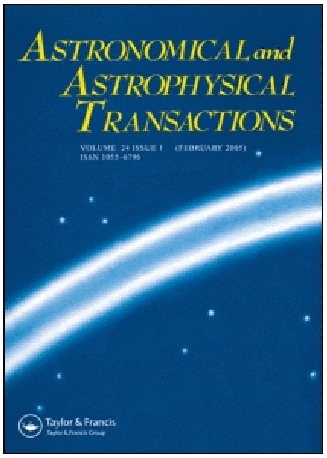

Astronomical & Astrophysical Transactions
The Journal of the Eurasian Astronomical Society
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713453505

Image restoration: Method-independent limit of efficiency and its realization

V. Yu. Terebizh [a]
[a] Crimean Station of Sternberg Astronomical Institute, Nauchny, Crimea

Online Publication Date: 01 January 1991
To cite this Article: Terebizh, V. Yu. (1991) 'Image restoration: Method-independent limit of efficiency and its realization', Astronomical & Astrophysical Transactions, 1:1, 3 - 29
To link to this article: DOI: 10.1080/10556799108244516
URL: http://dx.doi.org/10.1080/10556799108244516

Astronomical and Astrophysical Transactions, 1991,
Vol. 1, pp. 3–29
Reprints available directly from the publisher.

IMAGE RESTORATION: METHOD-INDEPENDENT LIMIT OF EFFICIENCY AND ITS REALIZATION

V. YU. TEREBIZH

Crimean Station of Sternberg Astronomical Institute, 334413, Nauchny, Crimea

Due to radiational and external noise, any image forming system can be described only in a frame of a stochastic model, with probability $f(N, S)$ of occurrence of the observed image N, for any searched object S. Thus, restoration problem must inevitably be considered as a statistical estimation of unknown parameters S. Proceeding from the appropriate definition of restoration efficiency, it is possible to show that the method-independent limit of efficiency and accuracy is set by the Rao–Cramer theorem. The most promising way to achieve the theoretical limit is based on a maximum likelihood image restoration (MLIR) method. Indeed, MLIR gives limiting accuracy when light intensity begins to exceed the mean level of external noise. The maximum entropy method includes non-necessary, and logically inconsistent, assumptions.

The concrete formulae for the calculation of an object estimate, and its accuracy, are given. Test cases are given, and examples, of using the proposed approach to different inverse problems.

KEY WORDS Image processing, image restoration.

1. INTRODUCTION

The problem of image restoration in its strict meaning assumes the existence of an unknown initial image, *object* $S(x)$, that was blurred by some image formation system with a *Point Spread Function* (PSF) and by random noise. It is necessary to restore the object as accurately as possible, on the basis of the *observed image*, and *a priori* information about the object, PSF and noise. This problem constitutes the particular case of the general class of *inverse problems*, and one can imply another meaning of the above functions at different circumstances. We will discuss such possibilities later (see Section 9). Now let us consider the image formation process with its specific feature: non-negative object generated by stochastic distribution of quantum events.

During more than fifty years of image restoration theory development, a number of relevant methods were proposed (see reviews by Frieden, 1979, by Vasilenko and Taratorin, 1986 and by Jain, 1989, ch. 8). The choice of one of the methods in a given circumstance, usually depends on a computer power, traditions or other attendant factors. From a rigorous point of view, just the existence of such a huge number of the image restoration methods, shows clearly the presence of serious difficulties in this field. Indeed, it is not a simple problem even to compare the power of various methods because it depends on a specific situation.

At the same time, it seems evident that *under given conditions of observations, including a priori information, the signal to noise ratio and the PSF form, a method-independent natural limit exists of image restoration efficiency*. The existence of such a limit allows us to introduce an *absolute* order in the power of various methods, and to give the quantitative estimate of the accessible errors of restoration.

The purpose of this paper is to discuss all aspects connected with the meaning and realization of the theoretical limit of restoration accuracy. The discussion proceeds from six papers of the author, and collaborators at *Astrofizika*, from which the first part was already published (Terebizh, 1990). For brevity, we designate the parts of this series as MLIR-1, . . . , MLIR-6. Further necessary references are given in the text.

The main inferences of subsequent discussion are as follows.

(i) Due to inner (radiational) and external noise, any image forming system can be described only in a frame of some *stochastic* model. The sole aim of this model is to provide the conditional probability Pr (*image* | *object*) $\equiv f(N, S)$ of occurrence of the observed image N, under a given object that we describe by a set of unknown *parameters* $S \equiv (S_1, \ldots, S_n)$, for example, object intensities at n pixels, or some structural parameters.

(ii) The stochastic nature of the image formation process *inevitably* brings the restoration problem to the estimation theory of unknown parameters S. We will name this method of thought for brevity, as *statistical parameters estimation* (SPE).

(iii) Proceeding from the appropriate definition of efficiency of a restoration method (for example, in a *root-mean-square* sense) it is possible to show that the *method-independent* limit of efficiency and accuracy is given by the Rao–Cramer theorem.

(iv) The most promising way to achieve the theoretical limit of efficiency is connected with the *maximum likelihood* principle. This statement is based on a theorem that maximum likelihood estimate coincides with the most efficient estimate, if this latter exists at all.

(v) In a frame of suitably chosen stochastic models of image formation, the *maximum likelihood image restoration* (MLIR) method demonstrates the limiting efficiency when light intensity becomes approximately equal to the mean level of external noise.

(vi) The comparative study of SPE, and the now widely used *maximum entropy* (MEM) approaches, show that MEM includes the non-necessary and logically inconsistent assumptions. Therefore, all versions of this method cannot avoid the investigator's influence on the choice of solution. On the contrary, the SPE approach takes into account *all* available information about the object, imaging system, and noise, and *only this* information.

The concrete expressions for calculation of the object estimate and its accuracy (*error corridor*) are also given.

To simplify notations we use one-dimensional version of the problem; it is easy to rewrite the final expressions at needed form.

2. IMAGE RESTORATION AS A STATISTICAL PROBLEM OF PARAMETERS ESTIMATION

Due to the quantum nature of light, we have dealt only with an ensemble of photoevents: the grains of photoemulsion, photoelectrons in the transparent cathodes, the electron-hole pairs in CCD's and so forth. Similar events take place in electronic images. Obviously, an ensemble of events is stochastic, and its statistical properties depend on the nature of incident radiation and detector. The events are obey the classical statistics in view of its spatial distinguishability.

The subsequent discussion will be much clearer if we firstly consider the simple models of image formation. This is when one attempts to find the object estimate, solely on the basis of the finite sample of events or under zero external noise. Since it is very difficult to avoid introducing external noise, these cases do not have practical significance, but from the theoretical point of view they are quite useful. After considering these "mental experiments", more realistic models are investigated with the remaining previous approach to the restoration problem. We will follow Terebizh (1990) and MLIR-5.

2.1. *Model A*

Let us introduce, at once, necessary discretization and consider the unknown theoretical brightness distribution at the object as non-negative vector $s \equiv (s_1, \ldots, s_n)$ normalized in the following way:

$$0 \leq s_k \leq 1, \qquad \sum_{k=1}^{n} s_k = 1. \tag{1}$$

If we have an ideal imaging system, the distribution s of events will result after the registration of infinitely large number of photons. But every image consists only from some finite number of events, for example, L. As a consequence, the real sampled distribution of events $v \equiv (v_1, \ldots, v_n)$ is different from the expected (mean) distribution $S_k = L \cdot s_k$. Of course,

$$\sum_{1}^{n} v_k = \sum_{1}^{n} S_k = L. \tag{2}$$

We have here the typical sequence of independent trials (Feller, 1966), so the probability $f(v, s)$ to obtain the given vector (v_k), that satisfies the condition (2), is obeyed to the multinomial distribution:

$$f(v, s) = \frac{L!}{v_1! \ldots v_n!} s_1^{v_1} \ldots s_n^{v_n}, \tag{3}$$

and the depending from unknown object part is:

$$\ln f(v, s) = \sum_{k=1}^{n} v_k \cdot \ln s_k + \text{const}. \tag{4}$$

As we will see later, just the expression for probability f is necessary in order to find and to compare different estimates of the object.

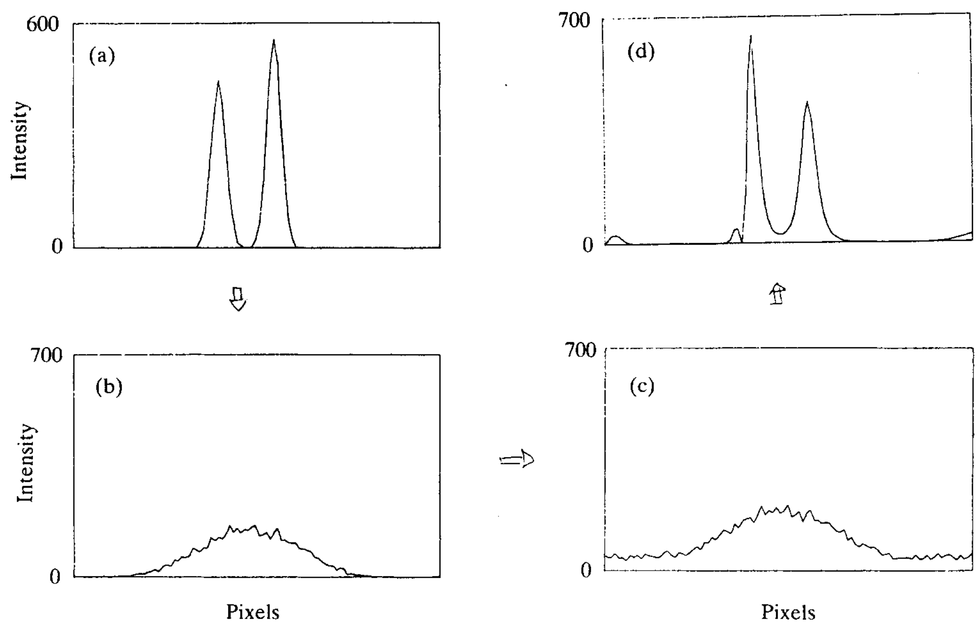

Figure 1 Example of blurring and restoring of an object; a—initial image (object), b—result of stochastic smoothing by an imaging system, c—observed image after adding external noise to previous image, d—result of restoration with the aid of MLIR method.

