Stability of model and selection of parameters with applications in image measurements

Eric Lantz, Jacques Duvernoy

Laboratoire d'optique P.M. Duffieux U.A. 214 CNRS U.F.R des sciences et techniques 25030 Besançon Cedex France

Abstract :

In optical microscopy, the a priori knowledge of the nature of the object to be imaged and of the transfer function of the optical system allows to improve the limit of resolution beyond classical bounds derived from the consideration of the optical transfer only. This communication presents a quantitative study of this improvement as a function of the object model and of the image noise. The method is derived from recent studies about the limit of resolution in image restoration. An application to linewidth measurement on integrated circuits is shown.

Résumé:

En microscopie optique, la connaissance a priori de la nature de l'objet imagé et de la fonction de transfert du système optique permet d'accroître considérablement la limite de résolution. On présente une étude quantitative de cette limite, en fonction du modèle adopté pour l'objet et du bruit dans l'image. La méthode d'analyse dérive des études récentes portant sur la limite en résolution de la restauration d'image. Une application à la mesure de la largeur de traits sur circuits intégrés est présentée.

Introduction :

The accuracy in metrological optical microscopy can be improved by modeling the object and the optical transfer. Widths of binary objects, like thin lines of integrated circuits have been measured {1} with an experimental accuracy of 0.02µm, using a model $m(p_1,\dots,p_n)$, where $p_1,\dots,p_n$ are a small number (n=3 for binary objects) of unknown parameters.

Using such an a priori information can be compared with "super-resolution" techniques where general a priori information, like localisation or positivity, is used. In these techniques, the resolution must be limited by means of regularisation operators {2}, in order to avoid instabilities in the reconstructed object. We will show that, with modeling techniques, the number of parameters must also be limited. The first part of this communication is devoted to the quantitative assessment of the accuracy obtained using a given model.

1) Theoretical study of the uncertainties of the model parameters

1.1) Expression of the uncertainties

A 1-D experimental image, $f(x), x=1,\dots,X$ is digitized in X pixels, in the image plane of an optical imaging system. A model $m(x,p_1,\dots,p_n)$ of this image is supposed to be available with a number, n, of parameters much lesser than the number, X, of pixels. This model accounts for the convolution of an object $O(x,p_1,\dots,p_n)$ and of the spread function of the optical system. Typically parameter p_1 will represent the line width, L. We assume that the model is an exact representation of the noise-free object :

$$f(x) = m(x,p_{01},\dots,p_{0n}) + b(x) \qquad [1]$$

where $p_{01},\dots,p_{0n}$ are the (unknown) exact values of the parameters that define the object. b(x) denotes the noise in the digitized image; it is assumed to be an additive, white, zero-mean gaussian process with variance σ^2.

A "best" set of parameters $p_1,\dots,p_n$, that minimizes a mean square criterion $\varepsilon(x)$,is obtained, using a non linear iterative method, (Gauss-Newton like) :

$$p_1,\dots,p_n \text{ such that } \sum_{x=1}^{X} [f(x)-m(x,p_1,\dots,p_n)]^2 = \sum_{x=1}^{X} \varepsilon(x)^2 \text{ minimum} \qquad [2]$$

Our purpose is to determine the relative uncertainties dp_i between the true and the computed parameters :

$$dP_i = (p_i-p_{0i})/p_i \quad ,i=1,\dots,n$$

To express them, it can be assumed that they are small enough (much lesser than 1) in order to expand the model $m(x,p_1,\dots,p_n)$ into a first order Taylor serie in the neighborhood of $(p_1,\dots,p_n)$:

$$f(x)=m(x,p_1,\dots,p_n)+\varepsilon(x) \equiv \sum_i (\partial m(x,p_1,\dots,p_n)/\partial p_i).(p_i-p_{0i})+\varepsilon(x) \qquad [3]$$

Equations [1] and [3] yield :

$$b(x)-\varepsilon(x) = \sum_i (\partial m(x,p_1,\dots,p_n)/\partial p_i).(p_i-p_{0i}) \qquad [4]$$

At this stage, the image model is linearized as a function of the uncertainties dp_i. Equation [4] can be expressed in a matrix form :

$$\mathbf{B_p} = \mathbf{B}-\mathbf{B}_\perp = \mathbf{A}\,\mathbf{dP} \qquad [5]$$

with : $\mathbf{B}_x=b(x)$, $\mathbf{B}_{\perp x}=\varepsilon(x)$, $\mathbf{A}_{xi} = p_i(\partial m(x)/\partial p_i)$, $\mathbf{dP}_i= (p_i-p_{0i})/p_i$

A is a matrix that represents a linear operator that maps the parameter space into the image (geometric) space. Both spaces are Hilbert spaces , of dimensions n and X, with the l_2 norm.

