes (19 in all their experiments), whose parameters are learnt from all image patches[1]. In the case of NLB, each patch is associated with a single Gaussian model, whose parameters are computed from similar patches chosen from a local neighborhood, i.e., a search window centered at the patch. We refer hereafter to this kind of per-patch modelling as *local*. The locality here refers to the fact that each patch has its own model, although this model is learnt in a non-local way. In contrast, PLE uses a fixed number of classes to represent all patches.

NLB outperforms PLE in the denoising task [15], mostly due to its local model estimation. Nevertheless, PLE obtains

[1]Actually, the authors report the use of 128×128 image sub-regions in their experiments, so we may consider PLE as a semi-local approach.

state-of-the-art results in other applications such as interpolation of missing pixels, deblurring and zooming. In particular, PLE yields very good results for the interpolation of random missing pixels with high masking rates. A variant of PLE for inpainting is proposed by Wang [16] (E-PLE), using a GMM initialized from natural images instead of using synthetic images of edges as it is done in PLE.

Zoran and Weiss [12] (EPLL) follow a similar approach, but instead of iteratively updating the GMM from image patches, they use a larger number of classes (200) that are fixed and learnt from a large database of natural image patches (2×10^6 patches). Wang and Morel [11] claim that, in the case of denoising, it is better to have fewer models that are updated with the image patches (as in PLE) than having a large number of fixed models (as in EPLL). Moreover, unlike the previous methods, EPLL restores image patches according to the GMM prior[2] while keeping the restored image close to the corrupted image for a given corruption model.

All of the previous restoration approaches share a common Bayesian framework based on Gaussian patch priors. As said above, relying on local priors [10], [11] has proven more accurate for the task of image denoising than relying on a mixture of a limited number of Gaussian models [12], [13]. However, such local priors are very difficult to estimate for more general restoration problems, especially when the image degradations involve missing pixels.

The main contribution of this work is twofold. First, we propose a robust framework enabling the use of such Gaussian local priors for solving more general restoration problems, by drawing on what is known in Bayesian statistics as a hyperprior. In particular, we show that the approach is well suited to problems involving a diagonal degradation operator, such as the interpolation of missing pixels, zooming or single image HDR image generation. An important restoration problem that is not addressed by our approach is image deblurring, since the discrete linear operator in the image formation model is no longer diagonal. Yu et al. [13] claim that PLE is capable of deblurring images contaminated with noise, however the results presented in their work are far less successful than those obtained for zooming or interpolation. When the blurring kernel support is significantly smaller that the patch size, PLE manages to reduce blur, but no theoretical analysis stating precisely under which conditions it is possible to obtain satisfactory results is provided. The experiments supporting this claim are not extensive either.

The second advantage of the proposed framework is its ability to deal with signal dependent noise, therefore making it adapted to realistic digital photography applications.

We propose to model image patches according to a Gaussian prior, whose parameters will be estimated locally from similar patches. The main challenge of this framework is to estimate the Gaussian parameters, i.e. the mean μ and the covariance matrix $\mathbf{\Sigma}$, from a set of patches with potentially high degradation levels. For example, in the case of the interpolation of random missing pixels with a high masking rate, the estimation problem becomes very ill-posed. In order to tackle this problem, we include prior knowledge on the model parameters making use of a hyperprior, i.e. a probability distribution on the parameters of the prior. In Bayesian statistics, μ and $\mathbf{\Sigma}$ are known as hyperparameters, since they are the parameters of a prior distribution, while the prior on them is called a hyperprior. Such a framework is often called hierarchical Bayesian modeling [17] and has been applied in particular to image unmixing problems [18]. The use of a hyperprior allows for a joint maximum a posteriori estimation of the unknown patches, μ and $\mathbf{\Sigma}$ from similar patches even if they present high degradation levels. The information provided by the hyperprior compensates for the patches missing information.

[2]EPLL does not impose a given prior, GMM is an option among others.

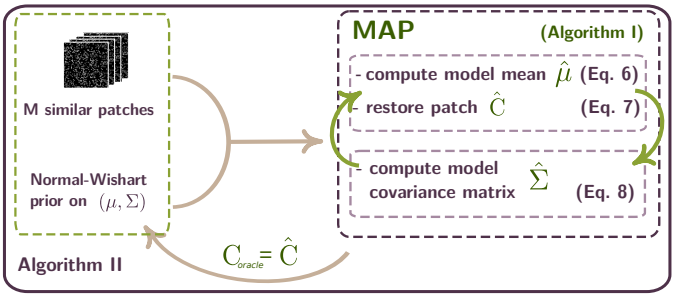


Fig. 1. Diagram of the proposed iterative approach.

Experiments conducted on both synthetic and real data show state-of-the-art results obtained by the proposed approach in various problems such as interpolation, denoising and zooming.

This general framework is also applied to the generation of high dynamic range (HDR) images. Due to physical limitations, current digital sensors cannot capture faithful representations of HDR scenes. HDR imaging tackles this problem and seeks to accurately capture and represent scenes with the largest possible irradiance range. As will be detailed in Section V, HDR imaging can be achieved from a single snapshot using specially modified sensors. We propose here a novel approach to single shot HDR imaging using the proposed restoration framework. This simple, yet powerful approach shows excellent performance in several examples both on synthetic and real data.

The article is organized as follows. Section II introduces the proposed approach while Section III presents the main implementation aspects. Supportive experiments are presented in Section IV. Section V is devoted to the application of the proposed framework to the HDR imaging problem. Last, conclusions are summarized in Section VI.

II. Hyperprior Bayesian Estimator

We describe here the proposed restoration method, called Hyperprior Bayesian Estimator (HBE). Following the idea of the recent Bayesian approach [10], we assume a Gaussian prior for image patches whose parameters μ and $\mathbf{\Sigma}$ are estimated from a group of patches similar to the current patch. The particularity of our method is that it uses a joint maximum a posteriori formulation to estimate both the image patches and the parameters μ and $\mathbf{\Sigma}$, thanks to a Bayesian hyperprior model on these parameters. This hyperprior permits

to stabilize the local estimation of the Gaussian statistics. As a consequence, we can exploit the accuracy of local model estimation, such as the one used in denoising by NLB, in general restoration problems.

A. Patch degradation model

The observed image z is decomposed into I overlapping patches $\{\mathrm{z}_i\}_{i=1,\dots,I}$ of size $\sqrt{n}\times\sqrt{n}$. Each patch $\mathrm{z}_i \in \mathbb{R}^{n\times 1}$ is considered to be a realization of the random variable Z_i given by

$$\mathrm{Z}_i = \boldsymbol{D}_i \mathrm{C}_i + \mathrm{N}_i, \tag{2}$$

where $\boldsymbol{D}_i \in \mathbb{R}^{n\times n}$ is a degradation operator, $\mathrm{C}_i \in \mathbb{R}^{n\times 1}$ is the original patch we seek to estimate and $\mathrm{N}_i \in \mathbb{R}^{n\times 1}$ is an additive noise term, not necessarily identically distributed as we will see.

In this paper, we are interested in different application scenarios, in which the degradation operator $\boldsymbol{D}_i$ (supposed to be known) can represent resolution change (zooming) or missing pixels (inpainting). As for the noise term N_i, we assume that it is well modeled by a Gaussian distribution $\mathrm{N}_i \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{N_i})$. Therefore, the distribution of Z_i given C_i can be written as

$$\begin{aligned} &p(\mathrm{Z}_i \mid \mathrm{C}_i) \sim \mathcal{N}(\boldsymbol{D}_i\mathrm{C}_i, \boldsymbol{\Sigma}_{N_i}) \\ &\propto |\boldsymbol{\Sigma}_{N_i}^{-1}|^{\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathrm{Z}_i - \boldsymbol{D}_i\mathrm{C}_i)^T \boldsymbol{\Sigma}_{N_i}^{-1}(\mathrm{Z}_i - \boldsymbol{D}_i\mathrm{C}_i)\right). \end{aligned} \tag{3}$$

In this noise model, the matrix $\boldsymbol{\Sigma}_{N_i}$ is only assumed to be diagonal (the noise is uncorrelated). It can represent a constant variance, spatially variable variances or even variances dependent on the pixel value (to approximate Poisson noise).

This degradation model is deliberately generic. We will see in Section IV that keeping a broad noise model is essential to properly tackle the problem of HDR imaging from a single image. It also includes a wide range of restoration problems generally studied in the literature, such as the particular case of multiplicative noise.

B. Joint Maximum A Posteriori

Following the direction of Bayesian patch-based methods, we assume a Gaussian prior for each patch, with unknown mean μ and covariance matrix $\boldsymbol{\Sigma}$

$$p(\mathrm{C}_i \mid \mu, \boldsymbol{\Sigma}) \sim \mathcal{N}(\mu, \boldsymbol{\Sigma}).$$

In the literature, μ and $\boldsymbol{\Sigma}$ are either estimated locally from a set of patches similar to C [10] , or chosen from a finite set of precomputed parameters that are learned globally [13]. The first solution is very accurate for Gaussian denoising, but not reliable when pixels are missing. The second one is more robust but yields smoother results.

The proposed approach uses a **joint Maximum a Posteriori** (MAP) in order to estimate both the patch C and its Gaussian parameters from a set of similar patches. To this aim, we make use of a hyperprior on μ and $\boldsymbol{\Sigma}$. On the one hand, using similar patches gives a spatially adaptive or local characterization of the patch [10]. On the other hand, including the hyperprior makes the parameter estimation more stable, which is critical when few similar patches are available or when some pixels are unknown (e.g. for interpolation and zooming).

To simplify calculations, we work with the precision matrix $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}^{-1}$ instead of the covariance matrix $\boldsymbol{\Sigma}$. As it is usual when considering hyperpriors, we assume that the parameters μ and $\boldsymbol{\Sigma}$ follow a conjugate distribution. In our case, that boils down to assuming a Normal-Wishart[3] prior for the couple $(\mu, \boldsymbol{\Lambda})$,

$$\begin{aligned} p(\mu, \boldsymbol{\Lambda}) &= \mathcal{N}(\mu|\mu_0, (\kappa\boldsymbol{\Lambda})^{-1})\mathcal{W}(\boldsymbol{\Lambda}|(\nu\boldsymbol{\Sigma}_0)^{-1}, \nu) \\ &\propto |\boldsymbol{\Lambda}|^{1/2} \exp\left(-\frac{\kappa}{2}(\mu-\mu_0)\boldsymbol{\Lambda}(\mu-\mu_0)^T\right) \\ &\quad |\boldsymbol{\Lambda}|^{(\nu-n-1)/2} \exp\left(-\frac{1}{2}\mathrm{tr}(\nu\boldsymbol{\Sigma}_0\boldsymbol{\Lambda})\right), \end{aligned} \tag{4}$$

with parameters μ_0, $\boldsymbol{\Sigma}_0$, scale parameter $\kappa > 0$ and $\nu > n-1$ degrees of freedom.

Now, assume that we observe a group $\{\mathrm{Z}_i\}_{i=1,\dots,M}$ of similar patches and that we want to recover the restored patches $\{\mathrm{C}_i\}_{i=1,\dots,M}$. If these unknown $\{\mathrm{C}_i\}$ are independent[4] and follow the same Gaussian model, we can compute the joint maximum a posteriori

$$\begin{aligned} &\underset{\{\mathrm{C}_i\},\mu,\boldsymbol{\Lambda}}{\arg\max}\ p(\{\mathrm{C}_i\}, \mu, \boldsymbol{\Lambda} \mid \{\mathrm{Z}_i\}) = \\ &= p(\{\mathrm{Z}_i\} \mid \{\mathrm{C}_i\}, \mu, \boldsymbol{\Lambda})\ p(\{\mathrm{C}_i\} \mid \mu, \boldsymbol{\Lambda})\ p(\mu, \boldsymbol{\Lambda}) \\ &= p(\{\mathrm{Z}_i\} \mid \{\mathrm{C}_i\})\ p(\{\mathrm{C}_i\} \mid \mu, \boldsymbol{\Lambda})\ p(\mu, \boldsymbol{\Lambda}). \end{aligned} \tag{5}$$

In this product, the first term is given by the noise model (3), the second one is the Gaussian prior on the set of patches $\{\mathrm{C}_i\}$ and the third one is the hyperprior (4). In the last equality we omitted the explicit dependence on μ and $\boldsymbol{\Lambda}$ in $p(\{\mathrm{Z}_i\} \mid \{\mathrm{C}_i\}, \mu, \boldsymbol{\Lambda})$, since these parameters are completely determined by the set $\{\mathrm{C}_i\}$.

