Precedent-based Low Count Rate Image Intensity Estimation using Maximum Likelihood Distribution Descriptions

Viacheslav Antsiperov
Kotelnikov Institute of Radioengineering and Electronics
of Russian Academy of Sciences
Moscow, Russia, 125009
antsiperov@cupire.ru

Abstract—The work is devoted to the recognition / identification of the objects of predefined shape on the images of a special class. This class of images is associated with registration of weak radiations and, accordingly, characterized by a relatively low rate of counts (low count rate images). The latter leads to a low signal-to-noise ratio, low contrast and fuzzy shapes of the objects. For this reason, the known methods, designed for traditional image recognition, are not effective enough in this case and new recognition approaches, oriented to a low count rate images, are required. In this paper we propose such an approach. It is based on the machine learning paradigm and designed for identifying (low count rate) objects given by a point–sets. So, in the frames of the approach the recognition problem is posed as the statistical inference via already observed data (precedents), rather than the traditional in statistics problem of hypothesis testing. For the implementation of this recognition approach, we had to reduce the problem of object recognition to the machine learning problem of maximizing the tested point–set likelihood with respect to the classes of modelling object shapes up to shape size and position. The resulting recognition algorithms have a structure close to the well-known EM algorithm in the Variational Bayesian approach interpretation.

Index Terms—point process intensity shape identification, formal shape description, inhomogeneous Poisson point process, finite mixture models, machine learning, effective computational schemes, EM, VB EM algorithms

I. INTRODUCTION

The main problem considered in the paper is connected with the statistical inference about the intensity of an inhomogeneous spatial Poisson point process (PPP), whose realization is given as a set of discrete points (events) \( \{x_1, x_2, \ldots, x_n\} \). It is well-known [1], that such processes are a good model of the images related to the registration of weak radiation and, accordingly, characterized by a relatively low rate of counts (low count rate images). Such images are common in astronomy (optical radiation), computer and positron–emission tomography (X-ray radiation), thermography (infrared radiation), terahertz applications.

There are many methods and approaches to the PPPs analysis. Among the first works devoted to the analysis and evaluation of PPP characteristics, in particular its intensity, we should mention the papers of Cox [2], [3]. These papers are characterized by the systematic application of classical statistical methods to the class of stochastic point processes. Many problems of evaluating the PPP intensity and comparing the intensities of several PPP have been considered and solved. But it should be noted, that most problems were solved only for the case of homogeneous Poisson processes. The inhomogeneous case is analytically greatly complicated, and therefore the problems of estimating the temporal dynamics of intensity were not practically considered (except for the problem of trend estimation).

With the development of computer technologies, statistical methods have become more and more algorithmic in nature and more in line with the principles of machine learning approaches [4], [5]. Several complex problems have been solved by successfully combining processing power of modern computers and effective computational schemes of some algorithms (usually of recurrent type) developed over the last 20 to 30 years [6]. For example, several problems of estimating the Poisson process intensity, that widely use methods of machine learning are set forth in the work [7]. The author considers parametric intensities and discusses approaches to estimating their parameters depending on the type of the problem - either it is related to the estimation of the intensity of a single process or to the mixture (superposition) of the processes. It is noted that in the case of mixtures the EM (Expectation–Maximization) algorithm for Gaussian mixtures is most effective. Several different approaches concerning the estimation of random intensities are presented in [8]. The author of the proceedings considers Poisson processes with random intensities as twice stochastic, or “Cox processes”, and proposes his own method for estimating the intensity based on the Monte Carlo methods for Markov processes.

In this paper, we also propose a new method for identifying the intensity shape of inhomogeneous Poisson process based on a certain variant of machine learning approach. In accordance with the principles of statistical (machine) learning, the problem considered is posed as the statistical inference via already observed data, precedents, rather than the traditional
statistic problem of hypothesis testing. So, it is assumed that a finite set of precedents (examples) of Poisson process intensity (training set) has been observed, for each precedent some description has been formed and it is required to determine to which of these precedents should be attributed the newly registered inhomogeneous Poisson point process, given by a set of discrete points \( \{ \bar{z}_i \} \).

II. MODEL

To proceed to the solution of the problem posed in the work in the spirit of the machine learning, it is necessary to clarify some initial assumptions. Let us consider two main groups of them. The first group of assumptions concerns the Poisson point process (PPP) characteristics accepted as its description for some database of precedents. The second group is connected with the type of correspondence between the identified, registered process and the corresponding target precedent from the database.

