

Optical tomography from focus

Frédéric Truchetet, Olivier Laligant

Le2i Lab. CNRS UMR 5158,
IUT, 12 rue de la Fonderie
71200 Le Creusot France

f.truchetet@iutlecreusot.u-bourgogne.fr

Abstract: A model and a method providing a 3D reconstruction of a given translucent object from a series of image acquisitions performed with various focus tunings is proposed. The object is imaged by transmission; refraction, reflection and diffusion effects are neglected. It is modeled as a stack of translucent parallel slices and the acquisition process can be described by a set of linear equations. We propose an efficient inversion technique with $O(n)$ complexity, allowing practical applications with a simple laptop computer in a very reasonable time. Examples of results obtained with a simulated 3D translucent object are presented and discussed.

© 2007 Optical Society of America

OCIS codes: (100.3020) Image Processing : Tomographic image processing; (100.3190) Image Processing : Inverse Problems.

References and links

1. Y. Y. Schechner, N. Kiryati, and R. Basri, "Separation of transparent layers using focus," Proc. of IEEE 6th Int. Conf. On Computer Vision 1061–1066 (1998).
2. T. S. Choi and J. Yun, "Three-dimensional shape recovery from the focused-image surface," Opt. Eng. **39** 1321–1326 (2000).
3. M. Noguchi and S. K. Nayar, "Microscopic shape from focus using a projected illumination pattern," Math. Comput. Modelling **24** 5/6 31–48 (1996).
4. S. Nayar and Y. Nakagawa, "Shape from focus," IEEE Trans. Pattern Anal. Machine Intell. **16** 824–831 (1994).
5. J. Ens and P. Lawrence, "An investigation of methods for determining depth from focus," IEEE Trans. Pattern Anal. Machine Intell. **15** 97–108 (1993).
6. J. Ens and P. Lawrence, "A matrix based method for determining depth from focus," Proc. CVPR 600–606 (1991).
7. J. Vitria and J. Llacer, "Reconstructing 3D light microscopic images using the EM algorithm," Pattern Recognition Letters **17** 1491–1498 (1996).
8. A. Pentland, T. Darell, M. Turk, and W. Huang, "A simple real-time range camera," IEEE Comput. Soc. Conf. Comput. Vision Patt. Recogn. 256–261 (1989).
9. A. Pentland, "A new sense for depth of field," IEEE Trans. Pattern Anal. Mach. Intell. **9** 522–531 (1987).
10. M. Asif and T. Choi, "Shape from focus using multilayer feedforward neural networks," IEEE Trans. on Image Processing **10** 1670–1675 (2001).
11. N. Dey, A. Boucher, and M. Thonnat, "Image formation model a 3-d translucent object observed in light microscopy," Proc. of IEEE ICIP (2002).
12. W. Pratt, "Vector formulation of 2d signal processing operations," Comp. Graph. and Image Proc. **4** 1–24 (1975).
13. G. Golub and C. V. Loan, *Matrix Computations*, Baltimore: John Hopkins University Press third ed. (1996).
14. H. H. Hopkins, "The frequency response of a defocused optical system," Proc. Royal Soc. London **231** A 91–103 (1955).
15. F. Truchetet, "3D translucent object reconstruction from artificial vision," Machine Vision Applications in Industrial Inspection XIV **6070** (2006).
16. C. Preza, M. I. Miller, L. J. Jr Thomas, J. G. McNally, "Regularized linear method for reconstruction of three-dimensional microscopic objects from optical sections," J. Opt. Soc. Am. A. **9** 2 219–228 (1992).
17. M. R. P. Homem, N. D.A. Mascarenhas, L. F. Costa and C. Preza, "Biological image restoration in optical-sectioning microscopy using prototype image constraints," Real-Time Imaging **8** 475–490 (2002).

18. S. Joschi, M. Miller, "Maximum a posteriori estimation with good's roughness for three-dimensional optical-sectioning microscopy," *J. Opt. Soc. Am. A.* **10** 5 1078–1085 (1993).
 19. The discrete convolution product is denoted as usual as $h * s(l, c) = \sum_j \sum_j h(i, j) s(l - i, c - j)$

