Variational Bayesian Inference Image Restoration using a Product of Total Variation-Like Image Priors

Giannis Chantas [1], Nikolaos Galatsanos [2], Rafael Molina [3], and Aggelos Katsaggelos [4]

[1] *Department of Computer Science, University of Ioannina, Ioannina, Greece 45110.*

[2] *Department of Electrical and Computer Engineering, University of Patras, Rio, Greece 26500.*

[3] *Department of Computer Science, University of Granada, Granada, Spain 18071.*

[4] *Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208-3118, USA.*

***Abstract*—In this paper a new image prior is introduced and used in image restoration. This prior is based on products of spatially weighted Total Variations (TV). These spatial weights provide this prior with the flexibility to better capture local image features than previous TV based priors. Bayesian inference is used for image restoration with this prior via the variational approximation. The proposed algorithm is fully automatic in the sense that all necessary parameters are estimated from the data. Numerical experiments are shown which demonstrate that image restoration based on this prior compares favorably with previous state-of-the-art restoration algorithms.**

I. Introduction

Total Variation (TV) was first introduced as a regularizer for image recovery in [1]. Since then it has been used extensively and with great success for inverse problems because TV has the ability to smooth noise in flat areas of the image and at the same time preserve edges.

Nevertheless, TV-based image restoration has certain shortcomings. One of them is the selection of the regularization parameter which to a large extent until recently has been ad-hoc. Bioucas-Dias *et al.* [3], using their majorization-minimization approach [4], propose a Bayesian method to estimate the original image and regularization parameter assuming that an estimate of the noise variance is available. Recently, a Bayesian inference framework which requires the approximation of the prior partition function and is based on the variational approximation was proposed to handle the simultaneous parameter and image estimation problems [5].

In this paper we contribute to the field of prior image modeling by utilizing and enhancing by combining the advantages of TV image modeling in [5] and Product of Experts (PoE) image modeling in [6]. The new image model has a number of novel features. First, unlike [5] and [3], it uses a spatially weighted version of the TV. These spatial weights provide the prior with the flexibility to model explicitly the *local salient features* of the image. Second, like in [6], it is in product form and has the ability to enforce simultaneously a number of different properties on the image. Due to the complexity of this model the inference methodology has to be derived for the new model of the restoration problem. We resort to the variational approximation for Bayesian inference [2], and develop several extensions of the methodologies used in [5] and [6].

The rest of this paper is organized as follows. In section II we present the imaging and image models. In section III we present the variational algorithm for Bayesian inference. In section IV numerical experiments are shown and in section V conclusions about this work are presented.

II. Imaging and image model

In what follows we use one dimensional notation for simplicity. Let $\mathbf{f}$ be the original image represented as an $N \times 1$ vector, blurred by a convolutional operator $\mathbf{H}$, of size $N \times N$. The degraded observation is given by

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}, \tag{1}$$

where $\mathbf{n}$ is the noise $N \times 1$ vector modeled as white Gaussian, i.e., $\mathbf{n} \sim N(\mathbf{0}, \beta^{-1}\mathbf{I})$, where $\mathbf{0}$ and $\mathbf{I}$ are the $N \times 1$ zero and $N \times N$ identity matrices, respectively, and β^{-1} represents the noise variance.

A. Modified Student's-t image prior

Image priors in product form are very attractive since they have the ability to enforce simultaneously many properties on an image; see for example [6]. For this purpose we propose herein a prior in product form for the image. To define such a prior we introduce P pairs of linear convolutional operators (filters) $(\mathbf{Q}_{2k-1}, \mathbf{Q}_{2k},)$ for $k = 1, 2, \dots, P$ of size $N \times N$ and assume that the filter outputs $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_{2P})$ are produced according to

$$\boldsymbol{\epsilon}_l = \mathbf{Q}_l \mathbf{f}, \; l = 1, \dots, 2P. \tag{2}$$

Then, for each pixel location i, it is assumed that each pair $\epsilon_{2k}(i)$ and $\epsilon_{2k-1}(i)$ is jointly distributed with probability density function

$$p(\epsilon_{2k}(i), \epsilon_{2k-1}(i) | a_k(i)) = \frac{\lambda_k^2 a_k(i)^2}{2\pi} \exp\left(-\lambda_k a_k(i) \sqrt{\epsilon_{2k}(i)^2 + \epsilon_{2k-1}(i)^2}\right). \tag{3}$$

with $k = 1, \dots, P$ and $i = 1, \dots, N$ where N the number of pixels in the image.

