OUKA

The University of Osaka
Institutional Knowledge Archive

Title	Optical Implementation of high-accuracy computing based on interval arithmetic and fixed point theorem
Author(s)	Sasaki, Toru; Tanida, Jun; Ichioka, Yoshiki
Citation	Optical Engineering. 1999, 38(3), p. 485-489
Version Type	VoR
URL	https://hdl.handle.net/11094/3401
rights	Copyright 1999 Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.
Note	

Optical implementation of high-accuracy computing based on interval arithmetic and fixed point theorem

Toru Sasaki
Jun Tanida
Yoshiki Ichioka
Osaka University
Graduate School of Engineering
Department of Material and Life Science
2-1 Yamadaoka, Suita
Osaka 565-0871
Japan
E-mail: sasaki@mls.eng.osaka-u.ac.jp

Abstract. We experimentally implement a new method for high-accuracy optical computing using interval arithmetic and the fixed-point theorem. In the implementation, a fixed point of an affine transformation is calculated by a TV feedback system. Employing interval arithmetic, the errors associated with an optical setup can be ignored. We evaluate two constraints on the optical system that seriously affect computing efficiency. The minimum value of computational accuracy achieved is 10^{-13}, which is determined by the control accuracy of the required affine transformation coefficients. © *1999 Society of Photo-Optical Instrumentation Engineers.* [S0091-3286(99)00703-5]

1 Introduction

Analog optical computing is an attractive field in which to explore the potential capabilities of optics.[1] Various optical information processing techniques, such as the Fourier transform and optical matched filtering, indicate the potential capability of analog optical computing. However, many inherent drawbacks also exist. A typical analog optical computing system exploits the intensity or locations of bright points as processed information. Therefore, the accuracy and dynamic range of the optical system are limited by the accuracy of quantization or the space-bandwidth product. For example, the dynamic range of typical optical systems dealing with light intensity is about 200, which is obtained by 8 bit quantization.[1] To overcome this problem, various optical computing algorithms, such as residue arithmetic[1,2] and neural networks,[1,3] have been considered.

As one of the approaches to solve the problem, a method was proposed for high-accuracy optical computing based on interval arithmetic and the fixed-point theorem.[4] Reference 2 describes the solution of a simultaneous equation with the proposed method. Using the fixed-point theorem, the procedure to solve a simultaneous equation is changed to an iterative process of an affine transformation, which is implemented easily by an optical fractal synthesizer. Although the accuracy and dynamic range are limited when the optical fractal synthesizer is used, interval arithmetic guarantees computational accuracy, and a coordinate conversion provides a wide dynamic range. The effectiveness of the proposed method is verified by a computer simulation. However, the proposed method is not carried out experimentally and several constraints on the optical system should be considered to obviate the errors associated with an optical setup.

In this paper, the proposed high-accuracy computing is implemented by an optical system and its operation is veri-

fied. For this purpose, a *plane of attention width* and a *display plane width* are defined as constraints to resolve the computational accuracy of the optical system with misalignment. Section 2 explains interval arithmetic and the fixed-point theorem as a mathematical basis of the method. Section 3 describes a procedure to implement interval arithmetic using an optical fractal synthesizer. Section 4 describes an optical implementation method, and the computational accuracy obtained by the optical experiment is evaluated.

2 Computational Algorithm

Interval arithmetic is a computational scheme in which a number is represented by an interval that includes the object number.[5,6] For example, a real number x is represented by a close interval $[a,b]$, where $\{x | a \leq x \leq b\}$. Four fundamental rules are described for the intervals as follows:

$$[a,b]+[c,d]=[a+c,b+d], \tag{1}$$

$$[a,b]-[c,d]=[a-d,b-c], \tag{2}$$

$$[a,b]*[c,d]=[\min(ac,ad,bc,bd),\max(ac,ad,bc,bd)], \tag{3}$$

$$[a,b]/[c,d]=[a,b]*[1/d,1/c]. \tag{4}$$

Using intervals with the preceding rules, we can exactly trace the behavior of calculations in spite of errors associated with rounding and quantization. We use interval arithmetic to calculate while ignoring the errors caused by optical setup.