1. Introduction

3D imaging of translucent objects is not a hot topic in the recent artificial vision literature. However, these objects, whose refraction index is almost constant and through which light rays follow a quasi straight line, are often observed. In microscopy for biology, for instance, a lot of objects show their internal structure only when dyes modifying only the absorption coefficient are applied. In scientific imaging of fluid movement, the same kind of behavior is observed. The short depth of focus of the optic devices used in microscopy is the main limitation on observing 3D object. However, this property can be used to estimate the thickness of the object, or more precisely, to acquire its 3D transparency function. For nontransparent objects this approach is well known as the "depth from focus" or "shape from focus" method [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. For diffusing objects this technique is well known as "optical-sectioning microscopy" [16, 17, 18]. For translucent or quasi-transparent objects, references are somewhat scarcer. Though its only concern is to provide a detailed model for the acquisition device transfer function, the work of Dey can be quoted [11]. Dey deals only with the direct problem; as far as we know, nobody has tried to solve the inverse problem of finding a 3D model for a translucent object from 2D imaging. We propose a model and a method allowing a 3D reconstruction of a given translucent object from a series of image acquisitions performed with various focus tunings. The method can be referred to as optical tomography from focus (OTFF). The robustness and the performance of the method are evaluated through simulated experiments. The results show the efficiency and the limits of the implemented algorithms.

2. Modeling a translucent 3D object

The object is imaged by transmission and is modeled as a stack of parallel translucent slices whose positions are denoted by abscissa x . Along its path through the object, the light is assumed to constitute a parallel ray flux (the object thickness and wideness are supposed to be small enough), which propagates through the stack of translucent slices with a normal incidence, as shown in Fig. 1. In this model, phenomena of refraction, reflection and diffusion are neglected; only the absorption due to the crossing of each slice is considered. This the main difference with "optical-sectioning microscopy" in which each point of the object is an omnidirectional light source.

Under these conditions, if the number K of planes is large, the transmission coefficient on plane x for the point (l, c) , $T_x(l, c)$ can be assumed to stay close to 1:

$$T_x(l, c) = 1 - a_x(l, c). \quad (1)$$

We introduce the absorption coefficient of the slice number i as following

$$a_{x_i}(l, c) = \int_{x_i}^{x_{i+1}} a_x(l, c) dx. \quad (2)$$

A discrete expression of the global transmission is

$$T(l, c) = (1 - a_{x_1})(1 - a_{x_2})(1 - a_{x_3}) \dots, \quad (3)$$

leading to a first-order discrete expansion of the transmission coefficient

$$T(l, c) = 1 - \sum_i a_{x_i}(l, c), \quad (4)$$

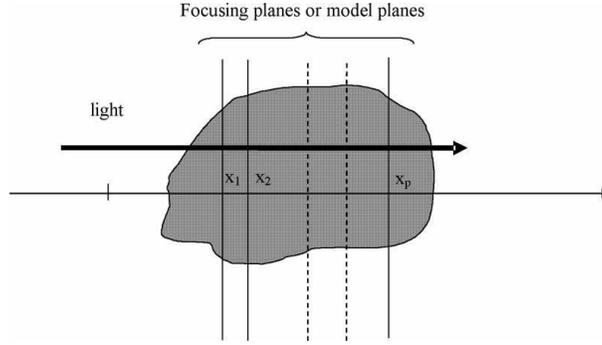


Fig. 1. Stack of parallel transparent slices model.

For the model to remain valid, the discrete transmission coefficient has to stay as close to one as possible and its variation within the slice must be small. To be more precise, a Taylor expansion shows that if the number of planes in the modeled object section is denoted by K , the term of order 2 is negligible if the absorption coefficient for each slice is around $a_{x_i} \sim \frac{0.2}{K-1}$. The experimental results presented by Dey [11] prove the validity of this model. The optical imaging device is assumed to be linear, with its blurring function (Point Spread Function) depending on the focusing distance. For a plane at abscissa x with respect to the focusing point, the impulse response is denoted as: $h_x(l, c)$. The image for a focusing distance y is given by[19]

$$s_y(l, c) = 1 - \int_x h_{x-y} * a_x(l, c) dx. \quad (5)$$

The continuous model is sampled by a first-order operator, obtaining:

$$s_{x_i}(l, c) = 1 - \sum_j h_{x_j-x_i} * a_{x_j}(l, c). \quad (6)$$

All the P images, acquired with various focusing distances are lexicographically sorted in a vector S [12] whose components are defined as

$$S_i = (1 - s_{x_i}(1, 1), 1 - s_{x_i}(1, 2), 1 - s_{x_i}(1, 3), \dots, 1 - s_{x_i}(2, 1), \dots, 1 - s_{x_i}(L, C))^t, \quad (7)$$

and the absorption coefficients of the object are similarly grouped in a vector E whose components are given by

$$E_j = (a_{x_j}(1, 1), a_{x_j}(1, 2), a_{x_j}(1, 3), \dots, a_{x_j}(2, 1), \dots, a_{x_j}(L, C))^t. \quad (8)$$

When using a matrix notation for the convolution, H_{ij} denotes the matrix associated with the impulse response $h_{x_j-x_i}$. H_{ij} is a Toëplitz-block-Toëplitz matrix, and its asymptotic limit obtained when boundary effects are neglected is circulant-blocks-circulant.