One drawback of the herein prior is the over-parameterization problem since PN unknowns $a_k(i)$ have to be estimated from N data points. In order to ameliorate this problem we assume each $a_k(i)$ to be a Gamma distributed hidden random variable [6] according to:

$$p(a_k(i)) = G(a_k(i); \nu_k/2, \nu_k/2) \tag{4}$$

for $k = 1, \ldots, P,\ i = 1, \ldots, N$.

The marginal distribution of $\epsilon_{2k}(i)$ and $\epsilon_{2k-1}(i)$ can be computed in *closed form* and is given by

$$p(\epsilon_{2k}(i), \epsilon_{2k-1}(i)) = \frac{\Gamma(\nu_k/2 + 1/2)}{\Gamma(\nu_k/2)} \left(\frac{\lambda_k}{\pi\nu_k}\right)^{1/2} \left(1 + \frac{\lambda_k \sqrt{\epsilon_{2k}(i)^2 + \epsilon_{2k-1}(i)^2}}{\nu_k}\right)^{-\nu_k/2 - 1/2} \tag{5}$$

for $k = 1, 2, \ldots, P$ and $i = 1, 2, \ldots, N$.

This density function is very similar in form to the Student's-t pdf [2], thus, in the rest of this paper we label it as *Modified Student's-t.*

At this point we note that we have not provided a prior for the image, $p(\mathbf{f})$. This was intentional, because we cannot compute it in closed form. More specifically, it is difficult to define a prior for the image $\mathbf{f}$ based on the prior in Eq. (3) because we cannot compute the partition function for such prior.

We thus introduce an alternative observation model, which is derived by applying the operators $\mathbf{Q}_l$ to the original imaging model in (1). This yields:

$$\mathbf{y}_l = \mathbf{H}\boldsymbol{\epsilon}_l + \mathbf{n}_l,\ l = 1, \ldots, 2P, \tag{6}$$

where $\mathbf{y}_l = \mathbf{Q}_l\mathbf{g}$, $\mathbf{n}_l = \mathbf{Q}_l\mathbf{n}$ and thus $\mathbf{n}_l \sim N(0, \beta^{-1}\mathbf{Q}_l\mathbf{Q}_l^T)$.

We finally arrive at the Bayesian modeling of our problem, that is,

$$p(\mathbf{y}, \boldsymbol{\epsilon}, \mathbf{a}; \theta) = p(\mathbf{y}|\boldsymbol{\epsilon})p(\boldsymbol{\epsilon}|\mathbf{a}; \theta)p(\mathbf{a}; \theta) \tag{7}$$

where $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_{2P})$, $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_{2P})$, with $\boldsymbol{\epsilon}_l = (\epsilon_l(1), \ldots, \epsilon_l(N))$, $\mathbf{a} = (\mathbf{a}_1, \ldots, \mathbf{a}_P)$, with $\mathbf{a}_k = (a_k(1), \ldots, a_k(N))$ $k = 1, \ldots, P$, and $\theta = (\lambda_1, \ldots, \lambda_P, \nu_1, \ldots, \nu_P)$, with the above probability distributions defined by

$$p(\mathbf{y}|\boldsymbol{\epsilon}) = \prod_{l=1}^{2P} p(\mathbf{y}_l|\boldsymbol{\epsilon}_l), \qquad p(\mathbf{y}_l|\boldsymbol{\epsilon}_l) = N(\boldsymbol{\epsilon}_l, \beta^{-1}\mathbf{Q}_l\mathbf{Q}_l^T), \tag{8}$$

$$p(\boldsymbol{\epsilon}|\mathbf{a}; \theta) = \prod_{k=1}^{P}\prod_{i=1}^{N} p(\epsilon_{2k}(i), \epsilon_{2k-1}(i)|a_k(i)), \tag{9}$$

and

$$p(\mathbf{a}; \theta) = \prod_{k=1}^{P}\prod_{i=1}^{N} p(a_k(i); \theta). \tag{10}$$

This Bayesian model will be used for inference in the next section where we treat $\boldsymbol{\epsilon}$ and $\mathbf{a}$ as *hidden variables* and θ as a parameter vector to be estimated.

