Title	Reconstruction of a high-resolution image on a compound-eye image-capturing system
Author(s)	Kitamura, Yoshiro; Shogenji, Rui; Yamada, Kenji et al.
Citation	Applied Optics. 2004, 43(8), p. 1719-1727
Version Type	VoR
URL	https://hdl.handle.net/11094/3338
rights	
Note	

Reconstruction of a high-resolution image on a compound-eye image-capturing system

Yoshiro Kitamura, Rui Shogenji, Kenji Yamada, Shigehiro Miyatake, Masaru Miyamoto, Takashi Morimoto, Yasuo Masaki, Noriyuki Kondou, Daisuke Miyazaki, Jun Tanida, and Yoshiki Ichioka

The authors have proposed an architecture for a compact image-capturing system called TOMBO (thin observation module by bound optics), which uses compound-eye imaging for a compact hardware configuration [Appl. Opt. **40,** 1806 (2001)]. The captured compound image is decomposed into a set of unit images, then the pixels in the unit images are processed with digital processing to retrieve the target image. A new method for high-resolution image reconstruction, called a pixel rearrange method, is proposed. The relation between the target object and the captured signals is estimated and utilized to rearrange the original pixel information. Experimental results show the effectiveness of the proposed method. In the experimental TOMBO system, the resolution obtained is four times higher than that of the unit image that did not undergo reconstruction processing. © 2004 Optical Society of America

OCIS codes: 110.2970, 110.4190, 100.3020, 230.0250, 230.3120.

1. Introduction

In recent years the demand for a portable image-capturing system has become as strong as the extension of digital imaging technology. For compact image-capturing hardware, an optical configuration that uses a microlens array is promising.[1–3] This idea is not new; it comes from the analogy of the multiple imaging system of arthropods.[4,5] A benefit of this hardware configuration is that a wide field of view can be achieved with specific imaging hardware. For example, an artificial visual sensor with 16×16 pixels was constructed with an array of microlenses and photodetectors.[1]

Regardless of the efforts dedicated to imitate the multiple imaging system, to our knowledge, no practical system capable of capturing high-resolution images has been developed. Most systems assume that the pixel number of the captured image is equal to that of the elemental imaging systems, which is the native form of the compound eye of arthropods. Therefore high-resolution imaging is problematic due to the difficulties in fabricating high-density imaging systems. A reasonable solution for this problem is using each microlens to image the object onto a subset of photodetectors and capturing multiple subsets of the image at the same time. Although the information acquired by a unit imaging system is low resolution, digital processing on the set of unit images enables us to retrieve a high-resolution image. To implement this idea, the authors propose a compact image-capturing system called TOMBO (thin observation module by bound optics).[6] TOMBO, implemented by a combination of multiple imaging systems and postdigital processing, provides a compact hardware configuration as well as processing flexibility.

In this paper an image reconstruction method to retrieve a high-resolution image from multiple low-resolution images captured by a compound-eye imag-
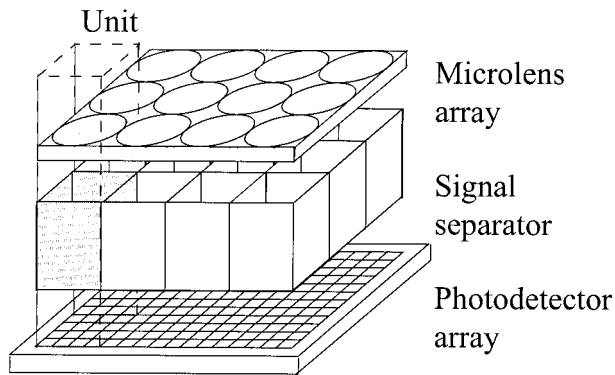
Fig. 1. Hardware configuration of TOMBO.

Fig. 2. Image acquisition by the TOMBO system.

ing system is presented. The method estimates the relation between the target object and the captured signals, which is based on a simple principle and is expected for fast processing. The proposed method is applied to the experimental TOMBO system, and the characteristics of the method are evaluated. In Section 2 the basic concept and the typical hardware configuration of the TOMBO architecture are described. In Section 3 the pixel rearrange method for image reconstruction is explained. In Section 4 the experimental results obtained by the proposed method are shown, and the effectiveness of the method is confirmed. In Section 5 issues on noise suppression and the relation to the supersampling problem are discussed.

2. TOMBO Architecture

Figure 1 shows the typical hardware configuration of the TOMBO architecture. TOMBO consists of three components: a microlens array, a signal separator, and a photodetector array. An elemental imaging system that corresponds to a microlens is called a unit, and the image captured by a unit is referred to as a unit image. The TOMBO system is considered to be a collective set of the units arranged in the array form. The signal separator is used to suppress the cross talk between the adjacent units while the hardware thickness is unchanged. The photodetector array collectively captures the signals of a set of unit images. This collective set of unit images is called a compound image.

