POISSON-HAAR TRANSFORM: A NONLINEAR MULTISCALE REPRESENTATION FOR PHOTON-LIMITED IMAGE DENOISING

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 a novel multiscale image representation belonging to the class of multiscale multiplicative decompositions, which we term Poisson-Haar transform. The proposed representation is well-suited for analyzing images degraded by signal-dependent Poisson noise, allowing efficient estimation of their underlying intensity by means of multiscale Bayesian schemes. The Poisson-Haar decomposition has a direct link to the standard 2-D Haar wavelet transform, thus retaining many of the properties that have made wavelets successful in signal processing and analysis. The practical relevance and effectiveness of the proposed approach is verified through denoising experiments on simulated and real-world photon-limited images.

Index Terms— Photon-limited imaging, Poisson noise, hidden Markov tree, Haar wavelet transform, Bayesian estimation.

1. INTRODUCTION

Photon detection is the basis of the image formation for a great number of imaging systems used in a variety of applications, including medical and astronomical imaging [1]. In such systems, image acquisition is accomplished by counting photon detections at different spatial locations of a sensor, over a specified observation period. For low intensity levels, one of the dominant noise sources responsible for the degradation of the quality of the captured images is the so-called quantum or shot noise. Quantum noise is due to fluctuations on the number of detected photons, an inherent limitation of the discrete nature of the detection process, and degrades such images both qualitatively and quantitatively. The resulting degradation can be proved a major obstacle preventing image analysis and information extraction. Thus, the development of methods and techniques to alleviate the arising difficulties is of fundamental importance.

The basic photon-imaging model assumes a Poisson distribution for the number of detected photons at each pixel location, thus the captured image can be considered as a realization of an inhomogeneous Poisson process. This process is characterized by a 2-D spatially varying rate function which equals the process mean and corresponds to the noise-free intensity image. The variability of the counts about the mean can be interpreted as noise. Since the Poisson process variance equals the rate function/noise-free image, the noise in the acquired image is spatially varying and signal-dependent.

Multiscale processing has emerged as an efficient tool for image processing and analysis, showing significant advantages in the representation of signals and is extensively used in applications such as image denoising, detection, and segmentation. In this paper we deal with a close relative of the wavelet transform, the multiplicative decomposition. The multiscale multiplicative decomposition for 1-D data was first proposed in [2, 3] and is a strong candidate for analyzing Poisson processes, since it provides more natural signal representations than the wavelet transform and makes full use of the Poisson noise properties. In addition, it is closely related to the Haar discrete wavelet transform (DWT) thus retaining many of the advantageous properties of wavelets in signal processing, namely sparsity, approximate decorrelation, and efficient coarse-to-fine processing.

Despite the benefits of using the multiplicative decomposition for analyzing Poisson processes, the existing 2-D representations [2, 4, 5] have poor directional selectivity properties and are not directly related to the Haar wavelet transform like their 1-D counterpart. Our main contribution is a novel multiscale image representation which we term Poisson-Haar decomposition and has a close link to the 2-D Haar DWT. Thus the proposed transformation better models the image edge structure and yields improved performance in image processing. In Section 2 we briefly review the 1-D multiplicative decomposition and indicate its relation to the Haar wavelet transform. Then in Section 3 we present the proposed Poisson-Haar transform and in Section 4 we provide an image intensity estimation method that exploits the properties of our representation and the Poisson noise statistics. Finally, in Section 5 we present image denoising results on simulated and real-world photon-limited images to demonstrate the effectiveness of the proposed methods.

This work was supported by grants ΠENEΔ 2003-EΔ 554 & 865 [co-financed by E.U.-European Social Fund (80%) and the Greek Ministry of Development-GSRT (20%)] and the EC projects ASPI and MUSCLE.