III. VARIATIONAL INFERENCE WITH THE MODIFIED STUDENT'S-T PRIOR

According to Bayesian inference we have to find the posterior distributions for the hidden variables ϵ and $\mathbf{a}$ given $\mathbf{y}$ and the parameter vector θ. However, the marginal of the observations which is required to find the posteriors of the hidden variables is hard to compute [2]. More specifically, the integral

$$p(\mathbf{y}; \theta) = \int_{\boldsymbol{\epsilon}, \mathbf{a}} p(\mathbf{y}, \boldsymbol{\epsilon}, \mathbf{a}; \theta) d\boldsymbol{\epsilon} d\mathbf{a} \tag{11}$$

is intractable.

The variational algorithm that we describe in what follows, bypasses this difficulty and maximizes a *lower bound* that can be found instead of the log-likelihood of the observations $\log p(\mathbf{y}; \theta)$ [2]. This bound is obtained by subtracting from $\log p(\mathbf{y}; \theta)$ the Kullback-Leibler divergence, which is always positive, between an arbitrary $q(\boldsymbol{\epsilon}, \mathbf{a})$ and $p(\boldsymbol{\epsilon}, \mathbf{a}|\mathbf{y}; \theta)$:

$$L(q(\boldsymbol{\epsilon}, \mathbf{a}), \theta) = \log p(\mathbf{y}; \theta) - KL(q(\boldsymbol{\epsilon}, \mathbf{a})||p(\boldsymbol{\epsilon}, \mathbf{a}|\mathbf{y}; \theta)), \tag{12}$$

and is equal to

$$\begin{aligned} L(q(\boldsymbol{\epsilon}, \mathbf{a}); \theta) &= \int_{\boldsymbol{\epsilon}, \mathbf{a}} q(\boldsymbol{\epsilon}, \mathbf{a}) \log p(\boldsymbol{\epsilon}, \mathbf{a}, \mathbf{y}; \theta) d\boldsymbol{\epsilon} d\mathbf{a} \\ &- \int_{\boldsymbol{\epsilon}, \mathbf{a}} q(\boldsymbol{\epsilon}, \mathbf{a}) \log q(\boldsymbol{\epsilon}, \mathbf{a}) d\boldsymbol{\epsilon} d\mathbf{a}, \end{aligned} \tag{13}$$

When $q(\boldsymbol{\epsilon}, \mathbf{a}) = p(\boldsymbol{\epsilon}, \mathbf{a}|\mathbf{y}; \theta)$, this bound is maximized and $L(q(\boldsymbol{\epsilon}, \mathbf{a}); \theta) = \log p(\mathbf{y}; \theta)$. Because the exact posterior $p(\boldsymbol{\epsilon}, \mathbf{a}|\mathbf{y}; \theta) = \frac{p(\boldsymbol{\epsilon}, \mathbf{a}, \mathbf{y}; \theta)}{p(\mathbf{y}; \theta)}$ cannot be found we use an approximation of the posterior. The mean-field approximation is a commonly used approach to maximize the variational bound w.r.t. $q(\boldsymbol{\epsilon}, \mathbf{a}); \theta$ [2]. According to this approach the hidden variables are assumed to be independent, i.e., $q(\boldsymbol{\epsilon}, \mathbf{a}) = q(\boldsymbol{\epsilon})q(\mathbf{a})$. However, for the herein model this is still not sufficient to obtain a closed form for $q(\boldsymbol{\epsilon})$ which is necessary for inference using this approach. More specifically, the square root in the joint $p(\boldsymbol{\epsilon}, \mathbf{a}, \mathbf{y}; \theta)$ which originates from the prior $p(\boldsymbol{\epsilon}|\mathbf{a})$ makes the definition of $q(\boldsymbol{\epsilon})$ intractable.

A. A Lower Bound for $L(q(\boldsymbol{\epsilon}, \mathbf{a}), \theta)$

For this purpose we introduce a *lower bound* on L also [5]. More specifically, we use the inequality

$$\sqrt{w} \le \frac{w + u}{2\sqrt{u}}, \tag{14}$$

which holds for $w \ge 0$ and $u > 0$. Notice that equality holds when $w = u$. This inequality is used at every pixel i by setting $w_k(i) = \epsilon_{2k}(i)^2 + \epsilon_{2k-1}(i)^2$, for $k = 1, 2, \ldots, P$, where $u_k(i)$ are auxiliary variables used for this approximation. Using this and the prior in Eq. (3) we have