Let us assume that each (k-th) precedent is the realization of spatial PPP with intensity \( \lambda_k(\bar{z}) \), defined in some plane region \( \bar{z} \in S \). We also assume that this intensity could be approximated by a superposition (mixture) of \( N_k \) Gaussian components (Gaussian mixture):

\[
\lambda_k(\bar{z}) = \Lambda_k \sum_{j=1}^{N_k} p_{k,j} \frac{\sqrt{\text{det}(A_{k,j})}}{2\pi} \exp \left\{ -\frac{1}{2} Q_{k,j}(\bar{z}) \right\},
\]

\( Q_{k,j}(\bar{z}) = (\bar{z} - \bar{m}_{k,j})^T A_{k,j}^{-1} (\bar{z} - \bar{m}_{k,j}), \quad (1) \)

where \( \Lambda_k \) is the integral over \( S \) of \( \lambda_k(\bar{z}) \) - total PPP intensity power, \( p_{k,j} \) is the fraction of power for j-th component, mean vector \( \bar{m}_{k,j} \in S \) and reverse covariance matrix \( A_{k,j} \) are the parameters of this component. Let us note, that \( \Lambda_k \) is also the parameter for Poisson distribution of \( n \) - total number of process points.

The power \( \Lambda_k \) and the set \( \{ (p_{k,j}, \bar{m}_{k,j}, A_{k,j}) \} \), \( j = 1, \ldots, N_k \) of the above listed parameters exactly describe intensity (1) of PPP. However, for comparing the intensities these parameters are not equally important. For example, the parameter \( \Lambda_k \), determining the overall "brightness" of the process does not say anything about intensity distribution, so it is not important when comparing the intensities in shape. Therefore, let us exclude \( \Lambda_k \) from the set of parameters describing the form of the mixture (1). As a result, the parameters describing each precedent will be the following: \( N_k \) - the number of mixture (1) components and the set of triples \( \{ (p_{k,j}, \bar{m}_{k,j}, A_{k,j}) \} \) in the number of \( N_k \) entities - parameters of all components of the mixture. Thus, setting \( \Lambda_k \) in (1) to unity, we come to a description of the precedents by parameters of Gaussian mixture which is the very popular probability distribution in statistics.

Regarding the introduced precedent feature (parameter) space, it is important to emphasize the following fact. For a fixed number \( n \) of Poisson process points, their coordinates \( \{ \bar{z}_i \} \) are the independent random variables identically distributed (iid) with the probability density \( \lambda_k(\bar{z})/\Lambda_k \) [7]. In this connection, the choice of the Gaussian mixture parameters space as a precedents feature space is important not only from the formal, discussed above point of view, but also from pragmatic position for implementing effective computational identification procedures. Indeed, within the framework of machine learning there are many effective algorithms that allow to find the maximum likelihood estimates of Gaussian component parameters for the sets of iid random points. This group of algorithms includes, for example, the popular EM-like algorithms [6] that recursively refine parameters \( \{ (p_{k,j}, \bar{m}_{k,j}, A_{k,j}) \} \) for Gaussian mixtures like (1).

Having determined the structure of the database of precedent descriptions and the way of its formation by EM-type algorithms, it is necessary to clarify the procedure for identifying any given test set of points whether it is an implementation of some description of precedent from the database. Theoretically, it would be possible to assume that there are precedents in the database of all possible shapes, so the description of the tested point-set will coincide with the normalized version (1) of one of them. Thus, if we form for \( n \) discrete samples \( \{ \bar{z}_i \} \) of the tested set the (logarithmic) likelihood function \( L_k(\{ \bar{z}_i \}) \) based on descriptions of all (k-th) precedents:

\[
L_k(\{ \bar{z}_i \}) = \ln \prod_{i=1}^{n} p(\bar{z}_i | k) = \ln \prod_{i=1}^{n} \sum_{j=1}^{N_k} p_{k,j} \frac{\sqrt{\text{det}(A_{k,j})}}{2\pi} \exp \left\{ -\frac{1}{2} Q_{k,j}(\bar{z}_i) \right\},
\]

then it would be possible to use the maximum likelihood (ML) principle to find the target precedent (K-th) as followed:

\[
K = \arg \max_k L_k(\{ \bar{z}_i \}).
\]

