Non local means under Poisson noise with automatic setting based on risk minimization

Moyennes non locales pour le bruit de Poisson avec réglage automatique fondé sur la minimisation du risque

Charles-Alban Deledalle
Florence Tupin
Loïc Denis

février 2010

Département Traitement du Signal et des Images
Groupe TII : Traitement et Interprétation des Images
Moyennes Non Locales pour le Bruit de Poisson avec Réglage Automatique fondé sur la Minimisation du Risque

Charles-Alban Deledalle, Florence Tupin et Loïc Denis

Résumé

Une extension des moyennes non locales (MNL) est proposée pour les images dégradées par du bruit de Poisson. La méthode proposée est guidée par l’image bruitée et une image pré-filtrée et est, de plus, adaptée au modèle statistique du bruit de Poisson. L’influence relative des deux images dépend de deux paramètres. Nous proposons un ajustement automatique de ces paramètres fondé sur la minimisation de l’erreur quadratique moyenne (EQM) estimée. Cette sélection utilise un estimateur de l’EQM pour le filtre MNL avec bruit de Poisson et une méthode de Newton pour trouver les paramètres optimaux en peu d’itérations.

Mots-clés

Moyennes non locales, bruit de Poisson, erreur quadratique moyenne, méthode de Newton

C. Deledalle et F. Tupin sont à l’Institut Telecom, Telecom ParisTech, CNRS LTCI, Paris, France, e-mail: charles-alban.deledalle@telecom-paristech.fr et florence.tupin@telecom-paristech.fr.

L. Denis est à l’École Supérieure de Chimie Physique Électronique de Lyon et au Laboratoire Hubert Curien, CNRS UMR 5516, St-Étienne, France, e-mail: loic.denis@cpe.fr.

February 15, 2010
Non Local Means under Poisson Noise
with Automatic Setting based on Risk Minimization

Charles-Alban Deledalle, Florence Tupin et Loïc Denis

Abstract

An extension of the non local (NL) means is proposed for images damaged by Poisson noise. The proposed method is guided by the noisy image and a pre-filtered image and is adapted to the statistics of Poisson noise. The influence of both images can be tuned using two filtering parameters. We propose an automatic setting to select these parameters based on estimated mean square error (MSE) minimization. This selection uses an estimator of the MSE for NL means with Poisson noise and a Newton’s method to find the optimal parameters in few iterations.

Index Terms

Non local means, Poisson noise, mean square error, Newton’s method

C. Deledalle et F. Tupin sont à l’Institut Telecom, Telecom ParisTech, CNRS LTCI, Paris, France, e-mail: charles-alban.deledalle@telecom-paristech.fr et florence.tupin@telecom-paristech.fr.
L. Denis est à l’École Supérieure de Chimie Physique Électronique de Lyon et au Laboratoire Hubert Curien, CNRS UMR 5516, St-Étienne, France, e-mail: loic.denis@cpe.fr.

February 15, 2010
I. Introduction

Poisson noise appears in low-light conditions when the number of collected photons is small, such as in optical microscopy or astronomy. Poisson noise is signal-dependent and, then, requires to adapt the usual denoising approaches.

NL means have been proposed by Buades et al. in [1] to denoise images damaged by additive white Gaussian noise. While local filters lead to biases and resolution loss, NL techniques are known to efficiently reduce noise and preserve structures. Instead of combining neighbor pixels, the NL means average similar pixels. NL means assume there are enough redundant pixels (pixel having identical noise-free value) in the image to reduce the noise significantly. Let $k_s$ be the observed noisy value at site $s$ and $\lambda_s$ its underlying noise-free value, NL means define the estimate $\hat{\lambda}_s$ as a weighted average:

$$\hat{\lambda}_s = \frac{\sum_t w_{s,t} k_t}{\sum_t w_{s,t}}$$

(1)

where $t$ is a pixel index and $w_{s,t}$ is a data-driven weight depending on the similarity between pixels with index $s$ and $t$. In practice, the pixels $t$ are located in a search window centered on $s$. For robustness reason, pixel similarity is evaluated by comparing surrounding patches around $s$ and $t$. Patch-similarity is classically defined by the Euclidean distance, leading to the following weight expression:

$$w_{s,t} = \exp \left(-\frac{\sum_b (k_{s+b} - k_{t+b})^2}{\alpha}\right)$$

(2)

where $s+b$ and $t+b$ denote respectively the $b$-th pixels in the patches $B_s$ and $B_t$ centered on $s$ and $t$, and $\alpha$ is a filtering parameter.

