An Adaptive Correlated Image Prior for Image Restoration Problems

Jakub Ševčík, Václav Šmídl, *Member, IEEE*, and Filip Šroubek, *Member, IEEE*

***Abstract*—Image restoration is typically defined as an ill-posed problem which has to be regularized to obtain an acceptable solution. In Bayesian interpretation, regularization is equivalent to prior model of the image. An added value of Bayesian point of view is the ability to form a hierarchical model and estimate the hyperparameters of the prior from the data. Many prior models are available, usually based on automatic relevance determination principle applied to the transformed image. However, the transformation (the most common is a differential operator) is assumed to be known. In this letter, we propose to relax this assumption and estimate the image transformation from the data. The resulting algorithm is analytically tractable using the variational Bayes method. Properties of the new prior are demonstrated on the problem of image superresolution.**

***Index Terms*—Adaptive image prior, image restoration, variational Bayes.**

I. Introduction

Many image restoration tasks, such as blind denoising, deconvolution, or superresolution, are ill-posed and thus formalized as an optimization problem. The quality of resulting images is determined by regularization terms. Conventional regularization techniques require to choose tuning coefficients and the results are often sensitive to their values. This shortcoming is addressed by Bayesian formalism, where the regularization term is interpreted as an image prior and the tuning coefficients as hyperparameters. By estimating the hyperparameters from data, Bayesian methods allow for self-tuning. For this reason, we focus on Bayesian methods in this letter.

Priors are not defined directly on image intensity values, but rather on transformed values, such as image derivatives, of which statistics depend less on the image content. The relation between a Gaussian prior and a quadratic form of ℓ_2 norm is well known [1]–[3]. When such a homogenous prior for all pixels is assumed, resulting images are overly smooth and lack sharp edges. It is possible to alleviate the problem by estimating variance of the prior for each pixel independently [4], which is known as the automatic relevance determination (ARD). However, the choice of the transformation on which to perform the ARD in [4] is still left to the user.

In the non-Bayesian approach, many successful regularization terms are nonquadratic, for example, sparse total variation (TV) regularization term [5], [6], ℓ_1 term [7], or ℓ_p term [8], $0 < p < 1$. If the tuning coefficients are correctly chosen, these methods outperform the quadratic penalizations. Bayesian versions of these methods were proposed to achieve estimation of the tuning parameters, for sparse TV [9], [10], ℓ_1 prior [11], ℓ_p prior [12], $0 < p < 1$, or so-called generalized- or super-Gaussian prior [13], [14]. The Bayesian inference is typically intractable in these cases and a local approximation of the regularization term by surrogate is necessary.

Recently, methods combining different priors appeared as an effective way to deal with image restoration. We mention an approach combining sparse priors (TV, ℓ_1) with nonsparse Gaussian prior [15], yet the estimation procedure assumes fixed weights between the priors. Other examples are combining 1:1 TV prior with a prior based on Frobenius norm of Hessian [16], or composing priors with the first or second-order derivatives in different directions [17]. More recent papers propose to integrate prior terms with an adaptively estimated norm inside the model [18], [19].

All of the above approaches are concerned with the choice of the transformed pixel residue norm and not with the transformation itself, which is used in the norm. A classical choice of the transformation is to apply homogenous (space-invariant) filters, such as the weighted combination of horizontal and vertical differences. To our knowledge, little work has been done on the use of heterogenous (space-variant) filters and estimation of coefficients therein. One of the few exceptions is the work of Katsuki *et al.* [20], where a standard Gaussian prior is assumed with estimated line processes that control local correlations among pixel values. However, only simple model structures are considered.

