Optical Engineering

Optical Engineering. SPIED igital Library.org

Special-purpose computer for digital holographic high-speed three-dimensional imaging

Yota Yamamoto Shintaro Namba Takashi Kakue Tomoyoshi Shimobaba Tomoyoshi Ito Nobuyuki Masuda



Special-purpose computer for digital holographic high-speed three-dimensional imaging

Yota Yamamoto, a,* Shintaro Namba, Takashi Kakue, Tomoyoshi Shimobaba, Tomoyoshi Ito, and Nobuyuki Masuda

^aChiba University, Graduate School of Engineering, Chiba, Japan ^bTokyo University of Science, Faculty of Industrial Science and Technology, Tokyo, Japan

Abstract. Digital holography is attracting attention because it can record instantaneous three-dimensional (3D) information and record dynamic phenomena. However, when recording high-speed phenomena, the frame rate ranges from tens of thousands to hundreds of millions, and the calculation time of the reconstructed images is a problem. We have developed a special-purpose computer for high-speed 3D imaging using digital holography. The developed special-purpose computer has four calculation modules and has achieved a calculation time 68 times faster than that of a personal computer with 48 cores. With the developed computer, a total of 32 reconstructed images can be calculated in 0.69 ms from four holograms of 128 × 128 pixels with eight varying depths. © *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.OE.59.5.054105]

Keywords: optics; digital holography; high-speed three-dimensional imaging; field-programmable gate array.

Paper 20200299 received Mar. 13, 2020; accepted for publication May 15, 2020; published online May 28, 2020.

1 Introduction

The digital holography field began with electronic recording of holograms in the 1960s. ^{1,2} Digital holography can dynamically capture information about the amplitude and phase of light. In addition, it is possible to record an instantaneous three-dimensional (3D) velocity field, ^{3–7} which is attracting attention. Hologram resolution and image quality have also improved significantly. Gigapixel digital holography ^{8,9} and large-scale holograms ^{10,11} have been manufactured. Large-scale holograms enable deep 3D measurement and simultaneous imaging of a large number of particle images. However, a huge amount of time is required to calculate the reconstructed image. Research about optical methods is also important; however, computation time remains a problem, but algorithm improvements and hardware approaches are very effective.

Our research group has been developing a special-purpose computer for digital holographic particle tracking velocimetry (DHPTV) named fast Fourier transform (FFT)-HORN 12,13 to realize fluid visualization. In DHPTV, the number of pixels is generally large (>1024 × 1024 pixels) and the number of frames can be tens to thousands of frames.

In this study, we deal with high-speed 3D imaging using digital holography. High-speed 3D imaging can be realized by combining it with high-speed imaging. 14-22 It is noteworthy that a faster camera produces a smaller number of pixels of the image sensor. When calculating reconstructed images, computational load is reduced using FFT. In FFT-based reconstructed image calculation, the calculation efficiency improves as the number of pixels increases, and the calculation load can be reduced. However, efficiency is reduced when the number of pixels is small. In other words, when the frame rate is improved and the number of pixels per hologram is reduced, calculation efficiency is reduced and more calculation time is required. When recording high-speed phenomena, the frame rate ranges from several thousand to several million, and a huge amount of time is required to calculate the reconstructed image.

^{*}Address all correspondence to Yota Yamamoto, E-mail: y-yamamoto@chiba-u.jp

To reduce computation time, we have developed a special-purpose computer for high-speed 3D imaging using a field-programmable gate array (FPGA). With a special-purpose calculation circuit, efficient calculations are possible even with a small number of pixels. In this study, 128×128 -pixel hologram data were input. We developed a system that can calculate 32 reconstructed images in parallel with eight depths from four holograms.

2 Convolution Diffraction Integration Method

The intensity of the reconstructed image $\phi(x_i, y_i)$ at a certain reconstructed distance z_i is calculated based on the Fresnel–Kirchhoff diffraction:

$$\phi(x_i, y_i) = \int_{-N/2}^{N/2} \int_{-N/2}^{N/2} I(x_\alpha, y_\alpha) g(x_i - x_\alpha, y_i - y_\alpha) dx_\alpha dy_\alpha, \tag{1}$$

$$g(x_i - x_\alpha, y_i - y_\alpha) = \frac{\exp(ikz_i)}{i\lambda z_i} \exp\left\{\frac{ik}{2z_i} [(x_i - x_\alpha)^2 + (y_i - y_\alpha)^2]\right\},\tag{2}$$

where $\phi(x_i, y_i)$ is the intensity of the reconstructed space at point (x_i, y_i, z_i) , $I(x_\alpha, y_\alpha)$ shows the intensity pattern of the hologram, λ is the wavelength of light, k is the wavenumber of light expressed as $k = \frac{2\pi}{\lambda}$, x_α and y_α are the coordinates on the hologram plane, and N is the number of pixels in the vertical and horizontal directions of the hologram surface and the reconstructed surface.

