Proceedings of the IASTED International Conference
Signal and Image Processing
October 27-31, 1998, Las Vegas, Nevada-USA

ULTIMATE RESOLUTION IN THE FRAMEWORK OF PARAMETER ESTIMATION

E. BETTENS, A.J. DEN DEKKER, D. VAN DYCK, J. SIJBERS
University of Antwerp (RUCA), Department of Physics
Groenenborgerlaan 171, B-2020 Antwerp, Belgium
Email: bettens@ruca.ua.ac.be

ABSTRACT

This paper describes two possible points of view on two-object resolution in the context of model-fitting theory. Thereby, it is stated that the precision and the accuracy with which the locations of the objects can be estimated will determine the attainable resolution.

The first approach describes the probability of resolution, that is the probability that the estimated locations will be distinct. The second leads to the maximally attainable precision. For both approaches, the special case of gaussian peaks is further investigated. It is shown that resolution is no longer possible for closely located peaks.

Keywords*- Resolution, Model Fitting, Statistical Precision, Probability of Resolution.*

1. Introduction

Resolution, expressing the ability to separate adjacent details, is widely used as a performance measure in the quality assessment of imaging systems. One of the most famous and widely used criteria is that of Rayleigh [1]. However, this criterion is no longer sufficient since it is only based on the limitations of the human visual system, and does not take into account, for example, the presence of noise. Nowadays, it is generally realized that resolution is ultimately limited by errors in the observations [2]. Therefore, in this paper, resolution is studied from the viewpoint of parameter estimation theory.

The studied experiment consists of counting events, for example, an electron hitting a pixel. The events are distributed by a probability density function, which describes the sum of two peak shaped objects. The locations of the objects under study are the parameters that have to be estimated from the experimental data. To which extent a reliable estimate of these parameters can be obtained will determine the attainable resolution.

In this work, it is shown that the Maximum Likelihood estimator (ML-estimator), which is known to fully exploit the knowledge of the probability density function of the observations, fails for objects relatively close together. It turns out that the ML-estimator is no longer accurate (i.e., a bias is introduced) for closely located objects: the estimates of the locations can coincide exactly. In this case we say that the objects are not resolvable. This kind of behavior of the ML-estimator can be explained with the aid of Catastrophe Theory [3]. It leads us to an expression for the probability of resolution, that is, the probability that the estimates of the objects' locations will not coincide. The developed theory is closely related to the work of Van den Bos and Den Dekker [4][5].

It will also be shown that even if an unbiased estimator would exist, this estimator would still fail to produce reliable estimates due to the poor attainable statistical precision. The attainable precision of an estimator is limited by the Cramér-Rao Lower Bound (*CRLB*). This bound defines the lower limit on the variance of an unbiased estimator. An easy to calculate expression for this lower bound will be derived that approximates the real expression sufficiently accurately.

The outline of this paper is as follows. In the next section, the used model is briefly described. In section 3, the behavior of the Maximum Likelihood estimator is studied. The limits on the attainable statistical precision are presented in section 4. In section 5, some results of simulated experiments are given.

2. Model

Consider an experiment that consists of counting events, for example, an electron hitting a detector. The events are distributed over a number of intervals, denoted as $\{x_i \, ; i=1,\dots M\}$, by a discrete density function. The observations are given by $\{n_i \, ; \, i=1,\dots M\}$, where n_i describes the number of counts or events in the interval x_i
The total number of counts is defined by N, *with* $N=\sum_i n_i$
The probability that an event occurs in the interval x_i will be denoted by $p(x_i;a_1,a_2)$, with:

$$p(x_i;a_1,a_2) = If(x_i - a_1) + (1-I)\,f(x_i - a_2) \qquad (1)$$

describing the sum of two peak shaped objects, centered around the locations a_1 and a_2. These locations are the parameters to be estimated. The factor I denotes the peak intensity ratio; its value is supposed to be known and lies between 0 and 1. The observations can be described as:

$$\underline{n}_i = Np(x_i;a_1,a_2) + \underline{\varepsilon}_i \qquad (2)$$

with $\underline{n}_i$ and $\underline{\varepsilon}_i$ stochastic variables; $\underline{\varepsilon}_i$ is the deviation from the expectation value $E[\underline{n}_i]$, which is given by:

$$E[n_i] = Np(x_i; a_1, a_2) \qquad (3)$$

These $E[\underline{n}_i]$ form the model of the observations: they describe the expected outcome of the experiment. Model fitting means that the parameters a_1 and a_2 are to be estimated such that the model fits the observations as good as possible in the sense of a suitable criterion of goodness of fit. If there would be no noise in the observations, and the model would be correct, the fit would be perfect and the parameters could be calculated exactly. However, this is never the case in a real experiment.

