

# An Intensity Similarity Measure in Low-Light Conditions

François Alter<sup>1,\*</sup>, Yasuyuki Matsushita<sup>2</sup>, and Xiaoou Tang<sup>2</sup>

<sup>1</sup> CMLA, École Normale Supérieure de Cachan

<sup>2</sup> Microsoft Research Asia

francois.alter@mines.org, {yasumat, xitang}@microsoft.com

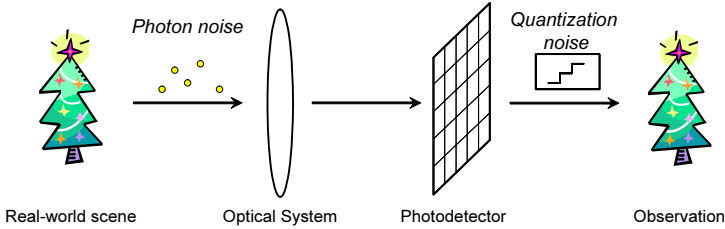
**Abstract.** In low-light conditions, it is known that Poisson noise and quantization noise become dominant sources of noise. While intensity difference is usually measured by Euclidean distance, it often breaks down due to an unnegligible amount of uncertainty in observations caused by noise. In this paper, we develop a new noise model based upon Poisson noise and quantization noise. We then propose a new intensity similarity function built upon the proposed noise model. The similarity measure is derived by maximum likelihood estimation based on the nature of Poisson noise and quantization process in digital imaging systems, and it deals with the uncertainty embedded in observations. The proposed intensity similarity measure is useful in many computer vision applications which involve intensity differencing, e.g., block matching, optical flow, and image alignment. We verified the correctness of the proposed noise model by comparisons with real-world noise data and confirmed superior robustness of the proposed similarity measure compared with the standard Euclidean norm.

## 1 Introduction

Noise is inevitable in any imaging device. A digital imaging system consists of an optical system followed by a photodetector and associated electrical filters. The photodetector converts the incident optical intensity to a detector current, or photons to electrons. During the process, the true signals are contaminated by many different sources of noise. In fact, the true signal itself has fluctuations in time due to the discrete nature of photons; the arrival of photons is not a steady stream and obeys the Poisson law [1]. It implies that no matter how accurately a computer vision experiment is performed, temporal fluctuation in intensity exists. The fluctuation becomes significant especially in low-light conditions where the number of incoming photons is limited, i.e., photon-limited conditions. In photon-limited conditions, quantization noise also becomes dominant due to the lack of intensity resolution in the limited dynamic range. Besides these noise sources, reset noise, dark current noise and read-out noise also become significant, and it is known that they can also be approximated by the Poisson noise model. Read-out noise is sometimes modelled by Gaussian noise; however, Gaussian noise with variance  $\sigma^2$  and mean  $\sigma^2$  is nearly identical to Poisson noise with

---

\* This work is done while the first author was visiting Microsoft Research Asia.



**Fig. 1.** Sources of noise in imaging process in low-light conditions. In low-light conditions, photon noise and quantization noise become dominant. In addition to them, there exist reset noise, dark current noise and readout noise that also cannot be ignored. It is known that they can be approximated by a Poisson distribution.

mean  $\sigma^2$  if  $\sigma$  is sufficiently large. Read-out noise satisfies this condition when operating a camera at room temperature and with a high read-out frequency.

Important low-light vision applications include night vision, medical imaging, underwater imaging, microscopic imaging and astronomical imaging. Even in daily situations, Poisson noise often becomes significant in high-speed imaging. In these situations, uncertainty in observations increases significantly due to Poisson noise and quantization noise. Particular operations under these conditions, such as template matching [8, 6, 12] and edge detection [5, 4, 2, 10], have been widely studied for many applications in fields such as object recognition and motion estimation. Image restoration is also one of the central problems since photon-limited images are usually severely degraded. Statistical methods [15, 14, 11, 13, 7], such as maximum likelihood (ML) estimation, are found to be effective since they can account for the special properties of the Poisson distribution. All of these techniques are found useful in low-light conditions; however, one fundamental question still remains open; *what is the similarity between two intensity observations with uncertainty?*

Intensity distance is often measured in many computer vision algorithms, and it is usually computed by Euclidean distance. Let  $k, l$  be two intensity measurements. The Euclidean distance  $d_E(k, l)$  is given by  $d_E^2(k, l) = (k - l)^2$ . It is correct for measuring the intensity distance between two signals when intensity noise is negligible, or a non-biased distribution of noise is assumed. However, they do not hold in photon-limited conditions where a significant amount of biased noise is added to the signal. Therefore, it is important to establish a new intensity similarity function which deals with the uncertainty embedded in observations.

In this paper, we describe a new noise model for low-light conditions and propose a new intensity similarity measure based upon the noise model. This paper has two primary contributions.

- **Poisson-quantization noise model:** A realistic noise model in low-light conditions is derived. The new noise model is built upon two inevitable noise sources: Poisson noise and quantization noise. We call the combined model of these two noise sources the Poisson-quantization noise, or PQ-noise in short. The proposed noise model is able to account for the uncertainty caused by

the nature of photon arrival and digitizing process. The correctness of the proposed noise model is confirmed by experiments with real-world data.

