PHOTON-LIMITED IMAGE DENOISING BY INFERENCE ON MULTISCALE MODELS

Stamatios Lefkimmiatis, George Papandreou and Petros Maragos

National Technical University of Athens, School of ECE, Zografou, Athens 15773, Greece.

Email:[sleukim,gpapan,maragos]@cs.ntua.gr

ABSTRACT

We present an improved statistical model of Poisson processes, with applications to photon-limited imaging. We build on previous work, adopting a multiscale representation of the Poisson process in which the ratios of the underlying Poisson intensities (rates) in adjacent scales are modeled as mixtures of conjugate parametric distributions. Our main novel contributions are (1) a rigorous and robust regularized Expectation-Maximization (EM) algorithm for maximum-likelihood estimation of the rate-ratio density parameters directly from the observed Poisson data (counts); (2) extension of the method to work under a scale-recursive Hidden Markov Tree model (HMT) which couples the mixture label assignments in consecutive scales, thus modeling inter-scale coefficient dependencies in the vicinity of edges; and (3) exploration of a fully 2-D quad-tree image partitioning, involving Dirichlet-mixture rate-ratio densities, instead of the conventional separable binary image partitioning involving Beta-mixture rate-ratio densities. Experimental intensity estimation results on standard images with artificially simulated Poisson noise and photon-limited images with real shot noise demonstrate the effectiveness of the proposed approach.

Index Terms— Photon-Limited Imaging, Poisson , Bayesian inference, Hidden Markov tree, Expectation-Maximization algorithm.

1. INTRODUCTION

Image acquisition by photon imaging systems is accomplished by counting photon detections at different spatial locations over a given observation period. The quality of captured images is degraded by the so called quantum or shot noise, due to fluctuations on the number of detected photons, a consequence of the discrete nature of the detection process. The number of detected photons in each pixel location follows a Poisson distribution, whose underlying intensity/rate corresponds to the desired clean source image. The described image formation model arises in various image processing applications, notably medical and astronomical imaging. Thus the problem of estimating the discretized underlying intensity of a Poisson process, given noisy observations, is of fundamental importance.

There is a variety of Poisson intensity estimation techniques proposed in the literature [1], with the Bayesian methods allowing incorporation of prior knowledge about the underlying intensity to be estimated. In particular, multiscale Bayesian methods are becoming increasingly popular since they can provide significant simplifications to the problem. In this work we follow the multiscale representation of Poisson processes of [2] and [3]. The adopted representation leads to a multiscale factorization of the Poisson process likelihood function, which, combined with a factorized conjugate parametric mixture model for the image intensity prior, substantially simplifies the intensity estimation problem, as discussed in Section 2. Inter-scale dependencies between mixture assignments can be efficiently modeled with a Hidden Markov Tree (HMT) structure [4, 5, 6].

A significant shortcoming of previous approaches [2, 3, 7] utilizing the above framework is that they do not properly address model parameter estimation from the noisy image. Our main contribution is an Expectation-Maximization (EM) technique for maximum-likelihood estimation of these parameters directly from the observed Poisson data, so that the model accurately matches the statistics of the source image. The robustness of the technique is enhanced by incorporating a simple regularization term, as we discuss in Section 3. The parameter estimation issue has been even more pronounced in the case of HMT-based approaches, where the HMT-specific parameters also need to be fitted. In Section 4 we extend our EM-based parameter estimation method to work with HMT-based models; this can provide further benefits in the intensity estimation problem. The proposed techniques can equally well apply to estimators based on either conventional separable dyadic-tree or fully 2-D quad-tree image partitioning which we explore in Section 2.2. Finally, in Section 5 we present image denoising experiments on both standard images with artificially simulated Poisson noise and photon-limited images, which demonstrate the effectiveness of the proposed methods.