The fixed point x^* of a mapping $f:X \rightarrow X$ is the point that satisfies $f(x^*)=x^*$ and its existence is proved if $f(X) \subset X$, where X is a space **R** (Refs. 6 and 7). The fixed

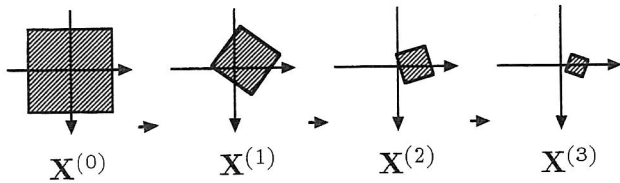**Fig. 1** Iterative transformation on a 2-D interval $\mathbf{X}^{(i)}$.

Fig. 2 Schematic diagram of the optical fractal synthesizer.

point $x^\star$ is calculated with a sequence of iterative operations $x_{i+1} = f(x_i)$. After a sufficient number of iterations, x_i goes to the fixed point $x^\star$.

Various computing algorithms have been developed with interval arithmetic and the fixed-point theorem.[8] As an instance of these algorithms, we explain a solution for a simultaneous linear equation. The target equation is expressed as

$$\mathbf{A}\mathbf{x} = \mathbf{a}, \tag{5}$$

where $\mathbf{A}$ is an $n \times n$ matrix and $\mathbf{x}$ and $\mathbf{a}$ are unknown and known n vectors, respectively. First, an approximate inverse matrix of $\mathbf{A}$ (denoted by $\mathbf{K}$) and an approximate solution (denoted by $\tilde{\mathbf{x}}$) are calculated. Equation (5) can be rewritten as the following form:

$$\mathbf{K}(\mathbf{a} - \mathbf{A}\tilde{\mathbf{x}}) + (\mathbf{I} - \mathbf{K}\mathbf{A})\mathbf{x}^\star = \mathbf{x}^\star, \tag{6}$$

where $\mathbf{I}$ is a unit matrix and $\tilde{\mathbf{x}} + \mathbf{x}^\star$ gives the correct solution of Eq. (5).

Referring to Eq. (6), if the right-hand side is considered as a mapping of a space $\mathbf{R}^n$, $\mathbf{x}^\star$ can be regarded as the fixed point of the mapping. Namely,

$$g(\mathbf{x}^\star) = \mathbf{K}(\mathbf{a} - \mathbf{A}\tilde{\mathbf{x}}) + (\mathbf{I} - \mathbf{K}\mathbf{A})\mathbf{x}^\star = \mathbf{x}^\star \tag{7}$$

indicates that $\mathbf{x}^\star$ is the fixed point of the mapping g.

Although Eq. (7) describes the mapping of the conventional number representation system, the mapping can also be applied to interval arithmetic:

$$g(\mathbf{X}) = \mathbf{K}(\mathbf{a} - \mathbf{A}\tilde{\mathbf{x}}) + (\mathbf{I} - \mathbf{K}\mathbf{A})\mathbf{X}, \tag{8}$$

where $\mathbf{X}$ is an n vector whose elements are intervals that represent the range of individual elements of the vector $\mathbf{x}^\star$. According to the fixed-point theorem, if $g(\mathbf{X}) \subset \mathbf{X}$, the existence of the fixed point $\mathbf{x}^\star \in \mathbf{X}$ is guaranteed.

To ensure convergence of the mapping, the following iterative calculation is applied:

$$\mathbf{X}^{(i)} = g[\mathbf{X}^{(i-1)}] \cap \mathbf{X}^{(i-1)},$$

$$\mathbf{X}^{(0)} = \mathbf{X}, \tag{9}$$

where the superscript indicates the iteration number. As i is increased, $\mathbf{X}^{(i)}$ converges to the fixed point $\mathbf{x}^\star$. Consequently, the solution of Eq. (5) is given by $\tilde{\mathbf{x}} + \mathbf{X}^{(i)}$ for sufficiently large i. When $\mathbf{X}^{(i)}$ consists of two intervals, $\mathbf{X}^{(i)}$ is represented as a rectangular area and is transferred, as shown in Fig. 1. In this case, the iterative calculation is
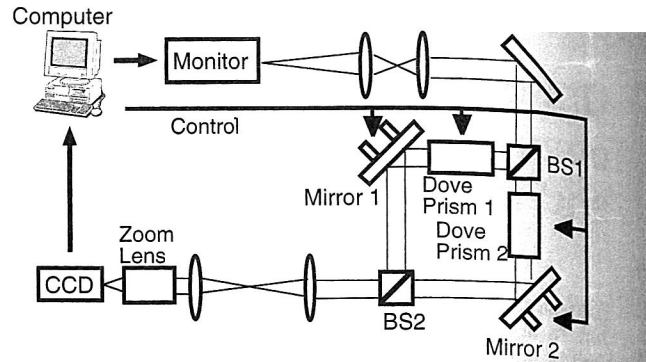
implemented easily by an optical TV feedback system called the *optical fractal synthesizer*, described in the next section.