In case of low signal-to-noise ratio images, it has been shown that the performances of the NL means can be improved by refining the weights using a pre-estimate $\hat{\theta}$ of the noise-free image [2]–[5]. When the weights are based only on the noisy image, they present a high variance since they are very sensitive to the noisy content. However, if the weights are also based on a pre-estimate $\hat{\theta}$ of the noise-free image, the weights are more suitable to increase the denoising performances are increased. The general expression of refined NL means is:

$$\hat{\lambda}_s = \frac{\sum_t w_{s,t} k_t}{\sum_t w_{s,t}}$$

(3)

with

$$w_{s,t} = \exp \left(-\frac{F_{s,t}}{\alpha} - \frac{G_{s,t}}{\beta}\right),$$

$$F_{s,t} = \sum_b f(k_{s+b}, k_{t+b})$$

and

$$G_{s,t} = \sum_b g(\hat{\theta}_{s+b}, \hat{\theta}_{t+b})$$

February 15, 2010
where $\alpha$ and $\beta$ are filtering parameters and $f$ and $g$ are two similarity criteria suitable respectively to compare noisy data and pre-estimated data. Note that the refined NL means include iterative NL means [2]–[5] when $\hat{\theta}$ is the result of the previous iteration. A typical choice for $f$ and $g$ is the squared difference: $f(x, y) = g(x, y) = (x - y)^2$. The choice of the filtering parameters for $\alpha$ and/or $\beta$ is a critical task already explored in [1], [2], [4], [5] and [6]. The purpose is to find automatically, from the image $k$ and the knowledge on the noise statistics, the parameters which provide the best denoising performances whatever the underlying image $\lambda$. According to our knowledge, there is no known method to jointly set $\alpha$ and $\beta$ in case of Poisson noise. Since Poisson noise is signal-dependent, the values for $\alpha$ and $\beta$ must be adapted to each observed image $k$.

II. PATCH-SIMILARITIES UNDER POISSON NOISE

Let $k$ be an image damaged by noise following a Poisson distribution with parameters described by the underlying noise-free image $\lambda$:

$$p(k_s|\lambda_s) = \frac{\lambda_s^{k_s}e^{-\lambda_s}}{k_s!}. \quad (4)$$

The probabilistic approach of [5] can be applied to extend the refined NL means in (3) to Poisson noise degradation model. The squared difference classically used for $f$ is replaced by:

$$f_L(k_1, k_2) = -\log \int p(k_1|\lambda_1 = \lambda)p(k_2|\lambda_2 = \lambda)d\lambda = \log \left(\frac{k_1!k_2!}{(k_1 + k_2)!}\right) + (k_1 + k_2 + 1) \log 2. \quad (5)$$

This similarity evaluates the joint likelihood for all possible values of the unknown parameter $\lambda$. Note that this criterion is very similar by its theoretical definition and by its behavior to the criterion defined in [7] also used in the case of Poisson noise. The squared difference generally used for $g$ is replaced by the symmetric Kullback-Leibler divergence:

$$g_{KL}(\hat{\theta}_1, \hat{\theta}_2) = D_{KL}(\hat{\theta}_1\|\hat{\theta}_2) = (\hat{\theta}_1 - \hat{\theta}_2) \log \frac{\hat{\theta}_1}{\hat{\theta}_2}. \quad (6)$$

This criterion is a good candidate to define similarities between estimated values since it can be considered as a statistical test of the hypothesis $\lambda_1 = \lambda_2$ [8].