2. MULTIPLICATIVE MULTISCALE DECOMPOSITION

Similarly to the wavelet transform, the multiplicative decomposition scheme of [2, 3] leads to a multiscale signal representation of 1-D signals. In fact, the multiplicative scheme is closely related to the Haar wavelet transform. More specifically, let $\boldsymbol{\lambda} = \left[\lambda(0), \lambda(1), \ldots, \lambda(N-1)\right]$ be a 1-D signal of length $N = 2^J$. Denoting with $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}$ the finest scale representation of $\boldsymbol{\lambda}$, a multiscale analysis with the Haar DWT is obtained through the recursions

$$\begin{aligned} u_j(k) &= \left(u_{j-1}(2k) + u_{j-1}(2k+1)\right)/\sqrt{2} \\ w_j(k) &= \left(u_{j-1}(2k) - u_{j-1}(2k+1)\right)/\sqrt{2}\,, \end{aligned} \tag{1}$$

for $j = 1, \ldots, J$ and $k = 0, \ldots, N/2^j - 1$, where $\mathbf{u}_0 = \boldsymbol{\lambda}_0$. In the above equations j denotes the scale of analysis (J is the coarsest scale) and k the position in the corresponding vector $\mathbf{u}_j$. For the multiplicative decomposition the recursions that yield the multiscale analysis of $\boldsymbol{\lambda}$ are

$$\begin{aligned} \lambda_j(k) &= \lambda_{j-1}(2k) + \lambda_{j-1}(2k+1) \\ \theta_j(k) &= \lambda_{j-1}(2k)/\left(\lambda_{j-1}(2k) + \lambda_{j-1}(2k+1)\right)\,. \end{aligned} \tag{2}$$

Note that while in the case of the Haar transform the coarse-to-fine signal refinement is encoded linearly in the additive detail

Fig. 1: Multiscale image decomposition at two scales of the cameraman image using the Poisson-Haar transformation.

coefficients $\mathbf{w}_j$, in the case of the multiplicative decomposition it is encoded nonlinearly in the *rate-ratio* coefficients $\boldsymbol{\theta}_j$, which can be interpreted as splitting factors [4] of coarse scale intensities. Although the Haar and the multiplicative multiscale transforms describe the detail information differently, there is a direct link between the signal representation $\boldsymbol{\lambda}$ in the multiplicative domain $[\lambda_J(0), \theta_J(0), \ldots, \theta_1(0), \ldots, \theta_1(N/2-1)]$ and the one in the Haar wavelet domain. More specifically, one can show that $\lambda_j(k)$ is a re-scaled version of the Haar scaling coefficient $u_j(k)$, $\lambda_j(k) = 2^{j/2} u_j(k)$, while $\theta_j(k)$ is a divisibly normalized and shifted by 0.5 version of the Haar wavelet detail coefficient $w_j(k)$, $\theta_j(k) = \frac{w_j(k)}{2u_j(k)} + 0.5$. This suggests that key wavelet properties such as sparsity, approximate decorrelation, and efficient coarse-to-fine processing are also inherited by the multiplicative decomposition.

3. POISSON-HAAR MULTISCALE DECOMPOSITION

For handling photon-limited images we need multiplicative decompositions tailored for 2-D signals. However, existing 2-D multiplicative decomposition schemes [2, 4, 5] have poor directional selectivity properties. In particular, the separable scheme of [2] can only represent horizontal and vertical image edges, while the nonseparable quadtree model described in [4, 5] has no explicit mechanism for capturing directional edge information. The proposed Poisson-Haar transform is a novel 2-D multiplicative multiscale decomposition which partly alleviates these shortcomings. Contrary to the existing schemes, the Poisson-Haar representation is closely related to the 2-D Haar DWT, inheriting its ability to also capture edges in the diagonal direction, resulting to improved orientation selectivity properties. Specifically, let $\boldsymbol{\lambda}$ be an intensity image of size $N_1 \times N_2$ and $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}$ its finest scale representation. Denoting for each jth-scale pixel location (k, ℓ) the 2×2 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 Poisson-Haar multiscale analysis of $\boldsymbol{\lambda}$ employs the recursions:

$$\begin{aligned} \lambda_j(k,\ell) &= \sum\nolimits_{(k',\ell') \in C_{k,\ell}} \lambda_{j-1}(k',\ell') \\ \theta_j^o(k,\ell) &= \lambda_j^o(k,\ell)/\lambda_j(k,\ell), \quad o \in \{h,v,d\} \end{aligned} \tag{3}$$

