Locally Adaptive Learning for Translation-Variant MRF Image Priors

Masayuki Tanaka and Masatoshi Okutomi
Tokyo Institute of Technology
2-12-1 O-okayama, Meguro-ku, Tokyo, JAPAN
mtanaka@ok.ctrl.titech.ac.jp, mxo@ctrl.titech.ac.jp

Abstract

Markov random field (MRF) models are a powerful tool in machine vision applications. However, learning the model parameters is still a challenging problem and a burdensome task. The main contribution of this paper is to propose a locally adaptive learning framework. The proposed learning framework is simple and effective learning framework for translation-variant MRF models. The key idea is to use neighboring patches as a locally adaptive training set. We use multivariate Gaussian MRF models for local image prior models. Although the Gaussian MRF models are too simple for whole natural image priors, the locally adaptive framework enables to express the prior distributions of the every observed image. These locally adaptive learning framework and the multivariate Gaussian translation-variant MRF models simplify the learning procedures. This paper also includes other two contributions; a novel iteration framework by updating the prior information, and a simple and intuitive derivation of the well-known bilateral filter. Experimental results of denoising applications demonstrate that the denoising based on the proposed locally adaptive learning framework outperforms existing high-performance denoising algorithms.

1. Introduction

Markov random field (MRF) models are a powerful tool in machine vision applications including image denoising [10], super-resolution [6], optical flow estimation [11], and others. In classical MRF models, the relation between neighboring pixels is modeled. Recently, local image patches or small regions, instead of two pixel cliques, are often modeled [5, 10, 12, 16, 17, 19]. This patch-based MRF model is called a high-order MRF model.

The typical local prior distribution is a Gaussian function. However, a Gaussian function is too simple for whole natural image priors. Actually, statistics of whole natural images are very non-Gaussian [19]. In many applications, the simple Gaussian MRF prior leads to over-smoothing.

This over-smoothing can be avoided by using more sophisticated prior models. Roth and Black proposed the Fields of Experts (FoE), which models the local prior distribution using the products of expert functions [10]. The expert functions are defined by filter outputs. In [10], Roth and Black use Student-t expert functions proposed in [20]. Weiss and Freeman proposed another FoE, which uses the Gaussian scale mixture for expert functions [19]. These FoE models are some of successful image prior models. However, the learning of those FoE models is still a hard problem. One reason is that those two FoE models are translation-invariant MRF models which must be a common global prior model for whole natural images. The common global prior model inherently needs to be a complex high-dimensional prior model because the simple prior model can not precisely express the prior distribution of the whole natural images. The learning algorithms of the high-dimensional prior models tends to be complex and computationally expensive. The learning of global prior models also requires a huge training set which represents the whole natural images. Although Tappen proposed a variational optimization to simplify the learning of these FoE models [16], the difficulties of the learning for the global prior models are still remained.

Another approach to avoid the over-smoothing of simple Gaussian prior models is an adaptive approach. Weighted Gaussian conditional random fields (WGCRF) is one of adaptive prior models [17]. In the WGCRF, the local prior distributions are modeled using a set of filter outputs. The products of the Gaussian function of filter outputs express the prior distributions of the local image patches. In addition, weights are assigned adaptively to each filter output of each image patch. However, the adaptive weights assignments are not easy tasks. In addition, the set of filters also needs to be designed appropriately. Elad and Aharon also use adaptive framework for sparse-coding based image priors [4]. The learning of sparse-coding based image priors is also a hard task.

Translation-variant MRF (TV-MRF, in short) models, or locally adaptive MRF models, are also mentioned very briefly in [10, 19]. However, the learning framework for TV-MRF models have not been discussed at all. In this paper, we propose a locally adaptive learning framework for TV-MRF models. The key idea of the proposed learning

978-1-4244-2243-2/08/$25.00 ©2008 IEEE

framework is to use neighboring patches as a locally adaptive training set. The local prior distributions are learned using this locally adaptive training set. This simple idea enables the effective learning of TV-MRF models. In addition, the proposed learning framework is applicable to other adaptive image prior models.

We use multivariate Gaussian TV-MRF models, where prior distributions of the local patches are modeled using a multivariate Gaussian function. Although a single multivariate Gaussian function is insufficient for whole natural image priors, the multivariate Gaussian TV-MRF models perform adaptively; thereby, we can deal with the whole natural images. This idea is similar to WGCRF models [17]. In addition, we show that the multivariate Gaussian prior models are closely related to FoEs and WGCRF models.

This paper also includes two other contributions; a novel iteration framework by updating the prior information and an intuitive derivation of the bilateral filter.

