Inverse Problems in Imaging: a Hyperprior Bayesian Approach

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

Inverse Problems in Imaging: a Hyperprior Bayesian Approach

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

***Abstract*—Patch models have proven successful to solve a variety of inverse problems in image restoration. Recent methods, combining patch models with a Bayesian approach, achieve state-of-the-art results in several restoration problems. Different strategies are followed to determine the patch models, such as a fixed number of models to describe all image patches or a locally determined model for each patch. Local model estimation has proven very powerful for image denoising, but it becomes seriously ill-posed for other inverse problems such as interpolation of random missing pixels or zooming. In this work, we present a new framework for image restoration that makes it possible to use local priors for these more general inverse problems. To this aim, we make use of a hyperprior on the model parameters which overcomes the ill-posedness of the local estimation and yields state-of-the-art results in problems such as interpolation, denoising and zooming. Moreover, taking advantage of the generality of the framework, we present an application to the generation of high dynamic range (HDR) images from a single snapshot. Experiments conducted on synthetic and real data show the effectiveness of the proposed approach.**

***Index Terms*—Non-local patch-based restoration, Bayesian restoration, Maximum a Posteriori, Gaussian Mixture Models, hyper-prior, conjugate distributions.**

I. Introduction

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

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

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 others [1].

Inspired by the patch-based approach for texture synthesis proposed by Efros and Leung [2], Buades et al. [3] 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. A major step forward in fully exploiting patches potential was taken with the introduction of patches 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 [4], [5], [6], [7], while others propose a more general framework for the solution of image inverse problems [8], [9], including for instance inpainting, deblurring and zooming. The work by Lebrun et al. [10], [6] presents a thorough and very interesting 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. [9], namely the piece-wise linear estimators (PLE), and the non-local Bayes (NLB) algorithm by Lebrun et al. [6]. 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, so the number of classes is not chosen a priori (one class per patch). NLB outperforms PLE in the denoising task [11], mostly due to its continuous classification model. Nevertheless, PLE obtains state-of-the-art results in other applications such as interpolation of missing pixels, deblurring and zooming. In particular, PLE yields very good results in interpolation of random missing pixels with high masking rates. A variant of PLE for inpainting is proposed by Wang [12] (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 [8] (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 [7] 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

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 Department of Signal and Image Processing, LTCI CNRS, Télécom ParisTech, 75634 PARIS Cedex 13, 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)

[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.

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. For the task of image denoising, relying on local (continuous) priors [6], [7] has proven more powerful and more accurate than relying on a limited number of Gaussian models [8], [9]. However, to the best of our knowledge, such continuous priors remain very difficult to estimate for more general restoration problems, especially when the image degradations involve missing pixels. The main contribution of this work is to propose a robust framework enabling the use of such Gaussian local priors for solving general restoration problems, by drawing on what is known as hyperprior in Bayesian statistics.

Following [6], [7], we propose to model image patches according to a Gaussian prior, whose parameters will be estimated locally from similar patches. The main challenge with 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 interpolation of random missing pixels with a masking rate of 70%, the patches used for the estimation of μ and $\mathbf{\Sigma}$ will lack 70% of the pixels, thus making the estimation problem very ill-posed. In order to tackle this problem, we include prior knowledge on the model parameters making use of what is known as 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. The use of a hyperprior allows to estimate 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. Finally, image patches are restored using the maximum a posteriori (MAP) estimator with the computed Gaussian model. Experiments conducted on both synthetic and real data, show the state-of-the-art results obtained by the proposed approach in various problems such as interpolation, denoising and zooming.

This general framework can also be applied to the generation of high dynamic range (HDR) images. Due to physical limitations, current digital sensors cannot capture faithful representations of high dynamic range 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 with Section III presenting 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.

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

II. Hyperprior Bayesian Estimator

We describe here the proposed restoration method, called Hyperprior Bayesian Estimator (HBE). Following the idea of the recent bayesian approach [6], we assume a Gaussian prior for image patches whose parameters μ and $\mathbf{\Sigma}$ are estimated locally 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. The maximization alternates two steps : first the log-likelihood is maximized in μ and $\mathbf{\Sigma}$ by combining similar patches and the hyperprior on them, and then the patch is restored under the Gaussian prior defined by μ and $\mathbf{\Sigma}$. Figure 1 shows a diagram of the proposed iterative approach.

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.

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 random missing pixels (inpainting). As for the noise term N_i, we assume that it is well modeled by a Gaussian distribution $\mathrm{N}_i \sim \mathcal{N}(0, \mathbf{\Sigma}_{N_i})$. The distribution of Z_i given C_i can thus be written

$$p(\mathrm{Z}_i \mid \mathrm{C}_i) \sim \mathcal{N}(\boldsymbol{D}_i\mathrm{C}_i, \mathbf{\Sigma}_{N_i}) \tag{3}$$

$$\propto |\mathbf{\Sigma}_{N_i}^{-1}|^{\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathrm{Z}_i - \boldsymbol{D}_i\mathrm{C}_i)^T \mathbf{\Sigma}_{N_i}^{-1}(\mathrm{Z}_i - \boldsymbol{D}_i\mathrm{C}_i)\right). \tag{4}$$

In this noise model, the matrix $\mathbf{\Sigma}_{N_i}$ is only supposed to be diagonal (the noise is uncorrelated). It can be constituted of a constant variance, 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 represent the problem of high dynamic range (HDR) estimation from a single image. It also includes a wide range of restoration problems generally studied in the literature [3].

[3]Observe that this model also includes the particular case of multiplicative noise.

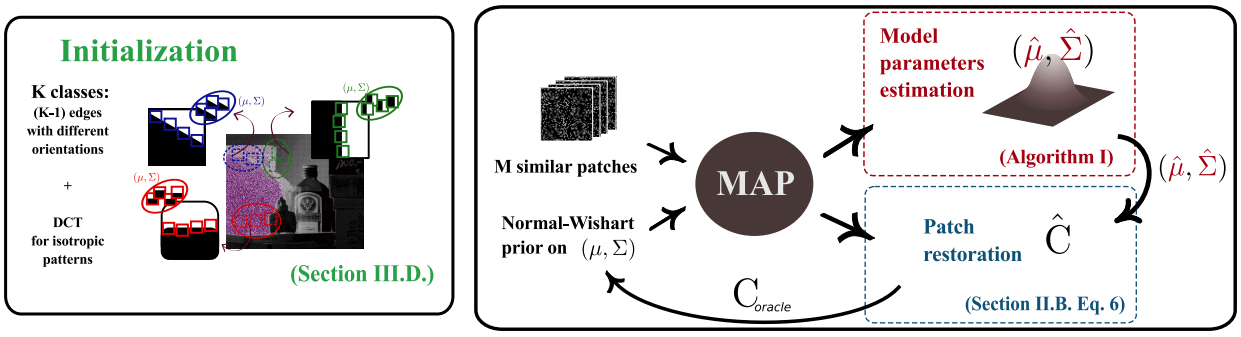


Fig. 1. Diagram of the proposed iterative approach with the corresponding initialization stage.

B. Joint Maximum A Posteriori

Following the direction of Bayesian patch-based methods, we assume in this paper a local Gaussian prior around each patch, with unknown mean μ and covariance matrix $\mathbf{\Sigma}$:

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

In the literature, the local parameters μ and $\mathbf{\Sigma}$ are either estimated from a set of similar patches [6], or chosen from a finite set of precomputed parameters [9]. 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 local Gaussian parameters from a set of similar patches. To this aim, we make use of a hyperprior on μ and $\mathbf{\Sigma}$. On the one hand, using similar patches gives a spatially adaptive or local characterization of the patch [6]. On the other hand, including the hyperprior makes the parameter estimation more robust, 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 $\mathbf{\Lambda} = \mathbf{\Sigma}^{-1}$ instead of the covariance matrix $\mathbf{\Sigma}$. As it is usual when considering hyperpriors, we assume parameters μ and $\mathbf{\Sigma}$ follow a conjugate distribution. In our case, that boils down to assuming a Normal-Wishart[4] prior for the couple $(\mu, \mathbf{\Lambda})$,

$$\begin{aligned} p(\mu, \mathbf{\Lambda}) &= \mathcal{N}(\mu|\mu_0, (\kappa\mathbf{\Lambda})^{-1})\mathcal{W}(\mathbf{\Lambda}|(\nu\mathbf{\Sigma}_0)^{-1}, \nu) \\ &\propto |\mathbf{\Lambda}|^{1/2} \exp\left(-\frac{\kappa}{2}(\mu-\mu_0)\mathbf{\Lambda}(\mu-\mu_0)^T\right) \\ &\quad |\mathbf{\Lambda}|^{(\nu-n-1)/2} \exp\left(-\frac{1}{2}\mathrm{tr}(\nu\mathbf{\Sigma}_0\mathbf{\Lambda})\right), \end{aligned} \tag{5}$$

where μ_0 is a prior on μ, $\mathbf{\Sigma}_0$ is a prior on $\mathbf{\Sigma}$, $\kappa > 0$ and $\nu > n-1$ are a scale parameter and the degrees of freedom of the Normal-Wishart respectively.