3 Optical Implementation

Figure 2 shows the setup of the optical fractal synthesizer.[9] In the experimental system, two affine transformations connecting a 2-D plane are implemented by dove prisms, mirrors, and a zoom lens. An affine transformation is written as follows:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}. \tag{10}$$

In the experimental system, the matrix in Eq. (10) is represented as a product of a reduction, a rotation, and a reflection matrix. Rotation and translation are performed by a dove prism and a tilted mirror. Reduction is carried out with a zoom lens. The dove prism and the mirror are controlled electrically. The CCD camera is employed to capture the transformation and to transfer it to the computer and the CRT display. One cycle of the iterative transformations expends about 30 ms, which is limited by the transfer rate of the CCD camera.

To precisely set the coefficients a to f, we develop a technique called the *three points method*. In the three points method, three different bright spots at (x_1, y_1), (x_2, y_2), and (x_3, y_3) are sequentially displayed on the CRT monitor, and we capture the locations of the points, (x_1', y_1'), (x_2', y_2'), and (x_3', y_3'), by the CCD camera. Then we can calculate the coefficients as follows:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{(x_1 - x_2)(y_2 - y_3) - (x_2 - x_3)(y_1 - y_2)}$$
$$\times \begin{bmatrix} x_1' - x_2' & x_2' - x_3' \\ y_1' - y_2' & y_2' - y_3' \end{bmatrix} \begin{bmatrix} y_2 - y_3 & -y_1 + y_2 \\ -x_2 + x_3 & x_1 - x_2 \end{bmatrix}, \tag{11}$$

$$\begin{pmatrix} e \\ f \end{pmatrix} = - \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_1' \\ y_1' \end{pmatrix}. \tag{12}$$

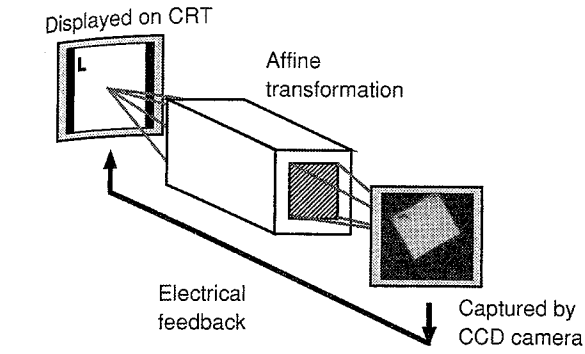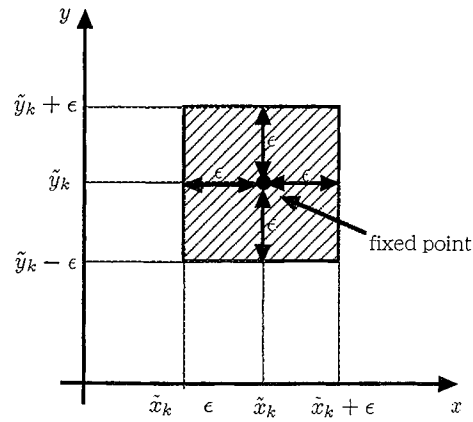To compensate differences between the target and the obtained coefficients, the parameters of the optical fractal syn-

Fig. 3 Conceptual diagram for iterative transformation using the optical fractal synthesizer.

a)

b)

Fig. 4 Procedure for POA generation and coordinate system conversion.

thesizer are controlled. By repeating these procedures, a system configuration with high precision is achieved.
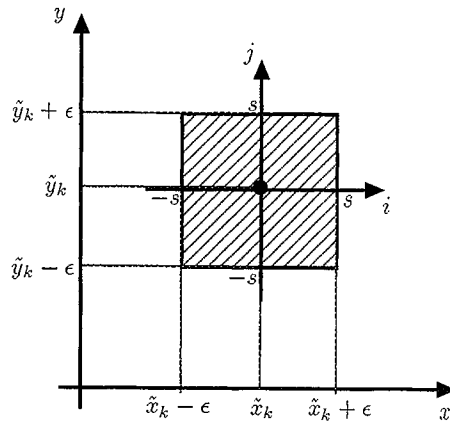
Figure 3 shows a conceptual diagram of the iterative calculation shown in Eq. (9) using the optical fractal synthesizer. Here $\mathbf{X}^{(i)}$ is displayed on the CRT display and is transferred to $\mathbf{X}^{(i+1)}$ by an affine transformation; $\mathbf{X}^{(i+1)}$ is captured by the CCD camera, which is directly returned to the CRT display. To execute interval arithmetic, many points in the interval must be calculated efficiently. In the optical fractal synthesizer, all points in the input plane are transformed simultaneously. Therefore, high computational capability is expected.