In some applications, image processing or filters are iteratively applied [4, 18]. This iteration is heuristically introduced, and it is experimentally known that the iteration is effective in many cases. We propose a novel iteration framework by updating the prior information, which we call the prior update. The point of the prior update is that the observation must remain unchanged, while the prior information should be updated.

The bilateral filter is well known as a denoising filter [18]. Some theoretical derivations of the bilateral filter in the context of a diffusion process are reported in the literature [1, 3]. We show that the bilateral filter can be simply derived based on the proposed locally adaptive learning framework.

As experiments, we apply to denoising the proposed locally adaptive learning framework. We demonstrate that the proposed denoising algorithm based on our simple MRF approach outperforms existing high-performance wavelet-algorithm and others.

2. Locally adaptive learning framework

We first describe general TV-MRF models. Then, we propose the locally adaptive learning framework to learn the TV-MRF models. A novel iteration framework, which we call prior update, is also proposed. Finally in this section, advantages and limitations of the proposed framework are summarized.

2.1. Translation-variant MRF models

In this paper, we specifically describe high-order TV-MRF models. High-order TV-MRF models are simply given as the translation-variant version of the translation-invariant high-order MRF models. The high-order TV-MRF models can also be considered as a variation of prior models introduced in [22]. The prior distribution of an image $\mathbf{X}$ under the high-order TV-MRF models can be written as

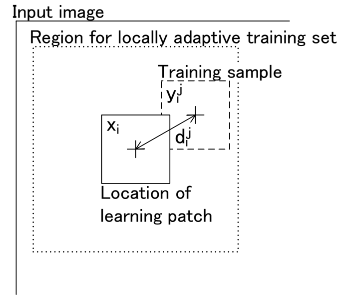


Figure 1. Location of the learning patch and training samples.

$$p(\mathbf{X};\mathbf{\Theta}) = \frac{1}{Z}\prod_{i=1}^{N}\varphi(\mathbf{x}_i;\boldsymbol{\theta}_i) = \frac{1}{Z}\exp\left[-\sum_{i=1}^{N}v(\mathbf{x}_i;\boldsymbol{\theta}_i)\right], \quad (1)$$

where $\mathbf{x}_i$ is the vectorized i-th local image patch, $\varphi(\mathbf{x};\boldsymbol{\theta})$ represents the prior distribution of the local image patch $\mathbf{x}_i$, $v(\mathbf{x};\boldsymbol{\theta})$ is the associated potential function, $\boldsymbol{\theta}_i$ is the prior distribution parameters of the i-th patch, $\mathbf{\Theta}$ is the vector which includes all parameters $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \cdots, \boldsymbol{\theta}_N)$, N is the number of patches which can be generated from the image $\mathbf{X}$, and Z is a normalization term. Note that $\mathbf{X}$ and $\mathbf{x}_i$ are random variables.

The translation-invariant MRF models are a special case of the TV-MRF models. The TV-MRF models with constant prior distribution parameters are translation-invariant MRF models. In this regard, the TV-MRF models are more general than the translation-invariant MRF models.

Without loss of generality, we can assume that the patch is a fixed size of the square patch. A non-square or other shaped patch can be applied.

2.2. Locally adaptive training set and learning

The idea of the TV-MRF models is simple and reasonable. However, the challenge is how to learn the prior distribution parameters of each local image patch appropriately. We propose a locally adaptive learning framework. The key idea is to use neighboring patches as a locally adaptive training set, as shown in Fig. 1. We assume that the statistics of the patches are changed locally, and smoothly. Under this assumption, it is reasonable to use the neighboring patches as the training set. Weights are assigned to each training patch corresponding to the distances between the center of the learning patch and training patches. The locally adaptive training set for a certain patch $\mathbf{x}_i$ can be expressed as $\{(w(d_i^j), \mathbf{y}_i^j)\}$, where $\mathbf{y}_i^j$ is the j-th patch in the input image of neighbor of the patch $\mathbf{x}_i$, d_i^j is the distance between the centers of patches $\mathbf{x}_i$ and $\mathbf{y}_i^j$, and $w(d)$ is a weighting function assigning higher weights for closer patches. The typical weight function is a Gaussian function. Other weighting

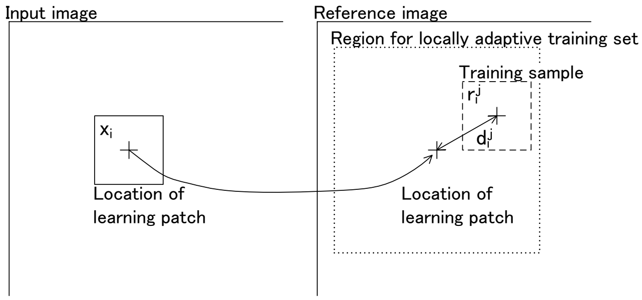


Figure 2. Locally adaptive training samples generated from the reference image.

schemes such as the bilateral filter [18] or the steerable kernel [14] can be applied.