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 and follow the same Gaussian model, we can compute the joint maximum a posteriori

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

In this product, the first term is given by the noise model (Section II-A), the second one is the Gaussian prior on the set of patches $\{\mathrm{C}_i\}_i$ and the third one is the hyperprior (5).

Proposition 1. *Assume that (μ, Λ) follow the Normal-Wishart distribution (5), that $\mathrm{C}_1, \dots, \mathrm{C}_n | \mu, \Lambda$ are independent realizations of $\mathcal{N}(\mu, \Lambda^{-1})$, and that $\mathrm{Z}_1, \dots, \mathrm{Z}_n | \mathrm{C}_1, \dots, \mathrm{C}_n$, are independent realizations of the respective distributions $\mathcal{N}(\boldsymbol{D}_i\mathrm{C}_i, \mathbf{\Sigma}_{N_i})$. Then, under the assumption that the noise covariance matrix $\mathbf{\Sigma}_{N_i}$ is independent from C_i for each i, if the values $(\{\widehat{\mathrm{C}_i}\}, \widehat{\mu}, \widehat{\mathbf{\Lambda}})$ maximize the unified log-likelihood $\log p(\{\mathrm{C}_i\}_i, \mu, \mathbf{\Lambda} \mid \{\mathrm{Z}_i\}_i)$, they must satisfy the following equations*

$$\widehat{\mathrm{C}_i} = \widehat{\mathbf{\Lambda}}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\widehat{\mathbf{\Lambda}}^{-1}\boldsymbol{D}_i^T + \mathbf{\Sigma}_{N_i})^{-1}(\mathrm{Z}_i - \boldsymbol{D}_i\widehat{\mu}) + \widehat{\mu}. \tag{6}$$

$$\widehat{\mu} = \frac{M\overline{C} + \kappa\mu_0}{M+\kappa}, \quad \textit{with } \overline{C} = \frac{1}{M}\sum_{i=1}^{M}\widehat{\mathrm{C}_i}. \tag{7}$$

and

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

Proof. See Appendix A. □

The expression of $\widehat{\mathrm{C}_i}$ in (6) is obtained under the hypothesis that the noise covariance matrix Σ_{N_i} does not depend on C_i. Observe that it can be shown 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 $\mathbb{E}[(\tilde{\mathrm{C}}_i - \mathrm{C}_i)^2]$.

Proposition 2. *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*

$$\tilde{\mathrm{C}}_i = \mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T + \mathbf{\Sigma}_{\mathrm{N}_i})^{-1}(\mathrm{Z}_i - \boldsymbol{D}_i\mu) + \mu. \tag{9}$$

[4]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 [13].

Proof. See Appendix A, paragraph *(f)*. □

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 $N_i = f(C_i)\varepsilon_i$ where ε_i is independent from C_i. This includes very different noise models, such as the one used for camera sensor acquisition for instance [1].

a) Interpretation of the MAP in μ, Λ: From (7), we find that the MAP estimator of μ can be seen as 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 covariance matrix estimated from the restored patches

$$\sum_{i=1}^{M}(\widehat{\mathrm{C}_i} - \widehat{\mu})(\widehat{\mathrm{C}_i} - \widehat{\mu})^T, \tag{10}$$

the covariance imposed by $\widehat{\mu}$ (since μ and $\mathbf{\Lambda}$ are not independent),

$$(\widehat{\mu} - \mu_0)(\widehat{\mu} - \mu_0)^T, \tag{11}$$

and the prior on $\mathbf{\Lambda}$ (or equivalently on $\mathbf{\Sigma}^{-1}$),

$$\mathbf{\Sigma}_0^{-1}. \tag{12}$$

If we inject expression (6) of the $\widehat{\mathrm{C}_i}$'s at the maximum into the previous expressions of $\widehat{\mu}$ and $\widehat{\mathbf{\Lambda}}$, we obtain

$$\widehat{\mu} = (\kappa \mathrm{Id} + \sum_{i=1}^{M} A_i \boldsymbol{D}_i)^{-1}(\sum_{i=1}^{M} A_i \mathrm{Z}_i + \kappa\mu_0), \tag{13}$$

and

$$\widehat{\mathbf{\Lambda}}^{-1} = \frac{\nu\mathbf{\Sigma}_0 + \kappa(\widehat{\mu} - \mu_0)(\widehat{\mu} - \mu_0)^T + \sum_{i=1}^{M} A_i(\mathrm{Z}_i - \boldsymbol{D}_i\widehat{\mu})(\mathrm{Z}_i - \boldsymbol{D}_i\widehat{\mu})^T A_i^T}{\nu + M - n} \tag{14}$$

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

b) Practical estimation: Since (13) depends on $\mathbf{\Lambda}$ and (14) depends on μ and $\mathbf{\Lambda}$, those are not closed-forms for the estimators. Hence, we propose to use an iterative approach to compute the parameters, as summarized in Algorithm 1. This algorithm results from the combination of two procedures. The outer loop follows from the classic EM estimation procedure for the mean and covariance (or precision) matrix. The inner one, which deals with the estimation of the precision matrix, converges if and only if the spectral norm of the precision matrix is less than one. In case this condition on the spectral norm of $\mathbf{\Lambda}$ holds, since EM and therefore the posteriors are guaranteed to converge to a local maximum, μ and $\mathbf{\Lambda}$ are ensured to converge to local maximizers. In practice, we observe that the algorithm converges after a single iteration of the outer loop with 3 to 4 iterations of the inner loop.

Algorithm 1: Computation of $\hat{\mu}$ and $\hat{\mathbf{\Lambda}}$.

Input: Z, $\boldsymbol{D}$, $\mu_0, \mathbf{\Sigma}_0, \kappa, \nu$ (see details in Section III-C)
Output: $\hat{\mu}$, $\hat{\mathbf{\Lambda}}$
Initialization: Set $\mathbf{\Lambda} = \mathbf{\Sigma}_0^{-1}$
for *it = 1 to maxIts*$_0$ **do**
  Compute $\hat{\mu}$ according to (13)
  Set $\mu = \hat{\mu}$.
  for *it = 1 to maxIts*$_1$ **do**
    Compute $\hat{\mathbf{\Lambda}}$ according to (14).
    Set $\mathbf{\Lambda} = \hat{\mathbf{\Lambda}}$.
  end
end

C. Summary of the proposed algorithm

The analysis previously presented leads to an iterative algorithm that implements the proposed approach. Two stages are alternated: the restoration step, where all patches are reconstructed, and the model estimation step, where the model parameters are updated (Figure 1). For this model estimation step, the result of the previous iteration is used as an "*oracle*". In practice, the algorithm is found to converge after 3 to 4 iterations. The procedure is summarized in Algorithm 2.

Algorithm 2: Summary of the proposed 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 $\hat{\mu}$ and $\hat{\mathbf{\Sigma}}$ following Algorithm 1.
    Restore the similar patches using (9) (see details in Section III-B).
  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 with L_2 distance to the current patch below a given threshold, which is given by a tolerance parameter times the distance to the closest neighbor. 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 ($\boldsymbol{D}_p^j = 1$) and to those originally missing and then restored ($\boldsymbol{D}_p^j = 0$). 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 according to

$$d(p,q) = \frac{\sum_{j=1}^{N}(\mathrm{c}_p^j - \mathrm{c}_q^j)^2\omega_{p,q}^j}{\sum_{j=1}^{N}\omega_{p,q}^j}, \tag{15}$$

with $\omega_{p,q}^j = 1$ if $\boldsymbol{D}_p^j = \boldsymbol{D}_q^j = 1$ and $\omega_{p,q}^j = 0.01$ otherwise [14]. With this formulation, known pixels are assigned a much higher priority than unknown ones. Variations of these weights could be explored.