TOMBO is characterized by the number of the unit and the number of the photodetector cells per unit. For a photodetector array with N pixels, the following equation is satisfied:

$$N = \mu\nu, \tag{1}$$

where μ and ν are the unit number and the cell number per unit, respectively. $\mu = 1$ is the case of the conventional imaging system with a single lens, and $\mu = N$ corresponds to the native compound-eye imaging system. As μ becomes large, the thickness of the system decreases because the focal length of the imaging lens can be shortened. However, recon-
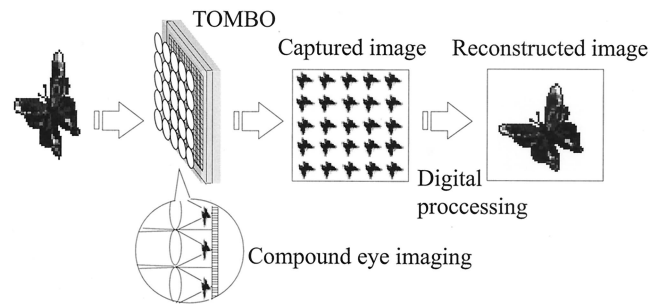
struction processing becomes difficult for large μ owing to overlap of the observed signals.

Figure 2 shows a schematic diagram of the image-capturing process in TOMBO. The object is focused onto the photodetector plane by the compound-eye imaging optics. The detected image is a compound image. The unit images composing the compound image are sampled at different positions for the individual units. A set of the unit images provides the object information equivalent to that obtained by oversampling. Therefore, combining the signals detected by all the units, we can retrieve the object image with high resolution.

3. Pixel Rearrange Method for Image Reconstruction

The problem of image reconstruction in the TOMBO architecture is reworded as retrieval of a high-resolution image from a set of low-resolution ones. The authors present the sampling method and the backprojection method[6] as solutions to this problem. Although these methods provide good results in simulations with an artificial compound image generated by a computer, processing results in the experimental system do not yield an acceptable performance. It is found that the major sources of image degradation are variation in the imaging properties of the individual lenses, misalignment of the components, and signal distortion in the analog circuits. To compensate for these undesired factors, the authors develop a new image reconstruction method called a pixel rearrange method.[7]

A. Overview

The principle of the pixel rearrange method is that the pixels in all the unit images are mapped onto a virtual image plane consisting of fine pixels, as shown in Fig. 3. For the pixel mapping, the registration parameters, which describe the correspondence between the pixels on the individual unit images and the virtual image plane, are estimated and are used to determine the pixel position. In this process one pixel value of the unit image is assigned only to a single pixel at the corresponding position of the virtual image plane without scaling of the pixel size. This operation is equivalent to neglecting the physical size of the pixel and the spatiality of the point-spread function of the imaging system, which is a key point of the proposed
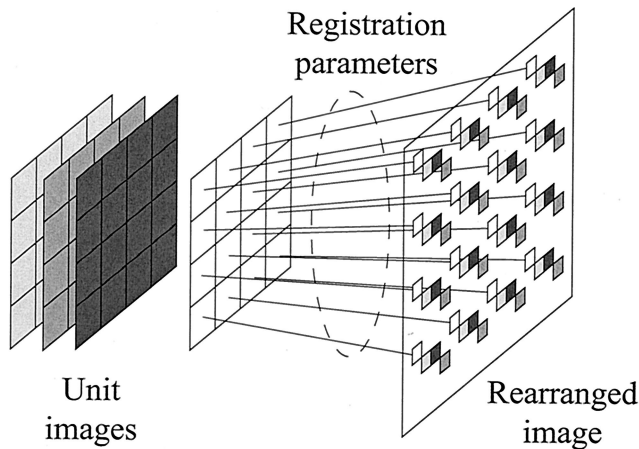
Fig. 3. Schematic diagram of the pixel rearrange method.

method. If the pixel is scaled in size and then placed on the virtual image plane, the observed pixel signal is dispersed and the image restoration becomes difficult. To avoid this situation, we approximate the observed pixel signal as the true value of the pixel center and map the signal onto the virtual image plane sparsely. Although this procedure requires that we use interpolation, the total number of calculations is expected to be reduced. Then a rearranged image composed of fine structuring cells is obtained. There are pixels that do not have a value assigned from the unit images, ensuring that interpolation processing is applied to the rearranged image. Finally, image restoration processing is executed to compensate for the blur due to the finite size of the photodetector cells and the diffraction of the imaging lens.