2. MULTISCALE MODELING OF POISSON PROCESSES

2.1. Dyadic-Tree Image Partitioning

Assume that the vector $\mathbf{x} = \begin{bmatrix} x_0 & x_1 & \dots & x_{N-1} \end{bmatrix}$ consists of observation samples of N independent random variables X_k, $k = 0, \dots, N-1$, following Poisson distribution with intensities $\boldsymbol{\lambda} = \begin{bmatrix} \lambda_0 & \lambda_1 & \dots & \lambda_{N-1} \end{bmatrix}$. We aim at estimating the underlying source image intensity $\boldsymbol{\lambda}$ from the noisy observations $\mathbf{x}$. If we represent with $\mathbf{x}_0$ and $\boldsymbol{\lambda}_0$ the finest representation of $\mathbf{x}$ and $\boldsymbol{\lambda}$ respectively, i.e $\mathbf{x}_0 = \mathbf{x}$ and $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}$, then a multiscale analysis of $\mathbf{x}$ and $\boldsymbol{\lambda}$ can be obtained through the following formulas, which resemble the recursions that yield the Haar wavelet transform scaling coefficients [2]:

$$x_{j,k} = x_{j-1,2k} + x_{j-1,2k+1} \tag{1}$$

$$\lambda_{j,k} = \lambda_{j-1,2k} + \lambda_{j-1,2k+1} \tag{2}$$

for $j = 1, \dots, J$ and $k = 0, \dots, N/2^j - 1$, and $J = \log_2(N)$. In the above equations j denotes the scale of analysis (J is the coarsest scale) and k the position in the corresponding vector $\mathbf{x}_j$. This decomposition is motivated by two fundamental properties of Poisson processes: (1) the number of counts over nonoverlapping intervals are independent, given the underlying intensities and (2) the sum of independent Poisson random variables remains Poisson. Thus, the random variable $X_{j,k}$ will remain Poisson distributed with intensity $\lambda_{j,k}$. Moreover, since for two Poisson random variables, $X_1 \sim \text{Pois}(\lambda_1)$ and $X_2 \sim \text{Pois}(\lambda_2)$, the conditional distribution $p(X_1 | X_1 + X_2)$ is Binomial, it holds that $p(x_{j-1,2k} | x_{j,k}) =$

This work was supported by grants ΠΕΝΕΔ 2003-ΕΔ 554 & 865 [co-financed by E.U.-European Social Fund (75%) and the Greek Ministry of Development-GSRT (25%)] and the EC projects ASPI and MUSCLE.

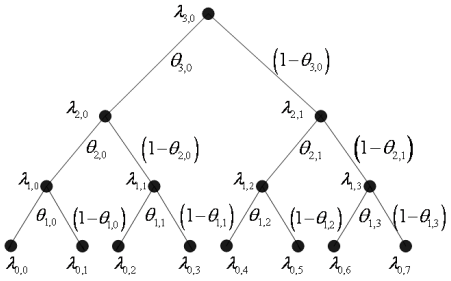


Fig. 1. Multiscale Dyadic Partitioning in 3 Scales.

$\text{Bin}\left(x_{j-1,2k}|x_{j,k}, \frac{\lambda_{j-1,2k}}{\lambda_{j,k}}\right)$. This statistical relation over adjacent scales permits the factorization of the likelihood function in the following form [3]:

$$p(\mathbf{x}|\boldsymbol{\lambda}) = p(x_{J,0}|\lambda_{J,0}) \cdot \prod_{j=1}^{J} \prod_{k=0}^{N/2^j-1} \text{Bin}(x_{j-1,2k}|x_{j,k}, \theta_{j,k}) \quad (3)$$

where $\theta_{j,k} = \frac{\lambda_{j-1,2k}}{\lambda_{j,k}}$ is the success rate of the Binomial distribution which can be interpreted as a splitting factor [5] that governs the multiscale refinement of the intensity $\boldsymbol{\lambda}$.