- **A new intensity similarity measure:** A new intensity similarity measure is proposed based on our noise model, which deals with the uncertainty caused by PQ-noise. The proposed similarity measure is useful in many computer vision applications which involve intensity differencing, e.g., block matching, stereo, optical flow and image alignment. The key advantage of the similarity function is that it can easily take place of existing intensity distance functions based on Euclidean distance. We compare the performance of the new intensity similarity function with the Euclidean distance function in order to verify the robustness of the proposed method against noise.

The outline of the paper is as follows: In Section 2, we briefly review the Poisson noise model and quantization noise model, and derive the PQ-noise model. Section 3 formalizes the intensity similarity function which measures the likelihood of two observations. The correctness and effectiveness of the PQ-noise model and the intensity similarity measure are verified with experiments described in Section 4.

## 2 Poisson-Quantization Noise Model

In this section, we first briefly review Poisson noise and quantization noise in Sections 2.1 and 2.2. We then formalize the Poisson-quantization noise model in Section 2.3.

### 2.1 Poisson Noise Model

Poisson noise is modelled by a Poisson distribution defined as follows.

**Definition 2.1:** A Poisson distribution [9] with parameter  $\lambda$  is defined for all  $k \in \mathbb{N}$  by the probability

$$p(k, \lambda) = \frac{\lambda^k}{k!} e^{-\lambda}, \tag{1}$$

with the mean  $E$  and variance  $V$  defined as follows.

$$E(\lambda) = \sum_{k=0}^{\infty} kp(k, \lambda) = \lambda, \quad V(\lambda) = \sum_{k=0}^{\infty} k^2p(k, \lambda) - E^2 = \lambda. \tag{2}$$

### 2.2 Quantization Noise Model

Quantization noise is the uncertainty caused by rounding observation amplitudes to discrete levels which occurs due to the finite amplitude resolution of any digital system. In analog-to-digital conversion, the signal is assumed to lie within a predefined range. Suppose the minimum number of electrons which is necessary to raise one level of observed intensity is  $q$ , and  $e_q$  is the quantization noise. The

count of electrons is proportional to the number of photons by the factor of the photon-electron conversion efficiency (the quantum efficiency of the sensor). A simple model of quantization noise can be described as

$$e_q = \frac{N}{q} - \left\lfloor \frac{N}{q} \right\rfloor, \tag{3}$$

where  $N$  is the number of electrons which are generated by the measurement. For a more detailed quantization error analysis in computer vision, readers are referred to [3].

### 2.3 Poisson-Quantization Noise Model

Now we formalize the PQ-noise model, which is the combination of Poisson noise and quantization noise.

**Definition 2.2:** A Poisson distribution with parameter  $\lambda$  and quantization  $Q = \{q_0 = 0, \dots, q_k, \dots, q_{n+1} = \infty\}$  is defined for all  $k \in \{0, \dots, n\}$  by the probability  $p(k, \lambda, Q)$  as

$$p(k, \lambda, Q) = \sum_{i=q_k}^{q_{k+1}-1} \frac{\lambda^i}{i!} e^{-\lambda}. \tag{4}$$

The quantization parameter  $q_k$  represents the minimum number of electrons which produces intensity level  $k$ .

Suppose the simple case where the quantization interval is constant, i.e.,  $q_k = kq$ , and the observations are far from saturation. The quantization interval is defined as the range of input values assigned to the same output level. With this simple model, we first observe the different behavior of PQ-noise from that of the Poisson noise model. We later relax the assumption to fit to a more realistic model. In this condition, the mean  $E(\lambda, q)$  and variance  $V(\lambda, q)$  of the PQ-noise model are given by<sup>1</sup>:

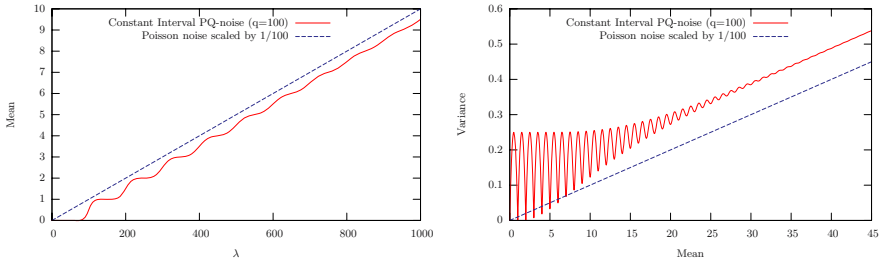
$$E(\lambda, q) = \frac{\lambda}{q} - \frac{1}{2} + \frac{1}{2q} + \frac{1}{q} \sum_{k=1}^{q-1} \frac{e^{\lambda(e^{\frac{2\pi ik}{q}} - 1)}}{1 - e^{-\frac{2\pi ik}{q}}}, \tag{5}$$

$$\begin{aligned} V(\lambda, q) = & \frac{\lambda}{q} + \frac{1}{12} - \frac{1}{12q^2} - \frac{2\lambda}{q^2} \sum_{k=1}^{q-1} \frac{e^{\frac{2\pi ik}{q}} e^{\lambda(e^{\frac{2\pi ik}{q}} - 1)}}{1 - e^{-\frac{2\pi ik}{q}}} + \frac{2}{q^2} \sum_{k=1}^{q-1} \frac{e^{-\frac{2\pi ik}{q}} e^{\lambda(e^{\frac{2\pi ik}{q}} - 1)}}{(1 - e^{-\frac{2\pi ik}{q}})^2} \\ & + \frac{1}{q} \sum_{k=1}^{q-1} \frac{e^{\lambda(e^{\frac{2\pi ik}{q}} - 1)}}{1 - e^{-\frac{2\pi ik}{q}}} - \frac{1}{q^2} \left( \sum_{k=1}^{q-1} \frac{e^{\lambda(e^{\frac{2\pi ik}{q}} - 1)}}{1 - e^{-\frac{2\pi ik}{q}}} \right)^2. \end{aligned} \tag{6}$$

As seen in the above equations and in Fig. 2, oscillation with an exponential decay  $e^{-\lambda(1-\cos(\frac{2\pi}{q}))}$  is observed in both mean and variance. The minimum intensity level of the linear range corresponds to  $\lambda \propto \frac{1}{1-\cos(\frac{2\pi}{q})} \approx \frac{q^2}{2\pi^2}$ .