C. Optimality conditions

Computing the joint maximum a posteriori is equivalent to minimize

$$\begin{aligned} f(\{\mathrm{C}_i\}, \mu, \boldsymbol{\Lambda}) &:= -\log p(\{\mathrm{C}_i\}, \mu, \boldsymbol{\Lambda} \mid \{\mathrm{Z}_i\}) \\ &= \frac{1}{2}(\mathrm{Z}_i - \boldsymbol{D}_i\mathrm{C}_i)^T\boldsymbol{\Sigma}_{N_i}^{-1}(\mathrm{Z}_i - \boldsymbol{D}_i\mathrm{C}_i) \\ &- \frac{\nu-n+M}{2}\log|\boldsymbol{\Lambda}| + \frac{1}{2}\sum_{i=1}^{M}(\mathrm{C}_i-\mu)^T\boldsymbol{\Lambda}(\mathrm{C}_i-\mu) \\ &+ \frac{\kappa}{2}(\mu-\mu_0)^T\boldsymbol{\Lambda}(\mu-\mu_0) + \frac{1}{2}\mathrm{trace}[\nu\boldsymbol{\Sigma}_0\boldsymbol{\Lambda}], \end{aligned}$$

over the set $\mathbb{R}^{nM}\times\mathbb{R}^n\times S_n^{++}(\mathbb{R})$, with $S_n^{++}(\mathbb{R})$ the set of real symmetric positive definite matrices of size n.

The function f is biconvex respectively in the variables $(\{\mathrm{C}_i\}, \mu)$ and $\boldsymbol{\Sigma}$, which means that it is convex in each of these two variables when the other is fixed. To minimize this

[3]The Normal-Wishart distribution is the conjugate prior of a multivariate normal distribution with unknown mean and covariance matrix. $\mathcal{W}$ denotes the Wishart distribution [19].

[4]We rely here on the classical independence assumption made in the patch-based literature, even if this assumption is false if some of the patches $\{\mathrm{C}_i\}$ are overlapping in the image.

energy for a given set of hyper-parameters $(\mu_0, \mathbf{\Lambda}_0)$, we will use an alternate convex minimization. At each iteration, f is first minimized with respect to $(\{\mathrm{C}_i\}, \mu)$ with $\mathbf{\Lambda}$ fixed and then with respect to $\mathbf{\Lambda}$ with $(\{\mathrm{C}_i\}, \mu)$ fixed.

Differentiating f with respect to each variable, we get explicit optimality equations for the minimization scheme. The proofs of the following propositions are available in Appendix A.

Proposition 1. *Assume that $\mathbf{\Lambda}$ is fixed and that the covariance $\mathbf{\Sigma}_{N_i}$ does not depend on the $\{\mathrm{C}_i\}$. The function $(\{\mathrm{C}_i\}, \mu) \mapsto f(\{\mathrm{C}_i\}, \mu, \mathbf{\Lambda})$ is convex on $\mathbb{R}^{n(M+1)}$ and attains its minimum at $(\{\hat{\mathrm{C}}_i\}, \hat{\mu})$, given by*

$$\widehat{\mu} = \left(\kappa \mathrm{Id} + \sum_{i=1}^{M} \mathbf{A}_i \boldsymbol{D}_i\right)^{-1} \left(\sum_{i=1}^{M} \mathbf{A}_i \mathrm{Z}_i + \kappa \mu_0\right). \quad (6)$$

$$\widehat{\mathrm{C}_i} = \mathbf{A}_i(\mathrm{Z}_i - \boldsymbol{D}_i\hat{\mu}) + \hat{\mu}, \qquad \forall i \in \{1, \dots M\} \quad (7)$$

with $\mathbf{A}_i = \mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T + \mathbf{\Sigma}_{N_i})^{-1}$.

Proposition 2. *Assume that the variables $(\{\mathrm{C}_i\}, \mu)$ are fixed. The function $\mathbf{\Lambda} \to f(\{\mathrm{C}_i\}, \mu, \mathbf{\Lambda})$ is convex on $S_n^{++}(\mathbb{R})$ and attains its minimum at $\hat{\mathbf{\Lambda}}$ such that*

$$\hat{\mathbf{\Lambda}}^{-1} = \frac{\nu\mathbf{\Sigma}_0 + \kappa(\mu - \mu_0)(\mu - \mu_0)^T + \sum_{i=1}^{M}(\mathrm{C}_i - \mu)(\mathrm{C}_i - \mu)^T}{\nu + M - n}. \quad (8)$$

The expression of $\widehat{\mathrm{C}_i}$ in (7) is obtained under the hypothesis that the noise covariance matrix $\mathbf{\Sigma}_{N_i}$ does not depend on C_i. Observe that it can be shown (see Appendix B) that under the somewhat weaker hypothesis that the noise N_i and the signal C_i are uncorrelated, this estimator is also the affine estimator $\tilde{\mathrm{C}}_i$ that minimizes the Bayes risk $\mathbb{E}[(\tilde{\mathrm{C}}_i - \mathrm{C}_i)^2]$. The uncorrelatedness of N_i and C_i is a quite reasonnable hypothesis in practice: this is for instance the case if the noise can be written as $N_i = f(C_i)\varepsilon_i$ where ε_i is independent of C_i. This includes very different noise models, including a reasonable approximation of the acquisition by a camera sensor in which the noise variance is an affine function of the signal [20].

From (6), we find that the MAP estimator of μ is a weighted average of two terms: the mean estimated from the similar restored patches and the prior μ_0. The parameter κ controls the confidence level we have on the prior μ_0. With the same idea, we observe that the MAP estimator for $\mathbf{\Lambda}$ is a combination of the prior $\mathbf{\Lambda}_0$ on $\mathbf{\Lambda}$, the covariance imposed by μ and the covariance matrix estimated from the patches C_i.

D. Alternate convex minimization of f

The previous propositions imply that we can derive a fully explicit alternate convex minimization for f, presented in Algorithm 1. Starting with a given value of $\mathbf{\Lambda}$, at each step, μ^l and C^l are computed according to Equations (6) and (7), then $\mathbf{\Lambda}^l$ is updated according to (8).

We show in Appendix A the following convergence result for the previous algorithm. The proof adapts the arguments in [21] to our particular case.

Algorithm 1: Alternate convex minimization for f

Input: Z, $\boldsymbol{D}$, $\mu_0, \mathbf{\Sigma}_0, \kappa, \nu$
Output: $\{\hat{\mathrm{C}}_i\}$,$\hat{\mu}$, $\hat{\mathbf{\Lambda}}$,
Initialization: Set $\mathbf{\Lambda}^0 = \mathbf{\Sigma}_0^{-1}$
for *l = 1 to maxIts* **do**
  Compute $(\mathrm{C}^l, \mu^l) = \arg\min_{(\mathrm{C},\mu)} f(\mathrm{C}, \mu, \mathbf{\Lambda}^{l-1})$ by equations (6) and (7)
  Compute $\mathbf{\Lambda}^l = \arg\min_{\mathbf{\Lambda}} f(\mathrm{C}^l, \mu^l, \mathbf{\Lambda})$ by Eq. (8)
end
$\{\hat{\mathrm{C}}_i = \mathrm{C}_i^{\text{maxIts}}\}$, $\hat{\mu} = \mu^{\text{maxIts}}$, $\hat{\mathbf{\Lambda}} = \mathbf{\Lambda}^{\text{maxIts}}$

Proposition 3. *As $l \to +\infty$ the sequence $f(\{\mathrm{C}_i^l\}, \mu^l, \mathbf{\Lambda}^l)$ converges monotonically and the sequence $\{\{\mathrm{C}_i^l\}, \mu^l l, \mathbf{\Lambda}^l\}$ has a subsequence that converges to a stationary point of f.*

In practice, we observe in our experiments that the algorithm always converges after a few iterations.

E. Full restoration algorithm

The full restoration algorithm used in our esperiments is summarized in Algorithm 2 and illustrated by Figure 1. It alternates two stages: the minimization of f thanks to Algorithm 1, and the estimation of the hyper-parameters $\mu_0, \mathbf{\Sigma}_0$. In order to estimate these parameters, we rely on an *oracle image* computed by aggregation of all the patches estimated on the first stage.

Algorithm 2: HBE algorithm.

Input: Z, $\boldsymbol{D}$, $\mu_0, \mathbf{\Sigma}_0, \kappa, \nu$ (see details in Section III-C)
Output: $\tilde{\mathrm{C}}$
Decompose Z and $\boldsymbol{D}$ into overlapping patches.
Initialization: Compute first oracle image C_{oracle} (see details in Section III-D)
for *it = 1 to maxIts*$_2$ **do**
  for *all patches not yet restored* **do**
    Find patches similar (L^2 distance) to the current z_i in C_{oracle} (see details in Section III-A).
    Compute μ_0 and $\mathbf{\Sigma}_0$ from C_{oracle} (see details in Section III-C).
    Compute $(\{\hat{\mathrm{C}}_i\}, \hat{\mu}, \hat{\mathbf{\Sigma}})$ following Algorithm 1.
  end
  Perform aggregation to restore the image.
  Set $\mathrm{C}_{oracle} = \tilde{\mathrm{C}}$.
end

III. Implementation details

A. Search for similar patches

The similar patches are all patches within a search window centered at the current patch, for which the L_2 distance to it is below a given threshold. This threshold is given by a tolerance parameter ε times the distance to the nearest neighbor (the most similar one). In all our experiments, the search window was set to size 25×25 (with a patch size of 8×8) and $\varepsilon = 1.5$. The patch comparison is performed in an oracle

image (i.e. the result of the previous iteration), so all pixels are known. However, it may be useful to assign different confidence levels to the known pixels and to those originally missing and then restored. For all the experimental results presented in Section IV, the distance between patches c_p and c_q in the oracle image C_{oracle} is computed as

$$d(p,q) = \frac{\sum_{j=1}^{N} (\mathrm{c}_p^j - \mathrm{c}_q^j)^2 \omega_{p,q}^j}{\sum_{j=1}^{N} \omega_{p,q}^j}, \tag{9}$$

where j indexes the pixels in the patch, $\omega_{p,q}^j = 1$ if $\boldsymbol{D}_p^j = \boldsymbol{D}_q^j = 1$ (known pixel) and $\omega_{p,q}^j = 0.01$ otherwise (originally missing then restored pixel) [22]. With this formulation, known pixels are assigned a much higher priority than unknown ones. Variations on these weights could be explored.

B. Optional speed-up by Collaborative Filtering

The proposed method computes one Gaussian model per image patch according to Equations (6) and (8). In order to reduce the computational cost, we can rely on the collaborative filtering idea previously introduced for patch-based denoising techniques [10], [23]. Based on the hypothesis that similar patches share the same model, instead of computing a different pair $(\mu, \boldsymbol{\Sigma})$ for each patch, we assign the same model to all patches in the set of similar patches (as defined in Section III-A). The restoration is thus performed for all similar patches according to the computed model.

The number of similar patches jointly restored depends on the image and the tolerance parameter ε, but it is often much smaller than what would result from the patch clustering performed by methods that use global GMMs such as PLE or EPLL. Performance degradation is observed in practice when using a very large tolerance parameter ($\varepsilon > 3$), showing that mixing more patches than needed is detrimental. The collaborative filtering strategy helps accelerate the algorithm up to a certain point, but a trade-off with performance needs to be considered.

C. Parameters setting

The four parameters of the Normal-Wishart distribution: κ, ν, the prior mean μ_0 and the prior covariance matrix $\boldsymbol{\Sigma}_0$, must be set in order to compute μ and $\boldsymbol{\Sigma}$.

a) Choice of κ and ν: The computation of μ according to (6) combines the mean $\sum_{i=1}^{M} \mathbf{A}_i \mathrm{Z}_i$ estimated from the similar patches and the prior mean μ_0. The parameter κ is related to the degree of confidence we have on the prior μ_0. Hence, its value should be a trade-off between the confidence we have in the prior accuracy vs. the one we have in the information provided by the similar patches. The latter improves when both M (*i.e.* the number of similar patches) and $P = \mathrm{trace}(\boldsymbol{D}_i)$ (*i.e.* the number of known pixels in the current patch) increase. These intuitive insights suggest the following rule to choose the value of κ:

$$\kappa = M\alpha, \quad \alpha = \begin{cases} \alpha_L & \text{if } P \text{ and } M > \text{threshold} \\ \alpha_H & \text{otherwise.} \end{cases} \tag{10}$$

A similar reasoning leads to the same rule for ν,

$$\nu = M\alpha + n \tag{11}$$

where the addition of n ensures the condition $\nu > n - 1$ required by the Normal-Wishart prior to be verified.