Equation (1) is a two-dimensional (2D) convolution integral about the x and y axes. It can be transformed using 2D Fourier transform; thus, Eq. (1) can be rewritten as follows:

$$\Phi(n,m) = \hat{I}(n,m)G(n,m),\tag{3}$$

where $\Phi(n,m)$ is the 2D Fourier transform of $\phi(x_i,y_i)$, $\hat{I}(n,m)$ represents the Fourier transform of $I(x_\alpha,y_\alpha)$, and G(n,m) represents that of $g(x_i-x_\alpha,y_i-y_\alpha)$. The convolution integral of $I(x_\alpha,y_\alpha)$ and $g(x_i-x_\alpha,y_i-y_\alpha)$ can be expressed as the product of $\hat{I}(n,m)$ and G(n,m) in the area after the Fourier transform by each Fourier transform. The intensity of the reconstructed surface can be obtained by performing an inverse 2D Fourier transform on the product calculation result and returning it to the original region. Here, FFT can be used to reduce the computational load.⁴

The pixel interval between the hologram and reconstructed surfaces is defined as ΔP in Eq. (2). The Fourier transformed G, which is $\Delta f = 1/(N\Delta P)$ as a minimum unit, is expressed as follows:

$$G(n\Delta f, m\Delta f) = \exp\left\{ikz_i - i\pi\lambda z_i \left[\left(\frac{n}{n\Delta P}\right)^2 + \left(\frac{m}{m\Delta P}\right)^2 \right] \right\}$$
$$= \exp(ikz_i) \exp\left[2\pi i \left(\frac{-\lambda z_i}{2N^2(\Delta P)^2}\right) (n^2 + m^2) \right]. \tag{4}$$

The 3D information of the measurement target is reconstructed by superimposing reconstructed images with different reconstructed distances (i.e., depths). Therefore, $\exp(ikz_i)$ in Eq. (4) can be considered a constant when reconstructing a single surface. In addition, 256 gradations are generated based on the minimum and maximum values of $\phi(x_i, y_i)$; thus, $\exp(ikz_i)$ can be omitted. Therefore, the reconstructed image is calculated as follows:

$$G'(n,m) = \exp\left[2\pi i \left(\frac{-\lambda z_i}{2N^2(\Delta P)^2}\right)(n^2 + m^2)\right].$$
 (5)

In the following, G'(n, m) in Eq. (5) is referred to as G(n, m). Here, the depth-related coefficient z_{param} is defined as follows:

$$z_{\text{param}} = -\frac{\pi \lambda z_i}{2N^2(\Delta P)^2}.$$
 (6)

Equation (5) is expressed by Eq. (6) as follows:

$$G(n,m) = \exp[2\pi i z_{\text{param}}(n^2 + m^2)].$$
 (7)

3 Implementation of Special-Purpose Computer

3.1 Hardware

In the calculation of the reconstructed image, the calculation load during FFT is very large. FFT is a simple calculation that repeats the butterfly operation and is suitable for hardware implementation. In this study, the Virtex-7 FPGA VC707 Evaluation Kit (VC707) provided by Xilinx was used in the initial research stage. It is an evaluation board with a Virtex-7 XC7VX485T-2FFG1761C. Table 1 shows the specifications of the logical FPGA installed in the VC707.

3.2 Implementation of Two-Dimensional Convolution Pipeline

Figure 1 shows a block diagram of the special-purpose calculation pipeline based on the convolution diffraction integration method. The numbers on the diagonal lines in the figure indicate the bit width of the data.

In this circuit, the intellectual property core (FFT v7.1) provided by Xilinx is used to perform the Fourier and inverse transforms of the hologram.²³ This core can select one-dimensional (1D)

Table 1 FPGA specifications: register is a small, high-speed data storage resource on the FPGA, and Block RAM is a large-capacity, low-speed data storage resource. The LUT is a resource for configuring a circuit.