---

<sup>1</sup> The derivation of the mean and variance is detailed in Appendix A.



**Fig. 2.** The mean and variance properties of PQ-noise. Due to quantization noise, strong oscillations are observed. Left: Evolution of the mean with respect to  $\lambda$  with  $q = 100$ . Right: Evolution of the variance with respect to the mean with  $q = 100$ .

Fig. 2 illustrates the difference between the ordinary Poisson noise model and the PQ-noise model. As shown in the figures, PQ-noise has oscillations due to the quantization noise component.

In practice,  $q_1$  does not equal to  $q$  due to the shift of the function caused by the offset voltage. Here we consider the more realistic model of the PQ-noise model with the following conditions:  $q_0 = 0$  and  $q_k = q_1 + (k - 1)q$ . It is possible to derive  $E(\lambda, q)$  and  $V(\lambda, q)$  for this condition by the same derivation used for Eqs. (5) and (6). When  $E(\lambda, q)$  is reasonably high, i.e.,  $E(\lambda, q) \gg \frac{q}{2\pi^2}$ , the following relationship<sup>2</sup> between  $E(\lambda, q)$  and  $V(\lambda, q)$  holds in the linear range.

$$V(E, q, q_1) = \frac{E}{q} + \frac{q^2 + 12q_1 - 6q - 7}{12q^2}. \tag{7}$$

This model has two unknown parameters  $q$  and  $q_1$ . These unknown parameters can be calibrated by fitting the observed noise data to Eq. (7).

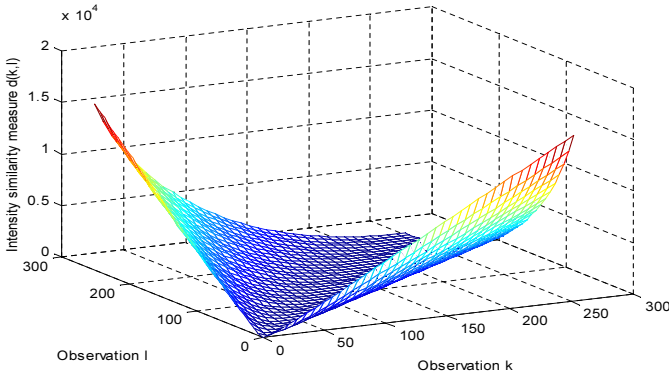
### 3 Derivation of Intensity Similarity Measure

Given two intensity observations  $k$  and  $l$ , what can we tell about the similarity between them? Usually, the similarity is measured by Euclidean distance with an assumption that  $k$  and  $l$  are *true* signals, or the noise model is non-biased. However, they do not hold in low-light conditions. We develop a new intensity similarity measure which is based upon the probability that two intensity observations come from the same source intensity. In this section, we first derive the intensity similarity measure for the Poisson noise case in order to make the derivation clear. We then develop the intensity similarity measure for the PQ-noise model. In fact, the Poisson noise model can be considered as the special case of the general PQ-model with the quantization parameter  $q = q_1 = 1$ .

#### 3.1 Poisson Noise Case

In the Poisson noise model, the case where two observed intensities arise from the same intensity distribution is equivalent to their sharing the same parameter

<sup>2</sup> The derivation of Eq. (7) is detailed in Appendix B.



**Fig. 3.** The intensity similarity function defined in Eq. (10)

$\lambda$ . If we assume that two intensity observations have the same parameter  $\lambda$ , the probability of obtaining two observations  $k$  and  $l$  is

$$P(k, l, \lambda) = p(k, \lambda)p(l, \lambda) = \frac{\lambda^{k+l}}{k!l!} e^{-2\lambda}. \tag{8}$$

This is obviously not sufficient to produce the actual probability because of the unknown parameter  $\lambda$ . However, the best case where  $\lambda$  maximizes the probability gives the measure which maximizes the similarity between  $k$  and  $l$ . Indeed, this approach corresponds to the ML estimation of the observation of the pair  $(k, l)$ . Therefore, the optimal  $\hat{\lambda}$  can be obtained by putting the first derivative  $\frac{\partial P}{\partial \lambda} = 0$ , and we obtain

$$\hat{\lambda} = \frac{k+l}{2}, \tag{9}$$

which maximizes the probability defined in Eq. (8). In this way, the intensity similarity function can be defined as

$$d(k, l) = -\ln(P(k, l, \hat{\lambda})) = (k+l) \left( 1 - \ln\left(\frac{k+l}{2}\right) \right) + \ln(k!) + \ln(l!). \tag{10}$$

Note that this similarity measure does not agree with the exact definition of distance because  $d(k, k) > 0$  if  $k > 0$ , but it produces the similarity between two observations.