This rule is used to obtain the experimental results presented in Section IV, and proved to be a consistently good choice despite its simplicity. However, setting these parameters in a more general setting is not a trivial task and should be the subject of further study. In particular we could explore a more continuous dependence of α on P, M, and possibly a third term $Q = \sum_{i=1}^{n} S_{ii}$ where $S = \sum_{j=1}^{M} \boldsymbol{\Lambda}^{-1} \boldsymbol{D}_j \boldsymbol{\Lambda}_j^* \boldsymbol{D}_j$. This third term Q estimates to what an extent similar patches cover the missing pixels in the current patch.

b) Setting of μ_0 and $\boldsymbol{\Sigma}_0$: Assuming an oracle image C_{oracle} is available (see details in Section II-E), μ_0 and $\boldsymbol{\Sigma}_0$ can be computed using the classical MLE estimators from a set of similar patches $(\tilde{\mathrm{c}}_1, \ldots, \tilde{\mathrm{c}}_M)$ taken from C_{oracle}

$$\mu_0 = \frac{1}{M} \sum_{j=1}^{M} \tilde{\mathrm{c}}_j, \quad \boldsymbol{\Sigma}_0 = \frac{1}{M-1} \sum_{j=1}^{M} (\tilde{\mathrm{c}}_j - \mu_0)(\tilde{\mathrm{c}}_j - \mu_0)^T. \tag{12}$$

This is the same approach followed by Lebrun et al. [10] to estimate the patch model parameters in the case of denoising. As previously stated, the method from [10] cannot be directly applied to zooming or interpolation since the presence of missing pixels results on poor estimates of the mean and the covariance matrix.

D. Initialization

A good initialization is crucial since we aim at solving a non-convex problem through an iterative procedure. To initialize the proposed algorithm we follow the approach proposed by Yu et al. [13] (described in detail in Appendix A in the supplementary material). They propose to initialize the PLE algorithm by learning the K GMM covariance matrices from synthetic images of edges with different orientations as well as the DCT basis to represent isotropic patterns. As they state, in dictionary learning, the most prominent atoms represent local edges which are useful at representing and restoring contours. Hence, this initialization helps to correctly restore corrupted patches even in quite extreme cases. The oracle of the first iteration of the proposed approach is the output of the first iteration of the PLE algorithm.

E. Computational complexity

With the original per-patch strategy, the complexity of the algorithm is defined by the step 3 of Algorithm 1, and is given by: $[(4n^3 + n^3/3)M + n^3/3]$, so the total complexity is $[(4n^3 + n^3/3)M + n^3/3] \times maxIts \times maxIts_2 \times T$ (where $T =$ total number of patches to be restored and assuming the Cholesky factorization is used for matrix inversion). The collaborative filtering strategy reduces this value by a factor that depends on the number of groups of similar patches, which depends on the image content and distance tolerance parameter ε setting. The main difference with the PLE algorithm complexity $((3n^3 + n^3/3) \times its_{PLE} \times T)$ is a factor given by the number of groups defined by the collaborative filtering approach and the ratio between its_{PLE} and $maxIts \times maxIts_2$. As mentioned by Yu et al. [13], computational complexity can be further

reduced in the case of binary masks by removing the zero rows and inverting a matrix of size $n^2/S \times n^2/S$ instead of $n^2 \times n^2$ where S is the masking ratio. Moreover, the proposed algorithm can be run in parallel in different image subregions thus allowing for even further acceleration in multiple-core architectures. The complexity comparison to NLB needs to be made in the case where the degradation is additive noise with constant variance (translation invariant degradation), which is the task performed by NLB. In that case, the complexity of the proposed approach (without considering collaborative filtering nor parallelization, which are both done also in NLB), is $11n^3/3 \times maxIts \times maxIts_2 \times T$ whereas that of NLB is $2 \times (4n^3/3)$.

IV. Image Restoration Experiments

In this section we illustrate the ability of the proposed method to solve several image inverse problems. Both synthetic (i.e., where we have added the degradation artificially) and real data (i.e., issued from a real acquisition process) are used. The considered problems are: interpolation, combined interpolation and denoising, denoising, and zooming. The reported values of peak signal-to-noise ratio ($PSNR = 20\log_{10}(255/\sqrt{MSE})$) are averaged over 10 realizations for each experiment (variance is below 0.1 for interpolation and below 0.05 for combined interpolation and denoising and for denoising only). Similar results are obtained with the structural similarity index (SSIM) which is included in the supplementary material (Appendix B).

A. Synthetic degradation

a) Interpolation: Random masks with 50%, 70% and 90% missing pixels are applied to the tested ground-truth images. The interpolation performance of the proposed method is compared to that of PLE [13], EPLL [12] and E-PLE [16] using a patch size of 8×8 for all methods. PLE parameters are set as indicated in [13] ($\sigma = 3$, $\varepsilon = 30$, $K = 19$). We used the EPLL code provided by the authors [24] with default parameters and the E-PLE code available in [16] with the parameters set as specified in the companion demo. The parameters for the proposed method are set to $\alpha_H = 1$, $\alpha_L = 0.5$ (α_H and α_L define the values for κ and ν, see Section III-C). The PSNR results are shown in Table I. Figure 2 shows some extracts of the obtained results, the PSNR values for the extracts and the corresponding difference images with respect to the ground-truth. The proposed method gives sharper results than the other considered methods. This is specially noticeable on the reconstruction of the texture of the fabric of Barbara's trousers shown in the first row of Figure 2 or on the strips that appear through the car's window shown in the second row of the same figure.

b) Combined interpolation and denoising: For this experiment, the ground-truth images are corrupted with additive Gaussian noise with variance 10, and a random mask with 20% and 70% of missing pixels. The parameters for all methods are set as in the previous interpolation-only experiment. Table I summarizes the PSNR values obtained by each method. Figure 3 shows some extracts of the obtained results, the PSNR values for the extracts and the corresponding difference images with respect to the ground-truth. Once again, the results show that the proposed approach outperforms the others. Fine structures, such as the mast and the ropes of the ship, as well as textures, as in Barbara's headscarf, are much better preserved.

c) Denoising: For the denoising task, the proposed approach should perform very similarly to the state-of-the-art denoising algorithm NLB [10]. The following experiments are conducted in order to verify this.

The ground-truth images are corrupted with additive Gaussian noise with variance $\sigma^2 = 10, 30, 50, 80$. The code provided by the authors [25] automatically sets the NLB parameters from the input σ^2 and the patch size, in this case 8×8. For this experiment, there are no unknown pixels to interpolate (the mask $\boldsymbol{D}$ is the identity matrix).

The results of both methods are very similar if HBE is initialized with the output of the first step of NLB [10] (instead of using the initialization described in Section II-E) and the parameters κ and ν are large enough. In this case, μ_0 and $\boldsymbol{\Sigma}_0$ are prioritized in equations (6) and (8) and both algorithms are almost the same. That is what we observe in practice with $\alpha_H = \alpha_L = 100$, as demonstrated in the results summarized in Table I. The denoising performance of HBE is degraded for small κ and ν values. The reason for this is that μ_0 and $\boldsymbol{\Sigma}_0$, as well as μ and $\boldsymbol{\Sigma}$ in NLB, are computed from an oracle image resulting from the first restoration step. This restoration includes not only the denoising of each patch, but also an aggregation step that highly improves the final result. Therefore, the contribution of the first term of (6) to the computation of $\hat{\mu}$ degrades the result compared to that of using μ_0 only (i.e. using a large κ).

d) Zooming: In order to evaluate the zooming capacity of the proposed approach, ground-truth images are downsampled by a factor 2 (no anti-aliasing filter is used) and the zooming is compared to the ground-truth. The results are compared with PLE, EPLL, E-PLE and Lanczos interpolation. Figure 4 shows extracts from the obtained results, the PSNR values for the extracts and the corresponding difference images with respect to the ground-truth. Again, HBE yields a sharper reconstruction than the other methods.

B. Real data

For this experiment, we use raw images captured with a Canon 400D camera (set to ISO 400 and exposure time 1/160 seconds). The main noise sources for CMOS sensors are: the Poisson photon shot noise, which can be approximated by a Gaussian distribution with equal mean and variance; the thermally generated readout noise, which is modeled as an additive Gaussian distributed noise and the spatially varying gain given by the photo response non uniformity (PRNU) [20], [26]. We thus consider in this case the following noise model for the non saturated raw pixel value $\mathrm{Z}(p)$ at position p

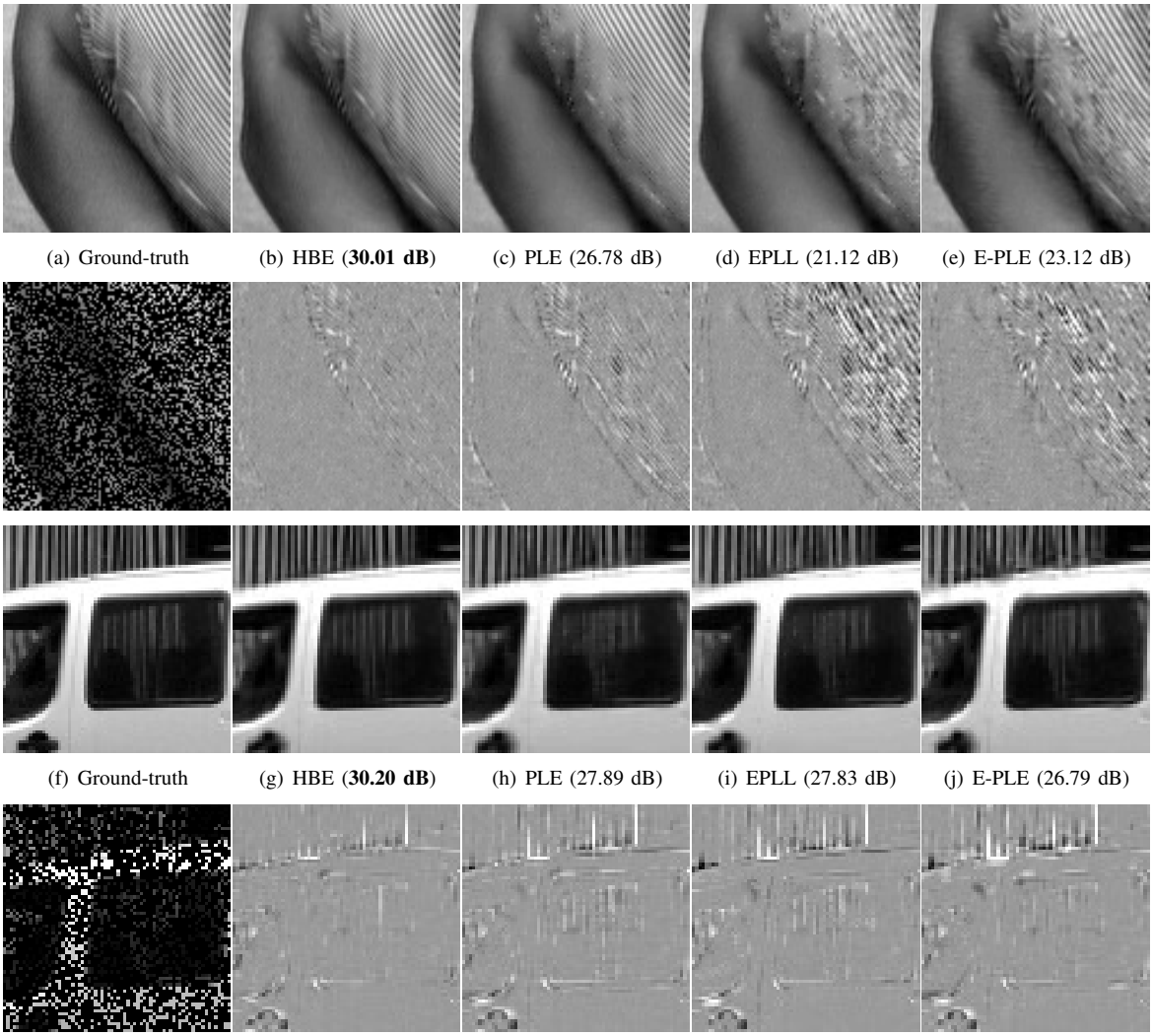
$$\mathrm{Z}(p) \sim \mathcal{N}(g a_p \tau \mathrm{C}(p) + \mu_R, g^2 a_p \tau \mathrm{C}(p) + \sigma_R^2), \qquad (13)$$

where g is the camera gain, a_p models the PRNU factor, τ is the exposure time, $\mathrm{C}(p)$ is the irradiance reaching pixel p, μ_R and σ_R^2 are the readout noise mean and variance. The camera parameters have to be estimated by a calibration procedure [20].