Observing the binary tree in Fig. 1 produced by the scale-recursion (2) on $\boldsymbol{\lambda}$, we notice that $\boldsymbol{\lambda}$ has an equivalent vector parameterization $\boldsymbol{\lambda}_e = \left[\lambda_{J,0}, \boldsymbol{\theta}\right] = \left[\lambda_{J,0}, \theta_{J,0}, \ldots, \theta_{1,0} \ldots, \theta_{1,N/2-1}\right]$. If we consider $\lambda_{J,0}$ and $\theta_{j,k}$ as the observation samples of a random variable $\Lambda_{J,0}$ and $\Theta_{j,k}$ respectively, where the random variables are statistically independent, then the prior distribution $p(\boldsymbol{\lambda})$ can be expressed in the following factorized form:

$$p(\boldsymbol{\lambda}) = p(\lambda_{J,0}) \cdot p(\boldsymbol{\theta}) = p(\lambda_{J,0}) \cdot \prod_{j=1}^{J} \prod_{k=0}^{N/2^j-1} p(\theta_{j,k}) \quad (4)$$

Since the Beta distribution is conjugate to the Binomial [8], it is convenient to express $p(\theta_{j,k})$ as mixture of Beta distributions, $p(\theta_{j,k}) = \sum_{m=1}^{M} \pi_{j,m} \text{Beta}(\theta_{j,k}|\alpha_{j,m}, \beta_{j,m})$, where $\pi_{j,m}$ is the mixture weight for the mth mixture component in jth scale of analysis, $\alpha_{j,m}$ and $\beta_{j,m}$ are the parameters of the Beta mixture component and M is the total number of mixture components in each scale.

The described statistical framework, resulting in the factorization of the likelihood function and the prior density, is appropriate for 1-D signals. For the image case, $\mathbf{x}$ and $\boldsymbol{\lambda}$ will be 2-D images consisting of the discrete observation samples and values of the intensity function, respectively. In this case, a possible multiscale analysis can be achieved by applying the 1-D model separably [2]; each decomposition level will consist of one decimation step across the horizontal and one across the vertical direction. Hereafter, we will refer to this multiscale framework as *separable* model.

2.2. Quad-Tree Image Partitioning

Instead of using the separable model to treat images, we also develop a fully 2-D multiscale decomposition on a quad-tree. Denoting for each jth-scale pixel location (k, ℓ) the set of children pixel locations at the next finer scale as $C_{k,l} = \{(2k, 2\ell), (2k, 2\ell+1), (2k+1, 2\ell), (2k+1, 2\ell+1)\}$, the decomposition formulas analogous to (1) and (2) will be:

$$x_{j,k,\ell} = \sum_{(k',\ell') \in C_{k,\ell}} x_{j-1,k',\ell'}, \quad \lambda_{j,k,\ell} = \sum_{(k',\ell') \in C_{k,\ell}} \lambda_{j-1,k',\ell'} \quad (5)$$

for $j = 1, \ldots J$, $k, \ell = 0, \ldots, N/2^j - 1$, and $J = \log_2(N)$. The random variable $X_{j,k,\ell}$ will remain Poisson distributed with intensity $\lambda_{j,k,\ell}$. Moreover it will also hold that, given $X_{j,k,l}$, the conditional distribution of the four children's random vector $\mathbf{X}^c_{j,k,\ell} = \{X_{j-1,k',\ell'}\}_{(k',\ell') \in C_{k,\ell}}$ will be Multinomial: $p\left(\mathbf{x}^c_{j,k,\ell}|x_{j,k,\ell}\right) = \text{Mult}\left(\mathbf{x}^c_{j,k,\ell}|x_{j,k,\ell}, \boldsymbol{\theta}_{j,k,\ell}\right)$, with $\boldsymbol{\theta}_{j,k,\ell} = \left\{\frac{\lambda_{j-1,k',\ell'}}{\lambda_{j,k,\ell}}\right\}_{(k',\ell') \in C_{k,\ell}}$. The likelihood function and prior distribution are then factorized as:

$$p(\mathbf{x}|\boldsymbol{\lambda}) = p(x_{J,0,0}|\lambda_{J,0,0}) \prod_{j=1}^{J} \prod_{k,\ell=0}^{N/2^j-1} \text{Mult}\left(\mathbf{x}^c_{j,k,\ell}|x_{j,k,\ell}, \boldsymbol{\theta}_{j,k,\ell}\right) \quad (6)$$

$$p(\boldsymbol{\lambda}) = p(\lambda_{J,0,0}) \cdot \prod_{j=1}^{J} \prod_{k,\ell=0}^{N/2^j-1} p(\boldsymbol{\theta}_{j,k,\ell}) \,. \quad (7)$$