In this letter, we introduce a hierarchical Bayesian model of Gaussian prior with unknown correlations between each pixel and its neighborhood. The correlation coefficients are estimated from the observed image itself, which can be interpreted as finding the most suitable transformation, more specifically, coefficients of heterogeneous filters. Additional advantages are that the proposed model allows an exact variational Bayesian (VB)

Manuscript received March 5, 2018; revised May 7, 2018; accepted May 9, 2018. Date of publication May 15, 2018; date of current version June 7, 2018. This work was supported in part by the Ministry of Education, Youth and Sports of the Czech Republic under the RICE New Technologies and Concepts for Smart Industrial Systems project LO1607; and in part by the Czech Science Foundation project GA18-05360S. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Marco Cagnazzo.
(Corresponding author: Jakub Ševčík.)

J. Ševčík is with the Regional Innovation Centre for Electrical Engineering, University of West Bohemia, Pilsen CZ-306 14, Czech republic (e-mail: jsevcik@rice.zcu.cz).

V. Šmídl and F. Šroubek are with the Czech Academy of Science, Institute of Information Theory and Automation, Prague CZ-182 08, Czech Republic (e-mail: smidl@utia.cas.cz; sroubekf@utia.cas.cz).
This letter has supplementary downloadable material available at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2018.2836964

solution without any approximation and the derived algorithm is free of tuning parameters.

II. Image Prior

Consider an image formation model, $\mathbf{y} = M\mathbf{x} + \mathbf{n}$, where vector $\mathbf{x} \in \mathbb{R}^N$ is a latent grayscale image lexicographically ordered of size $n_1 \times n_2 = N$, $\mathbf{y} \in \mathbb{R}^{\tilde{N}}$ is an acquired image or a set of images, M is a known matrix that performs degradation and $\mathbf{n}$ is an additive white Gaussian noise. Note that $\mathbf{y}$ and $\mathbf{x}$ are generally of different size. To use Bayesian formalism for estimation, it is necessary to define prior on the image $\mathbf{x}$.

Following Tzikas *et al.* [4], we formulate the image prior in a product form of H filtered images: $\boldsymbol{\epsilon}^{(h)} = Q_h\mathbf{x}$, where $N \times N$ square matrix Q_h represents the hth linear filter operator, $h = 1, \ldots, H$. For example, if horizontal and vertical first-order local differences are considered, then $H = 2$ and

$$\epsilon_i^{(1)} = x_i - x_{i+n_1}, \quad \epsilon_i^{(2)} = x_i - x_{i+1}. \tag{1}$$

Assuming an independent zero-mean Gaussian distribution for each filtered image $\boldsymbol{\epsilon}^{(h)}$, the prior for $\mathbf{x}$ takes the form

$$p(\mathbf{x}|A) \propto \mathcal{N}\big(\mathbf{x}|\mathbf{0}, (\hat{Q}^T A \hat{Q})^{-1}\big) \tag{2}$$

where $\hat{Q}^T = \left[Q_1^T, \ldots, Q_H^T\right]$, and A is a diagonal $HN \times HN$ matrix. Note that the majority of image priors are expressed similarly, see, for example, the TV prior in [9], [10].

Our goal is to relax the assumption of the known and fixed operator $\hat{Q}$. Direct estimation of coefficients of $\hat{Q}$ is problematic due to the determinant of covariance in the normalization constant of the Gaussian distribution. We propose to reparametrize the problem using the well-known LDL^T factorization. Specifically, we decompose $N \times N$ square matrix $\hat{Q}^T A \hat{Q}$ into a form $L\mathrm{diag}(\boldsymbol{\alpha})L^T$, where L is $N \times N$ lower unitriangular square matrix and $\boldsymbol{\alpha} = \left[\alpha_1, \ldots, \alpha_N\right]^T$.

This decomposition offers an intuitively appealing interpretation, since it forms a chain rule of pixel conditional probabilities

$$\begin{aligned} p(\mathbf{x}|A) &= \prod_{i=1}^{N-1} p(x_i|x_{i+1:N}, \alpha_i, L_{i+1:N,i}) \cdot p(x_N|\alpha_N) \\ &= \prod_{i=1}^{N-1} \mathcal{N}(x_i| - L_{i+1:N,i}^T x_{i+1:N}, \alpha_i) \cdot \mathcal{N}(x_N|0, \alpha_N) \end{aligned} \tag{3}$$

where the colon notation follows the Matlab syntax and extracts a range of elements from the vector $\mathbf{x}$ and matrix L. In this way, we create an order within pixels starting in the top-left corner, and assume that each pixel is correlated with pixels of higher indices.