The setting of the parameters $\alpha$ and $\beta$ in the case of Poisson noise is maybe a more critical problem than in other denoising tasks. In [1], [2] and [5], the authors propose to define the parameters according to the variance or the quantiles of the similarity criteria (subject to identical and independent distributed random variables). Unfortunately, in case of Poisson noise, these quantities depend on the unknown image $\lambda$ since the noise is signal-dependent. Van De Ville et al. propose a risk minimization approach.
for Gaussian noise [6]. Their method selects the parameters minimizing the risk (without any specific assumption on the underlying image $\lambda$). This kind of approach seems relevant in the case of Poisson noise. We will extend this idea in the next section to Poisson noise.

III. AUTOMATIC SETTING OF PARAMETERS BASED ON RISK MINIMIZATION

The parameters of the denoising technique can be selected as those that minimize the expected MSE:

$$
\mathbb{E} \left[ \frac{1}{N} \| \lambda - \hat{\lambda} \|^2 \right] = \frac{1}{N} \sum_s \left( \lambda_s^2 + \mathbb{E} \left[ \hat{\lambda}_s^2 \right] - 2\mathbb{E} \left[ \lambda_s \hat{\lambda}_s \right] \right)
$$

(7)

with $N$ the image size and $\mathbb{E}[.]$ the expectation operator. The MSE requires the knowledge of the noise-free image $\lambda$ but can still be estimated from the noisy image $k$ alone. Since the first term $\lambda_s^2$ in (7) is independent of $\hat{\lambda}_s$, it can be omitted when minimizing the MSE with respect to the denoising parameters. The Stein’s unbiased risk estimator (SURE) is an estimator of the MSE under Gaussian noise [9]. It is based on an estimator of $\mathbb{E}[\lambda_s \hat{\lambda}_s]$ which does not require $\lambda$. SURE has already been used successfully on images damaged by additive white Gaussian noise for wavelet filtering [10] and NL means filtering [6]. The main result in [6] is that SURE for NL means can be obtained in closed form. For Poisson noise, we use the result of Chen [11] to follow the same approach:

$$
\mathbb{E} \left[ \lambda_s h(k)_s \right] = \mathbb{E} \left[ k_s h(k)_s \right]
$$

(8)

with $k_t = \begin{cases} k_t & \text{if } t \neq s \\ k_t - 1 & \text{otherwise} \end{cases}$

where $h(.,)_s$ denotes the estimated value on site $s$ obtained by the application of the estimator $h$ on the given noisy image. Considering the estimator $h$ of the refined NL means defined in (3), we introduce the Chen and Stein’s unbiased risk estimator (CSURE) as:

$$
CSURE = \frac{1}{N} \sum_s \left( \lambda_s^2 + \hat{\lambda}_s^2 - 2k_s \bar{\lambda}_s \right)
$$

(9)

The value $\bar{\lambda}_s$ refers to the denoised value obtained by the application of the NL means on the noisy image $\overline{k}$, i.e:

$$
\bar{\lambda}_s = \frac{\sum_t w_{s,t} \overline{k}_t}{\sum_t w_{s,t}}
$$

(10)

with $w_{s,t} = \exp \left( -\frac{\overline{F}_{s,t}}{\alpha} - \frac{\overline{G}_{s,t}}{\beta} \right)$

and $\overline{F}_{s,t} = \sum_b f(\overline{K}_{s+b}, \overline{K}_{t+b})$.
According to (7), (8) and (9), it is straightforward to show that CSURE is unbiased: 

\[ E[CSURE(\lambda, \hat{\lambda})] = E[MSE(\lambda, \hat{\lambda})] \]

However, note that (10) holds by assuming that \( G_{s,t} \) (i.e., the pre-estimate \( \hat{\theta} \)) does not depend on the noise component of \( k \). To satisfy this assumption, the noise variance in \( \hat{\theta} \) has to be reduced significantly. This assumption simplifies drastically the expression of \( \lambda \).