Since the Dirichlet distribution is conjugate to the Multinomial [8] it is convenient to model $p(\boldsymbol{\theta}_{j,k,\ell})$ as mixture of Dirichlets, $p(\boldsymbol{\theta}_{j,k,\ell}) = \sum_{m=1}^{M} \pi_{j,m} \text{Dir}(\boldsymbol{\theta}_{j,k,\ell}|\boldsymbol{\alpha}_{j,m})$, where $\boldsymbol{\alpha}_{j,m}$ is the parameter vector of the mth mixture component in the jth scale of analysis. We will refer hereafter to this multiscale model as *quad* model.

2.3. Posterior Mean Estimation

Given the factorization of the prior and likelihood on multiple scales, it is straightforward to prove that the posterior distribution $p(\boldsymbol{\lambda}|\mathbf{x}) = p(\lambda_{J,0,0}|x_{J,0,0}) \cdot p(\boldsymbol{\theta}|\mathbf{x})$ also bears a multiscale factorization. The factorization of the posterior implies that the intensity estimation problem can be solved individually on each scale, instead of requiring a complicated high dimensional solution. For the quad model the Bayes posterior mean estimator $E[\boldsymbol{\theta}_{j,k,\ell}|\mathbf{x}]$ for each scale of analysis, will be given by:

$$\hat{\theta}^i_{j,k,\ell} = \sum_{m=1}^{M} \gamma_m(z_{j,k,\ell}) \left(\frac{x^{c,i}_{j,k,\ell} + \alpha^i_{j,m}}{x_{j,k,\ell} + \alpha^0_{j,m}}\right), \quad (8)$$

where the superscript $i = 1, \ldots, 4$ denotes the ith element of the corresponding vector, $\alpha^0_{j,m} = \sum_{i=1}^{4} \alpha^i_{j,m}$, and

$$\gamma_m(z_{j,k,\ell}) = \frac{\pi_{j,m} \text{Polya}\left(\mathbf{x}^c_{j,k,\ell}|x_{j,k,\ell}, \boldsymbol{\alpha}_{j,m}\right)}{\sum_{n=1}^{M} \pi_{j,n} \text{Polya}\left(\mathbf{x}^c_{j,k,\ell}|x_{j,k,\ell}, \boldsymbol{\alpha}_{j,n}\right)} \quad (9)$$

is the responsibility of the mth mixture component in the jth scale given the noisy observations $\mathbf{x}$, and Polya (also called Dirichlet-Multinomial) is the discrete multivariate distribution given by

$$\text{Polya}(\mathbf{x}|n, \boldsymbol{\alpha}) = \frac{n!}{\prod_{i=1}^{D} x_i!} \cdot \frac{\Gamma\left(\sum_{i=1}^{D} \alpha_i\right)}{\Gamma\left(n + \sum_{i=1}^{D} \alpha_i\right)} \cdot \prod_{i=1}^{D} \frac{\Gamma(x_i + \alpha_i)}{\Gamma(\alpha_i)} \quad (10)$$

and $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} \mathrm{d}t$ [9]. For the separable model the corresponding estimator is similar to (8) and is obtained from Eqs. (8), (9) by replacing $x^{c,i}_{j,k,\ell}$ with $x_{j-1,2k}$, $\alpha^i_{j,m}$ with $\alpha_{j,m}$, $\alpha^0_{j,m}$ with $(\alpha_{j,m} + \beta_{j,m})$, $\boldsymbol{\alpha}_{j,m}$ with the vector $(\alpha_{j,m}, \beta_{j,m})$, and removing the subindex ℓ.