2.2. *Model B*

In a practice, the imaging system redistributes events from each pixel to a wide set of other pixels, so the array of n initial pixels will be transformed to the array of $m \geq n$ pixels (Figure 1b). The concrete manner of this smoothing can be taken into account by a suitably chosen model. The linear smoothing without external noise was considered by Richardson (1972), Lucy (1974) and Kosarev *et al.* (1983).

Assume, according to Terebizh (1990), that in conditions of Model B the imaging system works in such a way. Any initial light event that for the ideal system should take place, say, at pixel k, may be distributed in real imaging system to pixel j with probability h_{jk}, where:

$$\sum_{j=1}^{m} h_{jk} = 1, \qquad m \geq n. \tag{5}$$

Each of the L events is distributed independently from the others. So, the imaging system is linear, and for any event, the probability to reach pixel j becomes:

$$p_j(s) = \sum_{k=1}^{n} h_{jk} \cdot s_k, \qquad \sum_{j=1}^{m} p_j = 1. \tag{6}$$

Now we have the sequence of independent trials with the probability array

(p_j), and the probability of occurrence, the numbers (v_k) is equal to:

$$f(v, s) = \frac{L!}{v_1! \ldots v_m!} p_1^{v_1} \ldots p_m^{v_m}. \tag{7}$$

Equations (6) and (7) define $f(v, s)$ for any (s_k). We have also:

$$\ln f(v, s) = \sum_{j=1}^{m} v_j \cdot \ln p_j(s) + \text{const.} \tag{8}$$

2.3. *Model C*

Of course, the most interesting is the case, when the image is blurred due to both inner fluctuations and external noise. The corresponding consideration was given by the author (Terebizh, 1990). We give here only the final results.

As regards to the smoothing, we assume the same linear stochastic process that was introduced in Models A and B. Matrix (h_{jk}) is assumed to be known, but it is easy to generalize the theory and input in the PSF some free parameters. If PSF is the Kronecker symbol, we have the problem of filtering a noisy image.

The external noise is defined by a random vector (ξ_j), that has to be added to the numbers of light events in corresponding pixels of smoothed image (Figure 1c). All variables (ξ_j) are supposed to be mutually independent (*white noise*) and have the Poisson distribution:

$$\Pr(\xi_j = r) = \exp(-b_j) \cdot \frac{(b_j)^r}{r!}, \qquad \begin{cases} r = 0, 1, \ldots; \\ j = 1, 2, \ldots, m. \end{cases} \tag{9}$$

Introduce designations $\mathbf{E}(\alpha)$ for mean value and $\mathbf{D}(\alpha)$ for variance of a random variable α. For Poisson distribution

$$\mathbf{E}(\xi_j) = \mathbf{D}(\xi_j) = b_j, \tag{10}$$

and the mean background may be both uniform ($b_i = b = \text{const}$) or arbitrarily varying.

At last, the observed image (N_j), $j = 1, \ldots, m$, is defined by the distribution of registered events including redistributed light events and noise. We observe:

$$N \equiv \sum_{j=1}^{m} N_j = L + \sum_{j=1}^{m} \xi_j \tag{11}$$

events at all. Notation $S_k \equiv L \cdot s_k$ is also used.

The strict expression for distribution $f(N, S)$ in the Model C is complicated. By using the Darwin–Fowler asymptotic approximation (valid for the number of events beginning from a few dozen), it was found:

$$\ln f(N, S) = \sum_{j=1}^{m} N_j \cdot \ln \left[\Lambda \cdot \sum_{k=1}^{n} h_{jk} S_k + b_j \right] - L \cdot \ln \Lambda - \sum_{k=1}^{n} S_k + \text{const}, \tag{12}$$

where $\Lambda(N, S)$ is the root of equation:

$$\sum_{j=1}^{m} \frac{N_j b_j}{\Lambda \cdot \sum_{k=1}^{n} h_{jk} S_k + b_j} = N - L. \tag{13}$$

One can expect from the central limit theorem that approximation $\Lambda \equiv 1$ will be better with more total numbers of events.

2.4. *Incoherent source*

We turn now to a very important practice case, when *a priori* information suggests an incoherent nature of the object's radiation.

Both the semi-classical approach and the strict quantum theory of time distribution of photoevents, shows that for incoherent radiation, the photoevents sequence is a *twice stochastic Poisson process* or *Cox–Mandel process* (Cox, 1955; Mandel, 1958, 1959; Loudon, 1973). This process follows from the usual Poisson process if its intensity in one's turn is assumed to be the stationary stochastic process. It is difficult to investigate the general properties of the Cox–Mandel process (Mehta, 1970; Terebizh 1991), but, fortunately, for a very wide sphere of phenomenons this process can be approximated by the usual Poisson sequence. It is necessary to have the exposition time much larger than the coherence time of light. Just this condition is fulfilled for typical astronomical or physical experiment. We consider, below, the temporal process of image formation for incoherent sources as a simple Poisson process with a constant intensity.

We can consider this model, by randomizing the number of light events L at Model C in accordance with Poisson law, or by randomizing the independent fluxes in pixels. The details one can find in MLIR-5, the final expression for density distribution is evident:

$$f(N, S) = \prod_{j=1}^{m} \exp(-\lambda_j) \cdot \frac{\lambda_j^{N_j}}{N_j!}, \tag{14}$$

where $S \equiv (S_k)$ is vector of mean intensities of the incident radiation (object) and:

$$\lambda_j(S) = \sum_{k=1}^{n} h_{jk} \cdot S_k + b_j \tag{15}$$

are the mean quantities of events in the blurred and noised image. We have, from the last two expressions, and (5):

$$\ln f(N, S) = \sum_{j=1}^{m} N_j \cdot \ln \lambda_j(S) - \sum_{k=1}^{n} S_k + \text{const.} \tag{16}$$

Again, this expression follows from (12) if we take $\Lambda = 1$. It should be noticed that functional (16) can be successfully used as an approximation to (12) even in the case, when Poissonian scheme is not valid; the reason for this is that it is connected with the law of large numbers.

2.5. *Formulation of Image Restoration Problem*

The number of model examples that are similar to those considered above can be easily increased. The first task of the investigator is choosing the appropriate image formation model that accounts data, and *a priori* information. The next step is connected with calculating the probability $f(N,S)$ to obtain an observed image for any object. At last, from the main characteristic of the concrete problem $f(N, S)$, we ought to find the method-independent accuracy that it is possible to achieve under the given data, and to provide a method for the realization of this accuracy.

It should be stressed, at first, an important fact. Namely, the finiteness of the sample of events, and the stochastic nature of the smoothing process, inevitably give rise to a distribution with fluctuations (see Figure 1). This inner, or radiational, noise plays a significant role, both for bright and faint images. Indeed, the statistics of events is usually close to the Poisson law, and for this case a relative value of radiational fluctuations $I_{\text{light}}^{-1/2}$ is negligible in comparison with additive noise fluctuations, $I_{\text{noise}}^{-1/2}$ only when the light intensity is much less than the mean level of noise. For this reason the restoration of images has a characteristic feature, in compare with a general inverse problem. The inner noise of the image *must* be taken into account. Particularly, the usual treatment of linear image smoothing by integral term $\int h(x, x')S(x')\,dx'$, is incorrect.

We now come to the natural formulation of the restoring problem: *it is necessary to find, some "good" estimates of unknown parameters on the basis of the observed image* $(N_1, \ldots, N_m)$, *and available a priori information about the object, imaging system, and noise.*

Up to now, *a priori* information about the object was confined only by requirements of its non-negativity and, perhaps, some information about its extent. The first property in a natural way was taken into account by the choice of probabilistic distributions. Very often there is some additional *a priori* information that allows us to significantly decrease the number of unknown variables. Let us consider, for example, the case when initial distribution concentrated in a few pixels, that is, looks like discrete spectrum:

$$S_k = I_k \cdot \delta_{k,k_t}, \qquad t = 1, 2, \ldots, T < n. \tag{17}$$

Then we only have $2T$ unknown parameters, and it is possible to achieve more deep restoration. Another case gives the relation:

$$S_k = \alpha \cdot U\left(\frac{k - \beta}{\gamma}\right), \tag{18}$$

where $U(x)$ is a given function, and α, β, γ are unknown parameters (for example, we can imagine that U describes the law of brightness distribution in elliptical galaxies, when we have dealt with these objects).

Supplementary information can be taken into account by a similar manner. Just from this information depends the ultimate quality of the restored image. For simplicity we imply further that unknown set of parameters S can include both some structural constants and object intensities, in arbitrary pixels. It is only important that the number of these parameters n is no more than the number of pixels of the observed image m. In the above discussed sense, the image restoration problem means the searching for a set of estimates of any parameters that define the model of image formation.

It is easy to understand that there are infinite methods of estimating an unknown set of parameters, that is, of restoring the blurred image. The best from them must have the least scattering near the true object; if a concrete model allows, (we will see later that it is possible only when the probability density $f(N, S)$ has a special form) the scattering achieves the theoretical Rao–Cramer lower limit. This latter case is usually characterized by a statement that the *efficient* estimate exists.