for $j = 1, \ldots J$, $k = 0, \ldots, N_1/2^j - 1$, $\ell = 0, \ldots, N_2/2^j - 1$ and $J = \min\{\log_2(N_1), \log_2(N_2)\}$. In Eq. (3), o is one of $\{h, v, d\}$ denoting the horizontal, vertical, or diagonal decomposition subband; the corresponding intermediate sums $\lambda_j^o(k,\ell)$ are defined as

$$\begin{aligned} \lambda_j^h(k,\ell) &= \lambda_{j-1}(2k,2\ell) + \lambda_{j-1}(2k,2\ell+1) \\ \lambda_j^v(k,\ell) &= \lambda_{j-1}(2k,2\ell) + \lambda_{j-1}(2k+1,2\ell) \\ \lambda_j^d(k,\ell) &= \lambda_{j-1}(2k,2\ell) + \lambda_{j-1}(2k+1,2\ell+1)\,. \end{aligned} \tag{4}$$

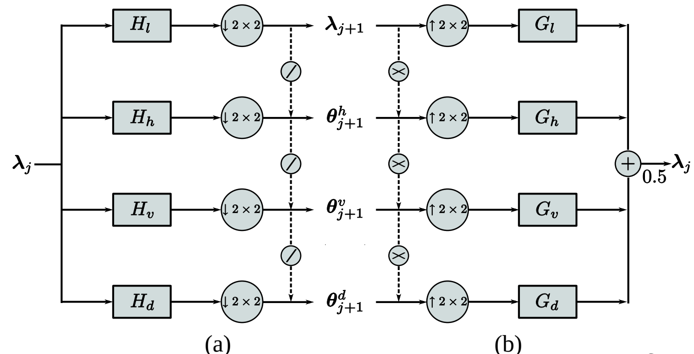


Fig. 2: Poisson-Haar image multiscale decomposition through filterbank processing. (a) Analysis filterbank, (b) Synthesis filterbank. The ratio coefficients $\boldsymbol{\theta}^h, \boldsymbol{\theta}^v, \boldsymbol{\theta}^d$ are sensitive to image edges at different orientations (horizontal, vertical, or diagonal).

Similarly to the 1-D case, the Poisson-Haar rate-ratio coefficients are directly related to the 2-D Haar wavelet detail coefficients w of the same subband:

$$\theta_j^o(k,\ell) = \frac{w_j^o(k,\ell)}{2u_j(k,\ell)} + \frac{1}{2}\,. \tag{5}$$

In addition, $\lambda_j(k,\ell)$ is a re-scaled version of the 2-D Haar scaling coefficient $\lambda_j(k,\ell) = 2^j u_j(k,\ell)$. Based on (5), the ratio coefficients $\boldsymbol{\theta}$ can be viewed as divisible normalized and shifted by 0.5 Haar wavelet detail coefficients. Thus, the Poisson-Haar decomposition explicitly represents horizontal, vertical, and diagonal edges in a symmetric fashion, yielding increased orientation selectivity compared to the previous multiplicative decomposition schemes. An example of the Poisson-Haar decomposition is shown in Fig. 1.

Similarly to 2-D Haar wavelets, the Poisson-Haar decomposition can also be viewed in terms of filterbank processing. In particular, to compute one stage of the Poisson-Haar transform of an image $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}$, we first convolve it with the filters $H_l = \left(\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right)$, $H_h = \left(\begin{smallmatrix}0&0\\1&1\end{smallmatrix}\right)$, $H_v = \left(\begin{smallmatrix}0&1\\0&1\end{smallmatrix}\right)$, and $H_d = \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$ and then decimate by two in each direction. Then the last three subband images of size $N_1/2 \times N_2/2$ are pointwise divisibly normalized by the first one as is schematically illustrated in Fig. 2. The filtering and downsampling process can be continued up to J times resulting in a multiscale representation of $\boldsymbol{\lambda}$. Since the transformation is one-to-one we can return back to the original domain by following a reverse procedure as depicted in Fig. 2. In this case the filters are $G_l = \left(\begin{smallmatrix}-1&1\\1&1\end{smallmatrix}\right)$, $G_h = \left(\begin{smallmatrix}1&1\\-1&-1\end{smallmatrix}\right)$, $G_v = \left(\begin{smallmatrix}1&-1\\1&-1\end{smallmatrix}\right)$, and $G_d = \left(\begin{smallmatrix}1&-1\\-1&1\end{smallmatrix}\right)$.