3. Inversion problem

Generally speaking, the output vector consisting of the concatenated images obtained from P shoots, each performed while focusing on one of P planes, chosen among the K planes of the

object model, is given by

$$\begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_P \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & \dots & H_{1K} \\ H_{21} & H_{22} & \dots & H_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ H_{P1} & H_{P2} & \dots & H_{PK} \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_K \end{bmatrix}. \quad (9)$$

The matrix constituted as described in the above given scheme is denoted as H and the acquisition process can be described by a set of linear equations: $S = HE$, where, under the classical asymptotic approximation for the convolution, H is the classical Toeplitz-block-Toeplitz matrix denoting a convolution operator with a finite impulse response. It is important to remark that if P images ($L \times C$), L being the number of rows and C the number of columns, with P different focusing distances, are acquired for an object modeled by K planes, the vector E has KLC components, S has PLC elements and the matrix H size is $KP(LC)^2$. As the reconstruction method is a global one, to avoid singularities, each model plane must contain its own information. Due to the PSF limited band, the object thickness and the camera resolution, the depth resolution and hence the number of planes are limited. An estimate for a simple pillbox model as PSF and standard optical material proposed in [15] shows that the number of planes must be limited to about 10. As the number of planes in the model is limited the boundary problem cannot be easily handled and globally diagonalizing this matrix by Fourier transformation is impossible. Even if the inversion problem is straightforward when the matrix is square and well conditioned, it is computationally costly. The complexity of the classical inversion algorithm is $O(n^3)$. Taking advantage from the peculiar structure of H we propose an efficient inversion technique with a complexity $O(n)$, allowing practical applications with simple laptop computer in very reasonable time. The inversion can be exact (mathematically speaking) or optimum in the least square approximation (a pseudo inversion is done in this case). The Toeplitz blocks in H do not have to be identical. The PSF can vary along the x axis ($h_{x_j-x_i}$ may depend on j) without any concern for the model and the inversion method.

3.1. 3-D object reconstruction

Reconstructing the object from the observation is the inverse problem, ie. find E given S . If the matrix H is square (ie. the number of images P is equal to the number of slices in the model K) and non-singular, then the solution can be simply expressed by $E = H^{-1}S$. However, even in this simple case, solving this equation with the classical methods of linear algebra is often not realistic, since the amount of data is too large. However, the special structure of the matrix H allows us to propose a technique dedicated to the problem that is efficient, robust and simple. The inversion process is performed in two stages. Indeed, since the images are widely greater than the convolution kernel, the asymptotic assumption is quite valid for each block H_{ij} , the boundary effects due to kernel windowing can be neglected, so passing into the Fourier domain in a first stage is possible for each block. In the second stage, the system inversion is thus finally reduced to a set of small elementary system inversions. The size of these reduced systems is fixed by the number of slices used in the model and in the acquisition process ($K \times P$). The regularization that can occur during the inversion process must be specifically considered. The algorithm is detailed in the following section.

3.2. A fast inversion method

An analytic or algebraic approach can be used for describing this method; we choose the first one since it seems more intuitive. The initial notations are simplified: the luminance of the pixel

(l, c) of the image shot while focussing on plane x_i is denoted as $s_i(l, c)$; the same notation is used for the absorption coefficient of the plane x_j , $a_j(l, c)$; and the impulse response of the optical blurring operator for the plane x_j , if the focus is on plane x_i , is denoted as h_{ji} . Letting $*$ denote the convolution operator, one can state:

$$s_i(l, c) = \sum_{j=1}^K h_{ji} * a_j(l, c). \quad (10)$$

Neglecting the boundary effects or assuming a periodization of the images, we can use the linearity of the DFT (Discrete Fourier Transform) and the theorem of Plancherel to write:

$$\hat{s}_i(n, m) = \sum_{j=1}^K \hat{h}_{ji}(n, m) \hat{a}_j(n, m), \quad (11)$$

with $\hat{s}_i(n, m)$ denoting the 2D DFT of $s_i(n, m)$. The inversion problem involves solving a linear system with P equations:

$$\hat{S}(n, m) = \hat{H}(n, m) \hat{A}(n, m). \quad (12)$$

$\hat{H}(n, m)$ is a $K \times P$ Toëplitz matrix whose elements are $\hat{h}_{ji}(n, m)$. This matrix is easy to invert;