For the coarsest scale of analysis a satisfactory choice is to estimate $\lambda_{J,0,0}$ by $\hat{\lambda}_{J,0,0} = x_{J,0,0}$ [2]. Alternatively, we can choose a Gamma prior for $\lambda_{J,0,0}$ and obtain the corresponding mean estimate.

3. EM-BASED MODEL PARAMETER ESTIMATION

The posterior mean estimates obtained in Section 2.3 require knowledge of the model parameters $\boldsymbol{\varrho} = \{\boldsymbol{\pi}_j, \boldsymbol{\alpha}_j\}$, i.e. the mixture weights and the Beta/Dirichlet mixture parameters governing the splitting factors $\Theta_{j,k,l}$. However, estimation of $\boldsymbol{\varrho}$ has not been

addressed satisfactorily in previous work. The authors in [2] model the random variables $\Theta_{j,k}$ with mixtures of 3 symmetric Betas where the Beta parameters are heuristically set for every scale of analysis. They also assume that all but one mixture weights are known and estimate it by the method of moments. The approach followed in [7] utilizes an auxiliary wavelet thresholding denoising method to obtain an estimate of the splitting factors $\boldsymbol{\theta}_{j,k}$, on which $\boldsymbol{\varrho}$ is fitted with EM; thus it is prone to a potential failure of the auxiliary method.

To account for these limitations, we pursue a Maximum-Likelihood (ML) estimator to infer model parameters $\boldsymbol{\varrho}$ *directly* from the observed Poisson data. Since the multiscale approach allows handling each scale independently, we drop index j for clarity. Also, instead of (k,ℓ), we use the index t to denote position, assuming that we have raster-scanned the image at each scale into a vector $\mathbf{x}$ of length T. The key idea is to integrate-out the unobserved splitting factors $\boldsymbol{\theta}_t$, and thus relate the model parameters $\boldsymbol{\varrho}$ directly to the observations $\mathbf{x}$; the integral can be computed in closed form, yielding:

$$\begin{aligned} p\left(\mathbf{x}|\boldsymbol{\varrho}\right) &= \prod_{t=1}^{T} \int_{\boldsymbol{\theta}_t} p\left(x_t, \boldsymbol{\theta}_t|\boldsymbol{\varrho}\right) \mathrm{d}\boldsymbol{\theta}_t \\ &= \prod_{t=1}^{T} \sum_{m=1}^{M} \pi_m \mathrm{Polya}\left(\mathbf{x}_t^c|x_t, \boldsymbol{\alpha}_m\right) \end{aligned} \tag{11}$$

where $\mathbf{x}_t^c$ are the Poisson counts of x_t's children.

The ML parameters maximize the log-likelihood $L\left(\boldsymbol{\varrho}\right) = \log p\left(\mathbf{x}|\boldsymbol{\varrho}\right)$. Optimizing $L\left(\boldsymbol{\varrho}\right)$ directly is hard, due to the multiple mixture components. We thus invoke the EM algorithm [8], and work with the joint log-likelihood, $L_c\left(\boldsymbol{\varrho}\right) = \log p\left(\mathbf{x}, \mathbf{z}|\boldsymbol{\varrho}\right)$, where $\mathbf{z}$ is the vector of the discrete mixture component assignments, i.e. $Z_t = m$, $m = 1, \ldots, M$, if the mth mixture is responsible for generating the observations $\mathbf{x}_t^c$. Parameter estimation then alternates between the Expectation (E-step) and Maximization (M-step) steps until convergence. In the E-step we compute the conditional expectation of the complete log-likelihood given the observations and the current estimates of the model parameters,