For inpainting applications, some patches include missing pixels. In these cases, patches without missing pixels are used for the locally adaptive training set. Another approach is to apply initial interpolation. The initial interpolation approach is useful for interpolation or super-resolution applications as well.

Once we obtain the locally adaptive training set, we can readily learn or estimate the local prior distribution parameters. A widely used learning algorithm is the maximum likelihood (ML) algorithm. The local prior distribution parameters estimated using the ML algorithm can be written as

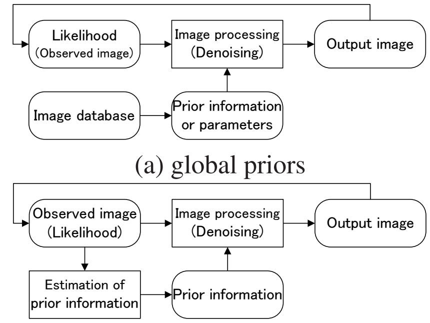
$$\hat{\boldsymbol{\theta}}_i = \underset{\boldsymbol{\theta}_i}{\arg\min} \sum_j w(d_i^j)\, v(\mathbf{y}_i^j; \boldsymbol{\theta}_i)\,. \tag{2}$$

Another estimation approach is a kernel density estimation [7]. Several kernel density estimation algorithms have been proposed. One example of kernel density estimation results is

$$\varphi(\mathbf{x}_i) = \sum_j w(d_i^j) k(\mathbf{x}_i - \mathbf{y}_i^j)\,, \tag{3}$$

where $k(\mathbf{x})$ is a kernel function. Other learning algorithms can be used.

Recently, several image processing techniques using image pairs have been proposed [8, 21]. The proposed locally adaptive learning framework can be extended straightforwardly to the image pair case. In the discussion provided above, the locally adaptive training set is generated from the input image. However, the image from which the locally adaptive training set is generated is not restricted to the input image. The locally adaptive training set can be generated from other images which we call reference images. We assume that the size of the reference image is equal to that of the input image, to simplify the problem. Figure 2 illustrates the relation between the training set and the reference image. The location of the learning patch is mapped into the reference image. Then, the locally adaptive training set is generated from the neighboring patches in the reference image. The locally adaptive training set with the reference image can be expressed as $\{(w(d_i^j), \mathbf{r}_i^j)\}$, where $\mathbf{r}_i^j$ represents the training samples in the reference images. The local prior distribution parameters can be learned with these training set generated from the reference image.



(a) global priors

(b) adaptive priors

Figure 3. Two types of heuristic iteration.

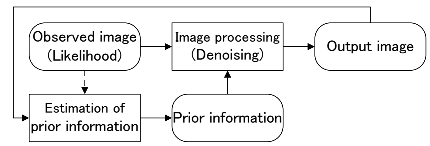


Figure 4. Proposed prior update iteration, where the dashed lines represent the first iteration.

The question is how to prepare the reference image. One answer is to use image pairs as in [8, 21]. The other is to apply iteration. In the next subsection, we describe a novel iteration framework.

2.3. Prior update

In some image processing applications, especially denoising, the image processing or the filters are applied iteratively [4, 18]. Although the iteration is useful for some applications in some circumstances, this kind of iteration is heuristically introduced without sufficient theoretical justification. We call this iteration heuristic iteration. Figure 3 shows two types of heuristic iterations; the iteration with global, or non-adaptive, priors and the iteration with adaptive priors. In both heuristic iterations, it can be considered that the observed image, or the likelihood distribution, is updated through the iterations. It seems unnatural that the observation is modified during image processing. The observation is already done before image processing and must remain unchanged during image processing. In this sense, it is natural that the observation is not modified during the image processing once the observation is obtained.

We propose the novel iteration framework which updates the prior information, while the likelihood or the observed image are unchanged. Figure 4 shows the blockdiagram of the proposed iteration framework. Prior information can be updated when we gain information about the true image. In this sense, the proposed iteration is reasonable compared to the heuristic iterations. We call the proposed iteration the prior update.

We can consider that the last estimation result provides more information about the true image. The proposed iteration framework can be applied to the adaptive image prior models. In the case of the locally adaptive learning framework, the proposed iteration can be performed by using the last estimation as the reference image.