To guarantee computational precision and to extend the dynamic range, the following procedure is carried out.

1. Configure the optical fractal synthesizer for a given problem. The orientation of the dove prism, the tilt angle of the reflection mirror, and the magnification of the zoom lens are adjusted to achieve the required affine transformation.

2. Apply the affine transformation to the input image and repeat it iteratively. Figure 1 shows the transition of the input plane during the iteration. Observe the location of a bright spot after a sufficient number of iterations. Then the transferred point is represented as $(\tilde{x}_k, \tilde{y}_k)$, where k is 0 at the initial state.

3. Generate a rectangle area around the obtained point $(\tilde{x}_k, \tilde{y}_k)$. The area is called the *plane of attention* (*POA*). Figure 4(a) depicts the POA. The POA is defined as $\{(x,y); \tilde{x}_k - \epsilon \leqslant x \leqslant \tilde{x}_k + \epsilon, \tilde{y}_k - \epsilon \leqslant y \leqslant \tilde{y}_k + \epsilon\}$, where ϵ is an estimated error caused by the optical setup. In the following, we refer to ϵ as the *POA width*.

4. Convert the coordinate system of the POA for the feedback process. Figure 4(b) shows the relationship between the coordinate system (x,y) and the display coordinate system (i,j). In the display coordinate system, the POA is represented as $\{(i,j); -s \leqslant i \leqslant s, -s \leqslant j \leqslant s\}$, where s is called the *display plane width*. As a result of the coordinate conversion, the POA is enlarged and is ready for the next transformation. The affine transformation for the calculation

is also modified as follows:

$$g\binom{i}{j} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\binom{i}{j} + \alpha\binom{e}{f} + \alpha\begin{bmatrix} (a-1) & b \\ c & (d-1) \end{bmatrix}\binom{\tilde{x}_k}{\tilde{y}_k}, \quad (13)$$

where $\alpha = s/\epsilon$.

5. Calculate the boundary of the resultant area represented by Eq. (14), which includes the correct solution of the problem.

$$\{(x,y); \tilde{x}_{center} - \alpha^{-k}\epsilon \leqslant x \leqslant \tilde{x}_{center} + \alpha^{-k}\epsilon, \tilde{y}_{center} - \alpha^{-k}\epsilon$$

$$\leqslant y \leqslant \tilde{y}_{center} + \alpha^{-k}\epsilon\},$$

$$(\tilde{x}_{center}, \tilde{y}_{center}) = \left(\sum_{l=0}^{k} \tilde{x}_l\alpha^{-l}, \sum_{l=0}^{k} \tilde{x}_l\alpha^{-l}\right). \quad (14)$$

6. If the result has acceptable precision, terminate the process; otherwise, increment k by 1 and return to step 2.

4 Experimental Verification

To verify the principle of the method, we executed an optical experiment. We consider the following two-variable simultaneous linear equation:
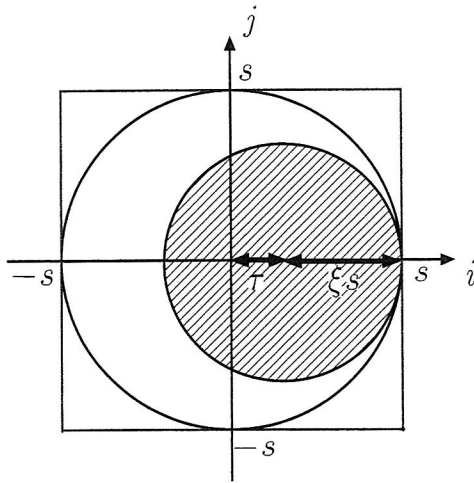
Fig. 5 Relationship between display plane width s and maximum value of translation τ. The large circular area is transformed to the shadowed area by the affine transformation.

$$\begin{bmatrix} 0.64 & 0 \\ 0 & 0.64 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -15\sqrt{3} \\ 12\sqrt{3} \end{pmatrix}. \tag{15}$$