3. NUMERICAL EXAMPLE

In order to compare different methods, and to introduce an absolute limit of their efficiency, we ought to consider rather complicated mathematical concepts. It

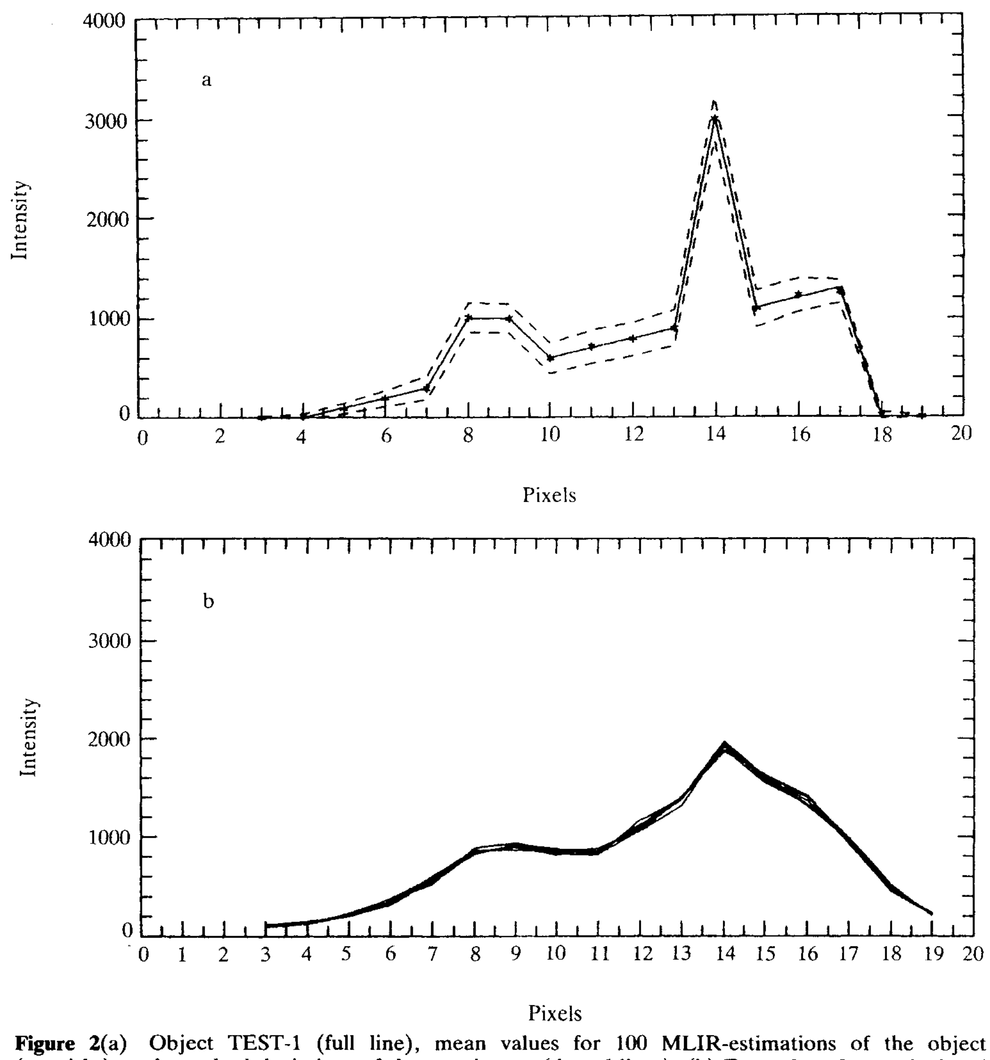

Figure 2(a) Object TEST-1 (full line), mean values for 100 MLIR-estimations of the object (asterisks), and standard deviations of these estimates (dotted lines). (b) Examples of smoothed and noised images of TEST-1. (c) Examples of restored with MLIR images of TEST-1. (d) Calculated standard deviations of MLIR-estimates according to 100 simulations (asterisks), and Rao–Cramer lower boundary at first approximation (dotted line).

seems desirable to analyze, at first, a simple example that shows the necessity and meaning of the chosen description. The full data concerning this example are given in MLIR-6.

The one-dimensional object TEST-1 (Figure 2) has extension $n = 17$, and its blurred and noised image has extension $m = 21$ pixels. The sampled intensities were considered as simulations of independent Poisson random variables, with mean values $S_1, \ldots, S_{17}$, the estimates of which should be found; the total mean

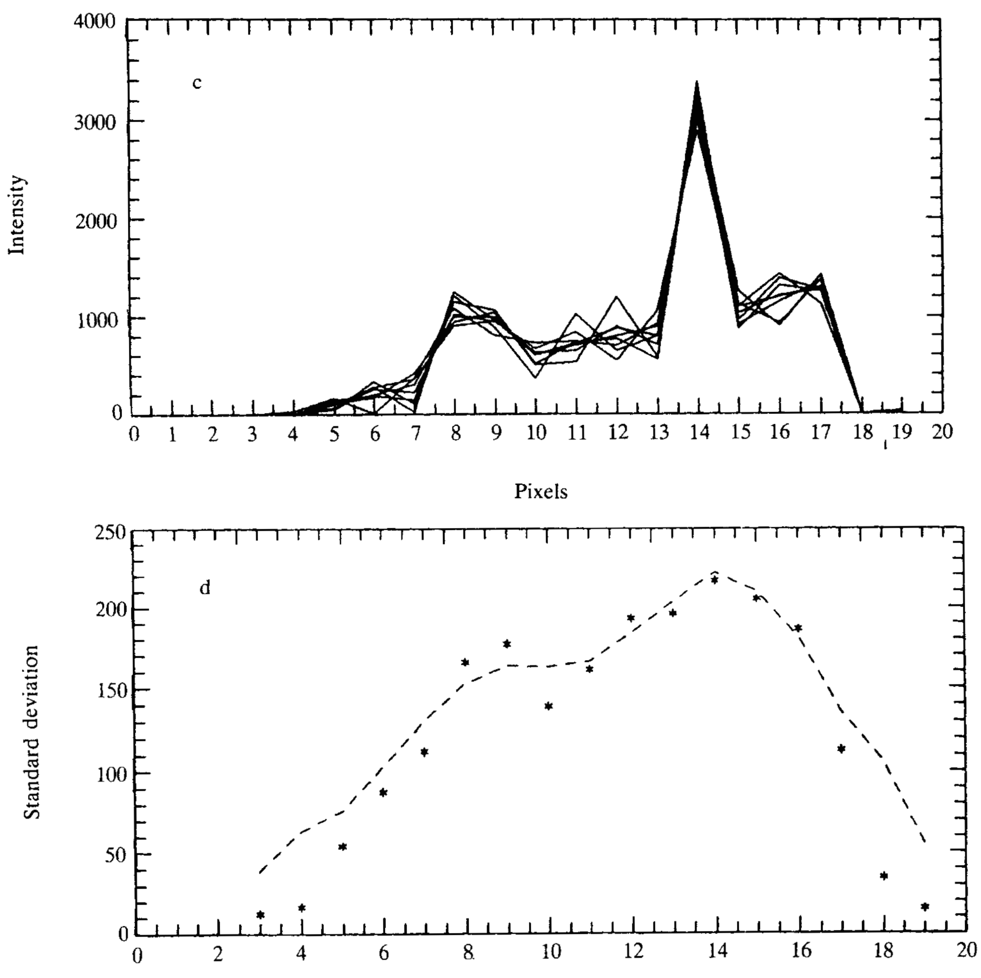

Figure 2(a) Object TEST-1 (full line), mean values for 100 MLIR-estimations of the object (asterisks), and standard deviations of these estimates (dotted lines). (b) Examples of smoothed and noised images of TEST-1. (c) Examples of restored with MLIR images of TEST-1. (d) Calculated standard deviations of MLIR-estimates according to 100 simulations (asterisks), and Rao–Cramer lower boundary at first approximation (dotted line).

brightness of the object is $L = 12{,}200$ events. 100 independent simulations of the object were performed, and each of them were further randomly smoothed and noised by Poisson external noise, with mean total intensity $B = 2100$. Each of the observed images were restored with the aid of MLIR, that was described at MLIR-1 and MLIR-5. The Poisson nature of the object and external noise, as well as the PSF form and mean level of noise were available as *a priori* information.

One can expect that the restoration of different random images of the same initial object lead to different object estimations. This feature is inherent to all methods of image restoration. A "good" method should give the *unbiased* estimate of the object, that is, its mean value from many restorations must be equal to the true object, and the *efficient* estimate, that is, variance of estimate, must be as small as possible (see for exact definitions the next section).

The bias of the mean from the 100 MLIR-estimates of TEST-1 (asterisks on Figure 2a) is considerably less than the standard deviation of these estimates (dotted lines). In other words, the bias is negligible in the considered case. The quality of any estimate depends mainly on its scattering near the true object. What can we say about this characteristic for MLIR-estimates of TEST-1? The observed scattering, and the least theoretically possible one, that follows from Rao–Cramer (RC) theorem (Section 5), are compared in Figure 2d. It should be noted that the *RC* limit was only approximately calculated here, without knowledge of bias, so, the theoretical boundary for scarce pixels near the edges of the image is slightly incorrect. Taking into account this fact, we see that the precision of MLIR-estimates can not be exceeded for this example—the *error corridor* for restoration by MLIR is as narrow as possible.

The high efficiency of MLIR is not accidental, it is a known result in the statistics that *if the efficient estimate exists, the maximum likelihood estimate coincides with it.*

4. DEFINITION OF RESTORATION EFFICIENCY

Figure 3 represents the main features of the image restoration procedure. It is considered at the frame of statistical parameters estimation approach.