Evaluation board	VC707
Family	Virtex-7
Device	XC7VX485T-2FFG1761C
Number of registers	607,200
Block RAM	37,080 Kb
Number of LUTs	303,600

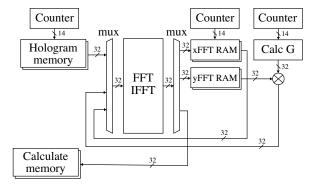


Fig. 1 Block diagram of special-purpose calculation pipeline for convolutional diffraction integration method.

Optical Engineering 054105-3 May 2020 • Vol. 59(5)

FFT and inverse FFT according to the input. It is noteworthy that it is necessary to perform both 2D and inverse FFT. When performing 2D FFT, we first perform 1D FFT in the *x* direction and then 1D FFT in the *y* direction. The same applies to the inverse FFT.

The distance z_i between the hologram and reconstructed surfaces only affects G(n,m). Therefore, by multiplying G(n,m) by distance z_i corresponding to each reconstructed plane as a parameter by Fourier transform $\hat{I}(n,m)$ of a single hologram, it is possible to obtain multiple reconstructed surface intensities in the depth direction. In this system, the initial value z_0 of the reconstructed distance z_i and the reconstructed distance interval dz are stored in the FPGA's Block RAM, and the values are handled as parameters.

The special-purpose pipeline is a state machine that uses four stages, where eight reconstructed images with different reconstructed distances are calculated from a single hologram.

- Stage 1: Hologram data are read from hologram memory, 1D FFT is performed (x direction), and the result is saved to xFFT RAM (Fig. 1). This process is repeated for 128 lines in the y direction. Simultaneously, the initial value z₀ of the reconstructed distance z_i is read.
- Stage 2: The FFT calculation result from stage 1 is read in the y direction from the xFFT RAM, 1D FFT is performed, and it is saved in yFFT RAM (Fig. 1). This process is repeated for 128 lines in the x direction. Simultaneously, the interval dz of the reconstructed distance z_i is read.
- Stage 3: The FFT calculation result from stage 2 is read from yFFT RAM and multiplied by the value of G. Then, 1D inverse FFT is performed. The result is stored in xFFT RAM (Fig. 1). This process is repeated for 128 lines in the y direction.
- Stage 4: The inverse FFT result from stage 3 is read from xFFT RAM in the y direction, 1D inverse FFT is performed, and it is saved in calculate memory (Fig. 1). This process is repeated for 128 lines in the x direction.

It is noteworthy that stages 3 and 4 are repeated while adding the interval dz to the reconstructed distance z_i . This repetition continues until the calculation of eight reconstructed planes is complete.

3.3 Implementation of Calc G Unit

Figure 2 shows a block diagram of the Calc G unit shown in Fig. 1. The Calc G unit calculates G(n, m) represented by Eq. (7). Here, ρ is defined as follows:

$$\rho = -\frac{\pi\lambda}{2N^2(\Delta P)^2},\tag{8}$$

where ρ is a constant that can be calculated in advance on the host personal computer (PC). Multiply the ρ calculated by the host PC and z_i then redefine Eq. (6) by Eq. (8) as follows:

$$z_{\text{param}} = \rho z_i. \tag{9}$$

From Euler's formula, $\exp(i\theta) = \cos \theta + i \sin \theta$ holds. Here, the lookup table (LUT) method is used for the cos and sin calculations, where precalculated values are stored in memory, and the values are obtained quickly by referencing read-only memory (ROM) addresses. In Fig. 2, the real and imaginary parts of G(n,m) are calculated using cos ROM and sin ROM.

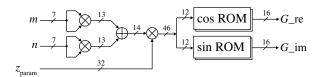


Fig. 2 Block diagram of Calc G unit.

Table 2 Circuit scale.

Resource name	Number of usages	Usage (%)
Register	16,460	2
Block RAM	24,844 Kb	67
LUT	13,283	4

Table 3 Calculation time.

Hardware	Calculation time (ms)	Acceleration ratio
FPGA: VC707	0.69	68
GPU: NVIDIA GeForce GTX 1080	15.90	2.97
CPU: Intel Xeon E5-2697	47.16	1.00

4 Results

4.1 Resource Usage

Table 2 shows the circuit scale of the special-purpose computer developed with the VC707. The operating frequency is 250 MHz. It was possible to install four units that calculate eight reconstructed images from a single hologram.