In principle, all elements of L can be estimated, however, we will consider only the nearest neighbors, i.e., the horizontal and vertical neighbor (if exists). Thus, we restrict L to be a lower unitriangular square matrix with ones on the main (zeroth) diagonal and with unknown coefficients only on the first and n_1th lower diagonals. This form of L assumes correlation between each pixel x_i and its lower neighboring x_{i+1} and right neighboring x_{i+n_1} pixel. To simplify notation, we introduce an overline vector $\bar{\mathbf{l}}_i$ which consists of elements $l_{i+1,i}$ and $l_{i+n_1,i}$ if these coefficients correspond to neighboring pixels of x_i. We define a selector matrix S_i such that $\bar{\mathbf{l}}_i = S_i L \mathbf{e}_i$, where $\mathbf{e}_i$ is the ith standard basis vector, so $S_i = \left[\mathbf{e}_{i+1}, \mathbf{e}_{i+n_1}\right]^T$.

In the conventional model of image residues, we define the space-variant image filter of the form

$$\epsilon_i = x_i + l_{i+1,i}x_{i+1} + l_{i+n_1,i}x_{i+n_1} = x_i + \bar{\mathbf{l}}_i^T S_i \mathbf{x} \tag{4}$$

where $l_{i+n_1,i}$ and $l_{i+1,i}$ are estimated parameters. Using this model, the proposed image Adaptive Correlated Prior (ACP) becomes

$$\begin{aligned} p(\mathbf{x}|\boldsymbol{\alpha}, L) &\propto \left[\prod_{i=1}^{N} \alpha_i^{1/2}\right] \exp\left[-\frac{1}{2}\mathbf{x}^T L\mathrm{diag}(\boldsymbol{\alpha})L^T\mathbf{x}\right] \\ &= \left[\prod_{i=1}^{N} \alpha_i^{1/2}\right] \exp\left[-\frac{1}{2}\sum_{i=1}^{N} \mathbf{x}^T \mathbf{l}_{:,i}\alpha_i \mathbf{l}_{:,i}{}^T\mathbf{x}\right]. \end{aligned} \tag{5}$$

Let us note that the prior term (5) is a generalized form of priors used in [21], [22] for a one-dimensional problem.

As mentioned before, we infer coefficients of matrix L during the estimation procedure. For this reason, it is also necessary to define a $l_{i+j,i}$ prior term, where $i \in 1 : N-1, j \in \{1, n_1\}$. Our hypothesis utilizes the Gaussian distribution with a constant-mean l_0 value

$$p(l_{i+j,i}|\psi_{i+j,i}) = \psi_{i+j,i}^{1/2} \exp\left[-\frac{\psi_{i+j,i}}{2}(l_{i+j,i} - l_0)^2\right]. \tag{6}$$

Parameter l_0 is set to the constant value -0.5 in the proposed ACP prior, however, its incorporation into the estimation process is also possible.

III. Superresolution Reconstruction

Let us demonstrate the use of the proposed prior in the superresolution reconstruction (SRR) problem, however, its application to other image restoration tasks is completely analogous. Consider an acquisition process that generates a set of K degraded low-resolution (LR) images $\mathbf{y} = \left[\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_K^T\right]^T$ from the high-resolution (HR) image $\mathbf{x}$. The size of LR is assumed to be $n_1/P \times n_2/P$, where $P > 1$ is referred to as SR factor. The image formation model relating the HR image to set of LR images consists of these typical degradations—warping $F_\bullet$, blurring effect (optical, motion) $H_\bullet$, downsampling D, and noise $\mathbf{n}_\bullet$. The model is common in SR literature [1] and takes the form

$$\mathbf{y}_k = DH_kF_k\mathbf{x} + \mathbf{n}_k = M_k\mathbf{x} + \mathbf{n}_k, \quad k = 1, \ldots, K. \tag{7}$$

We assume that matrices M_k are known. This may be difficult to guarantee in practice (see, for example, [9], [18] for including estimation of $F_\bullet$), yet our aim is to introduce and test the novel prior rather than handling various details of the SRR formation model.