		PSNR (dB)											
		HBE	PLE	EPLL	E-PLE	HBE	PLE	EPLL	E-PLE	HBE	PLE	EPLL	E-PLE
	% missing pixels	50%				70%				90%			
Interpolation	barbara	**39.11**	36.93	32.99	35.43	**34.69**	32.50	27.96	28.77	**24.86**	23.62	23.30	23.26
	boat	**34.92**	34.32	34.21	33.59	**31.37**	30.74	30.38	30.26	**25.96**	25.35	24.72	25.43
	traffic	30.17	30.12	**30.19**	28.86	**27.27**	27.12	27.13	26.64	**22.84**	22.34	21.85	22.27
Interpolation & Denoising	barbara	-	-	-	-	**33.34**	31.99	27.63	27.75	**24.57**	23.53	23.27	23.20
	boat	-	-	-	-	**30.61**	30.41	30.15	29.54	**25.78**	25.45	24.71	25.47
	traffic	-	-	-	-	26.99	26.98	**27.05**	26.35	**22.87**	22.43	22.21	22.35
		HBE	NLB	EPLL	HBE	NLB	EPLL	HBE	NLB	EPLL	HBE	NLB	EPLL
	$\boldsymbol{\sigma^2}$	10			30			50			80		
Denoising	barbara	**41.26**	41.20	40.56	**38.40**	38.26	37.32	**37.13**	36.94	35.84	**35.96**	35.73	34.51
	boat	**40.05**	39.99	39.47	36.71	**36.76**	36.34	35.41	**35.46**	35.13	**34.30**	34.33	34.12
	traffic	40.73	**40.74**	40.55	**37.03**	36.99	36.86	**35.32**	35.26	35.20	**33.78**	33.70	33.72

TABLE I

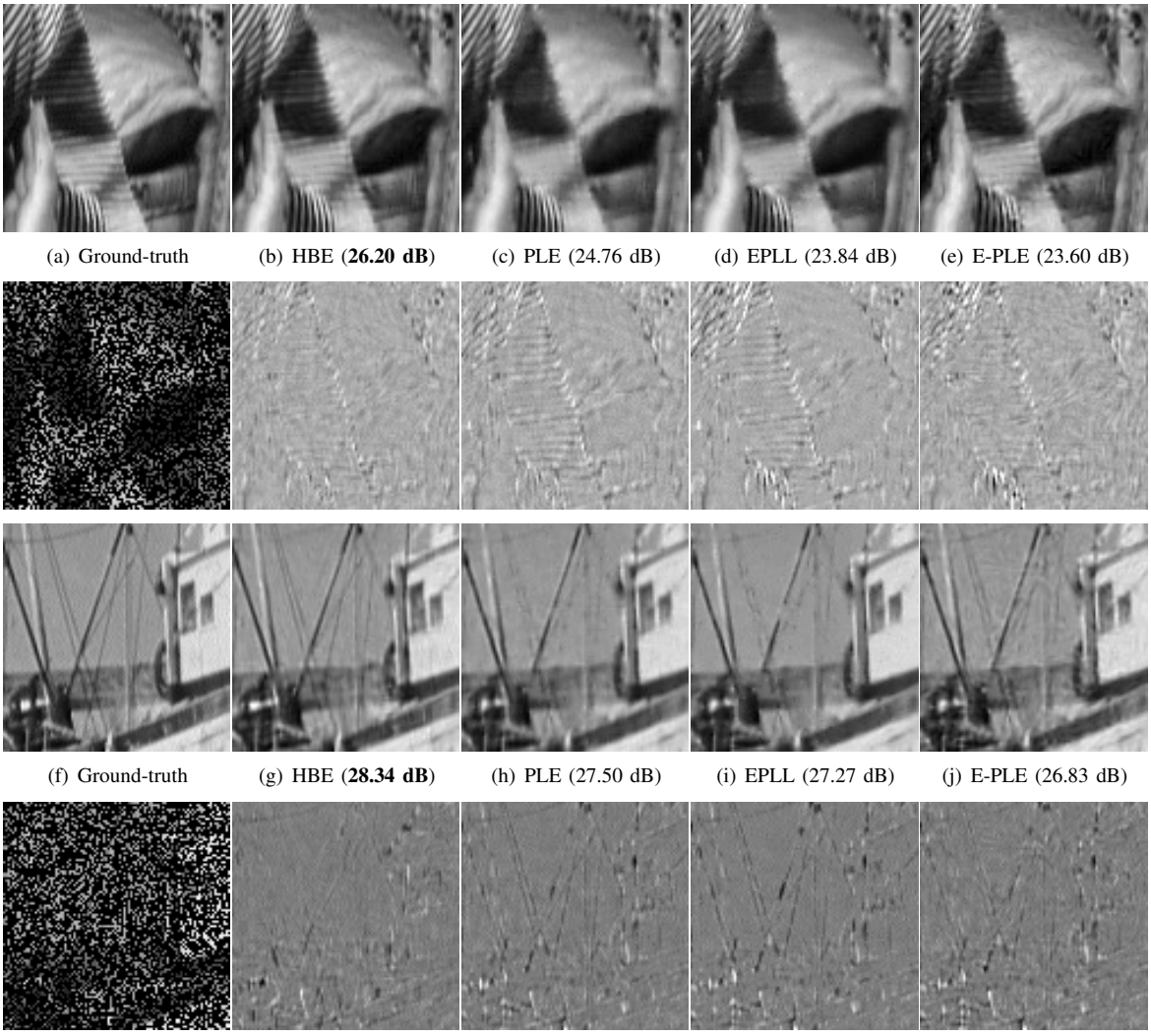
RESULTS OF THE INTERPOLATION, COMBINED INTERPOLATION AND DENOISING AND DENOISING TESTS DESCRIBED IN SECTION IV-A. PATCH SIZE OF 8×8 FOR ALL METHODS IN ALL TESTS. PARAMETER SETTING FOR INTERPOLATION AND COMBINED INTERPOLATION AND DENOISING, HBE: $\alpha_H = 1$, $\alpha_L = 0.5$, PLE: $\sigma = 3$, $\varepsilon = 30$, $K = 19$ [13], EPLL: DEFAULT PARAMETERS [24], E-PLE: PARAMETERS SET AS SPECIFIED IN [16]. PARAMETER SETTING FOR DENOISING, HBE: $\alpha_H = \alpha_L = 100$, NLB: CODE PROVIDED BY THE AUTHORS [25] AUTOMATICALLY SETS PARAMETERS FROM INPUT σ^2, EPLL: DEFAULT PARAMETERS FOR THE DENOISING EXAMPLE [24]



(a) Ground-truth (b) HBE (**30.01 dB**) (c) PLE (26.78 dB) (d) EPLL (21.12 dB) (e) E-PLE (23.12 dB)

(f) Ground-truth (g) HBE (**30.20 dB**) (h) PLE (27.89 dB) (i) EPLL (27.83 dB) (j) E-PLE (26.79 dB)

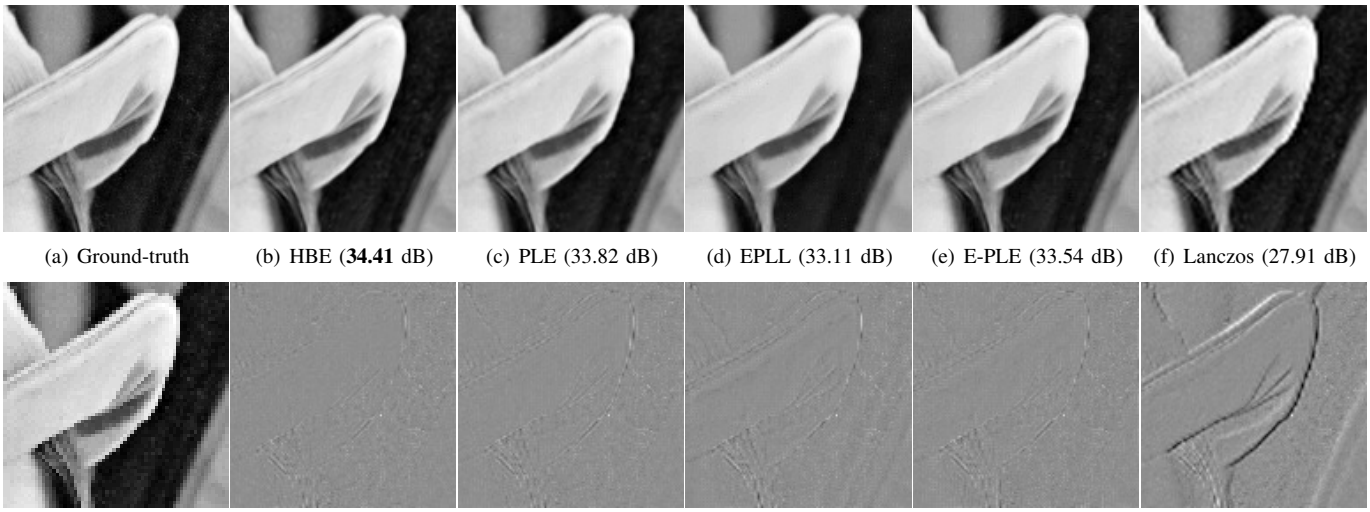
Fig. 2. **Synthetic data. Interpolation with 70% of randomly missing pixels**. **Left to right:** (first row) Ground-truth (extract of barbara), result by HBE, PLE, EPLL, E-PLE. (second row) input image, difference with respect to the ground-truth of each of the corresponding results. (third and fourth row) Idem for an extract of the traffic image. See Table I for the PSNR results for the complete images. Please see the digital copy for better details reproduction.



(a) Ground-truth (b) HBE (**26.20 dB**) (c) PLE (24.76 dB) (d) EPLL (23.84 dB) (e) E-PLE (23.60 dB)

(f) Ground-truth (g) HBE (**28.34 dB**) (h) PLE (27.50 dB) (i) EPLL (27.27 dB) (j) E-PLE (26.83 dB)

Fig. 3. **Synthetic data. Combined interpolation and denoising with 70% of randomly missing pixels and additive Gaussian noise ($\sigma^2 = 10$). Left to right:** (first row) Ground-truth (extract of barbara), result by HBE, PLE, EPLL, E-PLE. (second row) input image, difference with respect to the ground-truth of each of the corresponding results. (third and fourth row) Idem for an extract of the boat image. See Table I for the PSNR results for the complete images. Please see the digital copy for better details reproduction.



(a) Ground-truth (b) HBE (**34.41** dB) (c) PLE (33.82 dB) (d) EPLL (33.11 dB) (e) E-PLE (33.54 dB) (f) Lanczos (27.91 dB)

Fig. 4. **Synthetic data. Zooming $\times 2$. Left to right:** (first row) Ground-truth high resolution image (extract of lena). Result by HBE, PLE, EPLL, E-PLE, lanczos interpolation. (second row) Input low-resolution image, difference with respect to the ground-truth of each of the corresponding results. Please see the digital copy for better details reproduction.

Fig. 5. **Left. Real data.** JPEG version of the raw image used in the experiments presented in Section IV-B. The boxes show the extracts displayed in Figure 6. **Right. Synthetic data.** Ground-truth images used in the experiments presented in Section IV-A.

The noise covariance matrix $\mathbf{\Sigma}_N$ is thus diagonal with entries that depend on the pixel value $(\mathbf{\Sigma}_N)_p = g^2 a_p \tau \mathrm{C}(p) + \sigma_R^2$.

In order to evaluate the interpolation capacity of the proposed approach, we consider the pixels of the green channel only (i.e. 50% of the pixels in the RGGB Bayer pattern) and interpolate the missing values. We compare the results to those obtained using an adaptation of PLE to images degraded with noise with variable variance (PLEV) [27]. The results for the EPLL and E-PLE methods are not presented here since these methods are not suited for this kind of noise. Figure 6 shows extracts of the obtained results (see Figure 5 for a JPEG version of the raw image showing the location of the the extracts). As it was already observed in the synthetic data experiments, fine details and edges are better preserved. Compare for example the reconstruction of the balcony edges and the wall structure in the first row of Figure 6, as well as the structure of the roof and the railing in the second row of the same image.