Fig. 3 shows the intensity similarity function defined in Eq. (10). The function has a similarity to the  $l^2$  norm when two observed intensity levels are high. In fact, when  $k$  and  $l$  are sufficiently big,  $d(k, l)$  has a connection to the squared  $l^2$  norm. This can be shown by rewriting Eq. (10) with the approximation  $\ln(k!) \approx k \ln(k) - k$ :

$$\begin{aligned} d(k, l) &\approx k \ln(k) + l \ln(l) - (k+l) \ln\left(\frac{k+l}{2}\right) \\ &= f(k) + f(l) - 2f\left(\frac{k+l}{2}\right) \approx \frac{f''\left(\frac{k+l}{2}\right)(k-l)^2}{4} \approx \frac{(k-l)^2}{2(k+l)}. \end{aligned} \tag{11}$$

### 3.2 Poisson-Quantization Noise Case

Using the derivation in the previous section, we formalize the intensity similarity function for the PQ-noise model. We first define the joint probability  $P$ ; the probability of observing  $k$  and  $l$  having parameters  $\lambda_k$  and  $\lambda_l$  respectively with the quantization  $Q$  by

$$P(k, l, \lambda_k, \lambda_l, Q) = p(k, \lambda_k, Q)p(l, \lambda_l, Q). \tag{12}$$

We are seeking  $\hat{\lambda}$  ( $= \lambda_k = \lambda_l$ ) which maximizes the probability  $P$ . This is equivalent to maximizing the probability that two intensity observations share the same intensity source  $\hat{\lambda}$ . We denote  $P(k, l, \lambda, Q) = P(k, l, \lambda_k, \lambda_l, Q)$  when  $\lambda_k = \lambda_l$ . The optimal  $\hat{\lambda}$  always exists since the logged probability  $-\ln P$  is convex.<sup>3</sup>

1. If  $k = l$ , the maximum of  $P(k, k, \lambda, Q)$  is given by<sup>4</sup>

$$\hat{\lambda} = (q_k \dots (q_{k+1} - 1))^{\frac{1}{q_{k+1} - q_k - 1}}. \tag{13}$$

2. If  $k \neq l$ , we can obtain the optimal  $\hat{\lambda}$  by minimizing the convex function  $-\ln(P)$ . Here we describe a simple algorithm for finding the optimal  $\hat{\lambda}$  with a dichotomic search over the first derivative of  $P$ .

**Algorithm.** for finding the optimal  $\hat{\lambda}$   
 Input are  $k, l, q_k, q_{k+1}, q_l, q_{l+1}$  and  $n_{iter}$ .  
 Set  $\lambda_{min} = 0$  and  $\lambda_{max}$  sufficiently big, and  $n = 0$ ,  
 While  $n < n_{iter}$  and  $P' \neq 0$   
 do  
     Set  $\lambda \leftarrow \frac{\lambda_{min} + \lambda_{max}}{2}$  and  $n \leftarrow n + 1$ .  
     Compute  $P' = -\frac{\partial \ln(P(k,l,\lambda,Q))}{\partial \lambda}$   
     If  $P' < 0$  set  $\lambda_{min} \leftarrow \lambda$ .  
     If  $P' > 0$  set  $\lambda_{max} \leftarrow \lambda$ .  
 done  
 Set  $\hat{\lambda} \leftarrow \lambda$ .

The sign of  $P'$  can be determined by computing the sign of

$$\left( \frac{\lambda^{q_{k+1}-1}}{(q_{k+1}-1)!} - \frac{\lambda^{q_k-1}}{(q_k-1)!} \right) \sum_{i=q_l}^{q_{l+1}-1} \frac{\lambda^i}{i!} + \left( \frac{\lambda^{q_{l+1}-1}}{(q_{l+1}-1)!} - \frac{\lambda^{q_l-1}}{(q_l-1)!} \right) \sum_{i=q_k}^{q_{k+1}-1} \frac{\lambda^i}{i!}.$$

To find the optimal  $\hat{\lambda}$ , other descent methods such as gradient descent, Newton-Raphson, etc. can also be used alternatively.

<sup>3</sup> The convexity is proved in Appendix C.

<sup>4</sup> It can be derived from the first derivation of the function described in Appendix C.

In the above way, the optimal  $\hat{\lambda}$  is determined. The intensity similarity function is finally determined by plugging in the optimal  $\hat{\lambda}$  into the following function using Eqs. (4) and (12):

$$d(k, l, Q) = \min_{\lambda} \{-\ln(P(k, l, \lambda, Q))\} = -\ln(P(k, l, \hat{\lambda}, Q)) \quad (14)$$

$$= -\ln \left\{ e^{-2\hat{\lambda}} \left( \sum_{i=q_k}^{q_{k+1}-1} \frac{\hat{\lambda}^i}{i!} \right) \left( \sum_{j=q_l}^{q_{l+1}-1} \frac{\hat{\lambda}^j}{j!} \right) \right\}.$$

## 4 Experiments

In order to confirm our theoretical results, we performed experiments with real-world noise datasets. Our interests are 1) verifying the correctness of the proposed PQ-noise model and 2) confirming the superiority of the proposed intensity similarity measure over the standard  $l^2$  norm.