2.4. Advantages and limitations

In the TV-MRF models, we can use simple local prior models. Even if the local prior models are simple, locally adaptive framework allows us to deal with the whole natural images. The learning procedure of the simple local prior models is also simple. This simple learning procedure is one of advantages. In this paper, we use the multivariate Gaussian models for the local prior models. In this case, the learning procedure is given by a closed form as discussed in Section 3.2. In addition, this local learning procedure can be performed independently. This independent learning procedure is very suitable for parallel computing. Recently, parallel computing architectures are developing and improving rapidly. For practical implementations, the independent learning procedure is very useful.

One of other advantages is that the proposed learning framework does not require a huge training set. On the contrast, the learning of the common global MRF models, or the translation-invariant MRF models, requires a huge training set which represents the whole natural images. The reason is that the common global MRF models should essentially express the prior distribution of the whole natural images. Therefore, the training set also needs to represent the whole natural images. Preparing such an enormous training set is an expensive and exhausting task. In addition, learning with such a huge training set is computationally very expensive.

In other words, however, the training samples of the proposed learning framework are small. This small number of the training samples can be a limitation of the proposed learning framework. However, the small training samples is sufficient to learn the TV-MRF models because the local prior models of the TV-MRF models can be simple as discussed above. Therefore, we do not meet this limitation.

The prior update iteration is an additional advantage of the locally adaptive learning framework. The common global prior models are usually learned in advance and can not be updated through the iteration.

3. Multivariate Gaussian TV-MRF models

In this paper, we use the multivariate Gaussian TV-MRF models. We next discuss the reason why we use the multivariate Gaussian model. Then, a concrete learning procedure using the proposed locally learning framework is shown.

3.1. Why multivariate Gaussian?

The Gaussian prior models have been well studied in computer vision fields [13]. The Gaussian models are very effective for learning because the learning procedure can be performed in a closed form.

In the locally adaptive learning framework, the training samples for each local prior distribution are small. The learning of a high-dimensional model with small training samples leads to over-fitting. We must therefore use simple prior models to avoid over-fitting.

It is also known that statistics of whole natural images are very non-Gaussian, as described in [19]. The Gaussian prior models usually cause over-smoothing. However, we can overcome this weakness of the Gaussian prior models using the locally adaptive framework. This idea is similar to Gaussian conditional random fields [17].

For these reasons, we use a multivariate Gaussian model for local prior models of the TV-MRF models.

3.2. Learning procedure

The parameters of the multivariate Gaussian model are the average $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$. We use the ML algorithm to learn these Gaussian parameters. The ML algorithm of the Gaussian parameters has been well described in the literature. The parameters learned with the locally adaptive training set $\{(w(d_i^j), \mathbf{y}_i^j)\}$ are the following.

$$\hat{\boldsymbol{\mu}}_i = \frac{1}{s_i} \sum_j w(d_i^j)\, \mathbf{y}_i^j\,, \quad (4)$$

$$\hat{\boldsymbol{\Sigma}}_i = \frac{1}{s_i} \sum_j w(d_i^j)\, \mathbf{y}_i^j\, {\mathbf{y}_i^j}^T - \hat{\boldsymbol{\mu}}_i\, \hat{\boldsymbol{\mu}}_i^T\,, \quad (5)$$

where, T is the transpose operator and

$$s_i = \sum_j w(d_i^j)\,.$$

We must invert the covariance matrix $\hat{\boldsymbol{\Sigma}}_i$ to evaluate the Gaussian prior. However, in some cases, especially in the texture-less regions, the covariance matrix is singular. Therefore, we approximate the inverse of the covariance matrix as

$$\hat{\boldsymbol{\Sigma}}_i^{-1} \simeq \sum_{k=1}^{M} \frac{1}{\lambda_i^k + \varepsilon} \mathbf{u}_i^k {\mathbf{u}_i^k}^T\,, \quad (6)$$

where λ_i^k is the k-th eigenvalue of $\hat{\boldsymbol{\Sigma}}_i$, $\mathbf{u}_i^k$ is the k-th eigenvector of $\hat{\boldsymbol{\Sigma}}_i$, M is the dimension of $\hat{\boldsymbol{\Sigma}}_i$, and ε is a tiny positive value that is used to avoid divergence.

4. Relations to other algorithms

The proposed framework is closely related to other algorithms. We discuss relations to Fields of Experts and weighted Gaussian conditional random fields. The bilateral filter is also derived from the proposed learning framework.