B. Collaborative Filtering

The proposed method computes one Gaussian model per image patch according to Equations (13) and (14). In order to reduce the computational cost, we rely on the collaborative filtering idea previously introduced for patch-based denoising techniques [6], [15]. 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. The restoration is thus performed for all similar patches according to the computed model.

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 $\boldsymbol{\Sigma}$ using (13) and (14).

c) Setting of κ and ν: The computation of μ according to (13) combines the mean $\sum_{i=1}^{M} A_i \mathrm{Z}_i$ 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 on the prior accuracy vs. the one we have on the information provided by the similar patches. The latter improves when both M (*i.e.* the number of similar patches) and $P = \mathrm{trace}(D_i)$ (*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:

$$\kappa = M\alpha, \quad \alpha = \begin{cases} \alpha_L & \text{if } P \text{ and } M > \text{threshold} \\ \alpha_H & \text{otherwise.} \end{cases} \tag{16}$$

A similar reasoning leads to the same rule for ν,

$$\nu = M\alpha + n \tag{17}$$

where the additive 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. [5]

[5]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.

d) Setting of μ_0 and $\boldsymbol{\Sigma}_0$: Assuming an oracle image C_{oracle} is available (see details in Section II-C), 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}

$$\mu_0 = \frac{1}{M}\sum_{j=1}^{M}\tilde{\mathrm{c}}_j, \quad \boldsymbol{\Sigma}_0 = \frac{1}{M-1}\sum_{j=1}^{M}(\tilde{\mathrm{c}}_j - \mu_0)(\tilde{\mathrm{c}}_j - \mu_0)^T. \tag{18}$$

This is the same approach followed by Lebrun et al. [6] to locally estimate the patch model parameters in the case of denoising. As previously stated, the method from [6] cannot be directly applied to zooming or interpolation due to the presence of missing pixels.

D. Initialization

A good initialization is crucial since we aim at solving a non-convex problem through an iterative procedure. Yu et al. [9] propose to initialize the PLE algorithm 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.

Each covariance matrix $\boldsymbol{\Sigma}_k$, $k = 1, \ldots, K-1$, corresponds to one of $K-1$ orientations, uniformly sampled from directions zero to π. For a given orientation θ, a synthetic black-and-white image is generated and patches that touch the contour at different positions are randomly sampled from it. A covariance matrix is then computed from the sampled patches. The first eigenvector of the covariance matrix, which is almost constant, is replaced by a constant vector. This allows a class k of a given orientation to restore patches having different mean. Up to a certain gray level difference, dark or bright edges with the same orientation are correctly represented by the same class. A Gram-Schmidt orthogonalization is computed on the other eigenvectors to ensure the orthogonality of the basis. The eigenvalues of all bases are initialized with the same values, chosen to have a fast decay. At last, the DCT basis is added to represent isotropic image patterns, making a total of K classes. The mean of each class μ_k, $k = 1, \ldots, K$, is initialized to zero. The authors claim that they have found in practice that $K = 19$ classes (i.e. 18 orientations, 10 degrees apart) give a correct reconstruction and are a good compromise between performance and complexity for a patch size of 8×8. The fact that the algorithm is applied in regions of size 128×128, and therefore localized, also explains why this a priori small number of classes can be suitable to describe all image patches.

To initialize the proposed algorithm, we follow the approach by Yu et al. [9] and compute the K covariance matrices $\boldsymbol{\Sigma}_k$ as previously described. Then, each image patch z_i is reconstructed under each class $k = 1, \ldots, K$ as

$$\tilde{\mathrm{C}}_i^k = \boldsymbol{\Sigma}_k\boldsymbol{D}_i^T(\boldsymbol{D}_i\boldsymbol{\Sigma}_k\boldsymbol{D}_i^T + \boldsymbol{\Sigma}_{\mathrm{N}_i})^{-1}(\mathrm{Z}_i - \boldsymbol{D}_i\mu_k) + \mu_k. \tag{19}$$

The best suited class $\tilde{k}_i$ is chosen as the one maximizing the posterior probability of the patch $p(\mathrm{C}_i|\mathrm{z}_i, \mu_k, \boldsymbol{\Sigma}_k)$ over

k assuming $\mathrm{C} = \tilde{\mathrm{C}}_i^k$:

$$\tilde{k}_i = \arg\max_k \ln p(\mathrm{C}|\mathrm{z}_i, \mu_k, \mathbf{\Sigma}_k) \tag{20}$$

$$= \arg\min_k \left(\frac{\|\mathrm{z}_i - \boldsymbol{D}_i \tilde{\mathrm{C}}_i^k\|^2}{\sigma^2} + (\tilde{\mathrm{C}}_i^k - \mu_k)^T \mathbf{\Sigma}_k^{-1} (\tilde{\mathrm{C}}_i^k - \mu_k) \right. \tag{21}$$

$$\left. + \ln |\mathbf{\Sigma}_k| \right). \tag{22}$$

The first oracle is thus created by aggregating the estimations of all patches corresponding to the chosen classes. Figure 1 illustrates the proposed initialization and Algorithm 3 summarizes its steps.

Algorithm 3: Summary of the initialization procedure.

Input: Z, $\boldsymbol{D}$, K
Output: C_{oracle}

1. Decompose Z and $\boldsymbol{D}$ into overlapping patches.
2. Compute the covariance matrix of the K classes from synthetic images of edges plus the DCT.
3. Project all patches into the K classes using (19) and chose the best class using (22).
4. Compute the first oracle C_{oracle} aggregating the estimations of all patches for the chosen class.

IV. Image Restauration 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 denoising only).

A. Synthetic degradation

e) Interpolation: Random masks with 20%, 50% and 70% of missing pixels are applied to the tested ground-truth images. The interpolation performance of the proposed method is compared to that of PLE [9], EPLL [8] and E-PLE [12] using a patch size of 8×8 for all methods. PLE parameters are set as indicated in [9] ($\sigma = 3$, $\varepsilon = 30$, $K = 19$). We used the EPLL code provided by the authors [16] with default parameters and the E-PLE code available in [12] with the parameters set as specified in this 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.

f) 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 (see Figure 3).

g) Denoising: For the denoising task, the proposed approach should perform very similar to the state-of-the-art denoising algorithm NLB [6]. The following experiments are conducted in order to verify this. The experiments are performed with images corrupted with additive Gaussian noise with variance $\sigma^2 = 10, 30, 50, 80$. The code provided by the authors [17] automatically sets the algorithm 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 [6] (instead of using the initialization described in Section II-C) and the parameters κ and ν are large enough. In that case, μ_0 and $\mathbf{\Sigma}_0$ are prioritized in equations (13) and (14) and both algorithms are almost the same. That is what we observe in practice with $\alpha_H = \alpha_L = 100$, as exemplified 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 $\mathbf{\Sigma}_0$, as well as μ and $\mathbf{\Sigma}$ 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 (13) to the computation of $\hat{\mu}$ degrades the result compared to using μ_0 only (i.e. using a large κ).

h) 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 the Lanczos interpolation. Figure 4 shows extracts of the obtained results, the PSNR values for the extracts and the corresponding difference images with respect to the ground-truth. HBE yields a sharper reconstruction than the other methods.