4.2 Calculation Time

To compare calculation time, a program that calculates 32 reconstructed images from four holograms (128×128 pixels) was created on the central processing unit (CPU) and graphics processing unit (GPU). The comparison of calculation times is given in Table 3.

For the CPU comparison environment, an Intel Xeon E5-2697 2.70 GHz (48 cores) CPU was used. The program was created using CentOS 7.1 and the Intel C++ Compiler 16.0.1.150. In addition, FFTW 3.3.8²⁴ was used for the FFT calculation. It is noteworthy that all cores were used and executed in parallel.

For the GPU comparison environment, an NVIDIA GeForce GTX 1080 (2560 CUDA cores) was installed in the CPU comparison environment. The program was created using the NVIDIA CUDA compiler 9.1. In addition, the cuFFT library²⁵ in NVIDIA CUDA compiler 9.1 was used for the FFT calculation.

As can be seen in Table 3, the developed special-purpose computer is 68 times faster than the CPU and 23 times faster than the GPU.

4.3 Image Quality

The optical system shown in Fig. 3 was constructed and a hologram was captured. Here, a 532-nm laser was used as the light source. The linearly polarized light wave from the laser was propagated to the beam expander and split into two waves. One was propagated into an object. We have set falling granulated sugar particles as an object. Holograms were captured with an exposure time of $\sim 1/300,000$ s using a high-speed polarization camera. This camera has a polarization-detection function of a set of 2×2 pixels. Even though the function can select four polarization axes, we only used one axis in this experiment. Since captured holograms had vacant pixels, we interpolated the values of the vacant pixels before the reconstruction calculation. Figure 4 compares a reconstructed image calculated by the CPU with double precision

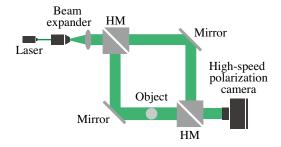


Fig. 3 Optical setup (HM stands half mirror).

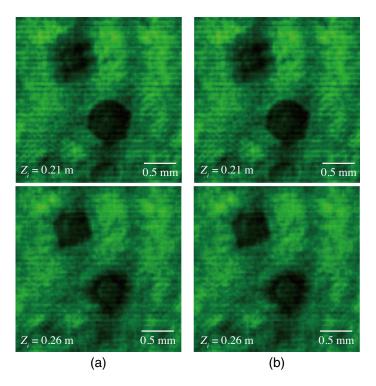


Fig. 4 Comparison of images reconstructed by (a) CPU and (b) FPGA (Video 1, MPEG, 302 KB [URL: https://doi.org/10.1117/1.OE.59.5.054105.1]).

and a reconstructed image calculated by the developed special-purpose computer. Table 4 shows the reconstructed conditions. The average peak signal-to-noise ratio and structural similarity with 32 reconstructed images were 43.8 and 0.993, respectively. As shown in Fig. 4, the special-purpose computer can calculate with image quality equivalent to the result calculated by the CPU.

Table 4 Reconstructed conditions.

Reconstructed distance	210 mm
Wavelength	532 nm
Pixel pitch	20 μm
Number of pixels	128 × 128

Optical Engineering 054105-6 May 2020 • Vol. 59(5)

5 Discussion and Future Work

We have developed a special-purpose computer for high-speed 3D imaging using digital holography. We have successfully developed a system that can calculate 32 reconstructed images in 0.69 ms from four holograms of 128×128 pixels. This special-purpose computer comprises four units and is 68 times faster than a 48-core CPU and 23 times faster than a GPU with 2560 CUDA cores.

Generally, a multicore system involves overhead, such as preprocessing and control in parallel execution. Even if the number of parallel processes is increased, increased speed cannot be obtained. In this study, we handled a 128 × 128-pixel hologram for high-speed imaging. Since the number of pixels is small, one hologram of processing is immediately terminated. The overhead of parallel execution was apparent and speedup reached its peak. However, such overhead does not exist in a special-purpose computer. If the conditions are met, a special-purpose computer can be sped up as the number of parallel processes increases. The results of this study suggest the usefulness of special-purpose computers in high-speed 3D imaging with a small number of pixels in tens of thousands to hundreds of millions of frames.