$$p(\epsilon_{2k}(i), \epsilon_{2k-1}(i)|a_k(i)) \ge M(\epsilon_{2k}(i), \epsilon_{2k-1}(i), u_k(i), a_k(i)) \tag{15}$$

where

$$M(\epsilon_{2k}(i), \epsilon_{2k-1}(i), u_k(i), a_k(i)) = \frac{\lambda_k^2 a_k(i)^2}{2\pi} \exp\left(-\frac{\lambda_k a_k(i)}{2} \frac{\epsilon_{2k}(i)^2 + \epsilon_{2k-1}(i)^2 + u_k(i)}{\sqrt{u_k(i)}}\right), \tag{16}$$

for $k = 1, \ldots, P$.

We also define $\mathbf{u}_k = (u_k(1), \ldots, u_k(N))$ and $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_P)$. Let us now define

$$L^b(q(\boldsymbol{\epsilon}), q(\mathbf{a}), \mathbf{u}, \theta) = \int_{\boldsymbol{\epsilon},\mathbf{a}} q(\boldsymbol{\epsilon})q(\mathbf{a}) \log \frac{F(\mathbf{y}, \boldsymbol{\epsilon}, \mathbf{a}; \mathbf{u}, \theta)}{q(\boldsymbol{\epsilon})q(\mathbf{a})} d\boldsymbol{\epsilon} d\mathbf{a}, \tag{17}$$

where

$$F(\mathbf{y}, \boldsymbol{\epsilon}, \mathbf{a}; \mathbf{u}, \theta) = p(\mathbf{y}|\boldsymbol{\epsilon}) \left[\prod_{k=1}^{P} \prod_{i=1}^{N} M(\epsilon_{2k}(i), \epsilon_{2k-1}(i), u_k(i), a_k(i)) \right] p(\mathbf{a}; \theta). \tag{18}$$

Then, since $F(\mathbf{y}, \boldsymbol{\epsilon}, \mathbf{a}; \mathbf{u}, \theta) \leq p(\mathbf{y}, \boldsymbol{\epsilon}, \mathbf{a})$ we have

$$L^b(q(\boldsymbol{\epsilon}), q(\mathbf{a}), \mathbf{u}, \theta) \leq L(q(\boldsymbol{\epsilon}, \mathbf{a}), \theta), \tag{19}$$

and consequently the bound becomes tight when

$$\max_{\mathbf{u}} L^b(q(\boldsymbol{\epsilon}), q(\mathbf{a}), \mathbf{u}, \theta) \leq L(q(\boldsymbol{\epsilon}, \mathbf{a}), \theta). \tag{20}$$

Notice that the new lower bound L^b is quadratic in the hidden variables $\boldsymbol{\epsilon}$, thus it is possible to find $q(\boldsymbol{\epsilon})$ that maximizes it. In contrast, the original bound L was not quadratic in $\boldsymbol{\epsilon}$.

B. A Constrained Variational Inference Algorithm

As we have already explained, $\boldsymbol{\epsilon}_l$, $l = 1, \ldots, 2P$, are used instead of $\mathbf{f}$ to avoid the computation of the normalization constant of the prior on $\mathbf{f}$. Thus, a question that needs to be addressed is how one finds $\mathbf{f}$ given the different $q(\boldsymbol{\epsilon}_l)$.

For this purpose the *constrained variational* approximation first proposed in [6] is applied. According to this approach, each $q(\boldsymbol{\epsilon}_l)$ is constrained to have the form:

$$q(\boldsymbol{\epsilon}_l) = N(\mathbf{Q}_l \mathbf{m}, \mathbf{Q}_l \mathbf{R} \mathbf{Q}_l^T), \tag{21}$$

where $\mathbf{m}$ is a $N \times 1$ vector, taken as the mean of the image, and $\mathbf{R}$ the $N \times N$ image covariance matrix.