Due to both radiational and external noise, any object, say O' or O'', from the *object space,* has inevitably random counterparts at the *image space.* If we define a stochastic model of image formation, it is only possible to calculate the probability distributions P' and P'' to observe images that were caused by the objects O' and O'' correspondingly. Similarly, the *true object* corresponds at the *image space* to the probability density P_t, and somewhere under this distribution is the really *observed image.*

Assume that we use some method to reach an *estimate of true object* E_t. Of course, it is very improbable to strictly get the *true object*; as a rule we will obtain an estimate at the *distance* $\rho > 0$ from it. As long as the *observed image* is random, so is ρ. From many "imaging and restoration" procedures of the same *true object,* we will obtain an ensemble of object estimates with different values of ρ. This random variable has some probability density $f_\rho(x)$, and can be

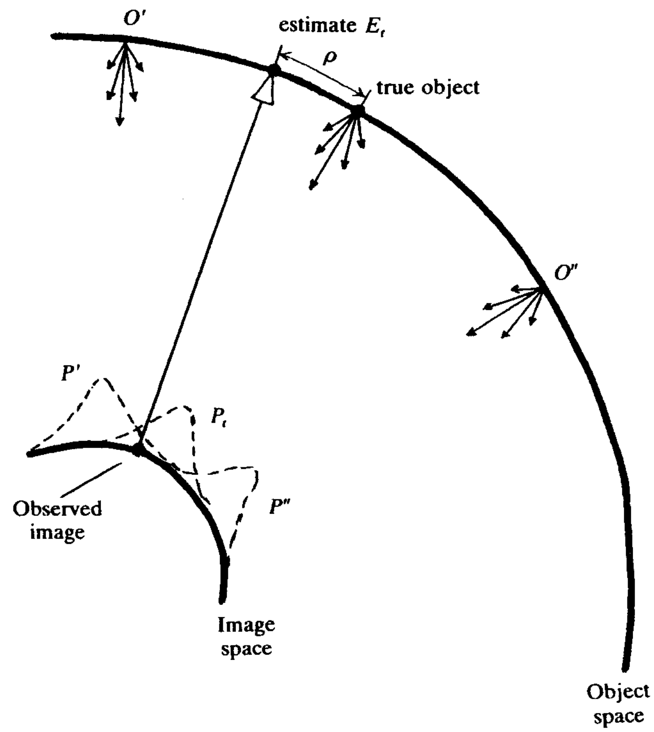

Figure 3 The scheme of quality evaluation of a random object estimate.

characterized by moments: mean value $\delta = \mathbf{E}(\rho)$, variance $\mathbf{D}(\rho) = \mathbf{E}[\rho - \mathbf{E}(\rho)]^2$ and so on. It is natural to consider, as a best, the estimate that has minimal *scattering* Ω near the true object:

$$\Omega \equiv \mathbf{E}(\rho^2) = \mathbf{D}(\rho) + \delta^2. \tag{19}$$

It should be stressed that scattering depends both from variance $\mathbf{D}(\rho)$ and *bias* δ.

The choice of quadratic measure of a distance is not only convenient, but is *sole* for any approximately normal sample distribution (Kramer, 1946, ch. 32).

Up to now, we have considered for the sake of simplicity, the one-dimensional case; in reality, both the object $S = (S_1, \ldots, S_n)$ and its estimate $S^* = (S_1^*, \ldots, S_n^*)$ have usually $n \gg 1$ dimensions, and the scattering depends on the direction in the object space. The foundation of general theory of parameters estimation was given by Ronald Fisher (see numerous references in Cramer, 1946, especially significant is the 1921 paper by Fisher). Let us turn to this general theory, following partly the handbooks of Cramer (1946), Kendall and Stuart (1969), Cox and Hinkley (1974), Borovkov (1984) and their cited literature.

In one-dimensional case scattering of the estimate S^* near the true value S is:

$$\Omega(S) = \mathbf{E}(S^* - S)^2 = \mathbf{D}(S^*) + \delta^2. \tag{20}$$

It seems, that optimal is an estimate, for which Ω is not larger than any other estimate scattering at *arbitrary* possible value of S (for all possible objects). A simple example $S^* = \text{const.}$ shows the meaningless of this definition (the broken clock once a day is more correct than any other one!). It is quite reasonable to narrow the class of possible estimates by considering, for example, only estimates with a given bias δ. We name *efficient* the one- or multidimensional estimate, for which scattering is minimal at a given subclass.

Consider the multidimensional images. For any model of image formation there is a probability density $f(N_1, \ldots, N_m; S_1, \ldots, S_n) \equiv f(N, S)$ to observe blurred and noised image N for any given object S. We denote by $\zeta \equiv (\zeta_1, \ldots, \zeta_m)$ the random image-vector, and by N its possible value. The searched estimates S^* are functions of ζ, and to find its moments we should use the distribution $f(N, S)$.

Statistics usually consider a *few* independent simulations of one- or multidimensional random variable, and parameter estimations are searched on the basis of the whole of this sample. But in the image restoration problem we usually have a *sole* image of a multidimensional object. Therefore, we consider here only this latter case, although the generalization is trivial.

Let us choose a direction γ in the n-dimensional object space. The scattering of estimate S^* in direction γ is:

$$\Omega_\gamma = \mathbf{E}\left[\sum_1^n (S_k^* - S_k) \cdot \gamma_k\right]^2 \geq 0, \tag{21}$$

and for the *efficient* estimate, the scattering is minimal *in all directions*. If we introduce the *scattering matrix*:

$$W = \|w_{ik}\| \tag{22}$$

with elements:

$$w_{ik} = \mathbf{E}[(S_i^* - S_i) \cdot (S_k^* - S_k)], \tag{23}$$

we can say that for efficient estimate the quantity:

$$\Omega_\gamma = \sum_{i,k=1}^n w_{ik} \cdot \gamma_i \gamma_k \tag{24}$$

is minimal for arbitrary vector γ.

In matrix form, the definitions (22) and (23), may be written as:

$$W = \mathbf{E}[(S^* - S)^T \cdot (S^* - S)], \tag{25}$$

where T denotes transposition. Introducing the estimate bias:

$$\delta(S) = \mathbf{E}(S^*) - S, \tag{26}$$

it is convenient to write the scattering matrix in a form:

$$W = \sigma^2 + \delta^T(S) \cdot \delta(S), \tag{27}$$

where:

$$\sigma^2 = \mathbf{E}[[S^* - \mathbf{E}(S^*)]^T \cdot [S^* - \mathbf{E}(S^*)]] \tag{28}$$

—the *variance matrix*. As one can see, expression (27) is quite similar to the one-dimensional formula (19).

It is easy to imagine all of the above quantities as appropriate descriptions (ellipsoid of scattering and so on) of points (restored images) located in n-dimensional object space near the true object.

5. RAO–CRAMER INEQUALITY

It seems quite evident that it is impossible to find an estimate of unknown parameters with an infinitely small scattering. Indeed, it was proved by R. Fisher, and other authors after him, that the scattering of any estimate has a lower limit (see references in the previous section). This statement is now usually known as the Rao–Cramer (RC) theorem. Perhaps, the first proposition to use the RC theorem in calculating errors of image restoration, was made by Slump and Ferwerda (1986); see also review by Adorf (1990). On the image restoration language, the multidimensional RC theorem gives the method-independent the most narrow error corridor after restoration. There is no need to explain the significance of this result.

The one-dimensional version of RC theorem has the following formulation: *under some regularity conditions R for any estimate S^* of unknown parameter S with bias $\delta(S)$, and limited variance the lower limit of scattering exists, that is given by inequality*:

$$\Omega(S) \geq \frac{[1 + \delta'(S)]^2}{I(S)} + \delta^2(S). \tag{29}$$

Here:

$$I(S) = \mathbf{E}\left[\frac{\partial}{\partial S} \ln f(\zeta, S)\right]^2 = -\mathbf{E}\left[\frac{\partial^2}{\partial S^2} \ln f(\zeta, S)\right] \tag{30}$$

—*Fisher information, $f(\zeta, S)$—sample probability density with change N to random variable ζ and the stroke means derivative with respect to S. The condition R assumes continuous differentiability of $f^{1/2}(N, S)$ with respect to S and also existence, non-negativity and continuity of $I(S)$.*

Consider a simple example. Assume that random variable ζ has Poisson distribution $f(N, S) = \exp(-S) \cdot S^N/N!$, $N = 0, 1, \ldots$ with unknown parameter S. We have: $\partial/\partial S \ln f(\zeta, S) = \zeta/S - 1$, $\partial^2/\partial S^2 \ln f(\zeta, S) = -\zeta/S^2$, and in view of $\mathbf{E}\zeta = S$ we obtain: $I(S) = S^{-1}$. The RC inequality for any method of estimation may be written as $\Omega(S) \geq [1 + \delta'(S)]^2 \cdot S + \delta^2(S)$. Particularly, the MLIR-estimate $\hat{S} = \zeta$ has zero bias, and for this estimate, RC inequality becomes $\Omega(S) \geq S$. On the other hand, the direct calculation gives $\Omega = \mathbf{D}(\hat{S}) = S$, so the scattering of MLIR-estimate reaches the RC boundary.