$$\begin{aligned} E[L_c\left(\boldsymbol{\varrho}\right)|\mathbf{x}, \boldsymbol{\varrho}^{(n)}] &= \sum_{t=1}^{T} \sum_{m=1}^{M} \gamma_m(z_t) \log\left(\pi_m\right) \\ &+ \sum_{t=1}^{T} \sum_{m=1}^{M} \gamma_m(z_t) \log \mathrm{Polya}\left(\mathbf{x}_t^c|x_t, \boldsymbol{\alpha}_m\right) \end{aligned} \tag{12}$$

where $\gamma_m(z_t)$ is defined in (9). In the M-step new estimates for the parameters of the model are obtained by maximizing (12) with respect to $\{\boldsymbol{\varrho}\}$. The updated mixture weight is $\pi_m^{(n+1)} = 1/T \sum_{t=1}^{T} \gamma_m(z_t)$. Maximizing (12) with respect to $\boldsymbol{\alpha}_m$ leads to a nonlinear equation which we solve by Newton-Raphson.

For images with flat cartoon-like content, the histogram of splitting factors $\boldsymbol{\theta}$ is strongly peaked (at 1/2 for the separable and 1/4 for the quad model), resulting to very large α_m^i parameters. The ML criterion then leads to overfitting, and the Newton-Raphson method is unstable. We have addressed this issue by adding a regularization term $-\epsilon \sum_{m=1}^{M} \sum_{i=1}^{4} \alpha_m^i$ in (12), with ϵ small positive constant. This term is interpretable as specifying a conjugate prior for $\boldsymbol{\alpha}_m$ and leads to a MAP instead of the standard ML estimation. The resulting penalized EM algorithm is extremely robust in practice.

4. INFERENCE ON HIDDEN MARKOV TREES

In Section 2 we adopted the simple assumption that the splitting factors $\Theta_{j,k,l}$ are independent, yielding the factorized prior $p\left(\boldsymbol{\lambda}\right)$ in the form of Eqs. (4) and (7). In this section we follow instead the HMT model first introduced in the context of signal denoising in [4], which better models inter-scale dependencies [4] and can thus provide additional benefits in the intensity estimation problem. While using the HMT model in conjuction with photon-limited imaging has been previously suggested by [5], wider adoption of the HMT model in this context has been hindered so far by lack of satisfactory solution to the model parameter estimation problem. We address this shortcoming by properly extending our EM-based technique of Section 3 to the HMT case.

More specifically, let $\boldsymbol{\varrho} = \{\boldsymbol{\pi}, A, \boldsymbol{\alpha}\}$ be the vector of HMT parameters corresponding to prior root node state probabilities, inter-scale state transition probabilities, and parameters for each mixture, respectively. Also let t index the $T+1$ nodes of the tree, where $t=0$ is the root node, $t=T$ is the last node of the finest scale, $p(t)$ is t's parent and $\mathbf{x}_t^c$ is the vector of x_t's children. Then the *apriori* probability of a hidden state path $\mathbf{z} = (z_0, \ldots, z_T)$ under the model is $P\left(\mathbf{z}\right) = \pi_{z_0} \prod_{t=1}^{T} A_{z_{p(t)} z_t}$. Similarly to the independent case of Section 3, the conditional expectation of the complete log-likelihood estimated in the E-step, will be of the form:

$$\begin{aligned} E[L_c\left(\boldsymbol{\varrho}\right)|\mathbf{x}, \boldsymbol{\varrho}^{(n)}] &= \sum_{t=0}^{T} \sum_{m=1}^{M} \gamma_m(z_t) \log \mathrm{Polya}\left(\mathbf{x}_t^c|x_t, \boldsymbol{\alpha}_m\right) \\ &+ \sum_{m=1}^{M} \gamma_m(z_0) \log\left(\pi_{0,m}\right) + \sum_{t=1}^{T} \sum_{m=1}^{M} \sum_{k=1}^{M} \xi_{m,k}(z_t, z_{p(t)}) \log A_{kmt} \end{aligned} \tag{13}$$