We now present the maximization method by giving the updates for the variables of the bound L^b in the j-th iteration. In the VE-step, the maximization of $L^b(q(\boldsymbol{\epsilon}), q(\mathbf{a}), \mathbf{u}, \theta)$ is performed with respect to $q(\mathbf{a})$, $\mathbf{m}$ and $\mathbf{R}$ keeping $\mathbf{u}$ and θ fixed, while in the VM-step, the maximization of $L^b(q(\boldsymbol{\epsilon}), q(\mathbf{a}), \mathbf{u}, \theta)$ is performed with respect to $\mathbf{u}$ and θ keeping $q(\mathbf{a})$, $\mathbf{m}$, and $\mathbf{R}$ keeping fixed. We have,

VE-step:

$$[\mathbf{m}^j, \mathbf{R}^j, q^j(\mathbf{a})] = \arg \max_{\mathbf{m}, \mathbf{R}, q(\mathbf{a})} L^b(q(\boldsymbol{\epsilon}), q(\mathbf{a}), \mathbf{u}^{j-1}, \theta^{j-1}) \tag{22}$$

VM-step:

$$[\mathbf{u}^j, \theta^j] = \operatorname{argmax}_{\mathbf{u},\theta} L^b(q^j(\boldsymbol{\epsilon}), q^j(\mathbf{a}), \mathbf{u}, \theta) \tag{23}$$

The updates for the VE-Step yield

$$q^j(\boldsymbol{\epsilon}_l) = N(\mathbf{Q}_l \mathbf{m}^j, \mathbf{Q}_l \mathbf{R}^j \mathbf{Q}_l^T), \tag{24}$$

where

$$\mathbf{m}^j = \beta \mathbf{R}^j \mathbf{H}^T \mathbf{g}, \tag{25}$$

$$\mathbf{R}^j = \left(\beta \mathbf{H}^T \mathbf{H} + \frac{1}{2P} \sum_{k=1}^{P} \lambda_k{}^{j-1} \left(\mathbf{Q}_{2k}^T \langle \mathbf{A}_k \rangle^{j-1} (\mathbf{U}_k^{-1/2})^{j-1} \mathbf{Q}_{2k} + \mathbf{Q}_{2k-1}^T \langle \mathbf{A}_k \rangle^{j-1} (\mathbf{U}_k^{-1/2})^{j-1} \mathbf{Q}_{2k-1} \right) \right)^{-1}. \tag{26}$$

From the above equations it is clear that $\mathbf{m}$ merges information from all filters $\mathbf{Q}_l$ and is used as the estimate of $\mathbf{f}$.

Finally, the approximate posterior of $\mathbf{a}$ in the VE-step is given by

$$q^j(a_k(i)) = G\left(a_k(i); \frac{\nu_k^{j-1}}{2} + 2, \frac{\nu_k^{j-1}}{2} + \lambda_k^{j-1} \sqrt{u_k^{j-1}(i)} \right)$$

for $i = 1, \ldots, N$ and $k = 1, 2, \ldots, P$. Thus, the expectation of $a_k(i)$ w.r.t $q^j(a_k(i))$ is

$$\langle a_k(i) \rangle_{q^j(\mathbf{a})} = \frac{\nu_k^{j-1} + 4}{\nu_k^{j-1} + 2\lambda_k^{j-1} \sqrt{u_k^{j-1}(i)}} \tag{27}$$

In the VM-step, the bound is maximized w.r.t to the parameters. To find $\mathbf{u}^j$ we have to solve

$$u_k^j(i) = \arg \min_{u_k(i)} \frac{\langle \epsilon_{2k}(i)^2 + \epsilon_{2k-1}(i)^2 \rangle_{q^j(\boldsymbol{\epsilon})} + u_k(i)}{\sqrt{u_k(i)}} \tag{28}$$

where $\langle . \rangle_{q^j(\boldsymbol{\epsilon})}$ represents the expectation w.r.t. $q^j(\boldsymbol{\epsilon})$, which produces

$$u_k^j(i) = \langle \epsilon_{2k}(i)^2 + \epsilon_{2k-1}(i)^2 \rangle_{q^j(\boldsymbol{\epsilon})} = \sum_{r=0}^{1} ((\mathbf{m}_{2k-r}^j(i))^2 + \mathbf{C}_{2k-r}^j(i,i)) \tag{29}$$

for $i = 1, \ldots, N$ and $k = 1, 2, \ldots, P$, where

$$\mathbf{m}_{2k-r}^j = \mathbf{Q}_{2k-r} \mathbf{m}^j, \ \mathbf{C}_{2k-r}^j = \mathbf{Q}_{2k-r} \mathbf{R}^j \mathbf{Q}_{2k-r}^T. \tag{30}$$

For λ_k the update formula is

$$\lambda_k^j = \frac{2N}{\sum_{i=1}^{N} \langle a_k(i) \rangle_{q^j(\mathbf{a})} \sqrt{u_k^j(i)}}. \tag{31}$$

Similarly, ν_k^j is the root of the function which is proportional to the derivative of $L^b(q^j(\boldsymbol{\epsilon}), q^j(\mathbf{a}), \mathbf{u}, \theta)$ with respect to ν_k.