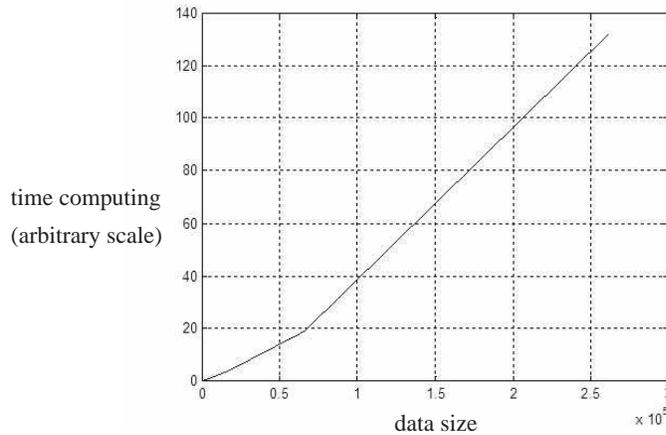


Fig. 2. Inversion computing time versus number of data.

as the number of slices in the model is limited, its size stays within reasonable limits (5 is a typical practical value for K). If the constraints about the blur resolution are taken into account, no conditioning problem can occur except for the mean value (origin in the Fourier plane), where all the K equations are the same. For this case, the inverse is simply taken to be the identity matrix, and the resulting error affects only the mean value of each reconstructed image. If the matrix $\hat{H}(n, m)$ is not square, which is the case when the shots are more numerous than the planes of the model ($P > K$), the problem is overdetermined and an inversion algorithm such as the pseudoinverse technique is used [13]. Then the result obtained is optimum in the mean square approximation sense. The inversion is iterated $L \times C$ times. Images are finally extracted by performing K inverse 2D DFTs of size $L \times C$. The overall complexity of the algorithm is $O(n)$, which is better than the best classical methods. The experimental curve in Fig. 2 shows that the experimental computation time is quasi-linearly dependent on the image size. Memory access operations are responsible for the light nonlinear discrepancy.

4. Simulations

4.1. Simulation model

A simple Gaussian model is used as the Point Spread Function in the simulations. In the context of real images, this function can be experimentally evaluated through, for example, the SML (Sum-Modified Laplacian) operator of Nayar et al. [4]. There already exists a huge literature

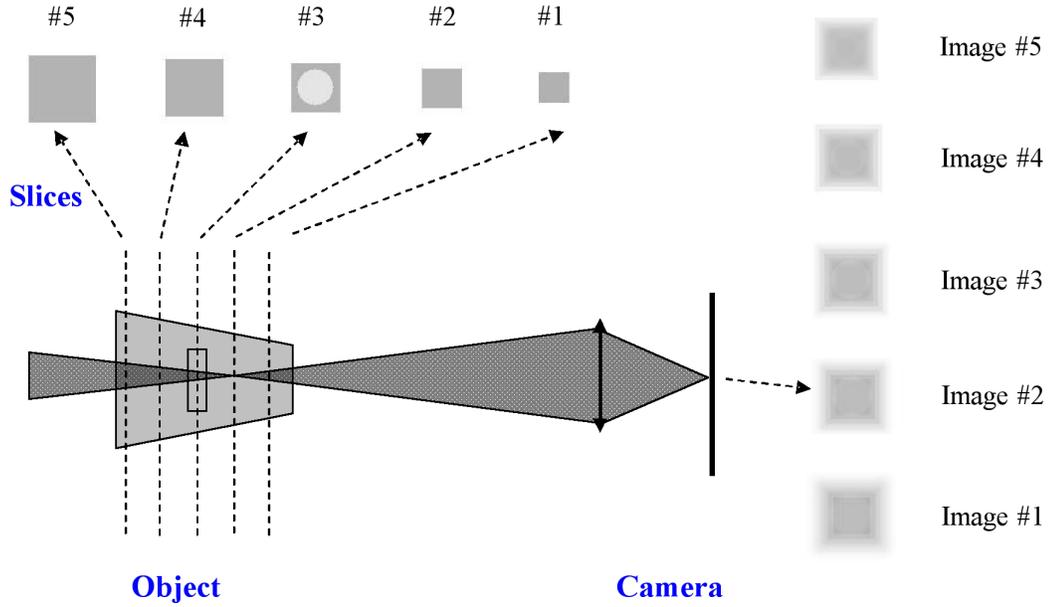


Fig. 3. Simulation device.

on this kind of measurement, and many other models, more complex, can be used (the most popular was proposed by Hopkins in 1955 [14]):

$$h_{ji}(l, c) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{l^2 + c^2}{2\sigma^2}\right) \quad (13)$$

with σ proportional to d_{ji} given by the pillbox model. This model, deduced from the elementary thin lens law, gives the blur spot size:

$$d_{ji} = \frac{fD}{x_j - f} \frac{|x_i - x_j|}{x_i} \quad (14)$$

where f is the focal length, D is the diaphragm aperture diameter, x_j is the focusing point and x_i is the object plane position.