4. BAYESIAN IMAGE INTENSITY ESTIMATION

In this section using the proposed Poisson-Haar decomposition and adopting the basic photon-limited imaging model, we describe a method to estimate the intensity image $\boldsymbol{\lambda}$ given a photon-limited version $\mathbf{x}$. Under this image formation model, $\mathbf{x}$ is considered as a 2-D array consisting of the observation samples (count data) of $N_1 \times N_2$ random variables $X(k,\ell)$ which are conditionally independent upon $\boldsymbol{\lambda}$, and each one follows a Poisson distribution with rate parameter $\lambda(k,\ell)$, denoted by $X(k,\ell)|\lambda(k,\ell) \sim \text{Pois}\,(\lambda(k,\ell))$.

In the wavelet literature, image denoising can be achieved by estimating the wavelet coefficients of the desired image from the noisy observations and then returning to the original domain through the inverse wavelet transform. Our method for estimating the underlying intensity image $\boldsymbol{\lambda}$ consists of the same two steps. First, we estimate the detail coefficients $\boldsymbol{\theta}$ of Eq. (3) and then we obtain our final result using the inverse Poisson-Haar transform. To estimate $\boldsymbol{\theta}$ we exploit some fundamental properties of Poisson processes [6]: (1) Given the underlying intensities, the counts over nonoverlapping intervals are independent. (2) The sum of independent Poisson random variables remains Poisson. Thus, the random variable $X_j(k,\ell)$, obtained as the sum of the Poisson random variables $X_{j-1}(k',\ell')$ lying in the 2×2 neighborhood $C_{k,l}$, will remain Poisson distributed with intensity $\lambda_j(k,\ell)$. (3) For two independent Poisson random variables, $X|\lambda_x \sim \text{Pois}(\lambda_x)$ and $Y|\lambda_y \sim \text{Pois}(\lambda_y)$, the conditional distribution of X given X+Y is binomial, namely $p(x|x+y) = \text{Bin}\left(x|x+y, \frac{\lambda_x}{\lambda_x+\lambda_y}\right)$.

Let us denote with $\mathbf{x}_0$ the finer representation of the photon-limited image $\mathbf{x}$. Next we decompose $\mathbf{x}_0$ in J scales using the recursions in Eqs. (3) and (4), retaining, respectively, the coarse scale values $x_j(k,\ell)$ and intermediate sums $x_j^o(k,\ell)$. Since each coefficient of $\mathbf{x}_j$ and $\mathbf{x}_j^o$ is derived as the sum of independent Poisson random variables, it will remain Poisson distributed. In addition, based on property (3) above, the conditional distribution $p\left(x_j^o(k,\ell)|x_j(k,\ell)\right)$, will be binomial

$$p\left(x_j^o(k,\ell)|x_j(k,\ell)\right) = \text{Bin}\left(x_j^o(k,\ell)|x_j(k,\ell),\theta_j^o(k,\ell)\right)\ , \quad (6)$$

where $\theta_j^o(k,\ell)$ corresponds to the detail coefficients of the image of interest $\boldsymbol{\lambda}$ defined in Eq. (3).