4.1. Fields of Experts

The multivariate Gaussian TV-MRF models can be expressed as a kind of FoE. Using Eq. 6, the multivariate Gaussian TV-MRF models can be transformed as

$$\begin{aligned}\varphi(\mathbf{x}_i;\boldsymbol{\mu}_i,\boldsymbol{\Sigma}_i) &= \exp\left[-\frac{1}{2}(\mathbf{x}_i-\boldsymbol{\mu}_i)^T\,\boldsymbol{\Sigma}_i\,(\mathbf{x}_i-\boldsymbol{\mu}_i)\right] \\ &= \prod_{k=1}^{M}\exp\left[-\frac{1}{2}\frac{||\mathbf{u}_i^{k^T}\,(\mathbf{x}_i-\boldsymbol{\mu}_i)||_2^2}{\lambda_i^k+\varepsilon}\right], \quad (7)\end{aligned}$$

where $||\cdot||_2^2$ represents the L_2 norm. In this sense, the multivariate Gaussian TV-MRF models can be considered as a variation of FoEs. We again emphasize that the proposed learning framework and MRF models is locally adaptive or translation-variant, whereas existing FoEs models and learning frameworks [19, 10] are only for translation-invariant models. In addition, the learning procedure of the proposed locally adaptive learning framework is given as the simple closed form solution. The learning procedures of existing FoEs are computationally very expensive and require a huge training set.

4.2. Weighted Gaussian conditional random fields

Weighted Gaussian conditional random field (WGCRF) models are also translation-variant models [17]. The local prior models of the WGCRF models can be written as

$$\varphi(\mathbf{x}_i;\boldsymbol{\omega}_i,\mathbf{F},\mathbf{t}) = \prod_k \exp\left[-\omega_i^k(\mathbf{f}^{k^T}\mathbf{x}_i-t_i^k)^2\right] \quad (8)$$

where $\mathbf{f}^k$ is a k-th filter, t_i^k is an estimation of the k-th filter output, and ω_i^k is a locally adaptive weight for the k-th filter.

These two models can be seen as fundamentally identical upon comparison of Eq. 8 with Eq. 7. The difference is that the filters $\mathbf{u}_i^k$ are translation-variant in the multivariate Gaussian TV-MRF models, whereas the filters $\mathbf{f}^k$ are common and must be designed appropriately in the WGCRF models. In this sense, the multivariate Gaussian TV-MRF models are more general than the WGCRF model.

4.3. Bilateral filter

We can derive the bilateral filter [18] based on the proposed locally adaptive learning framework. Several derivations of the bilateral filter in the context of the diffusion process have been reported [1, 3]. We give another derivation from the Bayesian framework. The bilateral filter can be derived as the expectation using the posterior distribution.

Let us consider a 1×1 sized patch for the local prior model in Eq. 1. Actually, a 1×1 sized patch is equivalent to one pixel. Consequently, we learn the local pixel prior distribution with the locally adaptive training set of $\{(w(d_i^j), y_i^j)\}$, where y_i^j is the pixel value of the neighbor pixel of x_i. The local pixel prior distribution is modeled using the kernel density estimation algorithm in Eq. 3. Assuming that the kernel function is the Dirac delta function, we can obtain the local pixel prior distribution as

$$\varphi(x_i) = \sum_j w(d_i^j)\,\delta(x_i-y_i^j), \quad (9)$$

where $\delta(x)$ is the Dirac delta function. Equation 9 can be considered as the direct mathematical formulation of the weighted local histogram. We also assume that the likelihood distribution can be expressed as $p(y_i|x_i)=g(y_i-x_i)$. The function $g(x)$ is a typical Gaussian function.

Using the prior distribution in Eq. 9 and the likelihood distribution $g(y_i-x_i)$, we can derive the definition of the bilateral filter [18] merely by calculating the expectation. The derivation is as shown in the following.

$$\begin{aligned}E_{p(x_i|y_i)}[x_i] &= \frac{\int x_i\,p(y_i|x_i)p(x_i)\,dx_i}{\int p(y_i|x_i)p(x_i)\,dx_i} \\ &= \frac{\int x_i\,g(y_i-x_i)\left\{\sum_j w(d_i^j)\,\delta(x_i-y_i^j)\right\}dx_i}{\int g(y_i-x_i)\left\{\sum_j w(d_i^j)\,\delta(x_i-y_i^j)\right\}dx_i} \\ &= \frac{\sum_j w(d_i^j)\,g(y_i-y_i^j)\,y_i^j}{\sum_j w(d_i^j)\,g(y_i-y_i^j)}. \quad (10)\end{aligned}$$

Equation 10 is identical to the definition of the bilateral filter. This derivation is very simple and intuitive. Many variations of the bilateral filter, such as a patch-based bilateral filter, can be developed based on this derivation.