B. Real data

Raw images are captured using a Canon 400D camera set to ISO 400 and exposure time 1/160 seconds. In this case, we use the complete noise model for CMOS raw pixel

		PSNR (dB)											
		HBE	PLE	EPLL	E-PLE	HBE	PLE	EPLL	E-PLE	HBE	PLE	EPLL	E-PLE
	% missing pixels	20%				50%				70%			
Interpolation	barbara	**45.57**	43.48	40.89	43.75	**39.11**	36.93	32.99	35.43	**34.69**	32.50	27.96	28.77
	boat	**41.43**	40.37	40.17	40.32	**34.92**	34.32	34.21	33.59	**31.37**	30.74	30.38	30.26
	traffic	35.66	35.53	**35.71**	35.10	30.17	30.12	**30.19**	28.86	**27.27**	27.12	27.13	26.64
Interpolation & Denoising	barbara	38.35	**38.37**	37.32	37.26	-	-	-	-	**33.34**	31.99	27.63	27.75
	boat	36.99	**37.02**	37.00	36.26	-	-	-	-	**30.61**	30.41	30.15	29.54
	traffic	34.07	34.13	**34.33**	33.50	-	-	-	-	26.99	26.98	**27.05**	26.35
	σ^2	HBE	NLB	EPLL	HBE	NLB	EPLL	HBE	NLB	EPLL	HBE	NLB	EPLL
			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: $\sigma = 3$, $\varepsilon = 30$, $K = 19$ [9], EPLL: DEFAULT PARAMETERS [16], E-PLE: PARAMETERS SET AS SPECIFIED IN [12]. PARAMETER SETTING FOR DENOISING, HBE: $\alpha_H = \alpha_L = 100$, NLB: CODE PROVIDED BY THE AUTHORS [17] AUTOMATICALLY SETS PARAMETERS FROM INPUT σ^2, EPLL: DEFAULT PARAMETERS FOR THE DENOISING EXAMPLE [16]

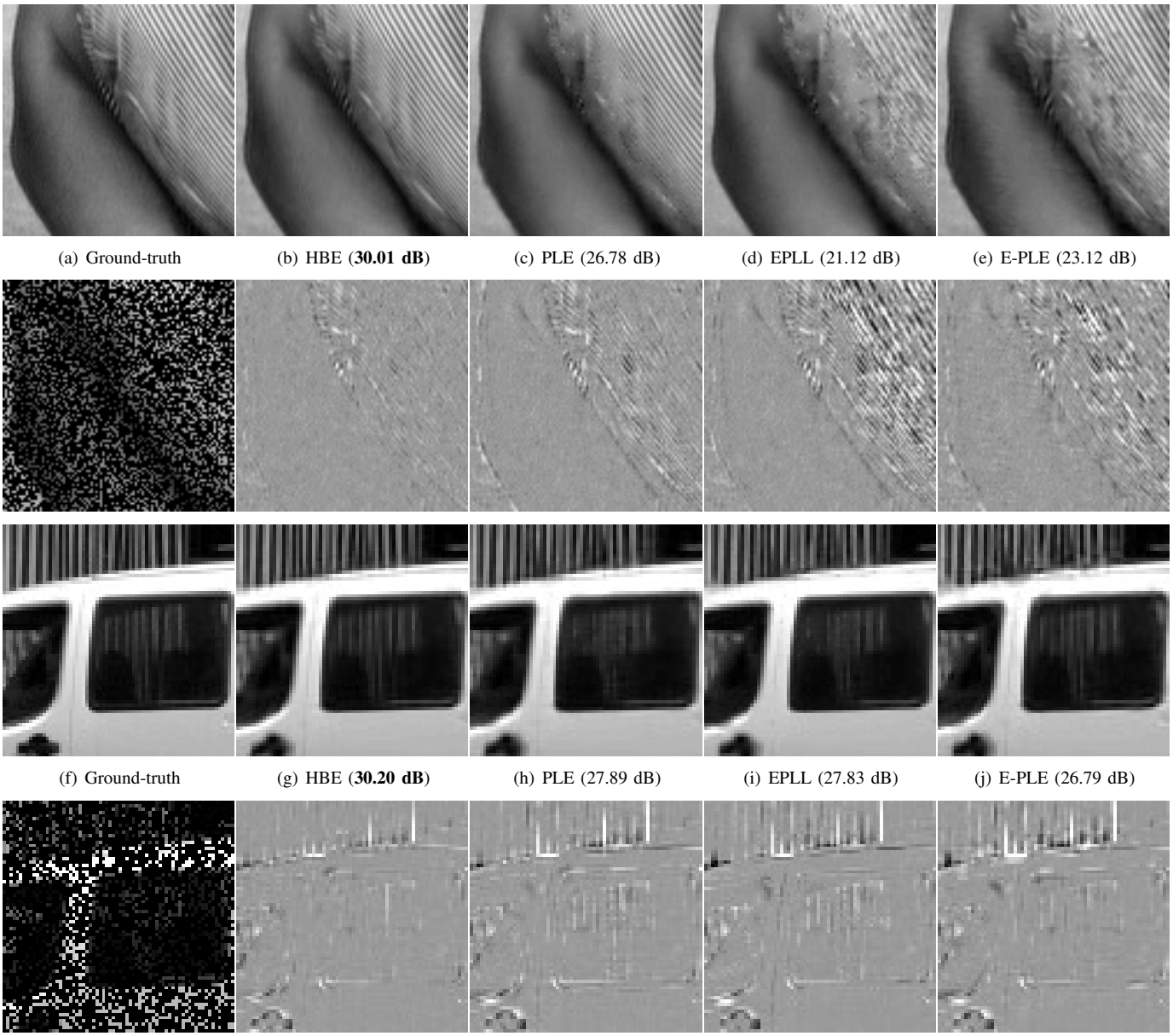
Fig. 5. **Real data.** JPEG version of the raw image used in the experiments presented in Section IV-B. The boxes show the extracts displayed in Figure 6.

values (23) introduced in Section V-A, whose parameters have to be estimated by a calibration procedure [1]. 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 [18]. The results for the EPLL and E-PLE methods are not presented 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 these 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 more sharp both in the 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 [9]), we find that 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 [19] 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,



(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)

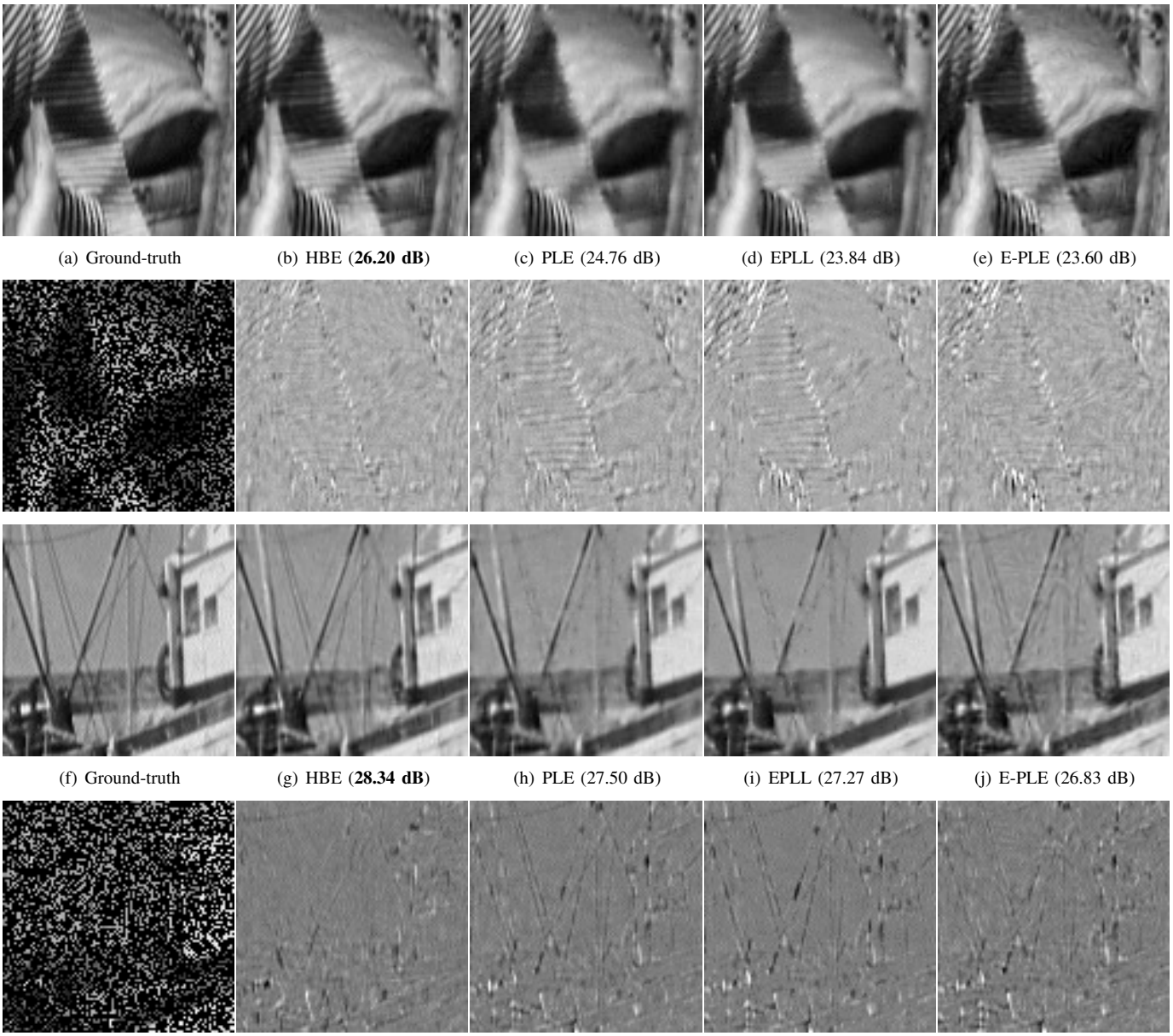
Fig. 2. **Synthetic data. Interpolation with 70% of randomly missing pixels**. **Left to right:** (first row) Ground-truth (extract of barbara), result by HBE, PLE, EPLL, E-PLE. (second row) input image, difference with respect to the ground-truth of each of the corresponding results. (third and fourth 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.

previously proven successful at describing patches [6], 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 [8]. 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 [7] claim that in the case of denoising, it is better having fewer models that are updated with the image patches (as in PLE) 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, it is here harder to tell if the improvement is due to the local model estimation performed from the similar patches or it is due to the different restoration strategies followed by these methods.