Given (7), the SRR problem is to estimate the HR image $\mathbf{x}$ from the set of K LR images $\mathbf{y_k}$, while having priors on noise $\mathbf{n_k}$, $k = 1, \ldots, K$, and HR image $\mathbf{x}$. Assuming that $\mathbf{n}_k$ is a white Gaussian noise with zero mean and precision β_k, statistically

TABLE I
APPROXIMATE POSTERIOR DISTRIBUTIONS AND THEIR SHAPING PARAMETERS

$\tilde{q}(\mathbf{x}) = \mathcal{N}(\mathbf{x}\vert\boldsymbol{\mu}_{\mathbf{x}}, \Sigma_{\mathbf{x}})$,	$\Sigma_{\mathbf{x}} = \left(\sum_{k=1}^{K} M_k^T\langle\beta_k\rangle M_k + \langle L\mathrm{diag}(\boldsymbol{\alpha})L^T\rangle\right)^{-1}$,	$\boldsymbol{\mu}_{\mathbf{x}} = \Sigma_{\mathbf{x}} \times \sum_{k=1}^{K}\left(M_k^T\langle\beta_k\rangle\mathbf{y}_k\right)$,
$\tilde{q}(\alpha_i) = \mathcal{G}(\alpha_i\vert\gamma_{\alpha_i}, \delta_{\alpha_i})$,	$\delta_{\alpha_i} = \delta_{0_\alpha} + \frac{1}{2}\mathbf{e}_i^T\langle L^T\mathbf{x}\mathbf{x}^T L\rangle\mathbf{e}_i$,	$\gamma_{\alpha_i} = \gamma_{0_\alpha} + \frac{1}{2}, \quad i \in 1:N$,
$\tilde{q}(\psi_{i+j,i}) = \mathcal{G}(\psi_{i+j,i}\vert\zeta_{i+j,i}, \eta_{i+j,i})$,	$\eta_{i+j,i} = \eta_0 + \frac{1}{2}\langle l_{i+j,i} - l_0\rangle^2$,	$\zeta_{i+j,i} = \zeta_0 + \frac{1}{2}, \quad i \in 1:N-1, j \in \{1, n_1\}$,
$\tilde{q}(\beta_k) = \mathcal{G}(\beta_k\vert\gamma_{\beta_k}, \delta_{\beta_k})$,	$\delta_{\beta_k} = \delta_{0_\beta} + \frac{1}{2}\left\langle(\mathbf{y}_k - M_k\mathbf{x})^T(\mathbf{y}_k - M_k\mathbf{x})\right\rangle$,	$\gamma_{\beta_k} = \gamma_{0_\beta} + \frac{n_1 n_2}{2P^2}, k \in 1:K$.

independent between each LR image $\mathbf{y}_k, k = 1, \ldots, K$, we can express the probability distribution of the set of LR images given by $\mathbf{x}$ as

$$p(\mathbf{y}|\mathbf{x}, \boldsymbol{\beta}) \propto \prod_{k=1}^{K} \beta_k^{\frac{N/P^2}{2}} \exp\left[-\frac{\beta_k}{2}\|\mathbf{y}_k - M_k\mathbf{x}\|_2^2\right] \quad (8)$$

where $\boldsymbol{\beta} = [\beta_1, \beta_2, \cdots, \beta_K]^T$.