One of the most important estimators is the Maximum Likelihood estimator (ML-estimator). The likelihood function $L(y;a_1,a_2)$ is defined by the joint probability of the observations y, and this as a function of the parameters. The ML-estimate of a parameter is then found by maximizing L with respect to these parameters. One of the most interesting properties of the ML-estimator is that if there exists an unbiased estimator that attains the *CRLB*, this estimator is given by the ML-estimator [6].

Maximizing the likelihood function L is equivalent to maximizing $ln(L)$, as the logarithm is a monotonic increasing function. We have:

$$lnL = \sum_i n_i \ ln \ p_i \quad with \quad p_i \equiv p(x_{i;} a_1, a_2) \qquad (4)$$

which is called the loglikelihood function.

3. Probability of resolution

Simulation results show that beyond a certain degree of overlap of the object functions, the ML-estimates of the locations of the objects will coincide exactly for a considerable amount of the experiments. This remarkable result can be explained with the aid of *Catastrophe theory* [3]. A complete description of all the steps involved in the analysis of the problem is outside the scope of this paper. For this, we refer to [3] and [4]. The main results are sketched below.

Suppose that a one-component function, $f(x_i-a)$, is fitted to the observations, with respect to the location parameter a, by maximizing the one-component loglikelihood function, and that $\hat{a}$ is the corresponding solution. Then, it can be shown that $(\hat{a},\hat{a})$, in the two-dimensional parameter space, is a stationary point of the two-component loglikelihood function, considering the same observations. This stationary point is crucial, since the structure of the loglikelihood function at this point will decide on whether two distinct locations will be found, or, only one. From Catastrophe theory we learn that it is useful to transform the parameters a_1 and a_2 into a new set of parameters, defined by: $A_1=Ia_1+(1-I)a_2$ and $A_2=a_1-a_2$, where A_1 is related to the center of mass of the peaks, and A_2 to the distance between the peaks. Now, only one of these two parameters is essential, that is, A_2. This means that the structure of the loglikelihood function can only change in the direction of A_2, and not in the direction of A_1. It can be shown that, in the direction of A_1, the loglikelihood function has always a maximum, in the one-component stationary stationary point $(2\hat{a},0)$, expressed in terms of A_1 and A_2. In the direction of A_2, $(2\hat{a},0)$ can be a maximum as well as a minimum. The possible structures of the loglikelihood function, as a function of A_2, around this stationary point, can be described by

$$a + bA_2^2 + cA_2^3 + dA_2^4$$

with

$$\begin{aligned} a &= \sum_i n_i (ln \ p_i)_{(2\hat{a},0)} \\ b &= \sum_i \frac{n_i}{2} \left(\frac{\partial^2 lnp_i}{\partial A_2^2} \right)_{(2\hat{a},0)} \\ c &= \sum_i \frac{n_i}{3!} \left(\frac{\partial^3 lnp_i}{\partial A_2^3} \right)_{(2\hat{a},0)} \\ d &= \sum_i \frac{n_i}{4!} \left(\frac{\partial^4 lnp_i}{\partial A_2^4} \right)_{(2\hat{a},0)} - \left(\left(\sum_i \frac{n_i}{3} \frac{\partial^3 lnp_i}{\partial A_1 \partial A_2^2} \right)^2 \Big/ \sum_i \frac{n_i}{8} \frac{\partial^2 lnp_i}{\partial A_1^2} \right)_{(2\hat{a},0)} \end{aligned} \qquad (5)$$

The possible situations in the direction of A_2 are given by:

- if $b>0$: the loglikelihood function has a minimum in $A_2=0$ and two further stationary points being maxima.
- if $b<0$ and $9c^2-32bd > 0$, the loglikelihood function has a maximum in $A_2=0$ and two other stationary points. However, it can be shown that the maximum in $A_2=0$ is very likely to be the absolute maximum. Also, the stationary points are so close that they are almost inseparable.
- if $b<0$ and $9c^2-32bd <0$, the loglikelihood function has only a maximum in $A_2=0$.