Most useful in practice, is the multidimensional RC theorem: *under regularity conditions R the scattering matrix* (25) *of any estimate S^* satisfies inequality*:

$$W \geq A, \tag{31}$$

where, by definition, matrix A is:

$$A \equiv [U + \Delta(S)] \cdot I^{-1}(S) \cdot [U + \Delta(S)]^T + \delta^T(S) \cdot \delta(S), \tag{32}$$

U is unit matrix, $\Delta(S) = \|\partial\delta_i(S)/\partial S_k\|$ *and* $I^{-1}(S)$ *is the inverse matrix to Fisher information matrix with elements*:

$$I_{ik}(S) = \mathbf{E}\left[\frac{\partial}{\partial S_i}\ln f(\zeta, S) \cdot \frac{\partial}{\partial S_k}\ln f(\zeta, S)\right]. \tag{33}$$

The R condition assumes continuous differentiability of $f^{1/2}(N, S)$ *with respect to* S_k*, continuity of Fisher matrix and difference of its determinant from zero. For unbiased estimates, evidently,*

$$W(S) \geq I^{-1}(S). \tag{34}$$

It is now necessary to give some explanations concerning the meaning of matrix inequalities. The condition $W \geq A$ equivalent to the non-negative deficiency of matrix $W - A$, that is, to usual inequality:

$$\sum_{i,k=1}^{n} w_{ik} \cdot \gamma_i \gamma_k \geq \sum_{i,k=1}^{n} a_{ik} \cdot \gamma_i \gamma_k \tag{35}$$

for any $\gamma_1, \ldots, \gamma_n$. In particular, substituting here γ_i equal to Kronecker symbol δ_{ip}, we obtain for diagonal elements:

$$w_{pp} \geq a_{pp}, \qquad p = 1, 2, \ldots, n. \tag{36}$$

In view of (20) and (23), let us define the *scattering of individual component of multidimensional parameter* by expression:

$$\Omega_k(S) = w_{kk} = \mathbf{E}(S_k^* - S_k)^2]. \tag{37}$$

The inequality (36) may be written at the final form:

$$\Omega_k(S) \geq a_{kk}, \quad k = 1, \ldots, n, \tag{38}$$

where a_{kk} are diagonal elements of matrix (32). For unbiased estimate:

$$\Omega_k(S) \geq I^{kk}(S), \tag{39}$$

where I^{ik} are elements of matrix I^{-1}, that is inverse to the Fisher information matrix (33).

In the frame of image restoration, these results mean *that the scattering of intensities of restored image relatively true one can not be less than some boundary values, that are defined by equations* (32), (33), (38) *and* (39).

Notice that (39) is a stronger inequality than simple generalization of the one-dimensional inequality $\Omega \geq 1/I_{kk}(S)$. Later we consider the meaning of the *RC* inequality for non-diagonal elements. Here we attract attention to the accuracy of restoration.

The RC theorem proves *existence,* and gives *value* of the accuracy limit, but it says nothing about very important practical questions: *when* is it possible to achieve the limit? and, *how* can it be done? The comprehensive answer to the first question is given by the following theorem (Darmois, 1935; for other references see Kendall and Stuart, 1969): *for estimate S^* of the vector parameter S, the equality in* (31) *can be achieved when, and only when, the distribution density $f(N, S)$ belongs to an exponential family, that is*;

$$\ln f(N, S) = \sum_{1}^{n} S_i^*(N) \cdot \varphi_i(S) + \psi(S) + \chi(N), \tag{40}$$

where scalar functions ψ and χ are arbitrary, and vector $\varphi(S)$ has derivative matrix:

$$\|\partial \varphi_i(S)/\partial S_k\| = [(U + \Delta(S))^{-1}]^T \cdot I(S). \tag{41}$$

For unbiased estimates, the right hand of (41) *is equal simply to Fisher information matrix $I(S)$. At last, in the one-dimensional case the* (40) *and* (41) *become*:

$$\begin{cases} \ln f(N, S) = S^*(N) \cdot \varphi(S) + \psi(S) + \chi(N), \\ \varphi'(S) = \dfrac{I(S)}{1 + \delta'(S)|}. \end{cases} \tag{42}$$

We discuss the answer to the second question, at the next section.

6. MAXIMUM LIKELIHOOD ESTIMATES

In view of what was said in a previous section, the importance of the following theorem is evident: *if the efficient estimate exists, and the regularity conditions R are valid, this estimate coincides with the maximum likelihood estimate $\hat{S}(N)$, that is, with the function that maximize density $f(N, S)$ under the given observed image N*:

$$f(N, S)|_{S=\hat{S}(N)} = \max. \tag{43}$$

In other words, the *maximum likelihood* estimate provides the maximal probability to the observed image N, under available information about the object, *PSF* and noise (see Fig. 3).

The maximum likelihood approach is the most promising way to reach the most accurate restoration. However, for this we must know the probability density $f(N, S)$, that is, to make concrete the imaging system.

Note that the generally accepted equations for calculating maximum likelihood estimate:

$$\frac{\partial}{\partial S_k} f(N, S)|_{S=\hat{S}(N)} = 0, \qquad k = 1, \ldots, n \tag{44}$$

are equivalent to (43), only if all maxima are situated *inside* of the object space $S_k \geq 0$. This is not the case for the faint parts of the image, therefore, the strict principle (43) should be used.

Consider the MLIR-estimates for models discussed at Section 2.

6.1. *Model A*

As far as the total number of observed events is equal to the number of light events L, the estimate of L is evident. Maximization of (4) under constraints (2) can be easily performed with the aid of Lagrange method, and we obtain the optimal MLIR-estimate:

$$\hat{s}_k = \nu_k / L, \quad \text{or} \quad \hat{S}_k = \nu_k. \tag{45}$$

This is the known result: the best estimate of theoretical distribution coincides with the sampling distribution.

The distribution density (4) belongs to the family (40), therefore, the efficient estimate exists. One can verify that this efficient estimate is coincident with (45).

6.2. *Model B*

The maximum likelihood approach for this model was proposed by Lucy (1974). As far as the existence of external noise is inevitable, and the presence of even such small noise drastically change the final picture, this model has mainly theoretical significance. It is worth mentioning that, as used by Richardson (1972) and by Lucy (1974), the ambiguous Bayesian approach is not necessary at all. Nevertheless, the Lucy (1974) paper was the first successive attempt to use the powerful maximum likelihood approach in the image restoration problem.

Some numerical methods of maximizing of (8) had been considered by Tarasko (1969), Richardson (1972), Lucy (1974) and Kosarev *et al.* (1983). Another approach to the problems of a similar type was discussed in MLIR-3.

Under $m > n$ density (8) cannot be written in the form (40). This means that the efficient estimate does not exist in this case. In reality the MLIR-estimate is very close to the efficient one; we further discuss this item for the Poisson model.

Note that the *inverse solution,* that is, the solution of the system of equations $p_j = \nu_j / L$, is not a positively defined, and for that reason it doesn't coincide with the desired optimal MLIR solution. There is also no guarantee that the inverse solution satisfies the second condition (1).

6.3. *Model C*

This model is quite simple, but at the same time it takes into account all the main features of real images. Algorithm for maximization (12) under constraints (1) was considered in MLIR-3; this paper also includes the discussion of test cases (the one at Figure 1 shows MLIR abilities just for Model C). The efficient estimate doesn't exist again.

6.4. *Poisson Model*

Consider the possibility to find the most theoretically accurate (or *efficient*) estimate for incoherent sources. It is evident that for $m > n$ distribution (16) has not the general form (40), when such an estimate can be found. Thus, we obtain the important conclusion: *for incoherent sources, there does not exist a method of image restoration that has Rao–Cramer limiting accuracy of restoration.*

At the same time, it is quite possible to find such a method that under some conditions has practically equal efficiency to the limiting one. In view of the results of Section 6, first of all, we can expect this from the maximum likelihood approach. Indeed, the strict solutions for the simplest cases, and numerical simulations for more complex ones, shows that MLIR has limiting efficiency when light intensity becomes approximately equal to the mean level of external noise (see Section 3 and Figure 2).

According to MLIR, we should obtain the estimate $\hat{S}$, that maximizes (16) under constraints:

$$S_k \geq 0, \qquad k = 1, 2, \ldots, n. \tag{46}$$

To calculate the RC boundary from (32), we must know Fisher information matrix (33) and bias (26). Information matrix for incoherent radiation can be easily found from (33), (15) and (16):

$$I_{ik}(S) = \sum_{j=1}^{m} \frac{h_{ji}h_{jk}}{\lambda_j(S)}, \qquad i,k = 1, 2, \ldots, n. \tag{47}$$

This expression is valid for an arbitrary method of estimating. From the rigorous point of view, we should know the object S beforehand to calculate $I(S)$, but beginning from the image intensity of order of a few dozen of events, per pixel, we can replace the *mean* intensities in (47) by *real observed* ones. The result is:

$$I_{ik} \cong \sum_{j=1}^{m} \frac{h_{ji}h_{jk}}{N_j}, \qquad i,k = 1, \ldots, n, \tag{48}$$

and it is possible to approximately calculate the limiting accuracy of restoration only on the basis of observational data.

As to the bias of an estimate, it is evident that $\delta(S)$ strongly depend on the method of restoration, and in many situations it is difficult to find strict results. In particular, we failed to find the general expression for the bias of MLIR yet. It seems very probable that the *bias of MLIR quickly tends to zero as light intensity exceeds the mean level of external noise.* If we take $\delta(S) = 0$, we obtain a slightly incorrect RC boundary, and for MLIR this factor has significance only in a very scarce region of image. This phenomenon one can see on the example of restoration of TEST-1 (Figure 2d).