The concrete procedure for the pixel rearrange method for image reconstruction is as follows:

1. Correction of the shading effect and spatial variation of the sensitivity in the captured compound image.
2. Extraction of the unit images from the captured compound image.
3. Estimation of the registration parameters describing the relation between the individual unit images and the target object.
4. Mapping of the pixels in all the unit images onto a virtual image plane according to their estimated registration parameters.
5. Interpolation of the pixel values for the blank pixels in the virtual image.
6. Application of image filtering to compensate for the point-spread function presumed for the image capturing process.

Using this procedure, the authors obtained the reconstructed image from the compound image. Note that, in the image reconstruction, step 4 is the essential process, which corresponds to the backpropagation process in the previous algorithm.[6]

B. Shading Correction and Unit-Image Extraction

The signal separator located between the microlens array and the photodetector array causes signal shading for the unit images in the compound image. Light intensity in the marginal area of each unit image is reduced owing to the shade of the signal separator. In addition, the spatial variation of the sensitivity in the photodetector array degrades the quality of the reconstructed image. To reduce these effects, preprocessing is performed before image capture.

First, two reference images are captured in the system.[8] One is the black reference, and the other is the white reference. The black reference is the image signal captured by the condition of no illumination. This reference image is used to remove the signal bias caused by the dark current of the photodetector. The white reference is the image signal of an object having uniform transmittance or reflectance. With these reference images the captured raw image $O(x, y)$ is converted to the corrected one $O'(x, y)$ by the following equation:

$$O'(x, y) = C \frac{O(x, y) - B(x, y)}{W(x, y) - B(x, y)}, \qquad (2)$$

where $B(x, y)$ and $W(x, y)$ are the black and white references, respectively, and C is the parameter for signal amplification. x and y are the pixel address.

Note that the above operation on the compound image compensates for the shading and sensitivity variance for all the unit images at the same time. The compensated compound image is divided into a set of unit images that refer to the pixel address of the image detected by the photodetector array.

C. Registration Parameter Estimation

To estimate the registration parameters of the unit images, two models are considered: a space-invariant shift model and an affine transform model. The former is a simplified model in which each unit image is assumed to be a laterally translated version of the reference unit image. The latter considers the misalignment of the optical system in which each unit image is treated as the resulting image of an affine transformation due to tilt and lateral shift of the imaging lens. The authors develop two parameter estimation methods for these models.

1. Correlation Method

For the space-invariant shift model, the relative shift value of each unit image with respect to the reference unit image is sufficient for determining the mapping address of the virtual image plane. As the reference unit image, the image obtained by the center unit is chosen to minimize the effect of system misalignment. The relative shift value of the unit image can be calculated by the cross correlation between the target and the reference unit images.

A normalized cross correlation $R(x, y)$ of two-variable functions $f(x, y)$ and $g(x, y)$ is defined as follows[9]:

$$R(x, y) = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y')g(x' - x, y' - y)dx'dy'}{\left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x', y')|^2 dx'dy'\right]^{1/2} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g(x', y')|^2 dx'dy'\right]^{1/2}}. \tag{3}$$

In this case $f(x, y)$ and $g(x, y)$ are assumed to have a similar distribution, except for the relative offset. Then the cross-correlation function is ideally symmetrical with respect to the peak point, as shown by the dotted curve a in Fig. 4. For simplicity, the figure is drawn along the x axis. A noteworthy point is that the correlation signals calculated from a couple of unit images are sparse due to a small pixel number in the unit images, as shown by the filled circles (plot b) in Fig. 4. Therefore, to improve the accuracy of the peak estimation, the sparse data points are interpolated as the solid curve c. The peak point $(x_{\text{peak}}, y_{\text{peak}})$ is determined as the solution of the following equations:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{x_{\text{peak}}} R(x, y)dxdy = \int_{-\infty}^{\infty} \int_{x_{\text{peak}}}^{\infty} R(x, y)dxdy, \tag{4}$$

$$\int_{-\infty}^{y_{\text{peak}}} \int_{-\infty}^{\infty} R(x, y)dxdy = \int_{y_{\text{peak}}}^{\infty} \int_{-\infty}^{\infty} R(x, y)dxdy. \tag{5}$$

For the interpolation the bicubic interpolation was used in our experiment, which is capable of treating a two-dimensional discrete function.[10] The bicubic interpolation requires the gradients and the cross derivative of the data points, as well as the values themselves. In this process small values of the cross-correlation signal are neglected, as indicated by the dashed curve d in Fig. 4. This procedure increases the accuracy of detecting the correlation peak caused by the lateral shift between two unit images. As a result, we can determine the lateral shift values with an accuracy that is finer than the sampling pitch of the pixel.

2. Maximum Likelihood Method

Although the correlation method is simple and fast to execute, the unit images obtained by the experimental TOMBO systems are much deformed owing to misalignment of the microlens and the photodetectors. To reduce this effect, a parameter estimation method is developed based on the maximum likelihood measure.[11]

We assume a likelihood function that gives the probability of each pixel value provided with the object model and the optical system configuration. Figure 5 shows the relation between the object plane and the image plane. Affine transformation is assumed to describe the projection caused by the misalignment associated with the shift and the rotation of the microlens array.