Note that in the HMT case of Eq. (13) observations from all scales are processed as whole. Utilizing the Upward-Downward algorithm [4] we can efficiently compute the conditional probability $\gamma_m(z_t) = P\left(z_t = m|\mathbf{x}\right)$ which is required in the posterior mean estimates of Eq. (8), as well as the joint-state probability $\xi_{m,k}(z_t, z_{p(t)}) = P\left(z_t = m, z_{p(t)} = k|\mathbf{x}\right)$. Then in the M-step new estimates for the parameters of the model are obtained. Treating $\boldsymbol{\pi}$ and A is done as in [4], while maximizing Eq. (13) with respect to $\boldsymbol{\alpha}$ is done exactly as in the independent case. Regularizing the solution is similarly important to achieve robustness.

5. RESULTS ON PHOTON-LIMITED IMAGING

Our interest in intensity estimation of inhomogeneous Poisson processes is motivated by the problem of photon-limited image denoising. We first present photon-limited image estimation experiments on standard images degraded by simulated Poisson noise. We compare our proposed methods with the popular Poisson denoising methods of [2] and [7], briefly described in Section 3. Since the two aforementioned methods assume 3-mixture symmetric distributions, we also adopt in these experiments symmetric 3-mixture densities for direct comparison among models; however our technique can equally well be applied to estimate more flexible nonsymmetric densities. The quality of the resulting images from all methods is measured in terms of PSNR and the perceptually-motivated mean structural similarity (MSSIM) index [10]. MSSIM index takes values between 0 and 1 and increases for better quality images. Reported experiments for all methods refer to their shift-invariant versions, which amounts to averaging the image estimates produced by each method over all circular shifts of the noisy image [2]. This averaging results to an image of higher quality with less blocking artifacts.

The results produced by all the methods for 2 testing images (Cameraman/Lena) are presented in Table 1. Three peak intensities are reported (20,40,60), corresponding to different Poisson noise levels. Using the model parameters obtained by our EM methods consistently yields quantitatively better results than the alternative techniques. Specifically, our independent mixture models SEP-IND

Table 1. Photon-limited intensity estimation results in PSNR and MSSIM for various peak intensities and 3-mixture distributions.

	PSNR (dB) / Methods							MSSIM / Methods						
Image/ Peak Int.	noisy	T-N [2]	L-K-A [7]	SEP IND	QUAD IND	SEP HMT	QUAD HMT	noisy	T-N [2]	L-K-A [7]	SEP IND	QUAD IND	SEP HMT	QUAD HMT
Cam./20	16.30	24.25	24.40	25.46	25.89	**26.39**	26.25	0.281	0.603	0.675	0.664	0.705	**0.765**	0.756
Cam./40	19.29	26.12	26.14	27.26	27.60	**27.93**	27.80	0.370	0.675	0.724	0.723	0.755	**0.796**	0.794
Cam./60	21.06	26.97	27.54	28.35	28.69	**29.01**	28.87	0.426	0.674	0.754	0.755	0.787	0.826	**0.827**
Lena/20	16.85	24.80	24.87	25.67	26.07	**26.48**	26.42	0.313	0.688	0.706	0.693	0.726	0.772	**0.777**
Lena/40	19.84	26.21	26.55	27.10	27.61	**28.01**	27.93	0.417	0.713	0.761	0.738	0.772	0.818	**0.820**
Lena/60	21.60	27.37	27.65	28.16	28.70	**29.08**	29.04	0.482	0.752	0.789	0.772	0.804	0.842	**0.845**