In order to obtain datasets, we mounted a video camera at a fixed position and captured an image sequence of a static scene in a low-light condition. Therefore, the only fluctuation in an image sequence is caused by noise. The images captured under the severe low-light conditions are almost totally black to human eyes, but they still contain intensity information. Fig. 4 shows one of such scenes used for the experiment. We captured raw image sequences by a Point Grey Dragonfly<sup>TM</sup> camera, and the intensities observed in the green channel are used for the entire experiment. For the illumination source, a DC light source is used to avoid intensity oscillations. We also used a small aperture and a short shutter speed to produce a low-light environment.

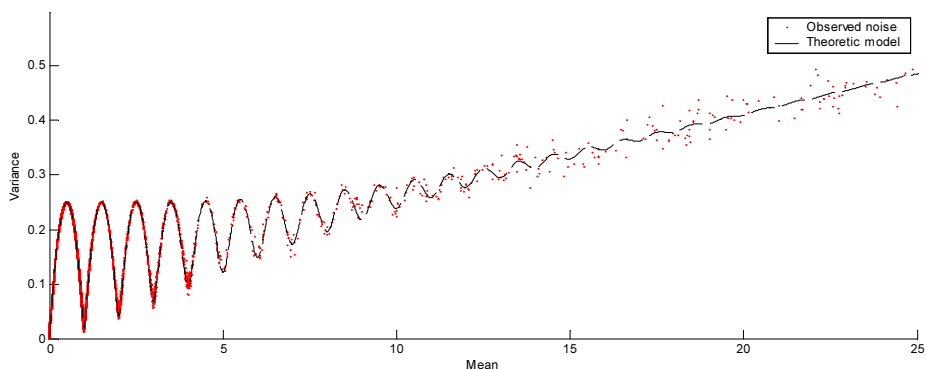
**PQ-noise model.** To verify the correctness of the PQ-noise model, we compared the mean-variance distribution of the real-world data with our analytic model described in Sec. 3.2. For the experiment, we have captured 1000 images of a static scene in a low-light condition. The unknowns  $q$  and  $q_1$  are both estimated by least squares fitting to the linear range of the PQ-noise model described in Eq. (7). Fig. 5 shows the plot of observations and the analytic model with estimated  $q$  and  $q_1$ . As shown in the figure, our analytic model well fits the actual observations, especially in the low intensity levels where the oscillation is observed clearly. The root mean-square error of observations from the theoretic curve is 0.0090 in Fig. 5.

**Intensity similarity measure.** To evaluate the robustness of the proposed similarity measure against noise, block matching is applied to the image sequences; if the block stays at the original position, the measure is not affected by noise. The same test is performed using the  $l^1$  and  $l^2$  norms over the same datasets, and we compared the outcomes of these norms with that of the proposed intensity similarity measure. We denote the  $l^1$  and  $l^2$  norm measure and our intensity similarity measure described in Eq. (14) as  $d_{l^1}$ ,  $d_{l^2}$  and  $d_{PQ}$  respectively. The parameters for  $d_{PQ}$ , i.e.,  $q$  and  $q_1$ , are calibrated beforehand by curve fitting as done in the previous experiment. The parameters  $q$  and  $q_1$  are estimated as  $q = 67$  and  $q_1 = 168$  in our experiment.





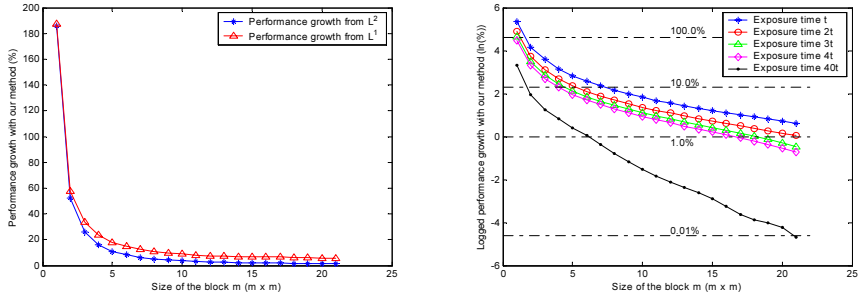
**Fig. 4.** One of the scenes used for the experiment. Left: The original input image in a low-light condition. Right: The left image is linearly scaled by 60.



**Fig. 5.** Evolution of the variance with the mean. The dotted line is the theoretical result with the calibrated parameters,  $q = 67$  and  $q_1 = 168$ . The dots are the measured noise which is obtained from 1000 images of a static scene under a low-light condition.

In the left image of Fig. 6, the performance growth obtained by our similarity measure in comparison with the  $l^1$  and  $l^2$  norm is shown. Our major interest is the comparison with the  $l^2$  norm; however, a comparison with the  $l^1$  norm is also given since the  $l^1$  norm is often used in practice because of its simplicity. In the left graph, 0% indicates the same performance as the  $l^1$  or  $l^2$  norm. The performance growth from the  $l^2$  norm, for example, is computed by  $\frac{N_c(d_{PQ}) - N_c(d_{l^2})}{N_c(d_{l^2})}$  where  $N_c(d_{PQ})$  and  $N_c(d_{l^2})$  are the number of correct matches with  $d_{PQ}$  and  $d_{l^2}$  respectively. The same computation is applied for the  $l^1$  norm case to obtain the performance improvement analysis as well. We can observe that our similarity measure significantly exceeds the  $l^1$  and  $l^2$  measures, especially when the block size is small. In fact, when the number of pixels in a block is large, the averaging of error reduces the biased property of the Poisson-quantization distribution. Therefore, the  $l^2$  norm which is adapted to Gaussian noise becomes effective as well.