For simplicity, we do not take the point-spread function of the optical system into consideration. When the center coordinate of each pixel in the unit image is projected on the object plane by use of a presumed registration parameter, the pixel value at the projected position can be treated as an estimated value. If the measurement error of the pixel is independent of the value of the neighborhood pixels, the likelihood of the entire unit image is a product of the likelihood of all the pixels belonging to the unit images. The value is conceptually represented as

$P[\text{all pixel values}|\text{input conditions, object model}]$

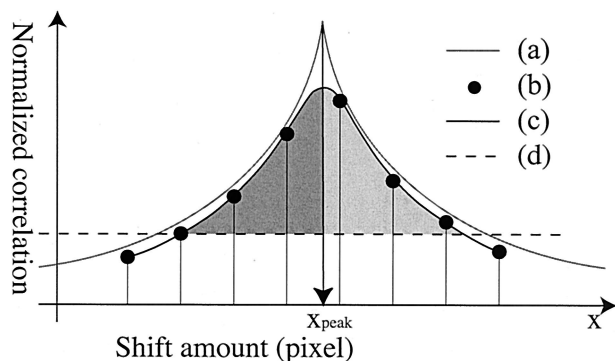$$= \prod_{k=1}^{K} \prod_{i=1}^{N} P[y_{i,k}|a_1, \ldots, a_M, y_k(x_i)], \tag{6}$$

Fig. 4. Explanation of the peak determination for the correlation between two resemble functions: (a) correlation curve between continuous functions, (b) series of correlation signals between discrete functions, (c) interpolated curve for (b), (d) threshold level for eliminating noises. The peak position is determined as the point equalizing the two shaded areas.
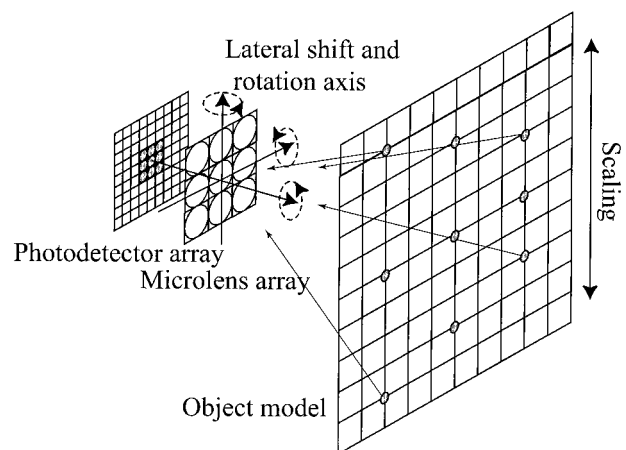

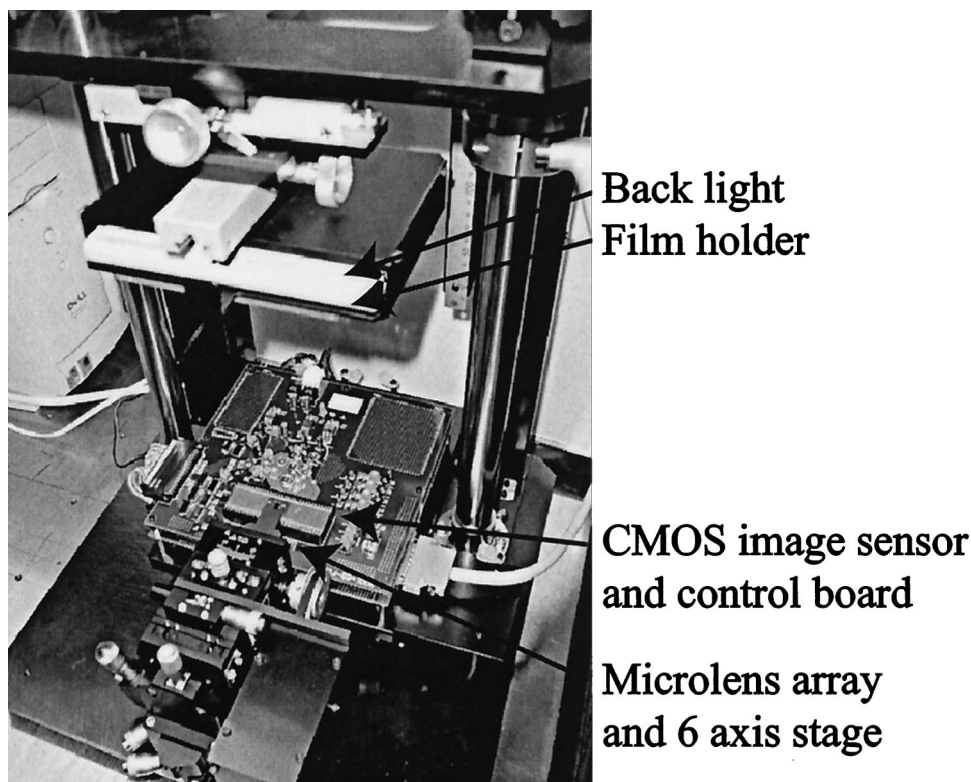Fig. 5. Relation between the object plane and the image plane.