$\mathbf{B_p}$ is the orthogonal projection , along $\mathbf{B}_\perp$, onto the range of **A**, viz the subspace of the image space spanned by columns vectors of **A**. This subspace is at most a n-dimensional space, as far as the columns of **A** are independent. For more details on the notion of projection, see ref.{2}.

In order to estimate dP, a back transformation is required to go from the image subspace to the parameter space. This is achieved by multiplying both sides of eq. [5] by the adjoint operator $\mathbf{A}^T$, that maps the image space into the parameter space . The resulting "normal" equation is :

$$\mathbf{A^T B_p = A^T A\, dP} \qquad [6]$$

Let $(\mu_1,...,\mu_n)$ be the eigenvalues of $\mathbf{A^T A}$, ranked by decreasing magnitude order, and $(\mathbf{V}_1,....,\mathbf{V}_n)$ the corresponding eigenvectors.

Let $\mathbf{dP}_j$ be the projection of **dP** upon $\mathbf{V}_j$ and let $\mathbf{B}_j = \mathbf{AdP}_j$. Then eq. [6] yields :

$$\mathbf{A^T B}_j = \mathbf{A^T A\, dP}_j = \mu_j \mathbf{dP}_j \qquad [7]$$

Since B is a gaussian random variable, dP_j is also a gaussian random variable, whose estimated norm is an estimation of the standard deviation.

Let $||\mathbf{V}||$ be the estimation of the norm of vector **V** . $\mathbf{B}_j$ being the projection of the noise **B** onto one-among-X degrees of freedom of the image, it follows that :

$$||\mathbf{B}_j|| = 1/\sqrt{X}\ .\ ||\mathbf{B}|| \qquad [8]$$

The roots $\sqrt{\mu_j}$ of the eigenvalues of $\mathbf{A^T A}$ are the singular values of both **A** and $\mathbf{A^T}$. Hence {2} :

$$||\mathbf{A^T B}_j|| = \sqrt{\mu_j}\ ||\mathbf{B}_j|| \qquad [9]$$

Equations [7],[8],[9] yield the norm of the projected uncertainties dP_j :

$$||\mathbf{dP}_j|| = ||\mathbf{B}|| \,/\, (\sqrt{\mu_j}\ .\ \sqrt{X}\,) \qquad [10]$$

The physical uncertainties $\mathbf{dP}_i$ are obtained from the projected uncertainties $\mathbf{dP}_j$ by implementing a change of basis reciprocal to that which allows to express eigenvectors $\mathbf{V}_j$ in the basis $\mathbf{P}_i$ of the physical parameters :

$$\mathbf{P}_i = \Sigma\, \alpha_{ij}\ \mathbf{V}_j \Rightarrow ||\mathbf{dP}_i|| = ||\mathbf{B}|| \,/\, \sqrt{X}\ (\ \Sigma\, \alpha_{ij}^2 / \mu_j)^{1/2} \qquad [11]$$

This final result is a quantitative assessment of the relative uncerainty, dP_i, on each parameter P_i of the image model. This expression holds if the relative uncertainties remain much lesser than one.

1.2) Physical interpretation

According to eq.[11], a small amount of noise may induce high relative uncertainties, dP_i,if :

- operator **A** presents at least one small singular value μ_n ;
- The projection, α_{in}, of parameter P_i onto the eigenvector V_n is strong.

Such a case may arise in two different ways :

1) A physical parameter $\mathbf{P}_i$ is nearly colinear to $\mathbf{V}_n$; the small value of μ_n means that parameter P_i has little influence on the image. Such a situation results in a high uncertainty on P_i, while the other parameters still present small uncertainties.

2) Two parameters, at least, simultaneously exhibit strong projections on eigenvectors associated with both a small and high singular value.

Figures depict such a situation, where parameters $\mathbf{P}_1$ and $\mathbf{P}_2$ have significant components, $(\alpha_{11},\alpha_{1n})$ and $(\alpha_{12},\alpha_{2n})$, on the eigenvectors $\mathbf{V}_1$ and $\mathbf{V}_n$. Figure 1 represents a plane spanned by eigenvectors $\mathbf{V}_1$ and $\mathbf{V}_n$ (i.e. a parameter subspace). The mapping of this subspace by the operator **A** is also a plane, in the image (geometric) space, spanned by the eigenvectors of $\mathbf{AA}^T$. This plane is represented in fig.2 .