In order to analyze the performance variation in different low-light conditions, we performed the same experiment over image sequences with different exposure



**Fig. 6.** Performance evaluation the proposed intensity similarity measure. Left: Performance growth of block matching using our intensity similarity measure in comparison with the  $l^1$  and  $l^2$  norms. Right: Variation of the performance growth in comparison with  $l^2$  norm under different low-light conditions. The logarithm is taken for a visualization purpose on the right graph.

settings. The dataset is obtained by capturing the same scene with changing exposure time, i.e.,  $t$ ,  $2t$ ,  $3t$ ,  $4t$  and  $40t$  where  $t = 15ms$ . For these datasets, the observed intensity values have a range of  $[0, 7]$ ,  $[0, 12]$ ,  $[0, 18]$ ,  $[0, 24]$  and  $[0, 235]$  respectively. The last setting  $40t$  is not considered a low-light condition, but we tested with this setting as well in order to see the behavior of our similarity measure in such a condition. The right graph of Fig. 6 shows the performance growth in comparison with the  $l^2$  norm using the different exposure settings. As shown in the graph, the proposed similarity measure is effective especially in severe low-light conditions. It can be seen that it also works for the ordinary condition (the  $40t$  setting), although the performance improvement from the  $l^2$  norm is almost zero.

## 5 Conclusion

In this work, we have proposed a new intensity similarity measure which is useful for low-light conditions where Poisson noise and quantization noise become significant. The intensity similarity measure is derived from the Poisson-quantization noise model which we develop as the combination of the Poisson noise model and the quantization noise model.

The correctness of the proposed PQ-noise model is verified by comparison with real-world noise data. The proposed intensity similarity measure is robust against Poisson-quantization noise, and is therefore effective in low-light conditions. The robustness is compared with the  $l^2$  norm using block matching, and we confirmed that the proposed method largely exceeds the performance of the  $l^2$  norm especially when the block size is small. Our intensity similarity measure is capable of achieving more accurate matching, especially in situations where large blocks cannot be used. The proposed noise model and intensity similarity measure are useful for many computer vision applications which involve intensity/image matching in photon-limited conditions.

## References

1. N. Campbell. Discontinuities in light emission. *Proc. Cambridge Phil. Soc.*, 15:310–328, 1909.
2. M. Das and J. Anand. Robust edge detection in noisy images using an adaptive stochastic gradient technique. In *Proc. of International Conference on Image Processing*, volume 2, pages 23–26, 1995.
3. B. Kamgar-Parsi and B. Kamgar-Parsi. Evaluation of quantization error in computer vision. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11(9):929–940, 1989.
4. A. Kundu. Robust edge detection. In *Proc. of Computer Vision and Pattern Recognition*, pages 11–18, 1989.
5. L. Liu, B.G. Schunck, and C.R. Meyer. Multi-dimensional robust edge detection. In *Proc. of Computer Vision and Pattern Recognition*, pages 698–699, 1991.
6. S.B. Lowen and M.C Teich. Power-law shot noise. *IEEE Trans. Information Theory*, 36:1302–1318, 1990.
7. H. Lu, Y. Kim, and J.M.M. Anderson. Improved poisson intensity estimation: Denoising application using poisson data. *IEEE Trans. on Image Proc*, 13(8), 2004.
8. G.M. Morris. Scene matching using photon limited images. *Journal of Opt. Soc. Am. A*, pages 482–488, 1984.
9. A. Papoulis. *Probability, Random Variables, and Stochastic Processes*, pages 554–576. McGraw-Hill, New York, 2nd edition, 1984.
10. P. Qiu and S.M. Bhandarkar. An edge detection technique using local smoothing and statistical hypothesis testing. *Pattern Recognition Lett.*, 17:849–872, 1996.
11. W.H. Richardson. Bayesian-based iterative method of image restoration. *Journal of Opt. Soc. of Am.*, 62:55–59, 1972.
12. R.E. Sequira, J. Gubner, and B.E.A. Saleh. Image detection under low-level illumination. *IEEE Trans. on Image Processing*, 2(1):18–26, 1993.
13. K.E. Timmermann and R.D. Nowak. Multiscale modeling and estimation of poisson processes with application to photon-limited imaging. *IEEE Trans. on Inform. Theory*, 45:846–862, 1999.
14. G.M.P. van Kempen, H.T.M. van der Voort, J.G.J. Bauman, and K.C. Strasters. Comparing maximum likelihood estimation and constrained tikhonov-miller restoration. *IEEE Engineering in Medicine and Biology Magazine*, 15:76–83, 1996.
15. Y. Vardi, L.A. Shepp, and L. Kaufman. A statistical model for positron emission tomography. *Journal of Amer. Statist. Assoc.*, 80:8–37, 1985.

## Appendix A

We derive the mean and variance for a PQ-distribution described in Eqs. (5) and (6). We begin with the following theorem about quantization.