Fig. 6. Photograph of the experimental TOMBO system.

where $y_{i,k}$ is the ith pixel of the kth unit image, $y_k(x_i)$ is the object model, and $a_1, \ldots, a_M$ are the registration parameters. N and K are the pixel number of the unit image and the number in the unit images, respectively.

Here we assume that the probability of an observed pixel value is normally distributed and that the standard deviation of the observed pixels from their expected value is the same for all the pixels. This deviation comes from the measurement error and the difference between the assumed model and the actual system. If these error sources are independent, the central limit theorem predicts that the resulting error distribution becomes close to normal. Therefore the maximum likelihood estimation of the registration parameters is converted to a problem of finding the registration parameters providing the minimum square error with respect to their projected values. We define the merit function E as follows:

$$E(a_1, \ldots, a_M) = \frac{1}{KN} \sum_{k=1}^{K} \sum_{i=1}^{N} [y_{i,k} - y_k(x_i)]^2. \quad (7)$$

In our experiment the optimal registration parameters were determined by use of the downhill simplex method, which is an algorithm suitable for minimization in multidimensional space.[12] The downhill simplex method is started with $N + 1$ initial points that are considered as the vertices of a simplex for N-dimensional space. With an evaluation of the target function at the vertices, the algorithm leads the simplex to the minimum point of the target function.

Although this algorithm is not necessarily the most effective one, we use it for simplicity. A benefit of this algorithm is that it finds a local minimum point without the use of differential functions.

D. Image Filtering

We apply image filtering to the interpolated image to compensate for image degradation caused by the point-spread function of the image-capturing system. In our experiment a pseudoinverse filter is applied to the interpolated image. The pseudoinverse filter is a popular image filter used to restore image degradation such as image blur.

A standard formulation of image degradation is given by

$$G(f_x, f_y) = F(f_x, f_y)H(f_x, f_y) + N(f_x, f_y), \quad (8)$$

where $F(f_x, f_y)$ is the Fourier transform of the ideal image, $H(f_x, f_y)$ is that of the point-spread function, and $N(f_x, f_y)$ is the spectrum of the additive noise. f_x and f_y are the spatial frequencies. The problem is to find a reasonable estimation of $F(f_x, f_y)$ from $G(f_x, f_y)$ and from the knowledge of $N(f_x, f_y)$. The pseudoinverse filter is an operator that provides the estimated image $\hat{F}(f_x, f_y)$ that minimizes the mean-square error against $F(f_x, f_y)$. The spectrum of the estimated image is obtained as follows[13]:

$$\hat{F}(f_x, f_y) = \frac{H(f_x, f_y)^2 G(f_x, f_y)}{H(f_x, f_y)^2 H(f_x, f_y) + C}, \quad (9)$$
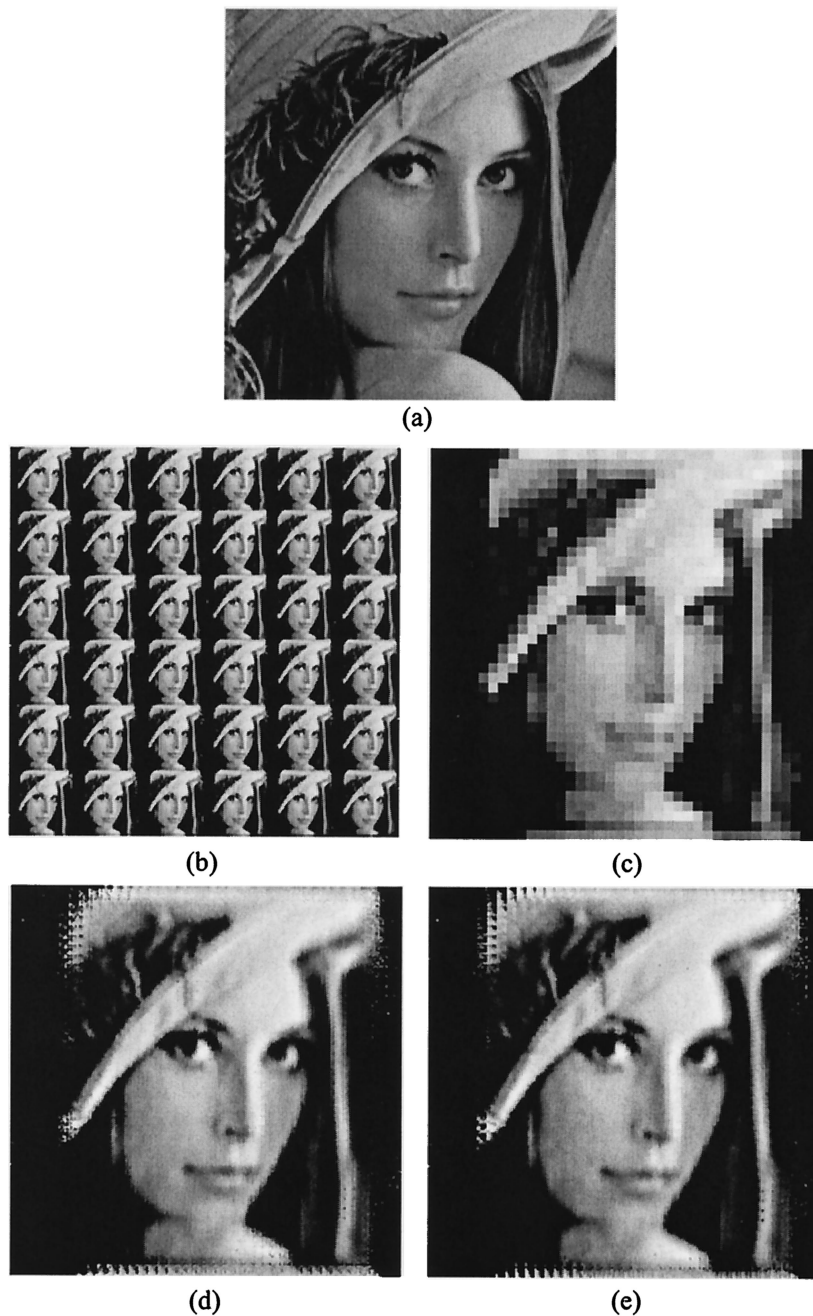
Fig. 7. Image reconstruction of Lena: (a) original image, (b) compound-eye image, (c) unit image, (d) reconstructed image obtained by the 2-D correlation method, (e) reconstructed image obtained by the maximum likelihood method.

where C is a constant depending on the signal-to-noise ratio of the power spectrum.

4. Experimental Results

Figure 6 shows a photograph of the experimental TOMBO system. In the system a refractive microlens array (Advanced Microoptic Systems GmbH, APO-Q-P500-AF1.3; focal length, 1.3 mm; diameter, 500 μm) is set in front of the photodetector array attached on the positioning stage with 6 degrees of freedom. The signal separator of nickel is made by a Lithograpie Galvanoformung Abformung process, to which antireflection coating is applied. The width of the wall is ~50 μm, the height is ~800 μm, and the pitch is 500 μm, which is equal to that of the microlens array. Considering future integration of post-digital processing, a complementary metal-oxide semiconductor image sensor is used as the photodetector array.[14] The pixel size is 12.5 μm, and the fill factor is approximately 30%. The characteristic parameters of the TOMBO system are $\mu = 6 \times 6$ and $\nu = 40 \times 40$. The target object was a transparent film mounted on the upper holder and illuminated from the back. The detected signals are quantized by 8-bit resolution and are sent to the host computer. The reconstruction processing is executed in the computer.
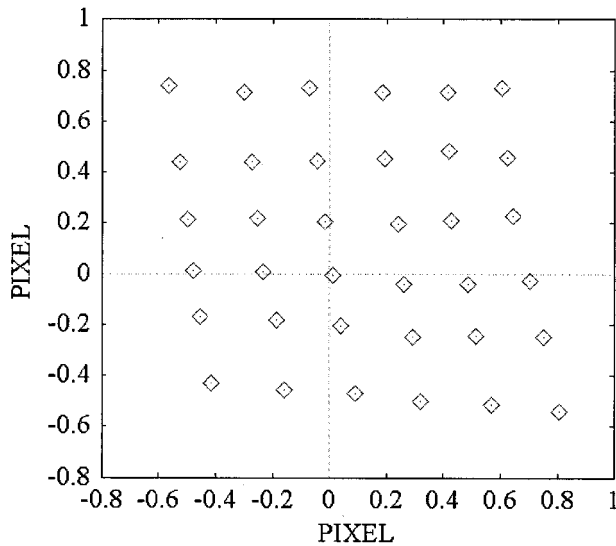
Fig. 8. Relative unit-image positions calculated by the correlation method.

Part	CM[a] (dB)	MLM[b] (dB)
Entire image (200 × 200 pixels)	10.74	10.87
Face (100 × 100 pixels)	10.44	10.50

[a]CM = correlation method.
[b]MLM = maximum likelihood method.