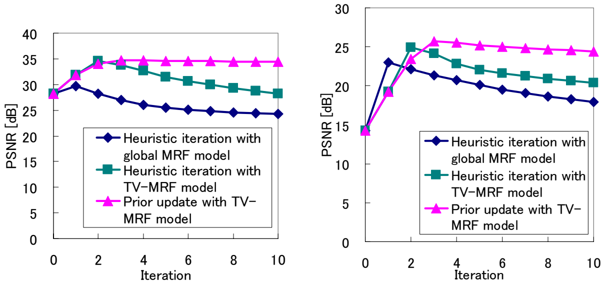
5. Denoising application and experiments

5.1. Denoising algorithm

We formulate a denoising algorithm using a multivariate Gaussian TV-MRF model. Denoising is performed using the maximum a posteriori (MAP) algorithm. The posterior distribution under the TV-MRF models can be expressed as

$$\begin{aligned}p(\mathbf{X}|\mathbf{Y}) &\propto p(\mathbf{Y}|\mathbf{X};\boldsymbol{\alpha})\,p(\mathbf{X};\boldsymbol{\Theta}) \\ &\propto \prod_{i=1}^{N} p(\mathbf{y}_i|\mathbf{x}_i;\boldsymbol{\sigma}_i)\,p(\mathbf{x}_i;\boldsymbol{\theta}_i), \quad (11)\end{aligned}$$

where $\mathbf{X}$ are random variables of the image, $\mathbf{x}_i$ are random variables of the i-th patch, $\mathbf{Y}$ is the observed or input image, $\mathbf{y}_i$ is i-th patch of the image $\mathbf{Y}$, $\boldsymbol{\alpha}$ are likelihood parameters of the image, $\boldsymbol{\sigma}_i$ are likelihood parameters of the i-th patch, $\boldsymbol{\Theta}$ are prior parameters of the image, $\boldsymbol{\theta}_i$ are prior parameters of the i-th patch, and N is the number of patches. We assume an independent identically distributed (i.i.d.) Gaussian whose standard deviation is σ for likelihood. Other likelihood models can be assumed. The MAP cost function



(a) Noise level is 10 (b) Noise level is 50

Figure 5. Comparison of iteration frameworks using Barbara.

to be minimized is the log posterior distribution. The cost function is

$$I = \sum_{i=1}^{N} \left[\frac{1}{\sigma^2} ||\mathbf{y}_i - \mathbf{x}_i||_2^2 + (\mathbf{x}_i - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_i) \right] .$$

Theoretically, the optimal solution of this cost function is given by a closed form solution. However, the calculation of the closed form solution is infeasible because it is a very large-scale problem. We use a conjugate gradient method for optimization of the cost function.

5.2. Comparisons of iteration frameworks

We compare three iteration types: heuristic iteration with the global MRF model, heuristic iteration with the TV-MRF model, and the prior update with the TV-MRF model. The multivariate Gaussian functions are used for the MRF models. The global MRF model is learned with the training images in [15]. The TV-MRF models are learned using the proposed locally adaptive learning framework. The ML algorithm described in Section 3.2 is applied to estimate Gaussian parameters. We use 4×4 sized of patches to model the local prior distribution. In the locally adaptive learning framework, the Gaussian function whose standard deviation is 4.0 [pixel] is used for the weighting function, $w(d)$.

Figure 5 shows comparisons of peak-signal to noise ratios (PSNRs) using the standard image of Barbara. A higher PSNR indicates a better result. Regarding both noise levels, the PSNRs of the TV-MRF models are higher than those of the global MRF model. The PSNR of the prior update with the TV-MRF model converges at high PSNR, whereas the PSNR of the heuristic iteration with the TV-MRF model decreases with the number of iterations. The prior update with the TV-MRF model is more useful than the heuristic iteration with the TV-MRF model because we do not need to set an appropriate iteration number.

5.3. Comparisons of denoising performance

For this study, we uses commonly used images in denoising experiments [9], as shown in Fig. 6.