It is not usually easy to see the general features of a process if we confine ourselves only to numerical simulations. Analytical examples are needed, even if they are not of practical significance. Here give a very simple example of incoherent source, only one pixel of the object and two pixels in the degraded image ($n = 1$, $m = 2$). In this case PSF is defined by values of h_{11} and $h_{21} = 1 - h_{11}$. One can immediately obtain the general MLIR-solution, but even at such restricted formulation, it has rather a complicated form. Thus, we further assume that mean values of external noise are an obeyed relation: $b_2 = [(1 - h_{11})/h_{11}] \cdot b_1$. The MLIR-estimate $\hat{S}_1$ of the real object intensity is:

$$\hat{S}_1 = \begin{cases} 0, & \text{if} \quad \omega \leq r, \\ \omega - r, & \text{if} \quad \omega > r, \end{cases} \tag{49}$$

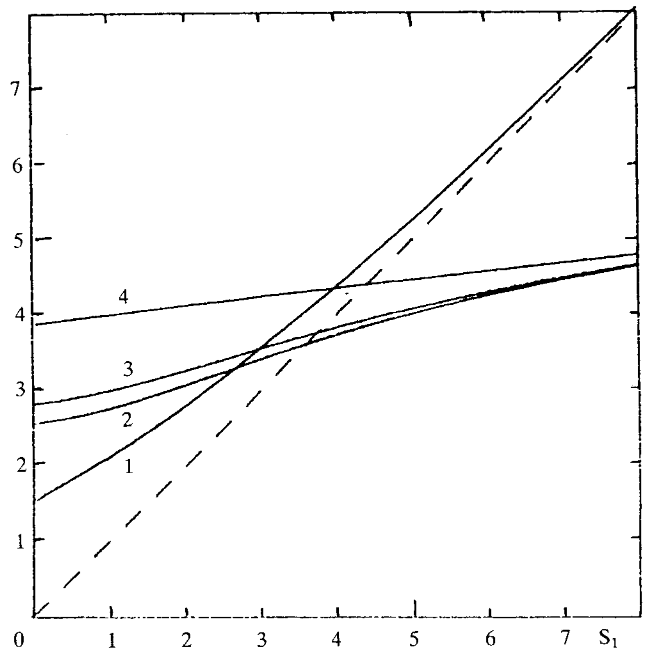

Figure 4 Main characteristics of MLIR-solution for analytical example with $r = 15$: the estimate of true intensity (1); square root from scattering accordingly to Rao–Cramer theorem with true (2) and zero (4) bias; the similar quantity for MLIR-estimate (3).

where $r = b_1/h_{11}$ and ω is Poisson random variable with mean value $\lambda = S_1 + r$. From this relation, and equations (20), (26), (29) and (30), one can obtain exact expressions for the bias and scattering of MLIR-estimate (49) and compare the last quantity with lower RC boundary $\Omega_{RC}S_1$). The full description is given in MLIR-6; we show here the results for the case $r = 15$ (Fig. 4).

The significative inferences from this example are following. (1) For zero external noise, the bias of MLIR-estimate is identically equal to zero. The possibility should be investigated that the bias of MLIR-estimates is completely caused by external noise. (2) The bias under discussion monotonously decreases as intensity S_1 increases, and relative value of bias is negligible when light intensity exceeds the mean level of external noise. (3) The scattering of MLIR-estimate slightly exceeds the RC boundary at small intensities, but very quickly becomes equal to it for $S_1 \geq r$. (4) If the estimate were unbiased, we would have accordingly to (29): $\Omega_{RC}(S_1) = [I(S_1)]^{-1} = S_1 + r$. By using this relation, we exaggerate the theoretical lower limit at small intensities, but for $S_1 \geq r$ the error is negligibly small (compare with Figure 2d).

7. CONTROVERSY WITH MAXIMUM ENTROPY APPROACH

The existence of a large number of image restoration methods raises the question concerning the place of the proposed statistical parameter estimation (SPE) approach, and its comparative power.

7.1. *Linear Methods*

As to linear methods of restoration, it has been known for a long time that they do not allow to account, in a natural way positivity, of the original image and the random character of the image formation process. For these reasons the linear methods cannot provide, in principle, the *superresolution* phenomenon. It means that such fine details of the object are restored, that it seems to be impossible, in view of the smoothing by PSF. The origin of this (inherent only to non-linear methods) phenomenon was completely ascertained by Schelkunoff (1943), Toraldo di Francia (1955), Wolter (1961), Harris (1964) and Frieden (1967). The practice also shows that "linear methods applied to Poisson-noise data immediately provide an error estimate" (Adorf, 1990b).

7.2. *Maximum Entropy Method (MEM)*

Much more attention now attracts the general maximum-entropy way of estimation on the basis of partial knowledge that was proposed by Jaynes (1957a,b). In conformity to image restoration MEM was developed by Frieden (1972), Skilling and Bryan (1984) and in many other investigations (see reviews by Frieden, 1979, and by Narayan and Nityananda, 1986). There are at least four definitions of "entropy" that have to be maximized and a lot of relevant numerical algorithms.

The main idea of MEM consists in removing the instability that is inherent to the solutions of inverse problems by choice of "maximal free", under given constraints solution. It was at first believed that the maximum entropy requirement (or maximal *degeneracy* in other words) of the solution automatically provides its maximal probability. Frieden (1985), in the example with rolls of a dice, has shown that such a view is erroneous. This fact was hard blow on validity of MEM.

Our general objection against MEM can be formulated in the following way: one cannot invent any *a priori* principle to solve a problem; it is only allowed to describe a real process as completely as possible, and to try to develop the most appropriate model. Any form of "entropy" is a human notion, and it can *follow* from correspondent consideration, but cannot *anticipate* it. By proposing the entropy principle, Jaynes introduced subjectivity into the whole subsequent consideration of the problems. This subjectivity reveals itself in various aspects of different versions of MEM.

To give a demonstration of such subjectivity, let us consider the strict application of Jaynes's principle in the frame of described in Section 2 Model C of image formation (similar inferences follow from discussion of the other models, in particular Poisson model for incoherent sources, see MLIR-5).

It will be recalled that we consider as an *object* some unknown normalized to 1 brightness distribution of events $s = (s_1, \ldots, s_n)$ together with its intensity L. Perfect (ideal) imaging system will produce *random* pattern $v = (v_1, \ldots, v_n)$ of events with mean $\mathbf{E}(v_k) = L \cdot s_k \equiv S_k$. Any real such system will give another random pattern $N = (N_1, \ldots, N_m)$, $m \geq n$, where N accounts the smoothing and external noise. In particular, for linear model, we have:

$$E(N_j) = \sum_{k=1}^{n} h_{jk} S_k + b_j. \tag{50}$$

The SPE approach implies the searching for the best, in some definite sense, estimates $S^* \equiv (S_k^*)$ of unknown parameters $S \equiv (S_k)$. These estimates are considered as the restored image.

The way of estimation in MEM is quite different. First of all, the equation (50) for *mean* values is replaced by equation:

$$N_j = \sum_{k=1}^{n} h_{jk} \nu_k + \xi_j \tag{51}$$

for *random* values: observed image N, object pattern ν and external noise ξ. Just some "best" estimate ν^* is considered here as a restored object. Secondly, in view of multiplicity and instability of the solutions of (51) the MEM chooses such solution ν^{E} of these equations that has maximal statistical weight, that is, it can be sampled by maximal number of ways. The last quantity equal to:

$$\Gamma = \frac{L!}{\nu_1! \cdot \ldots \cdot \nu_n!}, \tag{52}$$

so, the maximization of Γ is equivalent to the minimization of the production in the denominator of (52). In view of Stirling's approximation to factorials the last requirement becomes:

$$-\sum_{k=1}^{n} (\nu_k/L) \ln (\nu_k/L) = \max. \tag{53}$$

Thus, the searched estimate ν^{E} of the object should maximize the value of Shannon (1948), entropy (53), under constraints (51). Since the noise pattern ξ in (51) is unknown, the strict meaning of these constraints are necessary to specify.

The shortcomings of MEM approach are evident.

i) Equations (51) are not true, and the above mentioned replacing can sometimes be feasible, but sometimes it can introduce an appreciable error.

In SPE there are no outside equations; the strict relations (50) are *consequences* of the model description.

ii) As long as the concrete noise pattern ξ in (51) is unknown, there are infinite methods of accounting these equations; a number of such methods are really used in practice. This circumstance introduce subjectivity to the procedure of restoration.

iii) The way of introducing (51) as *constraints* is far from a natural way in SPE, where inner noise is an *inherent* feature of the model.

iv) Requirements of maximal probability, or degeneracy of the object pattern ν, are human notions that do not follow from the essence of the problem. These requirements do not coincide with the maximum likelihood principle.

Of course, this latter is also a subjective attempt to search for the solution, but in SPE maximum likelihood appears only at the final stage of the investigation as *one of the many possible estimation methods, perhaps the most promising among them.* It is possible (in SPE) to calculate method-independent limiting accuracy of restoration, after that one can use any method that allow us to achieve the Rao–Cramer limit.

v) The most degenerate solution differs from the most probable one (Frieden, 1983, 1985). It is easy to see this fact for the considered here model C. Indeed, the probability of the occurrence of pattern ν under constraints (2) is given by (3), and this latter function by no means is not maximal at $v = v^E$, when Γ is maximal. The most probable and most degenerate solutions coincide only for a uniformly gray object (s_k = const), but then "what is the purpose of imaging it?" (Frieden, 1983). The requirement of maximal degeneracy contradicts the *a priori* information and data. For instance, we may take as an alternative to the uniform distribution, the consequence of the large numbers law: $s_k \cong v_k/L$. Substituting these equations to (3) and using Stirling approximation we obtain condition of maximal degeneracy in the form:

$$-\sum_{k=1}^{n} \ln (v_k/L) = \max, \tag{54}$$

that is, Berg (1967) "entropy" has to be maximal.

It is difficult to avoid the conclusion that requirements to maximize Shannon, Berg or any other form of "entropy", are connected with available *a priori* information. This fact was stressed by Frieden (1983), who attempted to introduce this information in an explicit form. He suggested to *randomize* the distribution (3) with respect to (s_k):

$$f(v) = \frac{L!}{v_1! \cdot \ldots \cdot v_n!} \int \ldots \int s_1^{v_1} \cdot \ldots \cdot s_n^{v_n} \cdot p(s_1, \ldots, s_n)\, ds_1 \ldots ds_n, \tag{55}$$

where probability density $p(s)$ should take into account the physical information concerning the occurrence rate of s. In particular, in a case $p(s) = \delta(s - z)$ we come to models some of which were considered above. This time, by maximization of (55), we obtain, instead of Shannon or Berg functionals, the condition that conform to *a priori* information.