C. Discussion

In all the tested examples, the results obtained by HBE for every considered inverse problem outperform or are very close to those obtained by the other evaluated methods. Details are better reconstructed and the resulting images are sharper both in synthetic and real data examples. The improvement is more noticeable when comparing the difference images (available for the synthetic tests only), which present less structure in the result obtained by HBE.

Even if the PLE method can be considered as semi-local (since it is applied in 128×128 regions [13]), we find that in some cases, 19 classes are not enough to correctly represent certain image patches. This is mostly the case for patches that seldom appear in the image, such as certain edges or particular textures that appear in a few patches. This is quite noticeable in the extract of Barbara's trousers and in the interior of the car (Figure 2). The specific characteristics of these patches are buried in the PLE class update when combined with many other different patches.

A local model estimation as the one performed by HBE correctly handles those cases. The performance difference is much more remarkable for the higher masking rates. In those cases, two phenomena take place. On the one hand, very few pixels are known thus making the model selection less robust. On the other hand, the model accuracy is critical since a much larger part of the patch is to be restored. The proposed method tackles the model selection problem by limiting the model estimation to similar patches found on a local search window. It has been widely observed in denoising techniques based on the self-similarity principle [28] that performance improves when restricting the patch search space to a local search window instead of using the whole image. This strategy, in addition to the hypothesis of self-similarity in that neighborhood, restricts the possible models robustifying the model estimation, which is crucial for high masking rates. Furthermore, the local model estimation, previously proven successful at describing patches [10], gives a better reconstruction even when a very large part of the patch is missing.

EPLL uses more mixture components in its GMM model than PLE, where 200 components are learnt from 2×10^6 patches of natural images [12]. The results obtained by this approach, despite using a larger number of GMM components, are not very good for the restoration of certain patches. As previously mentioned, Wang and Morel [11] claim that in the case of denoising, it is better to have fewer models that are updated with the image patches (as in PLE) rather than having a large number of fixed models (as in EPLL). In this work, we observe that the proposed approach outperforms EPLL, not only in denoising, but also in inpainting and zooming. However, here it is harder to tell if the improvement is due to the local model estimation performed from the similar patches or if it is due to the different restoration strategies followed by these methods.

V. High dynamic range imaging from a single snapshot

In this section, we propose a novel approach to generate high dynamic range (HDR) images from a single shot based on the general framework introduced in Section II.

HDR imaging aims at reproducing an extended dynamic range of luminosity compared to what can be captured using a standard digital camera. The range of luminosity which a standard digital camera can capture is often not enough to produce a faithful representation of real scenes. In the case of a static scene and a static camera, the combination of multiple images with different exposure levels is a simple and efficient solution [29], [30], [26]. However, several problems arise when either the camera or the elements in the scene move [31], [32].

An alternative to the HDR from multiple frames was introduced by Nayar and Mitsunaga in [33]. They proposed to perform HDR imaging from a single image using spatially varying pixel exposures (SVE). An optical mask with spatially varying transmittance is placed adjacent to a conventional image sensor, thus controlling the amount of light that reaches each pixel (see Figure 7). This gives different exposure levels to the pixels allowing a single shot to capture an increased dynamic range compared to that of the conventional sensor.

The greatest advantage of this acquisition method is that it produces HDR imaging from a single image, thus avoiding the need for image alignment and motion estimation, which is the main drawback of the classical multi-image approach. Another advantage is that the saturated pixels are not organized

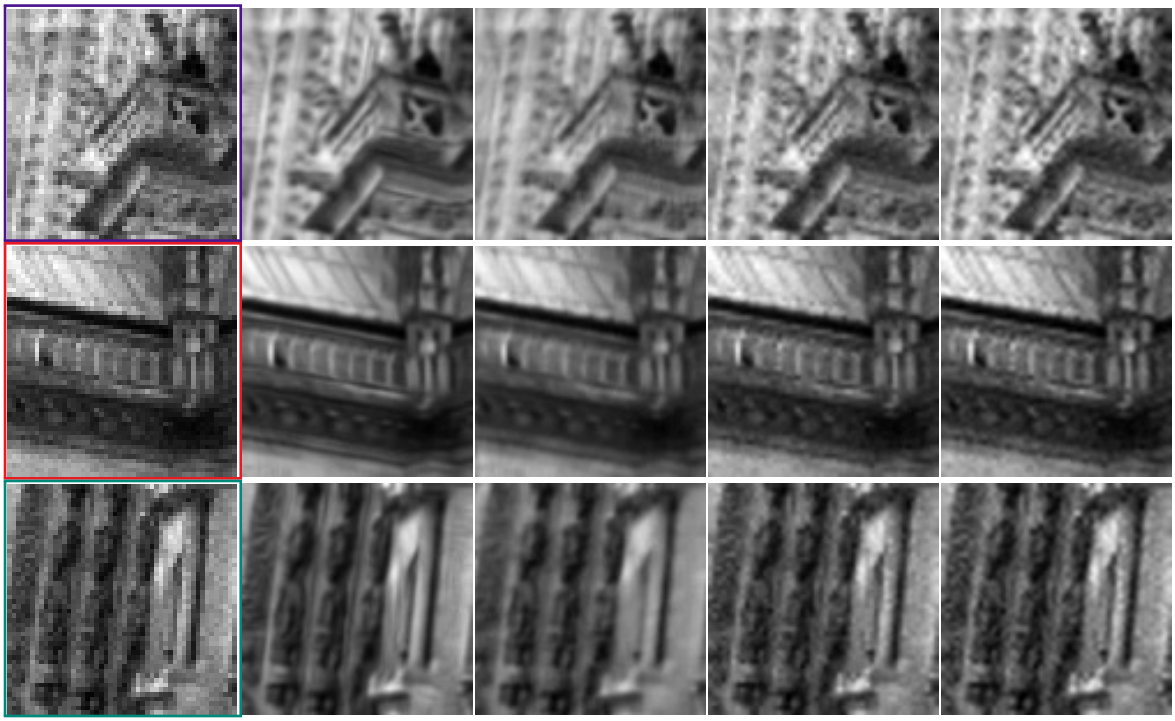


Fig. 6. **Real data. Zooming** $\times 2$**.** Interpolation of the green channel of a raw image (RGGB). **Left to right:** Input low-resolution image, result by HBE, PLEV [27], bicubic and lanczos interpolation.

in large regions. Indeed, some recent multi-image methods tackle the camera and objects motion problems by taking a reference image and then estimating motion relative to this frame or by recovering information from other frames through local comparison with the reference [34], [31]. A problem encountered by these approaches is the need to inpaint very large saturated and underexposed regions in the reference frame, since the information is completely lost in those areas. The SVE acquisition strategy avoids this problem since, in general, all scene regions are sampled by at least one of the exposures.

Taking advantage of the ability of the proposed framework to simultaneously estimate missing pixels and denoise well-exposed ones, we propose a novel approach to generate HDR images from a single shot acquired with spatially varying pixel exposures. The proposed approach shows significant improvements over existing approaches.

A. Spatially varying exposure acquisition model

As presented in [33], [35], [36] and recalled above, an optical mask with spatially varying transmittance is placed adjacent to a conventional image sensor to give different exposure levels to the pixels. This optical mask does not change the acquisition process of the sensor, whether using a conventional CCD or CMOS sensor. Hence, the noise model (13) can be adapted to the SVE acquisition by including the per-pixel SVE gain o_p[5]:

$$\mathrm{Z}(p) \sim \mathcal{N}(g o_p a_p \tau \mathrm{C}(p) + \mu_R, g^2 o_p a_p \tau \mathrm{C}(p) + \sigma_R^2). \quad (14)$$

In the approach proposed by Nayar and Mitsunaga [33], the varying exposures follow a regular pattern. Motivated by the aliasing problems of regular sampling patterns, Schöberl et al. [37] propose to use spatially varying exposures on a non-regular pattern. Figure 7 shows examples of both acquisition patterns. This fact led us to choose the non-regular pattern in the proposed approach.

B. Hyperprior Bayesian Estimator for Single Shot High Dynamic Range Imaging

1) Problem statement: In order to reconstruct the dynamic range of the scene we need to solve an inverse problem, that is, to find the irradiance values from the input pixel values. More precisely, we want to estimate the irradiance image C from the SVE image Z, knowing the exposure levels of the optical mask and the camera parameters.

For this purpose we map the raw pixel values to the irradiance domain Y with

$$\mathrm{Y}(p) = \frac{\mathrm{Z}(p) - \mu_R}{g o_p a_p \tau}. \quad (15)$$

We take into account the effect of saturation and underexposure by introducing the exposure degradation matrix $\boldsymbol{D}$,

[5]Some noise sources not modeled here, such as blooming, might have an impact in the SVE acquisition strategy and should be considered in a more accurate image modeling.

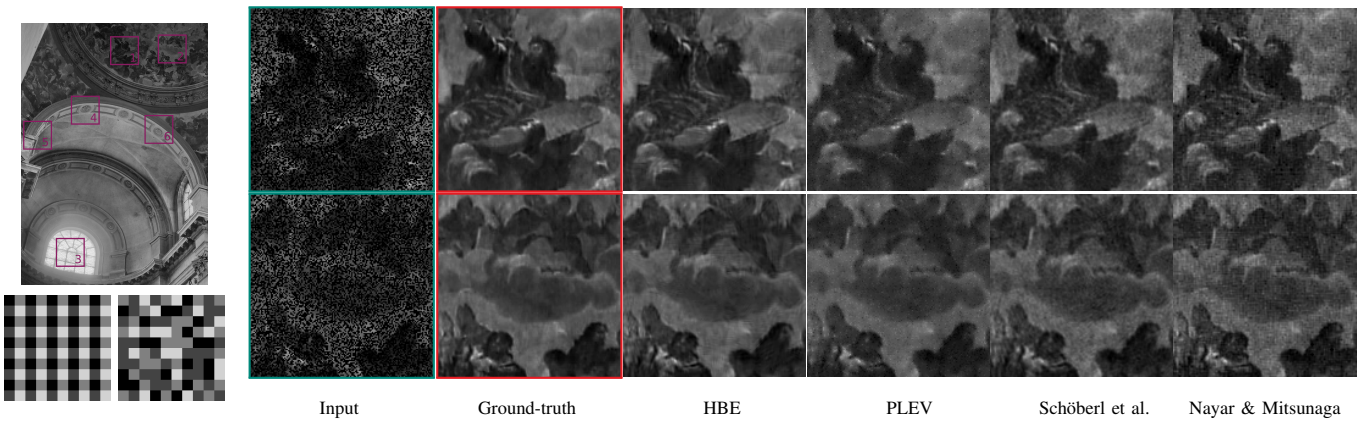


Fig. 7. **Synthetic data. Left:** (**top**) Tone mapped version of the ground-truth image used for the experiments in Section V-C1. Image from [38]. (**bottom**) Regular (left) and non-regular (right) optical masks for an example of 4 different filters. **Right:** Results for extracts 1 and 6. From left to right: Input image with random pattern, ground-truth, results by HBE, PLEV [27], Schöberl et al. [37], Nayar and Mitsunaga [33]. 50% missing pixels (for both random and regular pattern). See PSNR values for these extracts in Table II. Please see the digital copy for better details reproduction.

a diagonal matrix given by

$$(\boldsymbol{D})_p = \begin{cases} 1 & \text{if } \mu_R < \mathrm{Z}(p) < z_{sat}, \\ 0 & \text{otherwise,} \end{cases} \tag{16}$$

with z_{sat} equal to the pixel saturation value, thus eliminating the under or overexposed pixels. From (14) and (16), $\mathrm{Y}(p)$ can be modeled as

$$\mathrm{Y}(p)|(\boldsymbol{D})_p \sim \mathcal{N}\left((\boldsymbol{D})_p \mathrm{C}(p), \frac{g^2 o_p a_p \tau (\boldsymbol{D})_p \mathrm{C}(p) + \sigma_R^2}{(g o_p a_p \tau)^2}\right). \tag{17}$$

Notice that (17) is the distribution of $\mathrm{Y}(p)$ for a given exposure degradation factor $(\boldsymbol{D})_p$, since $(\boldsymbol{D})_p$ is itself a random variable that depends on $\mathrm{Z}(p)$. The exposure degradation factor must be included in (17) since the variance of the over or under exposed pixels no longer depends on the irradiance $\mathrm{C}(p)$ but is only due to the readout noise σ_R^2. From (17) we have

$$\mathrm{Y} = \boldsymbol{D}\mathrm{C} + \mathrm{N}, \tag{18}$$

where N is zero-mean Gaussian noise with diagonal covariance matrix $\boldsymbol{\Sigma}_\mathrm{N}$ given by

$$(\boldsymbol{\Sigma}_\mathrm{N})_j = \frac{g^2 o_p a_p \tau (\boldsymbol{D})_p \mathrm{C}(p) + \sigma_R^2}{(g o_p a_p \tau)^2}. \tag{19}$$

Then the problem of irradiance estimation can be stated as retrieving C from Y, which implies denoising the well-exposed pixel values ($(\boldsymbol{D})_p = 1$) and estimating the unknown ones ($(\boldsymbol{D})_p = 0$).