Since vector $\boldsymbol{\beta}$ in (8), vector $\boldsymbol{\alpha}$ in (5), and parameters $\{\psi_{i+j,i}\}$ in (6) are considered to be unknown, we define so-called hyperparameter distributions. Since the conjugate distribution for precision parameters of the Gaussian distribution is the Gamma distribution [23], we choose the Gamma distribution (with shape-rate parameterization) as a hyperparameter prior probability density

$$p(\alpha_i) = \mathcal{G}\left(\alpha_i|\gamma_{0_\alpha}, \delta_{0_\alpha}\right), \quad i \in 1:N \quad (9)$$

$$p(\beta_k) = \mathcal{G}\left(\beta_k|\gamma_{0_\beta}, \delta_{0_\beta}\right), \quad k \in 1:K \quad (10)$$

$$p(\psi_{i+j,i}) = \mathcal{G}\left(\psi_{i+j,i}|\zeta_0, \eta_0\right), \quad i \in 1:N-1, j \in \{1, n_1\}. \quad (11)$$

We can assign coefficients $\psi_{i+j,i}$ into matrix Ψ, which has the same structure as matrix L except ones on the main diagonal.

The desired joint distribution of the proposed model is now explicitly given as

$$p(\mathbf{y}, \mathbf{x}, L, \Psi, \boldsymbol{\alpha}, \boldsymbol{\beta}) = p(\mathbf{y}|\mathbf{x}, \boldsymbol{\beta})p(\mathbf{x}|\boldsymbol{\alpha}, L)\prod_{i=1}^{N} p(\alpha_i) \cdot \prod_{i=1}^{N-1}\prod_{j\in\{1,n_1\}} p(l_{i+j,i}|\psi_{i+j,i})p(\psi_{i+j,i})\prod_{k=1}^{K} p(\beta_k). \quad (12)$$

IV. VARIATIONAL BAYESIAN INFERENCE

We now apply the VB method (see e.g. [24]) to derive an estimation procedure for the model introduced in the previous section. Estimation of unknown HR image $\mathbf{x}$ and all unknown parameters can be theoretically obtained from (12) using Bayes' rule. Since this task is computationally intractable, following the VB framework, we seek *a posterior* distribution satisfying posterior conditional independence

$$q(\mathbf{x}, L, \Psi, \boldsymbol{\alpha}, \boldsymbol{\beta}) \approx \breve{q}(\mathbf{x})\prod_{i=1}^{N}\breve{q}(\alpha_i)\prod_{k=1}^{K}\breve{q}(\beta_k) \times \prod_{i=1}^{N-1}\left[\breve{q}(\bar{\mathbf{l}}_i)\prod_{j\in\{1,n_1\}}\breve{q}(\psi_{i+j,i})\right]. \quad (13)$$

Algorithm 1: VB Algorithm Using Adaptive Correlated Prior.

1) Initialization:
   i) Initialize parameters: $\beta_k = 1, k \in 1:K, \alpha_i = 10^{-5}$, $i \in 1:N, \psi_{i+j,i} = 10^{-5}, l_{i+j,i} = l_0 = -\frac{1}{2}$, $i \in 1:N-1, j \in \{1, n_1\}$.
   ii) Initialize hyperparameters: $\gamma_{0_\alpha} = \delta_{0_\alpha} = 10^{-8}$, $\gamma_{0_\beta} = \delta_{0_\beta} = 10^{-8}, \zeta_0 = \eta_0 = 10^{-2}$.
2) Iterate until convergence criterion is met
   i) Compute shaping parameters of distribution $\tilde{q}(\mathbf{x})$.
   ii) Evaluate value of vector $\boldsymbol{\alpha}$, $\alpha_i = \frac{\gamma_{\alpha_i}}{\delta_{\alpha_i}}$.
   iii) Compute shaping parameters of distributions $\tilde{q}(\bar{\mathbf{l}}_i)$, evaluate values of parameters $\psi_{i+j,i} = \frac{\zeta_{i+j,i}}{\eta_{i+j,i}}$, $i \in 1:N-1, j \in \{1, n_1\}$.
   iv) Evaluate value of vector $\boldsymbol{\beta}$, $\beta_k = \frac{\gamma_{\beta_k}}{\delta_{\beta_k}}$.
3) Put vector $\boldsymbol{\mu}_{\mathbf{x}}$ as the estimate of HR image out.