This is quite correct, but it should be mentioned that a) equation (55) is not the most general manner to set (in our language) a model of image formation: we can, for instance, randomize L as well, to introduce a parameters in $p(s)$ and so on; b) equation (55) is a way to make concrete the image formation model, but it has no bearing on the method of estimation of the object (maximization of $f(v)$ is additional requirement, as before, that is not the same as the maximum likelihood principle).

In conclusion, the shortcomings i–iv of MEM are inherent to the essence of this approach, and it is not possible to predict their consequences in any real situation.

8. SOME EXAMPLES

It should be confessed that since the proposal of the considered approach (Terebizh, 1990), we were working mainly with test cases, but not with the real data. The reasons for that are clear. First of all, only for the test case can we know the object completely and have an opportunity to perform an entire checking of the method. Secondly, to restore the real images we should prepare the corresponding two-dimensional software and have a powerful computer. In MLIR-1 to MLIR-6, one can find about a dozen one-dimensional test cases, so we give here a sole such example, but for two dimensions (Figure 5).

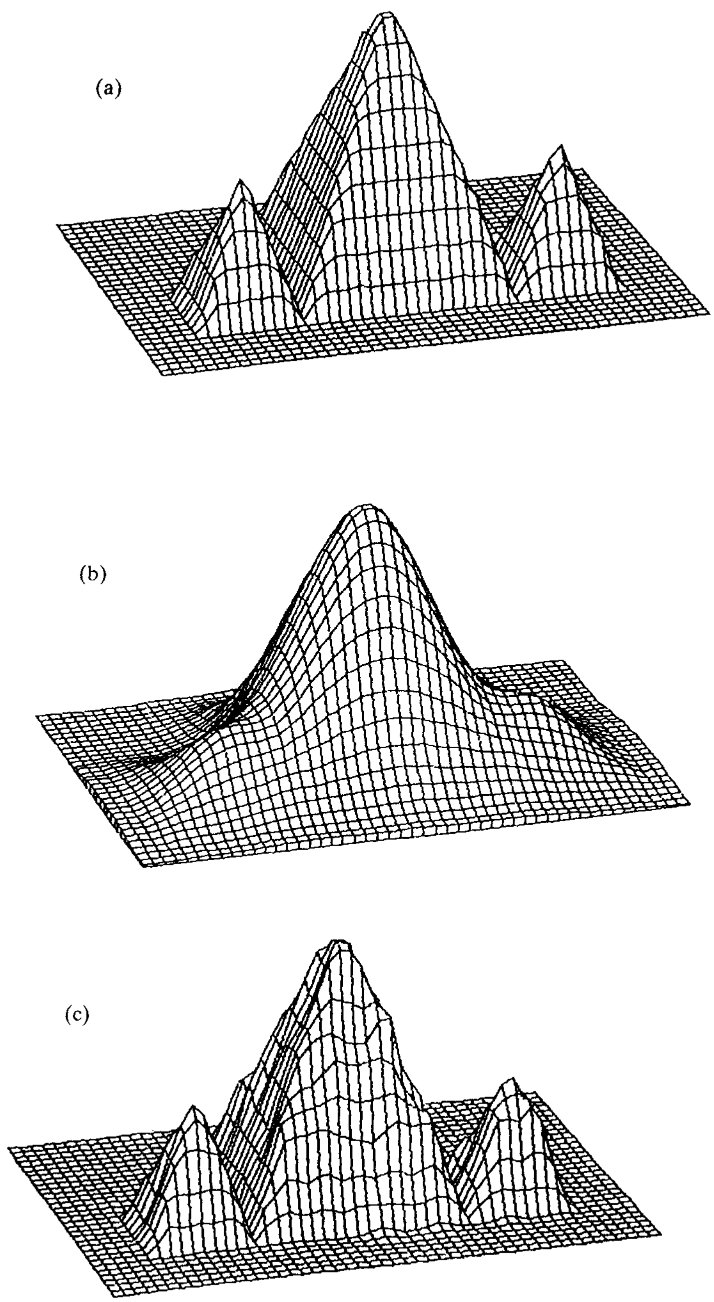

Figure 5 Test object (a), degraded due to smoothing and external noise image (b) and the MLIR-estimate with non-negativity as a sole *a priori* information (c).

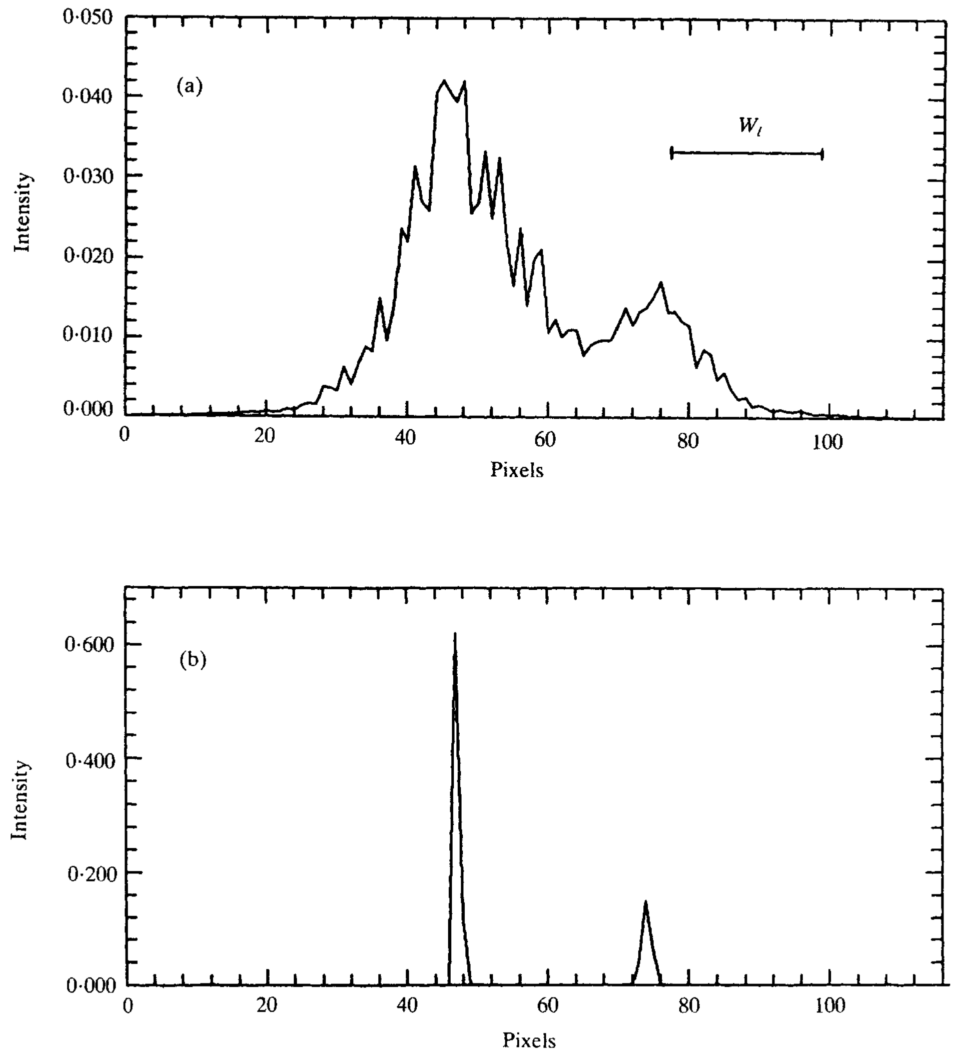

Figure 6 Photoelectric scan of double star π *Boo* with the components separation 5.6″, at very a bad seeing (a), and MLIR-estimate (b).

The second example deals with the real observations (Figure 6). The double star π *Boo* with components separation 5.6″ was observed at a very bad seeing with the strip photon counting photometer (Terebizh, 1981). The equivalent width W_l (or practically equal to it, the full width at half intensity) of a single star image was ≅4.5″, and even so wide a double star was partly overlapped. In the course of MLIR processing, the information about the double nature of the object was not used; only the non-negativity of an unknown object was assumed. The full width of the restored images at half intensity is approximately equal to 0.3″.

9. APPLICATIONS OF STATISTICAL APPROACH TO OTHER INVERSE PROBLEMS

The SPE approach has been discussed above in the context of optical images restoration. The characteristic feature of such problems is non-negativity of the object. Just this fact allows us to introduce the powerful probabilistic way of consideration. In the case of optical imaging this way corresponds to the physical nature of the problem, but it is easy to understand that SPE approach and MLIR can also be applied to many inverse problems, with the above mentioned feature and with the "events", as simply as an auxiliary model.

We consider here two examples, from which the first example concerns a concrete problem and the second one outlines an approach to the class of inverse problems.