An artificial model of a translucent 3D object has been constructed in order to validate the theoretical predictions and to test the proposed algorithms in a completely understood context. Those algorithms are independent from the object. The parameters are only driven by the measurement conditions; therefore, we chose to construct a rather simple object with a few planes, and the images of the absorption coefficients in each slice of the model are simple and clear. Those images are all different one from the other. The simulated 3D translucent object,

composed of a square base pyramid ($height = 80mm$), with a circular horizontal hole in the middle, has been cut at $K = 5$ planes parallel to the pyramid base (see Fig. 3). The P images of the model slices are presented in the same figure (in this case $P = K$). In the simulations the model is reconstructed from the blurred acquisitions given by a simulated measurement device. The standard optical parameters are: focal length is $f = 50mm$, optical aperture (in mm) is $D = 50/5.6$. This value can be tuned in order to change the divergence of the blur spot size, we abusively denote it as "optical gain". The spatial resolution of the camera is $1/a = 100pixels/mm$ and the image size is 128×128 . An example of $P = 5$ blurred images, provided by this device with the camera focused exactly on the centers of the model slices, is given in the Fig. 3. The exact inverse and the pseudoinverse methods are used to reconstruct the model images from these blurred acquisitions. For pseudoinverse method the number of input images ($P = 10$) is greater than the number of slices ($K = 5$) in the model. We will evaluate the robustness of the method against different perturbations, particularly optical blur parameters, noise, focusing accuracy and the depth resolution of the model. We show that there exists an optimal setting for the optics (focus and depth of field). The classical instability observed in almost all inverse problem appears here. In particular, reconstructions from noisy data typically manifest distortion in the low frequency components.

4.2. Simulation results

In a first test, the object is reconstructed without any perturbation of the data and with the exact values for all the parameters. In this case, the exact inverse ($K = P = 5$) and the pseudoinverse ($P = 10, K = 5$) methods give the same results. High quality images of the absorption coefficient planes are obtained with no visible defects. The PSNR of each image is about $50dB$ ($RMSE = 0.003$). The PSNR and the Root Mean Square Error (RMSE) ($-20 \log(A_M/RMSE)$)

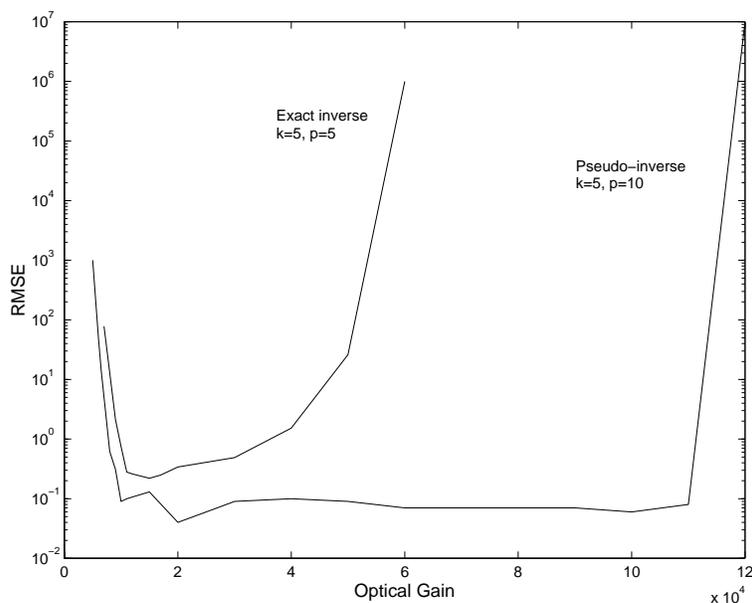


Fig. 4. Sensibility to the gain setting.

are computed from the difference between the reconstructed images and model images (A_M is the maximum pixel value). It is to be pointed out that, contrary to the fluorescence microscopy

image one[16], the quality of the images provided by CCD camera in optical tomography is not affected by Poisson noise. The PSNR of OTFF images can be expected to be usually above 55 dB and their noise can be simply modeled by a white Gaussian stochastic process. The real issue for the inversion method in OTFF is its robustness against the optical parameters. But, before testing that, we need to determine if there exists an optimal optical setting.