V. High Dynamic Range image generation 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 with that which can be captured using a standard digital camera. The range of luminosity which a standard digital camera can capture is often not enough to



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

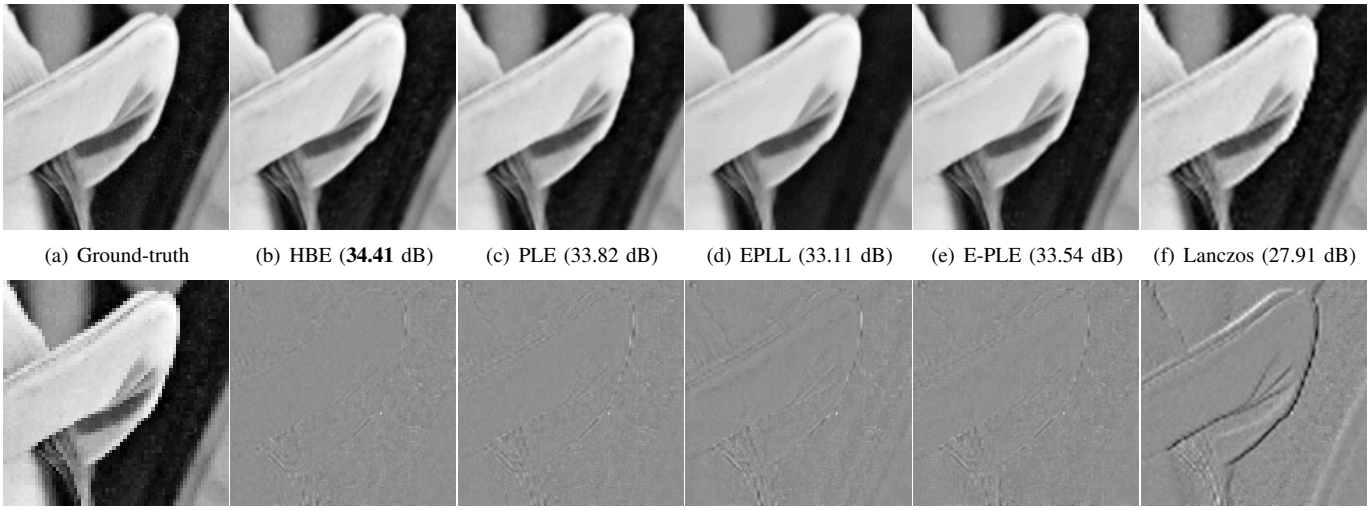
(f) Ground-truth (g) HBE (**28.34 dB**) (h) PLE (27.50 dB) (i) EPLL (27.27 dB) (j) E-PLE (26.83 dB)

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.

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 [20], [21]. However, several problems arise when either the camera or the elements in the scene move [22], [23].

An alternative to the HDR from multiple frames was introduced by Nayar and Mitsunaga in [24]. They propose to perform HDR imaging from a single image using spatially varying pixel exposures (SVE). An optical mask with spatially varying transmittance (see Figure 8) is placed adjacent to a conventional image sensor, thus controlling the amount of light that reaches each pixel. 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 allows HDR imaging from a single image, thus avoiding the need for 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 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 [25], [22]. A problem encountered by this approach is the need for inpainting saturated



(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.

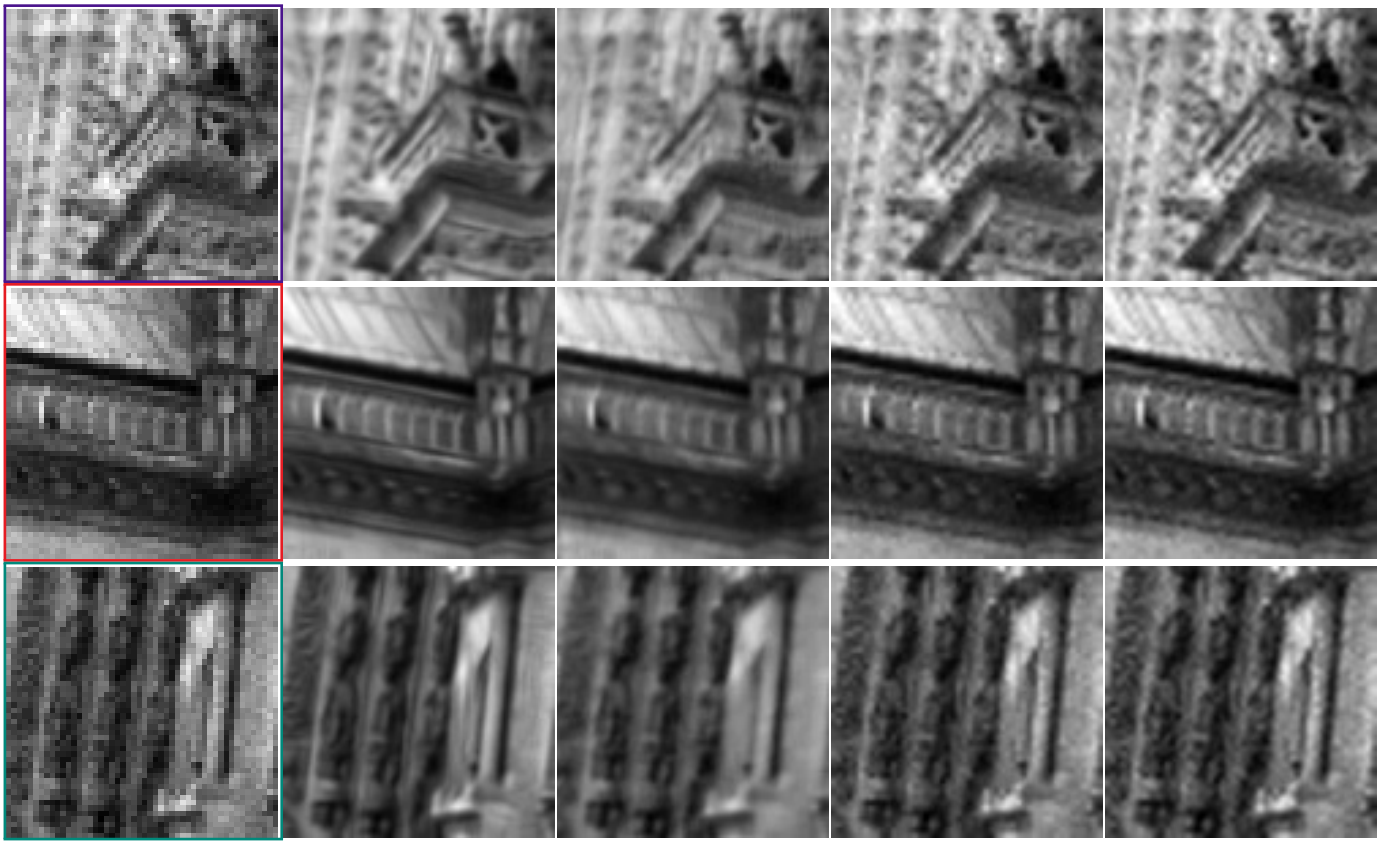


Fig. 6. **Real data. Zooming $\times 2$.** Interpolation of the green channel of a raw image (RGGB). **Left to right:** Input low-resolution image, result by HBE, PLE adapted to noise with variable variance [18], bicubic and lanczos interpolation.

Fig. 7. Example of the acquisition of an HDR scene using spatially varying pixel exposures. **Left:** Tone mapped HDR scene restored from the raw image. **Right top:** Raw image with spatially varying exposure levels. **Right bottom:** Mask of correctly exposed pixels (white) and under or over exposed pixels (black).