III. METHOD

Unfortunately, from practical point of view, the above straightforward approach to identification can’t be realized. Indeed, with this approach, for each shape of intensity (1) the shift of the coordinates \( \bar{z} \) origin, for example, will lead to a precedent different from the original one (with new parameters \( \bar{m}_{k,j} \) that differ from the old ones by the shift). Taking into account this and a number of other reasons, we consider the identification of the tested point-set and the target precedent from DB as a coincidence of their shapes \( \lambda(\bar{z}) \) and \( \lambda_k(\bar{z}) \) up to affine plane transformation (change in coordinates origin on \( \bar{c} \) and change of scale on \( s \)): \( \lambda(\bar{z}) \sim \lambda_k(s \bar{z} + \bar{c}) \). The log-likelihood function for this conditions should be rewritten as:
\[ L_k(\{ \tilde{x}_i \}) = \ln \left[ \int \rho_{\text{apr}}(\tilde{t}, s) \prod_{i=1}^{N_k} p(\tilde{x}_i | \tilde{t}, s, k) \, d\tilde{t} \, ds \right], \]

\[
p(\tilde{x} | \tilde{t}, s, k) = \sum_{j=1}^{N_k} p_{k,j} \frac{\sqrt{\det(A_k)}}{2\pi} s^2 \exp \left\{ -\frac{1}{2} Q_{k,j}(s \tilde{x} + \tilde{t}) \right\}, \tag{4}\]

where \( \rho_{\text{apr}}(\tilde{t}, s) \) is the a priori probability distribution of not related to the precedent description parameters \( \tilde{t} \) and \( s \).

In order to give the log-likelihood function (4) an appropriate EM algorithm form, in addition to the point-set of registered data \( \{ \tilde{x}_i \} \), let us introduce the hidden variables \( \{ z_i \} \), where \( z_i \in \{ 1, \ldots, N_k \} \) is an indicator of the belonging point \( \tilde{x}_i \) to the component \( z_i \) of the mixture. From this point of view, \( p(\tilde{x} | z, \tilde{t}, s, k) \) (4) is the marginal distribution of \( \tilde{x} \) from the joint distribution \( p(\tilde{x}, z | \tilde{t}, s, k) \) of the complete – observed and hidden data \( (\tilde{x}, z) \). Taking this into account, we rewrite (4) as follows:

\[
L_k(\{ \tilde{x}_i \}) = \ln \left[ \int \rho_{\text{apr}}(\tilde{t}, s) \prod_{z_1, \ldots, z_n, t_1=1}^{z_n} p(\tilde{x}_i, z_i | \tilde{t}, s, k) \, d\tilde{t} \, ds \right],
\]

\[
p(\tilde{x}_i, z_i | \tilde{t}, s, k) = p(\tilde{x}_i | z_i, \tilde{t}, s, k) = p_{k,s} \frac{\sqrt{\det(A_{k,z})}}{2\pi} s^2 \exp \left\{ -\frac{1}{2} Q_{k,z}(s \tilde{x} + \tilde{t}) \right\}.
\tag{5}\]

The computational aspects associated with calculating \( L_k(\{ \tilde{x}_i \}) \) (5) can be significantly complicated by the necessity to determine the sums for \( \{ z_i \} \) and integrals over \( \tilde{t} \) and \( s \). In order to avoid this complication, let us introduce the conditional distribution of \( \{ z_i \}, \tilde{t}, s \) (for a given \( \{ x_i \} \)):

\[
p \left( \{ z_i \}, \tilde{t}, s \mid \{ \tilde{x}_i \} \right) = \frac{p \left( \{ \tilde{x}_i \}, \{ z_i \}, \tilde{t}, s \mid k \right)}{p \left( \{ \tilde{x}_i \} \mid k \right)} = \frac{\rho_{\text{apr}}(\tilde{t}, s) \prod_{i=1}^{n} p(\tilde{x}_i, z_i | \tilde{t}, s, k)}{\int \rho_{\text{apr}}(\tilde{t}, s) \sum_{z_1, \ldots, z_n, t_1=1}^{z_n} p(\tilde{x}_i, z_i | \tilde{t}, s, k) \, d\tilde{t} \, ds}.
\tag{6}\]