Based on Eq. (6), we can work under a Bayesian framework and derive an optimal estimator for the detail coefficients. Under this framework, the detail coefficients $\boldsymbol{\theta}$ are considered as realizations of random variables and in order to obtain a solution an *a priori* knowledge about their distribution is required. In our previous work [5] and in [2] it has been shown that a suitable prior model for the rate-ratio coefficients of natural images is the Beta-mixture distribution. Due to the approximate decorrelation of the wavelet coefficients [7] and since our $\boldsymbol{\theta}$ ratio/detail coefficients are just normalized and shifted by 0.5 wavelet coefficients, it is also reasonable to assume them independent across subbands. Based on these assumptions, we adopt the following prior distribution

$$p\left(\theta_j^o(k,\ell)\right) = \sum_{m=1}^{M} \pi_{j,m}\,\text{Beta}\left(\theta_j^o(k,\ell)|\alpha_{j,m}^o,\beta_{j,m}^o\right)\ , \quad (7)$$

where $\pi_{j,m}$ is the mixture weight for the mth mixture component in jth scale of analysis, $\alpha_{j,m}^o$ and $\beta_{j,m}^o$ are the parameters of this beta mixture component for the specific subband and M is the total number of mixtures utilized at each scale. Using in Eq. (7) a mixture of beta densities instead of a single component better fits the model to natural image statistics.

Using the identity (can be easily verified by direct substitution)

$$\begin{aligned}\text{Bin}\left(x|n,\theta\right)\text{Beta}\left(\theta|\alpha,\beta\right) = \\ \text{Polya}\left(x|n,\alpha,\beta\right)\text{Beta}\left(\theta|x+\alpha,n-x+\beta\right)\ , \quad (8)\end{aligned}$$

where the bivariate Polya distribution [6] (also called Beta-binomial) is defined as $\text{Polya}(x|n,\alpha,\beta) = \binom{n}{x}\frac{B(x+\alpha,\,n-x+\beta)}{B(\alpha,\beta)}$ and $B(\cdot)$ is the Beta function, we can write the posterior density of $\theta_j^o(k,\ell)$ as:

$$\begin{aligned}p\left(\theta_j^o(k,\ell)|x_j(k,\ell),x_j^o(k,\ell)\right) = \sum_{m=1}^{M}\gamma_m(z_j^o(k,\ell)) \\ \times\,\text{Beta}\left(\theta_j^o(k,\ell)|x_j^o(k,\ell)+\alpha_{j,m}^o, x_j^{o,c}(k,\ell)+\beta_{j,m}^o\right)\ , \quad (9)\end{aligned}$$

where $x_j^{o,c}(k,\ell) = x_j(k,\ell) - x_j^o(k,\ell)$ and $z_j^o(k,\ell)$ indicates which mixture component is active and the corresponding posterior mixture assignment weights equal

$$\gamma_m(z_j^o(k,\ell)) = \frac{\pi_{j,m}\,\text{Polya}\left(x_j^o(k,\ell)|x_j(k,\ell),\alpha_{j,m}^o,\beta_{j,m}^o\right)}{\sum\limits_{n=1}^{M}\pi_{j,n}\,\text{Polya}\left(x_j^o(k,\ell)|x_j(k,\ell),\alpha_{j,n}^o,\beta_{j,n}^o\right)}\ .$$

Having at hand the posterior distribution (9) we can readily obtain a Bayesian estimator for the ratio coefficients. If we select the minimum mean squared error (MMSE) as the criterion to minimize, we end up with the posterior mean estimator

$$\hat{\theta}_j^o(k,\ell) = \sum_{m=1}^{M}\gamma_m(z_j^o(k,\ell))\left(\frac{x_j^o(k,\ell)+\alpha_{j,m}^o}{x_j(k,\ell)+\alpha_{j,m}^o+\beta_{j,m}^o}\right)\ , \quad (10)$$

which can be viewed as a thresholding technique of the noisy ratio coefficients, since the ratio $x_j^o(k,\ell)/x_j(k,\ell)$ corresponds exactly to the detail coefficient that we would have obtained if we had applied the Poisson-Haar decomposition (3) at the noisy image $\mathbf{x}$. Note that the estimation requires that we know the values of the parameters $\boldsymbol{\mu} = \{\boldsymbol{\pi}_j, \boldsymbol{\alpha}_j, \boldsymbol{\beta}_j\}$. These values can be inferred directly by the noisy observations using the EM-based technique we have proposed in [5]. Moreover, since the assumption of independence across scales is a weak one, we can also employ a hidden Markov tree (HMT) model [7] to capture the inter-scale coefficient dependencies in the vicinity of image edges for each subband image [5]. Finally, for the scaling coefficients at the coarsest scale of analysis a good choice is to estimate them by $\hat{\boldsymbol{\lambda}}_J = \mathbf{x}_J$; The signal-to-noise ratio (SNR) of a Poisson process increases linearly with the underlying intensity of the image, thus at the coarsest scale where the image intensity is accumulated into few scaling coefficients $\lambda_J(k,\ell)$, the Poisson noise will be significantly reduced.