From these observations it can be concluded that the factor b indicates whether or not the ML-estimate of the parameter A_2 will be zero, or thus, whether or not the estimates of the locations of the peaks will coincide. If $b>0$ the peaks can be resolved, if $b<0$ they can not. This coefficient b depends completely on the observations, given a certain model. This means that b can be considered as a stochastic variable, having a certain probability density function. If this probability density function is known, the probability of finding $b>0$, denoted by $P(b>0)$, can be derived. This probability is given by:

$$P(b>0) = P\left(\sum_i \frac{n_i}{2} \left(\frac{\partial^2 ln \ p_i}{\partial A_2^2} \right)_{(2\hat{a},0)} > 0 \right) \qquad (6)$$

which is called the *probability of resolution*. If the total number of counts, N, is large enough, the number of counts in different intervals can be considered as being independent. In this case, the *Central Limit Theorem* can be applied, so that distribution of the factor b can be considered as being a normal distribution. If the mean value of this distribution is given by $E[b]$, that is, the expectation value of Eq.(5), and the variance is given by

σ_b^2, the probability of having $b>0$ is given by:

$$P(b>0) = 1 - P\left(X < -\frac{E[b]}{\sigma_b} \right) \qquad (7)$$

where the stochastic variable X has a standard normal distribution. It is clear that the ratio between $E[b]$ and σ_b plays a crucial role in determining the probability of resolution.

It will be assumed that the intervals are small enough so that the sum in Eq.(5) can be replaced by an integral. For gaussian peaks this leads to:

$$E[b] = I^2(1-I)^2 \frac{NA_2^2}{2\sigma_o^2} \qquad (8)$$

with σ_o the width of the gaussian. The variance is found to be equal to:

$$\sigma_b^2 = I^2(1-I)^2 \frac{N}{2\sigma_o^4} + I^2(1-I)^2 \frac{A_2^2}{\sigma_o^6} \qquad (9)$$

For closely located peaks, the second term in Eq.(9) can be neglected, so that:

$$\frac{E[b]}{\sigma_b} = I(1-I)\frac{\sqrt{N}A_2^2}{\sqrt{2}\sigma_o^2} \qquad (10)$$

The larger this ratio, that is, the larger N and/or the larger the ratio $(A_2/\sigma_o)^2$, the larger the probability of separating the peaks, as could be expected. We also see that if the difference between the peak heights becomes larger, i.e., if the value of I goes from 0.5 towards 0 or towards 1, the probability of resolution decreases.

4. Statistical Precision

In the previous section, it was shown that the ML-estimator is not really suited to determine the locations of severely overlapping objects.

The next question that arises is whether an unbiased estimator, if one could find one, would do better. The maximal statistical precision an unbiased estimator can achieve is limited by the Cramér-Rao Lower Bound. The minimal variance of the estimator of the distance parameter A_2, i.e., $Var_{cr}(A_2)$, is given by:

$$Var_{cr}(A_2) = \frac{M_{11} + M_{22} + 2M_{12}}{M_{11}M_{22} - M_{12}^2} \qquad (11)$$

with

$$M_{kl} = -\sum_i E[\underline{n}_i]\left(\frac{\partial^2 \ln p_i}{\partial a_k \partial a_l} \right)_{(a_1^o, a_2^o)} \quad k = 1,2 \quad l = 1,2 \qquad (12)$$

a_1^o and a_2^o are the true values of the location parameters. This expression for the *CRLB* is rather complicated and not very easy to use. Therefore, a more simple and useful expression is derived, which will approximate the real values of *CRLB* sufficiently well. It will be assumed that A_2 is very small, as compared to the width of the objects, so that

$$\begin{aligned} p_i &\approx I f(x_i - A_1) + (1-I) f(x_i - A_1) \\ &\approx f(x_i - A_1) \end{aligned} \qquad (13)$$

Under this assumption, $M_{11}+M_{22}+2M_{12}$ can be approximated by

$$\sum_i \left(\frac{N(1-I)^2}{f(x_i - A_1)} \left(\frac{I}{1-I} f_{a_1}^{(1)}(x_i) + f_{a_2}^{(1)}(x_i) \right)^2 \right)_{(a_1^o, a_2^o)}$$

and $M_{11}M_{22}-M_{12}^2$ by

$$\left(\sum_i \frac{NI^2(1-I)^2}{f(x_i - A_1)} \left(f_{a_1}^{(1)}(x_i) \right)^2 \right)^2_{(a_1^o, a_2^o)} - \left(\sum_i \frac{NI^2(1-I)^2}{f(x_i - A_1)} f_{a_1}^{(1)}(x_i) f_{a_2}^{(1)}(x_i) \right)^2_{(a_1^o, a_2^o)}$$

with $f_{a_k}^{(1)}(x_i)$ a short notation for the first order derivative of $f(x_i - a_k)$ with respect to a_k. The partial derivatives will now be Taylor expanded about $A_2=0$:

$$f_{a_1}^{(1)}(x_i) \approx -f_{x_i}^{(1)}(x_i - A_1) + (1-I)A_2 f_{x_i}^{(2)}(x_i - A_1) \qquad (14)$$

$$f_{a_2}^{(1)}(x_i) \approx -f_{x_i}^{(1)}(x_i - A_1) - IA_2 f_{x_i}^{(2)}(x_i - A_1) \qquad (15)$$

so that

$$M_{11} + M_{22} + 2M_{12} \approx N \sum_i \frac{(f_{x_i}^{(1)}(x_i - A_1))^2}{f(x_i - A_1)} \qquad (16)$$

and

$$M_{11}M_{22} - M_{12}^2 \approx \frac{NI^2(1-I)^2}{4} \sum_i \left(\frac{A_2 \left(f_{x_i}^{(2)}(x_i - A_1) \right)^2}{f(x_i - A_1)} \right)_{(a_1^o, a_2^o)} \times \sum_i \left(\frac{\left(-2 f_{x_i}^{(1)}(x_i - A_1) + (1-2I) A_2 f_{x_i}^{(2)}(x_i - A_1) \right)^2}{f(x_i - A_1)} \right)_{(a_1^o, a_2^o)} \qquad (17)$$

If the objects are supposed to be gaussian peaks, we find for Eq.(16):

$$M_{11} + M_{22} + 2M_{12} \approx \frac{N}{\sigma_o^2} \qquad (18)$$

and for Eq.(17)

$$M_{11}M_{22} - M_{12}^2 \approx \frac{N^2 I^2 (1-I)^2}{4} \frac{2A_2^2}{\sigma_o^4} \left(\frac{4}{\sigma_o^2} + \frac{2(1-2I)^2 A_2^2}{\sigma_o^4} \right)_{(a_1^o, a_2^o)} \qquad (19)$$

If the higher order terms in A_2^2/σ_o^4 are neglected, the variance, see Eq.(11), is finally found to be:

$$Var_{cr}(A_2) = \frac{\sigma_o^4}{2NI^2(1-I)^2 A_2^2} \qquad (20)$$

Eq.(20) gives an approximation of the Cramér-Rao variance for the distance parameter, or, in other words, the ultimate precision any unbiased estimator of the distance can achieve. This variance is a function of the same factors that appeared in Eq.(10).

If this variance becomes too large, the peaks can no longer be located in an adequate way, making resolution no longer possible. Important is the ratio of the standard deviation of the estimated distance to the distance itself, which we define as the resolution factor R. Using Eq.(20), R is found to be given by

$$R \equiv \frac{\sigma_{A_2}}{A_2} = \frac{1}{\sqrt{2N}I(1-I)}\left(\frac{\sigma_o}{A_2}\right)^2 \qquad (21)$$

An, intuitive, obvious resolution criterion is that R must be smaller than 1, that is,

$$\frac{\sigma_o^2}{\sqrt{N}A_2^2} < \sqrt{2}I(1-I) \qquad (22)$$

which is a simple and useful rule of thumb.

5. Simulation Experiments

Table 1 lists the results of 3 experiments. Each experiment consisted in generating different sets of observations. The underlying model in all 3 experiments is the sum of two normalized gaussian functions, with as width $\sigma_o=30$, total number of counts $N=5000$ and $I=0.3$. However, for each experiment a different distance between the peaks was chosen. In each experiment, 10000 sets of observations were simulated. For each of these sets of observations, the b-coefficient, see Eq.(5), as well as the estimate of the distance parameter A_2 were calculated. The values of $E[b]$, σ_b and $P(b>0)$, as predicted by Eq.(8), Eq.(9) and Eq.(7) respectively, are listed in the table. Also, the experimental values for the expectation value of b, $<b>$, the standard deviation of b, s_b, and the number of negative b-values, $\#b>0$, are listed. It is concluded that there is a close agreement between the predicted values and the experimental values. The 95% confidence intervals of the found percentages contain the predicted values of the probabilities.

Fig.(1) shows the histogram of the estimated distances, of one of the experiments described above, that is, the one where the real distance was 10. It can be observed that a large number of the estimates are exactly zero. We observed also that, with decreasing distance, the mean value of the non-zero estimates will deviate more and more from the true value. This implies that the Maximum Likelihood fails for locating closely located objects.