2) Proposed solution: From (18), image Y is under the hypothesis of the HBE framework that we developed in Section II, so that we can apply the proposed patch-based reconstruction approach to HDR imaging. The proposed HDR imaging algorithm consists of the following steps: **1)** generate $\boldsymbol{D}$ from Z according to (16), **2)** obtain Y from Z according to (15), **3)** apply the HBE approach to Y with the given $\boldsymbol{D}$ and $\boldsymbol{\Sigma}_\mathrm{N}$.

C. Experiments

The proposed reconstruction method was thoroughly tested in several synthetic and real data examples. A brief summary of the results is presented in this section.

	PSNR (dB)					
	1 (Fig. 7)	2 (Fig. 7)	3	4	5	6
HBE	**33.08**	**33.87**	22.95	**35.10**	**36.80**	**35.66**
PLEV	29.65	30.82	22.77	33.99	36.42	34.73
Schöberl et al.	30.38	31.16	21.39	30.04	32.84	31.02
Nayar and Mitsunaga	29.39	30.10	**23.24**	25.83	30.26	26.90

TABLE II
PSNR VALUES FOR THE EXTRACTS SHOWN IN FIGURE 7.

1) Synthetic data: Sample images are generated according to Model (18) using the HDR image in Figure 7 as the ground-truth. Both a random and a regular pattern with four equiprobable exposure levels $o = \{1, 8, 64, 512\}$ are simulated. The exposure time is set to $\tau = 1/200$ seconds and the camera parameters are those of a Canon 7D camera set to ISO 200 ($g = 0.87$, $\sigma_R^2 = 30$, $\mu_R = 2048$, $\mathrm{v_{sat}} = 15000$) [26].

Figure 7 shows extracts of the results obtained by the proposed method, by PLEV [27] (basically an adaptation of PLE to the same single image framework) and by Schöberl et al. [37] for the random pattern and by Nayar et Mitsunaga [33] using the regular pattern. The percentage of unknown pixels in the considered extracts is 50% (it is nearly the same for both the regular and non-regular pattern). Table II shows the PSNR values obtained in each extract marked in Figure 7. The proposed method manages to correctly reconstruct the irradiance on the unknown pixels. Moreover, its denoising performance is much better than that of Schöberl et al. and Nayar and Mitsunaga, and still sharper than PLEV.

2) Real data: The feasibility of the SVE random pattern has been shown in [36] and that of the SVE regular pattern in [35]. Nevertheless, these acquisition systems are still not available for general usage. However, as stated in Section V-A, the only variation between the classical and the SVE acquisition is the optical filter, i.e. the amount of light reaching each pixel. Hence, the noise at a pixel p captured using SVE with an optical gain factor o_p and exposure time τ/o_p and a pixel captured with a classical camera using exposure time τ should be very close. We take advantage of this fact in order to evaluate the reconstruction performance of the proposed

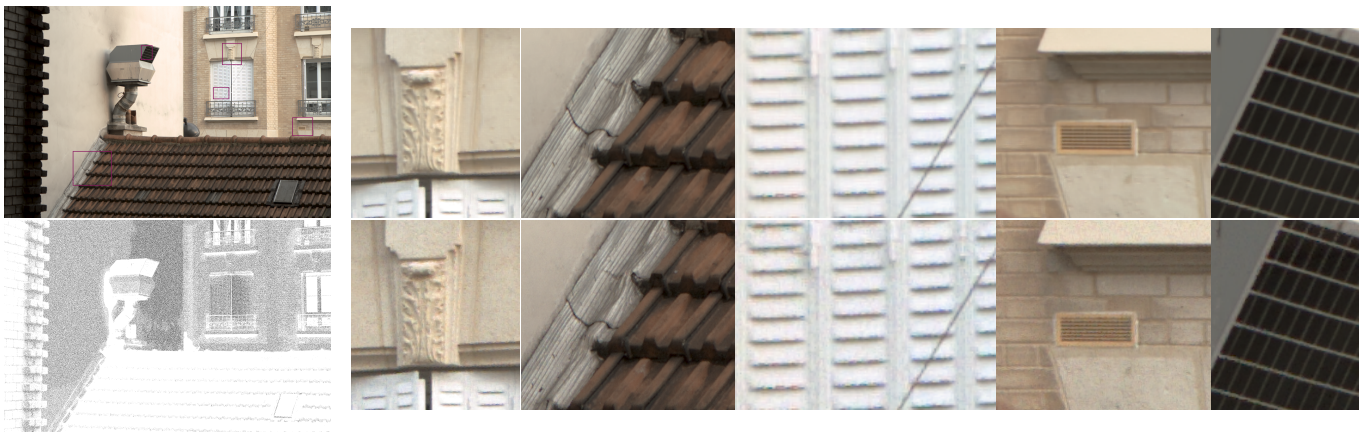


Fig. 8. **Real data. Left:** Tone mapped version of the HDR image obtained by the proposed approach and its corresponding mask of unknown (black) and well-exposed (white) pixels. **Right:** Comparison of the results obtained by the proposed approach (first row) and PLEV (second row) in the extracts indicated in the top image. Please see the digital copy for better details reproduction.

approach using real data.

For this purpose, we generate an SVE image z_{sve} drawing pixels at random from four raw images $\{\mathrm{z}^i_{raw}\}_{i=1,\dots,4}$ acquired with different exposure times. The four different exposure times simulate four different filters of the SVE optical mask. More precisely, the value at position (x, y) in z_{sve} is chosen at random among the four available values at that position $\{\mathrm{z}^i_{raw}(x, y)\}_{i=1,\dots,4}$. Notice that the Bayer pattern is kept on z_{sve} by construction. The images $\{\mathrm{z}^i_{raw}\}_{i=1,\dots,4}$ are acquired using a remotely controlled camera and a tripod so as to be perfectly aligned. Otherwise, artifacts may appear from the random sampling of the four images used to create the SVE frame.

This protocol does not allow us to take scenes with moving objects. Let us emphasize, however, that using a real SVE device, this, as well as the treatment of moving camera, would be a non-issue.

Figures 8 and 9 show the results obtained in two real scenes, together with the masks of well-exposed (white) and unknown (black) pixels (the SVE raw images are included in Appendix B in the supplementary material). Recall that among the unknown pixels, some of them are saturated and some of them are under exposed pixels. Square patches of size 6 and 8 were used for the examples in Figure 9 and Figure 8 respectively. For displaying purposes, the demosaicking method by Adams and Hamilton [39] is then used to obtain a color image from the reconstructed irradiance, followed by the tone mapping technique by Mantiuk et al. [40].

We compare the results to those obtained by PLEV. A comparison against the methods by Nayar and Mitsunaga and Schöberl et al. is not presented since they do not specify how to treat raw images with a Bayer pattern (how to treat color) and therefore an adaptation of their methods should be made in order to process our data. The proposed method manages to correctly reconstruct the unknown pixels even in extreme conditions where more than 70% of the pixels are missing, as for example the last extract in Figure 9.

These examples show the suitability of the proposed approach to reconstruct the irradiance information in both very dark and bright regions simultaneously. See for instance the example in Figure 9, where the dark interior of the building (which can be seen through the windows) and the highly illuminated part of another building are both correctly reconstructed (please see the pdf version of this article for better visualization).

VI. Conclusions

In this work we have presented a novel image restoration framework. It has the benefits of local patch characterization proven suitable by the NLB denoising methods, but manages to extend its use to more general restoration problems where the linear degradation operator is diagonal, by combining local estimation with Bayesian restoration based on hyperpriors. This includes problems such as zooming, inpainting and interpolation. In this way, all these restoration problems are set under the same framework. It does not include image deblurring or deconvolution, since the degradation operator is no longer diagonal. Correctly addressing deconvolution with large kernels with patch-based approaches and Gaussian prior models is a major challenge that will be the subject of future work.

We have presented a large series of experiments both on synthetic and real data that confirm the robustness of the proposed strategy based on hyperpriors. These experiments show that for a wide range of image restoration problems HBE outperforms several state-of-the-art restoration methods.

This work opens several perspectives. The first one concerns the relevance of the Gaussian patch model and its relation to the underlying image patches manifold. If this linear approximation has proven successful for image restoration, its full relevance in other areas remains to be explored, especially in all domains requiring to compare image patches. Another important related question is the one of the estimation of the degradation model in images jointly degraded by noise, missing pixels, blur, etc. Restoration approaches generally rely on the precise knowledge of this model and of its parameters. In practice however, we often deal with images for which the acquisition process is unknown, and that have possibly been affected by post-treatments. In such cases, blind restoration remains an unsolved challenge.

Fig. 9. **Real data. Left:** Tone mapped version of the HDR image obtained by the proposed approach and its corresponding mask of unknown (black) and well-exposed (white) pixels. **Right:** Comparison of the results obtained by the proposed approach (first row) and PLEV (second row) in the extracts indicated in the top image. Please see the digital copy for better details reproduction.

Finally, we have presented a novel application of the proposed general framework to the generation of HDR images from a single variable exposure (SVE) snapshot. The SVE acquisition strategy allows the creation of HDR images from a single shot without the drawbacks of multi-image approaches, such as the need for global alignment and motion estimation to avoid ghosting problems. The proposed method manages to simultaneously denoise and reconstruct the missing pixels, even in the presence of (possibly complex) motions, improving the results obtained by existing methods. Examples with real data acquired in very similar conditions to those of the SVE acquisition show the capabilities of the proposed approach.

APPENDIX A
MAXIMUM A POSTERIORI

A. Optimality equations

Proposition 4. *Assume that $\mathbf{\Lambda}$ is fixed and that $\mathbf{\Sigma}_{N_i}$ and C_i are independent. Then the function $(\{\mathrm{C}_i\},\mu) \to f(\{\mathrm{C}_i\},\mu,\mathbf{\Lambda})$ is convex on $\mathbb{R}^{n(M+1)}$ and has a unique minimizer $(\{\hat{\mathrm{C}}_i\},\hat{\mu})$, given by*

$$\hat{\mu} = \left(\kappa \mathrm{Id} + \sum_{i=1}^{M} A_i \boldsymbol{D}_i\right)^{-1} \left(\sum_{i=1}^{M} A_i \mathrm{Z}_i + \kappa\mu_0\right). \quad (20)$$

$$\hat{\mathrm{C}}_i = A_i(\mathrm{Z}_i - \boldsymbol{D}_i\hat{\mu}) + \hat{\mu}, \qquad \forall i \in \{1,\dots M\} \quad (21)$$

with $A_i = \mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T + \mathbf{\Sigma}_{N_i})^{-1}$.

Proof. Showing that f is convex in $(\{\mathrm{C}_i\},\mu)$ is a simple exercise, left to the reader. It is also easy to show that the minimum of the quadratic function f is unique. Indeed the expression of this unique solution involves the inversion of a positive definite matrix. Deriving f with respect to each C_i, we obtain

$$\frac{\partial f(\{\mathrm{C}_i\},\mu,\mathbf{\Lambda} \mid \{\mathrm{Z}_i\})}{\partial \mathrm{C}_i} = -\boldsymbol{D}_i^T\mathbf{\Sigma}_{N_i}^{-1}(\mathrm{Z}_i - \boldsymbol{D}_i\mathrm{C}_i) + \mathbf{\Lambda}(\mathrm{C}_i - \mu). \quad (22)$$

Equating to zero, the explicit solution must satisfy

$$\begin{aligned} \mathrm{C}_i &= \mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T + \mathbf{\Sigma}_{N_i})^{-1}(\mathrm{Z}_i - \boldsymbol{D}_i\mu) + \mu \\ &= A_i(\mathrm{Z}_i - \boldsymbol{D}_i\mu) + \mu. \end{aligned} \quad (23)$$

We recognize here a Wiener-like estimator for each C_i separately.