5. EXPERIMENTS AND APPLICATIONS

In order to validate the effectiveness of the proposed Poisson-Haar image decomposition coupled with the intensity estimation method, we provide experimental results for the task of image denoising. We compare our proposed method with our previous work presented in [5], which employs the previously existing separable (SEP) and quadtree (QUAD) multiplicative decomposition schemes. In addition, for the comparisons we also use methods that preprocess the Poisson data by variance stabilizing transforms (VST), in order to transform the noise statistics so that standard denoising methods designed for handling homogeneous Gaussian additive noise are appropriate. For this class of methods we use as VSTs the Anscombe [8] and Haar-Fisz [9] transforms and as denoising method the popular wavelet-domain SureShrink [10] employing Daubechies wavelets of 5 vanishing moments. For more extensive comparisons one can also refer to [11] where an extension of this work is presented. The quality of the resulting images is measured in terms of peak SNR (PSNR). In all cases we used 5 decomposition scales. To avoid blocking artifacts, we used shift-invariant versions of each method, obtained by averaging the image estimates of 32 random circular shifts of the noisy image.

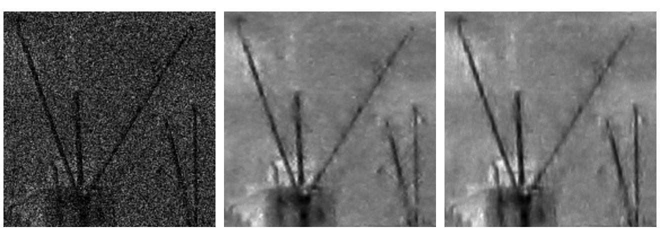
Table 1: Photon-limited intensity estimation results in terms of PSNR. The prior density models employ 3-mixture distributions.

	PSNR (dB) / Methods					
Image/ Peak Int.	noisy	Ansc. SURE	HF SURE	SEP HMT	QUAD HMT	PH HMT
Lena/5	9.95	24.44	24.69	25.59	25.68	**25.78**
Lena/10	12.96	26.15	26.29	27.03	27.09	**27.21**
Lena/15	14.72	27.02	27.12	27.87	27.85	**28.03**
Lena/20	15.97	27.65	27.73	28.45	28.41	**28.66**
Boat/5	9.93	23.19	23.51	24.08	24.04	**24.31**
Boat/10	12.94	24.66	24.73	25.32	25.21	**25.57**
Boat/15	14.70	25.45	25.45	26.04	25.92	**26.37**
Boat/20	15.95	26.04	26.05	26.59	26.48	**26.96**

The results produced by all methods, based on 10 independent trials in each case, for two test images (Lena/Boat) are presented in Table 1. Four peak intensities are reported (5,10,15,20), corresponding to different Poisson noise levels. Using the proposed Poisson-Haar (PH) multiscale image representation, coupled with the HMT model of [5], yields quantitatively better results than the other studied techniques. Specifically, our PH-HMT model gives a rough improvement of 1-1.25 dB in PSNR, compared to the method of fist stabilizing the data with the Anscombe or Haar-Fisz (HF) transforms, and then applying the SureShrink thresholding. Regarding comparison between our HMT variants, namely the ones using the separable and quadtree decomposition schemes and our newly proposed Poisson-Haar image representation, we note that the latter gives an improvement of about 0.25 dB, which is exclusively attributed to the image decomposition model; Similar results we have also obtained, but not report here due to space limitation, for our variants when independence across scales is assumed. The efficacy of our method relative to the alternative techniques can be visually appreciated from the representative Lena denoising example shown in Fig. 3. In Fig 4 we present a close up of the denoised boat image by the SEP and PH models and observe that the latter better preserves the image edges. Finally, in Fig. 5 we show the results obtained by applying the proposed PH-HMT model in an astronomical image with simulated Poisson noise and a nuclear medical image with real shot noise. The result in the latter case confirms that the proposed method is also efficient in “real-world” noisy conditions.