Note that the statistical independence of each column of matrix L, i.e., $\breve{q}(L) = \prod_{i=1}^{N-1}\breve{q}(\bar{\mathbf{l}}_i)$, is not imposed but it arises from the second equality in (5).

This approximate distribution is found by minimizing the Kullback–Leibler distance between (13) and the true posterior distribution. Using the VB method, we recognized the optimal approximate posterior densities. Their forms and shaping parameters are shown in Table I. Here, we highlight only the approximate posterior distribution of vectors $\bar{\mathbf{l}}_i$

$$\tilde{q}(\bar{\mathbf{l}}_i) = \mathcal{N}(\bar{\mathbf{l}}_i|\boldsymbol{\mu}_{\bar{\mathbf{l}}_i}, \Sigma_{\bar{\mathbf{l}}_i}), \quad i = 1, \ldots, N-1 \quad (14)$$

$$\text{where} \quad \Sigma_{\bar{\mathbf{l}}_i} = \left(\langle\alpha_i\rangle S_i\langle\mathbf{x}\mathbf{x}^T\rangle S_i^T + \mathrm{diag}\langle\overline{\boldsymbol{\psi}}_i\rangle\right)^{-1} \quad (15)$$

$$\boldsymbol{\mu}_{\bar{\mathbf{l}}_i} = \Sigma_{\bar{\mathbf{l}}_i}\left(-\langle\alpha_i\rangle S_i\langle\mathbf{x}\mathbf{x}^T\rangle\mathbf{e}_i + l_0\langle\overline{\boldsymbol{\psi}}_i\rangle\right). \quad (16)$$

$\langle\star\rangle$ denotes the expected value of $\star$ using its approximate distribution and $\overline{\boldsymbol{\psi}}_i$ is a vector selected from matrix Ψ as $\overline{\boldsymbol{\psi}}_i = S_i\Psi\mathbf{e}_i$.

A candidate solution of the set of implicit equations presented in Table I and (14)–(16) can be found iteratively using Algorithm 1.

Initial prior values of the shape and scale parameters γ_0 and δ_0 can be set to noninformative values, such as 10^{-8}. The choice of initial values of hyperparameters ζ_0 and η_0 has higher impact on the results and is set to 10^{-2} following [21].

V. EXPERIMENTAL RESULTS

We performed a number of experiments with various synthetically generated LR images with respect to the SR factor P and

Fig. 1. Original HR images used for synthetic experiments. © of the second image (face): Christiaan Briggs.

Fig. 2. Reconstruction of the "face" image—with SR factor 3 using four LR frames in upper row with SNR 10 dB, in lower row with SNR 45 dB.

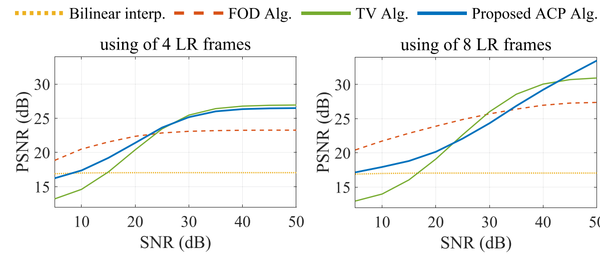

Fig. 3. Means of PSNR (averaged over 10 realizations of 4 different images—see Fig. 1) using four and eight LR frames with SR factor 3 under different levels of noise.