9.1. *Number of Flare Stars in Stellar Aggregates*

It was proposed by Ambartsumian (1968) that the full number of flare stars in a cluster can be estimated from the following arguments. Assume at first, that there are only two types of stars in a cluster: a) stable stars and b) flare stars, in a number of S_1 with the same frequency of flares f_1 hour^{-1}. It is very probable that under random moments of observations, and a comparatively large lag between flares, the temporal sequence of the *observed* flares of any flare star is close to the stationary Poisson process. Consequently, the mean number of flare stars that experienced k flares during time T is equal to:

$$M_k(T) = S_1 \exp(-f_1 T) \cdot (f_1 T)^k / k!, \qquad k = 0, 1, 2, \ldots . \tag{56}$$

It follows from these equations:

$$M_0 = \frac{M_1^2}{2M_2}. \tag{57}$$

If we equate the *mean* numbers M_1 and M_2 to the *observed* ones, we will obtain a possibility to estimate the number of flare stars that have not had flares during the considered exposition time, and then the full number of flare stars in the cluster S_1. This manner of estimation is known as the *moment method.*

As it turned out with time, the suggestion of a sole frequency of flares is too rough. So, Ambartsumian later proposed to introduce a frequency spectrum of flare stars (see for discussion and references Ambartsumian, 1988). If $g(f)\,df$ is the number of flare stars with frequencies of flares in the interval $[f, f + df]$, we can write instead of (56):

$$M_k(T) = \int_0^\infty \exp(-fT) \frac{(fT)^k}{k!} \cdot g(f)\,df. \tag{58}$$

Again with replacing M_k to the observed numbers we can try to restore the distribution $g(f)$. It was found that this inverse problem in its ordinary formulation is extremely difficult, and up to now the information about $g(f)$ is very uncertain.

We can know *all possible information on the basis of available observational data,* if we turn to the SPE formulation of the problem. Indeed, let us introduce the necessary discretization $S_k = \int_{\Delta f_k} g(f)\, df$, $k = 1, 2, \ldots$ and the Poisson "point spread function":

$$h_{jk}(T) = \exp(-f_k T)\frac{(f_k T)^j}{j!}, \qquad \begin{cases} j = 0, 1, 2, \ldots, \\ k = 1, 2, \ldots. \end{cases} \tag{59}$$

It then becomes possible to interpret as "smoothing", the process of flares occurrence, in which the "object" (S_k) is randomly transformed to the distribution (F_j) of stars with j flares during time T. At last, the background flare stars add the Poisson "noise" with some mean "intensity" (b_j). In this formulation of the problem we come to the model C, and can use not only the MLIR-solution, but calculate the error corridor according to (38) or (39). The concrete data will be considered elsewhere.

9.2. *Deterministic Smoothing*

A number of inverse problems that have no deal with images, in a strict meaning, suggest a non-random smoothing process. This situation constitutes the particular case of SPE, so, the general results are valid here as well. Let us consider, for instance, the *linear* deterministic smoothing and random external noise:

$$N(x) = \int h(x, x')S(x')\, dx' + \xi(x). \tag{60}$$

For a non-random object, S, the stochastic properties of the "image" N, that is, $f(N, S)$, are completely defined by external noise distribution.

In general, any inverse problem that permits to formulate a stochastic model of an investigated process with some density distribution $\Pr(N, N + dN \mid S) = f(N, S)\, dN$ can be considered in the frame of stochastic parameter estimation approach that was described in the previous sections.

10. CONCLUDING REMARKS

Perhaps, it is worthwhile to mention, once more, that the approach discussed here considers, as *independent* parts, the description of the image forming system, the statistical parameter estimation, and the concrete methods of restoration. It seems probable that only this way, allows us to avoid subjectivity—the most dangerous factor for power and reliability of restoration methods. As to the choice of the restoration method, the MLIR is preferable, in view of the fact that it provides the limiting efficiency and resolution power for a wide circle of conditions (only a rogue can cure *any* illness, but good physicians, in any given circumstance, do all that is possible).

The interest to the image restoration problem has been recently animated after the sad news that the Hubble Space Telescope's images strongly suffer from a spherical aberration. One can hope that intensive work in this field will provide a quick deepening of our understanding of relevant questions, so the Hubble Telescope, in a large measure, will return its power.

Acknowledgements

This work proceeds partly from the joint papers of V. V. Biryukov, O. K. Cherbunina and Yu. G. Cherbunin (MLIR-3,5,6). It is a pleasure to thank them for the discussions and calculations for this paper. I am grateful to professor B. Roy Frieden for valuable comments to the preliminary part concerning the maximum entropy method.

References

Adorf, H.-M. (1990a). Restoring HST images—an ST-ECF perspective, *ST-ECF Newsletter*, **14**, 8–12.

Adorf, H.-M. (1990b). Restoration of HST images and spectra, *ST-ECF Newsletter*, **14**, 15.

Ambartsumian, V. A. (1968). To the Statistics of Flare Objects. In: *Stars, Nebulae and Galaxies*, V. V. Sobolev, ed., Armenian Acad. Sci., Yerevan, 283–291.

Ambartsumian, V. A. (1988). Scientific Papers, v. 3, V. V. Sobolev, ed., Armenian Acad. Sci., Yerevan.

Borovkov, A. A. (1984). Mathematical Statistics, Nauka, Moscow.

Burg, J. P. (1967) Maximum Entropy Spectral Analysis, paper presented at 37th Ann. Soc. of Exploration Geoph. Meeting, Stanford Univ. Geoph. Dep., Oklahoma City.

Cox, D. R. (1955). Some statistical methods connected with series of events. *Journ. Roy. Statist. Soc.* **17B,** 129–164.

Cox, D. R. and Hinkley, D. V. (1974). *Theoretical Statistics,* Chapman and Hall, London.

Cramer, H. (1946). *Mathematical Methods of Statistics,* Princeton Univ. Press, New York.

Darmois, G. (1935). Sur les lois de probabilite a estimation exhaustive, *C.R. Acad. Sci.*, Paris, **200,** 1265.

Feller, W. (1966). *An Introduction to Probability Theory and Its Applicaitons, v. I,* Wiley, New York, London, Sydney.

Fisher, R. A. (1921). On the mathematical foundations of theoretical statistics, *Phil. Trans. Roy. Soc. London* **A 222,** 309.

Frieden, B. R. (1967). Band-Unlimited Reconstruction of Optical Objects and Spectra, *Journ. Opt. Soc. Amer.* **57,** 1013–1019.

Frieden, B. R. (1972). Restoring with maximum likelihood and maximum entropy, *Journ. Opt. Soc. Amer.* **62,** 511–518.

Frieden, B. R. (1979). Image enhancement and restoration. In *Picture Processing and Digital filtering,* T. S. Huang, ed., Springer, Berlin, New York, pp. 179–249.

Frieden, B. R. (1983). Unified Theory for Estimating Frequency-of-occurrence laws and optical objects, *Journ. Opt. Soc. Amer.* **73,** 927–938.

Frieden, B. R. (1985). Dice, entropy and likelihood, *Proc. IEEE,* **73,** 78–87.

Frieden, B. R. (1987). Maximum-probable restoration of photon-limited images, *Applied Optics* **26,** 1755–1764.

Harris, J. L. (1964). Diffraction and Resolving Power, *Journ. Opt. Soc. Amer.* **54,** 931–936.

Jain, A. K. (1989). *Fundamentals of Digital Image Processing,* Prentice-Hall, Englewood Cliffs.

Jaynes, E. T. (1957a). Information Theory and Statistical Mechanics, *Phys. Rev.* **106,** 620–630.

Jaynes, E. T. (1957b). Information Theory and Statistical Mechanics. II., *Phys. Rev.* **108,** 171–190.

Kendall, M. G. and Stuart, A. (1969). *The Advanced Theory of Statistics, v.* 2, Griffin, London.

Kosarev, E. L., Peskov, V. D. and Podolyak, E. R. (1983). *Journ. of Techn. Phys.*, **53,** 1101–1114.

Loudon, R. (1973). *The Quantum Theory of Light,* Clarendon Press, Oxford.

Lucy, L. (1974). An iterative technique for the rectification of observed distributions, *Astron. J.* **79,** 745–754.

Mandel, L. (1958). *Proc. Phys. Soc. (London),* **72,** 1037.

Mandel, L. (1959). *Proc. Phys. Soc. (London),* **74,** 233.

Mehta, C. L. (1970). Theory of Photoelectron Counting. In *Progress in Optics,* E. Wolf, ed., North-Holland, Amsterdam, London, **8,** pp. 373–440.

Narayan, R. and Nityananda, R. (1986). Maximum Entropy Image Restoration in Astronomy, *Ann. Rev. Astron. Astroph.* **24,** 127–170.

Richardson, W. H. (1972). Bayesian-based iterative method of image restoration, *Journ. Opt. Soc. Amer.* **62,** 55–59.

Schelkunoff, S. A. (1943). *Bell Syst. Tech. Journ.* **22,** 80.

Shannon, C. (1948). A mathematical theory of communications, *Bell. Syst. Techn. Journ.* **27,** 379–423, 623–656.

Skilling, J. and Bryan, R. K. (1984). Maximum Entropy Image Reconstruction: General Algorithm, *Monthly Not. R.A.S.* **211,** 111–124.

Slump, C. H. and Ferwerda, H. A. (1986). Statistical aspects of image handling in low-dose electron microscopy of biological material. In *Adv. Electronics and Electron Phys.* **66,** 201–308.

Tarasko, M. Z. (1969). *Prepr. FEI*-156, Obninsk.

Terebizh, V. Yu. (1981). Strip photometer, *Astron. Tsircular,* No. 1188, 3–6.

Terebizh, V. Yu. (1990). Maximum likelihood image restoration. I. Main equations, *Astrofizika* **32,** 327–339.

Terebizh, V. Yu. (1991). Time Series Analysis in Astrophysics, Nauka, Moscow.

Toraldo di Francia, G. (1955). Resolving Power and Information, *Journ. Opt. Soc. Amer.* **45,** 497–501.

Vasilenko, G. I. and Taratorin, A. M. (1986). Image restoration, Radio i svyaz, Moscow.

Wolter, H. (1961). On Basic Analogies and Principal Differences Between Optical and Electronic Information. In: *Progress in Optics,* E. Wolf, ed., v. I, 155–210.