**Theorem 1.** *Let  $X \in \mathbb{Z}$  be a discrete random variable which has a characteristic function  $\phi(t) = \mathbb{E}(e^{itX})$ . Its quantized version  $X_q$  can be defined by  $X_q = \lfloor \frac{X}{q} \rfloor$ . The characteristic function  $\phi_q$  of  $X_q$  is given by*

$$\phi_q(t) = \frac{1 - e^{-it}}{q} \sum_{k=0}^{q-1} \frac{\phi\left(\frac{t+2\pi k}{q}\right)}{1 - e^{-i\frac{t+2\pi k}{q}}}.$$

*Proof.* Let  $p_n$  be  $p_n = P(X = n)$ . The characteristic function  $\phi_q$  of  $X_q$  is

$$\phi_q(t) = \sum_{n=-\infty}^{\infty} p_n e^{i \lfloor \frac{n}{q} \rfloor t} = \sum_{n=-\infty}^{\infty} p_n e^{i \frac{n}{q} t} e^{i(\lfloor \frac{n}{q} \rfloor t - \frac{n}{q} t)}.$$

The function  $f : \begin{cases} \mathbb{Z} & \mapsto \mathbb{C} \\ n & \rightarrow e^{i(\lfloor \frac{n}{q} \rfloor t - \frac{n}{q} t)} \end{cases}$  is  $q$ -periodic, therefore it can be written

by a trigonometric polynomial  $\sum_{k=0}^{q-1} a_k e^{\frac{2\pi i k n}{q}}$ . Using a discrete Fourier transformation, we obtain:

$$\begin{aligned} a_k &= \frac{1}{q} \sum_{j=0}^{q-1} f(j) e^{-\frac{2\pi i j k}{q}} = \frac{1}{q} \sum_{j=0}^{q-1} e^{-i \frac{j}{q} t} e^{-\frac{2\pi i j k}{q}} = \frac{1}{q} \sum_{j=0}^{q-1} e^{-j \frac{i(t+2\pi k)}{q}} \\ &= \frac{1 - e^{-i(t+2\pi k)}}{q(1 - e^{-i \frac{t+2\pi k}{q}})} = \frac{1 - e^{-it}}{q(1 - e^{-i \frac{t+2\pi k}{q}})}. \end{aligned}$$

Therefore,

$$\begin{aligned} \phi_q(t) &= \sum_{n=-\infty}^{\infty} p_n e^{i \frac{n}{q} t} \sum_{k=0}^{q-1} a_k e^{\frac{2\pi i k n}{q}} = \sum_{k=0}^{q-1} \sum_{n=-\infty}^{\infty} p_n e^{i n \frac{t+2\pi k}{q}} \frac{1 - e^{-it}}{q(1 - e^{-i \frac{t+2\pi k}{q}})} \\ &= \sum_{k=0}^{q-1} \phi\left(\frac{t + 2\pi k}{q}\right) \frac{1 - e^{-it}}{q(1 - e^{-i \frac{t+2\pi k}{q}})}. \end{aligned}$$

*QED*

**Corollary 1.** *The mean  $E_q$  of  $X_q$  and variance  $V_q$  of  $X_q$  can be written by the mean  $E$ , the variance  $V$ , and the characteristic function  $\phi$  of  $X$ :*

$$\begin{aligned} E_q &= \frac{E}{q} - \frac{1}{2} + \frac{1}{2q} + \frac{1}{q} \sum_{k=1}^{q-1} \frac{\phi(\frac{2\pi k}{q})}{1 - e^{-\frac{2\pi i k}{q}}} \\ V_q &= \frac{V}{q} + \frac{1}{12} - \frac{1}{12q^2} - \frac{2E}{q^2} \sum_{k=1}^{q-1} \frac{e^{\frac{2\pi i k}{q}} \phi(\frac{2\pi k}{q})}{1 - e^{-\frac{2\pi i k}{q}}} + \frac{2}{q^2} \sum_{k=1}^{q-1} \frac{e^{-\frac{2\pi i k}{q}} \phi(\frac{2\pi k}{q})}{(1 - e^{-\frac{2\pi i k}{q}})^2} \\ &\quad + \frac{1}{q} \sum_{k=1}^{q-1} \frac{\phi(\frac{2\pi k}{q})}{1 - e^{-\frac{2\pi i k}{q}}} - \frac{1}{q^2} \left( \sum_{k=1}^{q-1} \frac{\phi(\frac{2\pi k}{q})}{1 - e^{-\frac{2\pi i k}{q}}} \right)^2. \end{aligned}$$

*These formulas are given by the computation of the derivatives of  $\phi_q$ , using the fact that  $E_q = -i\phi'_q(0)$  and  $V_q = -\phi''_q(0) - E_q^2$ .*

Eqs. (5) and (6) are the result of the previous formulas with a Poisson distribution which has the characteristic function  $\phi(t) = e^{\lambda(e^{it} - 1)}$ .

*Remark 1.* We can extend the previous result to the continuous random variable case. Let  $X$  be a random variable in  $\mathbb{R}$ , and  $\phi_X(t)$  its characteristic function. Let  $\lfloor X \rfloor$  be the quantized version of  $X$ . Using the Fourier series of the 1-periodic function defined over  $\mathbb{R}$ , i.e.,  $f(x) = e^{it(\lfloor x \rfloor - x)} = \sum_{n=-\infty}^{\infty} i \frac{e^{-i(t+2\pi n)} - 1}{t + 2\pi n} e^{2\pi i n x}$ , we can show that the characteristic function  $\phi_{\lfloor X \rfloor}$  of  $\lfloor X \rfloor$  is

$$\phi_{\lfloor X \rfloor}(t) = \sum_{n=-\infty}^{\infty} i \frac{e^{-i(t+2\pi n)} - 1}{t + 2\pi n} \phi_X(t + 2\pi n),$$

where the summation has to be done by grouping terms of  $n$  and  $-n$  together if it is not convergent. The mean  $E_{\lfloor X \rfloor}$  can be simply derived by

$$E_{\lfloor X \rfloor} = E_X - \frac{1}{2} + \sum_{n \neq 0} \frac{\phi_X(2\pi n)}{2\pi n}.$$