According to the algorithm, an approximate solution $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = (-81.25, 54.6875)$ and an approximate inverse matrix $\mathbf{K}$,

$$\mathbf{K} = \frac{1}{0.64} \begin{bmatrix} 0.7 & -0.3\sqrt{3} \\ -0.3\sqrt{3} & 1.3 \end{bmatrix}, \tag{16}$$

were used. From Eq. (7), the mapping for the given problem is as follows:

$$g\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} 0.3 & 0.3\sqrt{3} \\ 0.3\sqrt{3} & -0.3 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 40 \\ -50 \end{pmatrix}. \tag{17}$$

In the proposed method, the POA width ϵ and the display plane width s should be determined *a priori*. Because these constants are seriously affected by the system setup, such as optical aberration and the setting accuracy of the parameter, we determine them experimentally. The procedure for determining ϵ and s is as follows. First, we evaluate an error distribution function $\delta(e,f)$, where e and f are the translation parameters that appear in the affine transformation; $\delta(e,f)$ indicates the distance between the fixed point and the transferred $(\tilde{x}_0, \tilde{y}_0)$ for specific values of e and f. If ϵ is set so as to be larger than the maximum value of $\delta(e,f)$, ϵ is expressed as

$$\epsilon = \sup\{\delta(e,f); -\tau \le e \le \tau, -\tau \le f \le \tau\}, \tag{18}$$

where τ is selected to satisfy the constraint that ϵ is smaller than half of the CRT size.

Second, the display plane width s is determined from τ. Figure 5 shows the geometrical relationship between s and τ. Using a contraction factor ξ defined as $\xi = |ad - bc|^{1/2}$, s is written as follows:
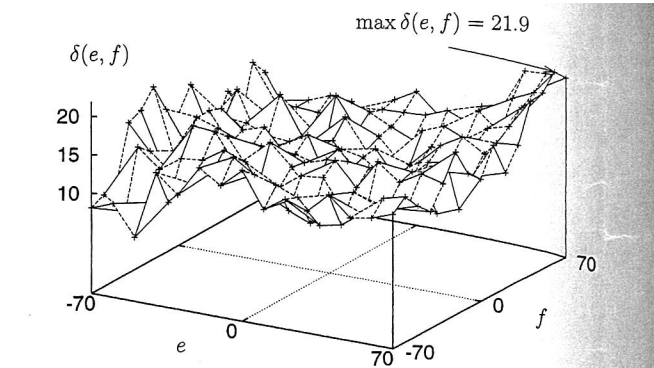

Fig. 6 Error distribution function for the experimental setup.

$$s = \frac{\tau}{1 - \xi}. \tag{19}$$

In this case, s is chosen to be smaller than half of the CRT size.

We examine the error distribution function $\delta(e,f)$ for the following affine transformation,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{bmatrix} 0.3 & 0.3\sqrt{3} \\ 0.3\sqrt{3} & -0.3 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}. \tag{20}$$

In the experimental system, the display plane consists of 512×485 pixels. The origin of the coordinate system is set at the center of the display plane. We change e and f every 10 pixels from -70 to 70, discretely. Although $\delta(e,f)$ should be examined for arbitrary transformation, we just tested a given transformation for simplicity. Figure 6 shows the obtained $\delta(e,f)$. From the result, we set $\epsilon = 25$ and $s = 175$. Therefore, α is determined as 7. As seen from Eq. (14), $\alpha^{-k}\epsilon$ determines computational accuracy. When α is small, the parameter k is forced to be large to achieve high precision. To increase computational efficiency, we must adjust the optical system strictly.

The fixed point for the case of $e = 40$ and $f = -50$ is obtained by the experimental system using the constants already obtained. Figure 7 shows the change of the display plane during the iterative transformation. Many points in the rectangular area at the initial stage converge in a bright spot after a sufficient number of iterations. Table 1 shows the transition of the intervals associated with k. In this case, $\{(x,y); 40.6550591976043 \le x \le 40.6550591976045, -22.2115473580836 \le y \le -22.2115473580834\}$ is obtained at the 17th stage, where $\alpha^k \epsilon \approx 1.07 \times 10^{-13}$ from Eq. (14). Consequently, the solution of Eq. (15) is $\{(x,y); -40.5949408023957 \le x \le -40.5949408023955, 32.4759526419164 \le y \le 32.4759526419166\}$. Comparing

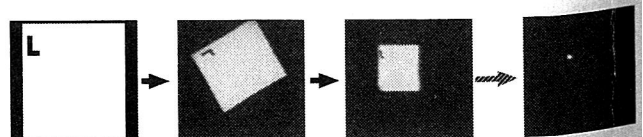**Fig. 7** Transition of 2-D interval during iterative transformation.