Substituting left side (6) denominator into \( L_k(\{ \tilde{x}_i \}) \) (5), averaging the result over arbitrary distributions \( v_{z_1, \ldots, z_n}, w(\tilde{t}, s) \) and splitting the final expression into two terms, we get:

\[
L_k(\{ \tilde{x}_i \}) = F_{k,vw}(\{ \tilde{x}_i \}) + D_{k,vw}(\{ \tilde{x}_i \}),
\tag{7}\]

\[
F_{k,vw}(\{ \tilde{x}_i \}) = \int \sum_{z_1, \ldots, z_n} v_{z_1, \ldots, z_n} w(\tilde{t}, s) \times \left[ \frac{\rho_{\text{apr}}(\tilde{t}, s) \prod_{i=1}^{n} p(\tilde{x}_i, z_i | \tilde{t}, s, k)}{v_{z_1, \ldots, z_n} w(\tilde{t}, s)} \right] d\tilde{t} ds,
\tag{8}\]

\[
D_{k,vw}(\{ \tilde{x}_i \}) = -\int \sum_{z_1, \ldots, z_n} v_{z_1, \ldots, z_n} w(\tilde{t}, s) \times \ln \left[ \frac{p \left( \{ z_i \}, \tilde{t}, s \mid \{ \tilde{x}_i \}, k \right)}{v_{z_1, \ldots, z_n} w(\tilde{t}, s)} \right] d\tilde{t} ds.
\tag{9}\]

The first term \( F_{k,vw}(\{ \tilde{x}_i \}) \) (8) in (7) is usually called the “free energy”, and it can be, with accuracy of averaging over \( v_{z_1, \ldots, z_n}, w(\tilde{t}, s) \) (and the information term – \( \langle \ln v_{z_1, \ldots, z_n} w(\tilde{t}, s) \rangle \)) formally obtained from \( L_k(\{ \tilde{x}_i \}) \) (5) by permuting the sum and integral operations with the logarithm. This, obviously, greatly simplifies the calculations. The second term \( D_{k,vw}(\{ \tilde{x}_i \}) \) (9) in (7) is the Kullback-Leibler divergence of the distribution density \( p \left( \{ z_i \}, \tilde{t}, s \mid \{ \tilde{x}_i \}, k \right) \) with respect to \( v_{z_1, \ldots, z_n} w(\tilde{t}, s) \). It is always non-negative and vanishes in the only case - if the first distribution coincides with the second. That is why \( F_{k,vw}(\{ \tilde{x}_i \}) \) (8) is always a lower bound estimate of \( L_k(\{ \tilde{x}_i \}) \) (5), and this estimate is as better as the free energy is greater (or divergence \( D_{k,vw}(\{ \tilde{x}_i \}) \) is smaller). Based on the above analysis, we can formulate the following variational method (VM) of the likelihood function calculation:

\[
L_k(\{ \tilde{x}_i \}) \approx \max_{v_{z_1, \ldots, z_n}, w(\tilde{t}, s)} F_{k,vw}(\{ \tilde{x}_i \}).
\tag{10}\]

Let us note, that the equality in (10) is possible only in the case when \( p \left( \{ z_i \}, \tilde{t}, s \mid \{ \tilde{x}_i \}, k \right) \) can be factorized to the product of the hidden variables \( \{ z_i \} \) and parameters \( (\tilde{t}, s) \) distributions.

Solving the variational problem for the functional \( F_{k,vw}(\{ \tilde{x}_i \}) \) (8) in the usual way (using the Lagrange multipliers method), we obtain the following system for optimal solutions:

\[
\frac{\rho_{\text{apr}}(\tilde{t}, s)}{\sum_{v} w(\tilde{t}, s)} = \exp \left\{ \sum_{z_1, \ldots, z_n} v_{z_1, \ldots, z_n} \ln \left\{ \prod_{i=1}^{n} p(\tilde{x}_i, z_i | \tilde{t}, s, k) \right\} \right\},
\tag{11}\]

\[
= \frac{1}{\sum_{v}} \exp \left\{ \int w(\tilde{t}, s) \ln \left\{ \prod_{i=1}^{n} p(\tilde{x}_i, z_i | \tilde{t}, s, k) \right\} d\tilde{t} ds \right\},
\]

where \( \sum_{w} \) and \( \sum_{v} \) are the normalization constants (partition functions). The system of equations (11) can be simplified to:

\[
w(\tilde{t}, s) = \frac{\rho_{\text{apr}}(\tilde{t}, s)}{\sum_{w}} \prod_{i=1}^{n} W(\tilde{x}_i | \tilde{t}, s),
\]

\[
W(\tilde{x} | \tilde{t}, s) = \exp \left\{ \sum_{z_1} V_z(\tilde{x}) \ln \left[ p \left( \tilde{x}, z \mid \tilde{t}, s, k \right) \right] \right\},
\]