Let ΔP_{01} and ΔP_{02} actual variations occuring on the "true" values, p_{01} and p_{02}, of p_1 and p_2 (eq.1). The mappings of these parameters variations are nearly colinear, close to $\mathbf{V'}_1$. Thus operator **A** has degenerated the parameter space into a nearly 1-D subspace of the geometric space (fig.2). The result of the measurement turns out to be very sensitive to the image noise associated with the small singular value, μ_n. Figure 3 illustrates the influence of a such noise on the measured values ΔP_1 and ΔP_2.

Angles between projections in the geometric space are an appropriate means of assessing the degeneracy. An angle between the mappings of two parameters, which strongly differs from 90°, will point out a model instability. In a such case, it will be necessary to discard one parameter. Such a removal is equivalent to the a priori loss of resolution induced by regularisation operators in image prodessing.

2) Application to line width measurement on integrated circuits :

2.1) Experimental set-up

A microscope with a high numerical aperture((N.A=0.95) images a line illuminated with partially coherent light. The image is sensed by a C.C.D. camera and digitized. The equivalent sampling step in the object plane is 0.1 µm. Signals are acquired in a 128 pixels window. In the following , images of isolated lines are processed.

2.2) Modeling:

The bilinear transfer function (BTF) in partially coherent light is defined for a pair of spatial frequencies . Nevertheless, in the case of binary thin objects, it has been shown {3} that the spectrum, I(u), of the image intensity takes a simple form :

$$I(u) = L\ (C+D)\ \mathrm{sinc}(\pi uL)\ TA_1(u) - C\ \cos(\pi uL)\ TA_2(u)$$

L : linewidth, $C=1+T_0-2\ T_0 \cos(\phi_0) = T_0-1+D$, T_0 : relative reflectance of the object with the respect of the substrate, ϕ_0 : relative phase shift, TA_1 and TA_2 are two "apparent" transfer functions computed from the microscope BTF

The model M(u) is the product of I(u) by the transfer function of the C.C.D. camera. This model is non linear with respect to the physical parameters L, T_0, and ϕ_0, even it accounts for a linear transfer of spatial frequencies. The inversion of the model is performed using Gauss-Newton method.

2.3) Experimental assessment of the measurement precision

a) Estimating the image noise

When the model is exact, the differences between experimental data and model are due to the orthogonal noise, $\mathbf{B}_\perp$, which has X-n degrees of freedom. Because the number, n, of parameters is much smaller than the number , X, of pixels, one has $||\mathbf{B}_\perp|| \cong ||\mathbf{B}||$.

b) Results

Table 1 presents the measurements performed on thin integrated circuits with nominal widths of 5 µm and 0.8 µm. The first column shows the measurements of the linewidth L, using MALT {1}. This method, that makes use of the abscisses of zeroes in the image spectrum, has been experimentally validated, yielding measurement errors less than 0.02µm. It can play the role of a reference as to the determination of L. The next column in Table 1 gives the measurements of the parameters L, T_0, ϕ_0 obtained by a Gauss-Newton method, and their theoretical precision derived from the analysis of section 1 .

The main results are

- a theoretical precision in a good agreement with the experimental one;
- a discrepancy between the linewidth given by MALT and that obtained by minimizing the difference between data and model.

A new model has therefore been proposed to reduce this discrepancy. This model is compatible with MALT since it yields the same location for the zeroes of the image spectrum. In fact, the actual transfer function can be worse than its theoretical values (because of possible aberrations) ; the new model includes a correcting factor made up of two gaussian functions :

$M5(u) = M(u).\ [1+g_5\ \exp(-g_4.u^2)]$, where g_5 and g_4 are two additional model parameters.

Table 2 shows the results obtained using this new model, M5. The width measurements seem to be better than those reported in Table 1; but parameters T_0 and ϕ_0 are wrong in the case of the first object (5µm wide). This result is confirmed by our study of the parameter uncertainties, which are theoretically greater than 20% (while that of the linewidth is about 0.5%).

A simplified model has then be introduced, which uses a transfer correction factor with a single parameter, g_4: $M4(u) = M(u)\ \exp(-g_4.u^2)$

Table 3 summarizes the results obtained with this four parameter model, M4. Altough the global difference between experimental data and model remains of the same order as that in Tables 1 and 2, the theoretical precision on the three physical parameters is improved, especially for T_0.

According to our study of the theoretical precision, it is possible to select model M4 as the one that yields the best results. But it should be kept in mind that the actual precision does not admit an analytic expression in the case where the differences between data and model not only preceed from the image noise, but also originate in a wrong model. Nevertheless the present study should prevent from introducing many new parameters in a given model, in order to try to alleviate discrepancies between data and model.