IV. Numerical Experiments

We demonstrate the value of the proposed restoration approach by testing it in experiments with four well known input images: *Lena*, *Cameraman*, *Barbara* and *USC-man*. Every image is blurred with two types of blur; the first blur has the shape of a Gaussian function with shape parameter 9, and the second is uniform with support a rectangular region of dimensions 9×9. The blurred signal to noise ratio ($BSNR$) was used to quantify the noise level:

$$BSNR = 10 \log_{10} \frac{\|\mathbf{Hf}\|_2^2}{N\sigma^2},$$

where σ^2 is the variance of the additive white Gaussian noise (AWGN). Three levels of AWGN were added to the blurred images resulting in $BSNR = 40$, 30 and 20 dB. Thus in total 24 image restoration experiments are presnted.

As performance metric, the improvement in Signal to Noise Ratio ($ISNR$) was used:

$$ISNR = 10\log_{10}\frac{\|\mathbf{f}-\mathbf{g}\|_2^2}{\|\mathbf{f}-\hat{\mathbf{f}}\|_2^2},$$

where $\mathbf{f}$, $\mathbf{g}$ and $\hat{\mathbf{f}}$ are the original, observed degraded and restored images, respectively.

In the implementation of our proposed restoration algorithm $P = 2$ was used. In other words, four filter outputs were used for the prior and it is a product with two terms. The operators $\mathbf{Q}_1$ and $\mathbf{Q}_2$ correspond to the horizontal and vertical first order differences. The other two operators $\mathbf{Q}_3$ and $\mathbf{Q}_4$ are obtained by convolving the previous horizontal and vertical first order differences filters with fan filters with vertical and horizontal pass-bands, respectively. In our experiments the fan filters in [7] were used.

We compared the herein proposed restoration method, abbreviated as CGMK from the first letter of each author's last name, with the *Lena* and *Cameraman* images with three recent TV-based algorithms: the algorithms in [3] and [4] abbreviated by BFO1 and BFO2, respectively, and the algorithm in [5] abbreviated as BMK. We also compared it with the variational Bayesian algorithm in [6] which is abbreviated as CGLS.

The *ISNR* results of this comparison are shown in Tables I and II for the experiments with uniform, and Gaussian blurs, respectively. We also present *ISNR* results for similar experiments for the *Barbara* and the *USC man* images in Tables III and IV. From the *ISNR* results in Tables I-IV one can say that out of the 24 experiments presented in this paper in 18 of them the herein proposed algorithm provided better *ISNR* from all other tested methods. Furthermore, for certain cases (*USC-man*, $BSNR = 20dB$, 9×9 blur) the $ISNR$ gain of the proposed algorithm over the best of its predecessors was quite significant ($0.44dB$).

In Figure 1 we show the images related to the experiment with the USC-man image, $BSNR = 20$, and uniform blur 9×9. It is clear that the image restored with the proposed prior is sharper than that with the TV prior (the same as here but whith inifinte degrees of freedom ν_k) and less "cartoon" than the algorithm in [6] based on a Student's-t. prior.

V. Conclusions

We presented a new promising image prior that is based on the Modified Student's-t pdf and a variational algorithm that estimates all the parameters of this model automatically and finds the restored image. We compared this restoration approach with previous state-of-the-art methods and found that it compares favorably to them. This prior can be used for a Bayesian setting for a variety of other image recovery problems, such as, super-resolution, blind-deconvolution, and tomographic reconstruction. Furthermore, it can be used in other imaging applications where a statistical model for the image is necessary, for example, detection of watermarks in images.

References

[1] L. I. Rudin, S. Osher and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *in Phys. D* Vol. 60, pp. 259-268, 1992.