4.2.1. Test for an optimal optical setting

In this test we evaluated the necessity of a particular and tight choice of the optical parameters for the shooting device. Therefore, the optical gain influence on the reconstruction quality was measured. For the test to be conducted we had to add some noise on the image. Therefore a white Gaussian noise was added to one image. When noise is involved in an experiment, each result is the mean of, at least, 10 tries. Figure 4 shows that there exists a tight optimal setting zone when the exact inverse method is applied, but the pseudoinverse method is more stable and allows a wide range of settings with fine reconstruction performance.

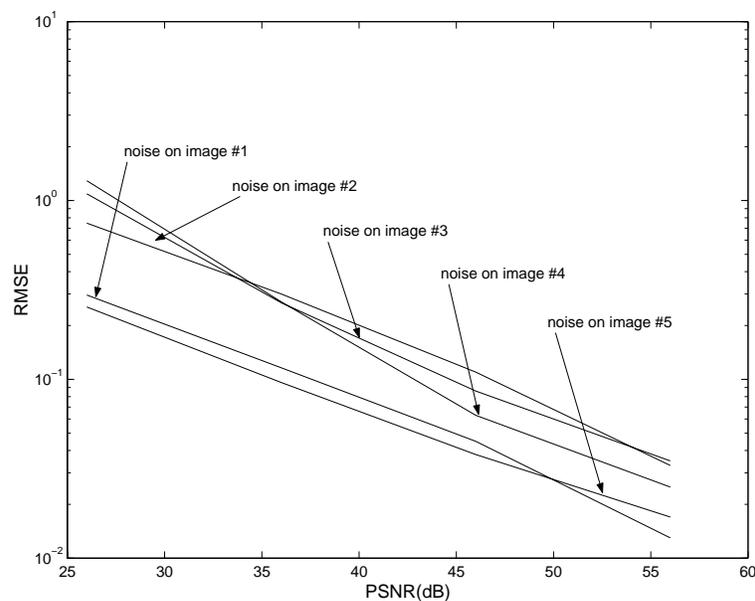


Fig. 5. Error in the reconstructed image from noisy measurements (exact inverse).

4.2.2. Test of robustness against optical parameters

The next simulation is dedicated to the study of the influence of errors on the estimate of the Point Spread Function (PSF). More precisely, it assumes that the real blur spot size is different than the one used in the matrix to be inverted. This blur spot size is defined by the standard deviation of the Gaussian model for the PSF. The curves of Fig. 7 show the evolution of the reconstruction quality when the PSF width error varies within an experimentally realistic range. Again, the pseudoinverse method is more robust than the exact inverse method. The range in which the error can be tolerated while maintaining an acceptable reconstruction quality ($RMSE < 0.04$ or $PSNR > 28dB$) is wide enough to suggest that these inversion methods can be used in practice. The quality is considered as acceptable if no visible defect can be seen.

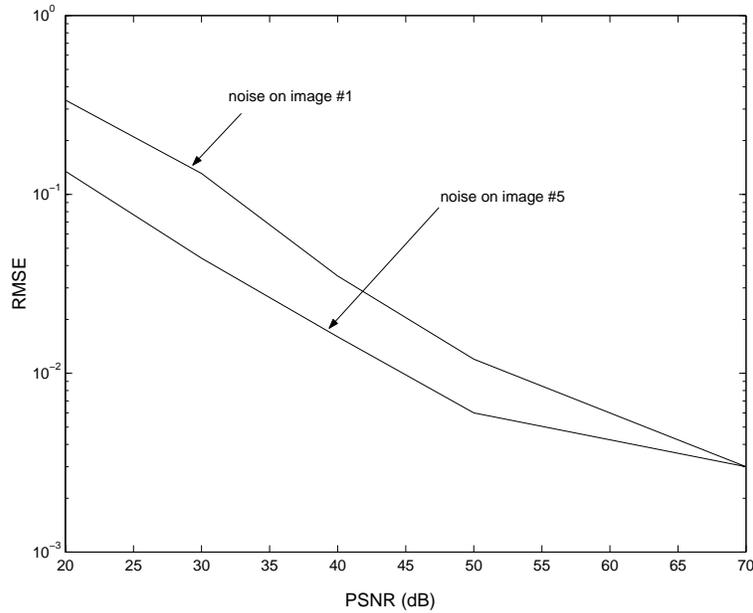


Fig. 6. Error in the reconstructed image from noisy measurements (pseudoinverse).

The next result concerns another possible discrepancy: an error on the focusing distance during the acquisition. As demonstrated by the curves of Fig. 8, the inversion methods appear to be especially sensitive to this experimental parameter, and even if the pseudoinverse method seems to be more secure, the focus offset cannot go beyond $\pm 2\%$ to stay within the acceptable reconstruction quality zone ($RMSE < 0.04$).