(a) (b) (c)

Fig. 3: Results on Lena image with peak intensity 10 and simulated Poisson noise. Close-up of: (a) Noisy image (PSNR=12.96). (b) HF-SureShrink result (PSNR=26.29). (c) Our PH-HMT result (PSNR=27.21).

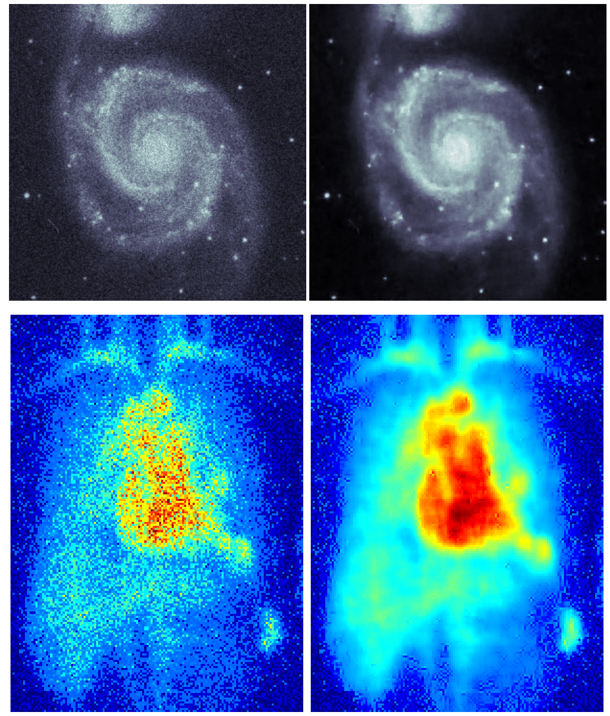


(a) (b) (c)

Fig. 4: Results on boat image with peak intensity 5 and simulated Poisson noise. Close-up of: (a) Noisy image (PSNR=9.93). (b) SEP-HMT result (PSNR=24.08). (c) PH-HMT result (PSNR=24.31).

6. CONCLUSIONS

The main contribution of our work is a multiscale image representation which is closely related to the 2-D Haar DWT, thus inheriting most of the benefits of the wavelet transform. Our multiscale Poisson-Haar image representation is shown to be well-suited for analyzing photon-limited images degraded by Poisson noise, better modeling the edge structure of images compared to existing 2-D multiscale multiplicative image decompositions.



Noisy images Our results (PH-HMT)

Fig. 5: Image intensity estimation on astronomical image with simulated noise and nuclear 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] 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.

[5] S. Lefkimmiatis, G. Papandreou, and P. Maragos, “Photon-limited image denoising by inference on multiscale models,” in *Int. Conf. Image Process. (ICIP)*, 2008.

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

[7] 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.

[8] F. J. Anscombe, “The transformation of Poisson, binomial and negative-binomial data,” *Biometrika*, vol. 35, no. 3, pp. 246–254, 1948.

[9] P. Fryźlewicz and G. P. Nason, “A Haar-Fisz algorithm for Poisson intensity estimation,” *J. Comp. Graph. Stat.*, vol. 13, pp. 621–638, 2004.

[10] D.L. Donoho and I.M. Johnstone, “Adapting to unknown smoothness via wavelet shrinkage,” *J. Amer. Stat. Assoc.*, vol. 90, pp. 1200–1224, 1995.

[11] S. Lefkimmiatis, P. Maragos, and G. Papandreou, “Bayesian inference on multiscale models for Poisson intensity estimation: Applications to photon-limited image denoising,” *IEEE Trans. Image Process.*, vol. 18, no. 8, pp. 1724–1741, 2009.