Derivating with respect to μ, we obtain

$$\frac{f(\{\mathrm{C}_i\},\mu,\mathbf{\Lambda} \mid \{\mathrm{Z}_i\})}{\partial \mu} = \mathbf{\Lambda}\sum_{i=1}^{M}(\mu - \mathrm{C}_i) + \kappa\mathbf{\Lambda}(\mu - \mu_0), \quad (24)$$

which is zero if and only if (assuming $\mathbf{\Lambda}$ is invertible)

$$\mu = \frac{M\overline{C} + \kappa\mu_0}{M + \kappa}, \quad \text{with } \overline{C} = \frac{1}{M}\sum_{i=1}^{M}\mathrm{C}_i. \quad (25)$$

Injecting the expression of the C_i's at the minimum into the previous expression of μ, we get

$$\mu = \frac{\sum_{i=1}^{M} A_i(\mathrm{Z}_i - \boldsymbol{D}_i\mu) + M\mu + \kappa\mu_0}{M + \kappa},$$

Grouping all the terms in μ on the left hand side of the equation, we obtain the expected explicit solution

$$\mu = (\kappa\mathrm{Id} + \sum_{i=1}^{M} A_i\boldsymbol{D}_i)^{-1}(\sum_{i=1}^{M} A_i\mathrm{Z}_i + \kappa\mu_0). \quad (26)$$

□

Proposition 5. *Assume that the variables $(\{\mathrm{C}_i\},\mu)$ are fixed. The function $\mathbf{\Lambda} \to f(\{\mathrm{C}_i\},\mu,\mathbf{\Lambda})$ is convex on $S_n^{++}(\mathbb{R})$ and its unique minimizer is*

$$\hat{\mathbf{\Lambda}} = \left(\frac{\nu\mathbf{\Sigma}_0 + \kappa(\mu - \mu_0)(\mu - \mu_0)^T + \sum_{i=1}^{M}(\mathrm{C}_i - \mu)(\mathrm{C}_i - \mu)^T}{\nu + M - n}\right)^{-1}. \quad (27)$$

Proof. When $(\{\mathrm{C}_i\},\mu)$ is fixed, the function f is convex in $\mathbf{\Lambda}$. Indeed, f is a sum of linear terms in $\mathbf{\Lambda}$ and a term in $-\log|\mathbf{\Lambda}|$ which is also convex [41]. In order to compute the partial derivative with respect to $\mathbf{\Lambda}$ (see for instance [42] for these derivations), observe that

$$\begin{aligned} f(\{\mathrm{C}_i\},\mu,\mathbf{\Lambda}) &= -\frac{\nu - n + M}{2}\log|\mathbf{\Lambda}| \\ &+ \frac{1}{2}\sum_{i=1}^{M}(\mathrm{C}_i - \mu)^T\mathbf{\Lambda}(\mathrm{C}_i - \mu) \\ &+ \frac{\kappa}{2}(\mu - \mu_0)^T\mathbf{\Lambda}(\mu - \mu_0) \\ &+ \frac{1}{2}\mathrm{trace}[\nu\mathbf{\Sigma}_0\mathbf{\Lambda}]. \end{aligned} \quad (28)$$

The derivative of the first term is ($\mathbf{\Lambda}$ is symmetric)

$$-\frac{\nu-n+M}{2}(\mathbf{\Lambda}^{-1})^T = -\frac{\nu-n+M}{2}\mathbf{\Lambda}^{-1}.$$

The derivative of the second and third terms are

$$\frac{1}{2}\sum_{i=1}^{M}(\mathrm{C}_i-\mu)(\mathrm{C}_i-\mu)^T$$

and

$$\frac{\kappa}{2}(\mu-\mu_0)(\mu-\mu_0)^T.$$

Finally, the derivative of the fourth term is

$$\frac{1}{2}\nu\mathbf{\Sigma}_0^T = \frac{1}{2}\nu\mathbf{\Sigma}_0.$$

Then it follows that

$$\begin{aligned}\frac{\partial f(\{\mathrm{C}_i\},\mu,\mathbf{\Lambda}\mid\{\mathrm{Z}_i\})}{\partial\mathbf{\Lambda}} &= -\frac{\nu-n+M}{2}\mathbf{\Lambda}^{-1}\\ &+ \frac{1}{2}\sum_{i=1}^{M}(\mathrm{C}_i-\mu)(\mathrm{C}_i-\mu)^T\\ &+ \frac{\kappa}{2}(\mu-\mu_0)(\mu-\mu_0)^T\\ &+ \frac{1}{2}\nu\mathbf{\Sigma}_0.\end{aligned}$$

Equating to zero, this yields the unique minimizer

$$\mathbf{\Lambda}^{-1} = \frac{\nu\mathbf{\Sigma}_0+\kappa(\mu-\mu_0)(\mu-\mu_0)^T+\sum_{i=1}^{M}(\mathrm{C}_i-\mu)(\mathrm{C}_i-\mu)^T}{\nu+M-n}. \tag{29}$$

□

B. Convergence of the alternate minimization scheme

We study in the following the convergence of this alternate minimization algorithm. To show the main convergence result, we need the following lemma

Lemma 1. *The function f is coercive on $\mathbb{R}^{n(M+1)}\times S_n^{++}(\mathbb{R})$.*

Proof. We need to show that

$$\lim_{\|(\{\mathrm{C}_i\},\mu,\mathbf{\Lambda})\|\to+\infty} f(\{\mathrm{C}_i\},\mu,\mathbf{\Lambda}) = +\infty.$$

Now, $\|(\{\mathrm{C}_i\},\mu,\mathbf{\Lambda})\|\to+\infty$ if and only if $\|\mathrm{C}_i\|\to+\infty$ or $\|\mu\|\to+\infty$ or $\|\mathbf{\Lambda}\|\to+\infty$.

The matrix $\mathbf{\Lambda}$ being positive-definite, the terms $\frac{1}{2}\sum_{i=1}^{M}(\mathrm{C}_i-\mu)^T\mathbf{\Lambda}(\mathrm{C}_i-\mu)$ and $\frac{\kappa}{2}(\mu-\mu_0)^T\mathbf{\Lambda}(\mu-\mu_0)$ are both positive. Thus

$$\begin{aligned}f(\{\mathrm{C}_i\},\mu,\mathbf{\Lambda}) &\geq -\frac{\nu-n+M}{2}\log|\mathbf{\Lambda}|\\ &+ \frac{1}{2}\mathrm{trace}[\nu\mathbf{\Sigma}_0\mathbf{\Lambda}].\end{aligned}$$

Now, this function of $\mathbf{\Lambda}$ is convex and coercive on $S_n^{++}(\mathbb{R})$, which implies that $f(\{\mathrm{C}_i\},\mu,\mathbf{\Lambda})\to+\infty$ as soon as $\|\mathbf{\Lambda}\|\to+\infty$. It also follows that the previous function of $\mathbf{\Lambda}$ as a global minimum that we denote by $m_{\mathbf{\Lambda}}$. We can now write

$$\begin{aligned}f(\{\mathrm{C}_i\},\mu,\mathbf{\Lambda}) \geq m_{\mathbf{\Lambda}} &+ \frac{1}{2}\sum_{i=1}^{M}(\mathrm{C}_i-\mu)^T\mathbf{\Lambda}(\mathrm{C}_i-\mu)\\ &+ \frac{\kappa}{2}(\mu-\mu_0)^T\mathbf{\Lambda}(\mu-\mu_0)\end{aligned}$$

and this function of $(\{\mathrm{C}_i\},\mu)$ clearly tends towards $+\infty$ as soon as $\|\mathrm{C}_i\|\to+\infty$ or $\|\mu\|\to+\infty$. □

Proposition 6. *The sequence $f(\{\mathrm{C}_i^l\},\mu^l,\mathbf{\Lambda}^l)$ converges monotonically when $l\to+\infty$.*

Proof. The sequence $f(\{\mathrm{C}_i^l\},\mu^l,\mathbf{\Lambda}^l)$ obviously decreases at each step by construction. The coercivity and continuity of f imply that this sequence is also bounded from below, and thus converges. □

Proposition 7. *The sequence $\{\{\mathrm{C}_i^l\},\mu_l,\mathbf{\Lambda}^l\}$ generated by the alternate minimization scheme has a subsequence that converges to a stationnary point of f.*

Proof. The function f being coercive, the fact that $f(\{\mathrm{C}_i^l\},\mu^l,\mathbf{\Lambda}^l)$ converges implies that the sequence $\{\{\mathrm{C}_i^l\},\mu_l,\mathbf{\Lambda}^l\}$ is bounded. It follows that it has at least one accumulation point $(\{\mathrm{C}_i^\star\},\mu^\star,\mathbf{\Lambda}^\star)$ and that there exists a strictly increasing sequence $(l_k)_{k\in\mathbb{N}}$ of integers such that $\{\{\mathrm{C}_i^{l_k}\},\mu^{l_k},\mathbf{\Lambda}^{l_k}\}_{k\in\mathbb{N}}$ converges to $\{\{\mathrm{C}_i^\star\},\mu^\star,\mathbf{\Lambda}^\star\}$.

Now, we can show that this accumulation point is a partial optimum of f, *i.e.* that $f(\{\mathrm{C}_i^\star\},\mu^\star,.)$ attains its minimum at $\mathbf{\Lambda}^*$ and $f(.,.,\mathbf{\Lambda}^*)$ attains its minimum at $(\{\mathrm{C}_i^\star\},\mu^\star)$. By construction,

$$f(\{\mathrm{C}^{l_k}\},\mu^{l_k},\mathbf{\Lambda}^{l_k}) \leq f(\{\mathrm{C}^{l_k}\},\mu^{l_k},\mathbf{\Lambda}),\quad \forall\mathbf{\Lambda}\in S_n^{++}(\mathbb{R})$$

which implies by continuity of f that

$$f(\{\mathrm{C}^*\},\mu^*,\mathbf{\Lambda}^*) = \min_{\mathbf{\Lambda}\in S_n^{++}(\mathbb{R})} f(\{\mathrm{C}^*\},\mu^*,\mathbf{\Lambda}).$$

Let G denote the function

$$\begin{aligned}G(\mathrm{C},\mu,\mathbf{\Lambda}) &= (\mathrm{C}',\mu',\mathbf{\Lambda}') \quad \text{with}\\ (\mathrm{C}',\mu') &= \min_{(\mathrm{C},\mu)} f(\mathrm{C},\mu,\mathbf{\Lambda})\\ \mathbf{\Lambda}' &= \min_{\mathbf{\Lambda}} f(\mathrm{C}',\mu',\mathbf{\Lambda}).\end{aligned}$$

The alternate minimization scheme consists in iterating G at each iteration. From Equations (20), (21), and (27), we see that G is explicit and continuous. Since $\{l_k\}_{k\in\mathbb{N}}$ is strictly increasing, for each $k\in\mathbb{N}^*$, $l_k\geq l_{k-1}+1$. The sequence $\{f(\{\mathrm{C}_i^l\},\mu^l,\mathbf{\Lambda}^l)\}_{l\in\mathbb{N}}$ decreases, so

$$\begin{aligned}f(G(\mathrm{C}^{l_{k-1}},\mu^{l_{k-1}},\mathbf{\Lambda}^{l_{k-1}})) &= f(\mathrm{C}^{l_{k-1}+1},\mu^{l_{k-1}+1},\mathbf{\Lambda}^{l_{k-1}+1})\\ &\geq f(\mathrm{C}^{l_k},\mu^{l_k},\mathbf{\Lambda}^{l_k})\\ &\geq f(G(\mathrm{C}^{l_k},\mu^{l_k},\mathbf{\Lambda}^{l_k})).\end{aligned}$$

Taking the limit when $i\to+\infty$ and using the continuity of G, it follows that

$$f(G(\mathrm{C}^*,\mu^*,\mathbf{\Lambda}^*)) = f(\mathrm{C}^*,\mu^*,\mathbf{\Lambda}^*).$$

Now, writing $(\mathrm{C}^{**},\mu^{**},\mathbf{\Lambda}^{**}) = G(\mathrm{C}^*,\mu^*,\mathbf{\Lambda}^*)$, we get

$$\begin{aligned}f(\mathrm{C}^*,\mu^*,\mathbf{\Lambda}^*) &\geq \min_{(\mathrm{C},\mu)} f(\mathrm{C},\mu,\mathbf{\Lambda}^*) = f(\mathrm{C}^{**},\mu^{**},\mathbf{\Lambda}^*)\\ &\geq \min_{\mathbf{\Lambda}} f(\mathrm{C}^{**},\mu^{**},\mathbf{\Lambda}) = f(\mathrm{C}^{**},\mu^{**},\mathbf{\Lambda}^{**}).\end{aligned}$$