4.2.3. Test of robustness against noise

The last experiment was performed to evaluate the stability of the inversion methods when the acquisition is perturbed by an additive white Gaussian noise which is a reasonable model to test robustness to noise. If the algorithm performs reasonably well at low SNR for Gaussian, it should hold well for other noise models as well. The curves of Fig. 5 and 6 show the impact on the reconstruction quality of such noise when it is added to one of the images in the measurement stage. For the exact inverse method, the result depends heavily on the focusing position of the corrupted acquisition. The results are worse when a central image is noisy. For the pseudoinverse method, a greater robustness and a relative insensibility against the place of the image affected by the noise can be denoted. It even seems that the method, in this case, is more sensitive to perturbations introduced on the extreme images. The perturbation induced on the reconstructed images when the measurements are noisy is more or less distributed amongst the planes, appearing as a low frequency distortion added to the images. Examples of reconstructed images and of corresponding error images (difference between the model and the reconstruction), given in Fig. 9 and 10, illustrates this phenomena. In these examples the reconstruction quality is not acceptable: $RMSE = 0.052$. The results are almost the same if the same amount of noise is dispatched on all the images. This is illustrated in Fig. 11 where the same amount of white Gaussian noise is added to each image in the measurement stage. As anticipated, the exact inverse method is highly sensitive to noise but the pseudoinverse method still gives satisfactory

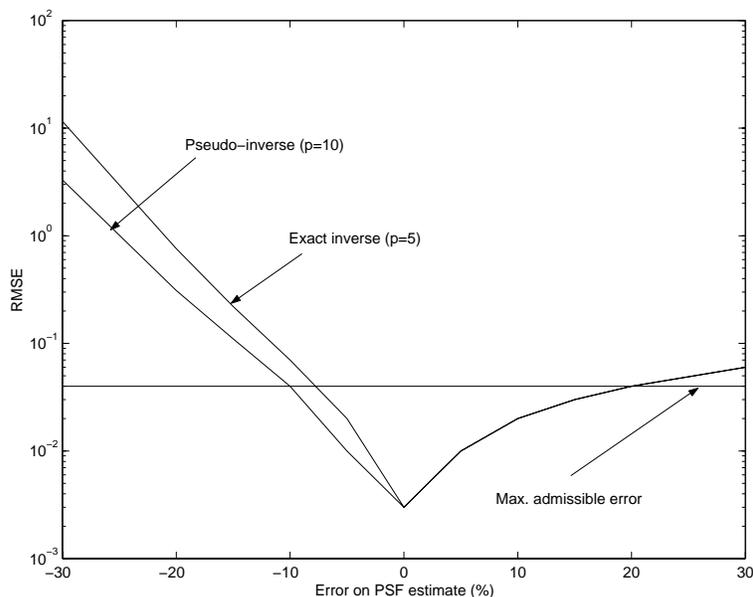


Fig. 7. Robustness against PSF width estimate.

results for PSNR in measured images higher than 45 dB. It must be remembered that PSNR of real images provided by CCD cameras is generally above 55 dB. Therefore, the pseudoinverse method is clearly able to provide a good quality reconstruction.

5. Conclusion

We presented a model and a new method providing a 3D reconstruction of a given translucent object from a series of image acquisitions performed with various focus tunings. The object is observed by transmission; refraction, reflection and diffusion effects are neglected. It is modeled as a stack of translucent planes. The acquisition process is assumed to be linear and characterized by a convolution matrix. Taking advantage from the peculiar structure of this matrix, we proposed an efficient inversion technique with a complexity of $O(n)$, allowing practical applications with simple laptop computer in very reasonable time. The inversion can be exact (mathematically speaking) or optimum in the least square approximation (a pseudo inversion is done in this case). A simple Gaussian model was used for the PSF in the simulations. Robustness of the method against different perturbations has been evaluated. The effect of optical blur parameters and noise have been explored and the experimental conditions for the best performance were determined. We demonstrated that there is an optimal setting for the optics (focus and depth of field). The pseudoinverse method benefits from the overdetermination, which gives robustness against the variation of the optical parameters. The classical instability observed in almost all inverse problem appears here in a particular way. The problem occurs with higher strength in the low spatial frequency components of the image. In our algorithm, the matrices to be inverted are composed of Fourier coefficients, and they are ill-conditioned, especially for the low spatial frequencies where the PSF response has the lowest discriminating power between the model planes. Some special regularization techniques should be developed to cope with this peculiarity. Simple Tikhonov method, or more sophisticated approaches as proposed in [16, 17, 18], should be effective. However the simplicity and the resulting fast