In terms of time complexity, we note as in [6] that the computation time is unchanged since the computation of CSURE can be incorporated within the core of the NL means. Moreover, the scan of the patches of \( k \) can be avoided thanks to the following relation:

\[
F_{s,t} = F_{s,t} + \begin{cases} 
  f(\bar{k}_s, \bar{k}_s) - f(k_s, k_s), & \text{if } s = t, \\
  f(\bar{k}_s, k_t) - f(k_s, k_t) \\
  + f(k_{2s-t}, \bar{k}_s) - f(k_{2s-t}, k_s), & \text{if } s \in B_t, \\
  f(\bar{k}_s, k_t) - f(k_s, k_t), & \text{otherwise.}
\end{cases}
\]

Selecting parameters that minimize CSURE gives parameters close to that minimizing the MSE. In the case of the classical NL means, the authors of [6] compute the optimal parameters by exhaustive search. Optimization techniques can be applied to reach the optimal parameters in few iterations. In [12], a gradient descent is performed to optimize SURE for wavelet shrinkage. We follow such a strategy here to optimize CSURE for the refined NL means by using the Newton’s method on the joint filtering parameters \( \alpha \) and \( \beta \). The Newton’s method iteratively refine \( \alpha \) and \( \beta \) with the update:

\[
\begin{pmatrix}
  \alpha^{n+1} \\
  \beta^{n+1}
\end{pmatrix} = \begin{pmatrix}
  \alpha^n \\
  \beta^n
\end{pmatrix} - H^{-1} \nabla
\]

with

\[
H^{-1} \nabla = \begin{pmatrix}
  \frac{\partial^2 CSURE}{\partial \alpha^2} & \frac{\partial^2 CSURE}{\partial \alpha \beta} \\
  \frac{\partial^2 CSURE}{\partial \beta \alpha} & \frac{\partial^2 CSURE}{\partial \beta^2}
\end{pmatrix}^{-1} \begin{pmatrix}
  \frac{\partial CSURE}{\partial \alpha} \\
  \frac{\partial CSURE}{\partial \beta}
\end{pmatrix}.
\]

To perform the optimization procedure in (11), the two first order differentials are required. Their expressions are given by substituting \( x \) and \( y \) by \( \alpha \) or \( \beta \) in the following equations:

\[
\frac{\partial CSURE}{\partial x} = \frac{2}{N} \sum_s \hat{\lambda}_s \frac{\partial \hat{\lambda}_s}{\partial x} - \frac{2}{N} \sum_s k_s \frac{\partial \hat{\lambda}_s}{\partial x},
\]

\[
\frac{\partial^2 CSURE}{\partial x \partial y} = \frac{2}{N} \sum_s \hat{\lambda}_s \frac{\partial^2 \hat{\lambda}_s}{\partial x \partial y} + \frac{2}{N} \sum_s \left( \frac{\partial \hat{\lambda}_s}{\partial x} \right) \left( \frac{\partial \hat{\lambda}_s}{\partial y} \right) - \frac{2}{N} \sum_s k_s \frac{\partial^2 \hat{\lambda}_s}{\partial x \partial y},
\]

February 15, 2010
Figure 1. The risk (MSE and CSURE) and their two first order variations (from top to bottom) with respect to the parameters $\alpha$ (left) and $\beta$ (right).

$$
\frac{\partial \hat{\lambda}_s}{\partial x} = \frac{\sum X_{s,t}w_{s,t}(k_t - \hat{\lambda}_s)}{x^2 \sum w_{s,t}},
$$

$$
\frac{\partial^2 \hat{\lambda}_s}{\partial x^2} = \frac{\sum X_{s,t}^2w_{s,t}(k_t - \hat{\lambda}_s)}{x^4 \sum w_{s,t}} - 2 \frac{\partial \hat{\lambda}_s}{\partial x} \frac{\sum (X_{s,t} + x)w_{s,t}}{x^2 \sum w_{s,t}},
$$

$$
\frac{\partial^2 \hat{\lambda}_s}{\partial x \partial y} = \frac{\sum X_{s,t}Y_{s,t}w_{s,t}(k_t - \hat{\lambda}_s)}{x^2 y^2 \sum w_{s,t}} - \frac{\partial \hat{\lambda}_s}{\partial x} \frac{\sum Y_{s,t}w_{s,t}}{y^2 \sum w_{s,t}} - \frac{\partial \hat{\lambda}_s}{\partial y} \frac{\sum X_{s,t}w_{s,t}}{x^2 \sum w_{s,t}}.
$$
where \( X = F \) (resp. \( Y = F \)) when \( x = \alpha \) (resp. \( y = \alpha \)) and \( X = G \) (resp. \( Y = G \)) when \( x = \beta \) (resp. \( y = \beta \)). The differentials for \( X \) are the same with respect to \( k, w \) and \( T \).