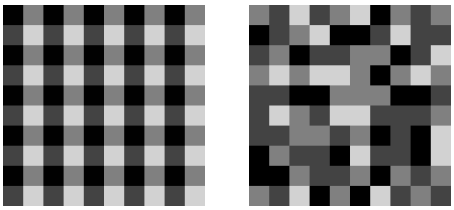


Fig. 8. Regular (left) and non-regular (right) optical masks for an example of 4 different filters.

and underexposed regions in the reference frame, since the information is completely lost in those areas. The SVE acquisition strategy prevents from having large saturated regions to inpaint. In general, all scene regions are sampled by at least one of the exposures thus simplifying the inpainting problem.

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 which shows significant improvements over existing approaches.

A. Spatially varying exposure acquisition model

As presented in [24], [26], [27], an optical mask with spatially varying transmittance can be 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. The main noise sources for this kind of 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) [1], [28]. Thus the following noise model for the non saturated raw pixel value $\mathrm{Z}(p)$ at position p is applicable for this new acquisition method[6]

$$\mathrm{Z}(p) \sim \mathcal{N}(go_p a_p \tau \mathrm{C}(p) + \mu_R, g^2 o_p a_p \tau \mathrm{C}(p) + \sigma_R^2), \quad (23)$$

where g is the camera gain, o_p is the per-pixel SVE 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.

In the approach proposed by Nayar and Mitsunaga [24], the varying exposures follow a regular pattern. Motivated by the aliasing problems of regular sampling patterns, Schöberl et al. [29] propose to use spatially varying exposures on a non-regular pattern. Figure 8 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

$$\mathrm{Y}(p) = \frac{\mathrm{Z}(p) - \mu_R}{go_p a_p \tau}. \quad (24)$$

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

$$(\boldsymbol{D})_p = \begin{cases} 1 & \text{if } \mu_R < \mathrm{Z}(p) < z_{sat}, \\ 0 & \text{otherwise} \end{cases} \quad (25)$$

with z_{sat} equal to the pixel saturation value, thus eliminating the under or overexposed pixels. From (23) and (25), $\mathrm{Y}(p)$ can be modeled as

$$\mathrm{Y}(p)|(\boldsymbol{D})_p \sim \mathcal{N}\left((\boldsymbol{D})_p \mathrm{C}(p), \frac{g^2 o_p a_p \tau (\boldsymbol{D})_p \mathrm{C}(p) + \sigma_R^2}{(go_p a_p \tau)^2}\right). \quad (26)$$

Notice that (26) is the distribution of $\mathrm{Y}(p)$ for a given $(\boldsymbol{D})_p$, since $(\boldsymbol{D})_p$ is itself a random variable that depends on $\mathrm{Z}(p)$. The exposure degradation factor must be included in (26) 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 (26) we have

$$\mathrm{Y} = \boldsymbol{D}\mathrm{C} + \mathrm{N}, \quad (27)$$

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

$$(\boldsymbol{\Sigma}_\mathrm{N})_j = \frac{g^2 o_p a_p \tau (\boldsymbol{D})_p \mathrm{C}(p) + \sigma_R^2}{(go_p a_p \tau)^2}. \quad (28)$$

[6]Some noise sources not modeled in [28], such as blooming, might have a considerable impact in the SVE acquisition strategy and should be considered in a more accurate image modeling.

Fig. 9. **Synthetic data. First column (top to bottom):** Ground-truth with indicated extracts, result obtained by the proposed approach, result by Schöberl et al. [29] **Second to fifth column (left to right):** Ground-truth, result by the proposed approach, Schöberl et al. [29], Nayar and Mitsunaga [24]. **Sixth column:** Random (top) and regular (bottom) mask for each extract. Black represents unknown and white known pixels. The percentage of unknown pixels for the first extract is 65% (it is nearly the same for both the regular and non-regular pattern). For the other two extracts, most pixels are known (99%) so that the proposed method mostly performs denoising in these extracts.

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 (27), image Y is under the hypothesis of the HBE framework so we can apply the proposed patch-based reconstruction approach to HDR imaging. The proposed HDR method consists in the following steps:

1) generate $\boldsymbol{D}$ from Z according to (25),
2) obtain Y from Z according to (24),
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.

1) Synthetic data: Experiments using synthetic data are carried out in order to be able to compare the reconstruction obtained by the proposed and previous methods against a ground-truth, which is not possible (or highly prone to errors) using real data. For this purpose, sample images are generated according to Model (27) using the HDR image in Figure 9 as ground-truth. Both a random and a regular pattern with four equiprobable exposure levels $o = \{1, 2, 5, 10\}$ are simulated. The exposure time is set to $\tau = 1/250$ seconds and the camera parameters are those of a Canon 400D camera set to ISO 200 ($g = 0.66$, $\sigma_R^2 = 17$, $\mu_R = 256$, $\mathrm{v_{sat}} = 4057$).

Figure 9 shows the results obtained by the proposed method and by Schöberlet al. for the random pattern and the results obtained by the bi-cubic interpolation proposed by Nayar et Mitsunaga using the regular pattern. Three extracts of the image are shown together with their corresponding masks of known (white) and unknown (black) pixels. The percentage of unknown pixels for the first extract is 65% (it is nearly the same for both the regular and non-regular pattern). For the other two extracts most of the pixels are known (99%) so that the proposed method mostly performs denoising in these extracts. Table II shows the PSNR values obtained in each extract by each method. The proposed method manages to correctly reconstruct the irradiance information from the input samples. Moreover, its denoising performance is much better than that of Schöberl et al. and Nayar and Mitsunaga, giving a similar reconstruction quality on the unknown areas.

	PSNR (dB)		
	extract 1 (green)	extract 2 (blue)	extract 3 (red)
Proposed method	35.7	49.3	41.9
Schöberl et al.	34.6	43.2	37.0
Nayar and Mitsunaga	35.9	43.9	35.4

TABLE II
PSNR VALUES FOR THE EXTRACTS SHOWN IN FIGURE 9.

We observed that for synthetic scenes with a very high dynamic range (e.g. 17 stops), the reconstructed HDR images could present some artifacts. This limitation never occurred in the real experiments we conducted. We suspect that the Gaussian mixture model used in the PLE approach is not fully adapted when the dynamic range of image patches is too large. We are currently working on a refinement of the stochastic model taking into account this specificity.

2) Real data: The feasibility of the SVE random pattern has been shown in [27] and that of the SVE regular pattern in [26]. 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 approach using real data. For this purpose we generate an SVE image drawing pixels at random from four raw images acquired with different exposure times. The four different exposure times simulate the different filters of the SVE optical mask. The images 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 to composite the SVE frame. Notice that the SVE image thus obtained is very similar to the one obtained if such an optical filter was placed adjacent to the sensor.

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.

Given the procedure we use to generate the SVE image from the input raw images, the Bayer pattern of the latter is kept in the generated SVE image. The proposed irradiance reconstruction method is thus applied to the raw SVE image with an overlap of $\sqrt{n} - 2$ between patches (i.e. a shift of two pixels) in order to compare pixels of the corresponding

Fig. 10. **Real data. Left:** Tone mapped HDR image obtained by the proposed approach (11.4 stops). **Middle top:** Raw image with spatially varying exposure levels. **Middle bottom:** Mask of unknown (black) and known (white) pixels. In the regions with unknown pixels, the percentage of missing pixels varies between 25% to 40%. **Right:** Extracts of the scene.

color channels. A patch size of 6 is used for the examples in Figures 11 and 12, and a patch size of 8 for the example in Figure 10. The demosaicking method by Adams and Hamilton [30] is then used to obtain a color image from the reconstructed irradiance. To display the results we use the tone mapping technique by Mantiuk et al. [31].

A comparison against the methods by Nayar and Mitsunaga and Schöberl et al. is not presented since they do not precise in their works 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.

Figures 10 to 12 show the results obtained in three real scenes, together with the input raw images and the mask of known (white) and unknown (black) pixels. Recall that among the unknown pixels, some of them correspond to saturated pixels and some of them to under exposed pixels. The proposed method manages to correctly reconstruct the unknown pixels even in extreme conditions where more than 70% of the pixels are missing.