\[
v_{z_1, \ldots, z_n} = \frac{\prod_{i=1}^{n} V_z(\tilde{x}_i)}{\sum_{v}},
\tag{12}\]

\[
V_z(\tilde{x}) = \frac{1}{\sum_{v}} \exp \left\{ \int w(\tilde{t}, s) \ln \left[ p \left( \tilde{x}, z \mid \tilde{t}, s, k \right) \right] d\tilde{t} ds \right\},
\]
and takes the final explicit form when \( p(\tilde{z}, z \mid \tilde{t}, t, s, k) \) is substituted into it from (5):

\[
\begin{align*}
w(t', s) &= \frac{s^{2n} \rho_{app}(\tilde{t}, s)}{\Sigma_w} \prod_{i=1}^{n} W(\tilde{z}_i \mid \tilde{t}, \tilde{s}, k), \\
\tilde{W}(\tilde{z} \mid \tilde{t}, s) &= \exp \left\{ -\frac{1}{2} \sum_{i=1}^{N_z} \gamma_i A_{k,x} X_{i} + s \tilde{z} + \tilde{t} \right\}, \\
v_{z_1, \ldots, z_n} &= \prod_{i=1}^{n} V(z_i, \tilde{z}_i), \\
V_{\tilde{z}}(\tilde{z}) &= \frac{p_{k, z} \sqrt{\det(\Sigma_{k, x})}}{\Sigma_V} \times \\
&\times \exp \left\{ -\frac{1}{2} \int w(t', \tilde{s}) Q_{k,x}(s \tilde{z} + \tilde{t}) d\tilde{d}s \right\}, \\
\end{align*}
\]

(13)

Sometimes, with a suitable choice of proper \( \rho_{app}(\tilde{t}, s) \), conjugate, for example, to \( s^{2n} \prod_{i=1}^{n} W(\tilde{z}_i \mid \tilde{t}, \tilde{s}, k) \), system of equations (13) can be reduced to (nonlinear) algebraic system and solved by well-known methods. However, it is possible to go along a different way – assuming that the number of points is large enough \( n \gg 1 \) and taking into account that with increasing \( n \) the division \( w(t', s) \) becomes increasingly narrow (of the order \( \sim 1/n \) of the analysed components characteristic sizes) – let us try to find it in the asymptotic form \( w(t', s) = \delta(t - \tilde{t}_m) \delta(s - s_m) \). In this case, obviously, the problem will be reduced to finding the extremum point \( (\tilde{t}_m, s_m) \) of the general expression for \( w(t', s) \) (13). However, before we find this solution, let us make a couple of simplifications, also related to asymptotics \( n \gg 1 \).

Namely, with respect to \( \rho_{app}(\tilde{t}, s) \), let us assume that its arguments are distributed independently and, for example, the distribution of \( \tilde{t} \) is well approximated by a Gaussian one with zero mean and a standard deviation \( \Delta_t \), that exceeds at least the sizes of the mixture components, and the distribution of \( s \) is well approximated by a gamma distribution with shape parameter \( k \), mean \( \bar{S} \) and, correspondingly, standard deviation \( \Delta_s = \sqrt{\bar{S}/k} \). We note that in the asymptotic case \( n \gg 1 \), the product \( s^{2n} \rho_{app}(\tilde{t}, s) \) in (13) will be well approximated by the old Gaussian distribution for \( \tilde{t} \) multiplied by also Gaussian distribution with the mean \( \bar{S}_n \sim 2n\bar{S}/k \) and standard deviation \( \Delta_{s,n} \sim \sqrt{2n}\Delta_s/\sqrt{k} \) for the \( s \). In these approximations, the logarithm of \( w(t', s) \) (13) becomes quadratic in the parameters \( (t, s) \), so its extremum can be easily found from the following linear system:

\[
P\left(\begin{array}{c}
s_m \\
\tilde{t}_m
\end{array}\right) = \left(\begin{array}{c}
\sigma \\
\tau
\end{array}\right)
\]

(14)

where:

\[
\begin{align*}
\gamma_x &= \sum_{i=1}^{N_z} \alpha_{x_i}(\tilde{z}_i) ; \\
\chi_x &= \sum_{i=1}^{n} \sum_{i=1}^{N_z} \alpha_{x_i}(\tilde{z}_i) \tilde{x}_i ; \\
B_x &= \frac{1}{\gamma_x} \sum_{i=1}^{N_z} \alpha_{x_i}(\tilde{z}_i) \tilde{x}_i \tilde{x}_i^T .
\end{align*}
\]