The Newton’s method finds in few iterations the best trade-off between the information brought by the noisy image and the pre-estimated image to define the weights. For instance, \( \beta \) will get a high value when the pre-estimated image has a poor quality, resulting to weights determined only from the noisy image. Reciprocally, \( \alpha \) will get a high value when the pre-estimated image has a high quality: the weights will be determined only from the well pre-estimated image.

### IV. Experiments and Results

The proposed extension of the NL means (Poisson NL means) is applied with a search window of size \( 21 \times 21 \) and patches of size \( 7 \times 7 \). The Newton’s method is performed until CSURE does not change
between two successive iterations. The pre-estimated image is obtained by a moving average (MA) filter with a $9 \times 9$ disk kernel. Using the optimization of [13], the computational time is of about 20s per iteration on a $256 \times 256$ image and C implementation on an Intel Pentium D 3.20GHz.

Figure 1 shows the risk and its two first order differentials with respect to $\alpha$ and $\beta$. These curves have been computed by applying the proposed method on a $150 \times 150$ noisy image for different values of the parameters. The MSE and its differentials have been computed from the noise-free image and finite differences. CSURE and its differentials have been evaluated using the expressions given in Section III.

Table I gives the signal-to-noise ratio (SNR) values obtained by different denoising methods on four $256 \times 256$ reference images damaged by synthetic Poisson noise with different degradation levels. The MA filter is applied with a $9 \times 9$ disk kernel. We performed a Poisson based total-variation minimization (Poisson TV) as proposed in [14]. Three versions of NL means are applied. NL means denotes the classical one i.e $f(x, y) = (x - y)^2$ and $\beta = \infty$. We call the refined NL means when $f(x, y) = g(x, y) = (x - y)^2$. Poisson NL means denotes our proposed method with $f$ and $g$ defined as in Section II. For all NL means versions, the optimal parameters are obtained by CSURE minimization using the Newton’s method. The table gives the optimal parameters $\alpha_{opt}$ and $\beta_{opt}$ and the number of iterations for the Poisson NL means. The refined and the Poisson NL means use both the pre-estimated image obtained by the MA filter. Poisson NL means provides the best performances with 6 to 13 iterations. The parameters behave as predicted with respect to the relative qualities of the noisy image and the pre-estimated image.

Figure 2 presents visual results obtained by Poisson TV and Poisson NL means on three images. The two first images are degraded by synthetic Poisson noise and the third one is an image\textsuperscript{1} of a mitochondrion\textsuperscript{2} sensed in low-light conditions by confocal fluorescence microscopy [15]. On both examples, Poisson NL means seem to better preserve the resolution while reducing the noise.

V. CONCLUSION

The methodology of [5] has been used successfully to extend the NL means to images damaged by Poisson noise. It is based on probabilistic similarities to compare noisy patches and patches of a pre-estimated image. An estimator of the risk for NL means, based on the idea of [6], has been derived for Poisson noise. This risk estimator is used in an optimization method to select automatically the filtering parameters in few iterations. Numerical results as well as visual results support the efficiency of the proposed method.

\textsuperscript{1}image courtesy of Y. Tourneur

\textsuperscript{2}Tetramethylrhodamine methyl ester (TMRM).
Figure 2. (a) Original images damaged by Poisson noise, denoised images obtained by (b) Poisson TV [14] and (c) the proposed Poisson NL means.

ACKNOWLEDGMENTS

The authors would like to thank Vincent Duval for its comments and for the reference [11].
REFERENCES