Table 1 Transition of the resultant area $\{(x,y); x_{min} \leq x \leq x_{max}, y_{min} \leq y \leq y_{max}\}$ associated with k.

k	x_{min}	x_{max}
0	26	76
3	40.6	40.7
6	40.654	40.655
9	40.65505	40.65506
12	40.6550591	40.6550592
15	40.65505919760	40.65505919761
17	40.6550591976043	40.6550591976045

k	y_{min}	y_{max}
0	-37	-13
3	-22.2	-22.1
6	-22.2116	-22.2112
9	-22.211547	-22.211546
12	-22.211547359	-22.211547355
15	-22.21154735808	-22.21154735807
17	-22.2115473580836	-22.2115473580834

the optically obtained result with a solution $(-40.5949408023955, 32.4759526419164)$ calculated by a workstation,* we can confirm that the computational accuracy of the optical implementation is of the order of 10^{-13}. This accuracy is limited by the control accuracy of the coefficients of the required affine transformation of the optical system. Although the accuracy of the result is limited by the setting accuracy of the coefficients of the affine transform, the proposed method was executed with high accuracy.

5 Conclusion

A method for high-accuracy analog optical computing based on interval arithmetic and the fixed-point theorem was implemented by the optical fractal synthesizer. By investigating the error distribution function, the computational efficiency was estimated. The accuracy of an example calculation using the experimental system was 10^{-13}, which is equal to the accuracy of the required affine transformation coefficients. The effectiveness of the interval arithmetic for high-accuracy optical computing was verified.

---

*This solution was calculated with Ref. 10, a library that supports the extended precision arithmetic whose maximum number of decimal digits in a conversion is 104.

References

1. D. G. Feitelson, *Optical Computing*, MIT, Cambridge (1988).
2. F. J. Taylor, "Residue arithmetic: a tutorial with examples," *Computer* **17**(5), 50–62 (1984).
3. J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci. USA* **79**, 2554–2558 (1982).
4. J. Tanida, W. Watanabe, and Y. Ichioka, "High-accuracy optical computing based on interval arithmetic and the fixed-point theorem," *Appl. Opt.* **35**, 1367–1371 (1996).
5. R. E. Moore, *Interval Analysis*, Prentice-Hall, Englewood Cliffs, NJ (1966).
6. G. Alefeld and J. Herzberger, *Introduction to Interval Computations*, Academic Press, New York (1983).
7. L. E. J. Brouwer, "Über Abbildungen von Mannigfaltigkeiten," *Math. Ann.* **71**, 97–115 (1912).
8. E. Kaucher and S. M. Rump, "E-Methods for fixed point equations $f(x)=x$," *Computing* **28**, 31–42 (1982).
9. J. Tanida, A. Uemoto, and Y. Ichioka, "Optical fractal synthesizer: concept and experimental verification," *Appl. Opt.* **32**, 653–658 (1993).
10. S. L. Moshier, ed. Cephes Mathematical Function Library, see http://netlib.belllabs.com/netlib for more detail.

Toru Sasaki received his BE and ME degrees in applied physics in 1996 and in material and life science in 1998 from Osaka University, where he is currently a PhD student in the Department of Material and Life Science. His research interests are optical computing with an emphasis on the optical feedback system to generate fractal images and the application of the optical computing system, including image compression and numerical computing.

Jun Tanida received his BE, ME, and DE degrees in applied physics in 1981, 1983, and 1986, respectively, from Osaka University, where he has been associated with the Department of Material and Life Science at Osaka University and is now an associate professor. His research interests are optical computing with an emphasis on the optical parallel digital computing system and the application of the optical computing system, including image processing, numerical computing, and fractals.

Yoshiki Ichioka received his BE, ME, and DE degrees in applied physics in 1960, 1962, and 1966, respectively, from Osaka University, Japan, where he has been associated with the Department of Material and Life Science since 1962 and has been a professor since 1985. His main research activities are digital image processing, optical information processing, optical computing, and the development of digital image processing systems and his present research interest is massively parallel information processing using optical and electronic techniques. Dr. Ichioka is a fellow of the Optical Society of America.