In Eq. (4), z_i should satisfy the following sampling condition:²⁷

$$z_i = \frac{N \times \Delta P^2}{\lambda}. (10)$$

Owing to the physical constraints of the experimental environment, we set z_i to 0.21 m, so the sampling condition is not satisfied. Zero padding is generally required to satisfy this sampling condition. However, it needs more memory and is not suitable for FPGA implementations, and it takes extra calculation time. Using the simulation results, we confirmed that zero padding had no significant effect on the quality of the reproduced image.

In the future, we plan to increase the density and integration of computing units. In the current study, the capacity and operating frequency of the Block RAM that temporarily stores calculation data was a bottleneck, and the installation of four units was a significant limitation. Our system uses 600-Kb memory per unit. In terms of memory capacity, the number of units can be increased by one or two. On the other hand, the distance between the computing units and memory block will increase, resulting in a decrease in the operating frequency. In this work, based on the trade-off between the number of computing units, memory capacity, and operating frequency, we chose four units operating at 250 MHz to obtain better computational performance. However, the latest FPGAs have large, high-speed on-chip UltraRAM storage with a maximum capacity of 500 Mb.²⁸ The latest FPGA board (Xilinx ALVEO U250) has 100 times more resources in terms of LUT; thus, it is expected to be nearly 100 times faster.

In addition, in this study, we focused on the calculation time of the special-purpose computer, and the communication of the special-purpose computer used a low-speed I/O port. The total communication time was 29.15 ms. However, the communication time can be reduced using direct memory access (DMA) transfer. The measured value of DMA communication time using the Xilinx ALVEO U250 is 0.017 ms to transfer four images and 0.078 ms to receive 32 images. Data reception can be ignored because it can be performed in parallel with the calculation. The actual communication time is only 0.017 ms at the time of data transmission, and, in future, it will be possible to develop a high-speed special-purpose computer with this communication time.

Acknowledgments

This work was partially supported by the Japan Society for the Promotion of Science KAKENHI Grant Nos. 19H01097 and 18K11328. The authors declare no conflicts of interest.

References

1. J. W. Goodman and R. Lawrence, "Digital image formation from electronically detected holograms," *Appl. Phys. Lett.* **11**(3), 77–79 (1967).