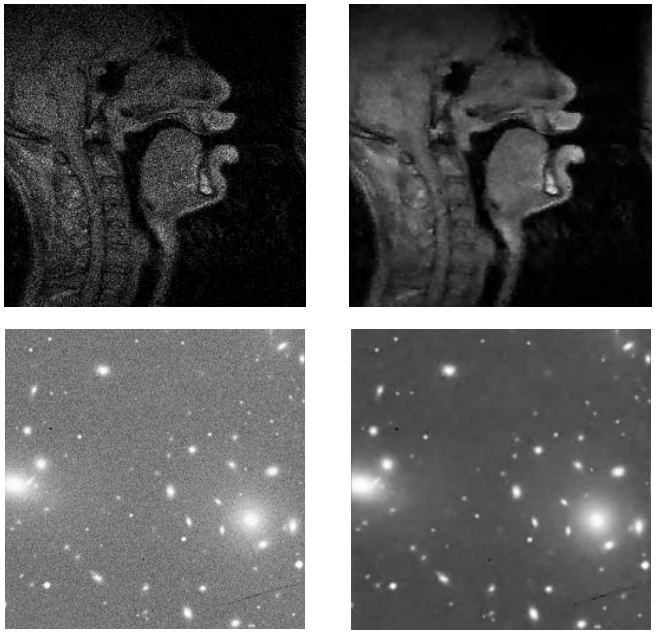
(separable) and QUAD-IND (quad) give roughly 1-1.5 dB improvement over T-N [2] and L-K-A [7]. Modeling scale dependencies with our SEP-HMT and QUAD-HMT HMT-based models gives a further 0.5-1 dB improvement. We can draw similar conclusions from the perceptual quality MSSIM index results. Regarding comparison between corresponding separable and quad variants, both perform equally well and the results do not reveal any clear advantage of either decomposition strategy over the other. The efficacy of our method relative to the alternative techniques can be visually appreciated from the representative Cameraman denoising example shown in Fig. 2. In Fig. 3 we show the results obtained by applying the proposed SEP-HMT model in a medical image with simulated Poisson noise and an astronomical image from [11] with real shot noise. The result in the latter case confirms that the proposed methods are equally efficient in "real-world" noisy conditions.

6. CONCLUSIONS

The main contribution of our work is an efficient EM-based approach for inferring the parameters involved in an important class of photon-limited imaging models. The proposed method can be applied equally well to all separable/quad multiscale partitioning and independent/HMT model variants. Being flexible and fully automatic, our technique can be used in a variety of applications involving shot noise and photon-limited imaging.

(a) (b) (c)

Fig. 2. Results on Cameraman image with peak intensity 20 and simulated Poisson noise. Close-up of (a) Noisy image (PSNR=16.30). (b) T-N result [2] (PSNR=24.25). (c) Our QUAD-HMT result (PSNR=26.25).



Noisy images Our results (SEP-HMT variant)

Fig. 3. Image intensity estimation on medical vocal tract image with simulated noise and astronomical image with real shot noise.

7. REFERENCES

[1] J.L. Starck and F. Murtagh, *Astronomical Image and Data Analysis*, Springer, 2nd edition, 2006.

[2] K. Timmerman and R. Nowak, "Multiscale modeling and estimation of Poisson processes with application to photon-limited imaging," *IEEE Trans. Inf. Theory*, vol. 45, pp. 846–862, 1999.

[3] E. Kolaczyk, "Bayesian multiscale models for Poisson processes," *J. Amer. Stat. Assoc.*, vol. 94, pp. 920–933, 1999.

[4] M. Crouse, R. Nowak, and G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. Signal Process.*, vol. 46, pp. 886–902, 1998.

[5] R. Nowak, "Multiscale hidden Markov models for bayesian image analysis," in *Bayesian Inference in Wavelet Based Models*, B. Vidacovic and P. Muller, Eds. Springer Verlag, 1999.

[6] A. Willsky, "Multiresolution Markov models for signal and image processing," *Proc. IEEE*, vol. 90, pp. 1396–1458, 2002.

[7] H. Lu, Y. Kim, and J. Anderson, "Improved Poisson intensity estimation: Denoising application using Poisson data," *IEEE Trans. Image Process.*, vol. 13, pp. 1128–1135, 2004.

[8] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[9] J. Bernardo and A. Smith, *Bayesian Theory*, Wiley, 2000.

[10] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, pp. 600–612, 2004.

[11] *Skyview Image Server*, http://skyview.gsfc.nasa.gov.