To evaluate the effectiveness of the proposed methods, we compared them by using a portrait picture printed on a transparent film. The size of the target object was 10 cm × 10 cm, and the distance between the object and the microlens array was 26 cm. Figure 7 shows the experimental results. Figure 7(a) is the original image, and Fig. 7(b) is the compound image with 240 × 240 pixels after correction for shading. A magnified view of a unit image with 40 × 40 pixels is shown in Fig. 7(c). The relative positions of the unit images obtained by the correlation method are depicted in Fig. 8. In this calculation, 30 × 30 pixels around the center of the unit image were used for the correlation to avoid the effect of the shade of the signal separator. The image reconstructed with the estimated shift values is shown in Fig. 7(d). The image consists of 200 × 200 pixels, and edge enhancement with an unsharp mask was executed on the image.

As seen from Fig. 8, the unit images are shifted anisotropically, whereas the ideal compound imaging system provides an isotropic image shift. Although such an anisotropic shift indicates misalignment of the optical system, precise adjustment is difficult and impractical. Because the correlation method cannot treat this type of distortion, gridwise noise appears along the edges in Fig. 7(d). On the other hand, the image reconstructed by the maximum likelihood method in Fig. 7(e) shows a good result. In this calculation 314 iterations were required to find an optimal result, which required 76.5 s by PowerPC G4 (700 MHz). Table 1 summarizes the signal-to-noise ratio of the obtained results. The maximum likelihood method shows a better performance than the correlation method, as shown in Table 1. We confirmed that the gridwise noise component is suppressed by the maximum likelihood method.

The resolution limit of the TOMBO was evaluated

with the Siemens's star chart. The diameter of the test chart was 8 cm, and the distance between the chart and the microlens array was 26 cm. Figure 9 shows the experimental results. Figure 9(a) is the original image, and Fig. 9(b) is the magnified image of a unit image whose resolution was measured as 38 line pairs per millimeter (lp/mm). The rearranged and interpolated image is shown in Fig. 9(c). The image size is 200 × 200 pixels. The spatial resolution was determined by measurement of the diameter of the contrast inversion on the Siemens's star. The observed resolution limits for the vertical and the horizontal directions are 153 lp/mm and 88 lp/mm, respectively. A narrow aperture of the photodetector provides high spatial resolution because the upper limit of the spatial frequency of the detectable signal is proportional to the ratio of the width of the aperture to the pixel pitch. The aperture shape of the photodetector is a rectangle of 5.1 μm × 9.0 μm, whereas the pixel pitch is 12.5 μm for both directions. Therefore the obtained results are in accordance with the aperture geometry. The image obtained after pseudoinverse filtering to eliminate the blur by the aperture shape is shown in Fig. 9(d). The Fourier transform of the aperture shape of the photodetector was used as the blurring function $H(f_x, f_y)$. Although the noise component is amplified, 153 lp/mm of the resolution limit is obtained for the vertical and the horizontal directions. The results confirm that the resolution achieved by the proposed method is four times higher than that of the unit image that did not undergo reconstruction processing.

5. Discussion

Although the proposed method provides a good result, the pseudoinverse filtering amplifies the gridwise noise components. This noise comes mainly from the spatial dispersion of the optical signals away from the optical axis. To improve the image quality, we should consider the space variance of the point-spread function over the image field.

Image reconstruction in the TOMBO system can be considered as a kind of super-sampling problem in which multiple images obtained by different sampling points are used to improve spatial resolution. For example, planetary surface reconstruction from multiple orbiter images was demonstrated.[11] Effective techniques for supersampling can be utilized in our problem. It should be stressed that the aim of TOMBO is to realize a compact imaging system with the help of digital processing. The TOMBO architecture and the supporting technologies are expected
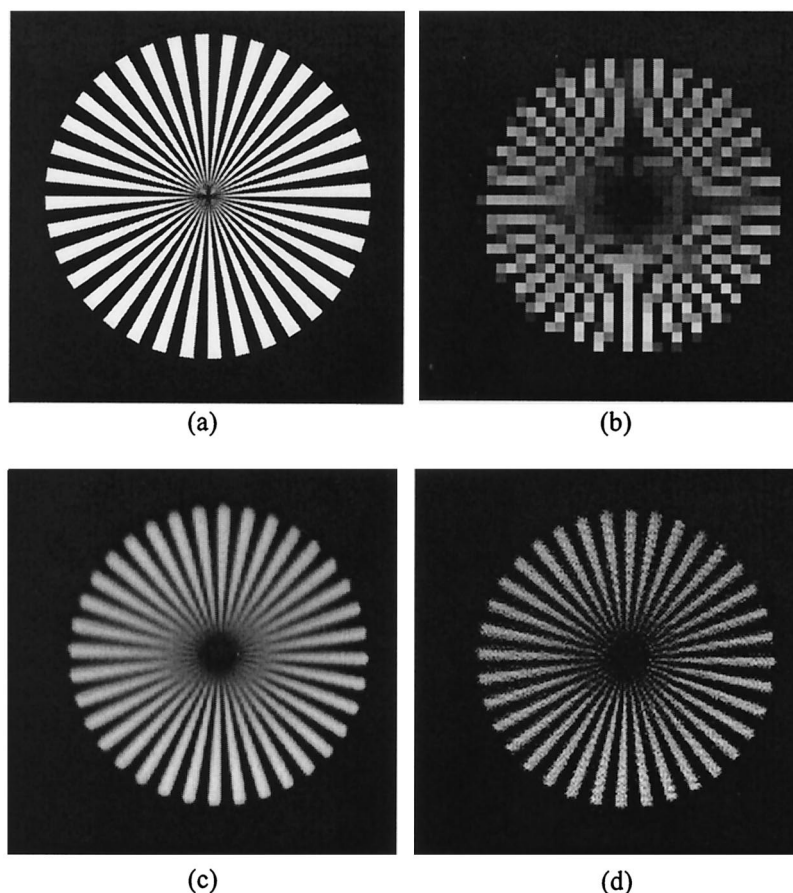
Fig. 9. Image reconstruction of the Siemens's star for resolution limit measurement: (a) original image, (b) unit image, (c) reconstructed image, (d) pseudoinverse filtered image.