We can conclude that all these terms are equal and in particular

$$f(\mathrm{C}^*,\mu^*,\mathbf{\Lambda}^*) = f(\mathrm{C}^{**},\mu^{**},\mathbf{\Lambda}^*) = \min_{(\mathrm{C},\mu)} f(\mathrm{C},\mu,\mathbf{\Lambda}^*).$$

Thus, the accumulation point $(\{\mathrm{C}_i^\star\}, \mu^\star, \mathbf{\Lambda}^\star)$ is a partial optimum of f and since f is differentiable, it is also a stationnary point of f. □

APPENDIX B
AFFINE RISK MINIMIZER

Proposition 8. *Assume that the noise has zero mean and is not correlated to the signal* C_i*. Then, the affine estimator* $\tilde{\mathrm{C}}_i$ *that minimizes the Bayes risk* $\mathbb{E}[(\tilde{\mathrm{C}}_i - \mathrm{C}_i)^2]$ *is given by*

$$\tilde{\mathrm{C}}_i = \mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T + \mathbf{\Sigma}_{\mathrm{N}_i})^{-1}(\mathrm{Z}_i - \boldsymbol{D}_i\mu) + \mu. \quad (30)$$

Proof. Let us first consider the case $\mu = 0$. If we consider linear estimators only, we look for the matrix $\tilde{\mathbf{W}}$ that verifies

$$\tilde{\mathbf{W}} = \arg\min_{\mathbf{W}} \mathbb{E}[(\mathbf{W}\mathrm{Z}_i - \mathrm{C}_i)^2]. \quad (31)$$

Hence, $\tilde{\mathbf{W}}$ must verify

$$\mathbb{E}[(\tilde{\mathbf{W}}\mathrm{Z}_i - \mathrm{C}_i)\mathrm{Z}_i^T] = 0, \quad (32)$$

and we have

$$\tilde{\mathbf{W}} = \mathbb{E}[\mathrm{C}_i\mathrm{Z}_i^T](\mathbb{E}[\mathrm{Z}_i\mathrm{Z}_i^T])^{-1}. \quad (33)$$

Since the noise N_i has zero mean and is not correlated to the signal C_i, the element (p, q) of matrix $\mathbb{E}[\mathrm{C}_i\mathrm{Z}_i^T]$ is given by

$$\begin{aligned}\mathbb{E}[\mathrm{C}_i\mathrm{Z}_i^T]_{p,q} &= \mathbb{E}[\mathrm{C}_i(\boldsymbol{D}_i\mathrm{C}_i + \mathrm{N}_i)^T]_{p,q} && (34)\\ &= \mathbb{E}[\mathrm{C}_i^p(\boldsymbol{D}_i\mathrm{C}_i)_q + \mathrm{C}_i^p\mathrm{N}_i^q] && (35)\\ &= (\mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T)_{p,q}. && (36)\end{aligned}$$

Also, the element (p, q) of matrix $\mathbb{E}[\mathrm{Z}_i\mathrm{Z}_i^T]$ is given by

$$\begin{aligned}\mathbb{E}[\mathrm{Z}_i\mathrm{Z}_i^T]_{p,q} &= \mathbb{E}[(\boldsymbol{D}_i\mathrm{C}_i + \mathrm{N}_i)(\boldsymbol{D}_i\mathrm{C}_i + \mathrm{N}_i)^T]_{p,q} && (37)\\ &= \mathbb{E}[(\boldsymbol{D}_i\mathrm{C}_i)_p(\boldsymbol{D}_i\mathrm{C}_i)_q^T + (\boldsymbol{D}_i\mathrm{C}_i)_p(\mathrm{N}_i)_q^T && (38)\\ &\quad + (\mathrm{N}_i)_p(\boldsymbol{D}_i\mathrm{C}_i)_q + (\mathrm{N}_i)_p(\mathrm{N}_i)_q^T] && (39)\\ &= (\boldsymbol{D}_i\mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T)_{p,q} + (\mathbf{\Sigma}_{\mathrm{N}_i})_{p,q}. && (40)\end{aligned}$$

Hence we have,

$$\tilde{\mathbf{W}} = \mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T + \mathbf{\Sigma}_{\mathrm{N}_i})^{-1}. \quad (41)$$

In the general case where $\mu \neq 0$, we can always consider the centered version of the patches $(Z_i - \boldsymbol{D}_i\mu)$ and apply the previous result. Therefore, the estimator of C_i that minimizes the risk function $\mathbb{E}[(\tilde{\mathrm{C}} - \mathrm{C}_i)^2]$ among all affine estimators under Model (2), is given by (30). □

ACKNOWLEDGMENT

The authors would like to thank the authors of [25], [16] and [24] of for kindly providing their code. This work has been partially funded by the French Research Agency (ANR) under grant nro ANR-14-CE27-001 (MIRIAM).

REFERENCES

[1] Henri Maitre, *From Photon to Pixel: The Digital Camera Handbook*. Wiley-ISTE, 2015.

[2] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.

[3] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Transactions on Image processing*, vol. 12, no. 11, pp. 1338–1351, 2003.

[4] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image processing*, vol. 15, no. 12, pp. 3736–3745, 2006.

[5] A. A. Efros and T. K. Leung, "Texture Synthesis by Non-Parametric Sampling," in *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, 1999, pp. 1033–.

[6] A. Buades, B. Coll, and J. M. Morel, "A Review of Image Denoising Algorithms, with a New One," *SIAM Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.

[7] G. Peyré, "Image processing with nonlocal spectral bases," *Multiscale Modeling & Simulation*, vol. 7, no. 2, pp. 703–730, 2008.

[8] S. Lyu and E. Simoncelli, "Modeling Multiscale Subbands of Photographic Images with Fields of Gaussian Scale Mixtures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 4, pp. 693–706, 2009.

[9] P. Chatterjee and P. Milanfar, "Patch-Based Near-Optimal Image Denoising," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1635–1649, 2012.

[10] M. Lebrun, A. Buades, and J. Morel, "A nonlocal bayesian image denoising algorithm," *SIAM Journal on Imaging Sciences*, vol. 6, no. 3, pp. 1665–1688, 2013. [Online]. Available: http://epubs.siam.org/doi/abs/10.1137/120874989

[11] Y.-Q. Wang and J.-M. Morel, "SURE Guided Gaussian Mixture Image Denoising," *SIAM Journal on Imaging Sciences*, vol. 6, no. 2, pp. 999–1034, 2013.

[12] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, 2011, pp. 479–486.

[13] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: From gaussian mixture models to structured sparsity," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2481–2499, 2012.

[14] M. Lebrun, M. Colom, A. Buades, and J. Morel, "Secrets of image denoising cuisine," *Acta Numerica*, vol. 21, no. 1, pp. 475–576, 2012.

[15] Y.-Q. Wang, "The Implementation of SURE Guided Piecewise Linear Image Denoising," *Image Processing On Line*, vol. 3, pp. 43–67, 2013.

[16] ——, "E-PLE: an Algorithm for Image Inpainting," *Image Processing On Line*, vol. 2013, pp. 271–285, 2013.

[17] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian data analysis*. Chapman & Hall/CRC Boca Raton, FL, USA, 2014, vol. 2.

[18] N. Dobigeon, J.-Y. Tourneret, and C.-I. Chang, "Semi-supervised linear spectral unmixing using a hierarchical bayesian model for hyperspectral imagery," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 2684–2695, 2008.

[19] H. Raiffa and R. Schlaifer, *Applied statistical decision theory*. Division of Research, Graduate School of Business Administration, Harvard University Boston, 1961.

[20] C. Aguerrebere, J. Delon, Y. Gousseau, and P. Musé, "Study of the digital camera acquisition process and statistical modeling of the sensor raw data," *Technical report hal-00733538v1*, 2012.

[21] J. Gorski, F. Pfeuffer, and K. Klamroth, "Biconvex sets and optimization with biconvex functions: A survey and extensions," *Mathematical Methods of Operations Research*, vol. 66, no. 3, pp. 373–407, 2007.

[22] P. Arias, V. Caselles, and G. Facciolo, "Analysis of a Variational Framework for Exemplar-Based Image Inpainting," *SIAM MMS*, vol. 10, no. 2, pp. 473–514, jan 2012. [Online]. Available: http://epubs.siam.org/doi/abs/10.1137/110848281

[23] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3d transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, 2007.

[24] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," http://people.csail.mit.edu/danielzoran/epllcode.zip, accessed: 29/09/2014.

[25] M. Lebrun, A. Buades, and J.-M. Morel, "Implementation of the "Non-Local Bayes" (NL-Bayes) Image Denoising Algorithm," *Image Processing On Line*, vol. 3, pp. 1–42, 2013.

[26] C. Aguerrebere, J. Delon, Y. Gousseau, and P. Musé, "Best Algorithms for HDR Image Generation. A Study of Performance Bounds," *SIAM Journal on Imaging Sciences*, vol. 7, no. 1, pp. 1–34, 2014. [Online]. Available: http://epubs.siam.org/doi/abs/10.1137/120891952

[27] C. Aguerrebere, A. Almansa, J. Delon, Y. Gousseau, and P. Musé, "Single shot high dynamic range imaging using piecewise linear estimators," in *Proceedings of IEEE International Conference on Computational Photography (ICCP)*, 2014.

[28] A. Buades, B. Coll, and J.-M. Morel, "Non-Local Means Denoising," *Image Processing On Line*, vol. 1, 2011.

[29] P. E. Debevec and J. Malik, "Recovering High Dynamic Range Radiance Maps from Photographs," in *Proceedings of the Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH, 1997, pp. 369–378. [Online]. Available: http://dx.doi.org/10.1145/258734.258884

[30] M. Granados, B. Ajdin, M. Wand, C. Theobalt, H. P. Seidel, and H. P. A. Lensch, "Optimal HDR reconstruction with linear digital cameras," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 215–222.

[31] C. Aguerrebere, J. Delon, Y. Gousseau, and P. Musé, "Simultaneous HDR image reconstruction and denoising for dynamic scenes," in *Proceedings of IEEE International Conference on Computational Photography (ICCP)*, 2013, pp. 1–11.

[32] D. Sidibé, W. Puech, and O. Strauss, "Ghost detection and removal in high dynamic range images," in *Proceedings of the European Signal Processing Conference*, 2009.

[33] S. Nayar and T. Mitsunaga, "High Dynamic Range Imaging: Spatially Varying Pixel Exposures," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, Jun 2000, pp. 472–479.

[34] P. Sen, N. K. Kalantari, M. Yaesoubi, S. Darabi, D. B. Goldman, and E. Shechtman, "Robust patch-based HDR reconstruction of dynamic scenes," *ACM Transactions on Graphics*, vol. 31, no. 6, pp. 203:1–203:11, 2012.

[35] F. Yasuma, T. Mitsunaga, D. Iso, and S. Nayar, "Generalized Assorted Pixel Camera: Post-Capture Control of Resolution, Dynamic Range and Spectrum," *IEEE Transactions on Image Processing*, vol. 99, Mar 2010.

[36] M. Schöberl, A. Belz, A. Nowak, J. Seiler, A. Kaup, and S. Foessel, "Building a high dynamic range video sensor with spatially nonregular optical filtering," in *Proceedings of Proc. SPIE 8499, Applications of Digital Image Processing XXXV*, vol. 8499, 2012, pp. 84 990C–84 990C–11.

[37] M. Schöberl, A. Belz, J. Seiler, S. Foessel, and A. Kaup, "High dynamic range video by spatially non-regular optical filtering," in *Proceedings of IEEE International Conference on Image Processing (ICIP)*, 2012, pp. 2757–2760.

[38] S. Hasinoff, F. Durand, and W. Freeman, "Noise-optimal capture for high dynamic range photography," http://people.csail.mit.edu/hasinoff/hdrnoise/, last accessed: 17/03/2014.

[39] J. Hamilton and J. Adams, "Adaptive color plan interpolation in single sensor color electronic camera," US Patent 5,629,734, 1997.

[40] R. Mantiuk, S. Daly, and L. Kerofsky, "Display adaptive tone mapping," *ACM Transactions on Graphics*, vol. 27, no. 3, pp. 68:1–68:10, 2008.

[41] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[42] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Technical University of Denmark, Tech. Rep., 2012.