This applies to any kind of distribution, for example, the mean of a Gaussian distribution can be derived as follows. Let  $X$  be a random variable following the Gaussian law  $\mathcal{N}(\mu, \sigma^2)$ . The characteristic function becomes  $\phi_X(t) = e^{i\mu t - \frac{\sigma^2 t^2}{2}}$ . Therefore,

$$E_{\lfloor X \rfloor} = \mu - \frac{1}{2} + \sum_{n \neq 0} \frac{e^{i\mu 2\pi n - \sigma^2 2\pi^2 n^2}}{2\pi n} = \mu - \frac{1}{2} + \sum_{n=1}^{\infty} \frac{\cos(2\pi n \mu) e^{-\sigma^2 2\pi^2 n^2}}{\pi n}$$

As the series on the right is decreasing exponentially with  $n$ , this gives a practical way to compute  $E_{\lfloor X \rfloor}$ .

## Appendix B

We derive the PQ-noise model in the linear range described in Eq. (7). With an assumption that observed intensities are far enough from saturation, the approximation  $n = \infty$  can be used. We use the previous theorem described in Appendix A and divide the problem into two cases.

1. If  $q_1 \leq q$ , let  $X(\lambda)$  be a discrete random variable on the shifted Poisson distribution defined by

$$\forall k \geq q_1 - q, P(X(\lambda) = k) = \frac{\lambda^{k+q_1-q}}{(k + q_1 - q)!} e^{-\lambda}.$$

In this case,  $\lfloor \frac{X(\lambda)}{q} \rfloor$  has the wanted PQ-distribution. It is straightforward to see that  $E_{X(\lambda)} = \lambda + q_1 - q$ , and  $V_{X(\lambda)} = \lambda$ .

2. If  $q_1 > q$ , let  $X(\lambda)$  be a discrete random variable which satisfies

$$\begin{cases} P(X(\lambda) = 0) = \sum_{j=0}^{q_1-q} \frac{\lambda^j}{j!} e^{-\lambda}, \\ P(X(\lambda) = k) = \frac{\lambda^{k+q_1-q}}{(k + q_1 - q)!} e^{-\lambda}. \quad \forall k \geq 1 \end{cases}$$

Then again,  $\lfloor \frac{X(\lambda)}{q} \rfloor$  has the wanted PQ-distribution. When  $\lambda$  is big, the approximation  $P(X = 0) \approx 0$  holds. Therefore,  $E_{X(\lambda)} \approx \lambda + q_1 - q$ , and  $V_{X(\lambda)} \approx \lambda$  are deduced.

Using Corollary 1 and the remark, which shows  $\lim_{\lambda \rightarrow \infty} \phi_{X(\lambda)}(2k\pi) = \lim_{\lambda \rightarrow \infty} \phi'_{X(\lambda)}(2k\pi) = 0$ , we are able to deduce that  $E(\lambda, q, q_1) = \frac{\lambda+q_1}{q} - \frac{1}{2} + \frac{1}{2q}$  and  $V(\lambda, q, q_1) = \frac{\lambda}{q^2} + \frac{1}{12} - \frac{1}{12q^2}$  in the linear range. Therefore,

$$V(\lambda, q, q_1) = \frac{qE(\lambda, q, q_1) - q_1 + \frac{q}{2} - \frac{1}{2}}{q^2} + \frac{1}{12} - \frac{1}{12q^2} = \frac{E(\lambda, q, q_1)}{q} + \frac{q^2 + 12q_1 - 6q - 7}{12q^2}.$$

### Appendix C

We show that the probability function  $-\ln P$  in Eq. (12) is convex.

**Proposition:**  $f : \lambda \mapsto -\ln\left(\sum_{i=m}^n \frac{\lambda^i}{i!} e^{-\lambda}\right)$  is convex.

*Proof.*

$$\begin{aligned} f'(\lambda) &= \frac{-\sum_{i=m}^n \left(\frac{\lambda^{i-1}}{(i-1)!} - \frac{\lambda^i}{i!}\right) e^{-\lambda}}{\sum_{i=m}^n \frac{\lambda^i}{i!} e^{-\lambda}} = \frac{\frac{\lambda^n}{n!} - \frac{\lambda^{m-1}}{(m-1)!}}{\sum_{i=m}^n \frac{\lambda^i}{i!}} \\ f''(\lambda) &= \frac{\sum_{i=m}^n \left(\frac{\lambda^{n-1}}{(n-1)!} - \frac{\lambda^{m-2}}{(m-2)!}\right) \frac{\lambda^i}{i!} - \left(\frac{\lambda^n}{n!} - \frac{\lambda^{m-1}}{(m-1)!}\right) \frac{\lambda^{i-1}}{(i-1)!}}{\left(\sum_{i=m}^n \frac{\lambda^i}{i!}\right)^2} \\ &= \frac{\sum_{i=m}^n \frac{\lambda^{n+i-1}}{n!i!} (n-i) + \frac{\lambda^{m+i-2}}{(m-1)!i!} (i-m+1)}{\left(\sum_{i=m}^n \frac{\lambda^i}{i!}\right)^2} > 0. \end{aligned}$$

Therefore  $f$  is convex.

*QED*