- 2. T. Tahara et al., "Digital holography and its multidimensional imaging applications: a review," *Microscopy* **67**(2), 55–67 (2018).
- 3. D. H. Barnhart, R. J. Adrian, and G. C. Papen, "Phase-conjugate holographic system for high-resolution particle-image velocimetry," *Appl. Opt.* **33**(30), 7159–7170 (1994).
- 4. H. Meng and F. Hussain, "In-line recording and off-axis viewing technique for holographic particle velocimetry," *Appl. Opt.* **34**(11), 1827–1840 (1995).
- 5. J. Sheng, E. Malkiel, and J. Katz, "Single beam two-views holographic particle image velocimetry," *Appl. Opt.* **42**(2), 235–250 (2003).
- 6. U. Schnars and W. Jüptner, "Direct recording of holograms by a CCD target and numerical reconstruction," *Appl. Opt.* **33**(2), 179–181 (1994).
- 7. S. Murata and N. Yasuda, "Potential of digital holography in particle measurement," *Opt. Laser Technol.* **32**(7–8), 567–574 (2000).
- 8. S. O. Isikman et al., "Giga-pixel lensfree holographic microscopy and tomography using color image sensors," *PLoS One* **7**(9), e45044 (2012).
- 9. T. Shimobaba et al., "In-line digital holographic microscopy using a consumer scanner," *Sci. Rep.* **3**, 2664 (2013).
- 10. H. Yoshikawa and T. Yamaguchi, "Review of holographic printers for computer-generated holograms," *IEEE Trans. Ind. Inf.* **12**(4), 1584–1589 (2016).
- 11. Y. Tsuchiyama and K. Matsushima, "Full-color large-scaled computer-generated holograms using RGB color filters," *Opt. Express* **25**(3), 2016–2030 (2017).
- 12. N. Masuda et al., "Special purpose computer system with highly parallel pipelines for flow visualization using holography technology," *Comput. Phys. Commun.* **181**(12), 1986–1989 (2010).
- 13. Y. Abe et al., "Special purpose computer system for flow visualization using holography technology," *Opt. Express* **16**(11), 7686–7692 (2008).
- 14. D. D. Aguayo et al., "Insect wing deformation measurements using high speed digital holographic interferometry," *Opt. Express* **18**(6), 5661–5667 (2010).
- 15. T. Kakue et al., "High-speed phase imaging by parallel phase-shifting digital holography," *Opt. Lett.* **36**(21), 4131–4133 (2011).
- 16. T. Tahara et al., "High-speed three-dimensional microscope for dynamically moving biological objects based on parallel phase-shifting digital holographic microscopy," *IEEE J. Sel. Top. Quantum Electron.* **18**(4), 1387–1393 (2012).
- 17. O. Matoba et al., "Optical voice recorder by off-axis digital holography," *Opt. Lett.* **39**(22), 6549–6552 (2014).
- 18. P. Xia et al., "One million FPS digital holography," *Electron. Lett.* **50**(23), 1693–1695 (2014).
- 19. K. Ishikawa et al., "High-speed imaging of sound using parallel phase-shifting interferometry," *Opt. Express* **24**(12), 12922–12932 (2016).
- 20. Y. Tanaka, S. Tani, and S. Murata, "Phase retrieval method for digital holography with two cameras in particle measurement," *Opt. Express* **24**(22), 25233–25241 (2016).
- 21. T. Fukuda et al., "Three-dimensional imaging of distribution of refractive index by parallel phase-shifting digital holography using Abel inversion," *Opt. Express* **25**(15), 18066–18071 (2017).
- 22. T. Kakue et al., "Digital holographic high-speed 3D imaging for the vibrometry of fast-occurring phenomena," *Sci. Rep.* **7**(1), 10413 (2017).
- 23. Xilinx, "Xilinx Product Specication LogiCORE IP Fast Fourier Transform v7.1," https://www.xilinx.com/support/documentation/ip_documentation/xfft_ds260.pdf (2011).
- 24. FFTW Home page, "FFTW," http://www.fftw.org/ (2018).
- 25. NVIDIA, "cuFFT," https://developer.nvidia.com/cufft (2020).
- 26. Y. Yamamoto et al., "Large-scale electroholography by HORN-8 from a point-cloud model with 400,000 points," *Opt. Express* **26**(26), 34259–34265 (2018).
- 27. D. Mas et al., "Fast algorithms for free-space diffraction patterns calculation," *Opt. Commun.* **164**(4–6), 233–245 (1999).
- 28. Xilinx, "UltraRAM," https://www.xilinx.com/support/documentation/white_papers/wp477-ultraram.pdf (2016).

Yota Yamamoto is a PhD student at Graduate School of Engineering, Chiba University. He received his BE and ME degrees from Tokyo University of Science in 2016 and 2018, respectively. His current research interests include high-performance computing and its applications for digital holography and electroholography. He is a member of the Optical Society of Japan (OSJ), the Information Processing Society of Japan (IPSJ), and the Institute of Electronics, Information and Communication Engineers (IEICE).

Shintaro Namba is a graduate student at Graduate School of Industrial Science and Technology, Tokyo University of Science. He received his BE and ME degrees from Tokyo University of Science in 2015 and 2017, respectively. His current research interests include high-performance computing and its applications for digital holography.

Takashi Kakue is an assistant professor at Chiba University. He received his BE, ME, and DE degrees from Kyoto Institute of Technology in 2006, 2008, and 2012, respectively. His current research interests include holography, 3D display, 3D measurement, and high-speed imaging. He is a member of SPIE, the Optical Society (OSA), the Institute of Electrical and Electronics Engineers (IEEE), OSJ, and the Institute of Image Information and Television Engineers (ITE).

Tomoyoshi Shimobaba is a professor at Chiba University. He received his PhD in fast calculation of holography using special-purpose computers from Chiba University in 2002. His current research interests include computer holography and its applications. He is a member of SPIE, OSA, IEICE, OSJ, and ITE.

Tomoyoshi Ito is a professor at Chiba University. He received his PhD in special-purpose computers for many-body systems in astrophysics and molecular dynamics from the University of Tokyo in 1994. His current research interests include high-performance computing and its applications for electroholography. He is a member of ACM, OSA, OSJ, ITE, IEICE, IPSJ, and the Astronomical Society of Japan.

Nobuyuki Masuda is a professor at Tokyo University of Science. He received his PhD in simulations of dynamical instability of accretion disks from the University of Tokyo in 1998. His current research interests include high-performance computing and its applications for digital holography. He is a member of IPSJ and IEICE.