These examples show the capacity of the proposed approach to reconstruct the irradiance information in both very dark and bright regions simultaneously. See for instance the example in Figure 11, 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. Figure 12 is another example of this where the dark region is particularly dark (please consult the pdf version of this chapter 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 general restoration problems such as zooming, inpainting and interpolation, by combining local estimation with Bayesian restoration based on hyperpriors.

Fig. 12. **Real data. Left:** Tone mapped HDR image obtained by the proposed approach (13.4 stops). **Right top:** Raw image with spatially varying exposure levels. **Right bottom:** Mask of unknown (black) and known (white) pixels. In the lamp area 70% percent of the pixels are unknown.

In this way, all the restoration problems are set under the same framework. We presented a large series of experiments on both synthetic and real data that confirm how sound the proposed strategy based on hyperpriors is. 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

Fig. 11. **Real data. Left:** Tone mapped HDR image obtained by the proposed approach (15.6 stops). **Right top:** Extract of the building which is almost saturated by the tone mapping algorithm and mask of the corresponding extract. **Right bottom:** Mask of unknown (black) and known (white) pixels. In the brightest part of the building 73% of the pixels are unknown. Despite this fact, the reconstructed HDR image exhibits no visible artifacts.

remains an unsolved challenge.

Moreover, we have presented a novel application of the proposed general framework to the generation of HDR images from a single 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
MAP ESTIMATION OF PARAMETERS

a) Unified MAP: We want to compute the maximum a posteriori

$$\underset{\{\mathrm{C}_i\},\mu,\mathbf{\Lambda}}{\arg\max}\ p(\{\mathrm{C}_i\}_i,\mu,\mathbf{\Lambda}\mid\{\mathrm{Z}_i\}_i) =$$
$$\underset{\{\mathrm{C}_i\},\mu,\mathbf{\Lambda}}{\arg\max}\ p(\{\mathrm{Z}_i\}\mid\{\mathrm{C}_i\},\mu,\mathbf{\Lambda})\,.\,p(\{\mathrm{C}_i\}\mid\mu,\mathbf{\Lambda})\,.\,p(\mu,\mathbf{\Lambda}^{-1}).$$

with $\forall i = 1\dots M$,

$$p(\mathrm{Z}_i\mid\mathrm{C}_i,\mu,\mathbf{\Lambda})\sim\mathcal{N}(\boldsymbol{D}_i\mathrm{C}_i,\mathbf{\Sigma}_{N_i})$$
$$\propto|\mathbf{\Sigma}_N^{-1}|^{\frac{1}{2}}\exp\left(-\frac{1}{2}(\mathrm{Z}_i-\boldsymbol{D}_i\mathrm{C}_i)^T\mathbf{\Sigma}_N^{-1}(\mathrm{Z}_i-\boldsymbol{D}_i\mathrm{C}_i)\right),$$

$$p(\mathrm{C}_i\mid\mu,\mathbf{\Lambda})\sim\mathcal{N}(\mu,\mathbf{\Lambda})$$
$$\propto|\mathbf{\Lambda}|^{\frac{1}{2}}\exp\left(-\frac{1}{2}(\mathrm{C}_i-\mu)^T\mathbf{\Lambda}(\mathrm{C}_i-\mu)\right),$$

and

$$p(\mu\mid\mathbf{\Lambda})\ \sim\mathcal{N}(\mu_0,\mathbf{\Lambda}^{-1}/\kappa)$$
$$\propto\ |\mathbf{\Lambda}|^{\frac{1}{2}}\exp\left(-\frac{\kappa}{2}(\mu-\mu_0)^T\mathbf{\Lambda}(\mu-\mu_0)\right)$$

$$p(\mathbf{\Lambda})\sim\mathcal{W}((\nu\mathbf{\Sigma}_0)^{-1},\nu)$$
$$\propto\ |\mathbf{\Lambda}|^{\frac{\nu-n-1}{2}}\exp\left(-\frac{1}{2}\mathrm{trace}[\nu\mathbf{\Sigma}_0\mathbf{\Lambda}]\right)$$

b) Maximization with respect to $\{\mathrm{C}_i\}_i$.: If we forget the dependence of $\mathbf{\Sigma}_N$ on C_i, we obtain

$$\frac{\partial\log p(\{\mathrm{C}_i\}_i,\mu,\mathbf{\Lambda}\mid\{\mathrm{Z}_i\}_i)}{\partial\mathrm{C}_i}$$
$$=\boldsymbol{D}_i^T\mathbf{\Sigma}_N^{-1}(\mathrm{Z}_i-\boldsymbol{D}_i\mathrm{C}_i)-\mathbf{\Lambda}(\mathrm{C}_i-\mu)\quad.$$

Equating to zero, the solution is given by the Wiener estimator for each i separately

$$\boxed{\mathrm{C}_i=\mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\mathbf{\Lambda}^{-1}\boldsymbol{D}_i^T+\mathbf{\Sigma}_{N_i})^{-1}(\mathrm{Z}_i-\boldsymbol{D}_i\mu)+\mu.}\quad(29)$$

c) Maximization with respect to μ.: Derivating with respect to μ, we obtain

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

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

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

d) Maximization with respect to $\mathbf{\Lambda}$.: The partial derivative with respect to $\mathbf{\Lambda}$ is slightly more complicated (see the Matrix Cookbook for these derivations [32]). Observe that

$$\log p(\{\mathrm{C}_i\}_i,\mu,\mathbf{\Lambda}\mid\{\mathrm{Z}_i\}_i) = \frac{\nu-n+M}{2}\log|\mathbf{\Lambda}| \quad(30)$$
$$-\frac{1}{2}\sum_{i=1}^{M}(\mathrm{C}_i-\mu)^T\mathbf{\Lambda}(\mathrm{C}_i-\mu) \quad(31)$$
$$-\frac{\kappa}{2}(\mu-\mu_0)^T\mathbf{\Lambda}(\mu-\mu_0) \quad(32)$$
$$-\frac{1}{2}\mathrm{trace}[\nu\mathbf{\Sigma}_0\mathbf{\Lambda}] \quad(33)$$

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 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

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

It follows that

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

Equating to zero, this yields

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

e) Computation of the maximum in $\{\mathrm{C}_i\}$, μ and $\boldsymbol{\Lambda}$.: If we inject the expression of the C_i's at the maximum into the previous expressions of μ and $\boldsymbol{\Lambda}$, we get

$$\begin{aligned}\mu &= \frac{M\overline{C}+\kappa\mu_0}{M+\kappa},\\ &= \frac{\sum_{i=1}^{M}\boldsymbol{\Lambda}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\boldsymbol{\Lambda}^{-1}\boldsymbol{D}_i^T+\boldsymbol{\Sigma}_{N_i})^{-1}(\mathrm{Z}_i-\boldsymbol{D}_i\mu)+M\mu+\kappa\mu_0}{M+\kappa}\end{aligned}$$

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

$$\begin{aligned}&(\kappa\mathrm{Id}+\sum_{i=1}^{M}\boldsymbol{\Lambda}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\boldsymbol{\Lambda}^{-1}\boldsymbol{D}_i^T+\boldsymbol{\Sigma}_{N_i})^{-1}\boldsymbol{D}_i)\mu\\ &=\sum_{i=1}^{M}\boldsymbol{\Lambda}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\boldsymbol{\Lambda}^{-1}\boldsymbol{D}_i^T+\boldsymbol{\Sigma}_{N_i})^{-1}\mathrm{Z}_i+\kappa\mu_0\end{aligned}.$$

In other words,

$$\boxed{\mu=(\kappa\mathrm{Id}+\sum_{i=1}^{M}A_i\boldsymbol{D}_i)^{-1}(\sum_{i=1}^{M}A_i\mathrm{Z}_i+\kappa\mu_0)}, \tag{34}$$

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

In the same way, replacing the C_i's in the value of $\boldsymbol{\Lambda}$ gives

$$\boxed{\boldsymbol{\Lambda}^{-1}=\frac{\nu\boldsymbol{\Sigma}_0+\kappa(\mu-\mu_0)(\mu-\mu_0)^T+\sum_{i=1}^{M}A_i(\mathrm{Z}_i-\boldsymbol{D}_i\mu)(\mathrm{Z}_i-\boldsymbol{D}_i\mu)^TA_i^T}{\nu+M-n}} \tag{35}$$

Since Equations (34) and (35) depends both on $\boldsymbol{\Lambda}$ and μ, we cannot obtain closed forms for the values of μ and $\boldsymbol{\Lambda}$ at the maximum. These two equations can however be seen as a fixed point problem. In practice, we use an iterative approach to find the values of μ and $\boldsymbol{\Lambda}$ from Equations (34) and (35), followed by Equation (29) to restore the patches.