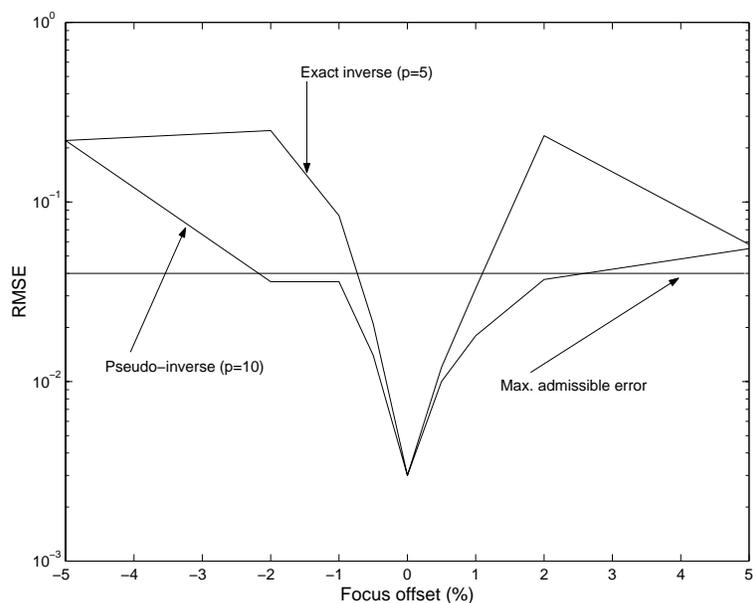


Fig. 8. Sensibility to focusing error.

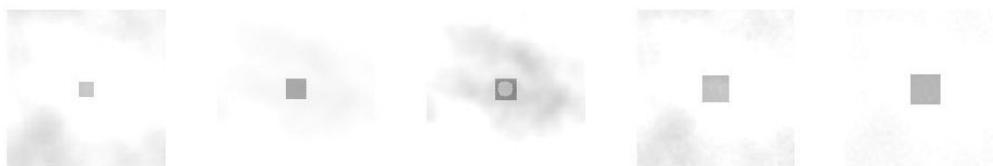


Fig. 9. Example of reconstructed images (rmse=0.052).

computing properties will be lost.

The inverse matrix coefficients can be computed once and for all, off-line, when the acquisition parameters are fixed. Therefore, a real-time deconvolution can be implemented in a software running on a simple laptop allowing an on-line use of the method. All these results must be confirmed in a real experiment conducted, for instance, with biological objects imaged by microphotography.

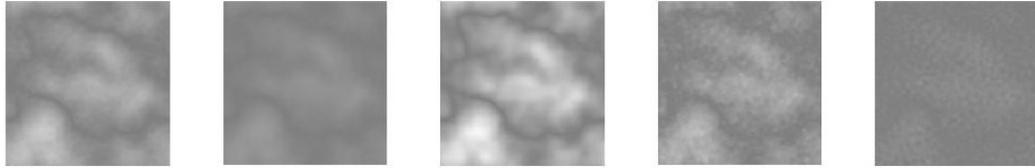


Fig. 10. Example of error images: $|reconstructed - model|$ (rmse=0.052).

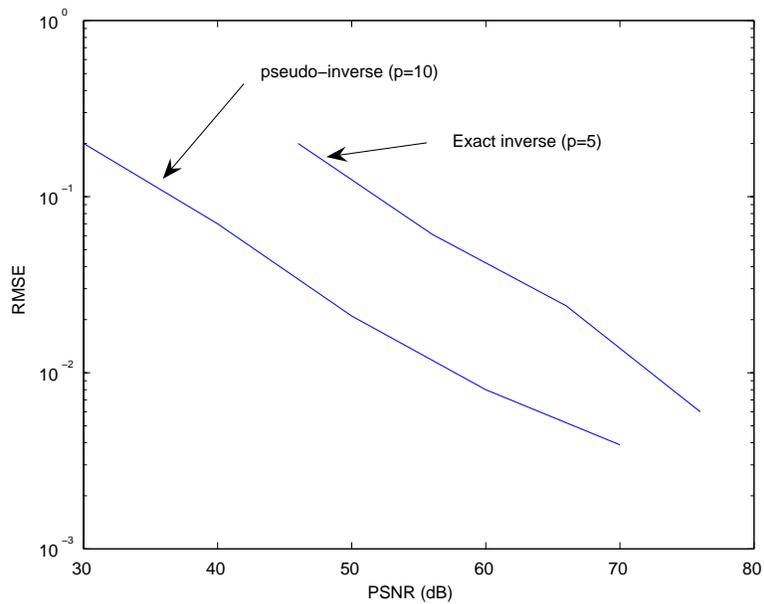


Fig. 11. Error in the reconstructed image from noisy measurements (the same amount of noise is added to each image in the measurement stage).