(15)

Thus, given an initial, neutral approximation for the parameters \( (t, s) \), for example, putting \( \tilde{t}_m(0) = 0, s_m(0) = 1 \), on the base of (16) we can find initial approximation to \( \{ V_z(\tilde{z}_i) \} \) for all \( \{ \tilde{z}_i \} \) and all corresponding \( \{ z_i \} \). Using them, we can compute by formulas (15) the matrix \( P \) and the right-hand side \( (\sigma, \tau) \) of (14). Getting the solutions of this linear system and putting them as the refinements \( (\tilde{t}_m, s_m) \) of the initial parameters, we can refine \( \{ V_z(\tilde{z}_i) \} \) (16), etc. So, we get a recursive algorithm in a form very reminiscent of EM:

\[
E: \quad \text{for } i \in \{1, \ldots, n\}, \quad \{ z \in \{1, \ldots, N_k\} \} \\
V_{z_i}^{(j)}(\tilde{z}_i) &= \frac{p(\tilde{z}_i, z \mid \tilde{t}_{(j-1)}, s_{(j-1)}, k)}{\sum_{z=1}^{N_k} p(\tilde{z}_i, z \mid \tilde{t}_{(j-1)}, s_{(j-1)}, k)} ;
\]

(17)

\[
M: \quad \text{for } z \in \{1, \ldots, N_k\} \text{ calculate} \\
\gamma_z &= \sum_{i=1}^{n} V_{z_i}^{(j)}(\tilde{z}_i) ; \quad \chi_x = \frac{1}{\gamma_x} \sum_{i=1}^{n} V_{z_i}^{(j)}(\tilde{z}_i) \tilde{x}_i ; \\
B_x &= \frac{1}{\gamma_x} \sum_{i=1}^{n} V_{z_i}^{(j)}(\tilde{z}_i) ; \\
\text{find } P, \sigma, \tau \text{ and solve} \\
P(\tilde{t}_m^{(j)}, s_m^{(j)}) &= \left(\begin{array}{c}
\sigma \\
\tau
\end{array}\right)
\]

The constructed algorithm (17) recursively, with increasing number of indexed by \( j \) iterations, refines the parameters \( (\tilde{t}_m^{(j)}, s_m^{(j)}) \). When the iterations begin to show signs of stabilization for any of the calculated values – \( \tilde{t}_m^{(j)} \), \( s_m^{(j)} \), or even \( \{ V_{z_i}^{(j)}(\tilde{z}_i) \} \), the algorithm can be stopped. On the basis of these values full solutions (13) \( v_{z_1, \ldots, z_n} = \prod_{i=1}^{n} V_{z_i}^{(j)}(\tilde{z}_i) \), \( w(t, s) = \delta(t - \tilde{t}_m) \delta(s - s_m) \) can be found. But, in fact, it is not needed. It turns out that the expression for the free energy \( F_{k,v,w}(\{ \tilde{z}_i \}) \) (8), which is the
goal of the computation, depends only on the estimates \( \tilde{t}(j) \), \( s(j) \). Indeed, substituting the solutions written above in (8) we obtain:

\[
F_{k,w}(\{\tilde{z}_i\}) = \sum_{x_1, \ldots, x_n} v_{x_1, \ldots, x_n} \times \\
\ln \left[ \rho_{apr}(\tilde{t}(j), s(j)) \prod_{i=1}^n p(\tilde{z}_i | \tilde{t}(j), s(j), k) \right] = (18)
\]

where it is taken into account that information \( \ln [w(\tilde{t}, s)] \) for the singular distribution \( w(\tilde{t}, s) \) is zero (at least after finite approximation). The value of \( \phi \) under the logarithm in (18) is the joint distribution of the recorded PPP data \( \{\tilde{z}_i\} \) and the current estimates of the parameters \( (\tilde{t}, s) \). Using all these data \( \ln \phi \) can be easily calculated starting directly from the original definitions (see (4)), it has a transparent interpretation and, therefore, as a criterion controlling the execution of iterations of the algorithm (17), \( \ln \phi \) is the best choose.

IV. CONCLUSIONS

In this paper a complete solution of the problem of Poisson process intensity identification is presented. For the implementation of the solution, a computational scheme is proposed. This scheme is very close in structure to the popular in machine learning EM algorithm. Since many technical aspects of the computer implementation of EM algorithms are now well-developed [6], one can expect that the implementation of the proposed solution will be also quick and effective.

REFERENCES