f) Affine risk minimizer:

Proposition 3. *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*

$$\tilde{\mathrm{C}}_i=\boldsymbol{\Lambda}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\boldsymbol{\Lambda}^{-1}\boldsymbol{D}_i^T+\boldsymbol{\Sigma}_{\mathrm{N}_i})^{-1}(\mathrm{Z}_i-\boldsymbol{D}_i\mu)+\mu. \tag{36}$$

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

$$\tilde{\mathbf{W}}=\arg\min_{\mathbf{W}}\mathbb{E}[(\mathbf{W}\mathrm{Z}_i-\mathrm{C}_i)^2]. \tag{37}$$

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

$$\mathbb{E}[(\tilde{\mathbf{W}}\mathrm{Z}_i-\mathrm{C}_i)\mathrm{Z}_i^T]=0, \tag{38}$$

and we have

$$\tilde{\mathbf{W}}=\mathbb{E}[\mathrm{C}_i\mathrm{Z}_i^T](\mathbb{E}[\mathrm{Z}_i\mathrm{Z}_i^T])^{-1}. \tag{39}$$

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

$$\begin{aligned}\mathbb{E}[\mathrm{C}_i\mathrm{Z}_i^T]_{p,q}&=\mathbb{E}[\mathrm{C}_i(\boldsymbol{D}_i\mathrm{C}_i+\mathrm{N}_i)^T]_{p,q} &(40)\\ &=\mathbb{E}[\mathrm{C}_i^p(\boldsymbol{D}_i\mathrm{C}_i)_q+\mathrm{C}_i^p\mathrm{N}_i^q] &(41)\\ &=(\boldsymbol{\Lambda}^{-1}\boldsymbol{D}_i^T)_{p,q}. &(42)\end{aligned}$$

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

$$\begin{aligned}\mathbb{E}[\mathrm{Z}_i\mathrm{Z}_i^T]_{p,q}&=\mathbb{E}[(\boldsymbol{D}_i\mathrm{C}_i+\mathrm{N}_i)(\boldsymbol{D}_i\mathrm{C}_i+\mathrm{N}_i)^T]_{p,q} &(43)\\ &=\mathbb{E}[(\boldsymbol{D}_i\mathrm{C}_i)_p(\boldsymbol{D}_i\mathrm{C}_i)_q^T+(\boldsymbol{D}_i\mathrm{C}_i)_p(\mathrm{N}_i)_q^T &(44)\\ &\quad+(\mathrm{N}_i)_p(\boldsymbol{D}_i\mathrm{C}_i)_q+(\mathrm{N}_i)_p(\mathrm{N}_i)_q^T] &(45)\\ &=(\boldsymbol{D}_i\boldsymbol{\Lambda}^{-1}\boldsymbol{D}_i^T)_{p,q}+(\boldsymbol{\Sigma}_{\mathrm{N}_i})_{p,q}. &(46)\end{aligned}$$

Hence we have,

$$\tilde{\mathbf{W}}=\boldsymbol{\Lambda}^{-1}\boldsymbol{D}_i^T(\boldsymbol{D}_i\boldsymbol{\Lambda}^{-1}\boldsymbol{D}_i^T+\boldsymbol{\Sigma}_{\mathrm{N}_i})^{-1}. \tag{47}$$

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 (36). □

ACKNOWLEDGMENT

The authors would like to thank the authors of [17], [12] and [16] 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] 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.

[2] 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–.

[3] 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.

[4] 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.

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

[6] 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

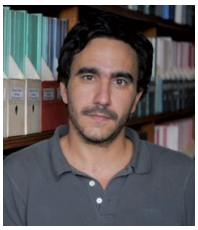
[7] Y.-Q. Wang and J.-M. Morel, "SURE Guided Gaussian Mixture Image Denoising," *SIAM Journal on Imaging Sciences*, vol. 6, no. 2, pp. 999–1034, 2013.
[8] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, 2011, pp. 479–486.
[9] 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.
[10] M. Lebrun, M. Colom, A. Buades, and J. Morel, "Secrets of image denoising cuisine," *Acta Numerica*, vol. 21, no. 1, pp. 475–576, 2012.
[11] Y.-Q. Wang, "The Implementation of SURE Guided Piecewise Linear Image Denoising," *Image Processing On Line*, vol. 3, pp. 43–67, 2013.
[12] ——, "E-PLE: an Algorithm for Image Inpainting," *Image Processing On Line*, vol. 2013, pp. 271–285, 2013.
[13] H. Raiffa and R. Schlaifer, *Applied statistical decision theory*. Division of Research, Graduate School of Business Administration, Harvard University Boston, 1961.
[14] 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
[15] 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.
[16] 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.
[17] 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.
[18] C. Aguerrebere, A. Almansa, J. Delon, Y. Gousseau, and P. Muse, "Single shot high dynamic range imaging using piecewise linear estimators," in *Proceedings of IEEE International Conference on Computational Photography (ICCP)*, 2014.
[19] A. Buades, B. Coll, and J.-M. Morel, "Non-Local Means Denoising," *Image Processing On Line*, vol. 1, 2011.
[20] 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
[21] 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.
[22] C. Aguerrebere, J. Delon, Y. Gousseau, and P. Muse, "Simultaneous HDR image reconstruction and denoising for dynamic scenes," in *Proceedings of IEEE International Conference on Computational Photography (ICCP)*, 2013, pp. 1–11.
[23] 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.
[24] 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.
[25] 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.
[26] 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.
[27] 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.
[28] C. Aguerrebere, J. Delon, Y. Gousseau, and P. Mus, "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
[29] 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.
[30] J. Hamilton and J. Adams, "Adaptive color plan interpolation in single sensor color electronic camera," US Patent 5,629,734, 1997.
[31] 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.
[32] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Technical University of Denmark, Tech. Rep., 2012.

Cecilia Aguerrebere received the B.Sc., M.Sc. and Ph.D. degrees in electrical engineering from Universidad de la República, Uruguay, in 2006, 2011 and 2014 respectively, the M.Sc. degree in applied mathematics from ENS Cachan, France, in 2011, and the Ph.D. degree in Signal and Image Processing from Télécom ParisTech, France, in 2014 (joint Ph.D program with Universidad de la República, Uruguay). From August 2015 she is with the Electrical and Computer Engineering Department, Duke University, where she holds a Postdoctoral Research Associate position.

Andrés Almansa received his HDR, Ph.D. and M.Sc./Engineering degrees in Applied Mathematics and Computer Science from Université Paris-Descartes, ENS Cachan (France) and Universidad de la Republica (Uruguay), respectively, where he is an Associate Professor since 2004. His current interests as a CNRS Research Scientist at Telecom ParisTech include image restoration and analysis, subpixel stereovision and applications to earth observation, high quality digital photography and film restoration.

Julie Delon studied mathematics from the École Normale Supérieure de Cachan and received the Ph. D. degree from ENS Cachan, France, in 2004 and the HDR degree from ENS Cachan in 2011. Between 2005 and 2013, she was a CNRS Researcher with Télécom ParisTech, Paris, France. Since 2013, she is a Professor of applied mathematics with Paris-Descartes University. Her current research interests include mono and multi-image restoration, optimal transport, and stochastic approaches in computer vision. She was coordinator of the young researcher ANR project FREEDOM between 2007 and 2011.

Yann Gousseau received the engineering degree from the École Centrale de Paris, France, in 1995, and the Ph.D. degree in applied mathematics from the University of Paris-Dauphine in 2000. He is currently a professor at Télécom ParisTech. His research interests include the mathematical modeling of natural images and textures, stochastic geometry, image analysis, computer vision and image processing.



Pablo Musé received his Electrical Engineering degree from Universidad de la República, Uruguay, in 1999, his M.Sc. degree in Mathematics, Vision and Learning and his Ph.D. in Applied Mathematics from ENS de Cachan, France, in 2001 and 2004 respectively. From 2005 to 2006 he was with Cognitech, Inc., Pasadena, CA, USA, where he worked on computer vision and image processing applications. In 2006 and 2007, he was a Postdoctoral Scholar with the Seismological Laboratory, California Institute of Technology, Pasadena, working on remote sensing using optical imaging, radar and GPS networks. Since 2008, he has been with the Division of Electrical Engineering, School of Engineering, Universidad de la República, where he is currently a Full Professor of signal processing.