Interestingly enough, the results summarized in Tables 1-3 show that the theoretical uncertainty on the width L is small : this parameter proves to be particulary stable. This fact accounts for the good experimental results obtained when measuring this parameter {1}.

To the contrary, the theoretical uncertainty of the phase shift ϕ_0 remains rather high. This parameter does not prove to be a very stable one. In fact its significal signification is not well established insofar as a partially coherent illumination is a combination of different plane waves, each of them being characterized by a different phase shift that depends on its angle of incidence.

References

{1} N.Noailly, "M.A.L.T Nouvelle méthode de mesure automatique de largeur de traits pour la microélectronique. Etude réalisation et validation". Thèse Institut National Polytechnique de Grenoble Juillet 85.

{2} A.Lannes, S.Roques, M.J.Casanove, "Stabilized reconstruction in signal and image processing. 1. Partial deconvolution and spectral extrapolation with limited field ", Journal of Modern Optics, Vol.34, No 2, pp.161-226 (1987)

{3} D.Courjon, D.Charraut, G.Bou Debs, "Simplifications of the bilinear transfer for microscopic binary objects", J. Opt. Soc. Am. A Vol.5, No 7, pp.1066-1072 (1988).

Table1 : experimental measurement and theoretical precision, 3-parameter model

MALT	width(M3)		T0(M3)		ϕ_0(M3)		stand.deviation
width	value	precision	value	precision	value	precision	exp-model
µm	µm	%	%	%	radians	%	arbit. units
4.72	4.83	0.4	37	2.0	1.62	6.9	0.23
0.67	0.69	1.2	83	2.0	2.17	2.7	0.10

Table2 : experimental measurement and theoretical precision, 5-parameter model

width(M5)		T0(M5)		ϕ_0(M5)		g_4	g_5	stand.deviation
value	precision	value	precision	value	precision	prec.	prec.	exp-model
µm	%	%	%	radians	%	%	%	arbit. units
4.75	0.4	89	20	0.66	91	65	198	0. 14
0.67	1.3	92	3.4	1.99	5.5	65	42	0.10

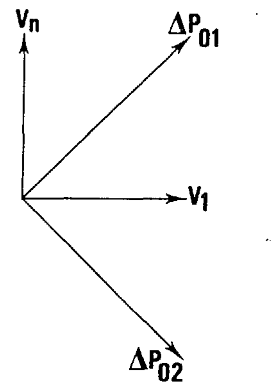

Fig.1: actual variations ΔP_{01} and ΔP_{02} of parameters P_{01} and P_{02} with significant components on V_1 and V_n associated with a high and a small singular value, respectively.

Table3 : experimental measurement and theoretical precision, 4-parameter model

width(M5)		T0(M5)		ϕ_0(M4)		g_4	stand.deviation
value	precision	value	precision	value	precision	prec.	exp-model
µm	%	%	%	radians	%	%	arbit. units
4.74	0.3	37	1.2	2.68	11	16	0. 14
0.67	0.9	87	1.6	2.5	4.0	20	0.07

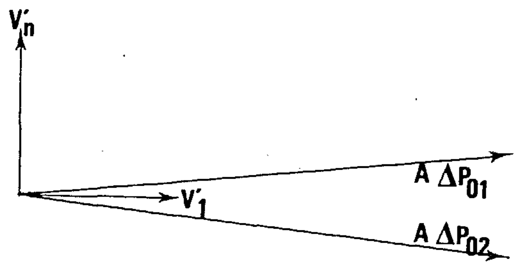

Fig.2 : mapping of the parameter plane of fig.1 into the (geometric) image space by **A.**

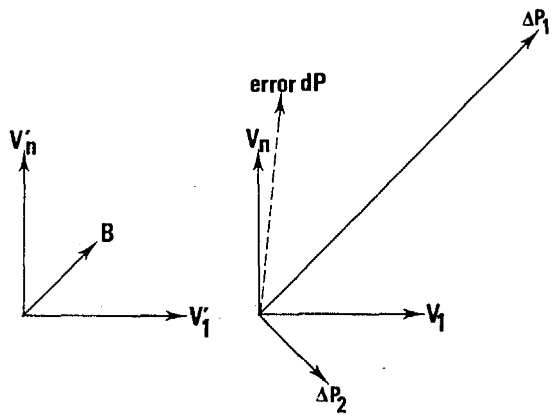

Fig.3 : influence of a white image noise on the back measurement of Δ**P**