The proposed denoising algorithm is the proposed prior update with the multivariate Gaussian TV-MRF models. The priors are learned with the locally adaptive learning framework. The iteration number is experimentally set as three from Fig. 5. Other parameters are same in Section 5.2. We also apply the algorithm described by Portilla *et al.* [9] to the same test images for comparison. The algorithm presented by Portilla *et al.* is known as a high-performance wavelet-algorithm; the code is available from their website. Table 1 provides PSNRs of the denoising results for the test images with various levels of additive Gaussian noise. We also put PSNR values of Roth and Black's results in [10] to Table 1. The bold numbers in Table 1 indicate the highest PSNR among the three algorithms. Because the range of images is 0–255, noise levels of 50 or higher are not practical situations. For practical noise levels, the proposed algorithm outperforms the other two algorithms. It is remarkable that the proposed simple MRF approach outperforms such sophisticated wavelet algorithms. Actually, Portilla *et al.* report that their algorithm is tuned to perform well on these test images [9]. Roth and Black also described that no MRF approach had outperformed such a wavelet algorithm on these test images at the time that their study was done [10].


Barbara Boat House Lena Peppers

Figure 6. Test images.

For objective assessments, Figs. 7 and 8 show denoising results of the proposed algorithm and that presented by Portilla *et al.* Both algorithms' results are effectively denoised, preserving edges. These experimental results demonstrate the high-performance of the proposed MRF-based denoising.

5.4. Comparisons of real color images

The proposed multivariate Gaussian TV-MRF models and learning framework can be simply extended to the color images. The proposed denoising algorithm can be extended as well. We demonstrate the denoising effect using real color images. The observed images are captured with high gain. Noisy observed images are denoised using three algorithms: those of Black *et al.* [2] and Portilla *et al.* [9], as well as the proposed algorithm. For the first two, we apply their algorithms for each color channel separately. Three denoising algorithms commonly require a noise level. We manually set the noise level as 16 because the true noise level is unknown. Figure 9 shows the noisy observed images and denoising results. These results demonstrate that the proposed algorithm can more effectively denoise while preserving edges than the other two algorithms.

Table 1. PSNR comparisons with test images, where PSNR is given in decibels, and bold numbers indicate the highest PSNR.

Noise	Barbara			Boat			House			Lena			Peppers		
level	[9]	[10]	Pro.	[9]	[10]	Pro.	[9]	[10]	Pro.	[9]	[10]	Pro.	[9]	[10]	Pro.
2	43.34	42.92	**43.76**	43.04	42.28	**43.19**	43.98	44.01	**44.33**	43.26	42.92	**43.71**	43.01	42.96	**43.37**
5	37.84	37.19	**38.35**	36.96	36.27	**37.26**	38.37	38.23	**38.97**	38.42	38.12	**38.70**	37.42	37.63	**37.82**
10	34.07	32.83	**34.73**	33.54	33.05	**33.64**	35.04	35.06	**35.52**	35.48	35.02	**35.59**	33.99	34.28	**34.30**
15	31.89	30.22	**32.60**	31.64	31.22	**31.67**	33.29	33.48	**33.62**	**33.76**	33.27	33.72	32.02	32.03	**32.28**
20	30.24	28.32	**30.93**	**30.33**	29.85	30.27	32.02	32.17	**32.25**	**32.51**	31.92	32.35	30.63	30.58	**30.84**
25	29.02	27.04	**29.64**	**29.32**	28.72	29.17	31.01	31.11	**31.12**	**31.55**	30.82	31.24	29.57	29.20	**29.70**
50	25.30	23.15	**25.64**	**26.36**	24.53	25.86	**27.82**	26.74	27.38	**28.48**	26.49	27.61	**26.38**	24.52	26.09
75	**23.50**	21.36	23.39	**24.63**	22.48	23.87	**26.18**	24.13	25.16	**26.76**	24.13	25.53	**24.57**	21.68	23.97
100	**22.49**	19.77	22.01	**23.57**	20.80	22.61	**24.88**	21.66	23.57	**25.55**	21.87	24.05	**23.33**	19.60	22.52

[9]: Portilla *et al*. [9], [10]: Roth and Black [10], Pro.: Proposed algorithm

(a) Original image (b) Noisy image (PSNR: 28.24 [dB]) (c) Image denoised using [9] (PSNR: 34.07 [dB]) (d) Image denoised using the proposed algorithm (PSNR: 34.73 [dB])

Figure 7. Comparisons of denoising results of Barbara, where the noise level is 10 (assessments by display are preferred to those by print).

(a) Original image (b) Noisy image (PSNR: 22.20 [dB]) (c) Image denoised using [9] (PSNR: 32.51 [dB]) (d) Image denoised using the proposed algorithm (PSNR: 32.35 [dB])

Figure 8. Comparisons of denoising results of Lena, for which the noise level is 20 (assessments by display are preferred to those by print).

6. Conclusions

We have proposed the locally adaptive learning framework. The proposed learning framework is simple and effective learning for the TV-MRF models. The key idea is to use neighboring patches as a locally adaptive training set. This simple idea enables effective learning of the TV-MRF models. In addition, a huge training set is not required for learning.