to provide a new method for reducing the size of optical hardware.

In the TOMBO architecture spatial resolution is inferior to conventional single-element optics that use intensive computer processing. However, the compactness of the system is attractive in specific applications, including imaging devices for cell phones and ultrathin cameras. These applications will make possible the development of information systems based on the sophisticated use of optics and electronics.

6. Conclusion

An image reconstruction method has been proposed for an image-capturing system called TOMBO that is based on compound-eye imaging and digital processing. This method estimates the relation between the target object and the captured signals and utilizes it to reconstruct the object image. Experimental results of the image reconstruction show the effectiveness of the proposed method. In the experimental TOMBO system, the obtained resolution is four times higher than that of the unit image that did not undergo reconstruction processing.

References

1. S. Ogata, J. Ishida, and T. Sasano, "Optical sensor array in an artificial compound eye," Opt. Eng. **33,** 3649–3655 (1994).
2. J. S. Sanders and C. E. Halford, "Design and analysis of apposition compound eye optical sensors," Opt. Eng. **34,** 222–235 (1995).
3. K. Hamanaka and H. Koshi, "An artificial compound eye using a microlens array and its application to scale invariant processing," Opt. Rev. **3,** 264–268 (1996).
4. G. A. Horridge, "Apposition eyes of large diurnal insects as organs adapted to seeing," Proc. R. Soc. London **207,** 287–309 (1980).
5. E. Hecht, *Optics*, 3rd ed. (Addison-Wesley, Reading, Mass., 1998), Chap. 3.
6. J. Tanida, T. Kumagai, K. Yamada, S. Miyatake, K. Ishida, T. Morimoto, N. Kondou, D. Miyazaki, and Y. Ichioka, "Thin observation module by bound optics (TOMBO): concept and experimental verification," Appl. Opt. **40,** 1806–1813 (2001).
7. J. Tanida, Y. Kitamura, K. Yamada, S. Miyatake, M. Miyamoto, T. Morimoto, Y. Masaki, N. Kondou, D. Miyazaki, and Y. Ichioka, "Compact image capturing system based on compound imaging and digital reconstruction," in *Micro- and Nano-Optics for Optical Interconnection and Information Pro-*

cessing, M. R. Taghizadeh, H. Thienpont, and G. E. Jabbour, eds., Proc. SPIE **4455,** 34–41 (2001).

8. M. W. Burke, *Image Acquisition* (The Alden Press, Oxford, 1996).

9. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1996).

10. W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes in C* (Cambridge U. Press, Cambridge, England, 1988).

11. P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, "Super-resolved surface reconstruction from multiple images," in *Maximum Entropy and Bayesian Methods*, G. R. Heidbre-der, ed. (Kluwer Academic, Dordrecht, The Netherlands, 1996), pp. 293–308.

12. J. A. Nelder and R. Mead, "A Simplex method for function minimization," Comput. J. **7,** 308–313 (1965).

13. H. C. Andrews and B. R. Hunt, *Digital Image Restoration* (Prentice-Hall, Englewood Cliffs, N.J., 1977).

14. S. Miyatake, K. Ishida, T. Morimoto, Y. Masaki, and H. Tanabe, "Transverse-readout CMOS active pixel image sensor," in *Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications II*, M. M. Blouke, J. Canosa, and N. Sampat, eds., Proc. SPIE **4306,** 128–136 (2001).