As a second example, the *CRLB* of the distance parameter, Eq.(11), and its approximation, Eq.(20), will be investigated for gaussian peaks.. The resolution factor R, see Eq.(21), is the most interesting quantity to investigate. We define R_{exact} as the non-approximated factor R, i.e., the root of Eq.(11) divided by A_2.
Fig.2 shows R_{exact} and R as a function of the ratio A_2/σ_o, for $I=0.1$ and $I=0.5$, with $N=5000$ for both situations. From this figure, it is concluded that the approximations predict the exact values sufficiently accurately.

Conclusion

If the Maximum Likelihood estimator is used to resolve two overlapping objects, the objects may collapse, in the sense that their estimated locations may coincide exactly. The probability that this happens can be calculated and depends only on a factor that is function of the distance of the objects, their widths, the number of counts and the intensity ratio of the objects. If this probability differs from zero, a bias is introduced. The attainable precision of an unbiased estimator, if one exists, depends on the same factor that determines the probability of resolution of the Maximum Likelihood estimator.
This implies that for severely overlapping objects, resolution is impossible in any way. In this paper, only the one-dimensional case was described, however, the developed theory can be extended to higher dimensions.

A_2	E[b]	<b>	σ_b	s_b	P(b>0)	#b>0
6	0.0049	0.0048	0.012	0.012	66.28%	6563
10	0.014	0.013	0.012	0.012	87.90%	8798
14	0.028	0.026	0.012	0.012	98.84%	9860

TABLE 1: FOR EACH DISTANCE A_2, 10000 SETS OF OBSERVATIONS WERE GENERATED, THE VALUES E[b], σ_b AND P(b>0) ARE GIVEN. THE CORRESPONDING EXPERIMENTAL VALUES <b>,s_b AND #b>0 ARE ALSO LISTED

References

[1] Lord Rayleigh, *Philosophical Magazine*, 47, 1874, 81-93, 193-205.

[2] A.J. den Dekker and A. van den Bos, Resolution: a survey, *Journal of the Optical Society of America*, A14 (3), 1997, 547-557.

[3] T. Poston, I.N. Stewart, *Catastrophe Theory and its Applications* (London, Pitman, 1978)

[4] A. van den Bos, Journal of the Optical Society of America, *Optical Resolution: an analysis based on catastrophe theory*, A4, 1987, 1402-1406.

[5] A.J. den Dekker, Model-Based Optical Resolution, *IEEE Transactions on Instrumentation and Measurement*, 46 (3), 1997, 798-802.

[6] A. van den Bos, *Parameter Estimation, Chapter 8 in Handbook of Measurement Science*, (Chichester, Wiley, 1982).

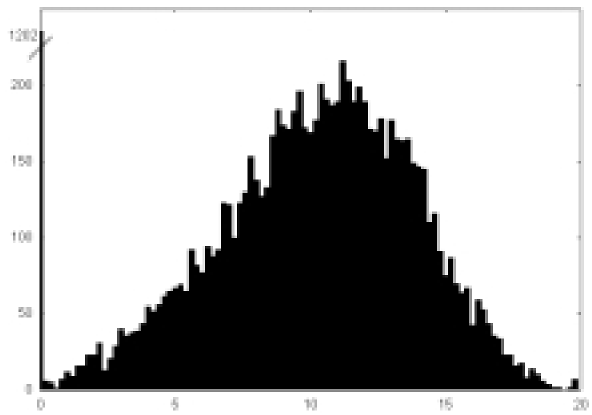

FIG.1: HISTOGRAM OF ESTIMATIONS OF A_2, TRUE VALUE OF A_2=10, σ_o =30, N=5000, I=0.3.

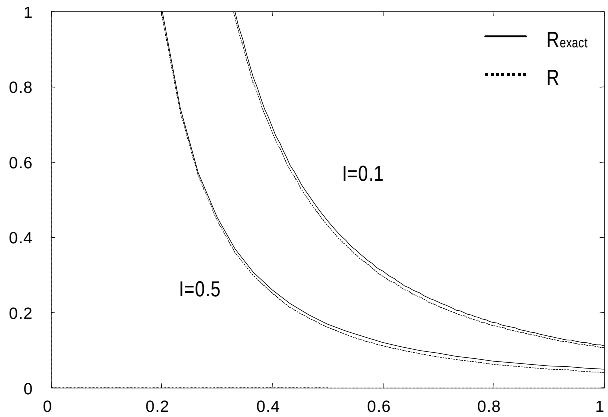

FIG.2: THE FACTOR R_{exact} AND ITS APPROXIMATION R, AS A FUNCTION OF A_2/σ_O, WITH N=5000, in case of I=0.1 AND I=0.5.