We use the multivariate Gaussian TV-MRF models. Although the Gaussian TV-MRF models are very simple, the locally adaptive learning framework allows us to deal with the whole natural images. These TV-MRF models and learning framework simplify the learning procedures.

We have also proposed the prior update iteration based on the Bayesian framework. In the proposed iteration, the prior information is updated, while the likelihood or the observed image is unchanged. It is reasonable compared to heuristic iterations.

We clarify the relation between the multivariate Gaussian TV-MRF models and other algorithms. The bilateral filter is simply and intuitively derived based on the proposed locally adaptive framework. The proposed framework can be said to be a very general framework.

The denoising applications based on the proposed TV-MRF models and learning framework are demonstrated using the standard test images and the real color images. We show that the proposed denoising application outperforms existing high-performance denoising algorithms under practical noise levels.

Our future works shall include development of other applications based on the proposed TV-MRF models and learning framework.

(a) Noisy observed images (b) Images denoised using [2] (c) Images denoised using [9] (d) Images denoised using the proposed algorithm

Figure 9. Examples of real color image denoising (assessments by display are preferred to those by print).

References

[1] D. Barash, "Fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 844–847, 2002. 2, 5

[2] M. Black, G. Sapiro, D. Marimont, and D. Heeger, "Robust anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 421–432, 1998. 6, 8

[3] M. Elad, "On the origin of the bilateral filter and ways to improve it," *IEEE Transactions on Image Processing*, vol. 11, no. 10, pp. 1141–1151, 2002. 2, 5

[4] M. Elad and M. Aharon, "Image Denoising Via Learned Dictionaries and Sparse representation," *Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, 2006. 1, 2, 3

[5] X. Lan, S. Roth, D. Huttenlocher, and M. Black, "Efficient Belief Propagation with Learned Higher-Order Markov Random Fields," *Lecture Notes In Computer Science*, vol. 3952, pp. 269, 2006. 1

[6] S. Park, M. Park, and M. Kang, "Super-resolution image reconstruction: a technical overview," *Signal Processing Magazine, IEEE*, vol. 20, no. 3, 2003, pp. 21–36. 1

[7] E. Parzen, "On Estimation of a Probability Density Function and Mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, 1962. 3

[8] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," *ACM Transactions on Graphics (TOG)*, vol. 23, no. 3, pp. 664–672, 2004. 3

[9] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1338–1351, 2003. 6, 7, 8

[10] S. Roth and M. Black, "Fields of experts: A framework for learning image priors," *IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp. 860–867, 2005. 1, 5, 6, 7

[11] S. Roth and M. Black, "On the Spatial Statistics of Optical Flow," *International Journal of Computer Vision*, vol. 74, no. 1, pp. 33–50, 2007. 1

[12] K. Shi and S.-C. Zhu, "Mapping Natural Image Patches by Explicit and Implicit Manifolds," *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1–7, 2007. 1

[13] R. Szeliski, "Bayesian modeling of uncertainty in low-level vision," *International Journal of Computer Vision*, vol. 5, no. 3, pp. 271–301, 1990. 4

[14] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel Regression for Image Processing and Reconstruction," *IEEE Trans. on Image Processing*, vol. 16, no. 2, pp. 349–366, 2007. 3

[15] D. Tal and J. Malik, "A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics," *Proc. Eighth Int'l. Conf. Computer Vision*, vol. 2, pp. 416–423, 2001. 6

[16] M. F. Tappen, "Utilizing Variational Optimization to Learn Markov Random Fields," *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1–8, 2007. 1

[17] M. F. Tappen, C. Liu, E. H. Adelson, and W. T. Freeman, "Learning Gaussian Conditional Random Fields for Low-Level Vision," *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1–8, 2007. 1, 2, 4, 5

[18] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," *Proc. of the Sixth International Conference on Computer Vision*, pp. 839–846, 1998. 2, 3, 5

[19] Y. Weiss and W. T. Freeman, "What makes a good model of natural images? *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1–8, 2007. 1, 4, 5

[20] M. Welling, G. Hinton, and S. Osindero, "Learning sparse topographic representations with products of student-t distributions," *Advances in Neural Information Processing Systems*, vol. 15, pp. 1359–1366, 2003. 1

[21] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, "Image Deblurring with Blurred/Noisy Image Pairs," *ACM SIGGRAPH 2007 papers*, vol. 26, no. 3-1, pp. 1–10, 2007. 3

[22] S. Zhu and D. Mumford, "Prior learning and Gibbs reaction-diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 11, pp. 1236–1250, 1997. 2