TABLE I: $ISNR$'s for the *Lena* and *Cameraman* images. Experiments using uniform 9×9 blur

Uniform blur 9×9		Lena	Cameraman
$BSNR(dB)$	Method	ISNR(dB)	
$BSNR = 40$	$CGMK$	**8.52**	**9.61**
	$CGLS$	8.49	9.53
	BMK	8.34	8.55
	$BFO1$	8.42	8.57
	$BFO2$	8.37	8.46
$BSNR = 30$	$CGMK$	**6.25**	**6.55**
	$CGLS$	6.10	6.29
	BMK	6.08	5.68
	$BFO1$	5.89	5.41
	$BFO2$	5.58	4.38
$BSNR = 20$	$CGMK$	**4.24**	**3.55**
	$CGLS$	3.98	3.33
	BMK	4.09	3.31
	$BFO1$	3.72	2.42
	$BFO2$	3.15	1.94

TABLE II: $ISNR$'s for the *Lena* and *Cameraman* images. Experiments using Gaussian blur (variance 9)

Gaussian blur variance 9		Lena	Cameraman
$BSNR(dB)$	Method	ISNR(dB)	
$BSNR = 40$	$CGMK$	4.64	3.49
	$CGLS$	**4.86**	3.45
	BMK	4.72	**3.51**
	$BFO1$	4.78	3.39
	$BFO2$	4.49	3.26
$BSNR = 30$	$CGMK$	**4.08**	2.81
	$CGLS$	3.89	2.74
	BMK	3.87	**2.89**
	$BFO1$	3.87	2.63
	$BFO2$	3.55	2.41
$BSNR = 20$	$CGMK$	**3.09**	2.07
	$CGLS$	2.76	1.86
	BMK	3.02	2.13
	$BFO1$	2.87	1.72
	$BFO2$	2.42	1.42

TABLE III: $ISNR$'s for the *Barbara* and *USC-man* images. Experiments using Gaussian blur (variance 9)

Gaussian blur		Barbara	USC-man
$BSNR(dB)$	Method	*ISNR(dB)*	
$BSNR = 40$	$CGMK$	**1.59**	**4.15**
	$CGLS$	1.53	3.91
	BMK	1.58	3.95
$BSNR = 30$	$CGMK$	**1.36**	**3.19**
	$CGLS$	1.30	2.95
	BMK	1.33	2.91
$BSNR = 20$	$CGMK$	**1.16**	**2.20**
	$CGLS$	1.00	1.92
	BMK	1.12	1.72

TABLE IV: $ISNR$'s for the *Barbara* and *USC-man* images. Experiments using uniform 9×9 blur

Uniform blur		Barbara	USC-man
$BSNR(dB)$	Method	*ISNR(dB)*	
$BSNR = 40$	$CGMK$	6.17	7.12
	$CGLS$	**6.23**	**7.70**
	BMK	6.29	7.50
$BSNR = 30$	$CGMK$	**3.86**	**5.26**
	$CGLS$	3.75	4.86
	BMK	3.75	4.89
$BSNR = 20$	$CGMK$	**1.37**	**3.13**
	$CGLS$	1.17	2.69
	BMK	1.20	2.63

Fig. 1: Experiment on *USC-man* image with uniform 9×9 blur and $BSNR = 20$; *ISNR* comparisons: (a) Degraded image, (b) Restored with Bayesian-TV using the proposed algorithm/prior with infinite degrees of freedom, $ISNR = 2.63$, (c) restored image with method in [6], $ISNR = 2.69$, (d) restored image with the proposed algorithm, $ISNR = 3.13$.

[2] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer Verlag, 2006.

[3] J. Bioucas-Dias, M. Figueiredo, J. Oliveira, "Adaptive Bayesian/total-variation image deconvolution: A majorization-minimization approach, " *in Proceedings of European Signal Processing Conference* , Florence, Italy, September 2006.

[4] J. Bioucas-Dias, M. Figueiredo, and J. Oliveira, "Total-variation image deconvolution: A majorization-minimization approach," *in Proceedings of International Conference on Acoustics and Speech and Signal Processing*, ICASSP 2006, May 2006.

[5] S. D. Babacan, R. Molina and A. K. Katsaggelos, "Parameter estimation in TV image restoration using variational distribution approximation," *IEEE Transactions on Image Processing*, vol. 17, no. 3, pp. 326-339, March 2008.

[6] G. Chantas, N. P. Galatsanos, A. Likas and M. Saunders, "Variational Bayesian image restoration based on a product of t-distributions image prior, " *IEEE Trans. on Image Processing, " IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1795-1805, October 2008.

[7] A. L. Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3089-3101, Oct. 2006.