Fig. 4. Illustration of the image prior via mean of Gaussian distribution truncated to positive support (covariance matrix is approximated by its main diagonal).

under different levels of noise. Specifically, we generated sets of LR images for each of the four original HR images displayed in Fig. 1. White Gaussian noise with different signal-to-noise ratio (SNR) from 5 to 50 dB and with 10 different noise realizations at each SNR level was added to each picture.

We present a brief comparison between VB restoration algorithm utilizing distinct priors: i) the proposed ACP prior, ii) the classic Gaussian prior (2) with two filter matrices representing horizontal and vertical first-order local differences (1)—further denoted as the first order difference (FOD) prior, iii) TV prior, i.e., algorithm in [9] without motion parameters estimation.

We made an effort to implement the VB methods for all compared priors as close as possible for a fair comparison. Note that our implementation of all compared VB methods perform exact inversion of matrix $\Sigma_{\mathbf{x}}$. Furthermore, we enforce positivity of pixels in all algorithms (after computing shaping parameters of distribution $\tilde{q}(\mathbf{x})$, negative values of $\hat{\mathbf{x}}$ are set to zero). More details can be seen in Matlab implementation freely available for download from http://www.utia.cas.cz/linear_inversion_methods.

Examples of obtained reconstructed images are displayed in Fig. 2. The proposed ACP algorithm subjectively outperforms other tested methods in the quality of reconstruction at low SNR (10 dB). Differences in the quality reconstruction for higher SNR (45 dB) are not recognizable for the ACP prior and the TV prior (both of them reach satisfactory quality).

Quantitative assessment of the results in terms of peak signal-to-noise ratio (PSNR) is presented in Fig. 3 where PSNR for each prior is averaged over 4 different images and 10 different noise realizations at each SNR level. Results are for four and eight LR images. More setups are presented in the supplementary material of this letter. The VB algorithm with the proposed ACP prior is comparable to that with the TV prior at high SNR, however it is systematically outperforming TV at lower SNR. The FOD prior yields the best PSNR at the lowest SNR, however, the results are subjectively unsatisfactory due to staircase artifacts (artificial contours, see Fig. 2).

Since all compared algorithms differ only in the structure of the covariance matrix of the prior on $\mathbf{x}$ and the prior is truncated to positive support, we can visualize its impact using mean of the truncated Gaussian prior. To achieve tractability, we approximate the covariance matrix only by its diagonal elements. The resulting means of the priors on $\mathbf{x}$ for image "city" (on the left side in Fig. 1) are displayed in Fig. 4. Note that only the proposed prior closely resembles the original image.

Execution time of one iteration of the proposed ACP algorithm is only mildly increased over the other considered VB methods. The most expensive operation (for all studied VB methods) is the exact inversion of matrix $\Sigma_{\mathbf{x}}$, which is of $O(N^2 \log N)$. The ACP algorithm consumes an extra time for evaluation of vectors $\bar{\mathbf{l}}_i$, each requiring inversion of 2×2 matrix, which is of complexity $O(N)$. The required number of iteration is image-, SNR-, and LR frames number-specific, it is comparable for all priors, see the supplement for details.

VI. Conclusion

The quality of image reconstruction is influenced by prior assumptions, typically chosen in the form of filter operators. In this contribution, we proposed an image prior with adaptive correlations, which corresponds to the estimation of filter coefficients automatically from the data. We demonstrated the use of the proposed prior on the problem of SRR where it outperformed methods based on FOD prior and TV prior. The prior can be readily applied to any other image processing application.

REFERENCES

[1] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.
[2] D. Capel and A. Zisserman, "Computer vision applied to super resolution," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 75–86, May 2003.
[3] R. Molina, M. Vega, J. Abad, and A. K. Katsaggelos, "Parameter estimation in Bayesian high-resolution image reconstruction with multisensors," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1655–1667, Dec. 2003.
[4] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "Variational Bayesian sparse kernel-based blind image deconvolution with student's-t priors," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 753–764, Apr. 2009.
[5] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D: Nonlinear Phenom.*, vol. 60, no. 1–4, pp. 259–268, 1992.
[6] Y. Zhu and Y. Shi, "A fast method for reconstruction of total-variation mr images with a periodic boundary condition," *IEEE signal Process. Lett.*, vol. 20, no. 4, pp. 291–294, Apr. 2013.
[7] K.-H. Yap, Y. He, Y. Tian, and L.-P. Chau, "A nonlinear l_1-norm approach for joint image registration and super-resolution," *IEEE Signal Process. Lett.*, vol. 16, no. 11, pp. 981–984, Nov. 2009.
[8] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Process. Lett.*, vol. 14, no. 10, pp. 707–710, Oct. 2007.
[9] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Variational Bayesian super resolution," *IEEE Trans. Image Process.*, vol. 20, no. 4, pp. 984–999, Apr. 2011.
[10] G. Chantas, N. P. Galatsanos, R. Molina, and A. K. Katsaggelos, "Variational Bayesian image restoration with a product of spatially weighted total variation image priors," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 351–362, Feb. 2010.
[11] S. Villena, M. Vega, R. Molina, and A. K. Katsaggelos, "Bayesian super-resolution image reconstruction using an ℓ 1 prior," in *Proc. IEEE 6th Int. Symp. Image Signal Process. Anal.*, 2009, pp. 152–157.
[12] B. Amizic, R. Molina, and A. K. Katsaggelos, "Sparse Bayesian blind image deconvolution with parameter estimation," *EURASIP J. Image Video Process.*, vol. 2012, no. 1, p. 20, 2012. [Online]. Available: http://link.springer.com/article/10.1186/1687-5281-2012-20
[13] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Generalized Gaussian Markov random field image restoration using variational distribution approximation," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2008, pp. 1265–1268.
[14] S. D. Babacan, R. Molina, M. N. Do, and A. K. Katsaggelos, "Bayesian blind deconvolution with general sparse image priors," in *Proc. Eur. Conf. Comput. Vis.* Springer, 2012, pp. 341–355.
[15] S. Villena, M. Vega, S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian combination of sparse and non-sparse priors in image super resolution," *Digit. Signal Process.*, vol. 23, no. 2, pp. 530–541, 2013.
[16] W.-Z. Shao, H.-S. Deng, and Z.-H. Wei, "Kullback–Leibler divergence based composite prior modeling for Bayesian super-resolution," *J. Sci. Comput.*, vol. 60, no. 1, pp. 60–78, 2014.
[17] S. Villena, M. Vega, R. Molina, and A. K. Katsaggelos, "A non-stationary image prior combination in super-resolution," *Digit. Signal Process.*, vol. 32, pp. 1–10, 2014.
[18] R. Jin, S. Zhao, X. Xu, and E. Song, "Multiframe super-resolution based on half-quadratic prior with artifacts suppress," *J. Vis. Commun. Image Represent.*, vol. 40, pp. 656–670, 2016.
[19] S. Zhao, H. Liang, and M. Sarem, "A generalized detail-preserving super-resolution method," *Signal Process.*, vol. 120, pp. 156–173, 2016.
[20] T. Katsuki, A. Torii, and M. Inoue, "Posterior-mean super-resolution with a causal Gaussian Markov random field prior," *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3182–3193, Jul. 2012.
[21] O. Tichý, V. Šmídl, R. Hofman, and A. Stohl, "Ls-apc v1. 0: A tuning-free method for the linear inverse problem and its application to source-term determination," *Geosci. Model Develop.*, vol. 9, no. 11, 2016, Art no. 4297.
[22] L. Ulrych and V. Šmídl, "Sparse and smooth prior for Bayesian linear regression with application to etex data," preprint arXiv:1706.06908, 2017.
[23] C. Robert, *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. Berlin, Germany: Springer Science & Business Media, 2007.
[24] V. Šmídl and A. Quinn, *The Variational Bayes Method in Signal Processing*. Berlin, Germany: Springer Science & Business Media, 2006.